Welsh School of Pharmacy - Cardiff University
Medicinal Chemistry Department

PhD Thesis

Computer aided design and synthesis of new anticancer and antiviral compounds

Maria Chiara Barbera

February 2007

UMI Number: U584142

UMI

Dissertation Publishing

ProQuest

El hombre que más ha vivido no es aquel que más años ha cumplido, sino aquel que más ha experimentado la vida (J.J.Rosseau).

Vencer sobre nuestras dificultades es lo más difícil, pero en el momento en que se decide combatir in realidad ya hemos vencido.

Haz lo que ames, porque así amarás lo que haces.

Acknowledgements

Table of Contents

Chapter 3

Introduction

Table of Figures

Chapter 3

Chapter 4

Chapter 5

Table of Schemes

Table of Tables & Charts

Abstract

Among all the recent advances in different but complementary scientific fields, the impressive progress and development of more powerful computational techniques play a key role in drug design. **Computer aided design** has become indispensable and complimentary to traditional and modern approaches by guiding medicinal chemists to improve and speed the discovery, design, synthesis and optimization of novel active molecules.[1,2,3,4]

The present study has been focused on the design of protein-protein and protein-DNA interacting small molecules inhibitors with the support of computational methods. An accurate investigation of protein-protein and protein-inhibitor interactions present in the microtubules complex[5] has been carried out by means of different computational methods: QSAR analysis, Molecular Docking, Flexible Alignment and Molecular Dynamics Simulation. The information derived from these computational studies have guided the design and chemical synthesis of new arylthioindole derivates as small molecule inhibitors of tubulin polymerization.[6,7] Based on structural data and synthesis work previously carried out,[i] two distinct databases of potential small molecule inhibitors of HCV helicase/NTPase enzyme[8] have been generated and evaluated by computational methods, thus guiding the selection of several structures. Successfully synthesised compounds have been sent for biological testing via an HCV replicon assay: some of them have shown interesting and promising inhibitory activity by blocking the virus replication *in vitro*. The refined homology model of the antiapoptotic protein Bcl-2[9,10] has led to the generation and virtual screening of small combinatorial libraries of drug-like compounds, carried out by means of different computational applications (pharmacophore and conformational searches, molecular docking, molecular dynamics) to find non-peptidic inhibitors of Bcl-2. By means of various molecular modeling applications (pharmacophore and conformational search, generation of combinatorial libraries, molecular docking, molecular dynamics), structurally diverse compounds have been designed and virtually screened; following the investigation, anisomycin derivates have resulted to be potential candidates as small molecule inhibitors of Mdm2/p53 complex.[9,10]

Chapter 1

Introduction

1.1 Medicinal Chemistry and Rational Drug Design

In many ancient cultures, chance and observation have led to the discovery and subsequently use of plants extracts as medical remedies. Although the effectiveness of many of these herbal potions on the health of the patients was due both to a placebo effect and the psychological impression derived from the witch rituals and religious ceremonies with which the remedies were administered, there are examples of ancient pharmacologically effective remedies from which modern drugs have been developed: **opium**, leaves from **coca bush**, bark from **cinchona tree**, *ipecacuanha*'s root.

In the nineteenth century, chemistry progress in the structural identification and analysis of the active principles contained in the herbal mixtures led to the synthesis of the first semi-synthetic and synthetic drugs.

The twentieth century saw Paul Ehrlich[11] establishing the principle of pharmacological selectivity (*the magic bullet theory*, fundamental concept of chemotherapy) and investigating the dose-dependent effect of a drug administered to explain the activity of a compound as a poison or a medicine.

Both the structural simplification of complex molecules and the purification of endogenous compounds promoting the synthesis of their corresponding analogs have been significant progresses in medicinal chemistry. While new classes of bacteriostatic and bactericidal agents were discovered (sulfonamides and penicillins, respectively), the structures of chemical molecules was finally revealed with the development of X-ray crystallography.

The 1945-1970s were milestones in the history of medicinal chemistry also effecting the society by improving health and life quality: novel antibiotics and semi-synthetic penicillins were developed against infections, while advanced psychotropic medicines were used to treat diseases of the central system. Other life saving molecules were discovered thus allowing more specific, effective and less toxic treatment of cardiovascular diseases, ulcers and asthma. The 1960s also saw the discovery and revolutionary introduction of oral contraceptives.

Although these thirty years signed the discovery of new potent drugs and the growth of pharmaceutical companies, still the mechanism of action of the active principles was not

clear and, moreover, medicinal chemistry was mainly based on **structure-based drug design.**

According to this drug design strategy, the identification of a lead compound and the synthesis and tests of new compounds (the design of which has traditionally derived from **structure activity relationship** – SAR – studies of the lead and its analogs) help to identify structural groups in the lead molecule that are crucial for the pharmacological activity (*pharmacophore*), and provide information about the potential structure of the binding site. These studies also showed that significant structural simplification of natural compounds does retain activity and that increased rigidity in simple molecules can improve their selectivity.

Outstanding advances in molecular biology (elucidation of signal transduction), genetic engineering (mapping of the human genome), chemistry (asymmetric synthesis, X-ray crystallography and 2-D nuclear magnetic resonance spectroscopy, modern purification techniques, solid/phase synthesis and combinatorial chemistry followed by microscale experimentation), and molecular modelling have guided and led medicinal chemistry to increasingly adopt a rational approach by improving the scientific knowledge of the human body at molecular level and, thus, explaining the mechanism of action of the drugs and the structure of their targets.

Cimetidine and *enalapril* are examples of the more efficient **target-based drug design** in medicinal chemistry, whose main objectives to achieve are now to increase the selectivity of a drug for a specific target in order to make it more effective and much less toxic, and to improve the activity of a compound, necessarily containing the proper pharmacophore, by optimizing and strengthen the interactions with its target. Feasibility in chemical synthesis of an active molecule and its capability of reaching the final specific target (*pharmacokinetics*) are additional but essential aims of target-based drug design.

Among all the advances in different but complementary scientific fields mentioned above, the impressive progress and development of more powerful computational techniques play a key role in drug design. **Computer aided design** has become indispensable and complimentary to other traditional and modern approaches by guiding, improving and speeding the discovery, design, synthesis and optimization of novel active molecules.

Nowadays the use of computers in the drug design process allows analysis and comparisons of small molecules as well as of proteins and their binding sites; design of

2

a lead compound and the identification of a *pharmacophore*, that can be successfully followed by combinatorial library design to quickly generate a large number of analogs; molecular docking, energy minimization and molecular dynamics simulation in order to investigate and optimize the binding between the best compounds and their targets; 3D-QSAR studies to correlate biological activity to chemical structure, just to mention the most used computational applications.

1.2 Molecular Modelling

Nowadays Medicinal chemistry is facing the challenging task of discovering new therapeutic molecules to treat human diseases. For a long time, chance observation and random screening, as well as experience and intuitution of the chemist, have played an important role in the laborious, time consuming and expensive "trial and error cycles" process. Although the concept of receptor and enzyme as molecular targets and the *lock-and-key* theory (P.Erlich,1909[11] and E.Fisher,1894[12]) have been always considered, it was only in the 1970s, when protein isolation, X-ray crystallography and high field NMR allowed to observe and study the detailed molecular structures together with their associated properties (geometry, electrostatic, hydrophobicity), that rational drug discovery and design could progress and molecular modelling could develop and be accepted as a complete and indispensable discipline to contribute to the success of drug research.

Based on similarity with known active compound (determined by means of superimposition procedures) or on complementarity with the tridimensional structure of the target's active site, molecular modelling allows the design of new molecules and the prediction of their potential biological properties by also taking into account both molecular interactions and binding energies specifically involved in the biological system under study.[13] Relevant advances in chemistry, molecular biology and computer technology have contributed to the growth, improvement and integration of molecular modelling in the research process, but this progress was mainly due to the increasing interest and need to overcome the limitation of the first generation rational drug design. In fact, in its very beginning this has been focused on the fixed 2D representation and visualization of the molecular entities, thus helping only the lead optimization of a given series.

By taking into consideration the more detailed and specific information provided by the 3D visualization of the ligand and its bioactive conformation bound to its protein (when available), molecular modelling has emerged and asserted itself in the challenging process of lead discovery and, particularly, in the following fields:

- Direct Drug design
- Indirect Drug Design
- Database Search
- 3D automated Drug design
- Molecular Mimicry

Thanks to all these areas covered by molecular modelling, chemical intuitition has been improved and more accurate insight into the biological mechanism involved in the ligand-protein complex under investigation has been provided. Although the current limitation of the exact calculation of the free energy of ligand-protein binding (that is fundamental for structure based drug design and to speed up the whole drug development process), the rapid evolution of CADD (Computer Assisted Drug Design) will allow not only to design new potential candidate drugs, but also to develop and further improve database and simulation techniques, registration, management and analysis of chemical or genetic structural data methods, thus effectively leading to and facilitate a major increase in productivity in the scientific research.[14]

1.3 Computer aided drug discovery

The present study aims to investigate protein-protein interactions of distinct cellular protein-protein complexes by means of different computational methods. The information derived from this investigation leads to the computer aided design and synthesis of small molecule inhibitors of the protein-protein interactions previously analyzed.

A specific methodology has been developed and successfully adopted to achieve the main purposes of the present project, a graphical exemplification of which has been given in Fig.1.1.
In all computational studies aimed to the analysis of the properties of ligand/protein complexes, it is inescapable to know the 3D structure of the target protein, that can be experimentally derived either by X-ray diffraction or NMR. If the protein structure has not been determined, the homology modelling (comparative protein modelling) helps to built a homolog based on a template structure. The higher the similarity between the template and the target protein, the higher the accuracy of the generated homology model.
The subsequent identification and analysis of the active site, usually located in the conserved region of the protein, and of the potential chemical nature of the ligand and its conformation inside the binding site, are the necessary steps for the molecular docking of either well known inhibitors or database of various compounds. This virtual screening is very important as it helps to verify the position of both active site and ligand previously estimated using computational

4

methods, and to observe which molecules have shown the more efficient and specific interaction with the protein.

Consequently, the flexible alignment of these best docked structures clearly reveals the presence of common patterns to use as scaffolds in the generation of combinatorial libraries, thus adding appropriate fragments to the main moiety by symmetric or asymmetric combinations.

The resulting databases can be directly subjected to a Finger Print study (that, based on principal components analysis, allows the identification of a set of descriptors for compounds classification, thus increasing the probability of identifying molecules with similar activity in virtual screening calculations) or can be accurately analyzed to find out the best candidates for chemical synthesis.

Once the synthesized molecules have been experimentally tested, their QSAR analysis and/or Finger Print screening can lead and guide both the further optimization of their synthesis and the drug design of new derivates.

Interestingly, if a database of well-known inhibitors and their corresponding biological activity values is available, the QSAR analysis can be used since the beginning of the pathway mentioned above, providing remarkable information to the whole study.
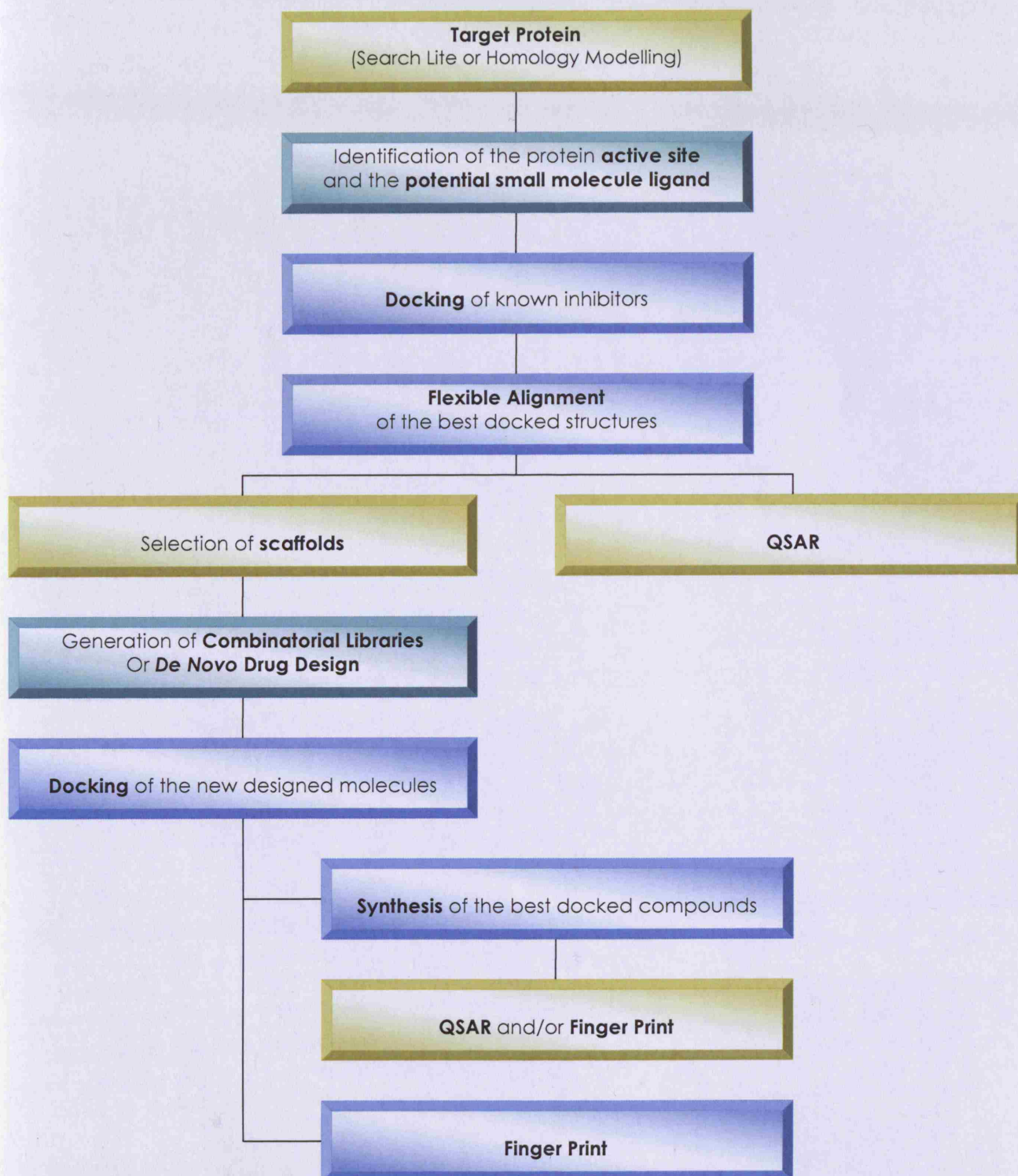
Fig.1.1 – Graphical visualization of the methodology developed and applied in the present project.

1.4 Computational Methods

1.4.1 3D Visualization

Molecular graphics is essential to other analytical techniques such as NMR spectroscopy and X-ray crystallography for the 3D visualization of small molecules, proteins or their complexes.

Indeed, first step in a molecular modelling study is the generation of 3D molecular structure of the compound of interest by means of cartesian coordinates defining the relative positions of atoms in the space. The main sources of these data are:

- X-ray crystallographic databases
- Fragment libraries
- Simply drawing ('sketching') on the computer screen a 2D structure that will be converted to 3D by automated processes.

1.4.2 Geometry optimization

Once a molecular structure has been generated using one of the above methods, its geometry always needs to be optimized in order to find the energy minimum state. *Molecular mechanics* is the most used computational method for calculating molecular geometries and energies as it is simpler and faster than *quantum mechanics* approaches (*ab initio* and *semiempirical molecular orbital* methods), chosen if the required force field parameters are not available and limited to small molecules.

More in detail, molecular mechanics considers electrons and nuclei of the atoms as a group of masses interacting with each other by harmonic forces.

Hooke's law is used for the calculation of the total energy of a molecule considering the deviations from reference "unstrained" bond lengths, angles, torsions, and non-bonded interaction. All these unstrained values, together with force-constants, are specifically known as a *force field*.

Keeping in mind that energy minimization algorithms find only local minima but not the global one, depending on the size of the system under study it is fundamental to choose the appropriate energy minimization method for its geometry optimization.

The most important minimization procedures are:

7

- steepest descent (as initial rough run to be followed by more advanced method) [15,16,17]

- conjugate gradient (for large systems, more efficient and time consuming than steepest descent) [16,17]

- Powell (similar to conjugate gradient but may alter torsion angles) [16,17]

- Newton-Raphson (requires major efforts and storage than the previous methods but converges rapidly to precise minima) [16,17]

The presence of charges or dipoles, solvation effects [15] and the use of solvent dielectric constants [18] must be taken into account when running minimization algorithms.

1.4.3 Conformational analysis

All atoms in a molecule continually move and vary their position and changes in torsion angles occur even at room temperature, thus causing continuous conformational transformation of one *conformer* (or *rotamer*) in another, usually in the low energy state.

Worth noticing is that the *bioactive conformation* of a drug is one of the low energy conformations, even if not necessarily the lowest one; therefore, it is imperative to search for this unique conformation able to bind to a particular active site and to guide the design of new selective and potent derivates [19,20].

The time required for an exhaustive conformational analysis is affected both by the number of rotatable bonds and the calculation method used, among which there are not only quantum mechanical or molecular mechanical procedures, but also Monte Carlo (random conformational search) [21] and molecular dynamics (based on the time and temperature-dependent motion of a molecule) [22] techniques.

1.4.4 Pharmacophore identification

The interaction between a pharmacologically active molecule and its macromolecular target (enzyme or receptor) requires the presence of a highly specific binding pocket in the latter one; drugs belonging to the same pharmacological class have to present identical functional chemical groups in positions sterically important for the binding to be effective, that is a common *pharmacophore* has to be present.

Once identified the functional elements to superimpose, molecular modelling techniques (*atom-by-atom* and *molecular field-based* methods)[23,24,25] enable the identification of the pharmacophore, thus also providing information about ligand-receptor interactions.

Superimposition of ligand molecules has to take into account the influence of hydrogen bonds between them and their target as well as the presence of ring systems and of rigid or semi-rigid moieties that could ease and improve both results and yield of the whole searching procedure.

1.4.5 QSAR analysis: CoMFA and QUASAR

Among all computational means supporting the chemistry side of indirect drug design projects, three-dimensional quantitative structure-activity relationship techniques (3D QSAR) are the most powerful. The classical Hansch QSAR is based on two-dimensional structure (2D), which is, however, not a valid description of a real chemical molecule defined by a three-dimensional (3D) arrangement of atoms. Therefore, in the recent decade 2D techniques have been supplemented by 3D methods enabling an efficient comparison of three-dimensional molecular structures. In particular, it is the Comparative Molecular Field Analysis (CoMFA) that both brought a new 3D strategy into molecular design and is still the most frequently used method for determining 3D QSAR; some new techniques have been developed recently. One of the most interesting approaches is probably the self-organizing neural network (SOM),[26] used by Zupan & Gasteiger to obtain two dimensional maps of selected properties of molecular surfaces. SOM is a trained network designed to reduce the dimensionality of the input objects while preserving their topology. Such maps were used for the visualization of the interactions of individual compounds with biological receptors, but this technique forms also a basis for comparison between molecules.

Despite their distinct approaches used for carrying out the QSAR analysis, the main aim of all these techniques is to establish a correlation of the biological activities of a series of structurally and biologically characterized compounds with the spatial fingerprints of several field properties of each molecule (steric demand, lipophilicity, electrostatic interactions).

A 3D QSAR analysis usually identifies the pharmacophoric arrangement of molecular fragments in space, thus providing information for the design of the next generation of compounds with enhanced and improved biological activity.

For the 3D QSAR study of the antitubulin agents database, both CoMFA[2] and QUASAR[3] methods have been used in order to compare the resulting data, establish a reasonable correlation structure-activity and explain the relevant biological effectiveness of the arylthioindole derivates as inhibitors of tubulin polymerization.

1.4.5.1 CoMFA method

The CoMFA method considers the steric and electrostatic fields to provide all the necessary information for understanding the biochemical properties of a set of compounds. Therefore, the molecules are located in a cubic grid and the interaction energies between the molecule and a defined probe are calculated for each point of the grid. Generally, only the steric and the electrostatic potentials (expressed in a Lennard-Jones function and a Coulomb function, respectively) are used within a CoMFA evaluation, with the disadvantage of not taking into account hydrophobic and entropic contributions governing many binding effects.

As with any QSAR method, there are crucial and important points to be considered for the analysis to provide realistic and reliable results:

- Comparable biological activities of all compounds studied
- A range of at least three orders of magnitude among all biological activity values of the compounds analyzed
- Realistic alignment of all molecules to a template structure in order to superimpose them in their putative bioactive conformation

Once all the above conditions have been satisfied, the CoMFA analysis constructs a lattice surrounding all the molecules and calculates the electrostatic and van der Waals energies with a selected probe at each intersection of the lattice, whose default spacing is equal to 2Å.

The relationship between the biological activities and the generated interaction fields are evaluated by the multivariate statistical technique of partial least square (PLS). In order to overcome the risk of chance correlation, the CoMFA study has to provide a final Q^2 value equal to or above 0.5, that is usually considered as statistically significant and acceptable. The final results of a CoMFA analysis can be visually inspected as steric and electrostatic contributions are contoured separately and shown in different colours. Although CoMFA contour maps can not be considered to mirror exactly the corresponding attributes of the target protein, if the ligand alignment is based on the geometry of protein-bound conformations, the steric and electrostatic contour maps may reflect at some extent the steric and electrostatic environment of the binding site.

1.4.5.2 QUASAR method

The QUASAR method is not based on properties calculated within a lattice, as in all CoMFA related approaches, but relies on properties calculated for discrete locations in the space at or near the union surface of the active ligands, thus analyzing experimental data and building numerical models of the data for prediction and interpretation purposes.

In order to build numerical models that depend on molecules it is necessary to construct a numerical representation of a molecule. To this end, molecules are described with n-vectors of numbers called molecular descriptors, or just descriptors, for the purposes of model building. A descriptor can be any quantity that depends on the molecule, such as molecular weight, van der Waals volume, dipole moment and number of carbon atoms.

A model is a numerical formula that is parameterized using the training set. This numerical formula accepts, as input, a vector of molecular descriptors which is used to predict an experimental result. Any number of formulae can be used to predict experimental results. The QUASAR method can run calculations by means of two distinct models:

• *Linear Regression Models*. A linear model is a model in which the experimental result is expressed as a linear combination of the descriptors plus a constant. The parameters, or coefficients, for the model are determined in such a way that the mean squared error between the training set's experimental results and the model's results is minimized.

• *Binary Models*. A binary model assumes that the experimental result is a binary value (1 or 0) representing Pass/Fail or Active/Inactive. The binary model is statistical in that it uses the training set to estimate the probability that a new molecule will have an experimental result of 1.

The quality of a model is reported in statistical terms (e.g., correlation coefficient or percent accurate). This means that an underlying assumption of model building is the fact that the training set is representative, i.e. sufficiently large sample of molecules. More precisely, the methods assume that the training set can be modeled with independent and identically distributed random variables. Unfortunately, it is impossible to determine empirically whether or not this is true for a given data set.

Once a model has been built, it is necessary to assess its predictiveness. A model that is capable of reproducing the training set experimental results accurately may not perform well on new molecules. This phenomenon is called overfitting and generally indicates that there are too

11

many parameters in the model or that the training set is not large enough. In order to remove outliers (training set molecules with very large model errors), the cross-validation procedure allows portions of the training set to be left out of the model building and subsequently used to simulate "new" molecules. Once a satisfactory model has been built, it can be used to predict experimental data for new molecules or analyzed in order to understand the important factors determining the experimental value.

1.4.6 Design of combinatorial libraries

Following the accurate analysis of the best docked structures and/or the more relevant alignments of well known compounds acting as inhibitors of the target protein, the more interesting molecules can be selected, simplified and reduced to a scaffold, which contains only the main patterns or moiety supposed to be crucial for the biological activity of the corresponding class of compounds previously docked and/or aligned.

By means of designing tools, the scaffold is added of one or more ports, i.e. specific sites for the addition of various fragments included in databases of distinct chemical groups. New databases with different fragments can be created according to the specific substitutions required for the inhibitor-protein complex and subsequently used for the building of combinatorial libraries.

The intuition of the researcher, the knowledge of the interactions involved in the ligand-receptor binding and the use of comparative computational methods can successfully guide either the location of the ports on the scaffold or the selection of the more reasonable fragments to insert onto the scaffold for a stable and efficient binding of the final compounds to the active site.

Computational applications may allow to choose between two distinct methods for the addition of the R-groups onto the scaffold, i.e. the symmetric and the asymmetric combination of the fragments, respectively. The former leads to the creation of small libraries that can be easily and quickly checked and analysed by molecular docking; with the latter, bigger and more heterogeneous databases can be obtained but they require a longer time consuming virtual screening.

1.4.7 Comparative protein modelling

The 3D structure of the target protein derived from crystal or NMR data and further refined is essential to identify distinct properties of the macromolecule or to analyze potential ligand-protein interactions. As the number of protein sequences is larger than that of solved 3D structures and, consequently, the amount of collected and stored data is impressive, urgent is the need to compare the sequences and observe similarities or differences. Among the most important databases there are SWISS-PROT[27] and Protein

Data Bank[28] (http://www.rcsb.org): the former provides protein sequences and relative highly accurate information, in the latter both protein and DNA data are stored. Despite the endless growth of data, only primary structure information are available from most of the current databases; therefore, alignment, knowledge-based and comparative modelling techniques are required to create a 3D protein model.

The genetic evolution has given origin to families of proteins containing similar amino acid sequences and 3D structures and, if an identical ancestor is identifiable, these are called *homologous* proteins.

The 3D structure of homologous proteins is highly conserved as it is crucial for the specific protein functionality. Dissimilarities are easy to be observed on the protein surface or in the loop regions, while residues in the core of the macromolecule as well as those of the secondary structure are conserved and do not vary frequently.

The high conservation of the overall conformantion of homologous proteins has led to the development of the comparative protein modelling (homology modelling) approach. Indeed, if one sequence is homologous to another whose 3D structure is known, homology modelling can predict the structure of the unknown protein by using all the available information about resolved proteins. Initially, the sequence of the new protein is compared by alignment methods with those structurally known present in a database.

When homologous proteins to the unknown are identified, these are used as templates for the structural prediction of the new protein, the success of which is strictly related to the sequence identity percentage as considering only structural homology cannot guarantee the prediction of a perfect model.[29,30,31,32] It is agreed that sequences containing more than 100 residues and being more than 25% identical (that is, identical amino acids being located in the same positions within the sequences) are likely related,[33] and that a reliable model is obtained if the sequence identity is above 50% (in fact, significant differences are noticed when this value is about 20% and, in this case, *threading methods*[34,35,36] are more indicated than homology modelling even if at the expense of the overall accuracy).[37]

In short, the comparative modelling technique is characterized by the identification of proteins related to the one under study (the most common methods for this purpose are FASTA[38] and BLAST[39]); once both structurally conserved regions (SCRs, mainly secondary structural elements - containing conserved residues in hydrophobic cores, salt bridges or in active sites, hydrogen bonds and disulfide bridges -, the accurate assignment of which depends on the number of available crystal structures of

homologous proteins as no reasonable comparison is possible in presence of only one template) [40] and structurally variable regions (SVRs, preferably located in loop regions, where addition or deletion of residues is more likely to occur, and little or no sequence homology is observed) [29,30,31] are detected, the former are used to guide the alignment of the sequence of the unknown protein with those of the template(s) by means of multiple alignment methods, for example CLUSTALW.[41]

Template(s)'s coordinates are used for the construction of the homology model's SCRs, to which SVRs creation and side chain modelling follow, the latter being even more challenging due to the considerable degrees of freedom and the number of possible low energy conformations. The model obtained is then subjected to structural refinement by energy minimization and molecular dynamics because of the likely presence of steric strains and overlaps as well as bad van der Waals contacts.

Conclusions followed the fourth CASP (Critical Assessment of techniques for protein Structure Prediction)[42] competition have stated the invaluable importance and role of homology modelling in producing useful and reliable models for the continuously growing sequence databases.[43]

1.4.8 Molecular Docking

Virtual screening can help and improve the selection of the most promising compounds from an electronic database for experimental screening; this useful computational method is carried out by analyzing database for molecules fitting either a pharmacophore or a 3D structure of the target protein or macromolecule. The present study has used the latter approach for the evaluation of all databases related to antitubulin and anti-HCV agents, Bcl2 and Mdm2/p53 projects, respectively.

As is very difficult to reproduce the conformational space accessible to a protein and this always requires an approximation, docking methods are generally classified into three categories depending on the approximation level:

- Rigid body docking (both protein and ligand are considered as rigid bodies)
- Semi-flexible docking (only the ligand is treated as flexible)
- Fully flexible docking (both protein and ligand are considered as flexible molecules)

Furthermore, it is worth noticing that currently available docking tools are different in terms of description of molecular interactions, algorithms used to create ligand structures and average running time per molecule. In addition, the algorithms are usually divided into those with deterministic and those with stochastic approaches; the former are reproducible, the latter are not completely reproducible as they imply a random factor.

Taking into account all the above considerations about molecular docking, distinct softwares implementing either a genetic algorithm based search method or other two different search protocols, known as Simulated Annealing and Tabu Search, have been chosen and used for the present study. The virtual screening by means of distinct algorithms and the accurate comparison of the deriving results have helped the assessment of the reliability of the final docked structures and, therefore, provided notable information for future molecular screening (see Fig.1.2).

1.4.8.1 Semi-flexible docking

Programs for flexibly docking ligands into binding sites may use incremental construction algorithm that builds the ligands in the binding site. It starts by extracting a core fragment from the ligand. The algorithm is dependent on selection of an appropriate base fragment, requiring one that makes enough specific contacts with the protein that a definitive preference for binding orientation can be determined.

Bond lengths and angles are kept invariant, using the values of the input ligand. The core is used as the base to which low energy fragment conformers are added, with conformers based on a statistical evaluation of fragments usually selected from the Cambridge Structural Database. In more detail, the base fragment (the ligand core) is automatically selected and is placed into the active site; the remainder of the ligand is built up incrementally from other fragments that are added in all possible conformations to all placements found in the previous iteration, but only the best placements are taken on to the next construction step. The conformational flexibility of the ligand is included by generating multiple conformations for each fragment and including all in the ligand building steps. Placement of the ligand is scored on the basis of protein-ligand interactions, as shape alone is a weak descriptor, especially for small or flexible ligands. Finally, the binding energy is estimated, and placements are ranked.

Distinct programs can search for favorable binding configurations between a small, flexible ligand and a rigid macromolecular target, which is generally a protein, differently. Searching is conducted within a specified 3D docking box, using one of the three available search protocols (genetic algorithm, simulated annealing or tabu search) and a molecular mechanics forcefield. All methods seek to optimize both purely spatial contacts as well as electrostatic interactions. The search can be restricted (and thereby reduced) by imposing angle, torsion and distance constraints or restraints on the ligand. During the calculations, the ligand molecule takes on conformations from the search trajectory.

1.4.8.2 Search protocols

- *Genetic Algorithm* – this computer program mimics the evolutionary process by manipulating data structures named chromosomes. Each of these chromosomes encodes a possible ligand-protein complex conformation and is assigned a fitness score based on the quality of its solution in terms of ligand-protein interactions.

An initial parent population of chromosomes, previously randomly generated, is subjected to cross-over and mutation repeatedly, until children chromosomes replace the least fit members of the population. Cross-over and mutation are two major genetic operators. The former requires two parents and produce two children, thus combining features from two different chromosomes into one; the latter needs one parent to produce one child and, therefore, introduces random perturbations. The evolutionary pressure into the algorithms derives from the random selection of only the best parent chromosomes from the existing population, thus ensuring that, over the time, the population move towards an optimal solution, i.e. to the correct, specific and efficient binding mode.

- *Simulated Annealing* - Simulated annealing, a global optimization technique based on the Monte Carlo method, is a special molecular dynamics simulation in which the system is cooled down at regular time intervals by decreasing the simulation temperature and thus allowing the system to be trapped in the nearest local conformation. Main disadvantages of this simulation are that the final results depend on the initial placement of the ligand and that the algorithm does not explore the solution space exhaustively because, in a Monte Carlo search, random movements sample the conformational space. In fact, the algorithm can calculate various states of a configuration space by generating small random changes in the current state and then accepting or rejecting each new state according to the Metropolis criterion [Metropolis 1953]. According to this criterion, moves that decrease the energy of the system are always accepted, while moves that increase the energy of the system are accepted according to a probability expression (Boltzmann equation). The probability of acceptance P is given as:

$$P = e^{(-\Delta E/kT)}$$

Where ΔE is the difference in energy from the previous step, T is the absolute temperature in Kelvin, and k is the Boltzmann constant. This means, the higher the temperature of the cycle, the higher the probability that the new state is accepted.

- **_Tabu Search_** - Tabu search is a stochastic searching algorithm that maintains a list of previously visited conformations. By preventing the search from revisiting these solutions, this procedure effectively encourages the search into unvisited areas of the search space. Only one current solution is maintained during the course of a search; at the beginning, the current solution is derived by randomizing the position and orientation of the ligand within a box around the active site. A mutation-like procedure results in a defined number of moves that are ranked according to a scoring function at the end of the calculation.

A new conformation is compared to the conformations in the list by calculating the root mean square (RMS) deviation between the Cartesian coordinates of the new conformation and those of every entry in the list. If the RMS deviation is 0.75Å or less, the conformations are considered to be the same and, therefore, the move is tabu. The highest ranked move is always accepted as new current solution if its energy is lower than the lowest energy obtained so far, thus replacing the previous best solution. Otherwise, the algorithm selects the best non-tabu move. If neither criterion can be chosen, the algorithm terminates and if a new current solution can be identified, this is added to the tabu list.

It is worth noticing a relevant difference between genetic and tabu search algorithms, respectively: while the former usually converges quickly at the close proximity of a global minimum but can be trapped in local minima, a tabu list helps in avoiding this drawback.

1.4.9 Optimization and energy minimization of the protein: Minimization and Molecular Dynamics Simulation

The present study has been focused on the optimization and model refinement of two proteins, Bcl-2 homology model and tubulin, respectively, as protein models derived from either comparative modelling or crystal structures need further refinement.

In fact, during the generation of proteins the loop and side chain conformations are usually chosen arbitrarily and, consequently, these do not correspond to energetically reasonable structures.

In order to remove the internal strain deriving from the crystal packing forces or to remove close contacts between hydrogen atoms or amino acid residues added to the crystal coordinates after structure determination, also crystal structures must be relaxed.

The *minimization algorithms* used for the geometry optimization usually find only the local minimum on the potential energy surface closest to the initial coordinates. Despite the relaxation of a crystal structure is often a straightforward procedure, sometimes crystal coordinates present several unfavourable atomic interactions that cause large Initial forces resulting in artificial movements away from the original structure at the beginning of the minimization process. These deviations can be avoided by allowing the protein model to relax gradually or by assigning tethering forces to all heavy atoms in the first stage of the minimization in order to fix atomic coordinates on predefined positions.

The former approach has been chosen for the geometry optimization of all the four projects undertaken in this study. In order to make the energy minimization processes accurate but quick, it has been avoided the treatment of solvated systems, known as very accurate but also long time consuming procedures.

Molecular Dynamics Simulation is a valuable tool for the investigation of the conformational behaviour of loop and side chains of protein models, especially if these are generated by comparative modelling techniques. The relaxed model obtained as result of a minimization process can be used as starting point for molecular dynamics simulation, being aware that the minimum structure found by minimization algorithms represents only one local minimum.
In order to search the conformational space more thoroughly, molecular dynamics simulation integrates the classical equations of motion over a period of time for the molecular system. The resulting trajectory is used to compute the average and time-dependent properties of the system.
Although molecular dynamics simulation has dramatically improved the understanding the dynamic process in protein at the atomic levels, limitations and problems arise with increasing size and associated number of degrees of freedom of molecular systems as well as with the choice of the more suitable simulation temperature.

18

Docking Methods

Incremental constructional
- Fragment decomposition
- Selection of a base fragment
- Adding the first fragment
- Reading the preferred torsional angle values
- Selecting the best
- Adding the second fragment
- Termination

Genetic Algorithm
- Selection of the operators
- Creation of the initial population
- Applying genetic operator
- Selecting Children chromosomes
- Children replacement of the least fit individuals
- Repeating the evolutionary process

Simulated Annealing
- Molecular Dynamics Simulation

Tabu Search
- Creating an initial solution
- Evaluating the current solution
- Updating the tabu list
- Generating and ranking x possible moves
- Examination of the ranked moves
- Acceptance or rejection of the move
- Repeating the process

Assessment of the reliability of the final best docking

Fig.1.2 – Molecular Docking methods.

Chapter 2

Introduction

2.1 Microtubules

As long, filamentous, tube-shaped protein polymers of the cytoskeleton, microtubules play a vital role in all eukaryotic cells: the development and maintenance of cell shape, intracellular transport of vesicles, mitochondria and other components, cell signalling as well as cell division and mitosis are all processes depending on the presence of microtubules. Composed of α-tubulin and β-tubulin heterodimers (Fig.2.1), microtubules are highly dynamic polymers, the polymerization of which is regulated both spatially and temporally.
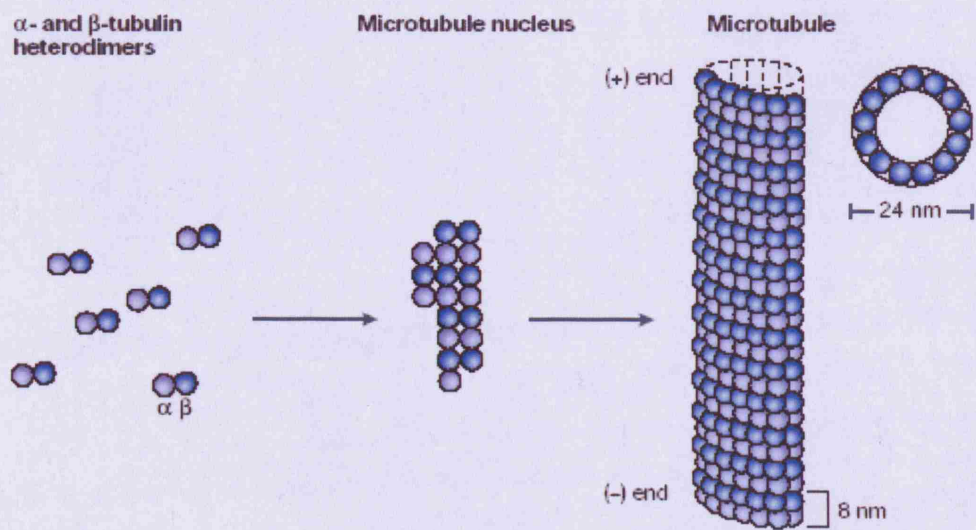
Fig.2.1 – Polymerization of microtubules.[14]

Binding of various regulatory proteins (including microtubule associated proteins, MAPs, such as dynein and kinesin motor proteins, as well as many microtubule-regulatory proteins, such as survivin, stathmin);[8-13]expression of different tubulin isotypes with distinct functions (human tubulin isotypes - 6 forms of α-tubulin and 7 forms of β–tubulin - at diverse levels in different cells and tissue) post-translational modifications of tubulin (polyglutamylation, polyglycylation, phosphorylation, acetylation, detyrosination/tyrosination and removal of the penultimate glutamic-acid residue of α-tubulins):[6,7] these are the ways by which microtubules acquire functional diversities, characteristic in certain cancer cells and sometimes developing drug resistance.

2.1.1 Microtubules as target

The fundamental role in mitosis and cell division make microtubules and their dynamics attractive targets for a chemically various group of anticancer drugs, referred to as antimitotic drugs (Table 2.1). The clinical success of these compounds has soon led to the conclusion that microtubules represents one of the best cancer targets currently known as well as the favourite one of natural, toxic molecules produced by diverse plants and animals. Indeed, *Vinca* alkaloids, isolated more than 40 years ago from periwinkle leaves (*Catharanthus Roseus*) are among the earliest identified antimitotic agents.

The need of orally available analogues, devoid of toxicity (monitored by therapeutic index values) and not developing resistance, has strongly motivated the discovery of new antimitotic drugs and the improvement of the existing ones.

Table 2.1 - Antimitotic drugs, their diverse binding sites on tubulin and their stage in clinical development.[14]

Binding domain	Related drugs or analogues	Therapeutic uses	Stage of clinical development
Vinca domain	Vinblastine (Velban)	Hodgkin's disease, testicular germ-cell cancer	In clinical use; 22 combination trials in progress
	Vincristine (Oncovin)	Leukaemia, lymphomas	In clinical use; 108 combination trials in progress
	Vinorelbine (Navelbine)	Solid tumours, lymphomas, lung cancer	In clinical use; 29 Phase I–III clinical trials in progress (single and combination)
	Vinflunine	Bladder, non-small-cell lung cancer, breast cancer	Phase III
	Cryptophycin 52	Solid tumours	Phase III finished
	Halichondrins (such as E7389)	–	Phase I
	Dolastatins (such as TZT-1027)	Potential vascular-targeting agent	Phase I; Phase II completed
	Hemiasterlins (such as HTI-286)	–	Phase I
Colchicine domain	Colchicine	Non-neoplastic diseases (gout, familial Mediterranean fever)	Appears to have failed trials, presumably because of toxicity
	Combretastatins (AVE8062A, CA-1-P, CA-4-P, N-acetylcolchicinol-O-phosphate, ZD6126)	Potential vascular-targeting agent	Phase I, II
	2-Methoxyestradiol	–	Phase I
	Methoxybenzene-sulphonamide (such as ABT-751, E7010)	Solid tumours	Phase I, II
Taxane site	Paclitaxel (Taxol), TL00139 and other analogues of paclitaxel	Ovarian, breast and lung tumours, Kaposi's sarcoma; trials with numerous other tumours	In clinical use; 207 Phase I–III trials in the United States; TL00139 is in Phase I trials
	Docetaxel (Taxotere)	Prostate, brain and lung tumours	8 trials in the United States (Phases I–III)
	Epothilones (such as BMS-247550, epothilones B and D)	Paclitaxel-resistant tumours	Phases I–III
	Discodermolide	–	Phase I
Other microtubule binding sites	Estramustine	Prostate	Phases I–III, in numerous combinations with taxanes, epothilones and Vinca alkaloids

2.1.2 Microtubules: polymerization dynamics

The complex polymerization dynamics of microtubules implies a *nucleation-elongation mechanism*: "the slow formation of a short microtubule 'nucleus' is followed by rapid elongation of the microtubule at its ends by the reversible, non-covalent addition of tubulin dimers[14]." The hydrolysis of GTP when tubulin with bound GTP adds to the microtubule ends provides the required energy for the two kinds of non-equilibrium dynamics characterizing and regulating microtubules functions.[15-18]

The former, called 'dynamic instability', is characterized by the switch between phases of growth and shortening occurring at single microtubule ends,[19] one of which, the plus end, grows and shortens faster and more extensively than the minus end. Four main

variables characterize and alternate each other in dynamic instability: the rate of microtubule growth; the rate of shortening; the frequency of transition from the growth or paused state to shortening (called a 'catastrophe'); and the frequency of transition from shortening to growth or pause (called a 'rescue').

'Treadmilling' is the latter dynamic behaviour, identified as growth at one microtubule end, which is in turn balanced by shortening at the opposite end[20-24]. During this process, tubulin subunits move from the plus end of the microtubule to the minus end. Tubulin isotype composition of a microtubule population, degree of posttranslational modification of tubulin and the activity of regulatory proteins may affect and explain the mechanism by which a microtubule population preferably shows one of the two behaviours.[25] Microtubule polymerization dynamics involves the gain and loss of a short region of tubulin–GTP (polymerized into straight protofilaments) or tubulin–GDP (placed into the microtubule core in a strained conformation) plus inorganic phosphate (Pi) at the two microtubule ends, known as GTP cap. It has been observed that presence of GTP or GDP–Pi cap allows for microtubules growth,[26,27] whereas loss of the cap causes rapid shortening of the microtubule end.[28,29] As host molecules regulating microtubule dynamics target assembled microtubule, antimitotic drugs are likely to function mimicking natural regulators controlling cellular dynamics processes.[8,25,30-36]

2.1.3 Microtubule dynamics and their role in mitosis

Impressive advances in visualization of dynamic mitotic-spindle microtubules have shown how rapidly mitosis occurs and, therefore, how important highly dynamic microtubules in the spindle are during all mitotic phases.[37-39]

Drugs such as paclitaxel (Taxol) and *Vinca* alkaloids suppress microtubule dynamics in order to block mitosis and kill tumour cells, more sensitive to antimitotic drugs because of their higher rate of division (accordingly to their own tissue specifity) compared with normal cells.

2.1.4 Antimitotic drugs

Acting on the rapid polymerization dynamics of spindle microtubules, antimitotic agents inhibit cell proliferation. Two are the main groups of microtubule-targeted antimitotic drugs (see Table 2.1 for a more detailed classification). The former group,

referred to as the *microtubule-destabilizing agents*, inhibits microtubule polymerization at high concentrations and includes: *Vinca* alkaloids (vinblastine, vincristine, vinorelbine, vindesine and vinflunine), cryptophycins, halichondrins, estramustine, colchicine and combretastatins — under clinical investigation or clinically used for anticancer therapy (Table1). In addition, this group comprises also compounds that have not been subjected to clinical development for cancer therapy, such as the anti-tussive noscapine,[40] maytansine, rhizoxin, spongistatins, podophyllotoxin, steganacins and curacins;[41] several herbicides that inhibit microtubule polymerization;[42] antifungal and antihelmintic agents;[43] and some psychoactive drugs.[44–46] The latter group, the *microtubule-stabilizing agents*, promote microtubule polymerization and include paclitaxel (the first agent to be identified in this class), docetaxel (Taxotere), the epothilones, discodermolide, the eleutherobins, sarcodictyins, laulimalide, rhazinalam, and certain steroids and polyisoprenyl benzophenones.[47,48]

The reason of this classification relies on the observation that generally high concentrations are required for antimitotic drugs to increase or decrease microtubule polymerization, whereas 10–100-fold lower concentrations of the same agents are sufficient to suppress microtubule dynamics, thus leading to kinetic stabilization of the microtubules (slowing or blocking of mitosis at the metaphase–anaphase transition) and induction of apoptosis, without changes in their mass.

Although their actions on tubulin and microtubules reveal important differences, the two classes of antimitotic drugs show similar antimitotic mechanisms (see also paragraph 2.2 for more detailed information regarding antimitotic agents).

2.1.5 Specific drug mechanisms of action

The **Vinca** *alkaloids.* Forty years ago the *Vinca* alkaloids have been successfully introduced into the clinic due to the active principles vinblastine and vincristine, but since the seventeenth century the leaves of the periwinkle plant have been used for their medicinal properties.

It was only in the late 1950s that two independent groups at Eli Lilly Research Laboratories and at the University of Western Ontario[49,50] discovered their antimitotic activity and peculiar chemotherapeutic potential against both childhood haematological and solid malignancies, and adult haematological malignancies. Subsequently, novel semi-synthetic analogues (vindesine, vinorelbine and vinflunine) have been developed

trying to reduce peripheral neuropathy and reversible myelosuppression, main side effects of this class of drugs.[51]

Vinca alkaloids specifically depolymerize microtubules and disrupt mitotic spindles at high concentrations leaving the dividing cancer cells blocked in mitosis with condensed chromosomes.

Paclitaxel and related drugs. The late twentieth century saw the discovery of paclitaxel and its semisynthetic analogue docetaxel. Slow was the development of paclitaxel, a complex molecule isolated from the bark of the yew tree in 1967 by Monroe Wall and Mansukh Wani,[52] until, in 1979, Peter Schiff and Susan Horwitz discovered that, contrary to the *Vinca* alkaloids, paclitaxel promoted microtubule polymerization.[53] Limited supplies of the natural compound halted its clinical use in anticancer regimens, but, following the achievement of its semi-synthesis,[54] it was approved for clinical use to treat breast and ovarian cancer, non-small-cell lung cancer and Kaposi's sarcoma. Like the *Vinca* alkaloids, neurotoxicity and myelosuppression are the main side effects.[55,56] The 3D crystal structure of tubulin complexed with paclitaxel,[57] has shed light on the binding site of this molecule. By diffusion through small openings in the microtubule or fluctuations of the microtubule lattice,[58] paclitaxel enters the β-subunit and binds the inside surface of the microtubule, thus stabilizing the microtubule and increasing microtubule polymerization, probably by means of a conformational change in the tubulin that may improve its affinity for neighbouring tubulin molecules.[58] Following the clinical success of the taxanes, other drugs able to increase microtubule polymerization have been identified, such as the epothilones, discodermolide, the sarcodictyins, eleutherobin and laulimalide. Interstingly, epothilones, discodermolide, eleutherobins and sarcodictyins compete with paclitaxel for binding to microtubules at or near the taxane site, while others, such as laulimalide, seem to bind to unique sites on microtubules.[59]

Colchicine. Clinically used in the treatment of gout, neither colchicine nor compounds binding to the colchicine site on tubulin have yet been significantly used in cancer treatment because of their potent toxicity (some of those extracted from natural sources are now under clinical trials for anticancer therapies). Like *Vinca* alkaloids, high concentrations of colchicines lead to microtubules depolymerization and low concentrations stabilize microtubule dynamics.

Once bound to soluble tubulin, colchicine stimulates tubulin to slowly change conformationin in order to form a reversible tubulin–colchicine complex,[16,60] that is able to copolymerize into microtubules ends[61] and to slow further tubulin addition as well as to regulate the gain and loss of the GTP or GDP–Pi cap.

In conclusion, although differences between the effects at high concentrations of the above discussed two classes of agents are evident, almost all of the microtubule-targeted antimitotic drugs achieve their anticancer effect likely by disrupting the normal microtubules dynamics, thus ultimately inducing apoptosis in cancer cells.

2.1.6 Antivascular activity

Anticancer treatments have been recently attracted by a new potential target, the tumour vasculature, because of its easy access to drugs and its vital role for cancer cells, the death of which is guaranteed if oxygen and nutrients are not supplied through the blood. Research has been focused on agents inhibiting the process of angiogenesis (cyclophosphamide is an example) and, more recently, on microtubule-targeted compounds able to destroy existing tumour vasculature.[62] In the late 1990s, the combretastatins and N-acetylcolchicinol-O-phosphate (both bind to the colchicine domain on tubulin), have been investigated as antivascular agents and some of them are in clinical trials (Table 2.1).

Once administered, these drugs cause disappearance of small blood vessels, slow blood flow, aggregation of red blood cells, haemorrhages from peripheral tumour vessels, vascular permeability increase, thus leading tumour cells to necrosis[62] without damaging normal tissues.[63]

Furthermore, diverse microtubule-targeted agents, including colchicine, vincristine and vinblastine, promote similar effects at high concentrations *in vitro* and damage tumour vasculature in animals.[14]

Interestingly, besides combretastatins with antivascular acitivity, two novel series of modified combretastatin derivates showing either inhibition of tubulin polymerization[64a] or induction of apoptosis[64b] have been recently reported. The former group of combretastatin analogues is characterized by a naphthalene replacing the 3,4,5-trimethoxyphenyl moiety, resulting in a significant decrease of cytotoxicity and

28

G2/M arrest of the cell cycle in human cancer cells.[64a] A five-membered heterocyclic (isoxazoline or isoxazole) or a six-membered ring (pyridine or benzene) instead replace the alkenyl motif of CA-4 in the latter group of combretastatin derivates producing potent and effective apoptosis-inducing agents.[64b]

2.1.7 Tumour sensitivity and resistance

Tissue specificity and development of drug resistance are due to factors influencing both the cells themselves and the pharmacological accessibility of the microtubule-targeted drugs to the tumour cells.[65] Overexpression of a class of membrane transporter proteins, referred to as ABC-transporters (ATP-dependent drug efflux pumps or ATP-binding cassettes) are often responsible for failure or resistance to chemotherapy with antimitotic agents by decreasing the intracellular drug concentration and leading to cross-resistance (multidrug resistance (MDR)) to structurally distinct molecules, e.g. paclitaxel and *Vinca* alkaloids. Efforts are concentrated on the identification of P-glycoprotein inhibitors, as this was the first to be discovered, and on the development of microtubule-targeted drugs resulting "unattractive" for these pumps.[66-69] Expression of regulatory proteins, post-translational modifications of tubulin and expression of different tubulin isotypes are some of the host determinants able to control microtubule-polymer levels and dynamics, thus conferring resistance or insensitivity to antimitotic drugs,[65,69-76] but in the same time offering themselves as unexplored potential targets involving microtubules for the design of novel anticancer drugs.[14]

2.1.8 Synergy of microtubule drugs

As many and diverse are the sites on tubulin and microtubules to which antimitotic agents bind, combinations of drugs stabilizing microtubule dynamics by different mechanisms are supposed to synergically improve efficacy and reduce the side effects derived from the high blood levels of a single compound.[14]

2.1.9 Future tasks

Increasing current knowledge about microtubule-targeted agents mechanism of action and improving their clinical effectiveness are the challenging tasks of the next future. Alternative approaches could significantly reduce the toxicity common to the majority of this class of drugs, such as combination therapy, use of weak microtubule-targeted molecules as adjuvants, administration of low concentrations of drugs for longer periods, combination of microtubule-targeted agents with 'molecularly targeted' ones.[14] Overcoming MDR and neuropathy, and exploiting the specificity of tumour tissue expression to design improved combinations of drugs, are additional aims extremely important to achieve, and likely to be realized given the promising results so far obtained from the currently available antimitotic compounds.[14]

2.2 Indole: core structure for novel tubulin polymerization inhibitors

As already mentioned above, the first natural β-tubulin interacting agents so far reported are colchicine (1), vincristine (2) and vinblastatine (3), isolated from *Colchicum autumnale* and *Catharanthus roseus*, respectively[77] and binding β-tubulin in distinct active sites, thus, destabilizing microtubules and inducing apoptosis[78].

Taxus brevifolia is the natural source of another class of compounds binding a different β-tubulin active site and including the anticancer molecule paclitaxel (4, Taxol). Microtubule polymerization and stabilization are promoted by high concentration of 4, lowering the concentration of which microtubules dynamic is inhibited.[79,80]

Same mechanism of action characterize also other molecules, for examples epothilones (5, 6),[81] laulilamide (7)[82], discodermolide (8)[83] (Chart 2.1).

Chart 2.1 - Tubulin polymerization agents.
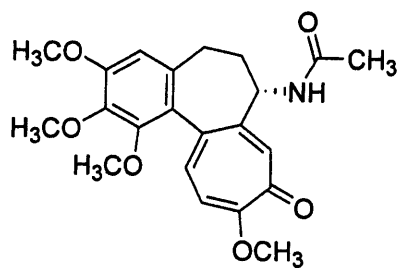
Colchicine (1)
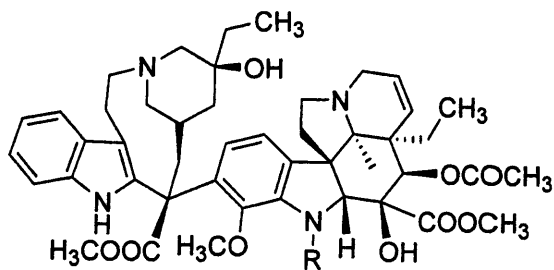
Vincristine (2) (R = CHO)
Vinblastine (3) (R = CH₃)
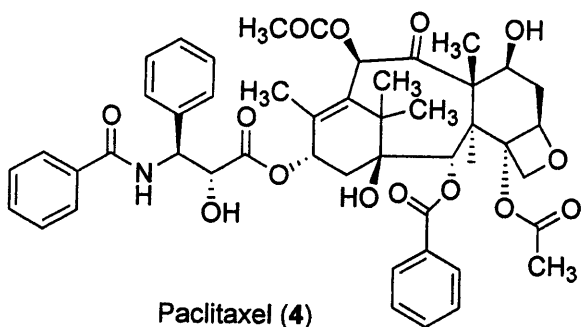
Paclitaxel (4)
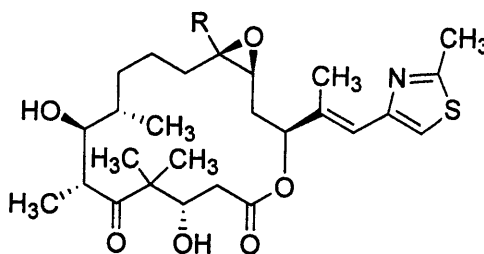
Ephotilone A (5) (R = H)
Ephotilone B (6) (R = CH₃)
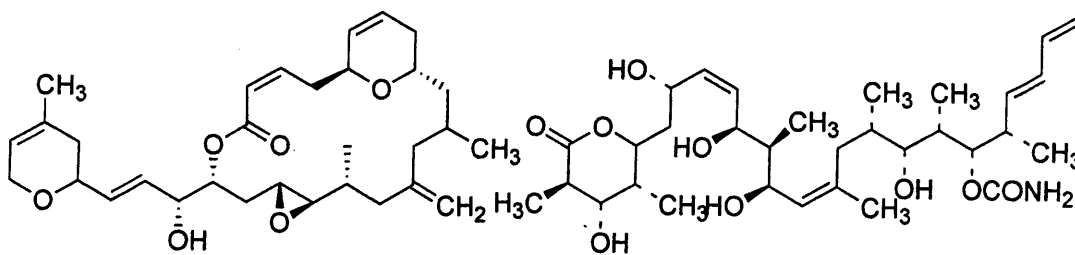
Laulimalide (7)

Discodermolide (8)

31

Hamel has classified antimitotic agents in four classes (Table 2.1).[84]

Table 2.1 - Classification of Antimitotic Agents according to Hamel.[84]

Class	Mechanism	Compounds
1	Microtubules hyperassembly and stabilization	Paclitaxel, epothilone B, discodermolide, eleutherobin
2	Binding to Cys239	t-BCEU, T138067, ottelione A
3	Microtubules destabilization, binding to the colchicine site	Combretastatin A-4, 2-methoxyestradiol, NSC-639829, mivobulin
4	Microtubules destabilization, binding to the vinblastine site	Maytansine and rhizoxine (competitive); macrocyclic polyethers, peptides, depsipeptides (non-competitive)

The first class (paclitaxel, epothilone B, discodermolide and eleutherobin) promotes hyperassembly and stabilizes microtubules by binding β-tubulin. Covalent reactions with sulfhydril residues of β-tubulin (Cys239 in particular) characterize the second class, among which there are t-BCEU, T138067 and ottelione A selectively reacting with Cys239 and inhibiting colchicine binding to tubulin.

Tubulin polymerization inhibition is observed for structurally unrelated small compounds included in the third class (for example, combretastatin A-4, 2-methoxyestradiol, NSC-639829, mivobulin act by binding to colchicine site).

Compounds interfering with vincristine and vinblastine binding sites form the fourth class to which two subclasses follow: the competitive inhibitors (maytansine and rhizoxine) and the non-competitive inhibitors ones (macrocyclic polyethers, peptides and depsipeptides, e.g. dolastatin).

Toxicity, drug resistance, complex formulations and limited bioavailability, obscuring promising results, gave great impulse to the discover of new tubulin inhibitors; in the last decade, both natural tubulin agents and tubulin polymerization inhibitors or stimulators derived from synthesis/semisynthesis have been published.[85]

2.2.1 Indoles: natural source and semisynthesis

Indole as core structure has been observed in several tubulin polymerization inhibitors derived from natural source or made by semisynthesis.

Tubulin polymerization inhibitors recognized since 1965, vincristine and vinblastine are still of clinical interest. The former (2) is used in combination in the treatment of acute lymphoblastic leukaemia and against both Hodgkin's and non Hodgkin lymphoma. The latter (3) is the treatment of choice for advanced Hodgkin's disease against germ cell cancer of the testes.[86] Both academic and industrial groups focused on the identification of more active and less cytotoxic *Vinca* alkaloid derivates, among which two semisynthetic molecules, vindesine (9) and vinorelbine (10), are employed in anti-cancer therapy[86] (Chart 2.2). A second generation alkaloid found more active then 10, vinflunine (11),[87] successfully passed Phase I and II trials as promising anticancer drug and has entered phase III studies.[88]

Secondary metabolites from *Aspergillus fumigatus* BM939 fermentation broth, Tryprostatin A (12) and B (13), showed antitubulin activity.[89] High concentration of the former acts on tubulin polymerization and M-phase of cell cycle without inhibiting the [^{3}H]colchicine or [^{3}H]vinblastine binding to β-tubulin, thus suggesting it is a distinct type of inhibitor. At a different site, compound 13 interferes in microtubule assembly.[90] *Celosia argentea* seeds are the natural source from which moroidin (14), a bicyclic peptide, was isolated and found more active (IC_{50} = 3.0 μM) than colchicine (IC_{50} = 10 μM) as tubulin polymerization inhibitor, whereas its hydrolysate was found as active as colchicine. Following these results and the hypothesis of bicyclic ring as important moiety for the activity,[91] further new indole bicyclic peptides, celogentins A-C, were extracted from *C. argentea*, identified as potent tubulin polymerization inhibitors and celogentin C was also found four times more active than moroidin.[92]

Anti-tubulin agent derived from marine sponge *Hemiasterella minor*, hemiasterlin (15) belongs to a recently discovered group of tripeptides, the stability and activity of which depend on unusual and sterically hindered amino acids.[93] Its inhibition of radiolabeled vinblastine binding to tubulin, encouraged the isolation of other two indole containing peptides, milnamide A (16) from *Auletta* cf. *constricta* sponge, and criamides A and B (17) from *Cymbastela* sp., both 16 and 17 found more potent than 23 in the cell lines analyzed.[84]

In 1991 diazonide A (**18**) and B, both binding to vinblastine active site, were extracted from *Diazona chinensis*.[94] The former showed to be very active (IC$_{50}$ = 0.33 µM) and equipotent with dolastatin[81] as tubulin polymerization inhibitor, the latter was less active than **27** against murine and human cell lines.

Chart 2.2 - Natural and semisynthetic indole sources.
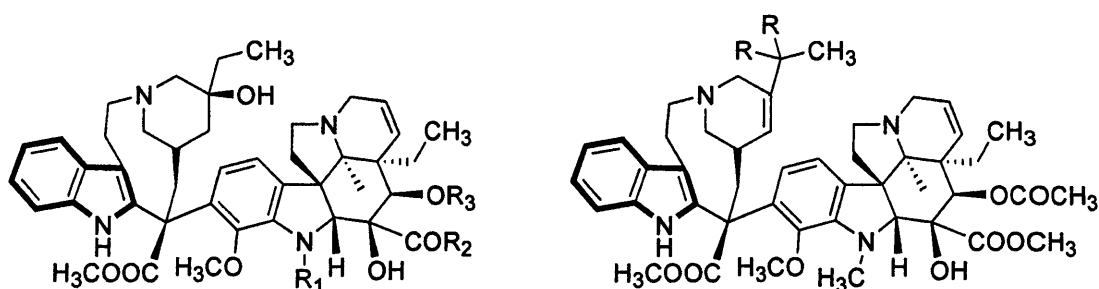
Vincristine (**2**) (R$_1$ = CHO, R$_2$ = OCH$_3$, R$_3$ = COCH$_3$)
Vinblastine (**3**) (R$_1$ = CH$_3$, R$_2$ = OCH$_3$, R$_3$ = COCH$_3$)
Vindesine (**9**) (R$_1$ = CH$_3$, R$_2$ = NH$_2$, R$_3$ = H)

Vinorelbine (**10**) (R = H)
Vinflunine (**11**) (R = F)
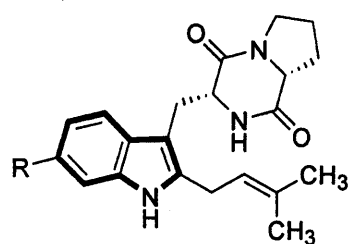
Tryprostatin A (**12**) (R = OCH$_3$)
Tryprostatin B (**13**) (R = H)
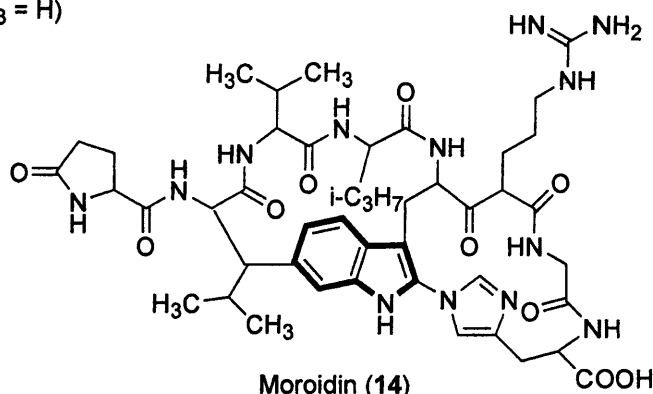
Moroidin (**14**)
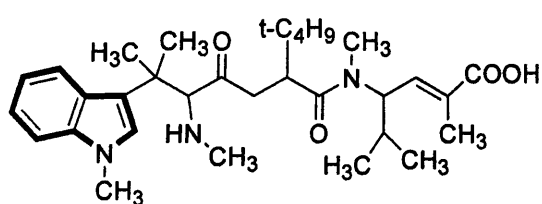
Hemiasterlin (**15**)
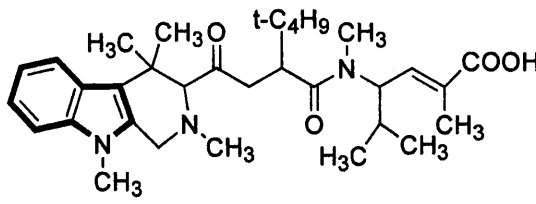
Milnamide (**16**)
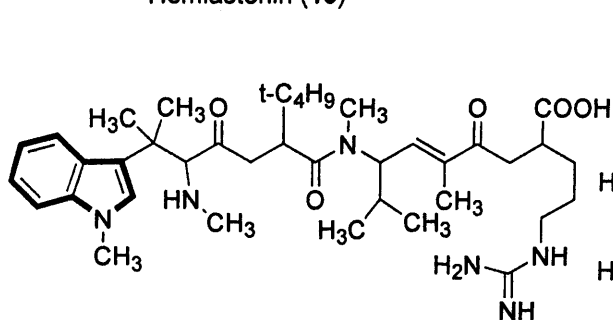
Criamide B (**17**)

Diazionamide A (**18**)

34
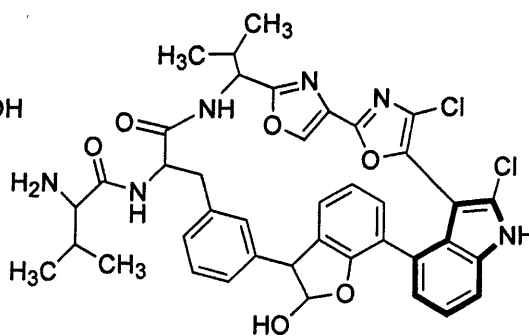
2.2.2 Indole: synthesis

Investigation about small molecules targeting tubulin has been very intense because of their easier synthesis and lower cost.[95] Therefore, diverse unrelated structures including indole derivates have been synthesized and analyzed. Examples of non-indole antitubulin agents are given in Chart 2.3: combretastatin A-4 (19),[96] T-138067 (20),[97,98] 2-methxyestardiol (21),[99] erbulozole (22),[100] NSC-639829 (23),[101] t-BCEU (24),[102] ER-344110 (25),[103] thienopyrrolizinone (26),[104]E-7010 (27),[105,98] and more recently, N1- and N4-substituted 3,5-dinitro sulfanilamides (a new class of antikinetoplastid agents that act on parasite tubulin),[106] a novel series of diarylsulfonylurea derivatives,[107]disorazol A1 (a macrocyclic polyketide compound),[108] new and improved antimitotic agents derived from Noscapine,[109] and 30,6-substituted 2-phenyl-4-quinolone-3-carboxylic acid derivatives as potential antimitotic agents.[110]

In the last decade great impulse and efforts have been given to the synthesis of novel tubulin polymerization inhibitors having indole ring as core nucleus; as several of these compounds contain a tricyclic system, they show a structural relation with colchicine or combretastatin A-4.

The most representative indole derivates - some of which shown in Chart 2.4 - are: 3-formyl-2-phenylindoles,[111,112] heterocombretastatins,[113,114] diarylindoles,[115,116,117] 2-aroylindoles,[117,118] D-24851,[119,120,121] 2-aryl-3-aroylindoles,[122,123,124] 3-aroyl- and 1-aroylindoles,[125,126,127] arylthioindoles.[128,129,130,131,132,133,134]

The amount of articles and patents so far published clearly documents the rising interest in indole derivates as potential anticancer drugs. Both the discovery of novel tubulin polymerization inhibitors and the improvement of existing classes with the support of molecular modelling, that can provide useful insights for rational structure-based drug design, are needed to develop new effective anticancer therapies.

Chart 2.3 - Non-indole antitubulin agents.
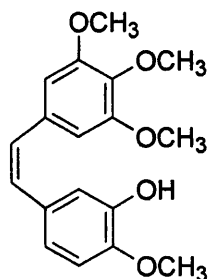
Combretastatin A-4 (**19**)
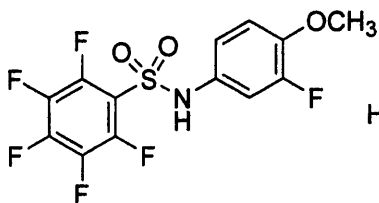
T-138067 (**20**)
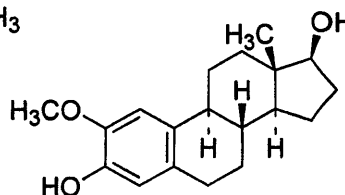
2-Methoxyestradiol (**21**)
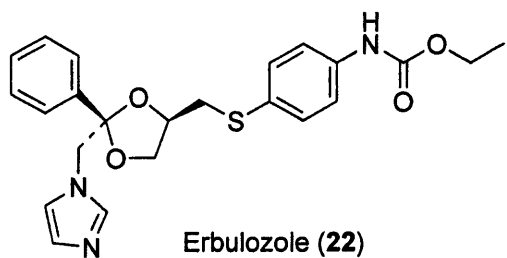
Erbulozole (**22**)
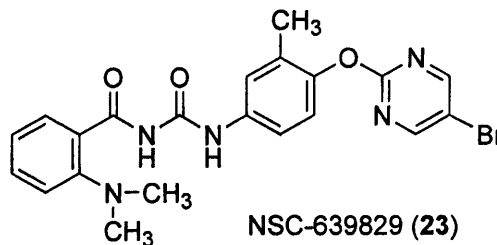
NSC-639829 (**23**)
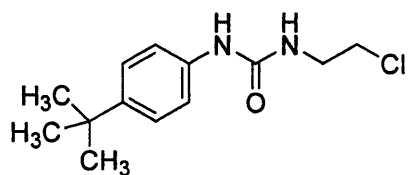
t-BCEU (**24**)
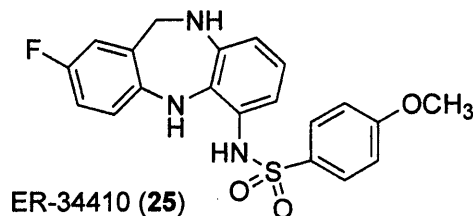
ER-34410 (**25**)
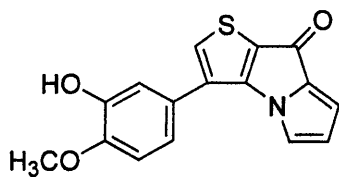
TPO (**26**)
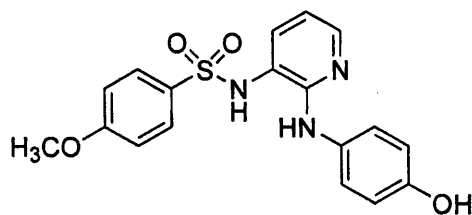
E-7010 (**27**)

Chart 2.4 - Most representative indole derivates.
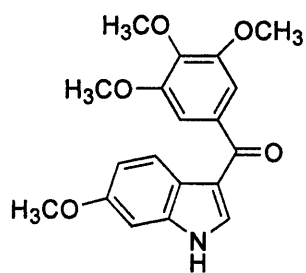
3-Aroylindoles
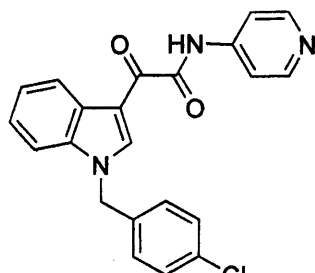
D-24851
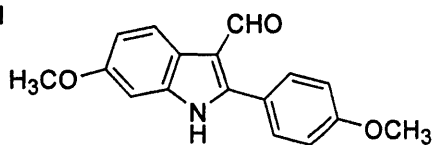
3-Formyl-2-phenylindoles
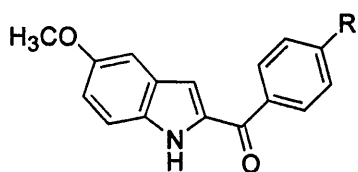
2-Aroylindoles (R=H,OCH₃)
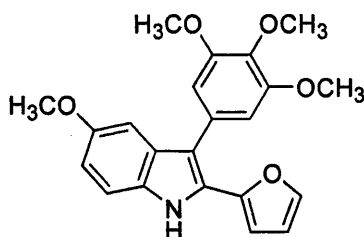
Diarylindoles
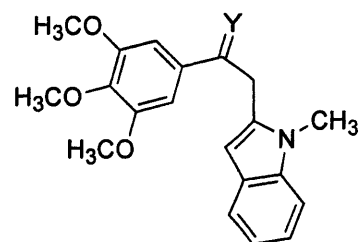
Heterocombretastatins
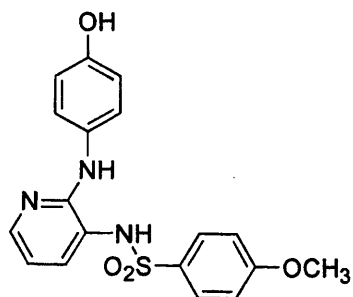
E7010
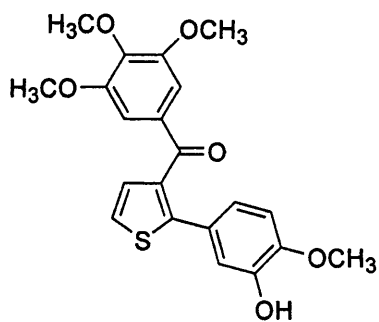
Thiophenes
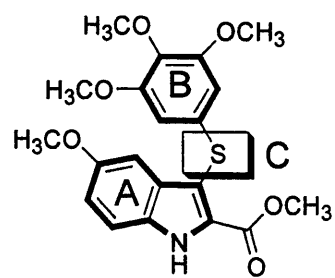
Arylthiolindoles

Aims & Objectives

Microtubules, essential components of the cytoskeleton in all eukaryotic cells, play a key role in the process of mitosis and cell division and, therefore, are among the most successful targets for anticancer therapy.[1] Microtubules, composed of α and β tubulin heterodimers, and their protein-protein dynamics are specifically targeted by chemically distinct antimitotic drugs, either currently used in the treatment of the cancer or in clinical development[1]. Furthermore, recent studies are investigating microtubule-active drugs that bind to the colchicine domain and function as vascular-targeting agents by depolymerization of microtubules of new vasculature to block and prevent blood supply to tumours.[1] Various synthetic molecules, either containing a sulfur atom[2] or an indole nucleus as core structure,[3] have shown destabilizing activity as they inhibit the binding of colchicine to tubulin. The aim of the present project is to design and synthesize novel arylthioindole derivates as antitubulin agents with the support of computational methods.

An accurate investigation of both protein-protein and protein-inhibitor interactions present in the microtubules complex has been carried out by means of different computational methods: QSAR analysis, Molecular Docking, Flexible Alignment and Molecular Dynamics Simulation. The information derived from these computational studies have guide and led to the design and chemical synthesis of new arylthioindole derivates as small molecule inhibitors of tubulin polymerization.

The 3D QSAR analysis of an initial database of arylthioindole derivates verified and explained the notable biological effectiveness of these compounds, and identified crucial structural patterns, thus guiding the design and synthesis of new molecules. The Molecular Docking and Flexible Alignment studies of the most active compounds and colchicine revealed the synthetic molecules to adopt and mimic the colchicine's binding mode.[4]

Given these relevant and promising results, the synthesis of new molecules with the indole as core structure has started aiming to increase the biological activity of this class of antimitotic drugs. Computational evaluation by means of Molecular Dynamics Simulation investigating the bioactive conformation of these anti-tubulin agents,[5] correlated well with the experimental data and will further direct and improve the design of new potential derivates.

Result and Discussion

Molecular Modelling

A series of arylthioindoles and arylsulfonylindoles[135] were designed as molecules joining structural features which characterize several potent tubulin polymerizarion inhibitors, such as (A) the indole nucleus (e.g.: 3-formyl-2-phenylindoles, 2-aroylindoles, heterocombretastatins, diarylindoles, D-24851, 2-aryl-3-aroylindoles, 3-aroyl- and 3-aroylindoles), (B) the 3,4,5-trimethoxyphenyl moiety (e.g., but not limited to: colchicine, combretastatin A-4), and (C) a sulfur atom (e.g.: sulfonamide E-7010,[136] thiophene[137] and benzothiophene[138] derivatives) (Chart 2.4).

Introduction of a methoxycarbonyl function at position 2 of the indole ring of **113** produced methyl 3-(phenylthio)-1H-indole-2-carboxylate (**114**), which was about twice as potent as **113** (IC$_{50}$ = 8.3 μM). Oxidation of the sulfur atom of **114** led to the inactive sulfone **115**. In contrast, a different behaviour was observed for the analogous pair **117/118**: the inhibitory activity of **117** was retained by the sulfone **118**. Despite their appreciable inhibition of tubulin activities, these compounds were unable to inhibit the growth of MCF-7 human breast carcinoma cells (Table 2.3).

Replacement of the phenylthio of **113** with the 3,4,5-trimethoxyphenylthio moiety furnished methyl 3-[(3,4,5-trimethoxyphenyl)thio]-1H-indole-2-carboxylate (**116**), which showed an IC$_{50}$ of 2.9 μM, 2.8 times superior to that of **113** and comparable with those of colchicine (IC$_{50}$ = 3.2 μM) and CSA4 (IC$_{50}$ = 2.2 μM). Most importantly, this compound inhibited the growth of the MCF-7 cells by 50% at 25 nM, a concentration only ca 2 fold higher than that observed with colchicine (IC$_{50}$ = 13 nM) and CSA4 (IC$_{50}$ = 17 nM).

The introduction of a methoxy group at position 5 of the indole gave **123** with enhanced inhibitory potency in both the biochemical and cytological assays. Compound **123** (IC$_{50}$ = 2.0 μM) was 1.6 times more active than colchicine and about as active as CSA4 in inhibiting tubulin polymerization. As inhibitor of the growth of MCF-7 cells (IC$_{50}$ = 13 nM) it was as potent as colchicine and CSA4. Among the tested derivatives, **123** was the most active.

It is interesting to note that, independently on the presence/nature of the substituent at position 5 of the indole, the 3,4,5-trimethoxyphenylthio moiety was crucial for the effective inhibition of the growth of MCF-7 cells.

Table 2.3 - Inhibition of tubulin polymerization, colchicine binding and growth of MCF-7 human breast carcinoma cells by compounds 113-123. [135]

113,114,116,117,119,120,122,123   **115,118,121**
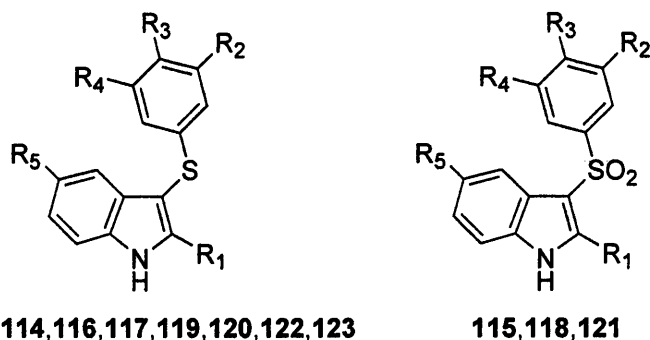
Compd	R_1	R_2	R_3	R_4	R_5	tubulin[a] $IC_{50} \pm SD$ (μM)	colchicine binding[b] (% $\pm$ SD)	MCF-7[c] $IC_{50} \pm SD$ (nM)
113	H	H	H	H	H	15 $\pm$ 0.7	13 $\pm$ 5	>2500
114	COOCH$_3$	H	H	H	H	8.3 $\pm$ 0.6	21$\pm$ 7	>2500
115	COOCH$_3$	H	H	H	H	>40	5.1 $\pm$ 1	>2500
116	COOCH$_3$	OCH$_3$	OCH$_3$	OCH$_3$	H	2.9 $\pm$ 0.1	81 $\pm$ 1	25 $\pm$ 1
117	COOC$_2$H$_5$	H	H	H	H	4.4 $\pm$ 0.3	19 $\pm$ 7	>1250
118	COOC$_2$H$_5$	H	H	H	H	4.4 $\pm$ 0.2	33 $\pm$ 0.6	>1000
119	COOCH$_3$	H	OCH$_3$	H	Cl	>40	3.6 $\pm$ 8	>2500
120	COOCH$_3$	OCH$_3$	OCH$_3$	OCH$_3$	Cl	2.5 $\pm$ 0.3	76 $\pm$ 0.2	42 $\pm$ 10
121	COOCH$_3$.	OCH$_3$	OCH$_3$	OCH$_3$	Cl	>40	1.6 $\pm$ 2	>2500
122	COOCH$_3$	H	H	H	OCH$_3$	6.2 $\pm$ 0.3	31 $\pm$ 8	>2500
123	COOCH$_3$	OCH$_3$	OCH$_3$	OCH$_3$	OCH$_3$	2.0 $\pm$ 0.2	93 $\pm$ 0.8 (68 $\pm$ 3)[d]	13 $\pm$ 3
Colch.						3.2 $\pm$ 0.4	-	13 $\pm$ 3
CSA4						2.2 $\pm$ 0.2	100 $\pm$ 0.9 (88 $\pm$ 0.4)[d]	17 $\pm$ 10

[a]Inhibition of tubulin polymerization. Tubulin was at 10 μM. [b]Inhibition of [^{3}H]colchicine binding. Tubulin was at 1 μM, both [^{3}H]colchicine and inhibitor were at 5 μM [c]Inhibition of growth of MCF-7 human breast carcinoma cells. [d]Inhibition of [^{3}H]colchicine binding. [^{3}H]Colchicine was at 5 μM, both tubulin and inhibitor were at 1 μM.

These results were confirmed by an extended SAR study on this class of compounds, which highlighted several crucial structural requirements indispensable for enhancing the effectiveness of the arylthioindoles: (i) the presence of a methoxy/ethoxycarbonyl group at position 2 of the indole nucleus; (ii) the 3,4,5-trimethoxyphenylthio group at position 3 of the indole; (iii) the sulfur atom in the sulfide oxidation state; (iv) the substituent at position 5 of the indole. The last seems to be especially important for inhibition of the growth of MCF-7 cells.[139]

Initial molecular modelling studies have revealed a possible binding mode for these novel inhibitors that could support the observed biological data.

QSAR analysis, Molecular Docking and molecular dynamics studies on colchicine and some ATIs have been performed using the 3D structure of tubulin co-crystallized with the colchicine analogue N-deacetyl-N-(2-mercaptoacetyl)-colchicine (DAMA-colchicine)[140] aiming not only to investigate further SAR rules and mechanism of action and to evaluate their activity against a wider range of tumor cell lines, but also to design and synthesize novel arylthioindole derivates as antitubulin agents (the results of this investigation have been already published).[139]

QSAR

The 3D quantitative structure-activity relationship of a database of about eighty arylthioindole derivates acting as potent and effective tubulin polymerization inhibitors was undertaken. Correlating the biological activity of a series of compounds with several properties of each molecule (steric demand, lipophilicity, electrostatic interactions), this computational analysis can verify and explain the relevant and notable biological effectiveness of these compounds, thus leading to the identification of crucial structural patterns and, therefore, to the design and synthesis of new indole derivates. Both CoMFA[141] and MOE-QSAR[142] methods have been used, trying to compare the results for an exhaustive comprehension of all the data obtained.

CoMFA results

Many attempts have been tried to obtained a CoMFA[141] model of statistical significance and a steric/electrostatic contour map consistent with the biological activity (expressed in terms of IC_{50}) of the antitubulin agents database.

None of these attempts have provided satisfactory results as the Q^2 value has always been very low and hence the model was unacceptable. Also the contour maps have been contradictory as the visual inspection could detect inactive or poorly active compounds to be evaluated as the most active inhibitors by the CoMFA analysis.

Useless has been the further attempt of modifying the lattice spacing trying to improve the energy.
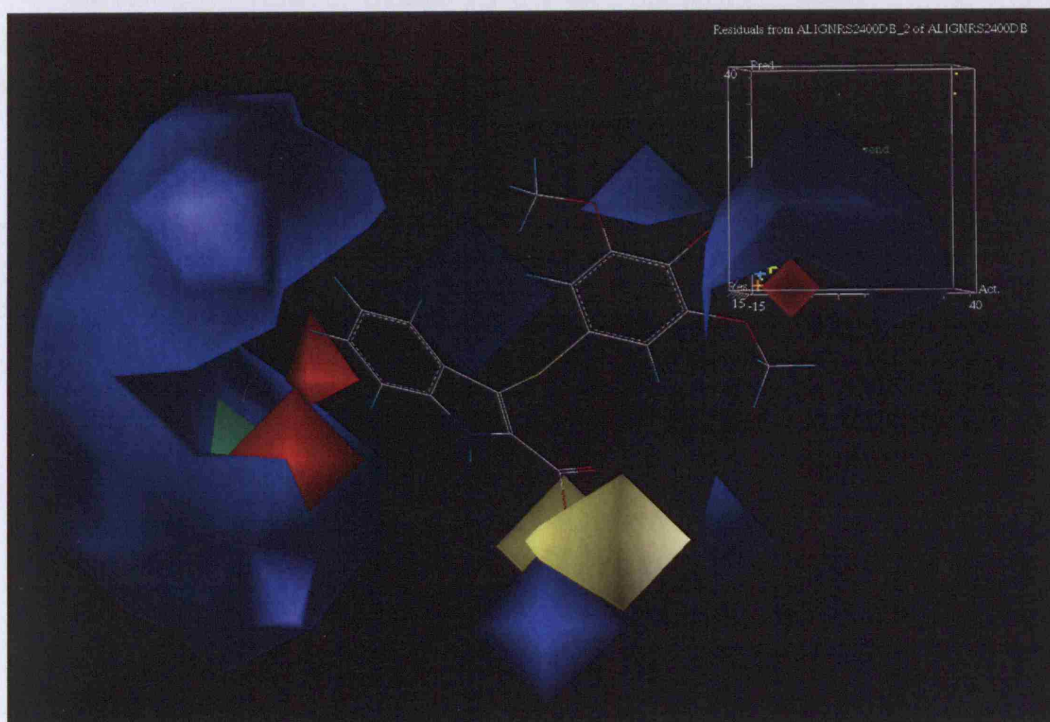
Fig. 2.2 - CoMFA results. **R2:** 0.970 **xR2** : 0.614 **Red**: negative potential **Blue**:positive potential **Green**: more bulk **Yellow**:less bulk

Possible reasons of these controversial results have been thought to be related to the hight homogeneity of the biological activities and to the missed alignment of all molecules in the database before starting the QSAR analysis.

In fact, mainly biological data spannig a range of at least three orders of magnitude should be used for a valid and robust CoMFA analysis. As up to date it has not been possible neither to increase the number nor to modify the homogeneity of the IC_{50} data, the only drawback that could be amended was the alignment of the molecules studied.

This is considered as the most crucial, difficult and important step in all CoMFA based methods as it is challenging to obtain a realistic alignment of the structures to analyze. As a prove of this, the preliminary alignment of all molecules in the antitubulin agents database to the template structure of RS2400 (one of the best docked structure resulted in the virtual screening) has led to a successful and interesting CoMFA analysis (Q^2=0.624; R^2=0.972; see Fig.2.2). Although these reasonable results, only 11 eleven compounds have been aligned and subjected to the QSAR study as the database may contain a too small number of entries. The availability of new tested compounds and, consequently, the increased range of biological activities will guide future and improved CoMFA studies on the arylthioindole derivates.

MOE-QSAR results

Despite the long time spent in studying and evaluated the best and appropriate descriptors for the molecules included in the antitubulin agents database, the MOE-QSAR method[142] has provided interesting and promising results since the first calculations. The main challenge has been to find out the smaller number of specific descriptors for the selected training test that could be able to correctly predict the biological activity of the corresponding test set and the identification of a heterogeneous and valid training test. For this QSAR method to work successfully, it is crucial to identify possible outliers (false positive and false negative) and, consequently, to modify both training and test sets until obtaining a final model able to provide reliable and consistent prediction also for new or not yet tested compounds. Furthermore, the availability of exactly determined IC_{50} values is fundamental for both the whole evaluation and the predictiveness of the final model.

It is important to note that, like in all CoMFA based methods, the wider the biological activity data range, the more successful and reasonable the model is.

The descriptors used for the final model obtained by the present study are listed as follows:

- a_IC : atom information content
- b_1rotN : number of rotable single bonds
- ASA_H : total hydrophobic surface area
- ASA_P : total polar surface area
- E_ele : electrostatic energy
- E_vdw : van der Walls energy
- Density : mass density
- KierFlex : molecular flexibility
- LogP (o/w) : Log octanol/water Partition Coefficient
- SlogP : log octanol/water partition coefficient
- PEOE_VSA_HYD : total hydrophobic vdw surface area
- SMR : molar refractivity
- TPSA : topological polar surface area
- Weight : molecular weight
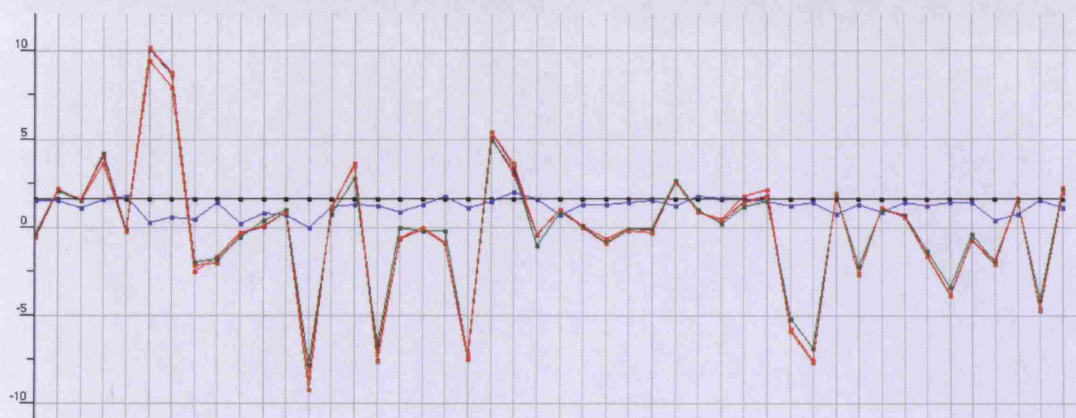
	MOE Model1	MOE Model2	MOE Model3
$R^{2(*)}$	0.949	0.880	0.880
$xR^{2(*)}$	0.646	0.563	0.578

Fig.2.3 – MOE-QSAR results. (*) R2 (correlation coefficient): indicator of the accuracy of the predictions; final xR2 value equal to or above 0.5: statistically significant and acceptable.

Flexible Alignment results

Following the interesting and relevant QSAR results, the most active compounds have been subjected to a flexible alignment (Moe Flexible Alignment).[142] The aim was to observe common structural and conformational features among all of them and, therefore, to explain and justify either the biological activity or the QSAR data obtained. Also the colchicine molecule has been aligned and, interestingly, many of the found conformations have shown an angle of about 59.9° between ring A and C (the natural compound has an angle of 56°).[143] The best colchicine conformation has been further aligned to that of RS2247, one of the most active arylthioindole derivate in the antitubulin agents database. The notable conformations derived from this last flexible alignment (see Fig.2.4) have provide further important information on the possible binding mode for these novel tubulin inhibitors. As their alignments and their corresponding docked structures have shown, they may mimic and adopt a similar orientation as the natural ligand to effectively bind to the binding site of the protein.
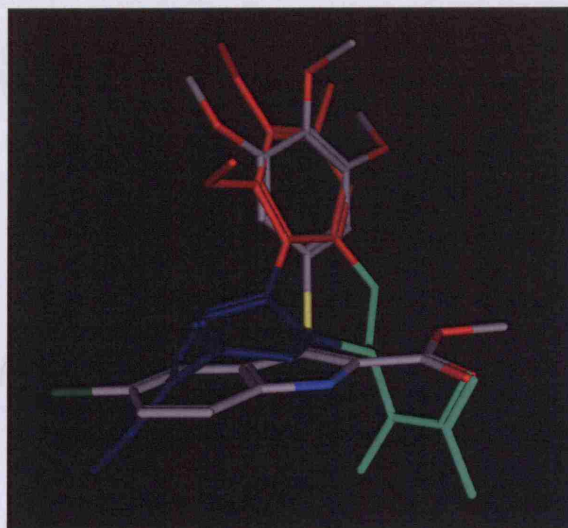
Fig.2.4 – Flexible alignment between colchicine and RS2247.

Molecular Docking

Along with QSAR study, Molecular Docking analysis of the antitubulin agents has been carried out, as this computational method allows the virtual screening of small molecule libraries that can be then tested against QSAR models; successful combining of both 3D-QSAR and docking results leads to the identification of the most promising compounds.

More in detail, the possible binding mode of these novel compounds have been investigated by docking studies of the arylthioindole **21** (see Table 2.4) in the colchicine binding site of tubulin, which was obtained from the recently reported 3D structure of tubulin cocrystallized with a colchicines analogue: *N*-deacetyl-*N*-(2-mercaptoacetyl)colchicines (DAMA-colchicine).[140]

Compound **21** was docked using a modified version of the docking tool of MOE,[142] which implements a genetic algorithm based search method.[144] The resulting protein/ligand complex was then minimized, revealing a very interesting and efficient binding model for the inhibitor to tubulin. Figure 2.5 clearly shows how the trimethoxy ring is well situated in proximity to Cys241 (residue numbers as in ref. 140), adopting an orientation very similar to that of the corresponding ring in the DAMA-colchicine of the crystallized structure. The methoxy substituent of the indole is also very close to the corresponding group on ring C of colchicine, leading to a very similar general binding of the two inhibitors. Furthermore, the indole moiety establishes one hydrogen bond between the NH and the backbone of Thr179 (shown in blue).

Several other hydrophobic contacts (not shown for the sake of clarity) stabilize further the binding of **21** to the protein. These observations are consistent with the highly efficient inhibition of [3H]colchicine binding that occurs with compound **21** (Table 2.5).

Table 2.4 – Structures of Arylthioindoles **11-21**. [135]

compd	R_1	R_2	R_3	R_4	R_5	S/SO$_2$
11	H	H	H	H	H	S
12	COOCH$_3$	H	H	H	H	S
13	COOCH$_3$	H	H	H	H	SO$_2$
14	COOCH$_3$	OCH$_3$	OCH$_3$	OCH$_3$	H	S
15	COOC$_2$H$_5$	H	H	H	H	S
16	COOC$_2$H$_5$	H	H	H	H	SO$_2$
17	COOCH$_3$	H	OCH$_3$	H	Cl	S
18	COOCH$_3$	OCH$_3$	OCH$_3$	OCH$_3$	Cl	S
19	COOCH$_3$	OCH$_3$	OCH$_3$	OCH$_3$	Cl	SO$_2$
20	COOCH$_3$	H	H	H	OCH$_3$	S
21	COOCH$_3$	OCH$_3$	OCH$_3$	OCH$_3$	OCH$_3$	S

Fig.2.5 - Proposed binding of **21**. DAMA-colchicine is represented in yellow, compound **21** in green. In red are represented the residues of the X-ray structure of tubulin, in cyan the same residues after minimization of tubulin with bound **21**. [135] Residue numbers are those used by Ravelli et al. [144]

Table 2.5 - Inhibition of Tubulin Polymerization, Colchicine Binding and Growth of MCF-7 Human Breast Carcinoma Cells by Compounds **11-21.** [135]

compd	tubulin[a] $IC_{50} \pm SD$ (μM)	colchicine binding[b] (% $\pm$ SD)	MCF-7[c] $IC_{50} \pm SD$ (nM)
11	15 ± 0.7	13 ± 5	> 2500
12	8.3 ± 0.6	21± 7	> 2500
13	> 40	5.1 ± 1	> 2500
14	2.9 ± 0.1	81 ± 1	25 ± 1
15	4.4 ± 0.3	19 ± 7	> 1250
16	4.4 ± 0.2	33 ± 0.6	> 1000
17	> 40	3.6 ± 8	> 2500
18	2.5 ± 0.3	76 ± 0.2	42 ±10
19	> 40	1.6 ± 2	> 2500
20	6.2 ± 0.3	31 ± 8	> 2500
21	2.0 ± 0.2	93 ± 0.8 (68 ± 3)[d]	13 ± 3
1 (colchicine)	3.2 ± 0.4		13 ± 3
4[e]	nd[g]	nd	160
6 (CSA4)	2.2 ± 0.2	100 ± 0.9 (88 ± 0.4)[d]	17 ± 10
7[f]	4.1 ± 0.6	28 ± 8	370 ± 2
8[f]	1.6	54 ± 0.7	45 ± 5

To further investigate the structural basis of the SAR that emerged from the biological results, additional docking studies on the entire series of compounds (see Table 2.6) have been carried out using the FlexX module included in SYBYL.[141] The results were evaluated using the built in scoring function as well as the functions included in the CScore module, with the aim of finding a good correlation between the experimental results and the in silico predictions. Some docking results were very similar to the proposed binding mode previously reported,[135] but, unfortunately, none of the scoring results correlated with the biological data (data not shown).

In particular, the scoring results for the *p*-methoxyphenyl analogues were often similar to or better than the corresponding results for the trimethoxyphenyl compounds. However, FlexX was able to identify an alternate possible binding conformation for this class of compounds compared with the one reported previously (Fig.2.6).

In this new pose, the indole moiety still forms a hydrogen bond with Thr179 (residue number based on the crystal structure used), but now the ester group is positioned deep in the binding pocket. In addition, the carbonyl of the ATI ester group forms another hydrogen bond with Lys352 (Fig.2.7). In comparing these two possible binding modes, the scoring functions employed were not able to clearly indicate which would be the preferred conformation. To solve this problem, molecular dynamics simulations (MD)

on the tubulin-ligand complexes have been carried out to evaluate the potential energy of the binding of the two different poses of **36** in the colchicines binding site.

Table 2.6 – Structure of Arylthioindoles **5 – 41** ([a] Ref. 135 [b] Ref. 145).[139]

compd	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	X
5	COO-n-Pr	H	H	H	H	H	H	H	S
6	COO-i-Pr	H	H	H	H	H	H	H	S
7	COO-n-Bu	H	H	H	H	H	H	H	S
8	COO-s-Bu	H	H	H	H	H	H	H	S
9	COO-t-Bu	H	H	H	H	H	H	H	S
10	COO-i-Pe	H	H	H	H	H	H	H	S
11	COOCH₂Ph	H	H	H	H	H	H	H	S
12	COOEt	OMe	H	H	H	H	H	H	S
13	COOEt	H	OMe	H	H	H	H	H	S
14	COOEt	H	H	OMe	H	H	H	H	S
15[a]	COOMe	H	OMe	OMe	OMe	H	H	H	S
16	COOEt	H	OMe	OMe	OMe	H	H	H	S
17	COOMe	OMe	H	H	H	H	Cl	H	S
18	COOMe	OMe	H	H	H	H	Cl	H	SO₂
19	COOMe	H	OMe	H	H	H	Cl	H	S
20	COOMe	H	OMe	H	H	H	Cl	H	SO₂
21[a]	COOMe	H	H	OMe	H	H	Cl	H	S
22	COOMe	H	H	OMe	H	H	Cl	H	SO₂
23[b]	COOMe	H	Me	H	Me	H	Cl	H	S
24[b]	COOEt	H	Me	H	Me	H	Cl	H	S
25	COOMe	H	OMe	H	OMe	H	Cl	H	S
26[a]	COOMe	H	OMe	OMe	OMe	H	Cl	H	S
27	COOEt	H	OMe	OMe	OMe	H	Cl	H	S
28	COOMe	H	H	H	H	H	H	OMe	S
29[a]	COOMe	H	H	H	H	H	OMe	H	S
30	COOMe	H	H	H	H	H	OMe	H	SO₂
31	COOMe	H	H	H	H	OMe	H	H	S
32	COOEt	OMe	H	H	H	H	OMe	H	S
33	COOEt	H	OMe	H	H	H	OMe	H	S
34	COOEt	H	H	OMe	H	H	OMe	H	S
35[a]	COOMe	H	OMe	OMe	OMe	H	OMe	H	S
36	COOEt	H	OMe	OMe	OMe	H	OMe	H	S
37	COOMe	H	OMe	OMe	OMe	H	NO₂	H	S
38	COOEt	H	H	H	H	H	OMe	OMe	S
39	COOEt	H	H	H	H	H	OMe	OMe	SO₂
40	COOMe	H	OMe	OMe	OMe	H	OMe	OMe	S
41	COOEt	H	OMe	OMe	OMe	H	OMe	OMe	S

Fig.2.6 - Two possible binding modes of **36**. In green the previously reported conformation (old pose);[135] in red the current proposed conformation (new pose).[139]
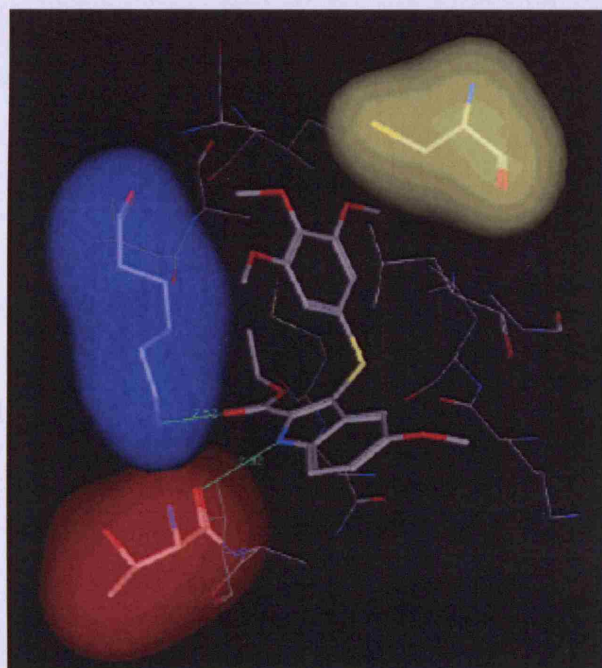
Fig. 2.7 - Proposed binding mode for **36**. Lys352 in blue, Thr179 in red, Cys241 in yellow. [139]

In the meantime, additional molecular modelling studies investigated the possible binding mode at the colchicine binding site of tubulin of newly designed arythioindoles with different R groups consisting of alkynes and corresponding alkanes (Table 2.7 below).

Table 2.7 - Structures of Different Alkynes and Alkanes.

R – Hydrocarbon Chain		R – Aryl ring	
Alkyne	Alkane	Alkyne	Alkane
A (11)	G		
B (12)	H	M (14)	P
C (13)	I		
D	J	N (15)	Q
E	K		
F	L	O (16)	R

As previously explained, some relevant amino acid residues can be observed in the tubulin active site; a cysteine residue (cysteine[241]) was identified to be important for

50

binding of the trimethoxy moiety in the active site. Leucine[248] and leucine[255] were responsible for holding the aryl ring of DAMA-colchicine in position. Other amino acid residues (lysine[352] and threonine[179]) were found to be in close proximity of the ligand and, although they play no part in the binding of DAMA-colchicine to the tubulin active site, these residues may play an important role in the binding of some of the arylthioindoles considered.

Minimised structures of arylthioindoles using MOE application were docked with FlexX (SYBYL) producing a database for thirty of the most stable conformations of each compound in the tubulin active site. Possible binding modes were visualized in MOE to observe important interactions in the active site and six compounds from Table 2.4 showed interesting binding in the colchicine active site.
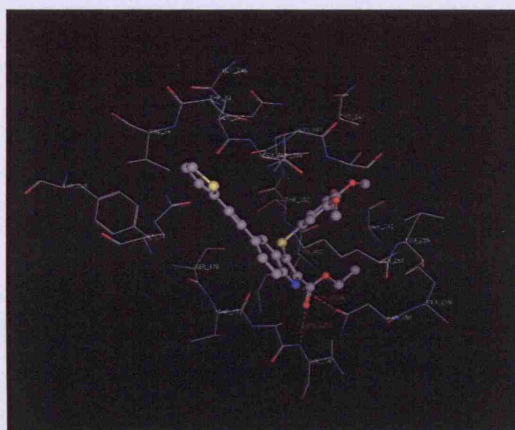
Fig. 2.8 - Arylthioindole with 3-ethynylthiourea as R group.

More in detail, an arylthioindole with 3-ethynylthiourea as R group (Fig.2.8) bound to valine[181] forming an ideal hydrogen bond at 2.7Å with the carbonyl oxygen of the ester moiety, while the nitrogen atom of the NH group formed a hydrogen bond with asparagine[258] (62%, 2.9Å). Like colchicine, this molecule formed important interactions with leucine[248] and leucine[255] that hold the aryl ring in position. Cysteine[241] was also found in close proximity to the 3,4,5-trimethoxy region of the aryl ring, thus cooperating on holding the aryl in its current position.
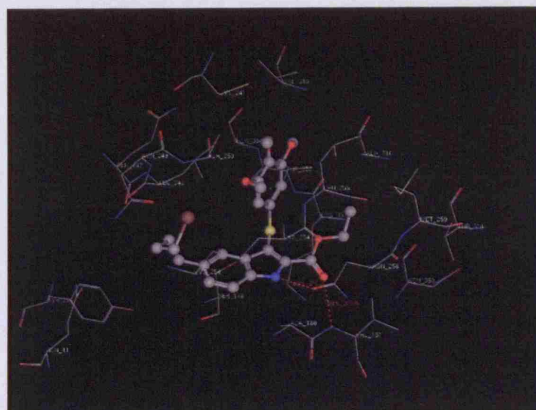
Fig. 2.9 - Arylthioindole with 1-chlorobutane as R group.

The arylthioindole shown in Fig.2.9 had 1-chlorobutane as R group. The 3,4,5-trimethoxy aryl ring was not as close cysteine[241], leucine[248] and leucine[255] as the molecule above, but residues were still in close proximity to hold the ring in place. Lysine[352] and threonine[179] were also found close to the ring but not involved in binding. Valine[181] formed a weak hydrogen bond with the carbonyl oxygen (59%, 3.1 Å) while asparagine[258] formed a hydrogen bond with the nitrogen of the NH group (63%, 2.8 Å).
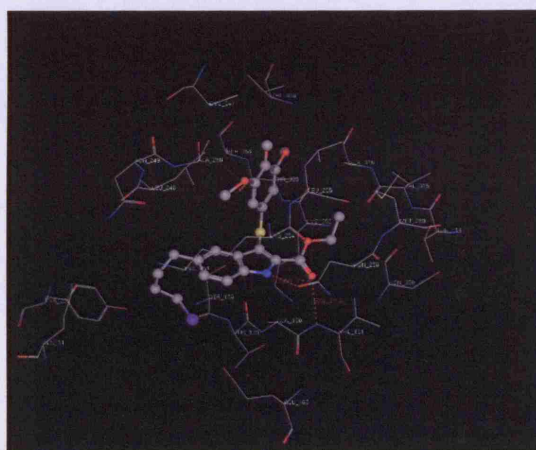
Fig. 2.10 - Arylthioindole with 1-bromobutane as R group.

In Fig.2.10 the R group of the arylthioindole is 1-bromobutane. The same amino acid residues were found to be in close proximity to the aryl moiety; clearly, cysteine[241], leucine[248] and leucine[255] are common residues involved in the interactions of all the investigated arylthioindoles. Valine[181] (59%, 3.1Å) and asparagine[258] (60%, 2.6Å)

formed hydrogen bonding with similar strengths as the molecule with the 1-chlorobutane (above). These residues also appear to be important common features.
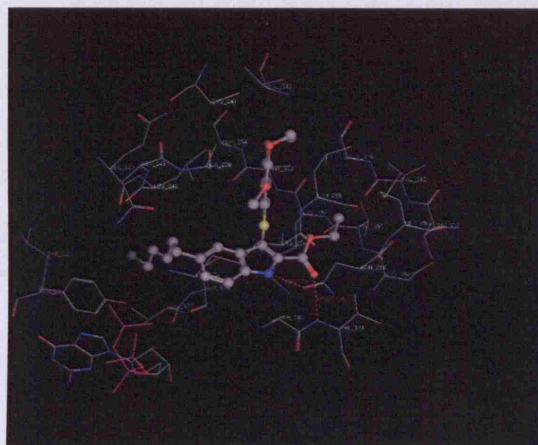
Fig. 2.11 - Arylthioindole with 1-chloropentane as R group.

A longer chained alkane as R group, 1-chloropentane, (Fig.2.11), highlighted the difference of one extra carbon in the alkane hydrocarbon chain modifying the binding mode to the active site. The alkane is closer to the threonine, glycine and tyrosine residues. A GTP molecule could also observed in this region. The aryl ring is in close proximity to the same residues noted above. The strength of the hydrogen bonds formed by asparagine[258] (60%, 2.8Å) and valine[181] (59%, 3.1Å) are similar again to the other alkane-containing compounds.
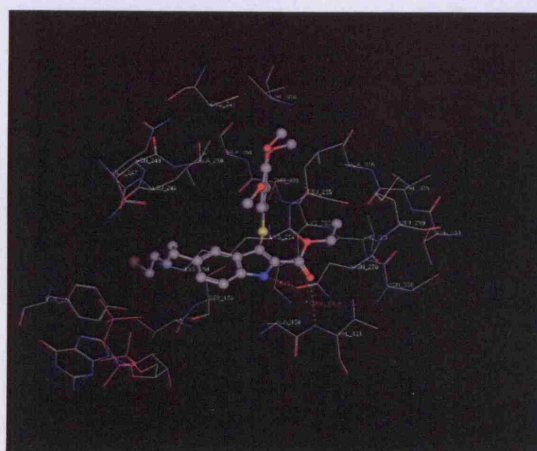
Fig. 2.12 - Arylthioindole with 1-bromobutane as R group.

The R group for the molecule shown in Fig.2.12, the alkane 1-bromopentane, is in close proximity to the tyrosine residue and GTP molecule. The usual hydrogen bonding with the valine[181] (58%, 3.1Å) and asparagine[258] (60%, 2.8Å) residues were observed. Probably due to the larger R group, this molecule was particularly close to the cysteine[241], leucine[248] and leucine[255] residues compared to the previous compounds.
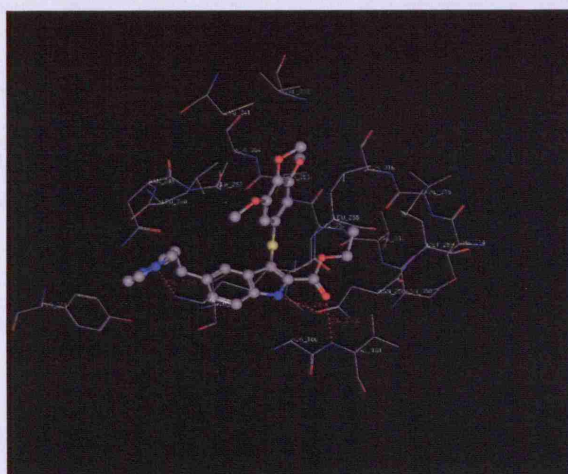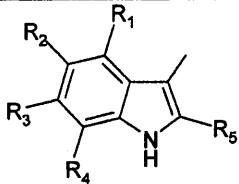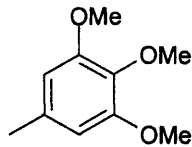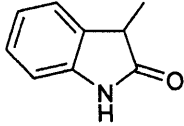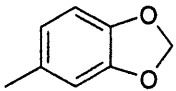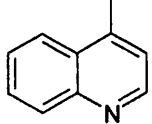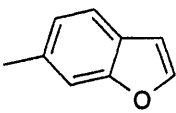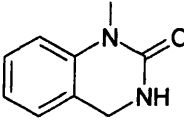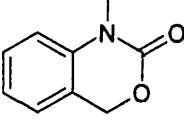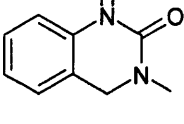
Fig. 2.13 - Arylthioindole with 5-ethyl-1-methyl-1H-imidazole as R group.

The compound with 5-ethyl-1-methyl-1H-imidazole as R group (Fig.2.13) showed the best binding mode to the tubulin active site out of the six compounds mentioned above. Other than the binding regions previously indicated, this compound formed an extra hydrogen bond interaction: lysine[264] bound to one of the imidazole nitrogens (55%, 2.7Å) near the tyrosine binding region.
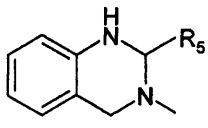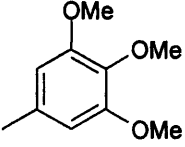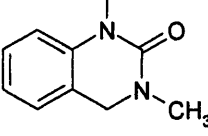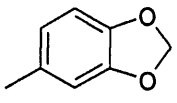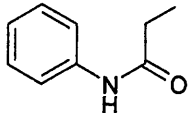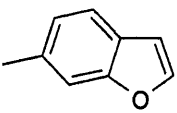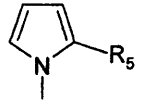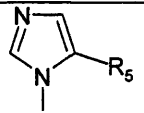
Molecular modelling results suggest the relevance of certain amino acid residues for binding to the tubulin active site: cysteine[241], leucine[248] and leucine[255] were important for holding the 3,4,5-trimethoxy aryl ring in its position, while valine[181] and asparagine[258] formed important hydrogen bond interactions with the carbonyl oxygen and indole nitrogen respectively. Since only one of the six compounds contained an alkyne (the other five compounds were alkanes), these series of alkane molecules designed as arylthioindole derivates showed more flexibility than the alkynes, thus better fitting the colchicine binding pocket.

With the purpose to design novel compounds with equal or improved active site interactions and simpler synthetic routes, a new database of modified compounds was designed and automatically generated in Molecular Operating Environment (MOE) from the scaffolds and fragments shown in Table 2.8 to investigate structure-activity relationships of novel antitubulin candidates characterized by a core structure different from the indole (A-1) as nucleus.

Table 2.8 – MOE new database of modified antitubulin agents.

A-B-C			
	A	**B**	**C**
1		S	
2		CH$_2$	
3		SO$_2$	
4			
5			
6			

A-B-C		
	A	**B**	**C**
7		S	
8		CH$_2$	
9		SO$_2$	
10			
11			

All compounds in the database were minimized using MOE application (MMFF94x was the force field applied) and then docked with FlexX program within Sybyl 7.0 (Multi Ligand Docking application was selected to run the docking) producing a database for thirty of the most stable conformations of each compound in the tubulin active site. The GTP molecule was kept in the protein as heteroatoms. Possible binding modes were visualized in MOE to observe important interactions in the active site.

Compounds derived from the combination of **B1**, **C1**, **C2**, **C3** fragments to the 1-methyl-1*H*-pyrrole nucleus (**A-10**) showed the most interesting binding in the colchicine active site. The results of these docking simulations highlighted how the rigidity and steric hinderance of the majority of the structures negatively influenced the interactions within the binding pocket, while the flexibility of 1-methyl-1*H*-pyrrole derivates positively favoured the desired docking conformation as these candidates better fit the colchicine binding pocket.

In addition, carboxylic and ester groups in position R_5 of the molecule formed additional hydrogen bonds with active site residues, while other substitutions (-H, -CH$_3$, -CHO) in the same position did not improve the interaction ligand/protein.

Interestingly, this molecular study further confirmed the relevance of certain amino acid residues for binding to the tubulin active site: cysteine,[241] leucine[248] and leucine[255] were important for holding the 3,4,5-trimethoxy aryl ring in its position.

In order to determine the reliability of the models realized up to this point and to gain a more accurate picture of how the compounds – and the enzyme – may behave *in vivo*, the configuration showing the most favourable interactions from the thirty generated by FlexX for each of the best compounds found as a result of this process, were selected and further minimized allowing the protein (previously kept rigid) to move around the compound and find a low-energy configuration (energy gradient 0.001).

The important data of this molecular modelling investigation suggested and inspired the chemical synthesis of novel antitubulin agents, the biological tests of which may lead to the discovery of more potent, effective and active tubulin polymerization small molecule inhibitors.

Molecular Dynamics

For the evaluation of the potential energy characterizing the binding of the two different poses of **36** (see Table 2.6 above) in the colchicines binding site, Molecular Dynamics calculations were performed using the MOE (Molecular Operating Environment)[142] software package, allowing free movement of the residues within a 15 Å distance of the ligand, keeping the rest of the protein fixed. This site was soaked in water, and the simulation was run in the NVT (number of particles, volume, and temperature of the system are kept constant during the simulation) environment, for a total of 600 ps at 300 K. The states generated during the first 100 ps were not considered in evaluating the results. The interaction potential energy values calculated by the dynamics simulations revealed a difference of almost 20 kcal/ mol between the two conformations (average values: -64.7 kcal/mol for the new pose, -44.9 kcal/mol for the old pose), suggesting that **36** would bind with the indole moiety oriented as in Figure 2.14.

Fig. 2.14 - Proposed binding mode for **36**. Lys352 in blue, Thr179 in red, Cys241 in yellow. [139]

These results encouraged to extend our molecular dynamics studies to other compounds and, in particular, to the *p*-, *m*-, and *o*-methoxyphenyl analogues (**32**, **33**, **34**). The simulation under the same conditions on colchicines was also performed to have a better understanding of its binding mode with tubulin. For the latter study, we modified the structure of the *N*-deacetyl- *N*-(2-mercaptoacetyl)colchicine cocrystallized with tubulin, and the resulting complex was energy minimized before the molecular dynamics experiment was performed. Colchicine established one hydrogen bond between the carbonyl group of ring C and Lys352 and another between the NH group and Thr179 (Fig. 2.15), the same residues involved in the binding of **36** (Fig. 2.14). During the simulation, the amide moiety changed its orientation, losing the hydrogen bond with Thr179 and forming another one between the amide carbonyl group and the hydroxyl group of Ser178. The trimethoxy ring A remains placed in a hydrophobic pocket where, in particular, two leucine residues (Leu248 and Leu255) interact strongly with the aromatic ring. Cys241, a key residue for the binding and biological activity of many colchicine analogues,19 is in close proximity to the C-3 methoxy group of ring A. The average value of the interaction potential energy for colchicine was -71.6 kcal/mol. The para-substituted arylthioindole **34**, on the other hand, does not establish a hydrogen

bond with Thr179 (Fig.2.16). The interaction with Lys352 remains stable during the simulation time, while the aromatic ring does not establish any contact with Cys241.
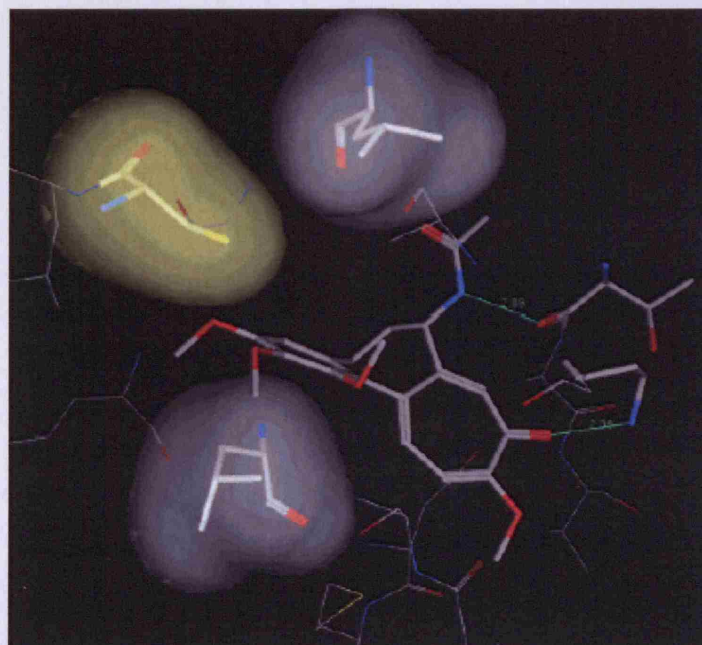
Fig. 2.15 - Proposed binding mode for colchicine. Leu248 and Leu255 in gray, Cys241 in yellow, Thr179 and Lys352 in stick representation. [139]
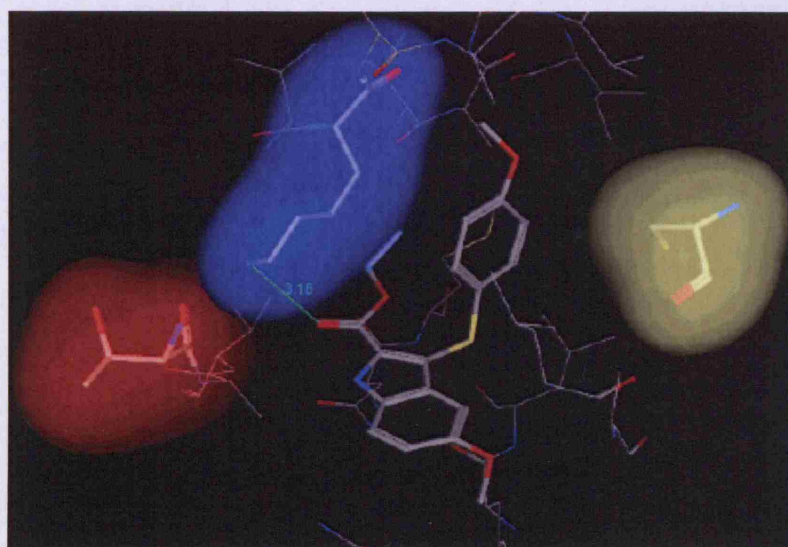
Fig. 2.16 - Proposed binding mode for **34**. Lys352 in blue, Thr179 in red, Cys241 in yellow. [139]

The energy value obtained for binding, -49.8 kcal/mol, is in agreement with the poor activity observed in the biological experiments. This is ~15 kcal/mol higher than the value obtained for **36**, indicating a reduced binding affinity for **34** relative to **36**. The results obtained for the ortho- and meta-substituted arylthioindoles (**32** and **33**) also correlated well with the experimental data. Both compounds established hydrogen bonds with Thr179 and Lys352 through the indole ring. The phenyl ring of compound **33** assumes a similar orientation as the trimethoxy ring of **36**, but, in the case of compound **32**, the o-methoxy aryl moiety is almost orthogonal to the corresponding aromatic ring of **36** (Figure not shown). The energy values obtained for **32** and **33** are in agreement with these observations and with the experimental results (Table 2.9).

Table 2.9 - Comparison between Average Potential Energy of Interaction (Uab) and Inhibition of Tubulin Polymerization for Compounds **32, 33, 34**, and **36**.[139]

compd	U_{ab} (kcal/mol)	tubulin $IC_{50} \pm SD$ (μM)
32	−57.6	16 ± 0.5
33	−60.1	3.1 ± 0.2
34	−49.8	>40
36	−64.7	2.4 ± 0.2

Chemistry

Current tubulin inhibitors, reported to be under preclinical and clinical investigation, are natural, semisynthetic and synthetic agents acting either by destabilizing or stabilizing mechanisms.[135] A notable number of tubulin polymerization inhibitors is characterized by owing an indole nucleus as the core structure. Baxter Oncology,[146] von Angerer,[147] Medarde[148] and Flynn[149] designed and prepared potent and effective antitubulin agents presenting the indole as the main moiety of their own compounds, thus further comfirming the promising results of the docking and QSAR studies of the antitubulin molecules analyzed in the present project and also encouraging the synthesis of novel arylthioindole derivatives.

In order to increase the effectiveness of the arylthioindoles, four main structural requirements, identified by a SAR study,[135,150] have to be taken into account in planning the design and synthesis of new derivatives:

1) a methoxy/ethoxycarbonyl group has to be at position 2 of the indole nucleus;

2) the presence of a 3,4,5-triethoxyphenylthio group is required at position 3 of the indole;

3) the sulphur atom has to be in the sulphide oxidation state;

4) the position 5 of the indole has to present a proper substitution.

The chemical procedure for the synthesis of the novel computer-aided designed arylthioindole derivatives resulted from the previous computational investigation is characterized by the following steps:

1) ethyl esterification of the 1H-indole carboxylic acid (fundamental step for the following two ones to work);[151]

2) synthesis of the ethyl 3,5-diiodo-1H-indole carboxylate;[152]

3) synthesis of the ethyl 5-diiodo-1H-indole carboxylate (deiodinizaton)
(step n.2 and step n.3 are a facile procedure employing a regioselective C3, C5-bisiiodination followed by zinc-mediated C3-dehalogenation);[152]

4) introduction of an alkyne substitution at position 5 of the ethyl 5-diiodo-1H-indole carboxylate (the alkynes being considered for the synthesis of arylthioindoles are shown in Table 2.10 below);

5) introduction of the 3-methoxythiophenol at position 3 of the ethyl 5-diiodo-1H-indole carboxylate.
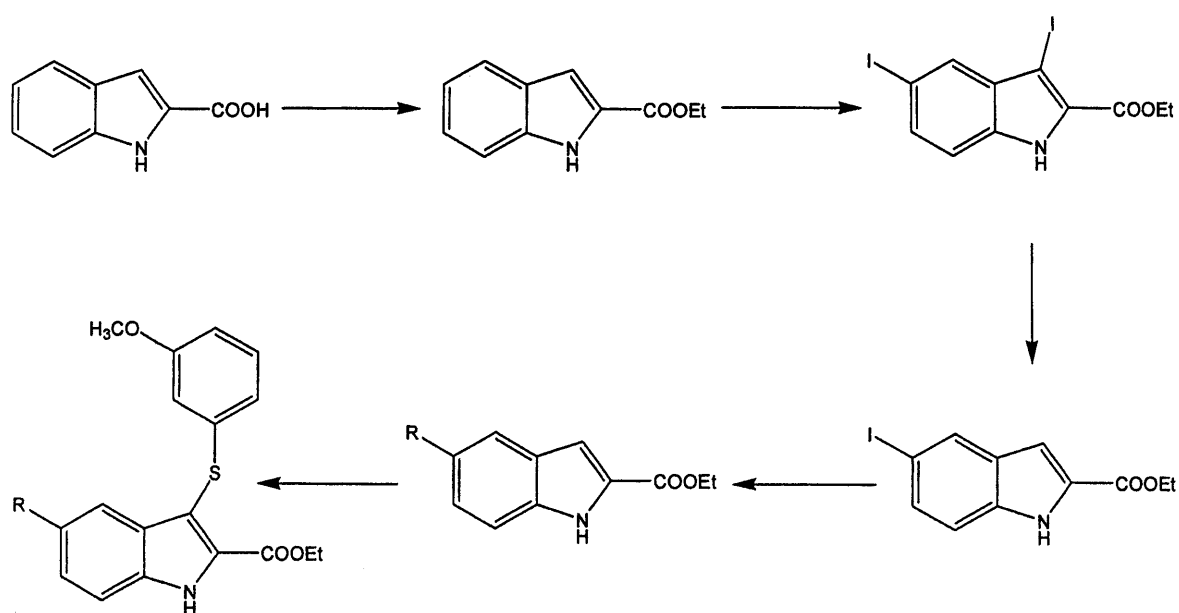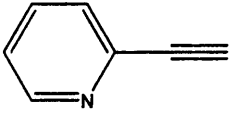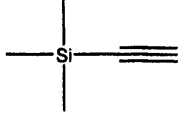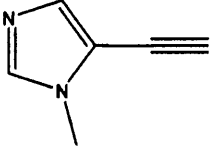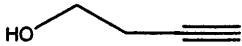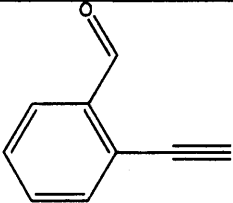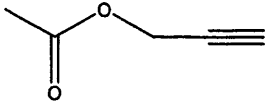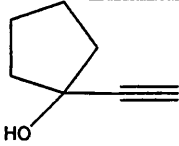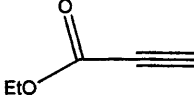
Fig. 2.17 - Five-Step Synthesis of the Arylthioindoles Structure.

Table 2.10 – Alkynes considered for the synthesis of arylthioindoles.

Arylthioindoles
Indole Nucleus

Alkyne (R)	

The synthesis of the indole nucleus as a core structure for the arylthioindoles required a three-step reaction scheme.
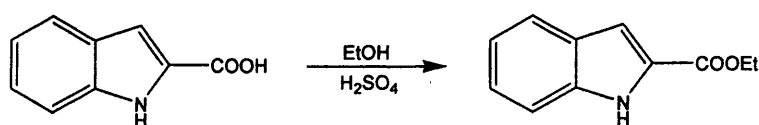
Scheme 2.1 - Synthesis of ethyl 1*H*-indole-2-carboxylate.

The first step to synthesise the indole involved a Fischer esterification process described by Fagan, G.P. *et al* (Scheme 2.1), during previous research on indoline analogues.[151] The reaction scheme shows the conversion of indole-2-carboxylic acid to the ester, ethyl 1*H*-indole-2-carboxylate via a nucleophilic acyl substitution reaction carried out under acidic conditions. Concentrated sulfuric acid is needed to make the carboxylic acid group more susceptible to nucleophilic attack by the alcohol (ethanol in this case)

yielding the desired ester, as both TLC (thin layer chromatography) and NMR analysis of the product with dichloromethane (DCM) also confirmed.
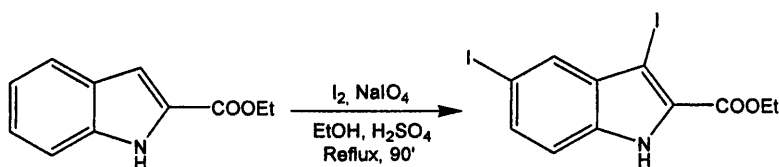
Scheme 2.2 - Synthesis of ethyl 3,5-diiodo-1H-indole-2-carboxylate.

The second step in the synthesis of the indole core structure involved the conversion of ethyl 1H-indole-2-carboxylate to ethyl 3,5-diiodo-1H-indole-2-carboxylate in a diiodination process devised by Beshore and Dinsmore.[152] Treatment of ethyl 1H-indole-2-carboxylate with I_2/NaIO$_4$/H$_2$SO$_4$/EtOH/reflux provided regioselectivity, thus ensuring iodination occurred specifically at both carbon 3 and 5, as shown by NMR spectra.
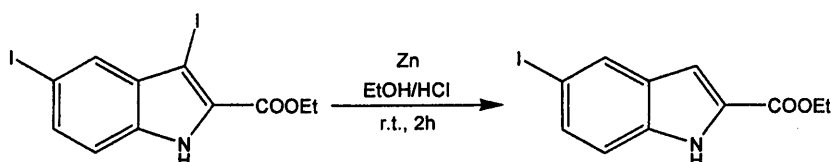
Scheme 2.3 - Synthesis of ethyl 5-iodo-1H-indole-2-carboxylate.

The third step in the synthesis of the indole core structure, also described by Beshore and Dinsmore in the same research (Scheme 2.3), required protodeiodination of ethyl 3,5-diiodo-1H-indole-2-carboxylate under reductive conditions affording the desired indole, ethyl 5-iodo-1H-indole-2-carboxylate. The procedure took place in the presence of powdered zinc, HCl, EtOH and ambient temperature.

The experimental procedure requires the product to be extracted four times with ethyl acetate. However, as the organic layer and aqueous layer took a long time to separate, the extraction process occurred at a total of thirteen times revealing to be quite time-consuming. The product was then washed to obtain pink crystals and allowed to dry overnight ready for use at the next stage. A more detailed NMR analysis, NOE

(Nuclear Overhauser Effect), was required to ensure that deiodination occurred correctly at carbon 3 and not at carbon 5.
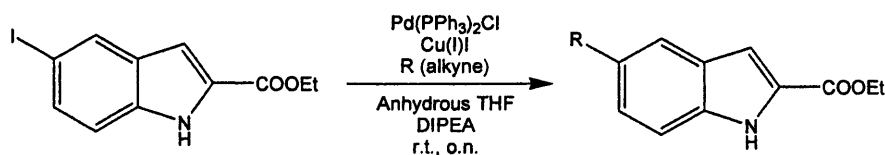
Fig. 2.18 - Attachment of the alkyne to the Indole Core Structure.

Attachment of the indole core structure to an alkyne (Fig.2.18) is the next step in the synthesis of arylthioindole derivates. The starting materials were treated with anhydrous THF (tetrafhydrofuran), DIPEA (N,N-diisopropylethylamine), Pd^{+2} (palladium (II) dichloride) and Cu(I)I (copper iodide). DIPEA accepts the hydrogen at carbon 1 of the alkyne and therefore acts as a base. Pd^{+2} acts as a catalyst and takes off the iodide at carbon 5 of the indole.
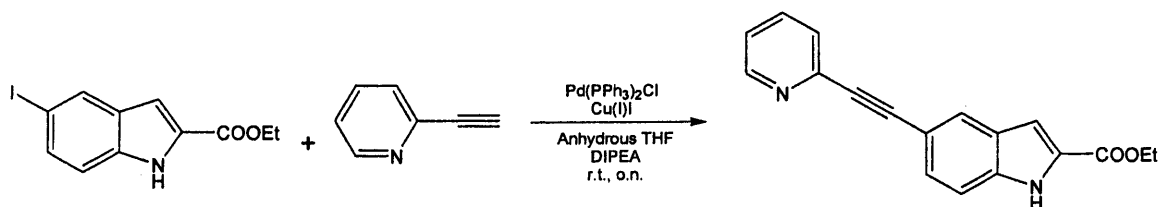
Scheme 2.4 - Synthesis of 5-pyridin-2-ylethynylyl-1H-indole-2-carboxylic acid ethyl ester.

The following ratio of materials was used: indole (1 eq.), Pd^{+2} (0.1 eq.), Cu(I)I (0.2 eq.) and 2-ethynyl pyridine (10 eq.) (Scheme 2.4). The reaction was left stirring overnight and then filtrated. Column chromatography was carried out to separate the four spots observed on TLC in the product mixture. ^{1}H NMR and ^{13}C NMR spectra suggested that indole-alkyne complex was not synthesized. Likely the alkyne coupled to itself during the reaction, thus indicating that the reaction kinetics favoured the alkyne-alkyne complex and not the alkyne-indole complex.

The reaction was repeated again with 2-ethynyl pyridine adding 2 eq. of alkyne instead of a 10 eq. with the purpose to improve reaction kinetics in favour of alkyne-indole

complex. A column chromatography was run to separate the three spots observed on TLC in the product mixture. This time both ^{1}H NMR and ^{13}C NMR spectra confirmed that the alkyne-indole complex was favoured over the alkyne-alkyne complex.
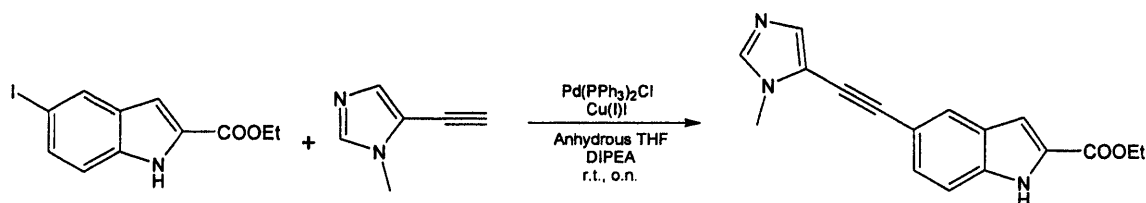
Scheme 2.5 - Synthesis of 5-(1-Methyl-1H-imidazol-2-ylethynyl)-1H-indole-2-carboxylic acid ethyl ester

For the synthesis of 5-(1-Methyl-1H-imidazol-2-ylethynyl)-1H-indole-2-carboxylic acid ethyl ester, 5-ethynyl-3-methyl-3H-imidazole was added to the indole in the presence of THF, DIPEA, Pd^{+2} and Cu(I)I (Scheme 2.5) according to the following ratio of materials: indole (1 eq.), Pd^{+2} (0.1 eq.), Cu(I)I (0.2 eq.) and alkyne (3 eq.). At the end of the reaction, left to run overnight, both ^{1}H NMR and ^{13}C NMR suggested the presence of the starting material 5-ethynyl-1-methyl-1H-imidazole but not that of the indole. A column chromatography was carried out trying to purify the mixture and, therefore, obtain clearer NMR spectra. Final ^{1}H NMR spectra confirmed that the desired product was not synthesized as lack of relevant peaks in the aromatic region clearly suggested. The reaction was successfully repeated following the same procedure and using the same conditions but paying attention to the time of the reaction; in fact, the reaction mixture was left stirring at room temperature under nitrogen atmosphere for 72 hours, thus probably favouring the equilibrium of the reaction towards completion.
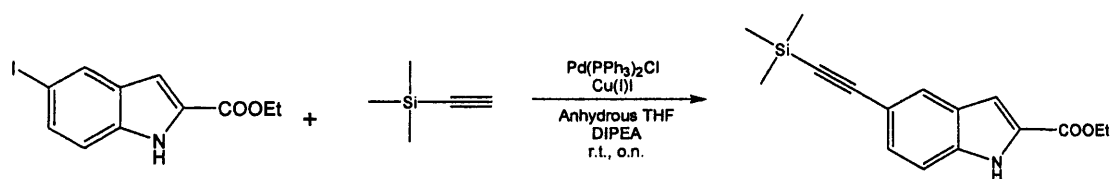
Scheme 2.6 - Synthesis of 5-trimethylsilanylethynyl-1*H*-indole-2-carboxylic acid ethyl ester.

For the synthesis of 5-trimethylsilanylethynyl-1*H*-indole-2-carboxylic acid ethyl ester, (Scheme 2.6) to a solution of anhydrous THF (8.5 mL) and ethyl 5-iodo-1H-indole-2-carboxylate (200.00 mg, 0.63 mmol, 1.00 eq) was added PdCl₂ (41.80 mg, 0.03 mmol, 5.7%), Cu(I)I (8.5 mg, 0.04 mmol, 7%), DIPEA (8.5 mL) and (trymethylsilyl) acetylene (119.85 mg, 0.17 mL, 1.26 mmol, 2eq) was added. The reaction was left stirring at 40 °C for 4 hours and then overnight at room temperature under a nitrogen atmosphere. The reaction mixture was filtrated and washed with THF, concentrated *in vacuo* and then easily purified by column chromatography using DCM as eluent. The purification yielded 131.5 mg of pure compound (72.6%).
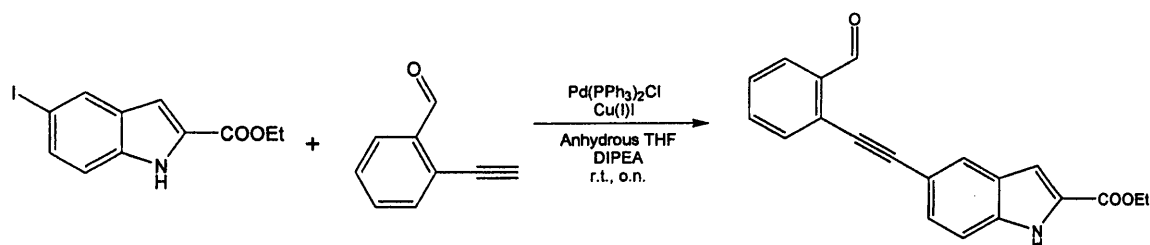
Scheme 2.7 - Synthesis of 5-(2-formyl-phenylethynyl)-1*H*-indole-2-carboxylic acid ethyl ester.

For the synthesis of 5-(2-formyl-phenylethynyl)-1*H*-indole-2-carboxylic acid ethyl ester (Scheme 2.7), to a solution of anhydrous THF (5 mL) and ethyl 5-iodo-1H-indole-2-carboxylate (100.00 mg, 0.32 mmol, 1.00 eq) was added PdCl₂ (22.3 mg, 0.03 mmol, 10%), Cu(I)I (12.00 mg, 0.06 mmol, 20%), DIPEA (5 mL) and 2-ethynylbenzaldehyde (123.9 mg, 0.10 mL, 0.95 mmol, 3 eq) was added. The reaction was left stirring overnight at room temperature under a nitrogen atmosphere. The reaction mixture was concentrated *in vacuo*, extracted with AcOEt and then purified by column chromatography (AcOEt/hexane 1:3). The purification yielded 32.2 mg of pure compound.
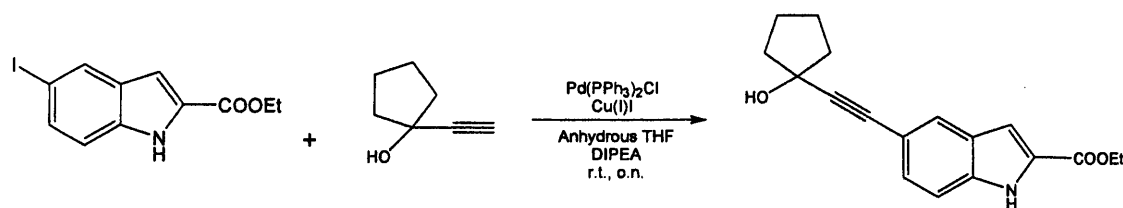
Scheme 2.8 - Synthesis of 5-(1-hyroxy-cyclopentylethynyl)-1*H*-indole-2-carboxylic acid ethyl ester.

For the synthesis of 5-(1-hyroxy-cyclopentylethynyl)-1*H*-indole-2-carboxylic acid ethyl ester (Scheme 2.8), to a solution of anhydrous THF (5 mL) and ethyl 5-iodo-1H-indole-2-carboxylate (100.00 mg, 0.32 mmol, 1.00 eq) was added PdCl₂ (22.3 mg, 0.03 mmol, 10%), Cu(I)I (12.00 mg, 0.06 mmol, 20%), DIPEA (5 mL) and 1-ethynyl-cyclopentanol (104.8 mg, 0.11 mL, 0.95 mmol, 3 eq) was added. The reaction was left stirring overnight at room temperature under a nitrogen atmosphere. The reaction mixture was concentrated *in vacuo*, extracted with AcOEt and then purified by column chromatography (AcOEt/hexane 1:3). The purification yielded 100.1 mg of pure compound.
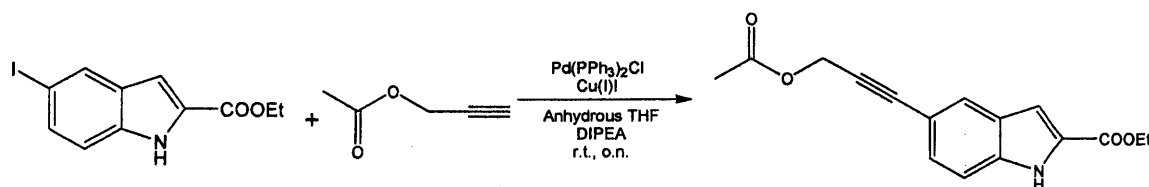
Scheme 2.9 - Synthesis of 5-(3-acetoxy-prop-1-ynyl)-1*H*-indole-2-carboxylic acid ethyl ester.

For the synthesis of 5-(3-acetoxy-prop-1-ynyl)-1*H*-indole-2-carboxylic acid ethyl ester (Scheme 2.9), to a solution of anhydrous THF (5 mL) and ethyl 5-iodo-1H-indole-2-carboxylate (100.00 mg, 0.32 mmol, 1.00 eq) was added PdCl₂ (22.3 mg, 0.03 mmol, 10%), Cu(I)I (12.00 mg, 0.06 mmol, 20%), DIPEA (5 mL) and propargylate (93.4 mg, 0.10 mL, 0.95 mmol, 3 eq) was added. The reaction was left stirring overnight at room temperature under a nitrogen atmosphere. The reaction mixture was concentrated *in vacuo*, extracted with AcOEt and then purified by column chromatography (AcOEt/hexane 1:3). The purification yielded 26.5 mg of pure compound.
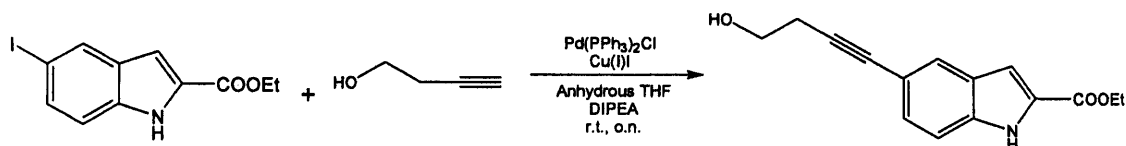
Scheme 2.10 - Synthesis of 5-(4-hydroxy-but-1-ynyl)-1*H*-indole-2-carboxylic acid ethyl ester.

For the synthesis of 5-(4-hydroxy-but-1-ynyl)-1*H*-indole-2-carboxylic acid ethyl ester (Scheme 2.10), to a solution of anhydrous THF (5 mL) and ethyl 5-iodo-1H-indole-2-carboxylate (100.00 mg, 0.32 mmol, 1.00 eq) was added PdCl$_2$ (22.3 mg, 0.03 mmol, 10%), Cu(I)I (12.00 mg, 0.06 mmol, 20%), DIPEA (5 mL) and 3-butyn-1-ol (222.43 mg, 0.24 mL, 3.17 mmol, 10 eq) was added. The reaction was left stirring four hours at room temperature under a nitrogen atmosphere. The reaction mixture was filtrated and washed with THF, concentrated *in vacuo* and then purified by column chromatography using AcOEt as eluent. The purification yielded 89.2 mg of pure compound (91.53%).
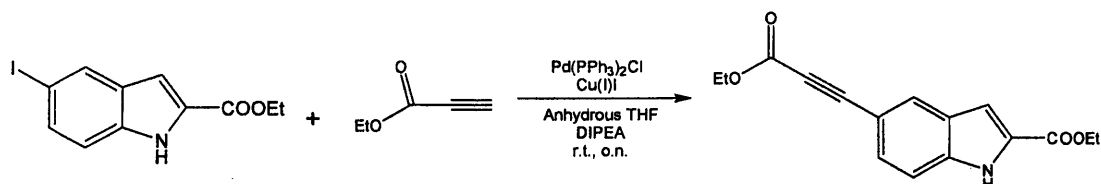
Scheme 2.11 - Synthesis of 5-ethoxycarbonylethynyl-1*H*-indole-2-carboxylic acid ethyl ester.

For the synthesis of 5-ethoxycarbonylethynyl-1*H*-indole-2-carboxylic acid ethyl ester (Scheme 2.11), it was necessary going back to the literature to study and find an alternative synthetic route as the one used for all the other derivates did not work, although several attempts and modifications had been tried. Interestingly, there is no record of successful direct reaction between aromatic ring containing iodine and this specific alkyne (ethyl propiolate).

According to the proper synthetic procedure, to a solution of ethyl propiolate (62.26 mg, 0.06mL, 0.63 mmol, 2eq) in THF (1.58 mL) was added 2.5M ButLi in hexane (0.25 mL, 2 eq) at -78°C. After stirring for 30 minutes at -78 °C, a solution of anhydrous zinc chloride (260.2 mg, 1.90 mmol, 6 eq) in THF (2.0 mL) was added.

The mixture was allowed to warm at room temperature and stirred for 1 hour. To the ice-cooled solution were added ethyl-5-iodo- 1*H*-indole-carboxylate (100 mg, 0.32 mmol, 1 eq) and PdCl$_2$ (11.10 mg, 0.01 mmol, 0.05 eq, 10%). After stirring at 50°C for 5 hours, the mixture was filtered through *celite* washing with THF. The filtrate was extracted with Et$_2$O, the combined extract dried (MgSO$_4$) and evaporated, and the residue purified by column chromatography (hexane/AcOEt 3:1) providing 145 mg of pure compound as brown oil.

At the end of the forth step for the synthesis of this series of arylthioindole candidates, it is worth highlighting and noticing both common features and problems observed during the experimental part of the present project.

Firstly, it was important to specifically find the right ratio of equivalents between the indole and the desired alkyne for the synthesis of the designed derivate to be successful. Similarly, crucial was the quantity and ratio of PdCl$_2$ and Cu(I)I.

Both these aspects required several attempts before finding the most efficient synthetic protocol. The work out was identical for most of the reactions and, relatively to the column chromatography, it was always observed the separation of the same three fractions:

- the first containing the indole or the alkyne
- the second containing a mixture of indole/alkyne and the desired compound
- the third, fluorescent, being the pure derivate

Worth mentioning is the problem and impossibility of synthesising two arylthioindole molecules because of the presence of two particular alkynes (3-ethynyl thiophene and propargylamine, respectively). In both cases, NMR spectra did not show traces of the expected compounds, and it was hypothesised that the interference of the two heteroatoms negatively affected the reaction.
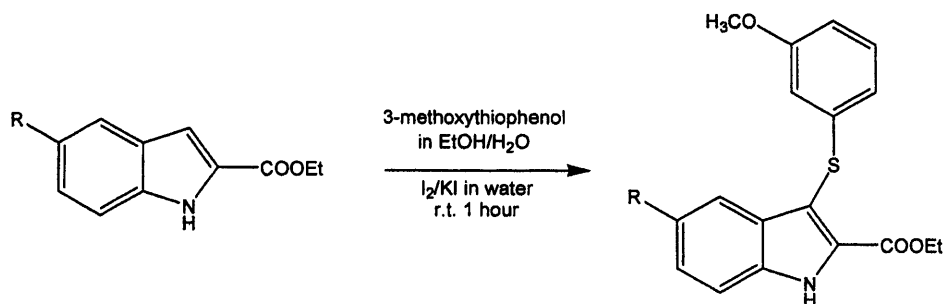
Fig. 2.19 - Attachment of the 3-methoxythiophenol to the Indole Core Structure.

The introduction of the 3-methoxythiophenol at position 3 of the ethyl 5-diiodo-1H-indole carboxylate is the fifth and final step of the synthesis of the arylthioindole molecules.

For all the arylthioindole derivates synthesised up to this point of the project, the following standard synthetic procedure (shown in Fig.2.19 above) was used for the synthesis of the corresponding final compounds:

to a stirring solution of the indole (1 eq.) and 3-methoxythiophenol (1 eq) in EtOH and water (1:1) is added a solution of I_2 (1 eq) and KI (4.8 eq) in water dropwise. The reaction mixture is left stirring at room temperature for one hour and, subsequently, H_2O, solid $NaHCO_3$ and $CHCl_3$ are added and the solution extracted. The separated organic layer is washed with a sodium thiosulphate solution, brine and then dried on $NaSO_4$. After concentration *in vacuo*, flash chromatography (hexane/AcOEt 9:1; TLC on silica gel and hexane/AcOEt 6:1 as eluent) is carried out to eliminate traces of starting materials; elution of the column with AcOEt or DCM follows to recover a mixture of starting materials and final compound, and then a last elution with MeOH affords the separation of the expected product. EtOH and excess of water are added to the last fraction (previously concentrated *in vacuo*) and the solution is kept in the fridge at least for 24 hours to easy the formation and precipitation of crystals, thus further purifying the desired final compound.

For all the final arylthioindole derivates synthesised following the above synthetic route, NMR investigation allowed the accurate analysis of both crystal and solvent recovered from the fridge to verify if any further purification of the final product was actually occurred.

71

Unfortunately, the NMR spectra of all arylthioindole revealed persistent presence of impurities. Due to the low quantity of the samples and time restraints, it was possible neither trying alternative techniques to isolate a sufficiently pure compound nor repeating the proposed synthetic scheme in a larger scale to improve the final yield and ease the purification of the desired product.

Conclusions and Future Work

The indole ring is the core nucleus of several potent tubulin polymerization inhibitors. Chart 2.5 shows the outstanding structural characteristics of the synthetic anti-tubulin indoles herein reviewed.

In the arylthioindole series, the relevant structural features were: (i) the methoxy/ethoxycarbonyl function at position 2; (ii) the 3,4,5-trimethoxyphenylthio group at position 3; (iii) the sulfur atom in the sulfide oxidation state; (iv) the substituent at position 5 of the indole. Arylthioindoles inhibited the binding of colchicine to tubulin, thus demonstrating that they bind to the colchicine binding site. Molecular modeling studies on some arylthioindole using the recently reported 3D structure of tubulin co-crystallized DAMA-colchicine were consistent with the highly efficient inhibition of [^{3}H]colchicine binding. The introduction of a methoxy group at position 5 of the indole was crucial for an effective inhibition in both biochemical and cytological assays. Compound **21** was more active than colchicine and about as active as CSA4 as inhibitor of tubulin polymerization, while it was as potent as colchicine and CSA4 as inhibitor of the growth of MCF-7 cells.

Several indoles are highly cytotoxic against many tumoral cell lines in the same range of concentration of several well-known antimitotic agents, such as colchicine, vincristine, vinblastine and paclitaxel. For the majority of tested compounds, experimental data are consistent with the inhibition of tubulin polymerization by binding to the colchicine binding site. However, some compounds did not bind at either the colchicine or vincristine colchicine binding site, thus suggesting a different binding site on tubulin. Small modifications of the chemical structure led often to changes of the molecular target, while the cytotoxity was retained.

Indole derivatives have great interest as potential drugs for cancer treatment, as it is well-documented by the volume of articles and patents recently published. It is evident the need to collect more information about these anti-tubulin polymerization agents. This goal can be pursued either by discovery of new tubulin polymerization inhibitors or by improvement of existing classes, also with the support of molecular modelling techniques.

Future work could include looking at a broader range of substitution groups at strategic point of the arylthioindoles and/or arylthiopyrrole as well as looking at more alkynes.

In the future, the biological activity of the newly designed and synthesised compounds could be evaluated and research could be extended by evaluating these compounds in a wide variety of cancer cell lines. Extensive molecular modelling studies would also be beneficial to further investigate other amino acid residues that could be important for binding in the colchicine active site, thus guiding and leading to the discovery of new potent and effective tubulin polymerization small molecule inhibitors.

Chart 2.5 - Main structural characteristics of anti-tubulin indoles.
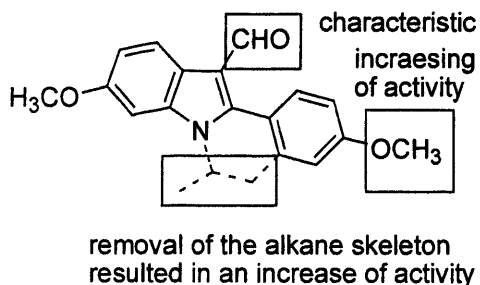
3-Formyl-2-phenylindoles
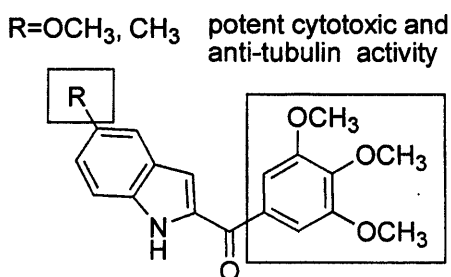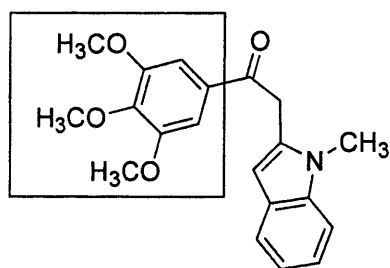
2-Aroylindoles

Heterocombretastatins
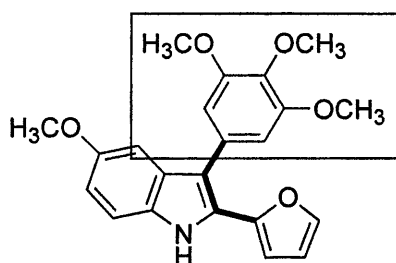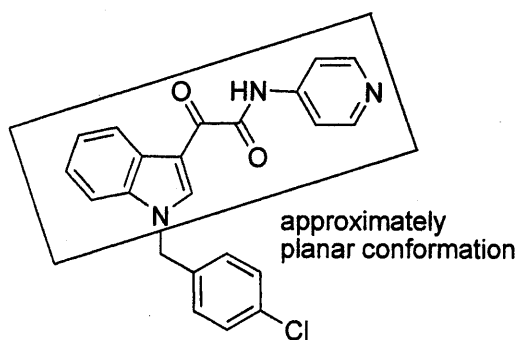
Diarylindoles

D-24851
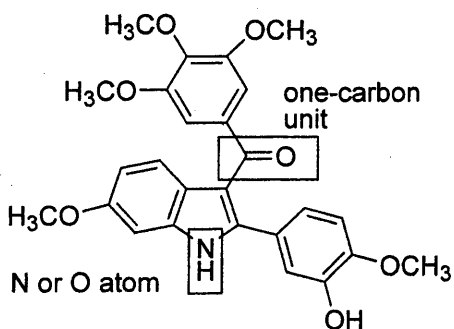
2-Aryl-3-aroylindoles
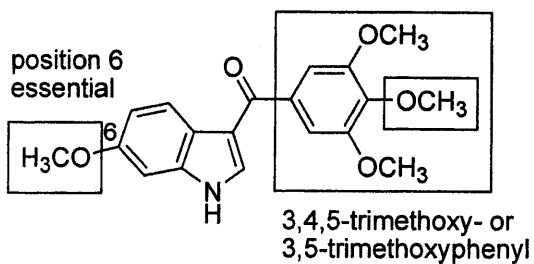
3-Aroylindoles
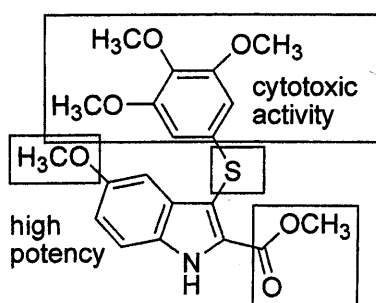
Arylthiolindoles

Chapter 3

Introduction

3.1 Hepatitis C Virus

As a noncytopathic hepatotropic member of the *Flaviviridae*, the hepatitis C virus (HCV) is responsible of acute and chronic hepatitis, and hepatocellular carcinoma (HCC),[1] currently infecting more than 170 million people.[2] Asymptomatic acute infection usually tends towards chronicity.

Like hepatitis B and human immunodeficiency (HIV) viruses, HCV is primarily transmitted percutaneously:[3] blood and related products[4] haemodialysis[5] and organ transplantation[6] have been the main means of infection, while nowadays HCV commonly affects injecting drug users and their sexual partners.[6] The hepatitis C virus persists in the major part of infected people causing illness and death (cirrhosis, end-stage liver disease and hepatocellular carcinoma). Limited medical treatments impose liver transplantation as unique solution to save patients' lives. Although the significant discovery of the virus and the sequencing of its genome, only recent efforts have made possible progress in the knowledge of the viral life cycle, protective response of the immune system, and viral ability to escape the host immune reaction.

Indeed, recent studies have shed light on the immune response to HCV revealing the central role of T cells in viral control and clearance.[7] Despite the relevance of this information, still the mechanisms by which T cells control HCV replication, the protective role of antibodies and how cellular and humoral immunity are contrasted in persistent infection remain to be explained.[7] New potential drugs have been designed and synthesised and the discovery of a vaccine may be close in the next future.

3.2 The virus and its life cycle

HCV uniquely infects humans and chimpanzees. Although until recently the absence of both cell culture and small-animal models of the infection has negatively influenced our knowledge about HCV, great progress has been done after the cloning of the HCV genome by Houghton and colleagues.[8]

3.3 The host–virus relationship

As persistent virus not killing its infected cells, HCV functions by prompting an immune-mediated inflammatory response (hepatitis) that can either clear the infection or cause the development of HCC depending on the efficiency of the antiviral immune response. Cell culture and small-animal models are required to investigate host-virus interaction and, although the former are just becoming available, causes and effects of HCV infection have started to be clarified (six different HCV genotypes with numerous subtypes, high replication and mutation rates, viral evasion strategies).[13,14,15,16,17]

3.4 Current treatments

Pegylated IFN-α plus ribavirin, standard treatment for HCV chronic infection[18,19] involving both immunostimulatory and antiviral mechanisms of action,[20] is not very effective, exerts toxicity[21] and often treatment success depends on unpredictable and still obscure viral and host factors. HCV replicon systems currently help develop drugs that inhibit viral replication aiming to less toxic and more active treatment regimens. Diverse promising small-molecule inhibitors of the NS3/4A protease and the NS5B polymerase are being developed but early tests have already shown the presence of escape variants.[22]

Targeting all aspects of the viral life cycle is becoming a reality thanks to the *in vitro* HCV infection model mentioned above.[10–12]

While some groups focus on the stimulation of innate immune system,[25] other groups are working on therapeutic vaccines hoping to activate the immune response,[23,24] which indeed makes patients able to clear their infection spontaneously and could eventually destroy drug-resistant viral clones.

3.5 HCV translation and polyprotein processing

It has taken about twenty years for the first HCV sequences to be reported (1989)[26] following the observation of non-A, non-B hepatitis contaminated blood in the

world's supply.[27] According to the current knowledge about HCV life cycle, once interacted with specific surface receptors of a cell, enveloped virus particles are easily internalized thanks to the acidic condition in the endocytic compartment, then a single-stranded (ss), positive-sense RNA genome is released into the cell cytoplasm to function as messenger RNA (mRNA) for translation of the viral proteins; as template for RNA replication; and as a nascent genome in new virus particles.

Formed into the endoplasmic reticulum (ER), virions leave the cell through the secretory pathway.

HCV genome lacks a 5'cap and its translation depends on an internal ribosome entry site (IRES) within the 5'-noncoding region (NCR). The HCV IRES binds 40S ribosomal subunits without needing for pre-initiation factors, and induces mRNA bound conformation in the 40S subunit.[28] The IRES–40S complex captures both the eukaryotic initiation factor (eIF) 3 and the ternary complex of Met-tRNA–eIF2–GTP to form first a 48S intermediate to which an active 80S complex follows.[29,30]

A large polyprotein is the product of the translation of the HCV genome and is subjected to a proteolytical cleavage into 10 viral proteins (Fig.3.2a). The virion structural proteins (the basic core (C) protein, and glycoproteins E1 and E2) are encoded by the amino-terminal one-third of the polyprotein; a small integral membrane protein, p7, working as an ion channel follows.[31,32] Nonstructural (NS) proteins NS2, NS3, NS4A, NS4B, NS5A and NS5B, responsible for the coordination of the intracellular processes of the virus life cycle are encoded by the remaining genome.

While signal peptidase cleavages between C/E1, E1/E2 and E2/p7 make active structural proteins, signal-peptide peptidase releases core from the E1 signal peptide. The NS region is involved in several processes: in addition to the p7/NS2 junction cleavage, the NS2 viral autoprotease cleaves at the NS2/3 junction and the NS3-4A viral serine protease cleaves at all downstream sites (Fig.3.2a, b). Ribosomal frame shifting into an alternative reading frame within the core gene can generate F (frame shift) or ARFP (alternative reading frame protein), a small protein virally encoded.[33]
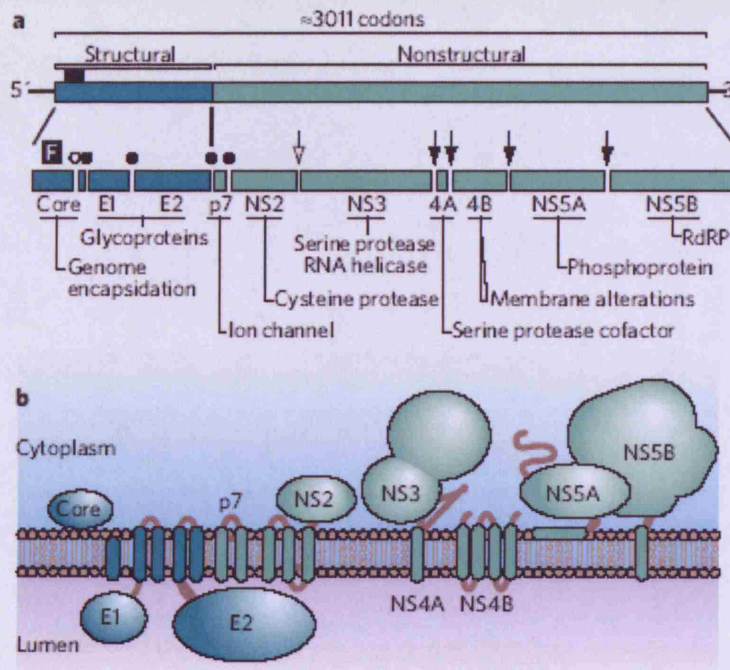
Fig.3.2 - HCV genes and gene products. **a**, The structure of the viral genome, including the long open reading frame encoding structural and nonstructural genes, and 5' and 3' NCRs. **b**, The topology of HCV proteins with respect to a cellular membrane.

HCV encodes two remarkable proteases; in fact, the catalytic triad of a cysteine protease is located in the carboxy-terminal two-thirds of NS2 where the presence of these residues is crucial for the cleavage of NS2/3 and the expression of NS3 serine protease, which may play a key role in the proper folding as suggested by the increased NS2/3 cleavage in presence of Zn^{2+}, ion able to structurally stabilize NS3 fold. As multifunctional protein, NS3 includes an N-terminal serine protease domain and a C-terminal RNA helicase/NTPase domain, the high-resolution structures of which have been solved.[34] The former has a typical chymotrypsin-like fold with three catalytically active residues at the surface interface between two β-barrel domains (Fig.3.3a). The intercalation of a β-strand present in NS4A[35-37] is necessary for complete folding and serine protease activity.[35-37] NS4A is a 54-amino-acid protein anchoring NS3 to cellular membranes through an N-terminal hydrophobic peptide.[37] As mentioned above, coordination of Zn^{2+} by three cysteine residues distant from the active site has to occur for a correct folding of the serine protease domain. Efforts have been focused on the discovery of specific inhibitors as NS3 is characterized by a shallow substrate-binding pocket for a serine protease, and have led to potent

molecules able to block NS3 serine protease activity, NS protein processing, and HCV RNA replication.[38]
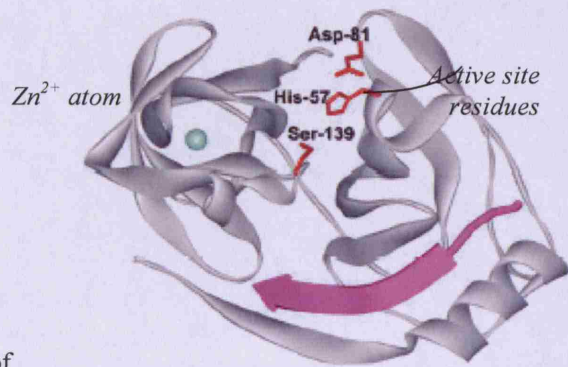
Fig.3.3a - Structure of   NS3 (protease only)-
4A Complex

A DExH/D-box RNA helicase encoded by the C terminus of NS3 uses the energy of NTP hydrolysis to unwind double-stranded RNA, while NS3 unwinds RNA and DNA homoduplexes and heteroduplexes in a 3' to 5' direction.[39] According to a recent kinetic study, the NS3 helicase behaves like a ratcheting two-stroke motor,[40] functioning as a dimer that incrementally rips apart 18-base-pair stretches of substrate RNA[41] and being regulated by the interactions between the serine protease and helicase domains of NS3.[42,43] HCV helicase, the folding of which is different from the new one adopted by the Dengue virus NTPase/helicase recently resolved at 2.4 Å resolution,[44] may be involved in the initiation of RNA synthesis on the HCV genome RNA.

Various cellular proteins, the fuction of which is still unclear, interact with NS5A protein,[45-47] that cellular kinases phosphorylate on multiple serine residues and, therefore, can be observed in hypophosphorylated (56 kDa) and hyperphosphorylated (58 kDa) forms. Although the identification of both the main phosphorylation sites,[48,49] and kinases phosphorylating NS5A,[50-53] still many are the questions and doubts about these aspects.

NS5A associates with membranes through an N-terminal amphipathic α-helix[54] and contains three distinct structural domains.[55] While no structural information is currently available for domain II and III, domain I (residues 1–213) is known to form a dimer with a novel fold and coordinate Zn^{2+} through a unique motif,[55,56] thus revealing both conserved external surfaces that might interact with other proteins, and

a highly basic channel that might be involved in RNA binding (Fig.3.3b). Interestingly, between cysteine residues 142 and 190, was discovered an usual modification for cytosolic proteins: a disulfide bond that could represent an alternative form of regulation.

THUMB
→ Keeps RNA in position for initiation of complementary strand manufacture and chain elongation

FINGERS & THUMB
→ Interact to form fully encircled active site which encourages co-ordinated & efficient processing of template and complementary strands

PALM
→ Acts as catalyst for nucleotidyl transfer reactions
→ Divalent metal cations (essential in nucleotide integration) chelated to highly conserved region of aspartic acid residues

FINGERS
→ Positions nucleotide triphosphates prior to integration together with the template strand bases they'll be paired to

Fig. 3.3b - Structure and function of the HCV NS5B RNA-dependent RNA polymerase (adapted from Francesco, R. D. et al., Approaching a new era for hepatitis C virus therapy: inhibitors of the NS3-4A serine protease and the NS5B RNA-dependent RNA polymerase. *Antiviral Res*, **58**, 1-16, 2003)

HCV RNA replication mainly depend on NS5B, a protein able to encode the RNA-dependent RNA polymerase (RdRP) and for which primer-dependent and *de novo* (unprimed) initiation of RNA synthesis have been observed. In fact, HCV RNA replication may involve *de novo* initiation by a multiprotein complex (replicase). With a typical 'right hand' polymerase structure, NS5A has catalytic sites in the base of the palm domain, surrounded by thumb and finger domains[34] entirely encircling the active site and creating a channel for binding to a ssRNA template.

A β-hairpin structure probably locates the template correctly as it protrudes from the thumb toward the active site.[57] Due to the high structural similarity between NS5B and RdRP of bacteriophage φ6,[58] their cocrystallization with model substrates and nucleoside triphosphates has provided a reliable model for *de novo* initiation.[59] With a low-affinity GTP-binding site supposed to function as allosteric regulator of the finger–thumb interaction,[60] NS5B is tethered to membranes by a C-terminal peptide

anchor[61] and, interacting with itself, it generates higher-order RdRP complexes potentially relevant to the membrane-bound replicase.[62]

Unclear aspects and processes regarding HCV replication and the host cellular environment as well as the evasion of intracellular host defence by the virus[63] may come to light following the characterization of cell-culture-adaptive mutations,[64] an important discovery subsequent to a previous breakthrough for the field of reverse genetics system for HCV, the selection of the first functional 'subgenomic' replicons in cell culture.[65] Indeed, the replicon system allowed to study HCV RNA replication and to evaluate potential antiviral compounds by means of a functional cell-based system.

Now that complete cell culture systems are available, genetic and biochemical approaches should shed further light on the processes of HCV entry, replication and virion production, thus providing information that can improve the development of specific antivirals targeting each stage in the viral life. New engineered small animal models of HCV infection and pathogenesis can help to progress and establish novel strategies to face and eradicate HCV and associated diseases.

3.6 Development of new therapies

The incidence of new infections has been sensibly reduced by current treatments; nevertheless the next future will see a constant prevalence of HCV infection.[66]

Additionally, current interferon (IFN)-based therapies are successful only on a minority of the patients at the expense of diverse adverse effects.[67,68] Therefore, more effective and tolerated treatments are needed and different are the approaches by research groups to achieve this objective. The majority of the efforts have focused on the discovery of agents able to inhibit specific steps in the life cycle of the virus.[69] Small-molecules, orally bioavailable inhibitors of the HCV enzymes as well as nucleic-acid based agents that attack the viral RNA, are referred to as 'HCV-targeted drugs'.[70] Recent investigations aim to the identification of compounds able to modulate the host immune response and to control and eventually eradicate HCV infection. Although the lack of accurate laboratory models of viral infection, some of these compounds resulted promising in early-phase clinical trials. However, the high rate of mutations (HCV has been also termed as "quasispecies")[70] make the viruses

resistant even to novel and more selective antiviral drugs, thus undermining the potentiality of a successful anti-HCV treatment and urging the identification of antiviral molecules able to suppress all HCV variants and to prevent the generation of new and more resistant ones. As a consequence, the discovery and optimization of potential antiviral candidates have to take into account both their spectrum of action on distinct genotypes and their resistance profile.

Since an efficient *in vitro* infection system is not available, and data obtained by HCV replicons cannot be directly extrapolated to viruses in infected individuals,[71] the combination of multiple drugs targeting both the virus and its host is considered more effective in chronic infections and likely to prevent drug resistance.

3.6.1 Small-molecule inhibitors of viral enzymes

Blocking vital viral enzymes is a straightforward approach to developing new anti-HCV compounds. Among all HCV enzymes, the NS3-4A serine protease and the NS5B RNA polymerase are the most targeted as shown by the development of competitive inhibitors of the NS3 protease as well as nucleoside, the promising clinical efficacy of which is giving renovate impulse to research in this field.

3.6.2 Inhibitors of the NS3-4A protease

The heterodimeric NS3-4A protease includes the amino-terminal domain of the NS3 protein and the small NS4A cofactor, and is fundamental for the generation of components of the viral RNA replication complex.[71] Its heterodimeric nature and the presence of a structural zinc atom make the NS3-4A protease different and unique within the chymotrypsin serine protease family to which it structurally belongs. Following a temporal sequence, the viral polyprotein is cleaved at four junctions by the enzyme that is also involved in blocking the host innate antiviral respons.[72] Taking into account the latter observation, inhibitors of this enzyme should stop virus replication and simultaneously increase antiviral efficacy.

Although the difference in substrate specificity of the NS3-4A protease from that of related enzymes[73] could ease the identification of inhibitors, their development as drugs resulted very challenging.

A long peptide substrate, with which to establish weak interactions along a vast surface is required by the enzyme, thus suggesting that only large molecules could inhibit the enzyme and, unfortunately, their dimensions represent an obstacle to their conversion into drugs. In addition, the 3D structure of the flat ligand-binding site of NS3-4A protease showed absence of cavities and clefts, essential features for the design of specific inhibitors of other proteases.[74-76] Nevertheless, efforts have been focused on the strong inhibition of the enzyme by the N-terminal peptide products released from enzymatically cleaved substrates.[77,78]

The first HCV protease inhibitor to enter clinical trials has been BILN 2061 (Ciluprevir; Boehringer Ingelheim), a macrocyclic inhibitor (Fig.3.4).[79] The development of BILN 2061 was built upon the observation that the carboxy-terminal carboxylic acid of the hexapeptide products provided an active-site affinity anchor and a moiety that warranted the desired selectivity for the HCV enzyme over the cellular proteases.[80,81] Structural information guided optimization of the side chains and the conversion of the linear peptidic leads into macrocyclic inhibitors with enhanced potencies and improved biopharmaceuticals properties,[82,83] culminating in the selection of BILN 2061 as a clinical candidate. Unfortunately, although an initial promising but transient effectiveness of the molecules, both cardiac toxicity[84] and development of resistance[85,86] determined the block of clinical investigation of BILN 2061.
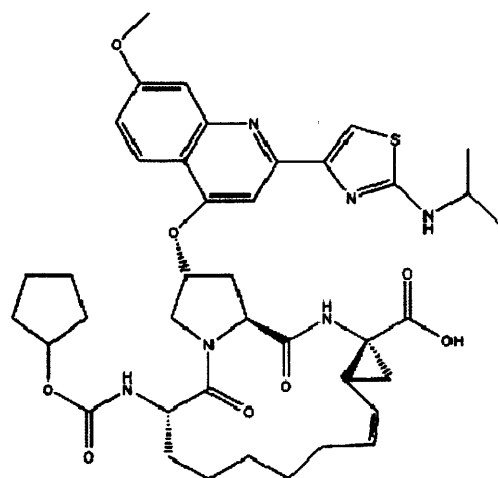
Fig. 3.4 - Chemical structure of the BILN 2061 (http://merops.sanger.ac.uk/smi/structures/biln2061.gif)

VX-950 (Vertex/Mitsubishi), a peptidomimetic inhibitor of the NS3-4A protease, stably interacts with the enzyme's active site thanks to the strategic inclusion of an α-ketoamide[87] (Fig.3.5), which is able to form a reversible covalent bond with the catalytic serine. Binding is potentially responsible for an unfavourable safety profile. As a consequence of the efficient inhibition of HCV replication in cell culture and of the promising results of phase Ia/Ib clinical trials,[88] further clinical studies have been encouraged and revealed development of viral resistance[86] and partial cross-resistance between BILN 2061 and VX-950.[89,90]
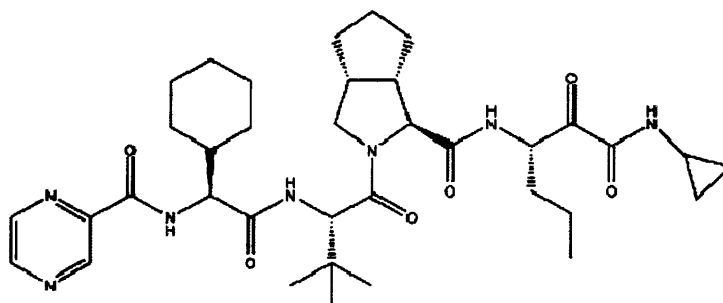
Currently other peptidomimetic inhibitors of the NS3-4A protease are being developed. In deed, significant efforts have been prompted towards the development of both pentapeptide α-ketoamide compounds showing weak HCV inhibition, and the 17-membered macrocyclic inhibitors, significantly more potent than the acyclic pentapeptide.[91] It has been also hypothesized that the macrocyclazion of 1,2,3,4-tetrahydroisoquinoline-3- carboxylamide (Tic) moiety to the P3 capping group could improve the binding through its extra contact with the Ala156 methyl group, and lead to the design of a less peptidic HCV inhibitor.[92]

However, it is fundamental to bear in mind that as the mutation of a single amino acid is sufficient to cause cross-resistance, the combination of protease inhibitors could result of scarce help in clinical practice.

3.6.3 NS5B polymerase inhibitors

The NS5B protein includes the RNA-dependent RNA polymerase (RdRp), the catalytic component of the HCV RNA replication[93] that, by synthesizing RNA using an RNA template uniquely in viral cells, can be targeted by selective antiviral agents. The structure of the NS5B enzyme is similar to that of a right hand.[9-96] The active site of the enzyme is placed in the palm domain whereas interactions with the RNA chain involve the fingers and the thumb, to which two loops extending from the fingers domain (the fingertips) are in contact, thus forming an active site similar to a tunnel where the RNA template, the nascent RNA strand and the nucleotide substrates are placed during the polymerization process.[70] The enzyme has been observed also in a more 'open' conformation, where there is no contact between the fingertips and the thumb and a wider entry to the catalytic site has replaced it.[97] This is the 'open', inactive conformation in which the enzyme appeared to be blocked by the interaction with allosteric inhibitors. Rational drug design identified various nucleoside analogues, whereas highthroughput screening discovered several non-nucleoside inhibitors (NNIs). The former class of inhibitors differentiates from the latter not only for distinct binding site and diverse pattern of resistance, but mainly for mechanism of action. Indeed, once converted by the host cell to the corresponding nucleotide and incorporated in the nascent RNA by the viral polymerase, the analogue leads to premature termination of the RNA synthesis. NNIs, instead, act as allosteric inhibitors by blocking the enzyme and preventing a conformational transition required for initiation of RNA synthesis.[98] Interestingly, these functional differences can result effective in combination therapy.

3.6.4 Nucleoside analogues

Up to day, the oral prodrug of 2'-C-methyl-cytidine (Fig.3.6), NM283 (Valopicitabine; Idenix/Novartis), is the only inhibitor of the NS5B polymerase, the antiviral activity at tolerated doses of which has been demonstrated in clinical trials.[99] This encouraging data suggested longer clinical trials in combination with pegylated IFN-α,[100] even if NM283 was not as effective or rapid as BILN 2061 and VX-950. Interestingly, it has been observed inhibition of NS5B enzymes and replicons derived from different HCV genotypes by 2'-C-methyl-nucleosides that consequently may

87

represent a potential treatment for all viral strains[101,102] if the shown development of resistance can be limited.[103,104] Surprisingly, it is the addition of a methyl group used to convert the enzyme substrates into inhibitors to be recognized by HCV and helps it to discriminate between these agents and the natural nucleotides.

Recently, the pyrimidine nucleoside beta-D-2'-deoxy-2'-fluoro-2'-C-methylcytidine has been developed as a hepatitis C virus RNA-dependent RNA polymerase (HCV RdRp) inhibitor. This molecule is able to increase inhibitory activity in the HCV replicon assay compared to 2'-C-methylcytidine and has low cellular toxicity.[105]

Furthermore, a series of adenosine 5'-phosphonate analogues have been designed to mimic naturally occurring adenosine monophosphate. Biological data have shown that some of these analogues function as chain terminators, whereas other derivates are competitive inhibitors with ATP.[106]
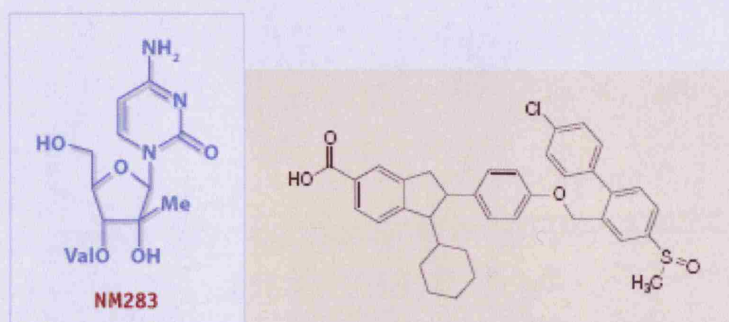
Fig. 3.6 - Chemical structure of NM283 (http://www.idenix.com/images/nm-283-sm.gif) and of the analogue of JTK-003, an orally active inhibitor of non-structuralprotein 5B (NS5B) (adapted from Seng-Lai Tan et al., Hepatites C Therapeutics: Current status and emerging strategies, *Nature Reviews, Drug Discovery*, Vol.**1** , pp. 867 – 881, 2002).

3.6.5 Non-nucleoside inhibitors

Few are the clinical data available on the three distinct classes of NNIs of the NS5B polymerase. JTK-109 and JTK-003 (Japan Tobacco) are the first oral NNIs of the NS5B polymerase entering clinical and belong to a heterogeneous series of 6,5-fused heterocyclic compounds based on a benzimidazole or indole core (Fig.3.6) (patent literature).[107] As allosteric inhibitors these drug candidates halt polymerase activity before elongation by impeding the conformational transition required to form a productive polymerase–RNA complex.[108,109] These inhibitors tend to bind on the

surface of the thumb domain, specifically in a pocket where one of the fingertips accommodates when the enzyme is free,[110] thus destroying the fingertips–thumb interactions and directing the enzyme into an 'open', inactive conformation. A single mutation within the inhibitor binding pocket is sufficient to develop resistance to this class of antiviral agents.

R803 (Rigel) and HCV-371, HCV-086 and HCV-796 (ViroPharma/Wyeth) are other polymerase NNIs to have entered clinical trials. Once identified and optimized using the replicon assay[111] and subjected to further investigation, R803 and related inhibitors have been found targeting the NS5B polymerase with a mechanism of action distinct from that shown by other known NNIs. Due to its scarce efficacy, R803 was soon excluded from clinical studies;[112] the development of new derivates and of a prodrug of R803 is progressing paying major attention to pharmacokinetic properties.

HCV-371, HCV-086 and HCV-796 are allosteric inhibitors of the NS5B polymerase but only HCV-796, a more potent analogue of HCV-086, has shown significant antiviral activity in an animal model of hepatitis C infection.[113]

Among other NS5B NNIs reported to inhibit HCV replication in tissue culture and currently in development,[96] a group of thiophene derivatives, acting as reversible allosteric inhibitors of the NS5B polymerase, has been investigated.[114,115] By binding to the enzyme in a cleft close to, but distinct from, the site occupied by benzimidazole-based inhibitors in the thumb domain,[95] these candidates induces the enzyme to adopt the 'open' conformation. Mutations is the main cause of resistance to this class of inhibitors but, following the observation that phenylalanine or dihydropyranone-based inhibitors share the same binding site, major efforts will point to further optimization.[116,117]

Benzothiadiazine-based allosteric inhibitors of the HCV polymerase[118] inhibit RNA synthesis before the formation of an elongation complex occurs[119,120] by binding to a different site and acting with distinct mechanisms from the allosteric inhibitors mentioned above.[120] The activity of these compounds has been found limited to a subset of enzymes and replicons.[101,102]

Following recent reports that indole-N-acetamides act as potent allosteric inhibitors of the HCV NS5B polymerase, optimized analogues devoid of PXR activation retain efficacy showing acceptable pharmacokinetic parameters in rat and dog.[121,122]

While traditional structure-activity relationships and structural studies have recently identified the mechanism of action for benzylidene substituted rhodanine derivates, several of which possess drug-like properties, as unique, reversible, covalent inhibitors of HCV NS5B,[123] molecular modeling approach resulted to be important and invaluable in designing new leads based on aryl dyketo acid scaffolds as HCV RdRp inhibitors,[124] and, through high-throughput screening (HTS), revealed a new class of HCV polymerase inhibitors, among which 1-(2 cyclopropylethyl)-3-(1,1-dioxo-2H-1,2,4-benzothiadiazin-3-yl)-6-fluoro- 4-hydroxy-2(1H)-quinolinone is the most representative inhibitor with excellent potency in biochemical and cellular assays as well as attractive molecular properties for advancement as a clinical candidate.[125]

3.7 Nucleic-acid-based antiviral agents

Lately more attention has been focused to the use of synthetic nucleic acids as drugs: antisense oligonucleotides, ribozymes and recently, siRNAs may represent potential new therapeutic agents in different areas. Nucleic-acid-based drugs' potentiality relies on the "efficient delivery of the synthetic polymers to the appropriate cells *in vivo*".[70] Both size and chemical nature make nucleic-acid polymers not orally but only parenterally bioavailable and, as systemic administration is ideal for the liver and less efficient for other organs,[126] HCV could successfully guide the identification and development of potent nucleic-acid-based drugs.

Considerable amount of data on IRES structure and function and its conservation among HCV genotypes have led to the identification of IRES as a valuable target for the development of RNA-based agents. The ribozyme RPI.13919 (Heptazyme; RPI) and the antisense oligonucleotide ISIS-14803 (ISIS Pharmaceuticals) demonstrated to reduce the HCV RNA translation and replication in cell culture, thus progressing to early-phase clinical studies in HCV-infected patients[127] then stopped due to the side effects or poor efficacy.

3.8 RNA interference approach to HCV antivirals

RNA interference (RNAi) is the latest nucleic acid based approach being investigated for HCV therapy.[128-132] As siRNA and short hairpin RNA (shRNA) directed against the viral genome in tissue culture, RNAi interferes with the replication of HCV replicons but, conversely to siRNA showing variable efficiency, siRNA acts on the whole HCV genome including conserved regions, and the most effective one can eradicate HCV from more than 98% of the replicon-bearing cells.[132] Following these noticeable *in vitro* results, the challenging task of delivering siRNA *in vivo* has been undertaken and partly slowed by the more promising use of chemically modified siRNA with enhanced stability and cell penetration and the ability to inhibit HCV replication both in tissue culture and in a mouse model of HCV infection.[133]

As the emergences of resistant replicons is diminished by the combination of two or more siRNAs,[70] the problem of resistance development can be overcome by using two or more siRNAs targeting different sequences of the viral genome. First clinical agent treating HCV infection through RNAi, BLT-HCV (Benitec) consists of three components targeting different HCV sequences and highlights the potentiality of a multi-targeting approach in preventing resistance development.[134]

3.9 Novel immunomodulatory agents

IFN-based treatments have clearly demonstrated that stimulation of the host innate and adaptive immunity can halt and eradicate HCV infection. As a consequence, synthetic agonists of Toll-like receptors (TLRs) 7 and 9 (molecular guards sensing the presence of invading microorganisms[135]and expressed by immune cells)[136] appear effective in controlling HCV infection in initial clinical studies. Innate and adaptive immune systems, dysfunctional in hepatitis-C infected patients, can be restored stimulating specific TLRs[137] because the immune response can be mimicked by synthetic agonists of TLRs.[70]

3.9.1 Antiviral activity of TLR-9 and TLR-7 agonists

Currently known potent agonists of TLR-9 are short synthetic oligonucleotides with one or more unmethylated CpG motifs flanked by specific sequences[138] and induce B-cell proliferation and antibody secretion, as the promising CPG-10101 (Actilon; Coley Pharmaceutical Group), has shown.[139]

Conversely, synthetic derivates structurally related to nucleic acids (imidazoquinolines, loxoribine (7-allyl-7,8-dihydro-8-oxoguanosine), and bropirimine (2-amin-5- bromo-6-phenyl-4(3)-pyrimidinone) are able to recognize TLR-7[135] and stimulate a potent, broad-spectrum antiviral response by inducing the release of inflammatory cytokines, especially IFN-α.

Once the immunostimulatory activity of ANA245 (7-thia-8-oxoguanosine or Isatoribine; Anadys Pharmaceuticals) was demonstrated,[140] the development of the corresponding prodrug, ANA975, has soon started[141] aiming to combine the broad spectrum efficacy of an immune-based therapy with the advantageous administration of an oral drug.[70]

3.10 Future challenges

To all novel drugs entering clinical evaluation to determine their clinical activity for HCV patients, safety and efficacy as well as the ability to inhibit all viral variants and prevent the emergence of resistance are required.

Currently, the best approach to control HCV infection and overcome the risk of resistance development is the administration of combinations of antiviral agents attacking different viral and possibly not host targets.

Development of novel antiviral agents inhibiting unexploited viral targets (p7 ion channel, NS2-3 cysteine protease and NS3 helicase) should be the challenge of the next future for researchers, in order to achieve and guarantee eradication of the HCV infection in as many patients as possible, possibly using effective combination schemes of new HCV-targeted drugs with eventual stimulation of the host immune system (immunomodulators, antiviral vaccines).

Aims & Objectives

HCV is a highly prevalent and damaging infection which urgently requires new and improved treatments. The present project is focuses on the helicase/NTPase enzyme, the crystal structure of which can be used for small-molecule inhibitor design,[142] and with this purpose two different approaches have been employed.

Based on inhibitor design and synthesis work previously carried out,[i] the former implied design and evaluation of a database of related candidate anti-HCV compounds by means of computational methods; for the most promising compounds, chemical synthesis has been attempted aiming to find a simple and straightforward procedure.

The latter approach focused on structural modifications and systematic alterations in structure to produce more active compounds and build upon data and SAR previously described.[143,144,145]

Those compounds successfully synthesised and tested through an HCV replicon assay together with additional structurally related derivatives have been evaluated by molecular modelling to rationalise any biological data obtained and identify those compounds with the greatest potential for further research. Considering the high degree of motif conservation between helicases of *flaviviruses*, the binding potential of the chosen structures has been evaluated against both HCV and dengue helicase[146] using molecular modelling and a predicted model of dengue helicase.[147]

Results and Discussion

HCV Helicase Inhibitors - I

Based on the general structural requirements suggested by Phoon et al. and the most active inhibitors of HCV helicase previously reported,[143,144] novel compounds to be synthesised were designed substituting the benzimidazole end-group with structurally related groups (Table 3.1a) with the purpose to evaluate their efficacy as limited data are available on the effect that substituent ring position has on efficacy.

The amide substitution in both para and meta aromatic positions (Table 3.1a) was selected for initial synthesis because it is closely related to the benzimidazole ring, relatively simple to synthesise, interesting for comparison with ester derivatives previously reported[145] and for providing evidence of the location of hydrogen bonding regions within the helicase enzyme.

The linker was chosen considering those which have been shown to be favourable and readily available as an acid dichloride to simplify synthesis (Table 3.1a, b). A variety of chain lengths were selected ranging from 4 to 6 carbons to determine the optimum chain length for helicase binding.

Table 3.1a – Benzimidazole ring replacements, substituent positions, linker reagents.

94

Linker Reagents		
Succinyl Dichloride	**Glutaryl Dichloride**	**Adipoyl Dichloride**

Table 3.1b – Substituent positions, linker reagents.

Linker Reagents				
Succinyl Dichloride	**Glutaryl Dichloride**	**Adipoyl Dichloride**	**Fumaryl Dichloride**	**Terephthaloyl Dichloride**
I	II	III	IV	V

Substituent positions (x = linker)		
A	B	C
D		

Chemistry - I

The general mechanism of the reactions attempted involves the formation of an amide bond through acylation of the primary aromatic amine, with an acid chloride. Important was the presence of a base to remove HCl formed as a by-product during the reactions to avoid the formation of an amine salt with the starting reagents, which would stop the reaction. According to the general mechanism of the reactions (Fig. 3.7) essential were anhydrous conditions (dry solvents and nitrogen atmosphere) to prevent hydrolysis of the acid chloride (Fig. 3.8) and, therefore, the formation of a less reactive carboxylic acid on the linker, lowering yields.
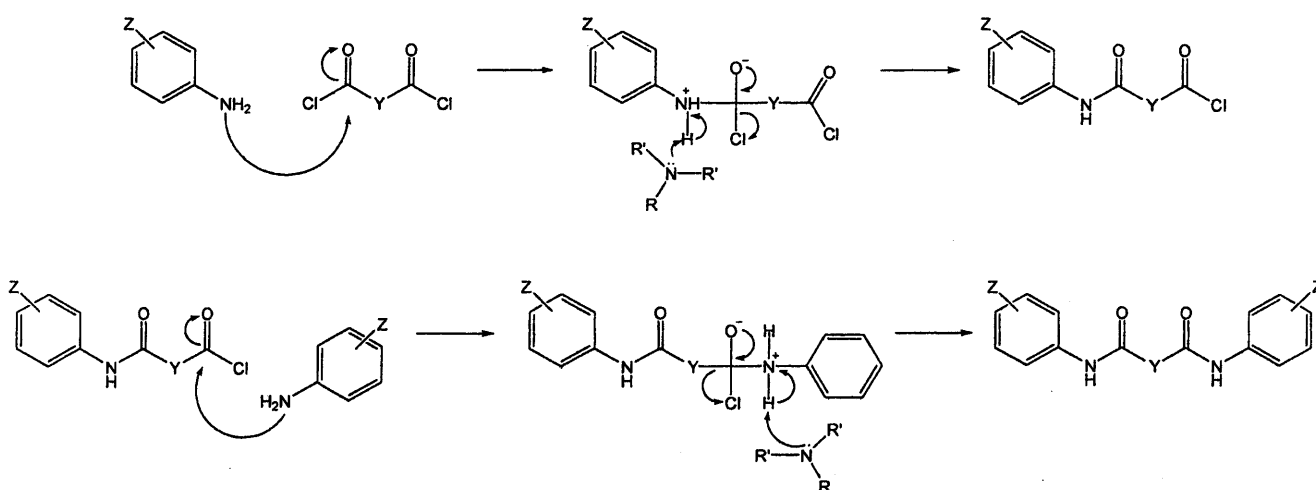
Fig. 3.7 - General mechanism of the attempted reactions.
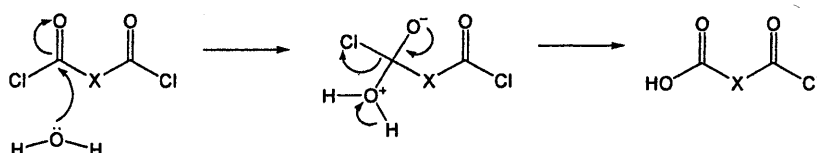
Fig. 3.8 - Inactivation of acid chloride linker by hydrolysis.
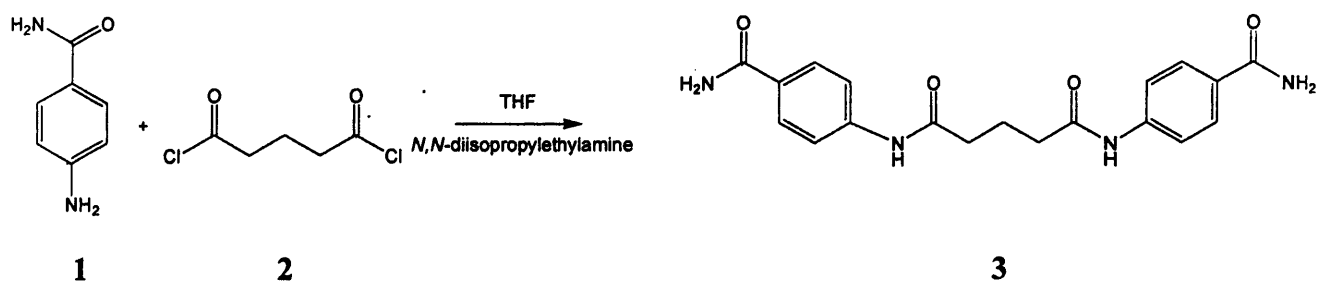
Scheme 3.1 - Synthesis of Pentanedioic acid bis-[(4-carbamoyl-phenyl)-amide]

Compound **3** was prepared by reacting **1** and **2** directly in a ratio of 2:1. As aromatic amines are more reactive, protective groups were not used to prevent reaction with the amide; in fact, compared to a carbon with phenylamines, the amide dissociates its lone pair of electrons to an oxygen that is more electronegative and forms a more stable conformation. Initially, a semi-solid formed within an opaque white solution and a white gas was observed above the solution, features common to all reactions and due to the hydrochloride salt of N, N' - diisopropylethylamine (DIPEA) (2.1 eq) and the by-product HCl respectively. After 24 hours TLC indicated a product had formed, the solution was filtered and washed with THF before column chromatography was used to isolate the required product from the mixture. A stationary phase of silica and an eluent of DCM were used but soon replaced by ethyl acetate as elution thorough the column was too slow. Although ^{1}H NMR indicated the sample remained impure, there were several characteristic peaks indicating compound **3** may have formed; removal of the remaining impurities was attempted by washing with petroleum ether, which precipitated a white solid, and diethyl ether to further purify the compound, the ^{1}H NMR and ^{13}C NMR of which showed the sample was sufficiently pure for biological testing on HCV helicase.
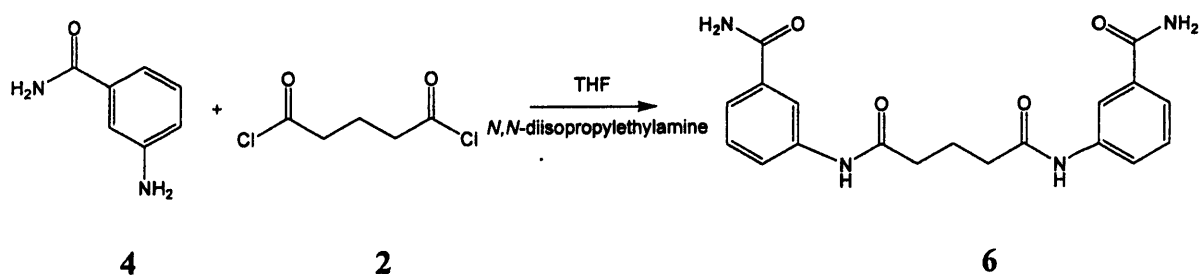
Scheme 3.2 - Synthesis of Pentanedioic acid bis-[(3-carbamoyl-phenyl)-amide]

Preparation of compound **6** resulted in the formation of an opaque white solution. After stirring at room temperature for 20 hours, TLC indicated a spot similar to that seen with compound **3** but less defined than previously. Filtration and washing with THF followed before ^{1}H NMR and TLC of the remaining oil clearly indicated still presence of impurities and, therefore, the necessity of purification by chromatography column. The spot presumed to contain compound **6** was isolated and the oil collected was washed with petroleum ether trying unsuccessfully to crystallise a solid, and analysed via ^{1}H NMR revealing that the sample was impure and very little if any of the expected compound was formed. Although NMR and TLC data suggested that the reaction may have been successful, the yields were too low to allow further purification or biological testing to be carried out.

As slow reaction kinetics might have been the explanation for the low yield obtained, the mixture of the second attempted synthesis of compound **6** was stirred at room temperature for 72 hours before purification. After initial purification - filtration, washing with THF and an extraction between organic and aqueous phases - chromatography was then used to separate the product from impurities. Various eluents were evaluated to determine which would achieve the optimum separation (among ethyl acetate, DCM, chloroform, DCM/hexane and chloroform/diethyl ether, ethyl acetate was best) and both flash chromatography and column chromatography were attempted without achieving any separation with either. Crystallisation with methanol precipitated a too small quantity of solid for further work, so the sample was discarded.

Moisture in the atmosphere other than slow reaction kinetics was supposed to be responsible for the low yield observed in the reaction. Therefore, in the third attempted synthesis, it was ensured anhydrous conditions were maintained and the formation of amide bonds was aided by increasing the quantity of base to remove HCl

more effectively and providing energy by heating the reaction at 40°C. After 48 hours, as TLC showed the yield was very low again, column chromatography was avoided and A crystallisation with methanol afforded a brown solid and light brown solution. Neither the solid nor the solution analysed by ^{1}H NMR showed presence of the expected product.

Reagents contaminated with water were proposed to cause the failure of previous attempts. Nevertheless, drying the reagents, using a new bottle of anhydrous THF, maintaining a nitrogen atmosphere throughout the reaction for 24 hours did not aid the reaction to achieve completion as both TLC and ^{1}H NMR proved that no product had formed.

It is unclear why the meta substituted derivative could not be synthesised; the closer proximity of functional groups compared to the para substituted compound 3, together with the already low yields, may have significantly inhibited the reaction.
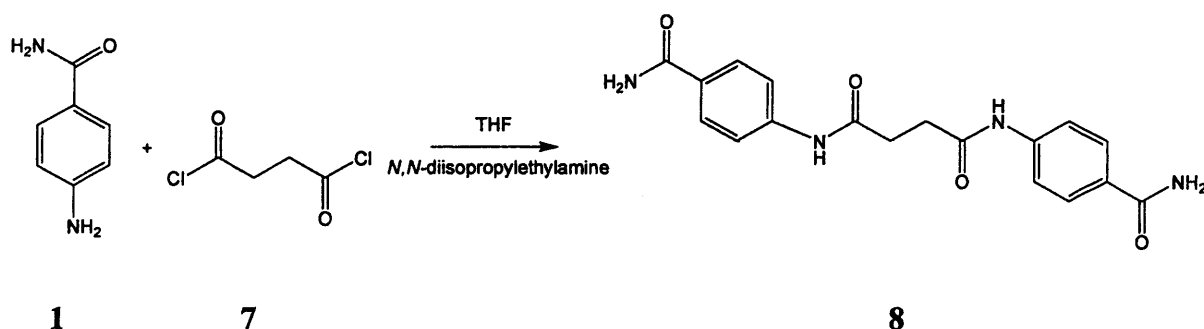
Scheme 3.3 - Synthesis of (N,N'-Bis-(4-carbamoyl-phenyl)-succinamide).

Synthesis of compound 8 resulted in a dark brown solution forming. After stirring the reaction overnight, a well-defined spot between the two reagents appeared on TLC; Purification via filtration, extraction and column chromatography successfully isolated compound 8, the presence of which was confirmed by ^{1}H NMR.

Washing the sample with diethyl ether was attempted to further purify the compound from some impurities revealed by the ^{1}H NMR; unfortunately, although compound 8 was precipitated as a relatively pure white solid, the quantity was insufficient for the HCV replicon assay and the reaction was repeated trying to improve the yield.

More in detail, a double quantity of base was used in the second synthesis to aid the formation of amide bonds. TLC indicated that some products had formed after

24 hours but it was attempted to improve the yield by leaving the reaction for a further 72 hours.

An extraction that left white and brown solids was the initial purification; in order to avoid column chromatography to maximise yield, that was·expected to be lower than previously as TLC spots were less intense, crystallisation with methanol was attempted unsuccessfully.

NMR data suggested presence of impure succinoyl dichloride that could have prevent the reaction to work achieving completion.
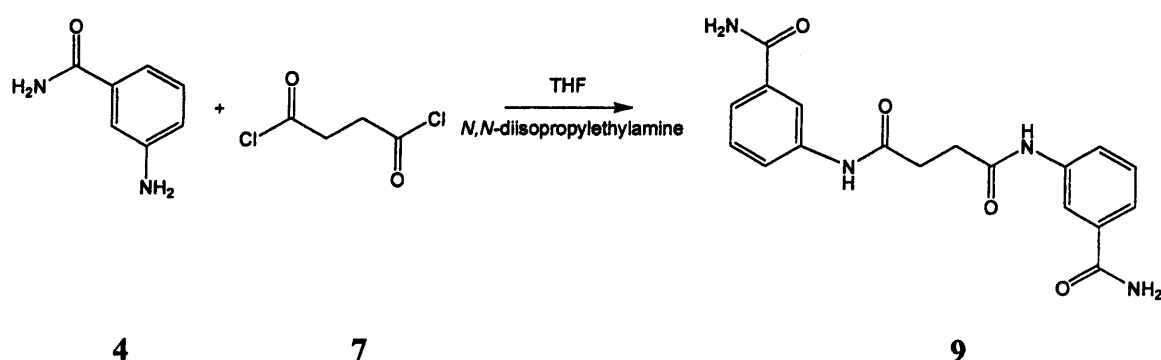
Scheme 3.4 - Synthesis of N,N'-Bis-(3-carbamoyl-phenyl)-succinamide.

Difficulties were anticipated with this compound as previous attempts to synthesise a meta substituted compound from this series were unsuccessful. After stirring for a longer period of 72 hours to maximise the yield, a light brown solution, that TLC revealed to contain a small amount of product, started to appear and was allowed to proceed for a further 24 hours before purification. Once the mixture was filtered, washed with THF and extracted with ethyl acetate, crystallisation with methanol precipitated a small amount of product. Alternative purification via column chromatography was considered impractical because the low yield would have made it difficult to monitor the elution of the compound and small losses would have been significant. Nevertheless, the quantity of solid was too small to be recovered from the filter paper and TLC indicated that no product remained in the filtered solution.
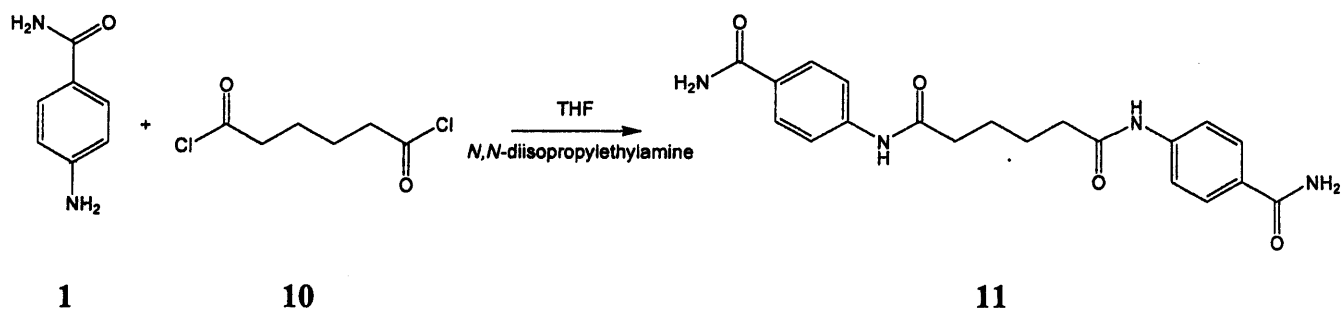
Scheme 3.5 - Synthesis of Hexanedioic acid bis-[(4-carbamoyl-phenyl)-amide]

Preparation of compound **11** resulted in an opaque orange solution after 24 hours when TLC analysis revealed that a product had formed. Following extraction with ethyl acetate and water, column chromatography isolated several products; the ^{1}H NMR on two of the separated products showed the latter potentially containing the expected compound with also few impurities, for the removal of which crystallisation with methanol was attempted.

This afforded only a small amount of solid, later discarded as it contained an insufficient yield for testing and was shown to be still impure by ^{1}H NMR. A second attempt was made to improve yield by doubling the quantity of base and prolonging the reaction period. As after 24 hours TLC monitoring indicated that only the reagents were present, the amount of base was increased further and the mixture allowed to react for an additional 6 days when TLC showed the presence of some product; however, the yield was again too low and the sample was abandoned.
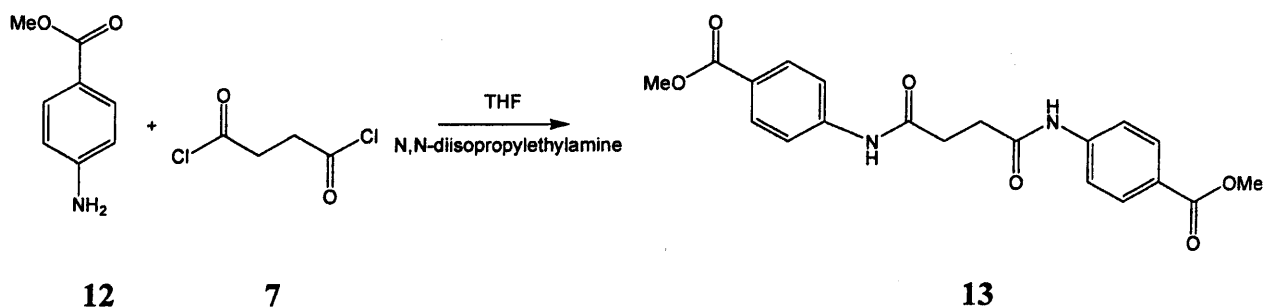
Scheme 3.6 - Synthesis of *N, N'*- bis -(4-methyl ester formoyl-phenyl)-succinamide

Compound **13** was prepared by reacting **12** and **7** directly in a ratio of 2:1 dissolved in THF anhydrous (15 mL). Initially, a semi-solid formed within an opaque white solution and a white gas was observed above the solution, features common to all reactions and due to the hydrochloride salt of *N, N'*- diisopropylethylamine (DIPEA) (2.1 eq) and the by-product HCl respectively. After stirring at room temperature for 24 hours, the solution was filtered and washed with THF, concentrated *in vacuo* and extracted (AcOEt), before column chromatography (AcOEt) was used to isolate the required product from the mixture. Removal of the remaining impurities was obtained by crystallization with MeOH; both ^{1}H NMR and ^{13}C NMR of the final product showed the sample was sufficiently pure for biological testing on HCV helicase.

Scheme 3.7 - Synthesis of Pentanedioic acid bis-[(4-methyl ester formoyl-phenyl)amide]

Compound **14** was prepared by reacting **12** and **2** directly in a ratio of 2:1 dissolved in THF anhydrous (15 mL). Also in this case, a semi-solid formed within an opaque white solution and a white gas was observed above the solution, features common to all reactions and due to the hydrochloride salt of *N, N'*- diisopropylethylamine (DIPEA) (2.1 eq) and the by-product HCl respectively. After stirring at room temperature for 24 hours, the solution was filtered and washed with THF, concentrated *in vacuo* and extracted (AcOEt), before column chromatography (AcOEt) was used to isolate the required product from the mixture. Removal of the remaining impurities was obtained by crystallization with MeOH; both ^{1}H NMR and ^{13}C NMR of the final product showed the sample was sufficiently pure for biological testing on HCV helicase.

Scheme 3.8 - Synthesis of Hexanedioic acid bis-[(4-methyl ester formoyl-phenyl)amide]

Compound **14** was prepared by reacting **X** and **10** directly in a ratio of 2:1 dissolved in THF anhydrous (15 mL). Again, a semi-solid formed within an opaque white solution and a white gas was observed above the solution, features common to all reactions and due to the hydrochloride salt of *N, N'*- diisopropylethylamine (DIPEA) (2.1 eq) and the by-product HCl respectively. After stirring at room temperature for 24 hours, the solution was filtered and washed with THF, concentrated *in vacuo* and extracted (AcOEt), before column chromatography (AcOEt) was used to isolate the required product from the mixture. Removal of the remaining impurities was obtained by crystallization with MeOH; ^{1}H NMR of the final product showed the sample was not sufficiently pure for biological testing on HCV helicase and the compound was abandoned.
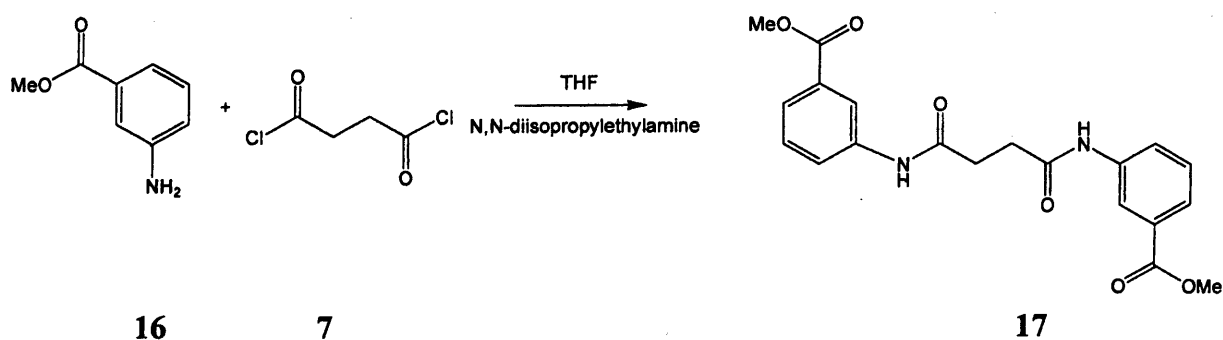
Scheme 3.9 - Synthesis of *N, N'*- bis -(3-methyl ester formoyl-phenyl)-succinamide

103

Compound **17** was prepared by reacting **16** and **7** directly in a ratio of 2:1 dissolved in THF anhydrous (15 mL). Again, a semi-solid formed within an opaque white solution and a white gas was observed above the solution, features common to all reactions and due to the hydrochloride salt of *N, N'*- diisopropylethylamine (DIPEA) (2.1 eq) and the by-product HCl respectively. After stirring at room temperature for 24 hours, the solution was filtered and washed with THF, concentrated *in vacuo* and extracted (AcOEt), before column chromatography (AcOEt) was used to isolate the required product from the mixture. Removal of the remaining impurities was obtained by crystallization with MeOH; both ^{1}H NMR and ^{13}C NMR of the final product were not completely clear and the sample was considered not sufficiently pure for biological testing on HCV helicase.

Scheme 3.10 - Synthesis of Pentanedioic acid bis-[(3-methyl ester formoyl-phenyl)-amide]

Compound **18** was prepared by reacting **16** and **2** directly in a ratio of 2:1 dissolved in THF anhydrous (15 mL). A semi-solid formed within an opaque white solution and a white gas was observed above the solution, features common to all reactions and due to the hydrochloride salt of *N, N'*- diisopropylethylamine (DIPEA) (2.1 eq) and the by-product HCl respectively. After stirring at room temperature for 24 hours, further DIPEA was added (2.1 eq) and the reaction was left stirring overnight. Subsequently, the solution was filtered and washed with THF, concentrated *in vacuo* and extracted (AcOEt). Removal of the remaining impurities was obtained only by crystallization with MeOH; both ^{1}H NMR and ^{13}C NMR of the final product showed the sample sufficiently pure for biological testing on HCV helicase.
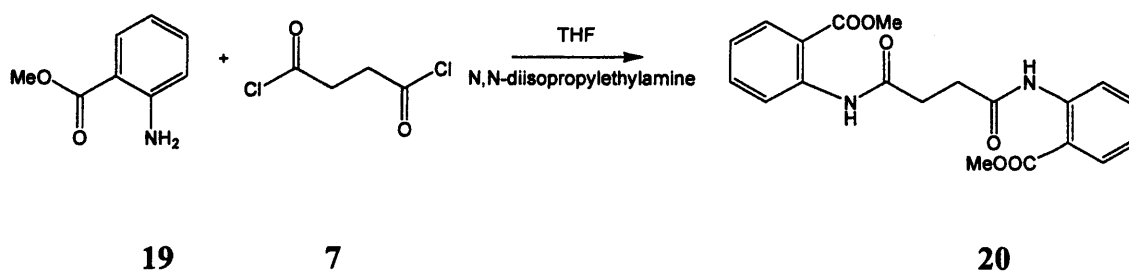
Scheme 3.11 - Synthesis of *N, N'*- bis -(2-methyl ester formoyl-phenyl)-succinamide

Compound **20** was prepared by reacting **19** and **7** directly in a ratio of 2:1 dissolved in THF anhydrous (15 mL). Features common to all reactions, a semi-solid formed within an opaque white solution and a white gas was observed above the solution, due to the hydrochloride salt of *N, N'*- diisopropylethylamine (DIPEA) (2.1 eq) and the by-product HCl respectively. After stirring at room temperature for 24 hours, the reaction mixture appeared as a green solution that was concentrated *in vacuo* and directly crystallized with MeOH providing a pure white solid. Both ^{1}H NMR and ^{13}C NMR of the final product showed the sample was pure for biological testing on HCV helicase.
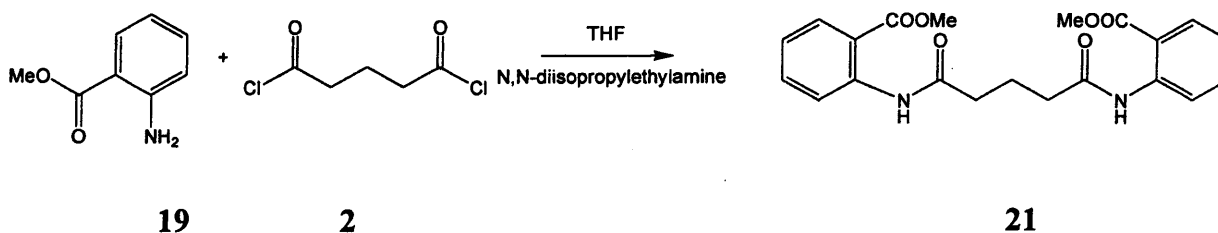
Scheme 3.12 - Synthesis of Pentanedioic acid bis-[(2-methyl ester formoyl-phenyl) amide]

Compound **21** was prepared by reacting **19** and **2** directly in a ratio of 2:1 dissolved in THF anhydrous (15 mL). Features common to all reactions, a semi-solid formed within an opaque white solution and a white gas was observed above the solution, due to the hydrochloride salt of *N, N'*- diisopropylethylamine (DIPEA) (2.1 eq) and the by-product HCl respectively. After stirring at room temperature for 24 hours, the reaction mixture appeared as a green solution that was concentrated *in vacuo* and directly crystallized with MeOH providing a pure white solid. Both ^{1}H NMR and

105

^{13}C NMR of the final product showed the sample was pure for biological testing on HCV helicase.

MeO ... NH$_2$ + Cl ... Cl → (THF, N,N-diisopropylethylamine) COOMe ... MeOOC

19   **10**   **22**
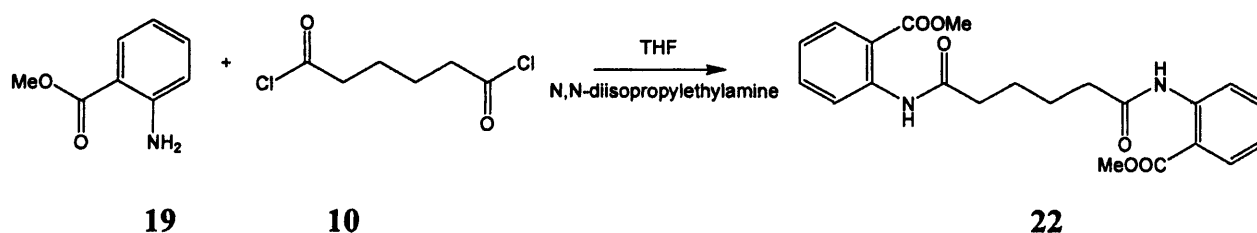
Scheme 3.13 - Synthesis of Hexanedioic acid bis-[(2-methyl ester formoyl-phenyl) amide]

Compound **22** was prepared by reacting **19** and **10** directly in a ratio of 2:1 dissolved in THF anhydrous (15 mL). Features common to all reactions, a semi-solid formed within an opaque white solution and a white gas was observed above the solution, due to the hydrochloride salt of *N, N'*- diisopropylethylamine (DIPEA) (2.1 eq) and the by-product HCl respectively. After stirring at room temperature for 24 hours, the reaction mixture appeared as a green solution that was concentrated *in vacuo* and directly crystallized with MeOH providing a pure white solid. Both ^{1}H NMR and ^{13}C NMR of the final product showed the sample was pure for biological testing on HCV helicase.

MeO ... NH$_2$ + Cl ... Cl → (THF, N,N-diisopropylethylamine) COOMe ... MeOOC

19   **23**   **24**
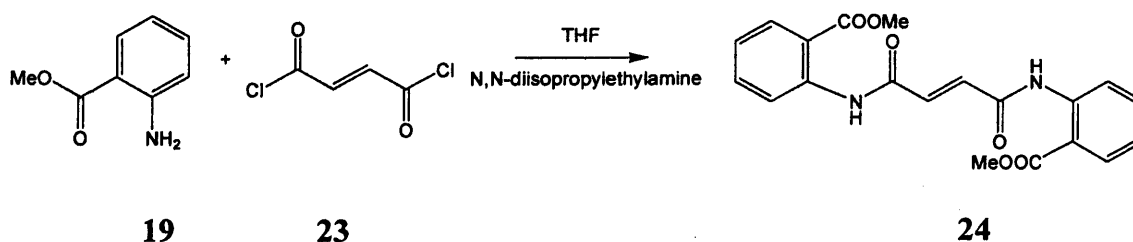
Scheme 3.14 - Synthesis of But-2-enedioic acid bis-[(2-methyl ester formoyl-phenyl)-amide]

Compound **24** was prepared by reacting **19** and **23** directly in a ratio of 2:1 dissolved in THF anhydrous (15 mL). Again, a semi-solid formed within an opaque white solution and a white gas was observed above the solution, due to the hydrochloride salt of *N, N'*- diisopropylethylamine (DIPEA) (2.1 eq) and the by-product HCl respectively. An instantaneous reaction was observed and further THF anhydrous needed to be added (15 mL). After stirring at room temperature for 24 hours, the

reaction mixture appeared as a green solution that was concentrated *in vacuo* and directly crystallized with MeOH providing a pure white solid. Both ^1H NMR and ^{13}C NMR of the final product showed the sample was pure for biological testing on HCV helicase.

Scheme 3.15 - Synthesis of Terephthalic acid bis-[(2-methyl ester formoyl-phenyl)-amide]

Compound **26** was prepared by reacting **19** and **25** directly in a ratio of 2:1 dissolved in THF anhydrous (15 mL). Again, a semi-solid formed within an opaque white solution and a white gas was observed above the solution, due to the hydrochloride salt of *N, N'*- diisopropylethylamine (DIPEA) (2.1 eq) and the by-product HCl respectively. An instantaneous reaction was observed and further THF anhydrous needed to be added (15 mL). After stirring at room temperature for 24 hours, the reaction mixture appeared as a green solution that was concentrated *in vacuo* and directly crystallized with MeOH providing a pure white solid. Both ^1H NMR and ^{13}C NMR of the final product showed presence of two unclear compounds; before planning future modifications of the synthetic procedure, it was proposed to wait for the biological test results on HCV helicase.
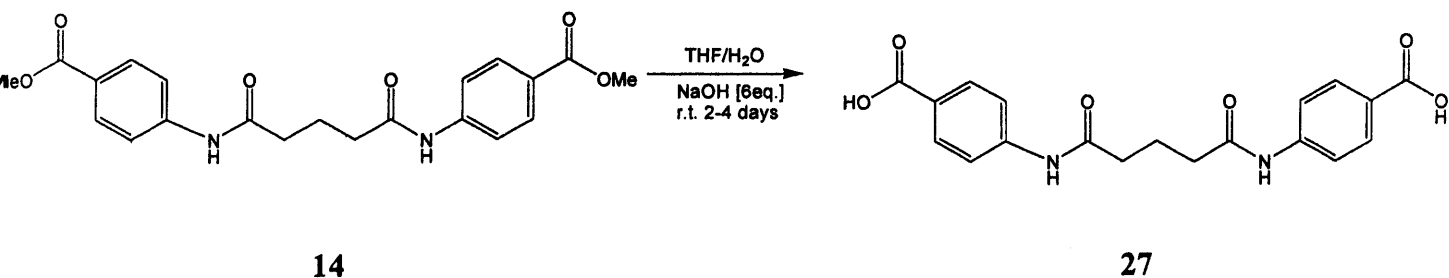
Scheme 3.16 - Synthesis of Pentanedioic acid bis-[(4-hydroxy formoyl-phenyl)-amide]

For the synthesis of the corresponding carboxylic acid of compound **14** as to increase the biological activity of this series of HCV helicase small molecule inhibitors, sodium hydroxide monohydrate (6eq) was added to a solution of compound 14 (1eq) in THF and water (1:1). The reaction mixture was then stirred at room temperature for four days. After dilution with water, the mixture was acidified with 2 N HCl until pH 2 was reached. The acid was extracted with AcOEt, washed with brine, and dried. Removal of the solvent by concentration *in vacuo* provided a pure compound as white solid.

Unique drawback of this protocol[148] could have been the long time the reaction required to reach the completion, but as there were no time restrictions and the synthetic procedure proved to be successful, it was decided to follow it also for the synthesis of the acidic form of both ester **20** and **22**.
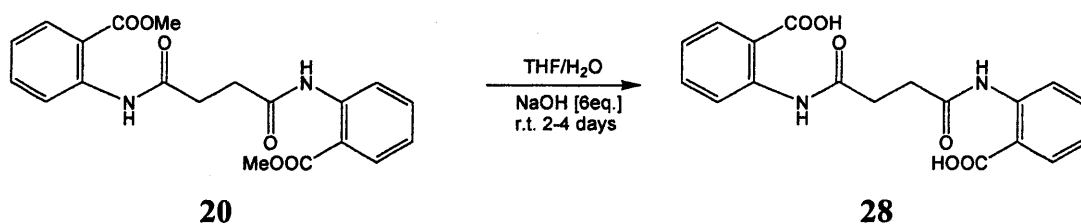
Scheme 3.17 - Synthesis of *N, N'*- bis -(2-hydroxy formoyl-phenyl)-succinamide

Also for the synthesis of the corresponding carboxylic acid of compound **20**, sodium hydroxide monohydrate (6eq) was added to a solution of compound 20 (1eq) in THF and water (1:1). Then the reaction mixture was stirred at room temperature for 48 hours, diluted with water, and acidified with 1 N HCl until pH 2 was reached. The acid was extracted with ethyl acetate, washed with brine, and dried. Removal of the

solvent by concentration *in vacuo* and precipitation from DCM to get rid of colour impurities afforded a pure compound.[148]
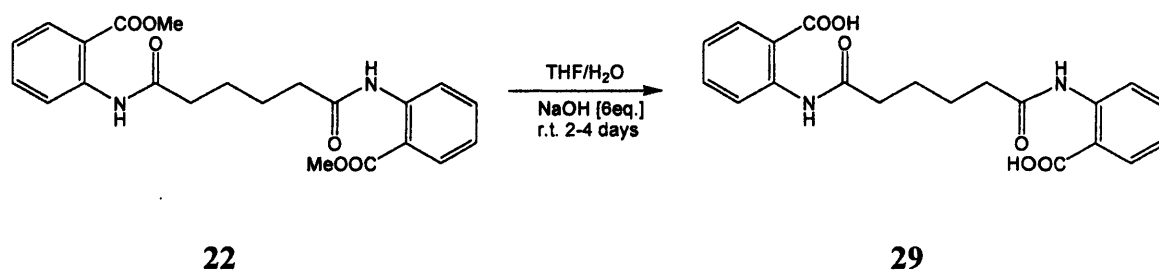
Scheme 3.18 - Synthesis of *N, N'*- bis -(2-hydroxy formoyl-phenyl)-succinamide

To synthesise the corresponding carboxylic acid of compound **22**, sodium hydroxide monohydrate (6 eq) was added to a solution of compound **22** (1eq) in THF and water (1:1). Then the reaction mixture was stirred at room temperature for 48 hours. After dilution with water, the mixture was acidified with 1 N HCl until pH 2 was reached. Extraction (AcOEt), removal of the solvent by concentration *in vacuo* and precipitation from DCM to get rid of colour impurities provided a pure white compound.[148]

Table 3.2a and b, illustrating compound structures, their corresponding names and yields, follow below.

Table 3.2a – Structure, name and yield of the synthesised compounds.

Structure	Name	Yield %
	CF-AB 14 **(14)**	42.03
	CF-AB 16 **(13)**	44.65
	CF-AB 18 **(15)**	24.56
	CF-AB 19 **(17)**	26.98
	CF-AB 20 **(18)**	4.27
	CF-AB 21 **(20)**	27.28
	CF-AB 22 **(21)**	47.27

Table 3.2b – Structure, name and yield of the synthesised compounds.

Structure	Name	Yield %
	CF-AB 23 **(22)**	4.84
	CF-AB 24 **(24)**	85.17
	CF-AB 17 **(3)**	39.67
	CF-AB 25 **(27)**	98.21
	CF-AB 26 **(28)**	31.13
	CF-AB-27 **(29)**	32.17

111

Focusing on the general structural requirements suggested by Phoon et al.,[143] it was attempted the synthesis of one of the most active inhibitors of HCV helicase previously reported,[144] with the purpose to design and synthesise a new series of related compounds if a feasible and successful synthetical pathway would have been found.
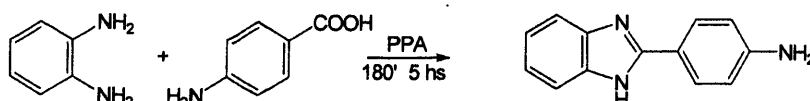
Scheme 3.19 - Synthesis of 4-(1H-benzo[d] imidazol-2-yl) benzenamine

For the synthesis of 4-(1H-benzo[d]imidazol-2-yl) benzenamine, intermediate to be used for the second final step of the proposed synthesis of the Vertex compound, N,N'-Bis-(1H-benzoimidazol-2-phenyl)-succinamide, a mixture of o-phenylenediamine (1eq) and aminobenzoic acid (1.06 eq) in polyphosphoric acid (13.5 eq) was heated with stirring in an oil bath at 180 °C for 5 hours, cooled to room temperature and poured into water. The resulting green precipitate needed to be added almost 500 mL of water to make a dark green sticky solid to go in solution, thus easing the precipitation of a light green powder and its filtering off. Subsequently, the solid was stirred with aqueous 10% sodium hydroxide and filtered off.[149]

For the synthesis of the Vertex compound, N,N'-Bis-($1H$-benzoimidazol-2-phenyl)-succinamide, it was decided to follow the same synthetic procedure used for the synthesis of the methyl 4-aminobenzote series as HCV helicase small molecule inhibitors.
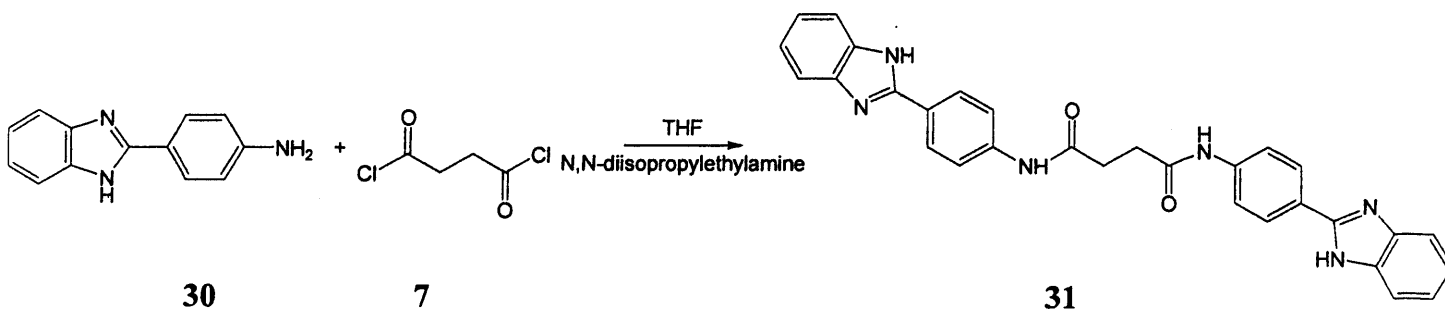
Scheme 3.20 - Synthesis of N,N'-Bis-($1H$-benzoimidazol-2-phenyl)-succinamide

Compound **31** was prepared by reacting **30** and **7** directly in a ratio of 2:1 dissolved in THF anhydrous (8 mL). After stirring at room temperature for 24 hours, the solution was filtered and washed with THF, concentrated *in vacuo* and precipitated with MeOH affording a brown solid, the ^{1}H NMR of which showed the sample was not pure. Following further crystallization (MeOH) to remove all remaining impurities, the ^{1}H NMR investigation revealed presence of the expected impure product in the MeOH filtrated; adding DCM to the MeOH solution, a new precipitate formed but still the desired compound was showed in the solvent by the ^{1}H NMR spectrum.

A column chromatography (CHCl$_3$/MeOH 5%) was therefore undertaken affording the isolation of the supposed final molecule, the NMR identification of which was quite difficult as its yield was too low and repeating the synthesis was required.

Table 3.3, illustrating the various approaches attempted from now on for the synthesis of N,N'-Bis-($1H$-benzoimidazol-2-phenyl)-succinamide, follows below.

Table 3.3 - Approaches attempted for the synthesis of *N,N'*-Bis-(1*H*-benzoimidazol-2-phenyl)-succinamide

SM	Linker	Base/Solvent	Act.Reagent	Condition
[2eq.]	Succinoyl dichloride [1eq.]	DIPEA [2.1eq.]/ THF anh.		r.t. o.n.
[2eq.]	Succinoyl dichloride [1eq.]	DIPEA [4eq.]/ DMF anh./ Pyridine 10%		r.t. o.n.
[2eq.]	Sodium Succinate [1eq.]	THF anh.	DCC [2.1eq.]	r.t. o.n.
[2eq.]	Sodium Succinate [1eq.]	DIPEA [2.92eq.]/ DMF anh.	BOP [1eq.]	r.t. 72 hs
[2eq.]	Succinoyl dichloride [1eq.]	DIPEA [2.1eq.]/ THF anh.	DMAP 10%	r.t. 72hs

A different synthetic pathway was followed in order to increase the scale and the overall yield of the reaction as well as to improve the solubility of **30** (thought to play a key role for the success of the synthetic procedure), thus easing the identification of the final product.
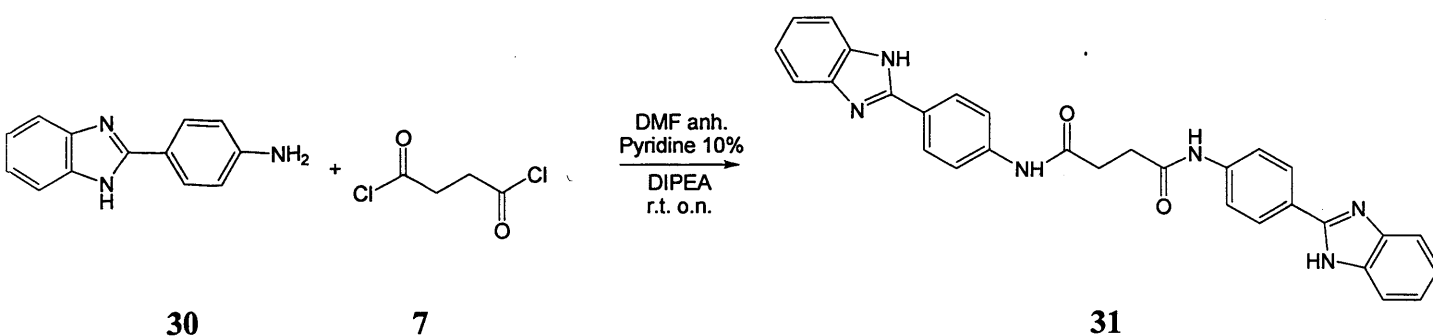
Scheme 3.21 - Synthesis II of *N,N'*-Bis-(1*H*-benzoimidazol-2-phenyl)-succinamide

Compound **31** was prepared by dissolving **30** (2 eq) in anhydrous dimethyl formammide (DMF) (2.24 mL) and anhydrous pyridine (10%) to increase the solubility of the starting material. Following the addition of DIPEA (4 eq) and **7**

extraction (AcOEt), both the organic and aqueous layers were analyzed by ^{1}H NMR; interestingly, the latter showed presence of the expected compound. Therefore, it was evaporated, added of hot MeOH to filtrate the eventual insoluble salt derived from the previous extraction and recover a yellow solution, the ^{1}H NMR of which revealed peaks of the desired product. As these signals were weak (probably due to the insolubility of the sample in both CDCl$_3$ and MeOD, and to traces of water), the compound was stored waiting for analyzing and comparing results coming from reactions set up using distinct approaches.

30   **7**   **31**
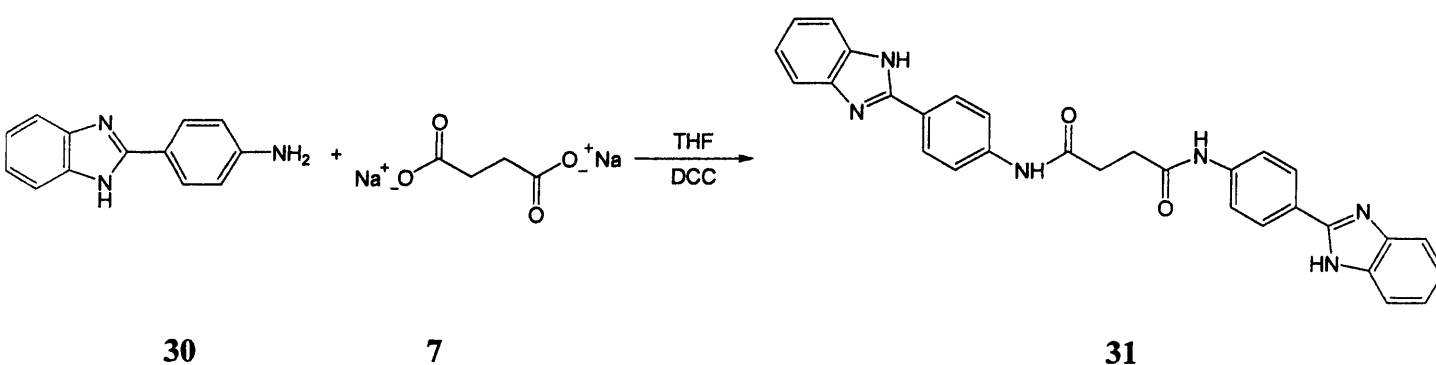
Scheme 3.22 - Synthesis III of *N,N'*-Bis-(1*H*-benzoimidazol-2-phenyl)-succinamide

Compound **31** was prepared by dissolving **30** (2 eq) in THF anhydrous (15 mL) and then added of a solution of sodium succinate (1 eq) and DCC (2.1 eq) in THF anhydrous. After stirring at room temperature for 48 hours, the reaction mixture was refluxed (70 °C) for 7 hours and left stirring for further 12 hours. As no spot on TLC was observed apart from that of the starting material, anhydrous DCM (4 mL) was added to improve both solubility and reactivity of DCC without obtaining any detectable results.

The presence of the salt instead of the acidic form of succinate was hypothesised being responsible for this attempt to fail. Consequently, the reaction was repeated according to the same protocol expect for the succinic acid that replaced sodium succinate.

Extraction (AcOEt) of the reaction mixture led to NMR and Mass Spectra investigation of the solid precipitated from the aqueous layer (MeOH); both analyses were neither clear nor consistent with the structural data expected.
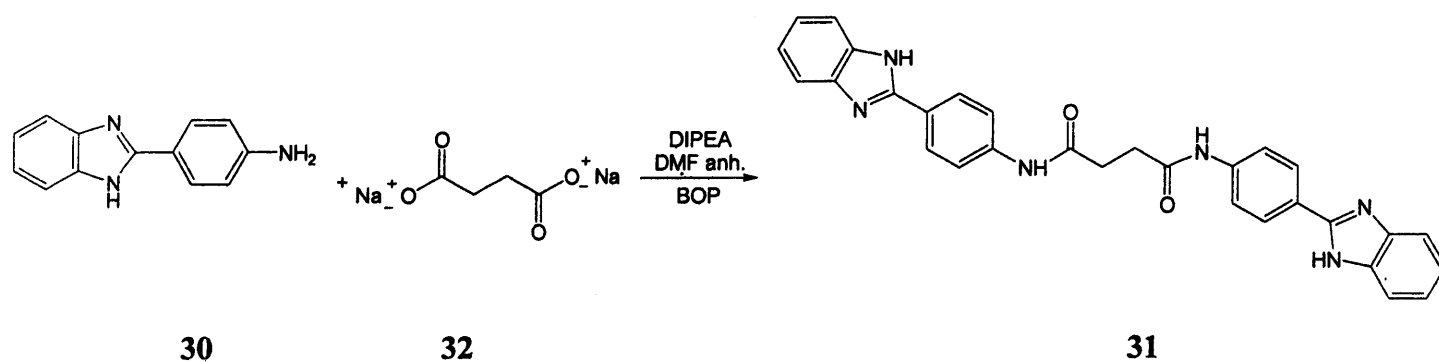
Scheme 3.23 - Synthesis IV of *N,N'*-Bis-(1*H*-benzoimidazol-2-phenyl)-succinamide

4-(1*H*-benzo[d] imidazol-2-yl) benzenamine (**30**) (2 eq) and sodium succinate (**32**) (1 eq) were dissolved in anhydrous dimethyl formammide (DMF) (12.8 mL); addition of DIPEA (2.92 eq) and BOP (1 eq) followed and the reaction mixture was then left stirring at room temperature for 72 hours. Although TLC did not show clear presence of the final product, the reaction mixture was filtrated and washed with water, concentrated *in vacuo* and subjected to column chromatography (hexane/AcOEt 1:4), checking the elution by TLC and short length waves as all the spots were visible only in this way.

The purification did not isolate any final product but eluted only impure starting material, thus explaining that the followed procedure did not allow the reaction to reach the completion. One last attempt was planned trying a distinct synthetic protocol.
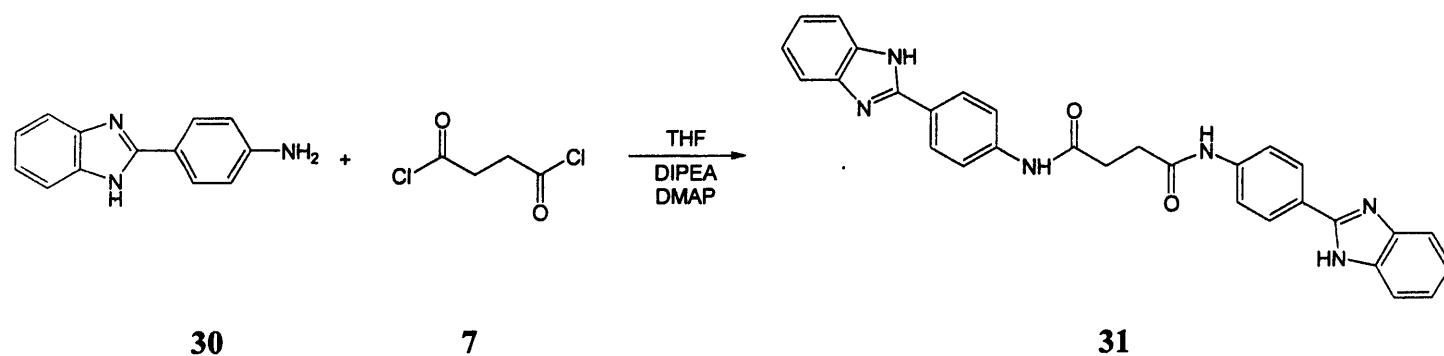
Scheme 3.24 - Synthesis V of *N,N'*-Bis-(1*H*-benzoimidazol-2-phenyl)-succinamide
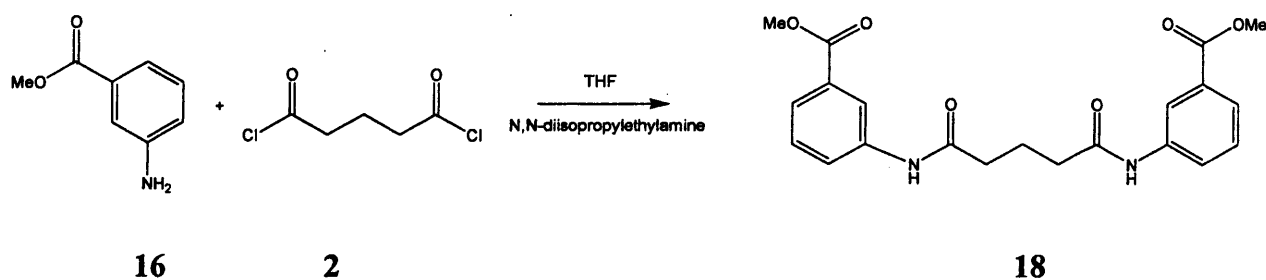
Compound **31** was prepared by reacting **30** and **7** directly in a ratio of 2:1 dissolved in THF anhydrous (8 mL). Following the addition of DIPEA (2.1 eq), DMAP (10%) was added and the reaction stirred at room temperature. After 72 hours and the observation of a faint new spot on TLC, further DMAP (10%) was added trying to force the equilibrium of the reaction towards completion. After other 24 hours, the reaction mixture concentrated *in vacuo* and precipitated with MeOH; the filtrated MeOH was added of DCM to ease the precipitation of another solid that was loaded on a column chromatography (AcOEt). As mentioned in the previous attempt, checking the elution by TLC was possible only using short length waves as all the spots were visible only in this way. Additionally, after five hours no product was eluted and only the use of MeOH allowed a quick separation of what revealed to be still impure starting materials.

Unfortunately, time restraints did not allowed to carry on the synthesis of this interesting molecule with different approaches and optimized technique of purification.

HCV Replicon Assay

At the time of synthesis an HCV replicon assay was available. The assay is performed on exponentially dividing Huh-5-2 cells and involves measuring cell growth and the formation of single stranded RNA.[150] Initially, baseline rates are determined for both in the absence of any compounds. The process can then be repeated with potential inhibitors and the results evaluated relative to the initial values.

The results of compounds with sufficient yield and purity to allow testing are shown in the following tables.

Table 3.4 - HCV helicase assay data
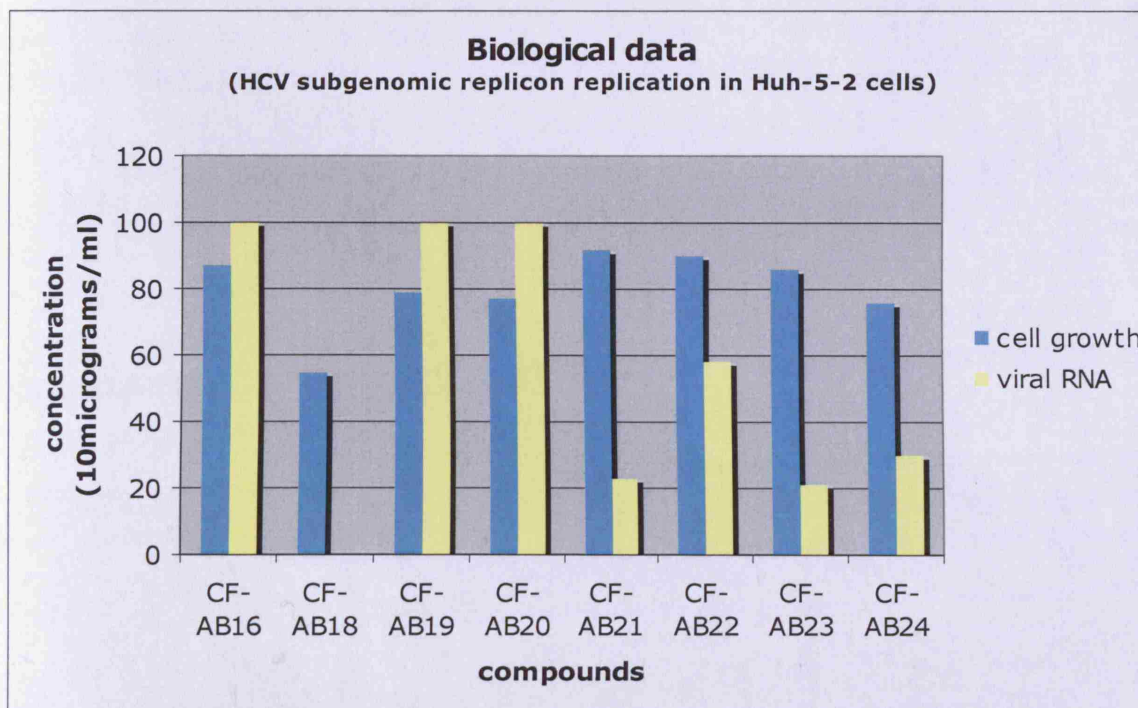
The collected data indicate that viral RNA formation relative to cell growth decrease with increasing concentrations of all compounds expect for compound **3**. Hence this compound is inactive. Interestingly, the synthesised carboxylic acid form of the compounds **14**, **20** and **22** showed important values in terms of percentage of inhibition in the HCV Helicase Assay.

EC50
(HCV subgenomic replicon replication in Huh-5-2 cells)
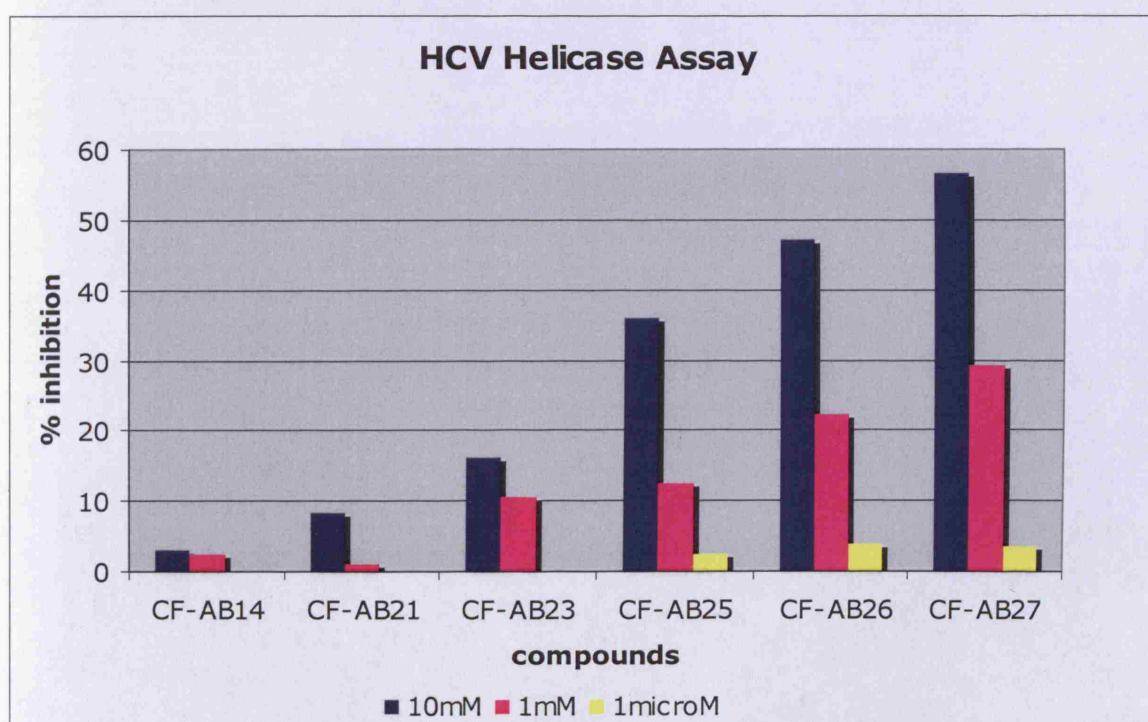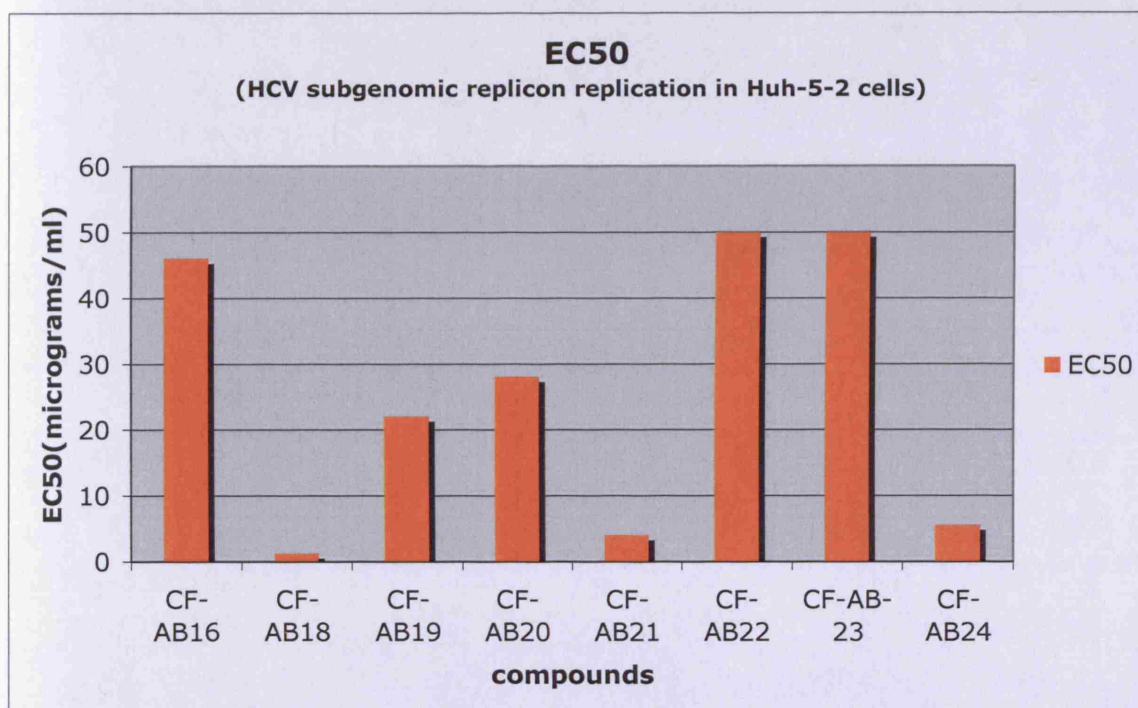
HCV Helicase Assay

Relatively to the last HCV helicase assay diagram, the assay (per well) was:[147]

(total mix volume is 90 µl per well)

11 nM of HCV

25 nM of MOPS

5 mM of ATP

2 mM of DTT

3 mM of MnCl2

100 mg/ml of BSA

2,5 ng of DIG+biotin labelled dsDNA

The reaction of the unwinding of the double-labelled dsDNA in the presence of the HCV helicase was allowed to proceed for exactly 60 minutes.

(All the above data have been kindly provided by Dr. Dimitrios Vlaclakis during our PhD collaboration).

Table 3.5, summarising compound structures, their corresponding data about cell growth, viral RNA, HCV helicase Assay and yields (%), follow below.
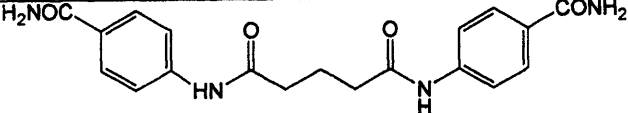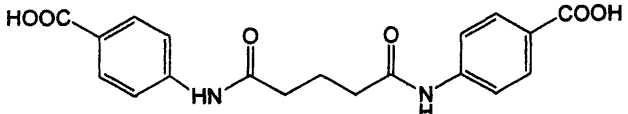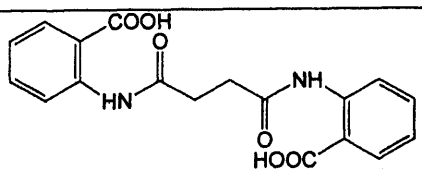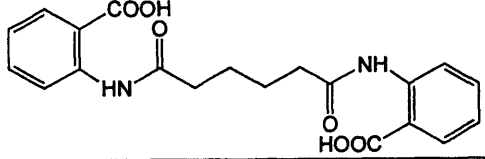
Table 3.5 - Compound structures, and corresponding cell growth, viral RNA, HCV helicase Assay, and yields (%) values (*continues*).

Structure	Name	Cell Growth[a]	Viral RNA[a]	HCV Helicase Assay[b]	Yield %
	CF-AB 14 (14)	82	29	3.2	42.03
	CF-AB 16 (13)	97	56	-	44.65
	CF-AB 18 (15)	78	30	-	24.56
	CF-AB 19 (17)	64	67	-	26.98
	CF-AB 20 (18)	77	100	-	4.27
	CF-AB 21 (20)	92	23	8.4	27.28
	CF-AB 22 (21)	90	58	-	47.27
	CF-AB 23 (22)	76	30	16.2	4.84
	CF-AB 24 (24)	90	58	-	85.17

a (% of untreated control): Interferon alfa -2b at 10.000 units/well reduced the signal in the viral RNA (luciferase) assay to background levels; without any cytostatic activity
b (% inhibition)

121

Table 3.5 - Compound structures, and corresponding cell growth, viral RNA, HCV helicase Assay, and yields (%) values.

Structure	Name	Cell Growth[a]	Viral RNA[a]	HCV Helicase Assay[b]	Yield %
H_2NOC ... $CONH_2$	CF-AB 17 (3)	-	-	-	39.67
$HOOC$ ... $COOH$	CF-AB 25 (27)	-	-	36.1	98.21 (41.3)
COOH ... HOOC	CF-AB 26 (28)	-	-	47.3	31.13
COOH ... HOOC	CF-AB-27 (29)	-	-	56.7	32.17

a (% of untreated control): Interferon alfa -2b at 10.000 units/well reduced the signal in the viral RNA (luciferase) assay to background levels; without any cytostatic activity

b (% inhibition)

Molecular Modelling - I

Molecular modelling aided to rationalise the biological results obtained and evaluate the potential of compounds within the same class that were not synthesised. MOE software[151] was used to create a database of 50 symmetrical compounds including all the compounds that syntheses were attempted for, the initial bioisosteric replacements suggested, and also an unsaturated and short chain aliphatic linker (Table 3.6).

Table 3.6 - Database of compounds used for docking with three-dimensional HCV helicase and Dengue helicase models.

The compound database was docked with models of helicase enzymes for HCV and DENV. A crystal structure for HCV helicase has been determined and the coordinates of this were downloaded from the RCSB protein databank.[152] At the time of this investigation, an equivalent crystal structure was not available for dengue helicase and

123

a structure predicted using homology modelling[147] was used. As the active site of this model was not described, the two enzymes were superimposed using MOE[151] and the active site from HCV helicase was used to define an equivalent region on the predicted model.

Using the FlexX suite within Sybyl,[153] the whole database was docked with HCV and dengue helicase models. The 30 most stable conformations for each was determined and evaluated.

Several theories exist about the mode of action of these inhibitors; potentially the compounds could provide steric hindrance to RNA attempting to bind, thus preventing the helicase enzyme from unwinding strands. Alternatively, they may prevent the movement necessary for activity; in fact, the helicase enzyme must 'open' and 'close' to separate RNA strands, and compounds binding strongly to the enzyme may be able bind it in a single conformation.[154]

Analysis of the compounds against HCV was based on these two theories.

The model of HCV helicase is crystallised with a single strand of DNA indicating where RNA would bind during unwinding. This allowed identification of both compounds providing steric hindrance and those binding deeply within the active site, through strong intermolecular interactions.

Evaluation of compounds against dengue helicase was based only on the extent of bonding, as this model lacks a strand of nucleic acid within the active site. Analysing the hydrophobic and hydrophilic interactions with a compound previously shown to bind strongly, several key amino acids were identified within the active site. In particular, the central arginine amino acid, considered relevant for activity because of its location, flexible nature and ability to form hydrogen bonds,[145] guided the selection of possible compounds.

The biological data was assessed via the structure of compound 3 that was docked with the HCV helicase model revealing that it binds at a site distant to the bound DNA. If the mode of action is via steric hindrance, then the apparent absence of activity could be explained accordingly.

The scoring function within MOE[151] indicated that, although several hydrogen bonds might form with nearby amino acids, in addition to a variety of hydrophobic

interactions, the compound does not bind deeply within the enzyme, thus suggesting that these interactions may have been lost during testing. It must be taken into account that the surface of the enzyme is exposed to the surrounding environment, including the solvent, able to break the bonds formed with the enzyme.

Assessing the conformations of each compound within the active site on their ability to obstruct the bound DNA, 7 of the 50 compounds evaluated achieved this, although several of these only provided a limited overlap. All the selected structures were characterized by a meta-substituted end group, thus suggesting that the conformations presenting 3-substitued phenyl rings better interact with the amino acids within the DNA binding site. Interestingly, the majority of these compounds contain either an imidazole ring or secondary amino groups, and also include both hydrophobic regions and groups able to form hydrogen bonds. This may infer that there are hydrophobic pockets relevant for binding close to the DNA. Due to the significant obstruction achieved, compound **16-f** showed the greatest potential.

Assessing the compounds on their ability to form strong intermolecular forces, many of these were observed to possess multiple interactions located on the surface of the model and exposed to the environment similarly to compound **3**, and, therefore, unable to inhibit the enzyme significantly if tested.

More in detail, compounds containing the amidine end group were poor at binding within the enzyme, whereas those with a meta-substituted end-group tended to form stronger complexes than para-substituted structures.

This supports the observation that conformations with meta end-groups achieve better alignment with the amino acids required for binding. It was also noted that compounds possessing a fumaryl linker often bound deeply within the active site. Further development of these compounds may be successful in developing an active helicase inhibitor.

Analysis of the database against dengue helicase revealed a variety of compounds theoretically capable to bind strongly with the amino acids in the active site. In contrast to HCV helicase, compound **14-b** (equivalent to compound **3**) appeared to bind very strongly with dengue helicase through six strong hydrogen bonds formed deeply within the active site, suggesting that biological activity may be more successful if tested against dengue.

Compounds, whose end groups were meta-substituted, often showed stronger bonding within the active site compared to the equivalent para-substituted forms, observation common to both HCV and dengue and may be worthwhile considering with further research as well as that compounds with a fumaryl linker tended to bind more strongly, this time with the arginine-107 residue located deep within the centre of the enzyme. Probably the extra rigidity conferred by a double bond allows the compounds to further extend to the region where hydrogen bonding can occur with arginine. Molecular modeling data confirm that future development of this series of compounds against flaviviruses should focus on meta-substituted endgroups and fumaryl linkers to be more successful.

Conclusions and Future Work -I

Various are the conclusions and recommendations for future work that can be drawn from the research presented.

Synthesis and biological testing of compound **3** revealed that this structure is not an inhibitor of HCV helicase. This provided a useful comparison for the symmetrical ester derivative, that was active, as well as for the others. The ester is expected to be hydrolysed to a carboxylic acid during the assay and so compound **3** differs only by an amine in place of a hydroxy group. This would suggest that the negative charge of the carboxylic acid is able to form a strong interaction with HCV helicase. Future work may benefit from incorporating negatively charged end-groups, such as carboxylic acids (compounds **27, 28, 29**) , into potential structures.

Molecular modelling has identified several structural features; in particular, meta substituted end-groups and fumaryl linkers appear to offer favourable conformations for binding strongly with both dengue and HCV helicase. Additionaly, the modelling has indicated that compound **3** forms many strong interactions with dengue helicase. Therefore, it would be appropriate to test this compound against this enzyme in the future if possible. Access to a dengue helicase assay could potentially reveal many active inhibitors from this series of compounds.

Finally, the research has demonstrated the need to optimise synthetic pathways if development is to proceed efficiently. Increasing the proportion of aminobenzamide to the other reagents may encourage amide formation. Alternatively improving the methods of purification to maximise the yield may be successful. This could be attempted by optimising the solvent used during crystallisations. However, if these changes are unsuccessful, better results may be achieved if a reaction with multiple steps is sought.

HCV Helicase Inhibitors - II

Molecular Modelling - II

Compound **32** was designed *in silico* as an anti-HCV helicase molecule to interact with four key amino acid residues within the HCV helicase active site – arginine-203, cysteine-238, aspartic acid-261 and arginine-288 (Fig.3.9) –,[i] but proved to be extremely hard to synthesise[i] and, therefore, used as a lead compound to design novel compounds with equal or improved active site interactions and simpler synthetic routes.
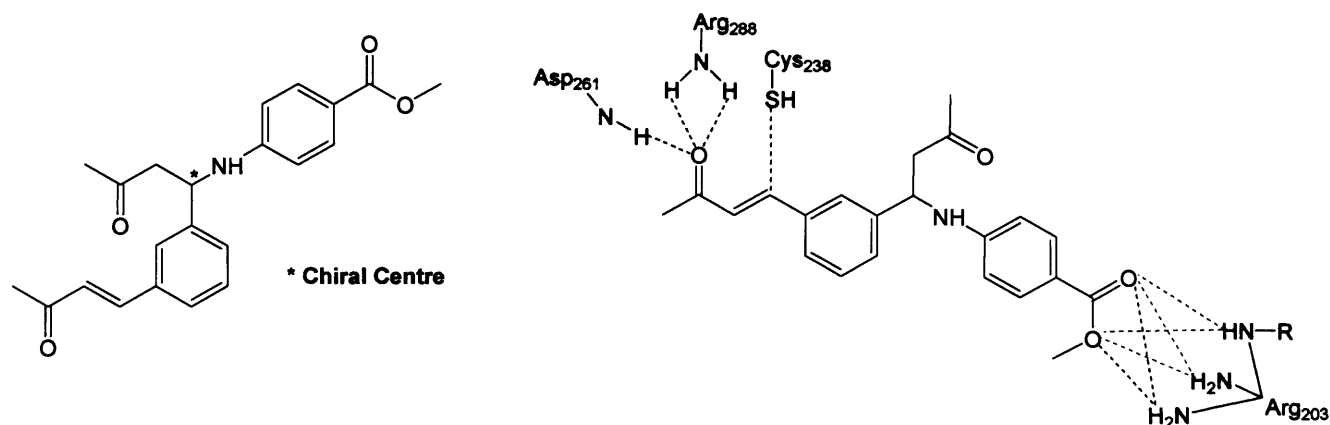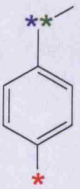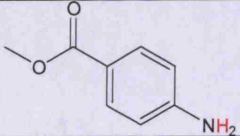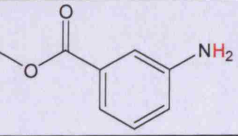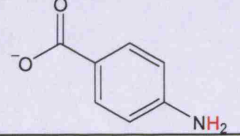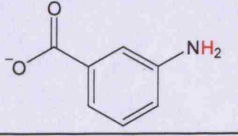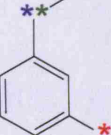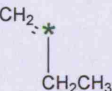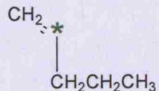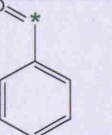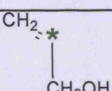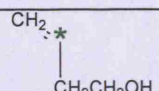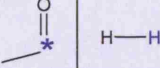
Fig.3.9 Compound 32 and its desired active site interactions with HCV helicase (red dashed lines: proposed hydrogen bonds; blue dashed/dotted line: proposed covalent bond; green structures – active site residues)

As shown in the above figure, the majority of the desired interactions are hydrogen bonds, and a covalent bond may form between the sulfhydryl residue of cysteine-238 and the double bond (present in both of lead compound **32** and the compounds designed below) via Michael addition, provided that the S-H residue and the double bond are parallel to each other and close enough for the reaction to occur.

Two databases of modified compounds were automatically generated in Molecular Operating Environment (MOE)[151] from the scaffolds and fragments shown in Table

3.7 to investigate structure-activity relationships. Attachment points for scaffolds and fragments are designated by different colours on the chemical structures.

Table 3.7 - Database components

All compounds from methyl ester derivates database have been docked into the active site of a model derived from a HCV helicase enzyme with DNA crystallised in the active site (DNA removed for modelling) using the FlexX program within Sybyl 7.0.[153] No configuration shown any of the expected interactions and conformations, probably because the active site in this form of the enzyme was crystallised around the natural substrate – a DNA molecule – which is much larger than the compounds generated and, consequently, the residues which these compounds need to bind to may be too far apart for the molecule to be able to concurrently form interactions (mainly hydrogen bonds) with all of them.

The process was repeated using a model of the enzyme minimised around the lead compound where the active site residues are closer together and, therefore, more

129

interactions with the active site and enable more useful compounds to be found. The results of these docking simulations confirmed both suppositions; several compounds resulted 'correctly' docked within the active site and many of these were in the ionized form. Additional features which appeared to favour the desired active site docking conformation included:

- *para* or *meta* substitution on either ring; two *para* substitutions or a *para* and a *meta* substitution more favourable than two *meta* substitutions

- no side chain (i.e. fragment = H)

- polar group at opposite end of the molecule to the ester or ionised group, which formed additional hydrogen bonds with active site residues.

In order to determine the reliability of the models realized up to this point and to gain a more accurate picture of how the compounds – and the enzyme – may behave *in vivo*, the configuration showing the most favourable interactions from the thirty generated by FlexX for each of the six best compounds (Table 3.8) found as a result of this process were selected and further minimized allowing the protein (previously kept rigid) to move around the compound and find a low-energy configuration (energy gradient 0.001).

Table 3.8 - The six database-generated compounds with the most favourable active site interactions.
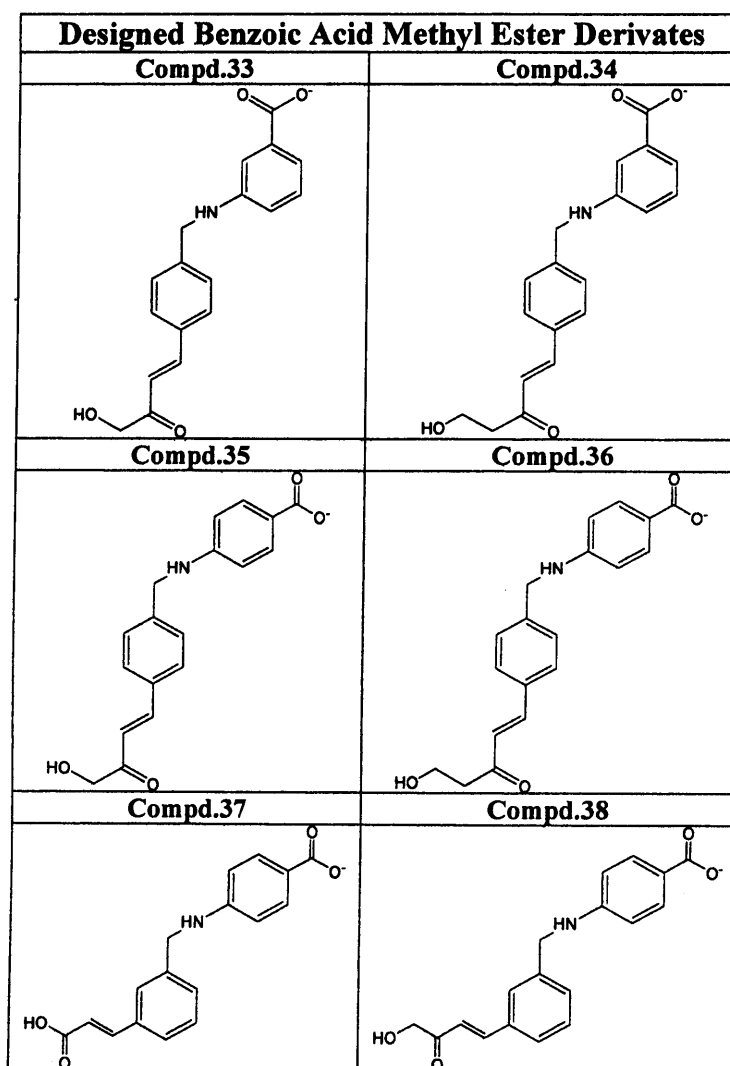
The most promising model (compound **35** showing the highest number of desired interactions) was subsequently added of water molecules to make it a more accurate reflection of the *in vivo* situation, trying to find configurations near to the energy minima considered to be closest to what is actually seen *in vivo*. The model aimed to simulate the molecular dynamics of the enzyme-compound interactions over one hundred picoseconds, resulting in a database of 200 configurations which could be run consecutively as a 'video' sequence. The interesting results obtained showed the compound remaining firmly anchored in the active site without being displaced by water molecules as the distances measured between key molecules during the

simulation also confirmed. The likelihood of hydrogen and covalent bonds forming has been evaluated taking into account only the last 100 database entries as these are more representative of the 'true' minimum energies which are likely to be seen *in vivo*, whereas the first 100 represent a period of stabilisation where the enzyme and compound are approaching their near-minimum energy configurations.

The lowest energy interaction configuration from the molecular dynamics simulation was then taken and subjected to the same MOE energy minimisation as the six best compounds were previously, albeit with a much smaller energy gradient (0.0001), in order to be further optimized.

Chemistry - II

The removal of the chiral centre found in the lead compound **32** greatly simplified the synthetic route and increased the number of methods possible. The reactions were planned to proceed in two steps:

1.   The synthesis of a precursor with the scaffold and methyl amino benzoic acid, ketone and hydrogen fragments (see molecular modelling above) in place but with an aldehyde residue at the point of attachment for the alkene chain fragment

2.   The reaction of this aldehyde residue with an aldehyde or methyl ketone which has the A_3 fragment residue attached to complete the compound.

The reaction scheme and conditions for the first step were adapted from a procedure for direct reductive amination using two starting materials, a catalyst and a reducing agent.[155] This reaction was chosen as it appeared to be the only single-step procedure, therefore avoiding the synthesis of an intermediate imine which many other routes required.[156]
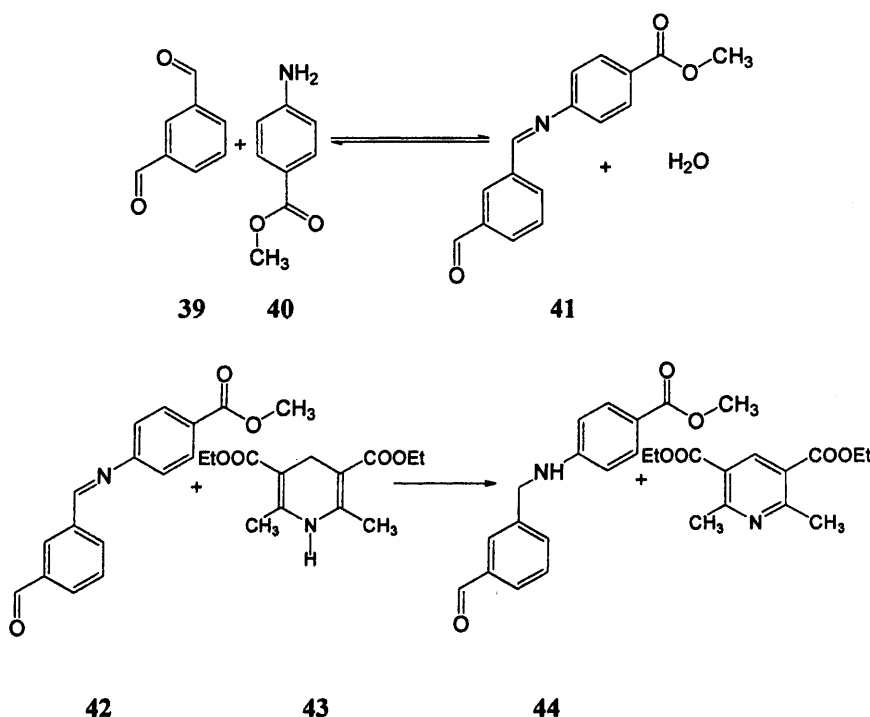
Fig.3.10 - Proposed mechanism for first synthetic step as carried out in the laboratory (NB this mechanism is transferable to the syntheses of the two other intermediates that would be required to produce compounds **33-36**)

133

Once complete, the product of this first reaction could act as a precursor for two final products – one with a CH₂OH following the double bonded oxygen (as in 33) and one with CH₂CH₂OH (as in 34) – in a procedure adapted from the work conducted prior to this project.[i]
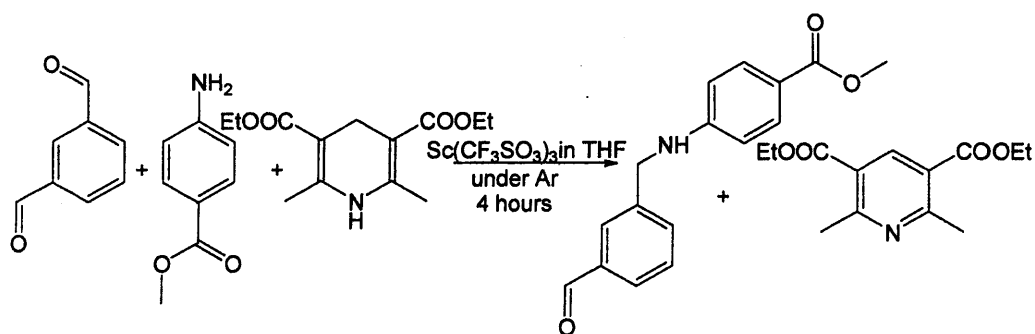
Scheme 3.25 - Reaction scheme for synthesis of methyl 4-[(3-formylbenzyl)amino]benzoate

This reaction was conceived as a test to determine if the procedure worked before synthesising precursors for compounds **35-38**. Therefore, only 100mg of dialdehyde **39** was used, to keep the reaction on a small scale. In an attempt to avoid the formation of an unwanted diimine product   (i.e. where the amine residues of two molecules of **40** react with both aldehyde residues of one molecule of **39**), **40** was dissolved in tetrahydrofuran (THF) and added dropwise over the first hour of the reaction.

Following extraction and drying of the organic layer from the reaction, a pale yellow solid was obtained. As thin layer chromatography (TLC) separation (CHCl₃/DEE 1:1) showed, 3 spots from the yellow solid, flash column chromatography (DCM/hexane 1:1) has been used without achieving any separation of the compounds. Flash chromatography was repeated (DCM/hexane 1:4) trying to slow the elution of the compounds and achieving the separation of a white solid (the first top spot seen on the TLC), the ¹H NMR spectrum of which clearly showed that it consisted of the starting dialdehyde **39**. A second ¹H NMR performed on the remaining mixture, which contained the second and third spots from the TLC separations, revealed a probable mixture of the amide/ester **40**, the desired product **43** and the oxidised reducing agent **44**.

Further purification of the mixture was therefore attempted. A smaller flash chromatography column (DCM/hexane 1:1) allowed separation of some of the second spot, much of this again co-eluted with the third spot. ^{1}H-NMR performed on the pure sample of the second spot showed presence of **44** contained in this sample but not that of the desired intermediate. Another flash chromatography column was attempted for the purification of the remaining mixture using different eluents and/or ratio (DCM/hexane in a 1:4, 2:3 and then 1:1 ratio; pure DCM) without elution of any compounds. To quickly elute and recover the entire mixture, ethyl acetate has been used and a column chromatography has been performed. Three spots were isolated from the mixture; ^{1}H-NMR and ^{13}C-NMR of the residue from fractions containing the second spot clearly showed the presence of the desired   intermediate **43**. However, as the amount recovered was not sufficient and somewhat impure for use in the next stage of the reaction, the reaction was repeated.

It is unclear why this reaction produced such a low yield as it bore great similarity to that carried out in the reference paper.[155] The only differences were in the reactants – a *meta* rather than *para* substituted dialdehyde and a methyl ester rather than methyl ether substitution on the amine derivative. The closer orientation of the aldehyde groups may make the carbon and oxygen of the non-reacting aldehyde group cause some steric hindrance when reacting with the amine, leading to a slower rate of reaction with a lower yield. Additionally, the presence of an ester para to the amine may lead to the deactivation of it.
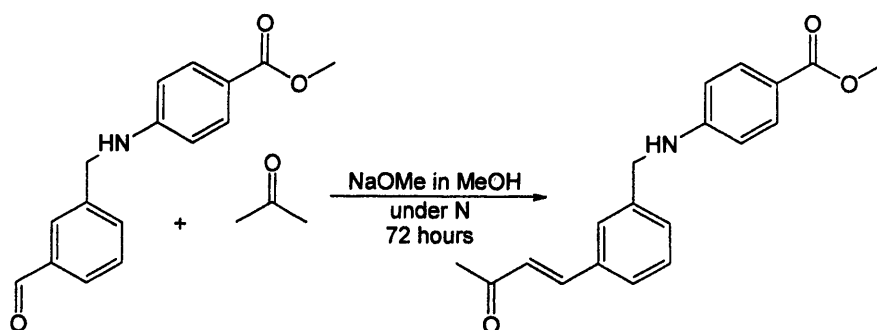
Scheme 3.26 - Preliminary reaction to test the feasibility of the second synthetic step

Although the insufficient product obtained from reaction one to carry out any quantitative step 2 reactions to form a final product, it was thought that enough was present for a qualitative test reaction to assess the feasibility of the next step. However, TLC analysis after 24, 48 and 72 hours of reacting showed nothing more than the starting materials and it was concluded that there simply was not enough of **43** for a successful reaction.
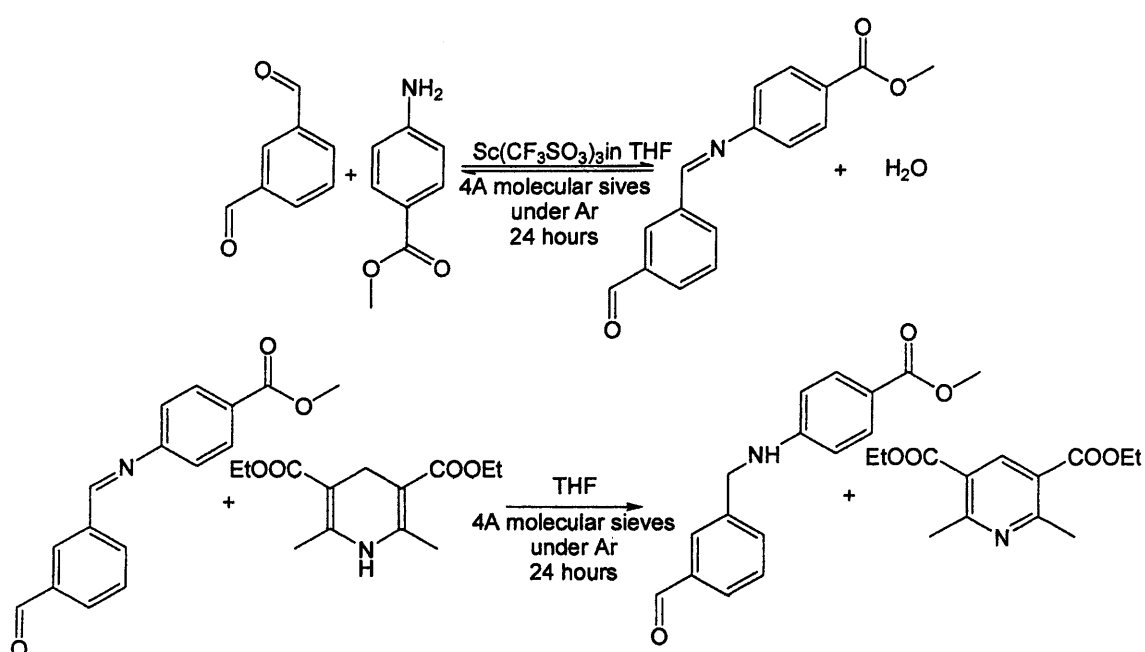
Scheme 3.27 - Two-step adaptation of the direct synthesis

To maximise yield and minimise impurities, several changes to the synthetic procedure were implemented when the reaction was repeated. A two-step process was introduced aiming to form the imine from the initial reactants before reducing this to the desired compound; in theory, this should both increase yield and reduce the formation of unwanted products. Molecular sieves were used to push the equilibrium of the first reaction to the right by removing water from the reaction environment, and remained in the scheme for the second step as the Hantzsch dihydropyridine **42** was added directly to the reaction vessel. Doubling the amount of the reactants and increasing the reaction time have been attempted to increase the amount and

proportion of precursor synthesised. Once the molecular sieves were filtered out and the organic phase extracted and dried as before, TLC analysis revealed five spots and column chromatography (CHCl₃/DEE 1:1) was carried out resulting only in minimal separation with a very unusual feature – a spot that suddenly appears *in between* two spots already eluting. The breakdown of remaining imine on the column during elution may be the explanation for this spot, as this compound is very unstable in presence of even a small amount of water. The fluorescent oxidised reducing agent from the other compounds has been the only purification afforded by this column.

Repeating column chromatography (CHCl₃/DEE 3:2) and slowing the elution of the compounds to enable more useful separation, afforded both the unusual appearance of the spot between two others, and a fraction containing the final two spots. This fraction was isolated and evaporated and the resultant solid/oil mixture crystallised with petroleum ether overnight hoping to separate them and isolate the product. ^{1}H-NMR of the crystals clearly showed that they contained the starting amine **40**, thus confirming the theory of the suddenly appearing spot – which this presumably was – being the result of the breakdown of the imine to starting materials. Additionally, the ^{1}H-NMR of the crystallisation residue contained in the petroleum ether revealed an unknown compound that was definitely not the desired product **41**.

Following the failure of the separation procedures conducted in Reaction Three, and supposing that a larger amount of product would aid in purifying the desired intermediate **43**, the procedure followed for Reaction Three has been repeated with three times the amount of each reactant, with the addition of a test TLC of the first reaction after 24 hours which showed a faint amine spot, indicating that most of the starting material had reacted at this stage. At the end of the reaction, the solution was sieved, extracted and evaporated as before. Although a larger (500mL) column was used for the column chromatography (CHCl₃/DEE 1:1) aiming to improve the separation of the compounds, the elution showed a pattern very similar to that previously experienced and afforded several fractions containing all spots except for the fluorescent spot (the reducing agent **42**). A second 500mL column was then performed using a new eluent (AcOEt/hexane 2:3) that, although producing poor separation of the second and third spots, allowed the purification of the first top spot, the ^{1}H-NMR and ^{13}C-NMR of which clearly showed to be the oxidised reducing agent **44**.

the ^{1}H-NMR and ^{13}C-NMR of which clearly showed to be the oxidised reducing agent **44**.

Aiming to maximise yield and minimise impurities, additional changes to the synthetic procedure were implemented when the reaction was repeated. The *two-step* approach was attempted other three times using different conditions (see Table 3.9 below) in order to form the imine from the initial reactants before reducing this to the desired compound as this should both increase yield and reduce the formation of unwanted products.

Table 3.9 – Experimental conditions for a two step synthesis of compound **43**.

n. STEPS	Benzoate	Aldehyde	Base/Solvent	Catalyst/ Red.Agent	Condition
One pot reaction	[1eq.]	[1eq.]	THF anh.	* Sc(OTf)$_3$ [2mol%]	Under Argon r. t. o. n.
I	[1eq.]	[1eq.]	DIPEA[1eq.]/ THF anh.		Reflux (12 hs)
II			THF anh./ MeOH anh.	HCl 1M in ether/ NaCNBH$_3$ [1.099eq.]	r. t. 10'
I * (65%yield)	[1eq.]	[1eq.]	EtOH absolute		Reflux (10 hs) r. t. o. n.
I * (45%yield)	[1eq.]	[2eq.]	THF anh./ MgSO$_4$ anh.		r. t. o. n. Reflux (5 hs)

DIPEA
THF anh.
Reflux
12 hours

NH₂

+ H₂O

THFanh./MeOH anh.
HCl 1M in ether
NaCNBH₃
r.t. o.n.

EtOOC — COOEt

EtOOC — COOEt
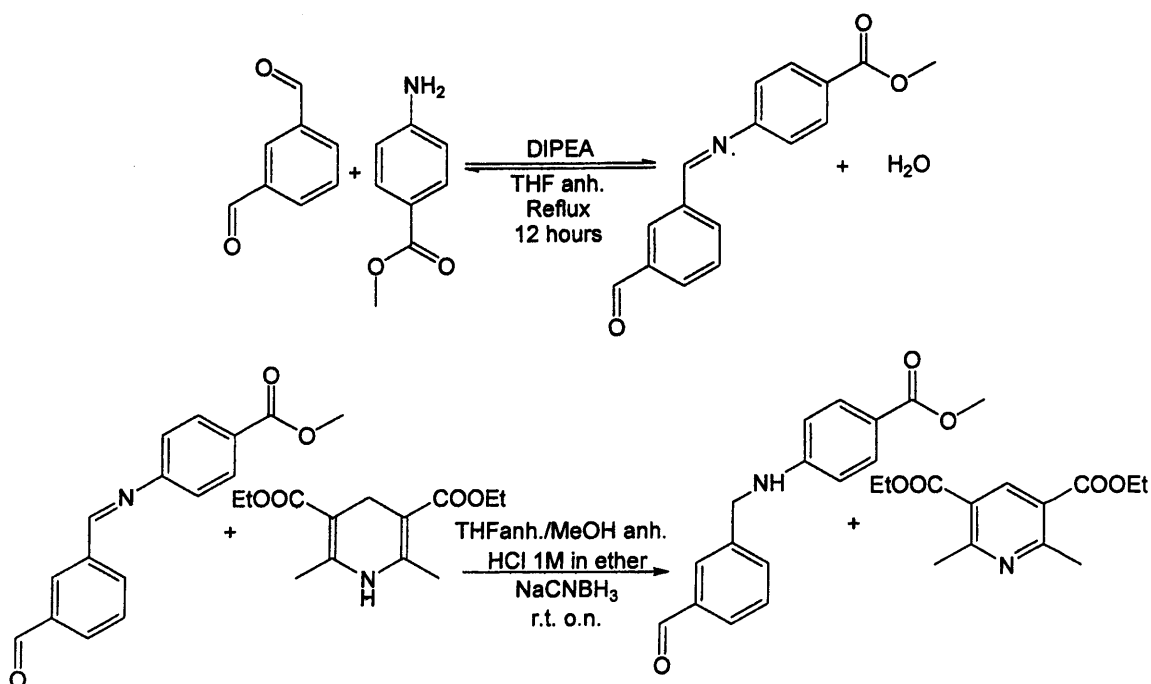
Scheme 3.28 - Two-step adaptation of the direct synthesis

Once dissolved in THF anhydrous, the initial reactants were added one equivalent of *N,N'*-diisopropylethylamine (DIPEA) leaving the reaction stirring for 12 hours; the reaction mixture was then concentrated *in vacuo,* dissolved in THF anhydrous/MeOH anhydrous (1:1), and, while stirring at room temperature, added HCl 1M in ether (1.21 mL) and NaCNBH₃ (2 eq) and left stirring overnight. Following extraction (DEE/water) and concentration *in vacuo* of the ether layer, the ¹H-NMR clearly showed the undesired reduction of the aldehyde to alcohol, thus suggesting to repeat the reaction slowly and carefully adding NaCNBH₃ in order to selectively obtain the imine reduction and carry on with the second step according to the Hantzsch procedure.

The reaction was then repeated, paying attention to the addition of the reducing agent; once extracted and concentrated *in vacuo,* the reaction mixture was subjected to column chromatography (CHCl3/DEE 10%; DEE) isolating the first spot, the ¹H-NMR of which clearly showed to be the dymerized compound of the imine (Fig.3.11).
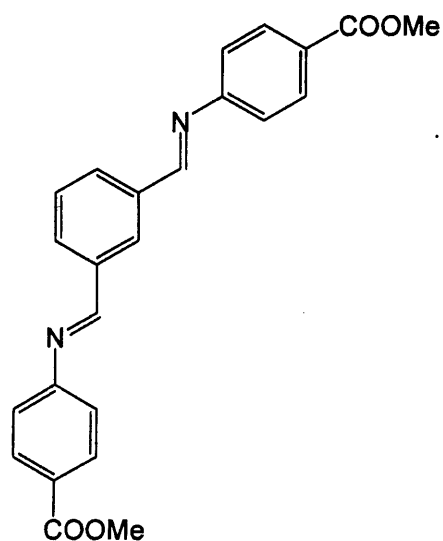
Fig. 3.11 - Dimerised compound of the expected imine

Given the failure of the above attempts, two distinct approaches (Reaction Scheme 3.28, I and II, respectively) were simultaneously tried for the synthesis of the imine, finally providing the desired intermediate subsequently subjected to the second step of the Hantzsch reduction.
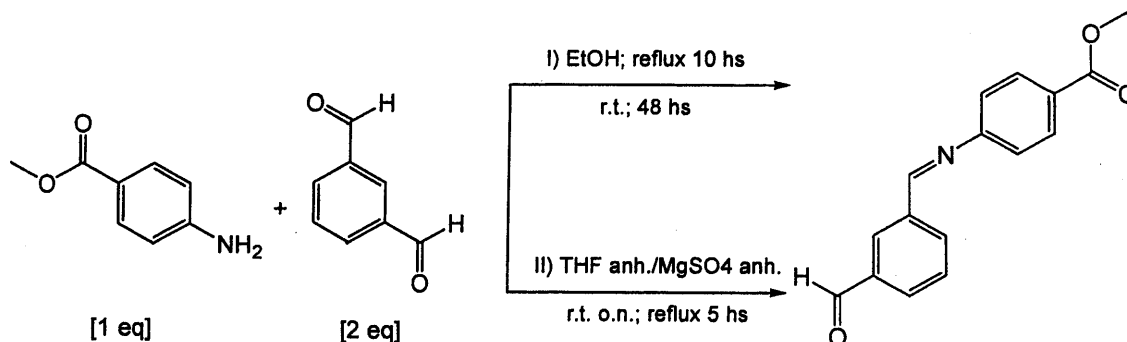
Scheme 3.29 - Two-step adaptation of the direct synthesis

More in detail, in the former case (Scheme 3.28-I)), once doubled the equivalent of the aldehyde and dissolved the reactants in absolute ethanol (EtOH) (20 mL), the reaction mixture was left stirring at room temperature for 48 hours after being refluxed for 10 hours.[157] The subsequent flash chromatography (hexane/DCM 9:1) allowed the isolation of the imine and its immediate use for the Hantzsch reaction

according to the previously mentioned conditions. Following extraction, drying of the organic layer from the reaction, flash column chromatography (hexane/CHCl₃ 9:1; hexane/CHCl₃ 1:2; AcOEt – conditions tested by TLC prior chromatography) and column chromatography (CHCl₃/hexane/AcOEt 1:1.5:0.5), the expected final amine was isolated and further purified by adding water and precipitating a pale yellow solid. As the ¹H-NMR of the precipitated showed the amine to be quite pure but also the presence of few impurities (probably trace of methyl 4-amino benzoate), further column chromatography (CHCl₃/MeOH 10%) and flash chromatography (hexane/MeOH/AcOEt) were carried out. Unfortunately, the purification with the latter technique led to the lost of the final compound as its amount was too low to be further columned. However, as the final compound was finally identified, it was possible to repeat the Hantzsch reaction using the imine synthesised through the latter approach of the *two step* adaptation of Reaction One (Scheme 3.28-II).

More specifically, once doubled the equivalent of the aldehyde and dissolved the reactants in THF anhydrous (20 mL), the reaction mixture was left stirring at room temperature overnight in a round flask containing MgSO₄ anhydrous (2.5g). Once refluxed for 5 hours, the subsequent flash chromatography (hexane/DCM 9:1) allowed the isolation of the imine and its immediate use for the Hantzsch reaction according to the previously mentioned conditions. Following extraction, drying of the organic layer from the reaction, column chromatography (CHCl₃/MeOH 5% - conditions tested by TLC prior chromatography), the expected final amine was isolated and further purified by a further column chromatography (CHCl₃/MeOH 10%). Although the ¹H-NMR of the purified amine showed still the weak presence of few impurities (probably traces of the dymerized compound), the compound was considered pure enough to set up the reaction previously only attempted as a qualitative test reaction to assess the feasibility of the synthesis of methyl 4-({3-[(1E)-3-oxobut-1-en-1-yl]benzyl}amino)benzoate (Fig. 3.12, Scheme 3.26).
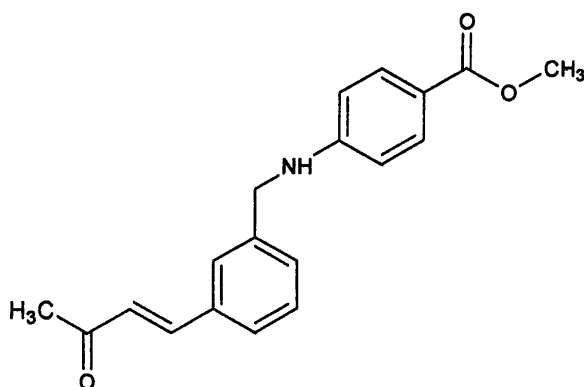
Fig.3.12 - Methyl 4-({3-[(1E)-3-oxobut-1-en-1-yl]benzyl}amino)benzoate

One equivalent of methyl 4-[(3-formylbenzyl)amino]benzoate was added to a round bottomed flask with 10 mL HPLC acetone and sodium methoxide (1.2 eq) and left stirring at room temperature under argon. Test TLC (CHCl₃/MeOH 10%) at 24, 48 and 72 hours after reaction began showed no reaction had taken place, although the addition of further 2.4 equivalents of sodium methoxide and the attempt of heating the reaction up. Unfortunately, time constraints prohibited any further efforts at further optimization of the synthetic pathway and relative purification techniques.

Conclusions & Future Work - IIa

The molecular modelling study of the present project was highly successful throughout. It demonstrated that the lead compound could be improved and new interactions with the active site generated; in addition, the most promising compound designed (35) showed very encouraging results in the molecular dynamics simulation.

Synthesising the compounds was very challenging. It was hypothesised that the removal of the chiral centre would simplify the synthetic process, as this had previously caused a lot of problems. Unfortunately, the procedure chosen for adaptation – which appeared to be extremely simple and showed high yields in the original paper – did not result in the same successes when applied to the compounds here. Therefore, no final compounds could be synthesised.

The compounds designed during the course of this project show great promise and deserve further attention in the future. The main target of any follow-up would therefore be the development and optimisation of a full synthesis. This may take the form of attempting the same basic synthesis for a different precursor molecule – for example for compounds 35 and 36 – which may be more successful if it is steric hindrance from the *meta* substituted starting material which is limiting yields. This could be combined with a more efficient purification procedure. Another possibility is using a different synthetic process, focusing on a two-step procedure (with a separate high-yield imine formation step followed by reduction of the imine bond). Once synthesised, the compounds would obviously need to undergo biological testing to assess whether the positive results from the molecular modelling process translate into success in reality.

During the present project, it was proposed to repeat the synthesis of compound **47** by direct Mannich Reaction using the same experimental conditions previously applied[i] apart from the solvent DMSO, the hygroscopicity of which was thought to be responsible of the previous failures.

Initially it was postulated that **47** could be synthesised via a direct, amine-catalysed, asymmetric, three-component Mannich reaction as shown by List,[158] and Notz et al.[159]
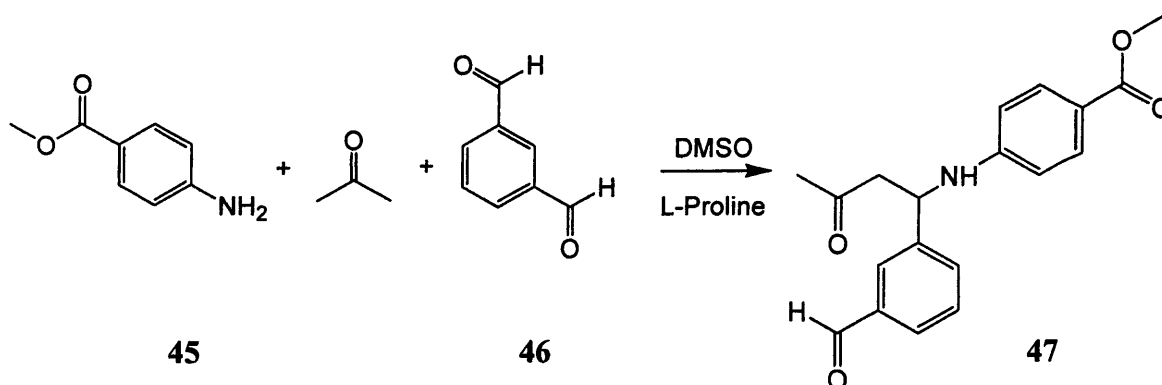
Scheme 3.30 - Mannich Reaction Producing 4-[1-(3-Formyl-Phenyl)-3-Oxo-Butylamino]-Benzoic Acid Methyl Ester (**47**)

In this reaction, L-proline acts as both base and catalyst, as it loses an acidic proton to form an intermediate with acetone, but hydrolysis returns it to its original state by the end of the reaction. The chiral nature of the catalyst also serves to "fix" the chiral centre of the product, essentially providing an enantioselective reaction, although a small amount of the alternative enantiomer may form.

The reaction is generally regarded as a direct synthetic pathway, although the proposed mechanism suggests that it occurs in two stages; the formation of an imine intermediate and the subsequent reaction of that intermediate with acetone and L-proline.

Initially, the lone pair from the amine attacks the partially nucleophillic carbon atom of the aldehyde moiety, forming a new carbon-nitrogen bond. Water is then eliminated, producing the imine intermediate. Subsequently, a lone pair from the oxygen of acetone forms a bond with the hydroxyl group of L-proline.

Subsequent geometric rearrangement of the structure results in a more stable tertiary amine conformation.

The final stage of the reaction begins with concurrent interactions between the imine intermediate and the catalyst. The product should be in a fixed chiral configuration, and as L-proline is being used, **47** should be the major enantiomer formed.
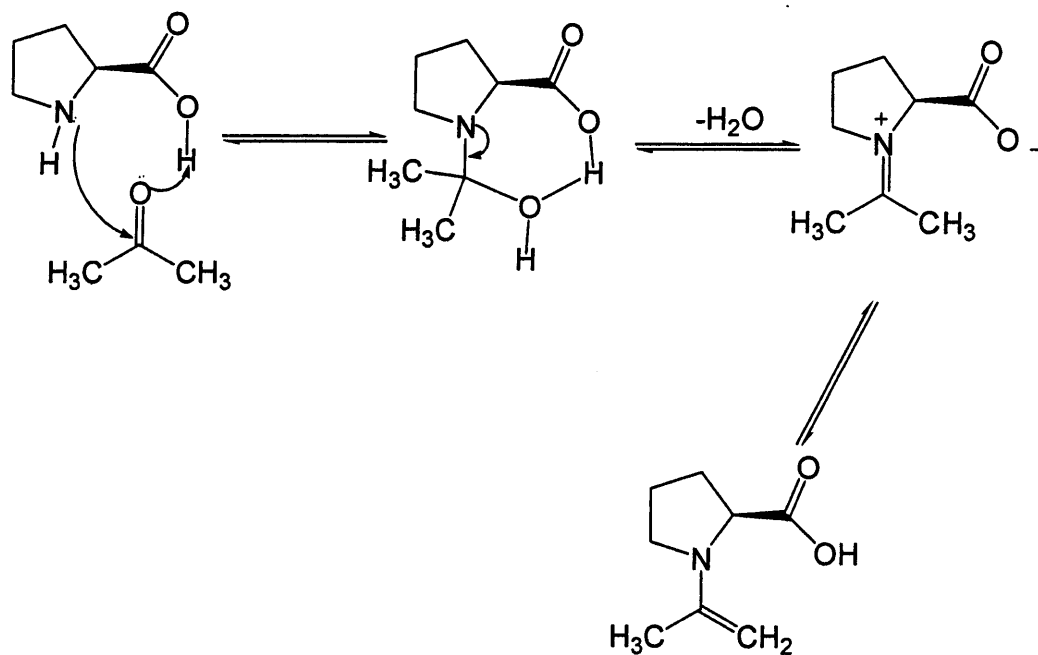
Scheme 3.31 - Mechanism of Reaction between L-Proline & Acetone

The imine bond is broken by the acceptance of an acidic proton from L-proline, which enables a new carbon-carbon bond to form with the L-proline derivative via a nucleophilic addition mechanism. Addition of water across the imine bond and subsequent rearrangement of the structure results in the formation of the final product and the restoration of L-proline to its original state (see Scheme 3.32).
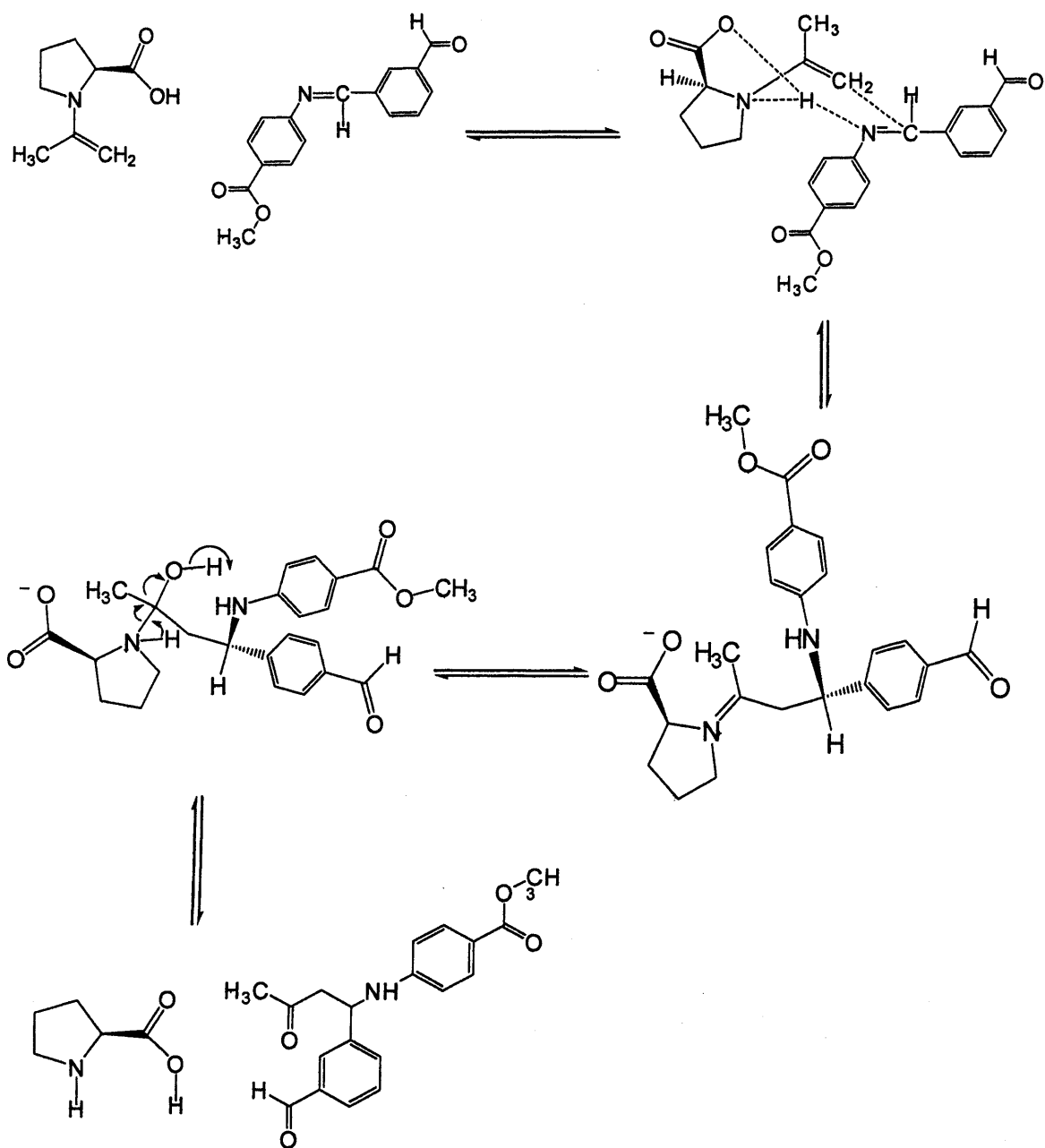
Scheme 3.32 - Mechanism of Aldol Condensation Reaction

146
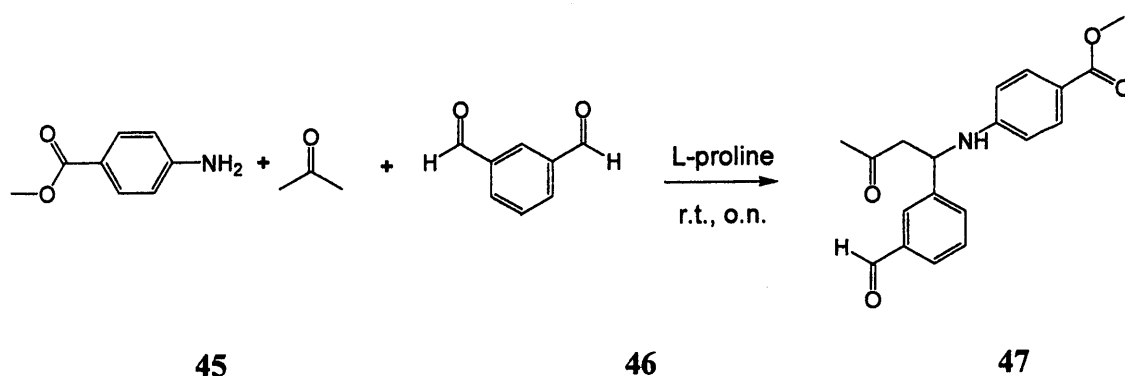
Scheme 3.33 - Direct Mannich Reaction to Synthesise **47**

In order to exclude any presence of water, methyl-4-aminobenzoate, isophthaldehyde and L-proline were dried overnight before setting the reaction. Methyl-4-aminobenzoate (2g, 13.23 mmol, 1 eq) and isophthaldehyde (1.95 g, 14.54 mmol, 1.09 eq) were added to a stirred solution of acetone (30 mL), followed by L-proline (304 mg, 2.64 mmol). The mixture was left stirring at room temperature for 3 days as no clear spot of the desired product was observed on TLC (silica/CHCl₃-MeOH 10%) apart from those of the starting materials. After further 28 hours, the reaction was stopped as it was observed on TLC the appearance of four spots in place of the one expected for the final compound, thus suggesting initial decomposition of the desired molecule. The solvent was evaporated in vacuo, the compound was recollected (4.16 g) and then purified using silica column chromatography (CHCl₃-MeOH 10%). All four spots were isolated and identified by MNR spectra thus elucidating their structures: the main spot corresponded to isophthaldehyde, while the others were impurities and possible molecules derived from the degradation of the product.

Because of lack of time, no further attempts could be tried to optimize the chosen synthetic pathway proved to be less than optimal for the synthesis of the designed potential inhibitor of the HCV helicase enzyme.

Conclusions and Future Work - IIb

This research work aimed to synthesise a potential inhibitor of the HCV helicase enzyme previously designed using molecular modelling software. Although a compound was designed and synthesised, the chosen synthetic pathway proved to be less than optimal.

Yields of the final compound was poor, probably due to occurrence of side reactions, inefficient purification techniques and probable loss of product during purification and experimental or operator error. The time required for completion of the reaction was lengthy, requiring at least 72 hours to enable the isolation of even a small amount of product. This would not be cost-effective on a larger scale, and therefore, alternative methods of synthesis should be sought; alternative starting materials or catalysts may prevent side reactions from occurring or require shorter reaction times. More efficient methods of purification should also be sought, as it is probable that a proportion of the compound was lost in the columns.

The biological activity of the designed compound has not been established, and consequently, it is not known whether the compound has the potential to inhibit the enzyme.

In conclusion, development of a more effective, more convenient treatment strategy for hepatitis C with an improved side effect profile is crucial. This project has attempted the optimization of the synthesis of a novel potential compound for use in hepatitis C therapy. However, further optimisation of its synthetic pathway and evaluation of its biological activity are necessary.

Chapter 4

Introduction

4.1 Apoptosis and its basic functions

Fig. 4.1 - Signaling pathways for programmed cell death (chroma.med.miami.edu/grad/biomed_feature.html).

Apoptosis is a natural process by which a cell is guided to programmed death. Its name, coined following investigations of the nematode Caenorhabditis elegans, is of Greek origin and describes leaves falling down plants. Condensation of the chromatin, cell shrinkage and fragmentation of its nucleus, degradation of DNA and disassembly into membrane-enclosed apoptotic vesicles are specific changes in cell morphology that characterize the course of apoptosis, clearly distinguishing this from the necrotic cell death, another form of cell death.[1]

Fig. 4.2 – apopotic cell (faculty.uca.edu/~benw/ biol1400/notes16.htm).

Based on a genetic program, apoptosis is essential for the development and function of an organism as it is able to eliminate specifically undesired or superfluous cells. Different are the conditions under which the programmed cell death is activated:

- **Tissue homeostasis**

Apoptosis has an essential role in homeostasis of tissues as in an organ or tissue the cell number must be kept constant within very narrow limits suitable for the proper development of the organism. Therefore, an increase in cell division is compensated for by the elimination of no longer functional, old cells and mainly of all cells defecting in signaling pathways regulating cell proliferation. Cells with an oncogenic potential derived from aberrant activation of proto-oncogenes can also be eliminated by programmed cell death, PCD. Pathological increase or decrease in the number of cells is the inevitable consequence of a defected apoptosis. Cancer and autoimmune diseases are associated with an increased rate of cell survival, while an increased apoptosis leads to AIDS and neurodegenerative diseases.[2]

- **Development and differentiation**

Apoptosis has a major role in development and differentiation events, particularly in the embryo, where it switches off all cells no longer required for embryonal morphogenesis and synaptogenesis.

- **Immune system**

During several events of their development, homeostasis and activation, T- and B-cells undergo apoptosis.[3] The elimination of virus-infected cells by cytotoxic T-lymphocytes as well as the elimination of autoreactive B- or T-lymphocytes, the physiological selection and elimination of cells in the thymus and bone marrow are the most relevant examples.

- **Cell damage**

If cell damage or stress occur, PCD can be activated to destroy damaged cells. In fact, before they can accumulate mutations and eventually degenerate into a tumor cell, cells containing damaged DNA are eliminated by apoptosis.

4.1.2 Apoptosis: overview

The apoptotic pathways and components in mammalian cells have been revealed and identified by means of genetic studies on apoptosis in *C. elegans*. A family of proteases known as caspases is involved in the initiation and execution of PCD and can be stimulated and made active by many stimuli via two main pathways, one involving mitochondria, the other using transmembrane receptors of the tumor necrosis factor α (TNFα) class. The former is an intrinsic pathway where stress signals, defects in signaling pathways and DNA damage signals are processed; the latter is an extrinsic pathway using external signaling proteins to activate apoptosis and is especially involved in developmental processes and in the immune system.

Following the caspase activation, cellular key enzymes and structural proteins are degraded, causing apoptosis. DNA damage, stress conditions, malfunction of specific pathways for cell proliferation are some of the main stimuli able to induce PCD. In a normal and health tissue, a perfect balance exists between pro-apoptotic signals activating apoptosis and anti-apoptotic signals suppressing the apoptotic program and promoting cell survival. Lack of survival signals or excess of pro-apoptotic signals can alter this equilibrium in favor of apoptosis.

Furthermore, cell survival can be increased in presence of defecting PCD or a surplus of proliferation signals. Most of the part of the cell death program exists in a latent, inactive form requiring only an apoptotic stimulus to make the program active and to start the apoptosis, thus occurring without activation of transcription.

Although this, forms of apoptosis depending on transcription have been also observed.

4.2 Bcl-2 proteins family

Apoptosis (programmed cell death) is often impaired in human cancer and is relevant in chemotherapy-induced tumour cell-killing, therefore targeting the protein-protein interactions involved in proapoptotic/antiapoptotic complexes may be a more effective way of treating cancer.

Bcl-2 and related cytoplasmic proteins play a key role in apoptosis, that regulate development, tissue homeostasis, and protection against pathogens.[4] Proteins functionally related to Bcl-2 inhibit other cellular members activating proteases (caspases) that degrade the cell; conversely, functionally distinct proteins stimulate apoptosis, probably by displacing the inhibiting proteins from the pro-survival ones.[4] Therefore, the balance between these opposing members determines cell fate. Maintenance of vital organ systems is guaranteed by Bcl-2 family and mutations eventually occurring on any of its members can cause cancer.[4]

Bcl-2 family components play a key role in the regulation and control of the mitochondria-mediated apoptosis. Among all members of this family, Bcl-2 protein was first observed as an oncoprotein coded by a gene subjected to translocations of chromosomes 14 and 18 in B cell lymphomas. Since it does not regulate the cell cycle, like other many oncoproteins function, it cannot be regarded as a classical oncogene, especially following the established homology with the Ced9 protein of *C. elegans*, which has an antiapoptotic role in the organism. Bcl-2 family include more than twenty members with negative or positive effect on the initiation of the programmed cell death. All components have at least one copy of BH motif (BH, BCL-2 homolog), of which four types are known to date (BH1-BH4).

The Bcl-2 family has been classified in three functionally and structurally distinct groups[5]:

- *Group I* – anti-apoptotic proteins harboring BH domains 1-4 and a hydrophobic C-terminal tail[6] used to span the cytosolic surface of various intracellular membranes, such as the outer mitochondrial membrane. All components are thought to bind and sequester the pro-apoptotic Bcl-2 proteins included in groups II and III, thus preventing cell death. Its antiapoptotic effect can explain the oncogenic activity of Bcl-2 protein that has been revealed during its overexpression.

- **Group II** – pro-apototic Bax[7] and Bak[8] proteins, structurally similar to group I except from the lacking N-terminal BH4 domain. They are essential for inducing mitochondria-mediated apoptosis. In particular, the insertion of Bax's $\alpha 9$ helix stimulates apoptosis and induces heterodimers formation and activity of antiapoptotic proteins.[9,10]

- **Group III** – large pro-apoptotic polypeptides (Bid, Bad, Bim[11], Noxa and Puma) containing a single BH3 domain[12] and binding to group I and/or group II members. As sensors of pro- and anti-apoptotic stimuli, they are assumed to play an important role in the life-or-death decision.

Structural studies performed on both anti- and pro-apoptotic members have shown a similar fold despite divergences in amino acid sequence and function (pro-apoptotic versus anti-apoptotic). The 3D structures of Bcl-2 family proteins revealed two central, mainly hydrophobic α-helices surrounded by six or seven amphipathic α-helices of different lengths and a long, unstructured loop between the first two α-helices. Evident is the similarity of the Bcl-2 family structures to the general fold of the pore-forming domains of bacterial toxins. In fact, Bcl-xL, Bcl-2, and Bax revealed to form pores in artificial membranes. A prominent hydrophobic groove on the surface of the anti-apoptotic proteins is the binding site for peptides mimicking the BH3 region of pro-apoptotic proteins (for example, Bak and Bad). Pro-survival members bind by means of an amphipathic α-helix and many hydrophobic contacts with the protein, the analysis of which has revealed the fundamental interactions for hetero-dimerization of Bcl-2 family members and also guided the discovery of small molecule inhibitors of Bcl-xL and Bcl-2 function.[13]

In particular, Bcl2 suppresses apoptosis, retards cell cycle entry at the G1 to S phase transition - a critical control point between cell cycle progression and induction of programmed cell death -,[14] but is also able to modulate the cytotoxicity of some anticancer drugs[15] and is responsible for drug resistance in certain types of cancer.[16]

Along with its unique function as apoptosis inhibitor, studies have shown Bcl-2 acting in neuronal differentiation, that resulted accelerated in Bcl-2 transfected rat pheochromocytoma cells (PC12) as well as it was for the polymerization between NF-L and NF-H (neurofilaments light and heavy).[17]

Furthermore, it has been noticed that in some cancers the programmed cell death is circumvented by overexpression of the antiapoptotic proteins Bcl-2 and/or Bcl-xL.[16,18,19] Therefore, Bcl-2 and Bcl-xL are attractive targets for the development of anticancer agents.[20]

Interestingly, within the flexible loop domain, single-site phosphorylation at serine 70 (S70) is needed to activate Bcl2, whereas multisite phosphorylation at threonine 69 (T69), S70, and S87 makes Bcl-2 inactivate.[21]

Indeed, it has been observed that Bcl-2 phosphorylation is strictly dependent on the presence of an intact loop domain, thus suggesting this region in Bcl-xL and Bcl-2 can halt the anti-apoptotic activity of these genes and may be a target also for regulatory post-translational modifications.[22]

Additionally, it has been reported that Bcl-2-related cytoprotectivity against paclitaxel is likely due to phosphorylation of the flexible loop in conjuncture with other factors.[23] In fact, the binding of paclitaxel to Bcl-2 does not only imply a conformational change but inactivates Bcl-2 and simultaneously phosphorylates residues in the loop domain.[24,25]

4.3 Small molecule Bcl-2 inhibitors

In general, it is a difficult challenge to find small molecule inhibitors of protein-protein interactions. Thus, even though several assays were available for compound screening, the discovery of small molecule inhibitors of Bcl-2 or Bcl-xL remained elusive until the last two years, when multiple, different classes of non-peptidic inhibitors were reported. The first inhibitor of Bcl-2 was reported by Wang et al.[26] in 2000 and was discovered by using a computer screening strategy. This method relies on the high-resolution three-dimensional structure of a targeted receptor protein and computer-aided techniques to search a large number of organic compounds for potential ligand molecules. Combining both the rational design by new computational methods and the diversity of existing compound databases, this strategy has become a powerful tool for nonpeptidic ligand discovery. Using this approach, Wang et al.[26] conducted a virtual screening on a computer of a large collection of more than 190,000 organic molecules and identified HA14-1 (see Fig.4.3).
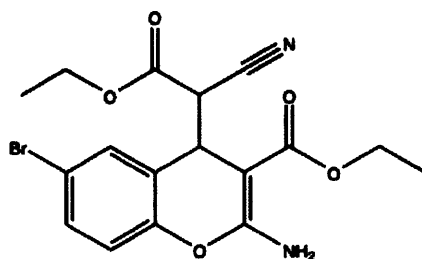
Fig. 4.3 – HA14-1 structure.[26]

In particular, they focused on potential ligands that interact with the site surrounded by a number of important residues, including F104, R107, Y108, D111, F112, A113, Q118, L119, T132, V133, E136, L137, G145, R146, I147, V148, A149, F153, and R207. Of 28 compounds that were finally selected and obtained for actual biological testing, HA14-1 was found to possess the desired biological activity in binding to the Bcl-2 surface pocket and inducing apoptosis of HL-60 cells. It should be noted that HA14-1 has two chiral centers located at the C4 position and the carbon atom attached with the cyano group, respectively. The compound used for biological experiments was a mixture of diastereomers as assessed by HPLC and NMR. The absolute configurations of the biologically active diastereomer need to be determined in further experiments. The data are consistent with a mechanism by which HA14-1 induces the activation of Apaf- and caspases, possibly by binding to Bcl-2 protein and inhibiting its function. The interaction of HA14-1 with the Bcl-2 surface pocket appeared to be specific for the chemical structure of the compound as a series of synthetic analogs derived from HA14-1 containing various modifications were found to have widely different Bcl-2 binding activities (N. Yu, D.L., S.S., and Z.H., unpublished work).

In 2001, three independent papers describing discoveries of small molecule inhibitors of Bcl-2 or Bcl-xL were published. Using the FP competitive binding assay to screen a chemical library of 16,320 compounds, Degtheterey et al.[27] identified two classes of small molecule ligands of Bcl-xL, termed BH3I-1 and BH3I-2 because they inhibit BH3 peptide binding to Bcl-xL. These compounds were also shown to inhibit BH3 peptide binding to Bcl-2. The affinity of these compounds is in the low micromolar range as determined by FP and NMR titration assay. The binding site of the compounds in the surface pocket of Bcl-xL was estimated by NMR experiments based on chemical-shift perturbations exhibited by residues on Bcl-xL upon compound binding. These compounds were shown to induce apoptosis in Jurkat cells overexpressing Bcl-xL. In a

subsequent study, Lugovskoy et al.[28] analyzed the structure-activity relationships of these compounds based on NMR and molecular modelling studies. In addition, they used computer screening to identify an analog of BH3I-1 that showed similar binding affinity to Bcl-xL (see Fig.4.4).
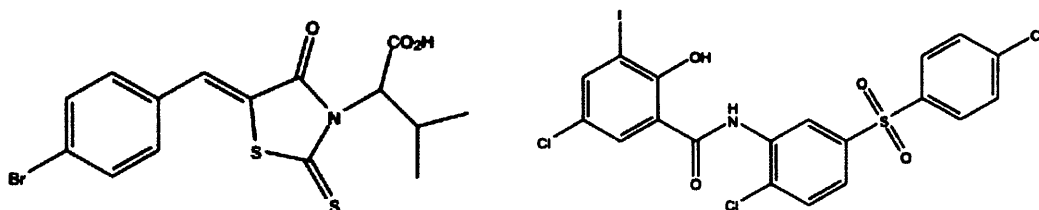
Fig. 4.4 – BH3I-1 (on the left) and BH3I-2 structures.[27,28]

A natural product, antimycin A, which is a known antibiotic that inhibits the mitochondrial electron transfer and binds cytochrome b, was reported by Tzung et al.[29] to have biological activity in mimicking the death-inducing BH3 domain (see Fig.4.5). Computational molecular docking analysis predicted that antimycin A interacts with the Bcl-2 homology domain 3 (BH3) binding the hydrophobic groove of BclxL. They demonstrate that antimycin A and a Bak BH3 peptide bind competitively to recombinant Bcl-2. Antimycin A and BH3 peptide both induce mitochondrial swelling and loss of $\Delta\Psi_m$ on addition to mitochondria expressing Bcl-xL. The 2-methoxy derivative of antimycin A3 is inactive as an inhibitor of cellular respiration but still retains toxicity for Bcl-xL$^+$ cells and mitochondria. Finally, antimycin A inhibits the pore-forming activity of Bcl-xL in synthetic liposomes, demonstrating that a small non-peptide ligand can directly inhibit the function of Bcl-2-related proteins and induces apoptosis in murine hepatocyte cell lines transfected with Bcl-xL. To address the issue of selectivity of antimycin A for the Bcl-2 or Bcl-xL targets, an analog, 2-methoxy-antimycin A3, which was previously known to be inactive as an inhibitor of cytochrome b-c1, was studied and shown to retain binding to Bcl-2. This suggested the feasibility of using antimycin A as a template to develop more specific inhibitors of Bcl-2 or Bcl-xL devoid of general mitochondrial toxicity.
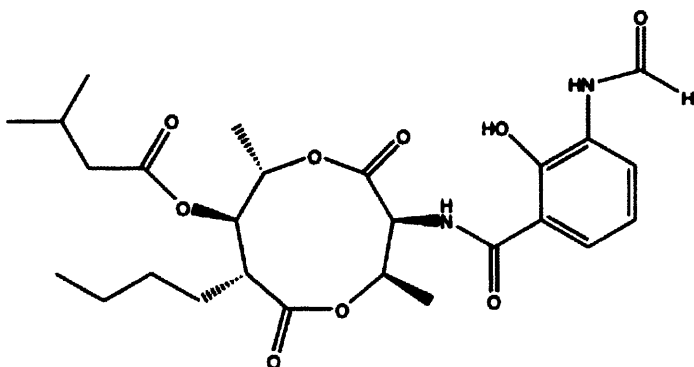
Fig. 4.5 – Antimycin A structure.[29]

Enyedy et al.[30] used the structure-based computer screening approach to search a collection of 206,876 compounds and identified seven compounds that had Bcl-2 binding with IC_{50} values by the FP assay ranging from 1.6 to 14.0 µM. Compound 6 was the most potent compound and had an IC_{50} value of 4 µM in reducing the viability of HL-60 cells expressing a high level of Bcl-2 (see Fig.4.6).

It was further shown that compound 6 was active in inducing apoptosis in HL-60 and MDA-231 cells that express Bcl-2 but had much less or no activity in T47D and MDA453 cells that express low or undetectable levels of Bcl-2. Finally, changes in chemical shifts as determined by NMR experiments showed those residues around the BH3 binding pocket of Bcl-xL to be involved in the binding of compound 6. Furthermore, using NMR methods, it was demonstrated that compound 6 binds to the BH3 binding site in Bcl-xL.

These results showed that small-molecule inhibitors of Bcl-2 such as compound 6 modulate the biological function of Bcl-2, and induce apoptosis in cancer cells with high Bcl-2 expression, while they have little effect on cancer cells with low or undetectable levels of Bcl-2 expression.

Therefore, compound 6 can be used as a valuable pharmacological tool to elucidate the function of Bcl-2 and also serves as a novel lead compound for further design and optimization. These data suggest that the structure-based computer screening strategy employed in the study is effective for identifying novel, structurally diverse, nonpeptide small-molecule inhibitors that target the BH3 binding site of Bcl-2.
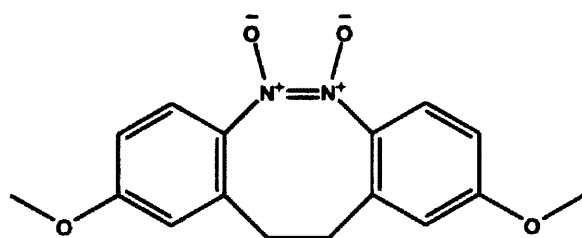
Fig. 4.6 – Compound 6 structure.[30]

The most recent addition to the reported list of nonpeptidic ligands of Bcl-2/Bcl-xL surface pocket was a series of terphenyl derivatives published by Kutzki et al.[31] Previously, this group showed that the terphenyl scaffold, in a staggered conformation, closely reproduces the projection of functionality on the surface of an α helix. Extending this scaffold for the mimicry of the BH3 helix that binds the Bcl-2/Bcl-xL surface pocket, they based their design on the crystal and solution structures of Bak/Bcl-xL complex, which show the helical Bak-peptide binding into a hydrophobic cleft formed by the BH1-BH3 domains of BclxL. From alanine scans of the Bak-peptide it is clear that four hydrophobic residues (Val74, Leu78, Ile81, Ile85) along one edge of the helix are involved in binding. In addition, Asp83 forms an ion pair with a lysine residue of Bcl-xL. A related 26-mer peptide derived from the Bad-protein binds better to Bcl-xL, exploiting larger hydrophobic residues (Tyr, Phe) to induce a slight structural change in the binding region of Bcl-xL. Furthermore, it has been shown that the R-helix propensity of these peptides is decisive for strong binding to Bcl-xL. On the basis of these structural requirements they designed a series of terphenyl molecules containing alkyl or aryl substituents on the three ortho positions (to mimic the key hydrophobic substituents on the helical exterior of Bak or Bad) and carboxylic acid substituents on either end (to mimic the additional ion pair). As determined by the FP competitive binding assay, the terphenyl molecule with two carboxylic acids and the isobutyl, 1-naphthalenemethylene, isobutyl sequence shows the strongest binding to Bcl-xL with a K_i value of 114 nM (see Fig.4.7). The less hydrophobic terphenyls show lower affinity, emphasizing the importance of hydrophobic interactions for binding to the recognition cleft in Bcl-xL. Scrambling the position of the substituents, leads to a significant loss in binding affinity, suggesting an effective shape complementarity for compound 4, as in the natural peptide. The importance of the hydrophobic groups is further confirmed by the weak binding of an analogue of compound 4 lacking the naphthyl and two isobutyl substituents.

160
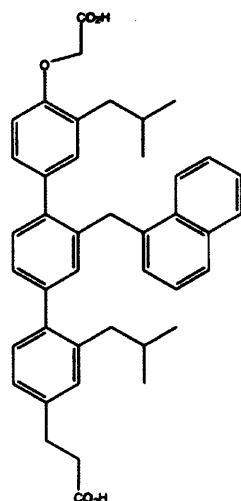
Fig. 4.7 – The most active terphenyl structure.[31]

Finally, the role of the carboxylate groups was probed by partial removal or conversion to positively charged groups, leading in both cases to significant loss of activity. Further NMR experiments and computational docking studies showed that this compound binds to the same hydrophobic surface pocket on Bcl-xL as the Bak BH3peptide.

Among different natural compounds showing antiapoptotic activity (i.e. helenalin A, dykellic acid, etoposide), it was explored the pontential of irofulven (hydroxymethylacylfulvene, HMAF),[32] a novel DNA- and protein-reactive anticancer drug, to overcome the antiapoptotic properties of Bcl-2 in HeLa cells with controlled Bcl-2 overexpression (see Fig.4.8). Irofulven treatment resulted in rapid (12hr) dissipation of the mitochondrial membrane potential, phosphatidylserine externalization, and apoptotic DNA fragmentation, with progressive changes after 24 hr. Bcl-2 overexpression caused marginal or partial inhibition of these effects after treatment times ranging from 12 to 48 hr. Both Bcl-2 dependent and –independent responses to irofulven were abrogated by a broad-spectrum caspase inhibitor. Despite the somewhat decreased apoptotic indices, cell growth inhibition by irofulven was unaffected by Bcl-2 status. In comparison, Bcl-2 overexpression drastically reduced apoptotic DNA fragmentation by etoposide, acting via topoisomerase II-mediated DNA damage, but had no effect on apoptotic DNA fragmentation by helenalin A, which reacts with proteins but no DNA. Irofulven retains its pro-apoptotic and growth inhibitory potential in cell lines that have naturally high Bcl-2 expression. Collectively, the results implicate multiple mechanisms of apoptosis induction by irofulven, which

may differ in time course and Bcl-2 dependence. It is possible that the sustained ability of irofulven to induce profound apoptosis and to block cell growth despite Bcl-2 overexpression may be related to its dual reactivity with both DNA and proteins.
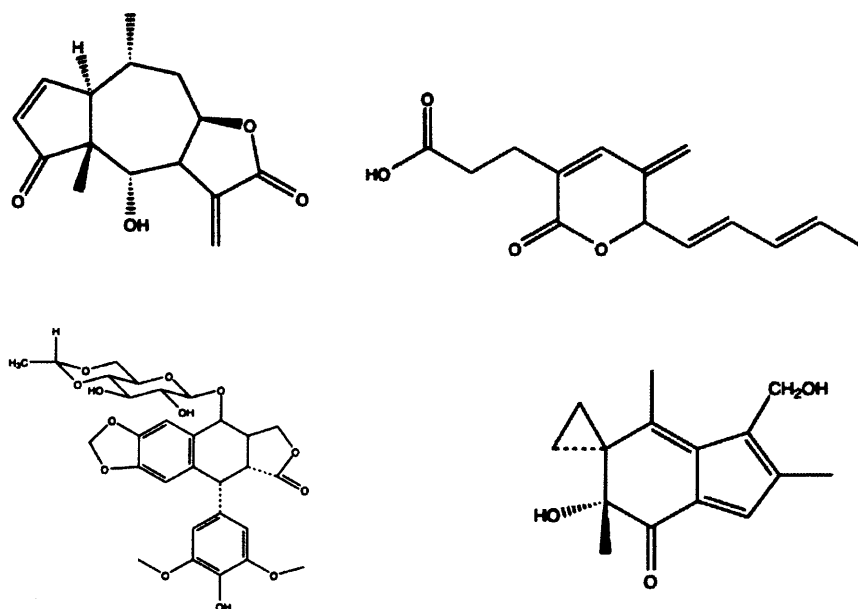
Fig. 4.8 – Helenalin (on the top left), dykellic acid (on the top right), etoposide (on the bottom left) and irofulven[32] (on the bottom right) structures.

Recently, new natural compounds (polyphenols, cathechins, theaflavins) have been discovered, characterized and studied for their direct inhibition of the antiapoptotic Bcl-2 family proteins, and of the BH3 binding to Bcl-xL in particular: apogossypol,[33] gossypol,[34] purporogallin,[34] theaflavin-3'-gallate,[35] theaflavin and theaflavanin[35] (see Fig.4.9, 4.10 and 4.11).

162
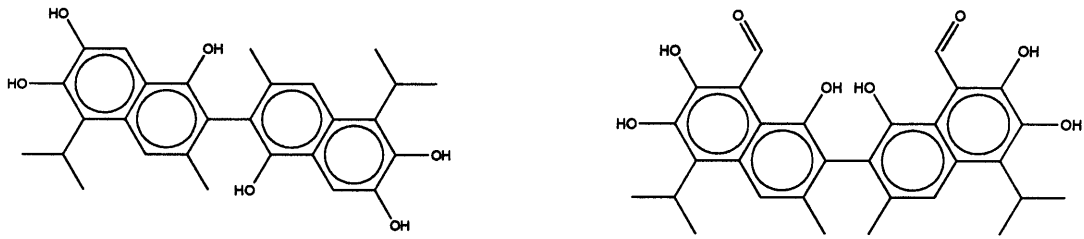
Fig. 4.9 – Apogossypol (on the left) and gossypol (on the right).
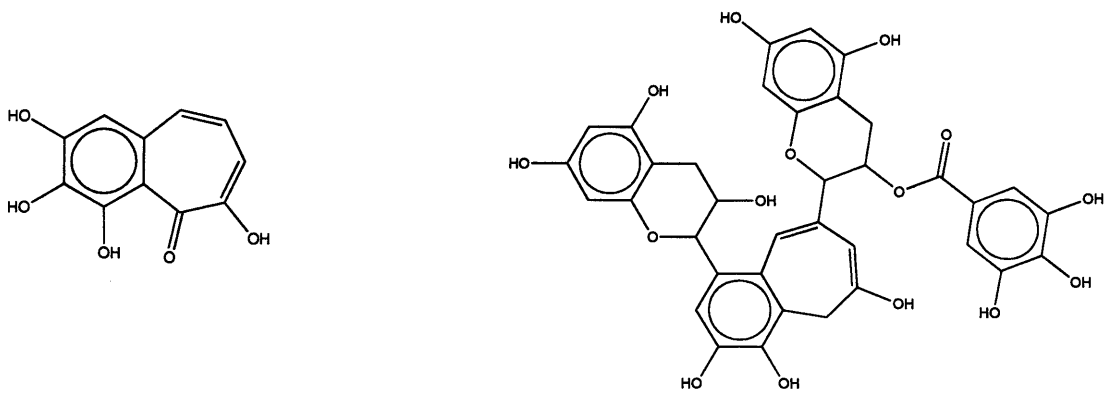
Fig.4.10 – Purporogallin (on the left) and theaflavin-3'-gallate (on the right).
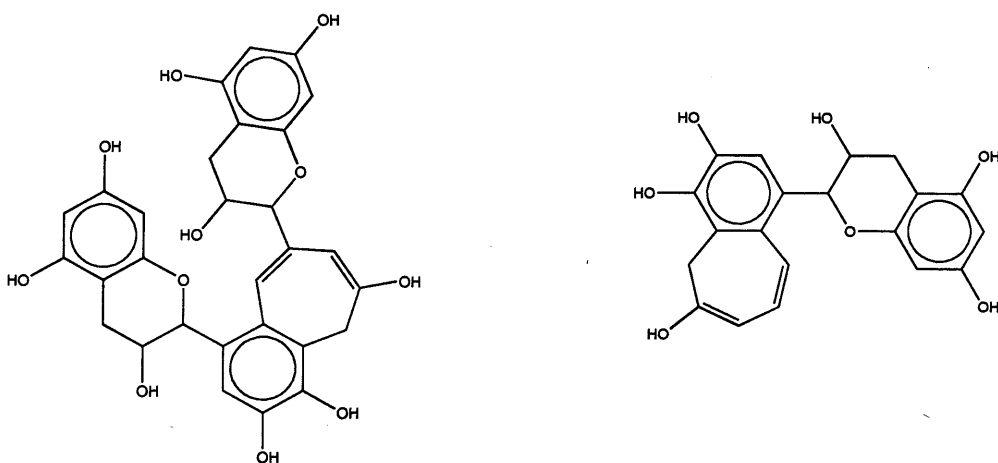
Fig.4.11 – Theaflavin (on the left) and theaflavanin (on the right).

4.4 Apoptosis: links and cellular signaling pathways

Extracellular activating or suppressing signals from other cells as well as intracellular stimuli and pathways regulate and control the programmed cell death. Cell-cell interaction, growth-factor-controlled signalling pathways, cell cycle and DNA check-point system, are few examples of a series of links detectable at distinct levels of apoptosis.

Fatal is the suppression of the programmed cell death in tumorigenesis; in fact, many links between malfunction of apoptotic proteins and tumor formation have been discovered and studied, even if a detailed knowledge about the involved mechanisms still is limited to a few examples.

Indeed, interesting targets involved in the production of malignancy and worth to be subjected to future studies, are the tumor suppressor factor p53, whose genetic mutation causes more than 50% of human tumours, and PI3K, operating along with the signalling pathways of the cells. Both of them require protein-protein recognition and interactions in order to assure a perfect function in the most relevant biological processes.

Interestingly, loss of p53 enhances cellular survival if damaged DNA is present leading to enhanced tumour progression; in addition, it exerts its anti-neoplastic activity primarily by inducting of apoptosis. In fact, it can activate the proapoptotic protein Bax permeabilizing mitochondria and engaging the apoptotic program in absence of other proteins and when it accumulates in the cytosol.[7]

Aims & Objectives

Current treatments of cancer disease include different methods according to the nature of the cancer and how far it is advanced (surgery, radiotherapy, photodynamic therapy, immunotherapy and vaccines, chemotherapy). The discovery of the main biochemical differences between healthy and tumour cells could allow the development of new effective anticancer agents by a more rational and accurate drug design rather than the empirical methods commonly used to evolve and synthesise most of the drugs.[1] Apoptosis (programmed cell death) is often impaired in human cancer and is relevant in chemotherapy-induced tumour cell-killing, therefore targeting the protein-protein interactions involved in proapoptotic/antiapoptotic complexes may be a more effective way of treating cancer.[2] Consequently, the aim of this project is to build a homology

model of the antiapoptotic protein Bcl-2 and to design and synthesise new potential small molecules acting as Bcl-2 inhibitors.

Due to the high percentage of amino acid sequence identity and similarity, the x-ray structure of the antiapoptotic protein Bcl-xl[3] has been used as a template to generate a homology model of the protein Bcl-2, which belongs to the same protein family and shares antiapoptotic activity in the cell along with Bcl-xl. The refined homology model of the antiapoptotic protein Bcl-2[4,5] has been further analyzed to define the active site and locate critical residues for hydrophobic and electrostatic interactions with potential non-peptidic inhibitors of Bcl-2. The investigation has led to the generation and virtual screening of small combinatorial libraries of drug-like compounds, carried out by means of different computational applications (pharmacophore and conformational searches, molecular docking, molecular dynamics) in order to identify a specific lead compound as a reference for starting the chemical synthesis of novel potent proapoptotic agents.

Results and Discussion

Using Composer implemented with Sybyl software package[36] for the homology model of the antiapoptotic protein Bcl-2, three models were built starting from the same template (1MAZ, Bcl-xl X-ray structure) and different query sequences (1G5M - human Bcl-2 isoform 1-, 1GJH - human Bcl-2 isoform 2, both sharing 52,4% of identity with 1MAZ, and 1LXL, Bcl-xl NMR structure).

It is worth noticing that the Bcl-xl protein presents a large, unstructured loop between residues 26-76 that complicate the homology modelling; as a consequence, due to the lack of a loop between residues 23-88, the elaboration of the models was quite difficult: they showed a rigid, straight and unusual chain in this position on the surface of the protein. Inevitable, on the other hand, was the inclusion of these residues in the homology modelling process because of the main biological function of this chain in the binding to natural ligands and in the apoptotic activity of the complex Bcl-2 protein/ligand. An alternative program for comparative protein modelling implemented with MOE software package[37] was subsequently used to elaborate further Bcl-2 homology models aiming to build more accurate models and with the purpose to compare the different results obtained and to improve them. Using the same template and query sequences utilized previously, ten intermediate protein structures were built for each couple of query-template sequences as result of the permutational selection of

different loop candidates and sidechain rotamers. At the end of the calculations, the output database included the 10 models and a minimized average model. For each model were written both a Protein Report collecting a variety of stereochemical measurements useful in the assessment of protein models, and a Protein Consensus isolating outliers. It is very important to examine the protein report and observe whether the residues of importance from a biological perspective have many exceptions (outlier). If these regions contain many exceptions, the model under consideration may be insufficiently minimized or it may be inadequate and require that another candidate model be considered. According to the Protein Report obtained, all ten MOE-homology models of the most relevant couple of query-template sequence (1GJH-1MAZ) needed to be refined with a first minimization without restraints (MMFF94), followed by a Molecular Dynamic Simulation, and a second minimization with dyhedral restraints changing angle, bond, torsion parameters according to how required by each model.

At the end of the refining process, three MOE-homology models without any outliers were obtained:

- model **R1** (energy = -3.7704 kcal/mol - 3 outliers (Ser78, Ala81, Arg82) before refining)

- model **R3** (energy = -3.6918 kcal/mol - no outlier before refining)

- model **R8** (energy = -3.7646 kcal/mol - 1 outlier (Glu50) before refining)

Despite the refinement, still the models showed the rigid and straight chain on the surface of the protein as observed for the Composer models previously elaborated.

The stereochemical quality of all homolog-protein structures was analyzed examining their overall and residue-by-residue geometry through PROCHECK[38] Ramachandran Plots Statistics, according to which a good quality model would be expected to have over 90% in the most favoured regions.

Two of the homology models elaborated by Composer - the former derived from the 1GJH sequence and yielded a PROCHECK score equal to 89%, the latter, obtained considering the 1G5M sequence, yielded a PROCHECK score of 84.9% - were further subjected to minimization and refinement with Composer application (minimization with Conjugate Gradient). The subsequent superimposition between the refined and minimized models and the template revealed that the former had been broken in different chains during the optimization; therefore, only the latter was considered for a final refinement and subjected to an additional minimization in vacuum on Sybyl using a different and stronger algorithm (minimization with Conjugate Gradient), thus

resulting in an interesting RMS value (1.87Å) observed superimposing the model and its template, a higher PROCHECK score (86.3%), a reduced bond length between main chains (from −1.38 Å to 0.20 Å), but also in two negative values referred to dihedral angles, probably due to the presence of the rigid and straight chain on the surface of the protein structure. In order to overcome this undesired structural result, the homology modelling procedure was repeated with Composer using the full amino acid sequence of Bcl-xl as template, and that of Bcl-2 - deprived of the residues corresponding to those between residues 23-88 in the unstructured part of Bcl-xl protein - as query sequence. The resulting model was first minimized on MOE for a fast minimization, and then, once superimposed with 1MAZ structure, further minimized in presence of water keeping the template fixed, thus trying to improve both flexibility and folding of the homolog elaborated. Noticeable were the final results obtained: there was no trace of the previous rigid chain on the surface of the model, PROCHECK score was equal to 88.9%, and the only four amino acids (Thr81, His143, Leu134, Arg142 - referred to as "disallowed regions" in the relative Ramachandran Plot) were in the sidechains and, therefore, not relevant for the biological activity of the Bcl-2 homology model. Along with the homology modelling of the Bcl-2 protein, a database of 443 compounds, including both molecules synthesized by a collaborative group at "La Sapienza" University in Rome, and molecules designed by the author of the present thesis considering the chemical structures of well known antiapoptotic compounds, was built subjecting each structure to minimization, Conformational Search, and Pharmacophore Search using default parameters available in MOE.

FlexX, a molecular docking program implemented with Sybyl,[36] was used to dock all compounds included in the mentioned database into the Bcl-2 homolog. Initially, "multi ligand docking" option was attempted with the purpose to quickly observe possible interactions between molecules and protein; unfortunately, all docked compounds were found outside the protein, i.e. on the surface of the active site and the external loop. Running additional FlexX docking calculations by strategically modifying some of the parameters required for the selection of the active site, did not improve the ligand/receptor interaction either. A different computational approach, provided with Sybyl software package,[36] was then chosen to dock the whole database: FlexiDock. Unable to locate the ligand in the binding site of the receptor by itself, this program requires the user locates the ligand in the binding site manually and roughly. Interestingly, at the end of the calculations, almost all compounds were found internally

the active site interacting with the main amino acids in the core of the receptor (only few molecules were located borderline, behind or on the bottom of the protein), but the high energy values relative to each conformation could clearly determine instability of the ligand/receptor complex and inefficiency of the biological binding.[39,40]

Unsuccessful were also the data derived from molecular docking studies using MOE:[37] the energy values were low enough to guarantee a stable binding, but all molecules were located externally and distant from the binding site making unlikely any interaction between inhibitor and protein. Following the mentioned data, it was supposed that compounds with no specific chemical features for an anti-Bcl-2/Bcl-xl activity may have negatively affected the molecular docking studies, thus suggesting *de novo* drug design of new potential Bcl-2 inhibitors[41] as alternative approach to specifically target the antiapoptotic protein. As LigBuilder[42] requires proper "seeds" to "grow" in the core of the receptor for designing specific ligands, a Multi-Fragment Search on MOE was run to select potential fragments that best fit in the active site. The parameters used for this calculation were as follows:

- partial charges : Kollman '89

- forcefield : MMFF94

- amino acid sequence used as active site: 97-104

- all fragments present in the MOE Multi-Fragment panel were chosen

2-Butyne, acetate ion, methane, methanol, propyne, methylthiol were the selected seeds among a database of 2312 ones evaluated by MOE used by LigBuilder to design various ligands generating a database for each of them including several chemical information (M.W.; cLogP; pkd; C.S.) referred to all single molecules. Disappointingly, the subsequent FlexX docking calculations of the designed candidates did not reveal any improvement compared to the results previously obtained.

The accurate analysis of all known Bcl-2 antiapoptotic proteins/Bcl-2 proapoptotic peptides complexes also by means of alternative computational methods (i.e. FT-DOCK[43] for the molecular docking of protein-protein complexes) aimed to compare all the collected data, identify differences and explain the results obtained.

In particular, the homology alignment between the elaborated Bcl-2 homology model and its template revealed that both location and amino acid sequence of 1MAZ active site were different from those observed in the homolog, thus suggesting that Bcl-2 homology model active site was located in a distinct part from the well known conserved region of the Bcl-2 family proteins. Indeed, as a computational elaboration the homolog protein could show differences in sites of biological relevance for the interaction and activity compared to those of the rest of its family.

This finding successfully led to the evaluation and identification of the Bcl-2 homolog potential active site by means of a specific MOE application, α-Site Finder.[37] Contact statistic and molecular surface studies further aided the analysis of the protein and its binding potentiality; more in detail, the former application calculates preferred locations for hydrophobic and hydrophilic ligand atoms taking into account the 3D atomic coordinates of the receptor. The latter allows togenerate a visual representation of atomic properties projected onto a molecular surface.

Both calculations confirmed the nature and the position of the potential ligand represented with dummy atoms by α-Site Finder.

Among 33 apoptosis inducer molecules[44] included in a small database, 17 compounds were selected considering their chemical functional groups and their similarity to the potential ligand, and then docked into the discovered active site using three different molecular docking programs: FlexX, FlexiDock (both implemented with Sybyl) and 'more_dock.svl' (available with MOE package).[37a]

In contrast with the not relevant results derived from the first two methods, those obtained running MOE application showed the majority of apoptosis inducers fitting exactly the active site and matching the position of the potential ligand elaborated by α-Site Finder at negative energy values.

Following these interesting data, 21 novel compounds were designed based on the pattern and moiety of the apoptosis-inducers with the best docking results in terms of energy and location in the active site. All reference structures were simplified and various hydrophilic functional groups were introduced aiming to increase potential interactions among ligand and receptor. Subsequent 'more_dock.svl' docking of the newly designed proapoptotic candidates clearly showed the potentiality of both bisbenzimide and curcumin derivates as Bcl-2 small molecule inhibitors, since their best

conformations were found in the active site interacting with important amino acid residues.

In order to verify the interesting results deriving from the Genetic Algorithm (GA) based Moe docking of the Apoptosis Inducer database to Bcl-2 homology model active site, previously identified along the flexible loop of the protein by MOE α-Site Finder, the same database was docked to the corresponding binding site of 1MAZ, the template used for the homology modelling of Bcl-2, following identical procedure. All docked structures to 1MAZ matched the α-spheres (representing the chemical nature of the potential ligand inside the binding site) of the active site evaluated by MOE, as it was previously observed with the homolog, thus confirming the promising results of the docking to the Bcl-2 active site.

As bisbenzimide and curcumin showed the more relevant docked conformations, they were selected as reference structures to simplify and modify into scaffolds to build small combinatorial libraries of new derivates.

MOE CombiGen[37] application was used for this purpose applying both symmetric and asymmetric combinations in order to create as many various as possible molecules (Table 4.1). Due to their smaller number of entries, both the bisbenzimide and curcumin derived symmetric combinatorial libraries were rapidly docked to the Bcl-2 homology model active site, previously used for the Apoptosis Inducer database virtual screening.

While checking and analysing all docked structures obtained from the docking, it was very important to observe and identify if there is any potential interaction (hydrogen bonds, electrostatic and hydrophobic interactions) between the molecule and the main amino acids along the loop (Thr47, Ser49, Ser75 according to the MOE homology model enumeration) involved in the mono- and multisite phosphorylation (inhibition) of Bcl-2.[45]

Interestingly, the discovery that taxol binds strongly to the above mentioned residues in the flexible loop thus inhibiting the antiapoptotic activity of Bcl-2,[46] further confirmed the reliability of the α-Site Finder evaluation of the Bcl-2 active site. The taxol structure was docked to the same homolog active site mentioned above by means of GA based Moe docking (undoubtedly, the use of the same location and of the same algorithm in a docking screening allows reasonable and reliable comparison of the results deriving from the calculations). The observation of the taxol final docked structures adopting the same orientation along the loop as bisbenzimide and curcumin, and matching the

position of the α-spheres, encouraged to investigate and study the loop and improve its interaction with potential ligands in more detail.

To increase the probability and occurrence of an efficient and stable binding of the ligand to the loop, new modified combinatorial libraries were created introducing more hydrophilic fragments to the bisbenzimide and curcumin scaffolds, respectively (MOE CombiGen – symmetric combination - Table 4.1). Subsequently, these libraries were docked following the procedure adopted for the previous libraries. The analysis of the docking data revealed different results for the bisbenzimide and curcumin databases. In fact, none of the curcumin derivates was placed in the active site but out or distant from this; all bisbenzimide derivates, instead, were found in the expected position along the loop.

Another approach aiming to improve the interaction ligand/loop, was to investigate potential Ser49 and Thr47 rotamers (MOE Rotamer Explorer),[37] trying to adjust the orientation of their own sidechain and, thus easing the formation of hydrogen bonds with the ligand. In the same time, the Bcl-2 homology protein was subjected to a molecular dynamics simulation, keeping the α-Site Finder active site unfixed and able to find a more stable conformation, and the rest of the protein fixed (MOE: MMFF'94x force field, solvatation not included). The best conformation resulted from the simulation was selected and, after the substitution of the Ser49 with its corresponding best rotamer (Thr47 did not need any change as the molecular dynamics simulation moved it in a favourable position), minimized (MOE: MMFF'94x force field, solvatation not included) and subjected to a further molecular dynamics simulation (MOE: MMFF'94x force field, solvatation not included). After each dynamics simulation, it was noticed the loop laying more down and becoming narrower and curved with respect to the initial protein structure. In order to reduce the dimension of the hole between loop and rest of protein, specific restraints (Table 4.2 below) were introduced in the best conformation resulted from the first molecular dynamics simulation before running a last energy minimization.

Table 4.1 – MOE-CombiGen Libraries *(continues)*

	R1 Fragment Structure	Fragment Name	Modified R1 Fragment Structure	Fragment Name
1		benzyl		2-methylbenzoic acid
2		2-butyl		butan-1-amine
3		N-methylaminobutyl		3-methoxy-N-methylpropan-1-amine
4		iso-butyl		2-methylpropan-1-ol
5		iso-propyl		methoxyethane
6		4-hydroxybenzyl		5-hydroxy-2-methylciclohexane carboxylic acid
7		ethyl		methanethiol
8		propyl		propionaldehyde
9		butyl		propylformate
10		N-methylaminopropyl		4-(methylamino)butanamide
11		2-naphtylmethyl		1-methoxy-3-methyl-decahydronaphtalene

Table 4.1 – MOE-CombiGen Libraries *(continues)*
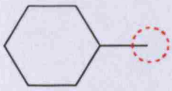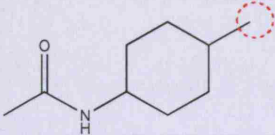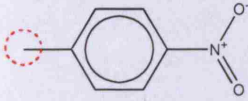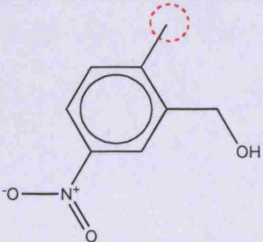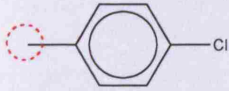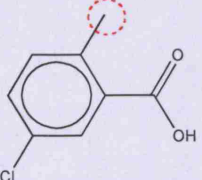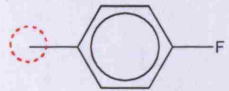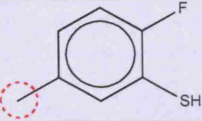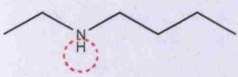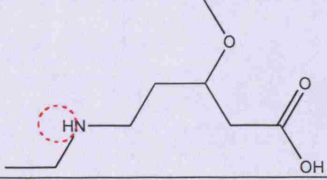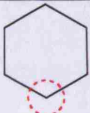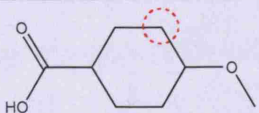
	R1 Fragment Structure	Fragment Name	Modified R1 Fragment Structure	Fragment Name
12		cyclohexylmethyl		N-(4-methylcyclohexyl)acetamide
13		4-nitrobenzyl		(2-methyl-5-nitrohexyl)methanol
14		4-chlorobenzyl		5-chloro-2-methylcyclohexanecarboxylic acid
15		4-fluorobenzyl		2-fluoro-5-methylcyclohexanethiol
16		N-ethylaminobutyl		5-(ethylamino)-3-methoxypentoic acid
17		cyclohexyl		4-methoxycyclohexane carboxylic acid

Table 4.1 – MOE-CombiGen Libraries
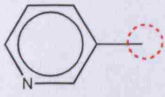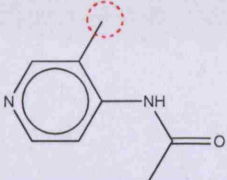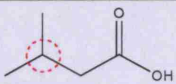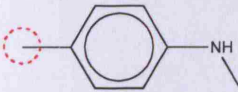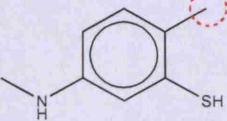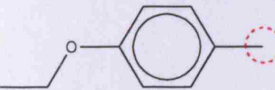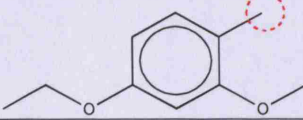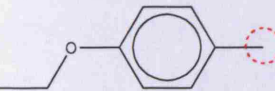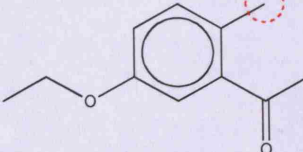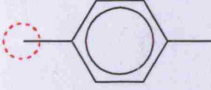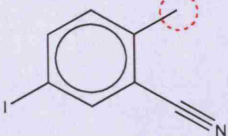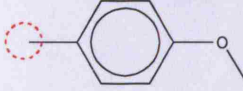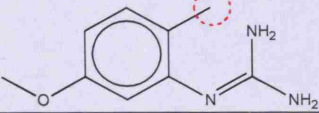
	R1 Fragment Structure	Fragment Name	Modified R1 Fragment Structure	Fragment Name
18		3-pyridylmethyl		N-(3-methylpyridin-4-yl)acetamide
19		t-butyl		3-methylbutanoic acid
20		4-N-methyl amino benzyl		2-methyl-5-(methylamino)cyclohexanethiol
21		4-ethoxybenzyl		4-ethoxy-2-methoxy-1-methylcyclohexane
22		4-ethoxybenzyl		1-(5-ethoxy-2-methylcyclohexyl) ethanone
23		4-iodobenzyl		5-iodo-2-methylbenzonitrile
24		4-methoxybenzyl		2-(5-methoxy-2-methylphenyl)guanidine

Table 4.2 – List of the restraints applied to the protein for the second Molecular Dynamics Simulation.

Residues	Actual Distance	Restraints
Thr47-Pro44	4.59 Å	4.0-4.3 Å
Ser49-Glu45	7.10 Å	6.5-6.8 Å
Val51-Leu96	4.84 Å	4.5-4.7 Å
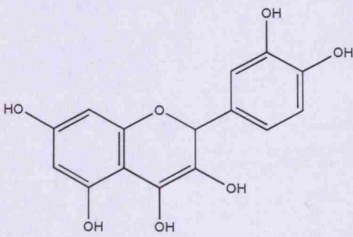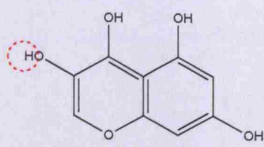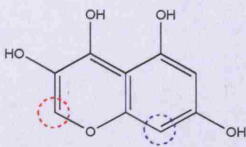
Again, the comparison between the restrained and minimized protein to the initial one clearly revealed the tendency of the loop to "fall down and curve" instead of "going upper" and, thus, reducing the hole between loop and rest of protein. Both the molecular dynamics and the restraints approaches provided unsuccessful results: the final resulting

174

hole became wider and larger, making the ligand free to move in any direction and unable to interact specifically and strongly with the main residues of loop.

Next efforts focused on refining and improving the design of small molecule inhibitors (adjusting size and length of the bisbenzimide derivates, adding more specific chemical substituents, docking different compounds known as antiapoptotic agents) and, subsequently, on the optimization of the inhibitor/Bcl-2 homolog model complex by means of molecular dynamics simulations, fundamental technique to simulate the dynamical properties of the ligand-protein complex.

Interestingly, among the docked structures of different antiapoptotic compounds, those of quercetin (antiapoptotic molecule acting as PI3K inhibitor) and its derivates (obtained by means of combinatorial design – Table 4.3) were found matching exactly the expected position along the loop. Furthermore, these compounds were able to interact and form more than one hydrogen bond with the most relevant amino acids (Glu48, Glu50, Ser49, Thr47, Pro44, Leu96) for the inhibition of the protein.

Table 4.3 – Quercetin: MOE-CombiGen Library.

Additionally, the molecular docking of the latest Bcl-xl inhibitors (apogossypol,[33] gossypol,[34] purpurogallin,[34] theaflavin-3'-gallate,[35] theaflavin,[35] theaflavanin;[35] see

introduction) strongly confirmed the flexible loop as potential active site and proved the key role played by the mentioned amino acids for the inhibition of Bcl-2 protein.

Major attention was therefore paid to an accurate analysis and refinement of these important results.

Once the most promising candidates derived from the molecular docking of quercetin and its designed derivates were selected, they were subjected to an energy minimization and then evaluated considering their ability to conserve interactions and hydrogen bonds formed with the most relevant amino acids (Glu48, Glu50, Ser49, Thr47, Pro44, Leu96) during the docking as prove of their potentiality as Bcl-2 inhibitors. Taking into account the presence of essential hydrogen and covalent bonds and the conformation adopted in the binding pocket as parameters to further reduce the number of molecules and identify the final candidates, 12 compounds were subjected to a first molecular dynamics simulation (MOE),[37] meant as a fast *selective filter* to be followed by a longer, more accurate and definitive molecular dynamics study.

More in detail, each molecule was added of water molecules to make the model a more accurate reflection of the *in vivo* situation, trying to find configurations near to the energy minima considered to be closest to what is actually seen *in vivo*. Indeed, the model aimed to simulate the molecular dynamics of the protein-compound interactions over one hundred picoseconds, resulting in a database of 200 configurations which could be run consecutively as a 'video' sequence. The interesting results obtained showed seven compounds remaining firmly anchored in the active site without being displaced by water molecules as the distances measured between key molecules during the simulation also confirmed. The likelihood of hydrogen and covalent bonds forming has been evaluated taking into account only the last 100 database entries as these are more representative of the 'true' minimum energies which are likely to be seen *in vivo*, whereas the first 100 represent a period of stabilisation where the enzyme and compound are approaching their near-minimum energy configurations.

The lowest energy interaction configuration from the molecular dynamics simulation was then taken and subjected to MOE energy minimisation with a smaller energy gradient (0.0001), in order to be further optimized.

Energy minimization and analysis of both interactions and bonds formed, and conformations adopted by the designed structures, followed with the purpose to compare the results obtained from both molecular docking and molecular dynamics studies.
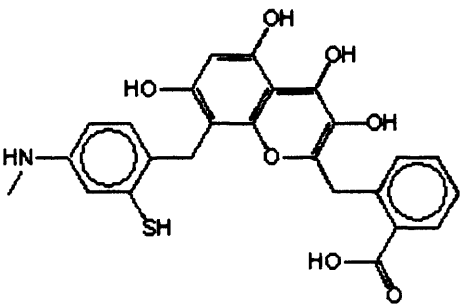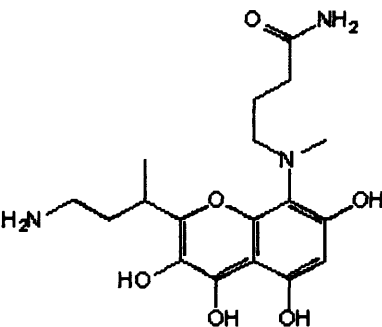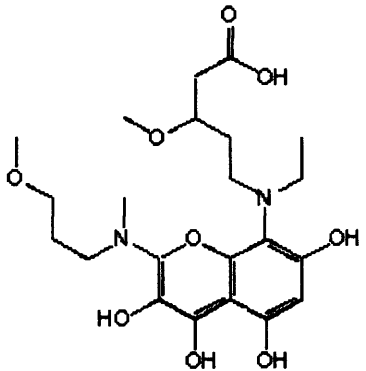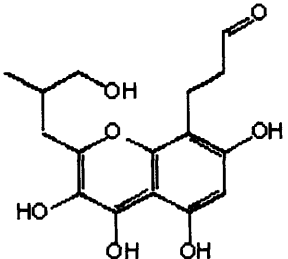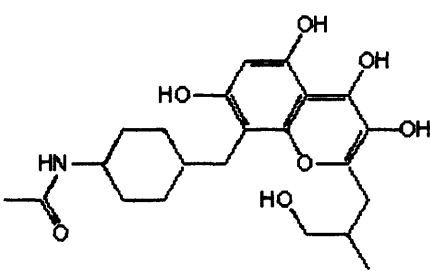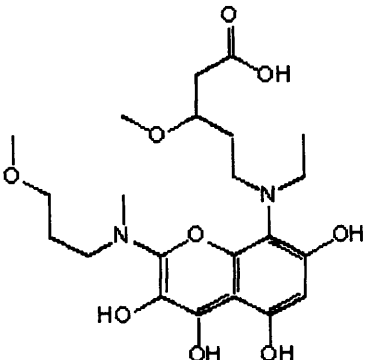
Six quercetin derivates were observed to maintain the same interactions and conformations throughout the mentioned computational investigations, thus confirming their reliable potentiality as stable small molecule Bcl-2 inhibitors. Therefore, these final six compounds (**4.1**; **4.2**; **4.3**; **4.4**; **4.5**; **4.6**; see Table 4.4 for structures) were subjected to a last molecular dynamics simulation using a different and more detailed computational program (GROMACS)[47], aiming to compare and eventually confirm the previous results. The accurate evaluation of the main energy values involved in the protein-ligand binding (LJ, Coulomb, ΔG bind – see Table 4.4 for results), that followed energy minimization and analysis of bonds formed and conformations adopted, revealed new and interesting aspects about the small molecule-Bcl2 complexes under investigation.

Indeed, five candidates (**4.1**; **4.2**; **4.3**; **4.4**; **4.5**) were found to maintain interactions and conformations within the active site very similar to those observed in the previous molecular dynamics study; the corresponding complexes also showed reasonable ΔG energy values for a stable and strong binding. Surprisingly, despite a favourable position in the active site, the most promising candidate resulted from the MOE simulation, **4.6**, lost all the interactions and hydrogen bonds formed and kept during the molecular docking and the first molecular dynamics study.

Worthy of notice was the fifth structure, **4.5**, that conserved not only interactions and conformation similar to those previously observed, but also hydrogen bonds with specific and essential residues for the protein inhibition. Given these final results and the corresponding interesting ΔG energy values, structure **4.5** was finally identified as the most promising candidate to be used as a lead compound to guide the design and optimization of a novel class of Bcl-2 inhibitors, and promote their chemical synthesis.

It could be also noticed the efficiency of both the methodology applied and the computational methods chosen and used to investigate and design new small molecule inhibitors of protein-protein interactions, thus suggesting the same approach to be used in similar projects.

Table 4.4 – GROMACS Molecular Dynamics Simulation: structures and results

Compound	Results	Compound	Results
 4.1	**LJ:** -131.297 **Coul.-SR:** -551.853 **DG bind:** -128.607	 **4.2**	**LJ:** -105.841 **Coul.-SR:** -462.794 **DG bind:** -173.846
 4.3	**LJ:** -130.600 **Coul.-SR:** -525.639 **DG bind:** -482.880	 **4.4**	**LJ:** -115.192 **Coul.-SR:** -115.349 **DG bind:** -23.522
 4.5	**LJ:** -157.39 **Coul.-SR:** -152.313 **DG bind:** -13.206	 **4.6**	**LJ:** -115.142 **Coul.-SR:** -560.631 **DG bind:** -96.721

Conclusion and Future Work

Apoptosis (programmed cell death) is often impaired in human cancer and is relevant in chemotherapy-induced tumour cell-killing, therefore targeting the protein-protein interactions involved in proapoptotic/antiapoptotic complexes may be a more effective way of treating cancer.[48] Consequently, the homology model of the antiapoptotic protein Bcl-2 and the design and synthesise of new potential small molecules acting as Bcl-2 inhibitors can effectively trigger these targets.

Due to the high percentage of amino acid sequence identity and similarity, the X-ray structure of the antiapoptotic protein Bcl-xL[49] was used as a template to generate a homology model of the protein Bcl-2, which belongs to the same protein family and shares antiapoptotic activity in the cell along with Bcl-xL. The model of Bcl-2 was further refined using different current computational methods. Subsequently, the virtual screening of small libraries of drug-like compounds was carried out by means of different computational applications (pharmacophore and conformational searches, identification of the active site, flexible alignment, Molecular Docking, Molecular Dynamics Simulation) in order to find non-peptidic inhibitors of Bcl-2.

An accurate and specific analysis of the data collected revealed a new potential active site in the model and the main chemical features of a putative ligand, then confirmed by biological studies.[50] Newly designed small molecule libraries, including derivates of known apoptosis-inducers, were docked in the predicted active site identifying some structures which actually bound effectively in the pocket of the model.

These results will guide future studies to further define the active site, searching critical residues for hydrophobic and electrostatic interactions and design a specific and accurate lead compound as a reference for starting the chemical synthesis of novel potent proapoptotic compounds.

Furthermore, given these relevant and promising data, future computational evaluation by means of Molecular Dynamics Simulation will investigate the bioactive conformation of these small molecule inhibitors thus further directing and improving the design of new potential derivates.

Chapter 5

Introduction

5.1 p53 and apoptosis

The tumor suppressor protein p53 is characterized by two distinct activities that are mediated by different pathways: growth-inhibiting and pro-apoptotic effects (see Fig. 5.1).

The former is regulated mainly by p21^{CIP1}, which is in turn controlled by p53 at the level of expression; the latter can be exerted either by transcription-dependent or transcription-independent reactions, both leading to apoptosis.
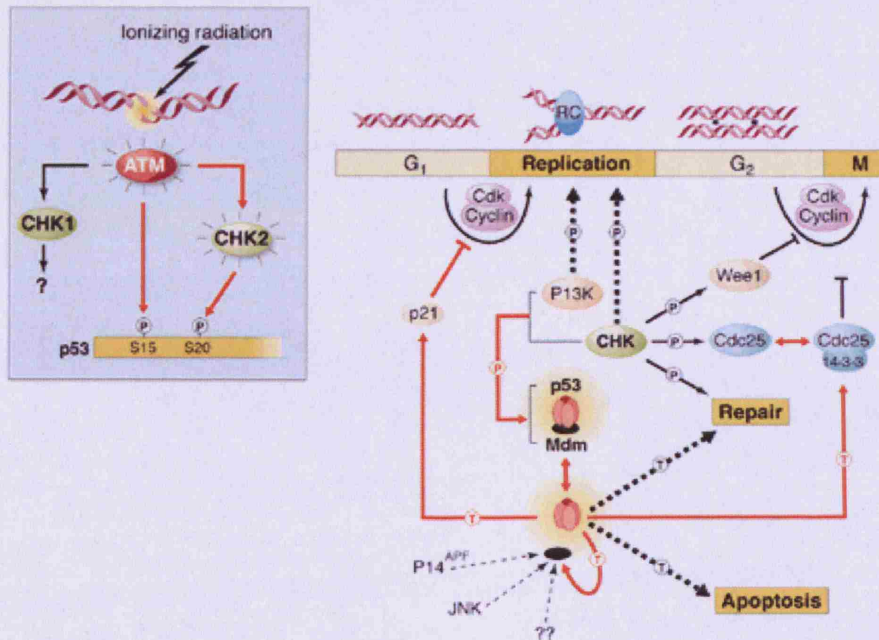
Fig. 5.1 – p53 pathway (www.pharmazie.uni-halle.de/ pb/mgst/Genetik5/).

5.1.2 p53: activation of apoptotic genes

Members of Bcl-2 family proteins, death receptors and their ligands, are few examples of the genes activated by p53 and known to be relevant for programmed cell death (PCD).

In addition, p53 represses transcription of the anti-apoptotic Bcl-2. Although none of these proteins has been identified as pivotal to apoptosis, it has been proved the highly

specific contribution of the mentioned p53-controlled proapoptotic genes to p53-mediated apoptosis.

Therefore, according to the cellular context, post-translational modifications of p53 (phosphorylation or acetylation) can affect the expression pattern of apoptotic genes as well as a specific cell-type binding to distinct transcriptional factors may influence the choice of p53 target genes. Changes in the levels of the main pro-apoptotic proteins is thought to be due to a loss of p53 activity, thus leading to survival of damaged cells that would otherwise be eliminated by PCD.

5.1.3 p53: transcription-independent induced apoptosis

As previously mentioned, p53 acts as tumor suppressor through both transcription-dependent and transcription-independent processes. The transcription-dependent activity of p53 has been extensively studied, whereas the mechanism for transcription-independent p53-mediated tumor suppression has been less characterized.[1] An example of transcription-independent induced apoptosis is the p53-mediated redistribution of Fas receptor from the cytosol to the cell membrane. In addition, p53 is able to partecipate to apoptosis by directly signalling to the mitochondria.

Recently, it has been reported that by complexation of the DNA-binding region of p53 with the anti-apoptotic proteins Bcl-xL and Bcl-2,[2] p53 can directly induce mitochondrial permeabilization and then promote apoptosis. The observed interactions involves the p53 region also used by the protein to contact DNA. It is worth noticing the block of the interaction of p53 with Bcl-xL by the binding of a 25-residue peptide derived from the BH3 region of the pro-apoptotic protein Bad.[1]

5.1.4 p53/Mdm2 complex

As a cellular inhibitor of the p53 tumor suppressor, the Mouse Double Minute (MDM2) oncoprotein binds the transactivation domain of p53 and downregulates its ability to activate transcription.[3] In certain cancers, MDM2 amplification is a common event and contributes to the inactivation of p53. The crystal structure of the 109-residue amino-terminal domain of MDM2 bound to a 15-residue transactivation domain peptide of p53 revealed that MDM2 has a deep hydrophobic cleft on which the p53 peptide binds as an amphipathic helix. The interface relies on the steric complementarity between the

MDM2 cleft and the hydrophobic face of the p53 helix and, in particular, on a triad of p53 amino acids (Phe19, Trp23, and Leu26) which accommodate into the MDM2 cleft. These same p53 residues are also involved in transactivation, supporting the hypothesis that MDM2 inactivates p53 by concealing its transactivation domain.

Since it has been confirmed that wild-type p53 plays a crucial role in the prevention of cancer and that dysfunction of p53 can be caused by increased levels of the protein MDM2, inhibition of MDM2-p53 interaction can stabilize p53 and may offer a novel strategy for cancer therapy. Potent and selective antagonists of MDM2 have been identified and their mode of action have been confirmed through the crystal structures of complexes.[4] These compounds bind MDM2 in the p53-binding pocket and activate the p53 pathway in cancer cells, leading to cell cycle arrest, apoptosis, and growth inhibition of human tumor xenografts in nude mice. The discovery and structure determination of a fungal metabolite, chlorofusin, which antagonizes the p53/MDM2 interaction has been reported[5] (see Fig.5.2).
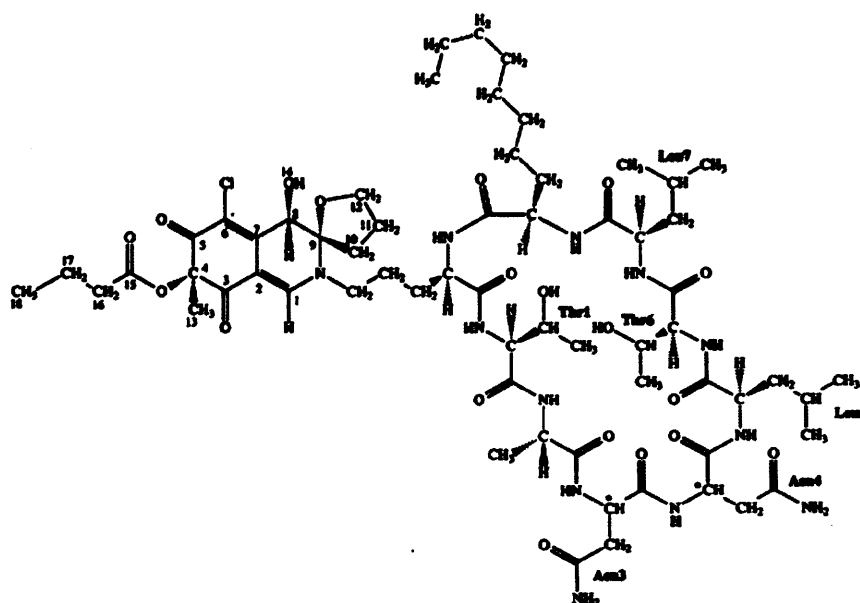
Fig. 5.2 – Chlorofusin structure.[5]

Taking into account that several peptidic MDM2 inhibitors have demonstrate to effectively inhibit the interaction between p53 and MDM2, thus providing a therapeutic strategy for some tumors, non-peptidic HDM2 (the human homologue of MDM2)

inhibitors have been obtained by some research groups using computer-aided design and subsequently synthesized by chemical method (see Fig.5.3).[6] Bio-evaluation showed that some of these inhibitors have affinity with HDM2, and can cause death of some tumor cells which express wild-type p53. Cellular assays showed that one of these compounds, syc-7, can activate the p53 pathway in some of these tumor cell lines, and induce apoptosis. The results suggest that developing non-peptidic small-molecule HDM2 inhibitors is a promising way for new antitumor drug discovery – it is indeed worth reminding how likely a peptidic drug is to be made inactive by enzymatic processes.
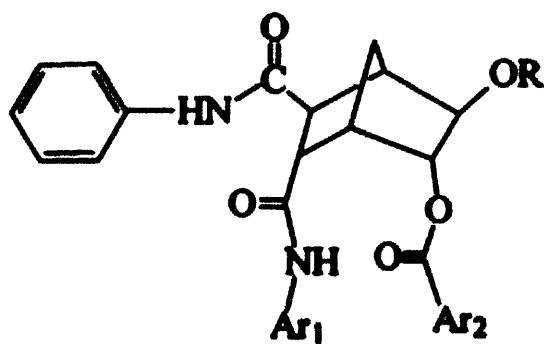
Fig. 5.3 – example of non-peptidic small-molecule HDM2 inhibitor structure.[6]

5.1.5 p53 and its interactions with Bcl-2 family proteins

The tumor suppressor p53 exerts its anti-neoplastic activity primarily through the induction of apoptosis. Cytosolic localization of endogenous wild-type or trans-activation–deficient p53 has been found to be necessary and sufficient for apoptosis. p53 directly activated the proapoptotic Bcl-2 protein Bax in the absence of other proteins to permeabilize mitochondria and engage the apoptotic program.[7] p53 also released both proapoptotic multidomain proteins and BH3-only proteins [proapoptotic Bcl-2family proteins that share only the third Bcl-2homology domain (BH3)] that were sequestered by Bcl-xL. The transcription-independent activation of Bax by p53 occurred with similar kinetics and concentrations to those produced by activated Bid. It has been proposed that when p53 accumulates in the cytosol, it can function similarly to the BH3-only subset of proapoptotic Bcl-2 proteins to activate Bax and trigger

185

apoptosis.[7] Furthermore, recent studies have indicated that p53 has a direct signalling role at mitochondria in the induction of apoptosis, although the mechanisms involved are not completely understood. After cell stress, p53 appears to interact with the proapoptotic mitochondrial membrane protein Bak.[8] Interaction of p53 with Bak causes oligomerization of Bak and release of cytochrome c from mitochondria. Notably, formation of the p53–Bak complex has been shown to coincide with loss of an interaction between Bak and the anti-apoptotic Bcl-2-family member Mcl1. These results are consistent with a model in which p53 and Mcl1 have opposing effects on mitochondrial apoptosis by interacting with, and modulating the activity of, the death effector Bak.

5.1.6 p53 and its interactions with HIPK2

Transcriptional activity of p53, a central regulatory switch in a network controlling cell proliferation and apoptosis, is modulated by protein stability and post-translational modifications including phosphorylation and acetylation. The human serine/threonine kinase homeodomain-interacting protein kinase-2 (HIPK2) has been found to colocalize and interact with p53 and CREB-binding protein (CBP) within promyelocytic leukaemia (PML) nuclear bodies.[9] In fact, following activation by ultraviolet (UV) radiation, HIPK2 selectively phosphorylates p53 at Ser46, thus facilitating the CBP-mediated acetylation of p53 at Lys382, and promoting p53 dependent gene expression. Accordingly, the kinase function of HIPK2 mediates the increased expression of p53 target genes, which results in growth arrest and the enhancement of UV-induced apoptosis. Interference with HIPK2 expression by antisense oligonucleotides impairs UV-induced apoptosis. Results obtained from recent studies imply that HIPK2 is a novel regulator of p53 effector functions involved in cell growth, proliferation and apoptosis.

Aims & Objectives

In the last decade antitumor drug discovery has started to be more interested in the Mdm2/p53 complex; in fact, Mdm2 binds p53 tumor suppressor protein and negatively regulates its transcriptional and antitumour activity as well as its stability through a feedback mechanism. As many human tumors show overexpression of MDM2,

inhibition of MDM2-p53 interaction and consequent activation of p53 offer a novel strategy for cancer therapy. Despite the identification of few small-molecule antagonists of MDM2, development of nonpeptidic MDM2 antagonists has been challenging.

By *in silico* molecular modelling tecniques, the present study aimed to carefully analyze the interface between these two proteins to spot the main residues involved in the binding and, subsequently, to dock distinct compounds into the active site.

The most interesting docked molecule has been chosen as lead compound to create small combinatorial libraries of related structures, that have been further subjected to molecular docking and molecular dynamics simulation to identify potential non-peptidic small-molecule HDM2 inhibitors to synthesize by chemical method.

Results and Discussion

As MOE α-Site Finder did not locate any potential active site in the same position of the imidazoline Mdm2 inhibitor within the X-ray structure of the human protein (1RV1),[10] the p53/Mdm2 complex (1YCR)[10] has been selected as target for computer aided design of protein-protein interaction inhibitors.

Indeed, evaluating the latter protein structure by means of the same MOE tool, two interesting active sites have been identified. These are very closed to each other in the same location of the p53 α-helix that binds to Mdm2 protein.

The study was focused on observing which of these two potential active sites could mimic the natural one, and which of the two related potential ligands (shown as α-spheres by MOE) could lead to the design of small molecule inhibitors.

Following an accurate analysis of both the identified active sites and the p53 amino acid sequence matching the computationally located α-spheres, both p53-Trp23 and –Leu26 residues have been noticed fitting and matching exactly the mentioned α-spheres. Therefore, a new binding site derived from the union of the two identified ones has been considered as target for the virtual screening of small molecule ligands.

Interestingly, the flexible alignment of two compounds, camptothecin and 4-hydroxy tamoxifen (both included in the Apoptosis Inducer database that has been previously mentioned in the Bcl-2 project) to the p53 scaffold (i.e. Trp23, Lys24, Leu25, Leu26), has led to the docking of the whole Apoptosis Inducer database by means of two different softwares: the modified version of the MOE docking[11] tool implementing a

genetic algorithm (GA) based search method, and the FlexX multi-ligand[12] tool using an incremental construction approach.

Although none of the docked structures evaluated by the latter virtual screening has been found in the proposed active site, the former docking has provided very promising results, as all docked compounds, anisomycin and curcumin in particular, matched the position of the α-spheres in an energetically favourable mode. A careful evaluation of all docked structures and scoring values regarding their potentiality in effectively interacting with the target protein, has guided the selection of anisomycin as molecule to simplify and reduce to a scaffold for the building of combinatorial libraries by symmetric combination (MOE-CombiGen).[11] The smallest of these combinatorial libraries has been subjected to a GA based MOE docking and the results obtained have revealed the need of increasing the number of hydrophilic groups in the chemical structure of the designed molecules in order to make the interaction ligand-protein more effective and stronger. In addition, the notable positions of all docked structures in the united active site and their low or negative total energy values have further confirmed both the previous relevant results related to the MOE docking of the Apoptosis Inducer database and the promising choice of anisomycin as scaffold for drug design of small molecule inhibitors (see Fig.5.4).

Consequently, more hydrophilic fragments have been introduced in the creation of new combinatorial libraries by symmetric combination (MOE-CombiGen)[11] and the resulting databases have been docked again by means of the GA based MOE docking. The subsequent best docked and scored structures have been further subjected to two distinct MOE docking tools, one based on a Simulated Annealing simulation and one implementing a Tabu Search algorithm.

The purpose of these two additional dockings was to study and compare the behaviour of the designed compounds using different algorithm based search methods and, interestingly, the results obtained not only have confirmed all the relevant and promising data observed previously, but also shown a closer and stronger binding between the designed small molecule inhibitors and Mdm2 than that present between the known Mdm2 inhibitor, Nutlin,[13] and the protein (see Fig.5.5).

The best compounds found as a result of this process were taken and subjected to a more in-depth molecular modelling process in an attempt to determine the reliability of the models constructed up to this point and to gain a more accurate picture of how the compounds – and the protein – may behave *in vivo*. The configurations showing the

most favourable interactions from those generated by GA based MOE docking were minimized (MOE minimization, MMFF94x force field, energy gradient 0,0001), added water to accurately reflect the *in vivo* situation, and molecular dynamics simulations were run. This stage of the modelling has tried to find configurations near to the energy minima (as calculated by MOE), as this should be closest to what is actually seen *in vivo*; more in detail, the model simulates the molecular dynamics of the protein-compound interactions over one hundred picoseconds, resulting in a database of 200 configurations which could be run consecutively as a 'video' sequence. The aim of this additional study was to prune the number of structures and select a lead compound as starting point for optimization process and chemical synthesis of improved derivates. Only the last 100 database entries have been taken into account as they are more representative of the 'true' minimum energies which are likely to be seen *in vivo*. The first 100, on the other hand, represent a period of stabilisation where protein and compound are approaching their near-minimum energy configurations.

In contrast with the data provided by the docking calculations, the results from this simulation were not encouraging, as they showed the compounds moving out the active site and being displaced by water molecules despite the additional minimization of the complex always required after simulation process (MOE minimization, MMFF94x force field, energy gradient 0.0001).
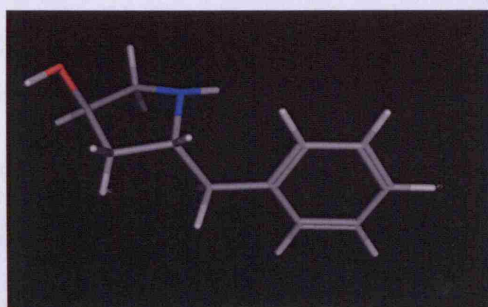
Fig. 5.4 – Anisomycin scaffold.
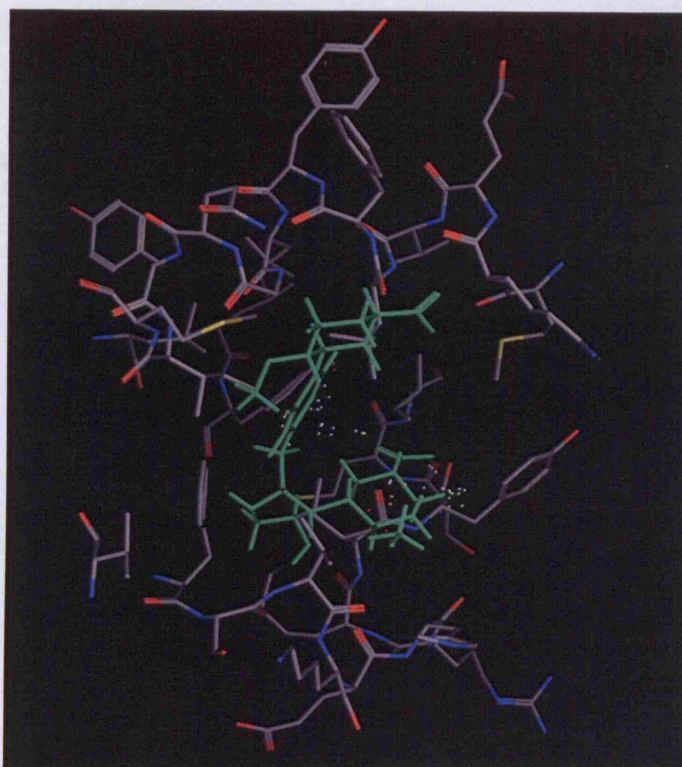
Fig. 5.5 – Simulated Annealing docking results of anisomycin derivates.
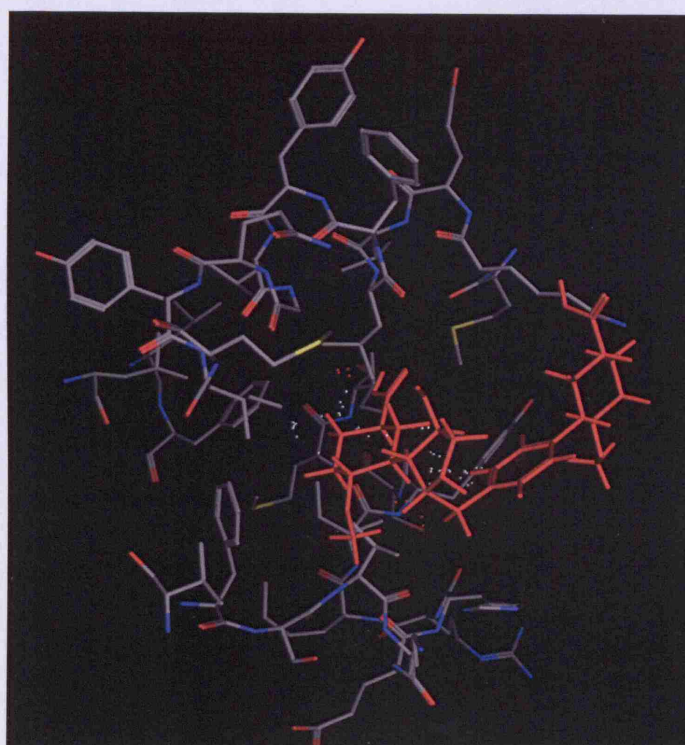
Fig. 5.6 – Tabu Search docking results of anisomycin derivates.

Conclusion and Future Work

Both the design of anisomycin derivates as small molecule p53/Mdm2 interaction inhibitors by combinatorial libraries generation, and the molecular docking studies of the derived structures, have proved to be successful and accurate procedures to identify potential candidates, whereas molecular dynamics simulation has highlighted the need to further refine and optimize them.

Modifying the position of the hydrophilic substitutes or varying both R-groups and position on the anisomycin scaffold may be alternative approaches for structure refinement, and ligand-protein affinity and binding improvement.

Molecular docking of newly generated compounds can function as *first filter* for the selection of potential p53/Mdm2 inhibitors, the most promising of which are to be further evaluated and, eventually confirmed, by a *second final filter* such as molecular dynamics simulation.

Chapter 6

Experimental Section

Molecular Modelling

General Information

All molecular modelling studies were performed on a RM Innovator Pentium IV 2.4GHz running Linux Fedora Core 3 using Molecular Operating Environment (MOE) 2004.03[1] and FlexX module in SYBYL 7.0[2] molecular modelling software.

All the minimisations were performed with MOE until RMSD gradient of 0.05 Kcal mol^{-1} Å^{-1} with the MMFF94x force field and the partial charges were automatically calculated.

Docking experiments were carried out using the FlexX docking programme of SYBYL with the default. The output of FlexX docking was visualised in MOE and the scoring.svl script was used to identify interaction types between ligand and protein.

Molecular dynamics was performed with MOE using the NVT environment for 100 ps and constant temperature of 300 K using the MMFF94x force field with all other default settings in MOE-dynamics chosen. The lowest energy configuration from the molecular dynamics calculations was subjected to energy minimisation calculations using the MMFF94x force field. The water molecules and defining box were left in place and an energy gradient of 0.0001 chosen.

The Combigen function within the Molecular Operating Environment (MOE) program was used to generate databases from fragments given in the text.
Minimisation of the generated structures was performed within the MOE Database Viewer (settings used: rebuild from scratch, calculate force field partial charge, preserver chirality, add hydrogens to structures). Structures which were subsequently found to have incorrect configuration were manually altered within MOE.

Experimental information on computational studies on antitubulin agents (Chap. 2) – I[3]

All the calculations were performed on a RM Innovator computer with a 2.8 GHz Pentium IV processor running Linux Fedora Core 1. All the simulations were carried out using the Molecular Operating Environment (MOE) 2004.03 software package.[1]

All the energy minimizations were performed using the MMFF94s force field, adding the corresponding partial charges to all atoms, and they were terminated when the RMS value fell below 0.001. The docking experiment was performed using the more_dock.svl script available from http://svl.chemcomp.com. A genetic algorithm global search method, in conjunction with the MMFF94s force field, was used with these parameters: Generations per run 1500; Birth rate 7; Mutation frequency 3; Number of runs 5, starting from a random conformation.

The tubulin structure was obtained from http://www.rcsb.org/pdb/index.html (PDB code: 1SA0). Hydrogen atoms were added and their orientation optimized by fixing the heavy atoms and minimizing the structure. The active site between Chain C and Chain D was used for the docking simulation after the removal of the colchicine analogue. The desired arylthioindole (21) was built with Molecule Builder, minimized and docked into the binding pocket. The second highest scored pose was selected, and the protein/ligand complex was minimized, leading to the structure discussed in the manuscript.

Experimental information on computational studies on antitubulin agents (Chap. 2) – II[4]

All molecular modelling studies were performed on a RM Innovator with Pentium IV 2.8 GHz processor, running Linux Fedora Core 3 using Molecular Operating Environment (MOE) 2004.03[1] and the FlexX module in SYBYL 7.0.[2] The tubulin structure was downloaded from the PDB data bank (http://www.rcsb.org/pdb/index.html PDB code: 1SA0).[5]

All the minimizations were performed with MOE until RMSD gradient of 0.05 kcal mol-1 Å-1 was reached with the MMFF94x force field. The partial charges were automatically calculated. Docking experiments were carried out using the FlexX docking program of SYBYL 7.0 using the default settings. The output of FlexX docking was visualized in MOE, and the scoring.svl script[6] was used to identify interaction types

between ligand and protein. Molecular dynamics was performed with MOE using the NVT environment for 600 ps and constant temperature of 300 K using the MMFF94x force field with a time step of 2 fs. Residues within 15 Å of the ligand were allowed to move freely, keeping the rest of the protein fixed. The binding site was soaked in a water sphere of 25 Å radius from the sulfur atom of the ligand, and the total charge of the system included in the water droplet did not require any adjustment. The water molecules were energy minimized keeping the coordinates of the protein-ligand complex fixed before the MD simulation. A distance restraint of 25 Å with a weight of 100 between the oxygen atoms of the water molecules and the sulfur atom of the ligand was also applied.

Chemistry

General Information

All chemicals, reagents and solvents were purchased from Aldrich or purified by standard techniques.

THIN LAYER CHROMATOGRAPHY

Silica gel plates (Merck Kieselgel 60F$_{254}$) were used and were developed by the ascending method. After solvent evaporation, compounds were visualised by irradiation with UV light at 254nm and 366nm.

COLUMN CHROMATOGRAPHY

Glass columns were slurry packed in the appropriate eluent under gravity, with Woelm silica (32-63mm). Samples were applied as a concentrated solution in the same eluent. Fractions containing the product were identified by TLC, combined and the solvent removed in vacuo.

FLASH CHROMATOGRAPHY

Pre-filled ISOLUTE silica columns were washed with the appropriate solvent, and samples were applied as a concentrated solution in the same eluent. The samples were run on a FlashMaster Personal (K10607, Jones Chromatography Ltd.).

NMR SPECTROSCOPY

^{1}H, ^{13}C, DEPT NMR spectra were recorded on a Bruker AVANCE 500 spectrometer (500MHz and 75MHz respectively) and auto calibrated to the deuterated solvent reference peak. Chemical shifts are given in δ relative to tetramethylsilane (TMS); the coupling constants (J) are given in Hertz. The spectra were recorded in CDCl$_3$ at room temperature; TMS served as an internal standard (δ = 0 ppm) for ^{1}H NMR and CDCl$_3$ was used as an internal standard (δ = 77.0 ppm) for ^{13}C NMR.

Synthesises of Antitubulin and Anti-HCV agents
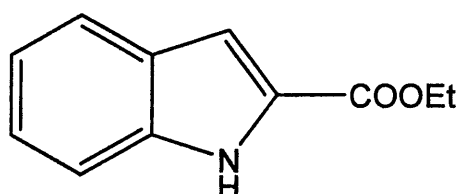
Synthesis of ethyl 1H-indole-2-carboxylate

To a solution of indole-2-carboxylic acid (27 g, 0.17 mmol) was added ethanol (400 mL) and concentrated sulphuric acid (4 mL) which was heated under reflux for 18 hours. The resulting mixture was cooled and concentrated *in vacuo*. Dilute sodium hydroxide solution was added to form an aqueous layer for extraction. The product was extracted using dichloromethane as the solvent. The organic layer was washed with water and MgSO₄ was used to dry the organic layer. The resulting mixture was filtrated and evaporated to quantitatively yield the product, ethyl 1*H*-indole-2-carboxylate.[151]

151. Gay P. Fagan et al., Indoline analogues of idazoxan: potent α2-antagonists and α1-agonists, *Journal of Medicinal Chemistry*, 1988, Vol.31, pp. 944-948.
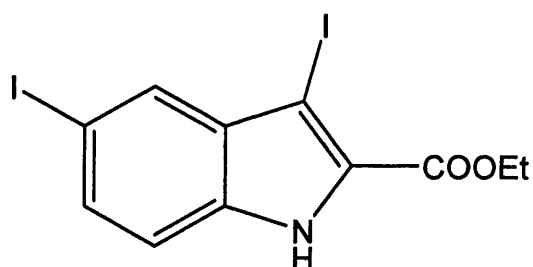
Synthesis of Ethyl 3,5-diiodo-1H-indole-2-carboxylate

I_2 (13.42 g, 52.8 mmol. 2.0 eq), NaIO4 (5.66 g, 26.40 mmol, 1.00 eq) and H_2SO_4 (5.88 mL, 105.60 mmol, 4.00 eq) was added to stirring solution of ethyl-1*H*-indole-2-carboxylate (10.00 g, 52.80 mmol, 2.00 eq) in 100 mL of absolute EtOH. The mixture was refluxed for 90 minutes and then cooled to an ambient temperature. The mixture as then poured into saturated aqueous Na_2SO_3 before being extracted three times with AcOEt. The organic layer was washed with brine and dried with Na_2SO_4. The resulting mixture was filtered and concentrated, *in vacuo*, to yield ethyl 3,5-diiodo-1H-indole-2-carboxylate (85.0% yield).

^{1}H NMR (500MHz, CDCl$_3$) δ9.22 (s, 1H), δ7.84 (s, 1H), δ7.54 (dd, J=7.1Hz and J=1.6Hz, 1H), δ7.10 (d, J=8.7Hz, 1H), δ4.42 (q, J=7.9Hz, 2H) and δ 1.40 (t, J=7.1Hz, 3H) ppm.
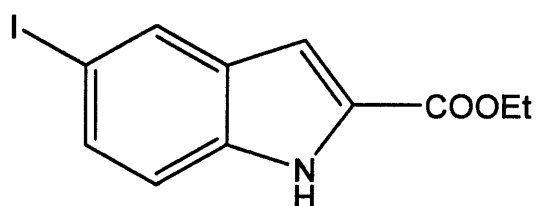
Synthesis of Ethyl 5-iodo-1H-indole-2-carboxylate

Concentrated aqueous hydrogen chloride (35.60 mL, 430.00 mmol) was added to a stirring suspension of ethyl 3,5-diiodo-1H-indole-2-carboxylate (18.91 g, 43.00 mmol, 1.63 eq) in EtOH (405 mL). The reaction was left stirring for 90 minutes. Zinc dust (43.60 g, 668 mmol) was added portion wise over the 90 minutes. The mixture was diluted with water after stirring for 30 minutes. This was the extracted four times with AcOEt. Aqueous saturated NaHCO₃ and brine was added once to wash the combined extracts and then dried with Na₂SO₄. The resulting mixture was filtered and concentrated, *in vacuo*, before crystallization thrice using EtOAc/hexanes to yield ethyl 5-iodo-1H-indole-2-carboxylate as pink crystals (56.0% yield).

^{1}H NMR (500MHz, CDCl₃) δ8.90 (s, 1H), δ8.06 (d, J=0.5Hz, 1H), δ7.59 (dd, J=7.1Hz and J=1.6Hz, 1H), δ7.21 (d, J=8.7Hz, 1H), δ7.16 (d, J=1.0Hz, 1H), δ4.44 (q, J=7.15Hz, 3H), δ1.44 (t, J=7.15Hz, 2H) ppm.
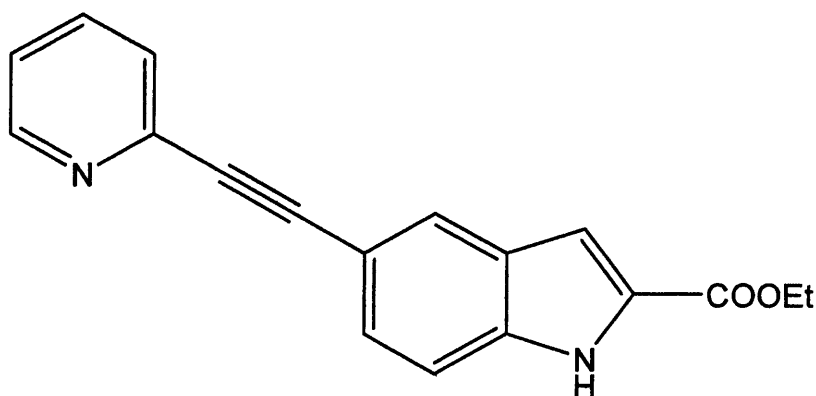
Synthesis of 5-Pyridin-2-ylethynyl-1H-indole-2-carboxylic acid ethyl ester

To a solution of anhydrous THF (7.5 mL) and ethyl 5-iodo-1H-indole-2-carboxylate (100.00 mg, 0.32 mmol, 1.00 eq) was added PdCl$_2$ (22.30 mg, 0.03 mmol, 0.1 eq), Cu(I)I (12.10 mg, 0.06 mmol, 0.2 eq), DIPEA (5.0 mL) and 2-ethynylpyridine (32.70 mg, 0.03mL, 2.0 eq) was added. The reaction was left stirring overnight at room temperature under a nitrogen atmosphere. Water was added to the reaction mixture and then extracted with AcOEt. The organic layer was dried on MgSO$_4$, filtered and the solvent evaporated. Purification of the final compound was carried out by column chromatography using AcOEt as eluent (32.34% yield).

^{1}H NMR (500mHz, CDCl$_3$) δ9.37 (s, 1H), δ8.56 (d, J=4.4Hz, 1H), δ7.89 (s,1H), δ7.61 (t, J=7.8Hz and J=1.5Hz, 1H), δ7.48(t, J=8.35Hz and J=1.2Hz, 1H), δ7.34 (d, J=8.55Hz, 1H), δ7.19 (s, 1H), δ7.15(d, J=0.6Hz,1H), δ4.34 (q, J=7.13Hz, 2H) and δ1.34(t, J=7.13Hz, 3H) ppm.

^{13}CNMR (500MHz, CDCl$_3$): 161.75 (CO), 150.02 (aromatic CH), 143.86 (aromatic C), 135.79 (aromatic CH), 136.13 (aromatic CH), 128.90 (aromatic CH), 128.70 (aromatic CH), 127.35 (aromatic CH), 127.12 (aromatic CH), 126.98 (aromatic CH), 122.43 (aromatic CH), 114.46 (aromatic C), 112.19 (aromatic CH), 108.70 (aromatic C), 77.19 (C≡C), 67.96 (C≡C), 60.80 (CO$_2$CH$_2$CH$_3$) and 14.29 (CO$_2$CH$_2$CH$_3$) ppm.
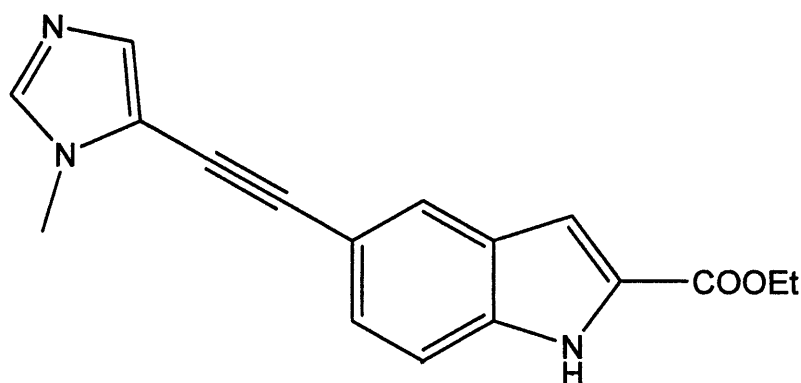
201

Synthesis of 5-(1-Methyl-1H-imidazol-2-ylethynyl)-1H-indole-2-carboxylic acid ethyl ester

To a solution of anhydrous THF (5.0 mL) and ethyl 5-iodo-1H-indole-2-carboxylate (100.00 mg, 0.32 mmol, 1.00 eq) was added $PdCl_2$ (22.30 mg, 0.03 mmol, 0.1 eq), Cu(I)I (12.10 mg, 0.06 mmol, 0.2 eq), DIPEA (5.0 mL) and 5-Ethynyl-1-methyl-1H-imidazole (101.00 mg, 0.10 mL, 3.00 eq) was added. The reaction was left stirring overnight at room temperature under a nitrogen atmosphere. Water was added to the reaction mixture and then extracted with AcOEt. The organic layer was dried on $MgSO_4$, filtered and the solvent evaporated. Purification of the final compound was carried out by column chromatography using AcOEt as eluent (19.0% yield).

^{1}H NMR (500mHz, CDCl$_3$) δ7.78 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.69 (d, J=7.5Hz, 1H), δ8.10 (d, J=1.5Hz and J=1.5Hz, 1H), δ8.40 (s, 1H), δ7.22 (d, J=1.5Hz, 1H), δ7.76 (d, J=1.5Hz, 1H), δ7.84 (d, J=1.5Hz, 1H), δ3.86 (s, 3H), δ4.41 (a) d, J=2.0Hz, t, J=6.0Hz, 2H), δ4.36 (b) d, J=2.0Hz, t, J=6.0Hz, 2H) and δ1.25 (t, J=6.0Hz, 3H) ppm.
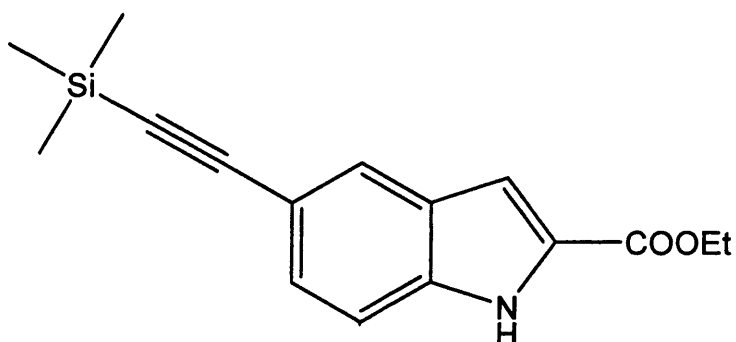
Synthesis of 5-Trimethylsylanylethynyl-1H-indole-2-carboxylic acid ethyl ester

To a solution of anhydrous THF (8.5 mL) and ethyl 5-iodo-1H-indole-2-carboxylate (200.00 mg, 0.63 mmol, 1.00 eq) was added $PdCl_2$ (41.80 mg, 0.03 mmol, 5.7%), Cu(I)I (8.5 mg, 0.04 mmol, 7%), DIPEA (8.5 mL) and (trymethylsilyl) acetylene (119.85 mg, 0.17 mL, 1.26 mmol, 2eq) was added. The reaction was left stirring at 40°C for 4 hours and then overnight at room temperature under a nitrogen atmosphere. The reaction mixture was filtrated and washed with THF, concentrated *in vacuo* and then purified by column chromatography using DCM as eluent (48.5% yield).

^{1}H NMR (500mHz, $CDCl_3$) δ7.66 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.07 (d, J=7.5Hz, 1H), δ7.90 (d, J=1.5Hz and J=1.5Hz, 1H), δ8.40 (s, 1H), δ7.22 (d, J=1.5Hz, 1H), δ0.08 (s, 3H), δ4.29 (a) d, J=2.0Hz, t, J=6.0Hz, 2H), δ4.29 (b) d, J=2.0Hz, t, J=6.0Hz, 2H) and δ1.30 (t, J=6.0Hz, 3H) ppm.
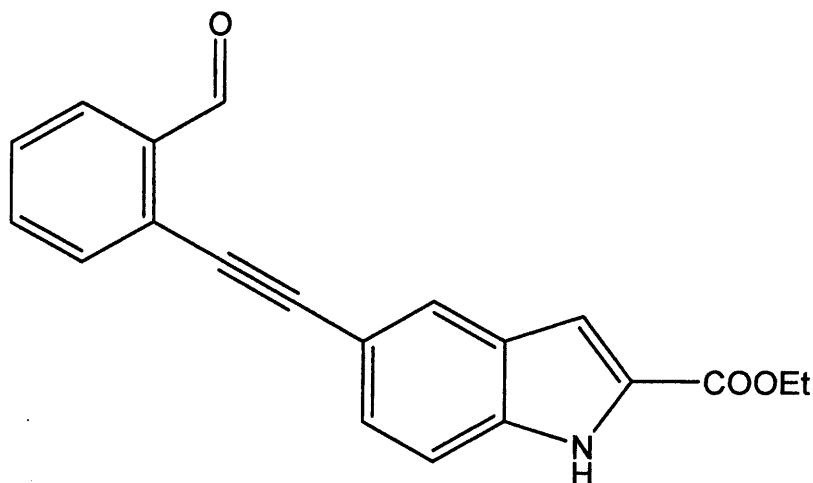
O

COOEt

N
H

Synthesis of 5-(2-Formyl-phenylethynyl)-1H-indole-2-carboxylic acid ethyl ester

To a solution of anhydrous THF (5 mL) and ethyl 5-iodo-1H-indole-2-carboxylate (100.00 mg, 0.32 mmol, 1.00 eq) was added PdCl₂ (22.3 mg, 0.03 mmol, 10%), Cu(I)I (12.00 mg, 0.06 mmol, 20%), DIPEA (5 mL) and 2-ethynylbenzaldehyde (123.9 mg, 0.10 mL, 0.95 mmol, 3 eq) was added. The reaction was left stirring overnight at room temperature under a nitrogen atmosphere. The reaction mixture was concentrated *in vacuo*, extracted with AcOEt and then purified by column chromatography (AcOEt/hexane 1:3) (23.45% yield).

^{1}H NMR (500mHz, CDCl₃) δ7.66 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.59 (d, J=7.5Hz, 1H), δ7.98 (d, J=1.5Hz and J=1.5Hz, 1H), δ8.37 (s, 1H), δ7.19 (d, J=1.5Hz, 1H), δ7.92 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.45 (d, J=7.5Hz, J=7.5Hz and J=1.5Hz, 1H), δ7.54 (d, J=1.5Hz, J=7.5Hz and J=7.5Hz, 1H), δ7.81 (d, J=1.5Hz and J=7.5Hz, 1H), δ10.29 (s, 1H) δ4.38 (d, J=1.1Hz, t, J=6.0Hz, 2H) and δ1.26 (t, J=6.0Hz, 3H) ppm.
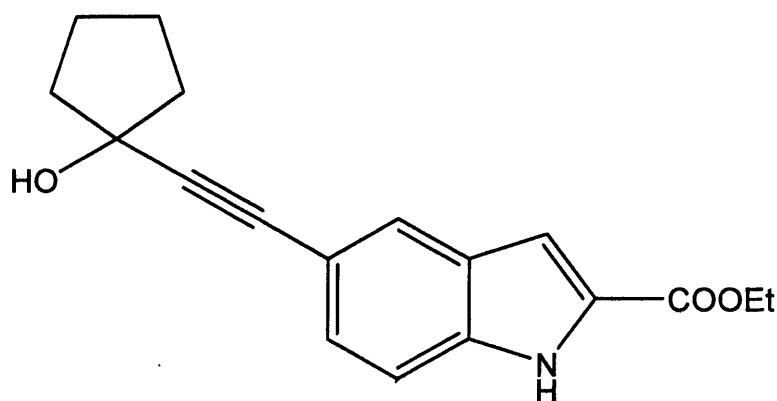
Synthesis of 5-(1-Hydroxy-cyclopentylethynyl)-1H-indole-2-carboxylic acid ethyl ester

To a solution of anhydrous THF (5 mL) and ethyl 5-iodo-1H-indole-2-carboxylate (100.00 mg, 0.32 mmol, 1.00 eq) was added PdCl₂ (22.3 mg, 0.03 mmol, 10%), Cu(I)I (12.00 mg, 0.06 mmol, 20%), DIPEA (5 mL) and 1-ethynyl-cyclopentanol (104.8 mg, 0.11 mL, 0.95 mmol, 3 eq) was added. The reaction was left stirring overnight at room temperature under a nitrogen atmosphere. The reaction mixture was concentrated *in vacuo*, extracted with AcOEt and then purified by column chromatography (AcOEt/hexane 1:3) (37.6% yield).

¹H NMR (500mHz, CDCl₃) δ7.50 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.50 (d, J=7.5Hz, 1H), δ7.85 (d, J=1.5Hz and J=1.5Hz, 1H), δ8.35 (s, 1H), δ7.14 (d, J=1.5Hz, 1H), δ1.29 (s, 1H), δ2.19 (a) d, J=12.5Hz, J=4.8Hz and J=8.5Hz, 2H) δ1.83 (b) d, J=12.5Hz, J=8.6Hz and J=5.4Hz, 2H,), δ1.64 (a) d, J=4.8Hz, J=8.6Hz, J=12.5Hz, J=3.8Hz and J=9.5Hz, 2H) δ1.65 (b) d, J=8.5Hz, J=5.4Hz, J=12.5Hz, J=9.6Hz and J=5.4Hz, 2H), δ1.68 (a) d, J=3.8Hz, J=9.6Hz, J=12.5Hz, J=4.5Hz and J=7.9Hz, 2H) δ1.62 (b) d, J=9.5Hz, J=5.4Hz, J=12.5Hz, J=7.9Hz and J=6.4Hz, 2H), δ1.87 (a) d, J=4.5Hz, J=7.9Hz, and J=12.5Hz, 2H) δ2.14 (b) d, J=7.9Hz, J=6.4Hz and J=12.5Hz, 2H), δ4.40 (a) d, J=3.2Hz, t, J=6.0Hz, 2H), δ4.34 (b) d, J=3.2Hz, t, J=6.0Hz, 2H) and δ1.24 (t, 3H) ppm.
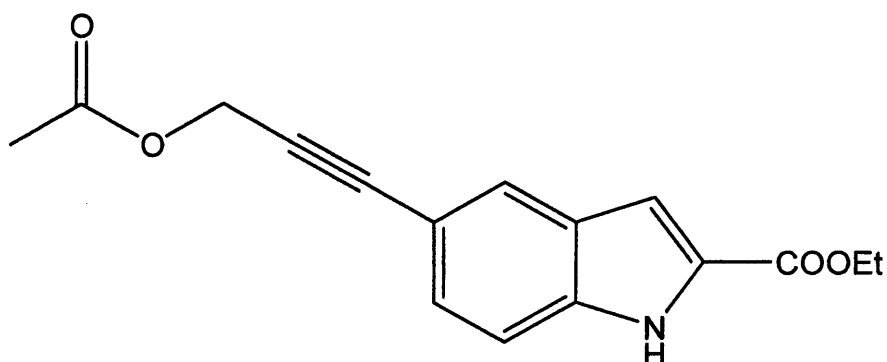
Synthesis of 5-(3-Acetoxy-prop-1-ynyl)-1H-indole-2-carboxylic acid ethyl ester

To a solution of anhydrous THF (5 mL) and ethyl 5-iodo-1H-indole-2-carboxylate (100.00 mg, 0.32 mmol, 1.00 eq) was added $PdCl_2$ (22.3 mg, 0.03 mmol, 10%), Cu(I)I (12.00 mg, 0.06 mmol, 20%), DIPEA (5 mL) and propargylate (93.4 mg, 0.10 mL, 0.95 mmol, 3 eq) was added. The reaction was left stirring overnight at room temperature under a nitrogen atmosphere. The reaction mixture was concentrated *in vacuo*, extracted with AcOEt and then purified by column chromatography (AcOEt/hexane 1:3) (18.45% yield).

^{1}H NMR (500mHz, CDCl$_3$) δ7.51 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.45 (d, J=7.5Hz, 1H), δ7.83 (d, J=1.5Hz and J=1.5Hz, 1H), δ8.01 (s, 1H), δ6.94 (d, J=1.5Hz, 1H), δ4.91 (a) s, 2H), δ4.94 (b) s, 2H), δ2.10 (s, 3H), δ4.26 (a) d, J=1.6Hz, t, J=6.0Hz, 2H), δ4.16 (b) d, J=1.6Hz, t, J=6.0Hz, 2H) and δ1.30 (t, J=6.0Hz, 3H) ppm.
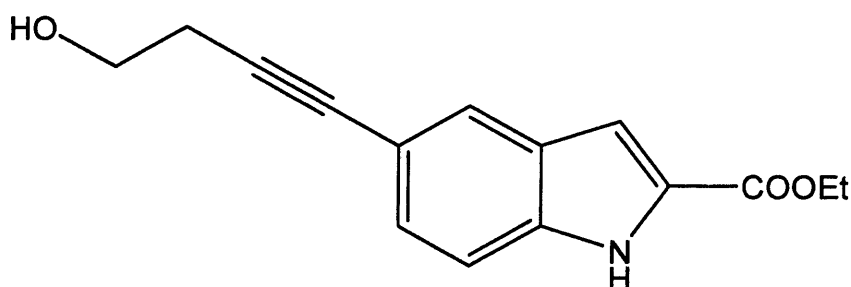
Synthesis of 5-(4-Hydroxy-but-1-ynyl)-1H-indole-2-carboxylic acid ethyl ester

To a solution of anhydrous THF (5 mL) and ethyl 5-iodo-1H-indole-2-carboxylate (100.00 mg, 0.32 mmol, 1.00 eq) was added PdCl$_2$ (22.3 mg, 0.03 mmol, 10%), Cu(I)I (12.00 mg, 0.06 mmol, 20%), DIPEA (5 mL) and 3-butyn-1-ol (222.43 mg, 0.24 mL, 3.17 mmol, 10 eq) was added. The reaction was left stirring four hours at room temperature under a nitrogen atmosphere. The reaction mixture was filtrated and washed with THF, concentrated *in vacuo* and then purified by column chromatography using AcOEt as eluent (12.80% yield).

^{1}H NMR (500mHz, CDCl$_3$) δ7.49 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.48 (d, J=7.5Hz, 1H), δ7.80 (d, J=1.5Hz and J=1.5Hz, 1H), δ8.31 (s, 1H), δ7.14 (d, J=1.5Hz, 1H), δ2.56 (a) d, J=12.5Hz, J=7.1Hz and J=2.3Hz, 2H), δ2.58 (b) d, J=12.5Hz, J=2.2Hz and J=6.2Hz, 2H), δ4.11 (a) d, J=7.1Hz, J=2.2Hz and J=11.5Hz, 2H), δ4.12 (b) d, J=2.3Hz, J=6.2Hz and J=11.5Hz, 2H), δ1.41 (s, 1H), δ4.43 (a) d, J=0.9Hz, t, J=6.0Hz, 2H), δ4.31 (b) d, J=0.9Hz, t, J=6.0Hz, 2H), and δ1.25 (t, 3H) ppm.

Synthesis of 5-Ethoxycarbonylethynyl-1H-indole-2-carboxylic acid ethyl ester

To a solution of ethyl propiolate (62.26 mg, 0.06mL, 0.63 mmol, 2eq) in THF (1.58 mL) was added 2.5M ButLi in hexane (0.25 mL, 2 eq) at -78°C. After stirring for 30 minutes at -78°C, a solution of anhydrous zinc chloride (260.2 mg, 1.90 mmol, 6 eq) in THF (2.0 mL) was added. The mixture was allowed to warm at room temperature and stirred for 1 hour. To the ice-cooled solution were added ethyl-5-iodo- 1H-indole-carboxylate (100 mg, 0.32 mmol, 1 eq) and $PdCl_2$ (11.10 mg, 0.01 mmol, 0.05 eq, 10%). After stirring at 50°C for 5 hours, the mixture was filtered through *celite* washing with THF. The filtrate was extracted with Et_2O, the combined extract dried ($MgSO_4$) and evaporated, and the residue purified by column chromatography (hexane/AcOEt 3:1) providing a brown oil (10.15% yield).[I]

^{1}H NMR (500mHz, $CDCl_3$) δ7.58 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.50 (d, J=7.5Hz, 1H), δ7.92 (d, J=1.5Hz and J=1.5Hz, 1H), δ8.04 (s, 1H), δ6.96 (d, J=1.5Hz, 1H), δ4.15 (a) d, J=3.3Hz, t, J=6.0Hz, 2H), δ4.17 (b) d, J=3.3Hz, t, J=6.0Hz, 2H), δ1.13 (t, J=6.0Hz, 3H), δ4.21 (a) and b) d, J=0.6Hz, t, J=6.0Hz, 2H) and δ1.30 (t, J=6.0Hz, 3H) ppm.

I. Yi Li and Manfred Hesse, The synthesis of cyclic spermine alkaloids: analogues of buchnerine and budmunchamine C, *Helvetica Chimica Acta*, 2003, Vol.86, pp.310-323.
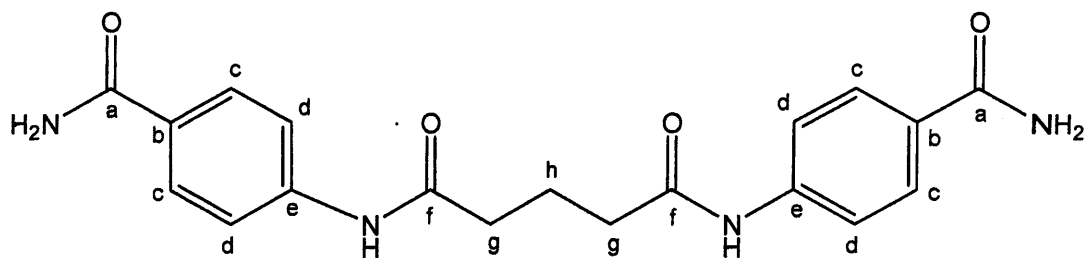
Synthesis of Pentanedioic acid bis-[(4-carbamoyl-phenyl)-amide]

To a solution of 4-aminobenzamide (500 mg, 3.7 mmol, 2 eq) in THF (20 mL) and *N,N*-diisopropylethylamine (0.67 mL, 3.8 mmol, 2.1 eq) was added glutaryl dichloride (0.24 mL, 1.9 mmol, 1 eq).

The solution was stirred at room temperature for 22 hours before the resulting semi-solid was filtered and washed with THF. The solution collected was concentrated *in vacuo* to leave a brown oil. Purification of this was attempted through column chromatography using a stationary phase of silica and an eluent of ethyl acetate. The compound collected was precipitated and further purified by washing with petroleum ether and finally diethylether (39.67% yield).

^{1}H NMR (500mHz, CDCl$_3$) δ6.86 (s, 1H), δ2.69 (a) d, J=12.5Hz, J=8.3Hz, J=3.7Hz and J=0.6Hz, 2H), δ2.73 (b) d, J=12.5Hz, J=3.7Hz, J=7.1Hz, J=0.7Hz and J=0.6Hz, 2H), δ2.09 (a) d, J=8.3Hz, J=3.7Hz, J=12.5Hz, J=8.6Hz and J=1.7Hz, 2H), δ2.13 (b) d, J=3.7Hz, J=7.1Hz, J=12.5Hz, J=1.7Hz and J=9.1Hz, 2H), δ2.61 (a) d, J=0.6Hz, J=0.7Hz, J=8.6Hz, J=1.7Hz and J=12.5Hz, 2H), δ2.56 (b) d, J=0.6Hz, J=1.7Hz, J=9.1Hz and J=12.5Hz, 2H), δ7.52 (s, 1H), δ7.11 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.61 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.61 (d, J=1.5Hz and J=7.5Hz, 1H), δ7.11 (d, J=1.5Hz and J=7.5Hz, 1H), δ7.05 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.51 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.51 (d, J=1.5Hz and J=7.5Hz, 1H), δ7.05 (d, J=1.5Hz and J=7.5Hz, 1H), δ8.90 (a) s, 2H), δ6.05 (b) s, 2H), δ7.02 (a) s, 2H) and δ6.09 (b) s, 2H) ppm.

^{13}C NMR (CDCl3) δ: 176.05 (Cf=O), 172.29 (Ca=O), 139.27 (Ce, Ar), 133.12 (Cc, Ar), 129.77 (CdH, Ar), 118.08 (CbH, Ar), 50.34 (CgH2), 32.94 (ChH2).

Synthesis of *N, N'*- bis -(4-methyl ester formoyl-phenyl)-succinamide

To a solution of methyl 4-aminobenzoate (500mg, 3.3mmol, 2eq) in anhydrous THF (20mL) and *N,N* diisopropylethylamine (0.60 mL, 3.4 mmol, 2.1eq) was added succinoyl dichloride (0.18mL, 1.6mmol, 1eq).

The reaction was left stirring overnight at room temperature under a nitrogen atmosphere; the solution collected was concentrated *in vacuo* to leave a brown solid that was filtered and washed with THF. Water was added to the reaction mixture and then extracted with ethyl acetate. The organic layer was dried on $MgSO_4$, filtered and the solvent evaporated.

Purification was carried out through column chromatography using a stationary phase of silica and ethyl acetate as eluent. The compound collected was further purified by crystallization with absolute MeOH (44.65% yield).

^1H NMR (500mHz, $CDCl_3$) δ10.38 (s, 1H), δ3.08 (a) d, J=12.5Hz, J=8.9Hz and J=1.1Hz, 2H), δ2.98 (b) d, J=12.5Hz, J=1.1Hz and J=8.6Hz, 2H), δ2.90 (a) d, J=8.9Hz, J=1.1Hz and J=12.5Hz, 2H), δ2.99 (b) d, J=1.1Hz, J=8.6Hz and J=12.5Hz, 2H), δ9.44 (s, 1H), δ7.59 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.71 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.71 (d, J=1.5Hz and J=7.5Hz, 1H), δ7.59 (d, J=1.5Hz and J=7.5Hz, 1H), δ7.66 (d, J=7.5Hz and J=1.5Hz, 1H), δ8.01 (d, J=7.5Hz and J=1.5Hz, 1H), δ8.01 (d, J=1.5Hz and J=7.5Hz, 1H), δ7.66 (d, J=1.5Hz and J=7.5Hz, 1H), δ3.78 (s, 3H) and δ3.91 (s, 3H) ppm.

^{13}C NMR (CDCl3) δ: 165.9 (Cf=O), 174 (Ca=O), 142.3 (Ce, Ar), 124.7 (Cc, Ar), 130.9 (CdH, Ar), 120.1 (CbH, Ar), 32.1 (CgH2).

Synthesis of Pentanedioic acid bis-[(4-methyl ester formoyl-phenyl)-amide]

To a solution of methyl 4-aminobenzoate (500mg, 3.3mmol, 2eq) in anhydrous THF (20mL) and *N,N* diisopropylethylamine (0.60 mL, 3.4 mmol, 2.1eq) was added glutaryl dichloride (0.21mL, 1.6mmol, 1eq).

The reaction was left stirring overnight at room temperature under a nitrogen atmosphere; the solution collected was concentrated *in vacuo* to leave a brown solid that was filtered and washed with THF. Water was added to the reaction mixture and then extracted with ethyl acetate. The organic layer was dried on MgSO$_4$, filtered and the solvent evaporated.

Purification was carried out through column chromatography using a stationary phase of silica and ethyl acetate as eluent. The collected compound was evaporated *in vacuo* affording a pure white solid similar to cotton without needing of further purification (42.03% yield).

^{1}H NMR (500mHz, CDCl$_3$) δ6.82 (s, 1H), δ2.64 (a) d, J=12.5Hz, J=11.4Hz and J=3.2Hz, 2H), δ2.71 (b) d, J=12.5Hz, J=3.5Hz, J=4.5Hz, J=0.7 and J=1.0Hz, 2H), δ2.15 (a) d, J=11.4Hz, J=3.5Hz, J=12.5Hz, J=11.4Hz and J=3.7Hz, 2H), δ2.18 (b) d, J=3.2Hz, J=4.5Hz, J=12.5Hz, J=1.5Hz and J=6.4Hz, 2H), δ2.55 (a) d, J=0.7Hz, J=11.4Hz, J=1.5Hz and J=12.5, 2H), δ2.58 (b) d, J=1.0Hz, J=1.7Hz, J=6.4Hz and J=12.5, 2H), δ7.41 (s, 1H), δ7.23 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.74 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.74 (d, J=1.5Hz and J=7.5Hz, 1H), δ7.23 (d, J=1.5Hz and J=7.5Hz, 1H), δ7.01 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.48 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.48 (d, J=1.5Hz and J=7.5Hz, 1H), δ7.01 (d, J=1.5Hz and J=7.5Hz, 1H), δ3.69 (s, 3H) and δ3.82 (s, 3H) ppm.

^{13}C NMR (CDCl3) δ: 165.9 (Cf=O), 173.9 (Ca=O), 142.3 (Ce, Ar), 124.7 (Cc, Ar), 130.9 (CdH, Ar), 120.1 (CbH, Ar), 38.1 (CgH2), 22.4 (ChH2).

Synthesis of Hexanedioic acid bis-[(4-methyl ester formoyl-phenyl)-amide]

To a solution of methyl 4-aminobenzoatè (500mg, 3.3mmol, 2eq) in anhydrous THF (20mL) and *N,N* diisopropylethylamine (0.60 mL, 3.4 mmol, 2.1eq) was added adipoyl chloride (0.24mL, 1.6mmol, 1eq).

The reaction was left stirring overnight at room temperature under a nitrogen atmosphere; the solution collected was concentrated *in vacuo* to leave a brown solid that was filtered and washed with THF. Water was added to the reaction mixture and then extracted with ethyl acetate. The organic layer was dried on MgSO4, filtered and the solvent evaporated.

Purification was carried out through column chromatography using a stationary phase of silica and ethyl acetate as eluent. The compound collected was further purified by crystallization with absolute MeOH (24.56% yield).

^{1}H NMR (500mHz, CDCl3) δ7.02 (s, 1H), δ2.70 (a) d, J=12.5Hz, J=12.4Hz, J=1.0Hz and J=0.8Hz, 2H), δ2.10 (b) d, J=12.5Hz, J=1.2Hz, J=6.2Hz and J=1.1, 2H), δ2.25 (a) d, J=12.4Hz, J=1.2Hz, J=12.5Hz, J=13.2Hz and J=2.4Hz, 2H), δ1.82 (b) d, J=1.0Hz, J=6.2Hz, J=12.5Hz, J=2.4Hz, J=4.3Hz, J=0.8Hz and J=1.1Hz, 2H), δ2.25 (a) d, J=13.2Hz, J=2.4Hz, J=12.5Hz, J=12.4 and J=1.2Hz, 2H), δ1.82 (b) d, J=0.8Hz, J=1.1Hz, J=2.4Hz, J=4.3Hz, J=12.5Hz, J=1.0Hz and J=6.2Hz, 2H), δ2.69 (a) d, J=0.8Hz, J=12.4Hz, J=1.0Hz and J=12.5Hz, 2H), δ2.09 (b) d, J=1.1Hz, J=1.2Hz, J=6.2Hz and J=12.5Hz, 2H), δ7.05 (s, 1H), δ7.11 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.70 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.70 (d, J=1.5Hz and J=7.5Hz, 1H), δ7.05 (d, J=1.5Hz and J=7.5Hz, 1H), δ7.06 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.46 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.46 (d, J=1.5Hz and J=7.5Hz, 1H), δ7.06 (d, J=1.5Hz and J=7.5Hz, 1H), δ3.75 (s, 3H) and δ3.87 (s, 3H) ppm.

^{13}C NMR (CDCl3) δ: 165.9 (Cf=O), 172.4 (Ca=O), 142.3 (Ce, Ar), 124.7 (Cc, Ar), 130.9 (CdH, Ar), 120.1 (CbH, Ar), 38.2 (CgH2), 26.6 (ChH2).
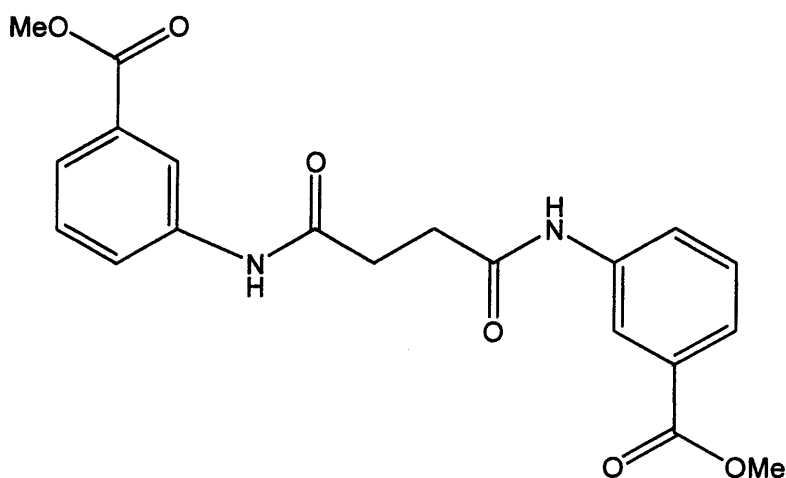
Synthesis of _N, N'_- bis -(3-methyl ester formoyl-phenyl)-succinamide

To a solution of methyl 3-aminobenzoate (500mg, 2.6mmol, 2eq) in anhydrous THF (20mL) and _N,N_ diisopropylethylamine (0.49mL, 2.8mmol, 2.1eq) was added succinoyl dichloride (0.15mL, 1.3mmol, 1eq). The reaction was left stirring for two days at room temperature under a nitrogen atmosphere; the second day further 1.50mL of _N,N_ diisopropylethylamine was added to boost the reaction to completion; the solution collected was concentrated _in vacuo,_ then lightly warmed with absolute MeOH and filtrated (26.98% yield).

^{1}H NMR (500mHz, CDCl$_3$) δ11.34 (s, 1H), δ3.10 (a) d, J=12.5Hz, J=8.9Hz and J=1.2Hz, 2H), δ3.20 (b) d, J=12.5Hz, J=1.2Hz and J=8.5Hz, 2H), δ2.93 (a) d, J=8.9Hz, J=1.2Hz and J=12.5Hz, 2H), δ2.84 (b) d, J=1.2Hz, J=8.5Hz and J=12.5Hz, 2H), δ8.53 (s, 1H), δ7.72 (d, J=1.5Hz and J=1.5Hz, 1H), δ7.76 (d, J=1.5Hz, J=7.5Hz and J=1.5Hz, 1H), δ7.61 (d, J=7.5Hz and J=7.5Hz, 1H), δ8.08 (d, J=1.5Hz, J=1.5Hz and J=7.5Hz, 1H), δ7.95 (d, J=1.5Hz and J=1.5Hz, 1H), δ7.55 (d, J=1.5Hz, J=7.5Hz and J=1.5Hz, 1H), δ7.54 (d, J=7.5Hz and J=7.5Hz, 1H), δ7.29 (d, J=1.5Hz, J=1.5Hz and J=7.5Hz, 1H), δ3.91 (s, 3H) and δ3.90 (s, 3H) ppm.

^{13}C NMR (CDCl3) δ: 166.9 (Cf=O), 174.0 (Ca=O), 137.2 (Ce, Ar), 132.6 (Cc, Ar), 124.8 (CdH, Ar), 129.6 (CbH, Ar), 124.8 (CgH, Ar), 120.8 (ChH, Ar), 32.1 (CiH2).
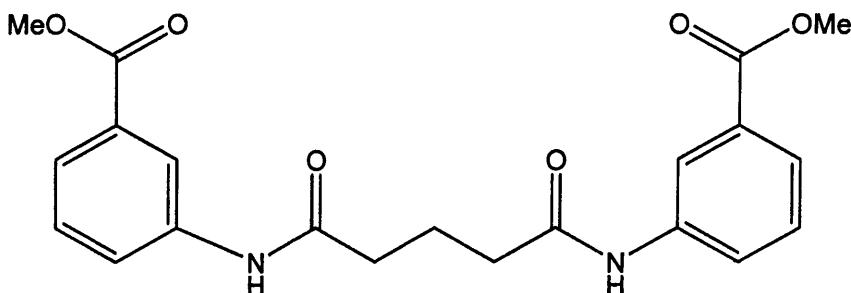
Synthesis of Pentanedioic acid bis-[(3-methyl ester formoyl-phenyl)-amide]

To a solution of methyl 3-aminobenzoate (500mg, 2.6mmol, 2eq) in anhydrous THF (20mL) and *N,N* diisopropylethylamine (0.49mL, 2.8mmol, 2.1eq) was added glutaryl dichloride (0.15mL, 1.3mmol, 1eq). The reaction was left stirring for two days at room temperature under a nitrogen atmosphere; the second day further 2.00mL of *N,N* diisopropylethylamine was added to boost the reaction to completion; the solution collected was concentrated *in vacuo*, then lightly warmed with absolute MeOH and filtrated (4.27% yield).

^{1}H NMR (500mHz, CDCl$_3$) δ7.74 (s, 1H), δ2.52 (a) d, J=12.5Hz, J=8.7Hz, J=3.0Hz and J=0.6Hz, 2H), δ2.64 (b) d, J=12.5Hz, J=3.0Hz, J=7.5Hz and J=0.7Hz, 2H), δ2.17 (a) d, J=8.7Hz, J=3.0Hz, J=12.5Hz, J=9.4Hz and J=4.6Hz, 2H), δ2.11 (b) d, J=3.0Hz, J=7.5Hz, J=12.5Hz, J=4.4Hz and J=5.1Hz, 2H), δ2.56 (a) d, J=9.4Hz, J=4.4Hz and J=12.5Hz, 2H), δ2.74 (b) d, J=0.6Hz, J=0.7Hz, J=4.6Hz, J=5.1Hz and J=12.5Hz, 2H), δ8.24 (s, 1H), δ7.55 (d, J=7.5Hz, J=1.5Hz and J=1.5Hz, 1H), δ7.47 (d, J=7.5Hz and J=7.5Hz, 1H), δ7.67 (d, J=1.5Hz , J=7.5Hz and J=1.5Hz, 1H), δ7.17 (d, J=1.5Hz and J=1.5Hz, 1H), δ6.97 (d, J=1.5Hz and J=1.5Hz, 1H), δ7.39 (d, J=1.5Hz, J=7.5Hz and J=1.5Hz, 1H), δ7.43 (d, J=7.5Hz and J=7.5Hz, 1H), δ7.68 (d, J=1.5Hz, J=1.5Hz and J=7.5Hz, 1H) δ3.82 (s, 3H) and δ3.88 (s, 3H) ppm.

^{13}C NMR (CDCl3) δ: 166.9 (Cf=O), 182.6 (Ca=O), 137.2 (Ce, Ar), 132.6 (Cc, Ar), 124.8 (CdH, Ar), 129.6 (CbH, Ar), 124.8 (CgH, Ar), 120.8 (ChH, Ar), 38.1 (CiH2), 22.4 (CjH2).
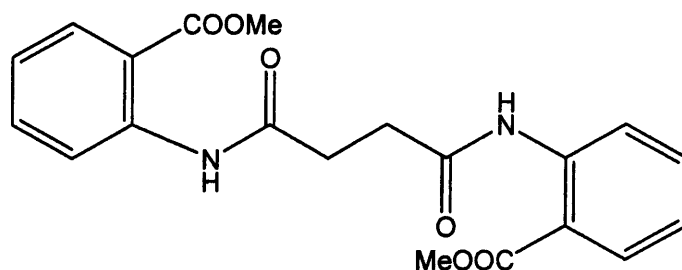
Synthesis of *N, N'*- bis -(2-methyl ester formoyl-phenyl)-succinamide

To a solution of methyl 2-aminobenzoate (0.43mL, 3.3mmol, 2eq) in anhydrous THF (20mL) and *N,N* diisopropylethylamine (0.60 mL, 3.4 mmol, 2.1eq) was added succinoyl dichloride (0.18mL, 1.6mmol, 1eq).

The reaction was left stirring overnight at room temperature under a nitrogen atmosphere; the green solution collected was concentrated *in vacuo* to leave a green solid that was crystallized with absolute MeOH (27.28% yield).

^{1}H NMR (500mHz, CDCl$_3$) δ11.11 (s, 1H), δ3.16 (a) d, J=12.5Hz, J=7.2Hz and J=1.5Hz, 2H), δ3.27 (b) d, J=12.5Hz, J=1.5Hz and J=10.0Hz, 2H), δ2.88 (a) d, J=7.2Hz, J=1.5Hz and J=12.5Hz, 2H), δ3.02 (b) d, J=1.5Hz, J=10.0Hz and J=12.5Hz, 2H), δ11.33 (s, 1H), δ7.68 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.36 (d, J=7.5Hz, J=7.5Hz and J=1.5Hz, 1H), δ7.56 (d, J=1.5Hz, J=7.5Hz and J=7.5Hz, 1H), δ8.28 (d, J=1.5Hz and J=7.5Hz, 1H), δ7.63 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.37 (d, J=7.5Hz, J=7.5Hz and J=1.5Hz, 1H), δ7.59 (d, J=1.5Hz, J=7.5Hz and J=7.5Hz, 1H), δ7.77 (d, J=1.5Hz and J=7.5Hz, 1H), δ4.04 (s, 3H) and δ3.78 (s, 3H) ppm.

^{13}C NMR (CDCl3) δ: 169.1 (Cf=O), 175.0 (Ca=O), 140.3 (Ce, Ar), 117.6 (Cc, Ar), 130.5 (CdH, Ar), 120.9 (CbH, Ar), 135.6 (CgH, Ar), 120.8 (ChH, Ar), 32.1 (CiH2).
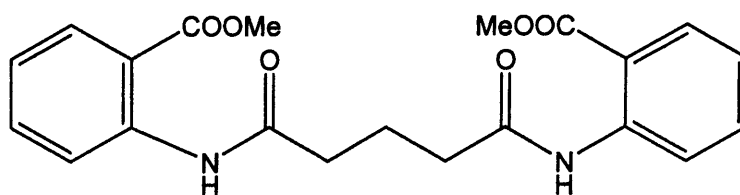
Synthesis of Pentanedioic acid bis-[(2-methyl ester formoyl-phenyl)-amide]

To a solution of methyl 2-aminobenzoate (0.43mL, 3.3mmol, 2eq) in anhydrous THF (20mL) and *N,N* diisopropylethylamine (0.60mL, 3.4mmol, 2.1eq) was added glutaryl dichloride (0.21mL, 1.6mmol, 1eq).

The reaction was left stirring overnight at room temperature under a nitrogen atmosphere; the green solution collected was concentrated *in vacuo* to leave a green solid that was crystallized with absolute MeOH (47.27% yield).

^{1}H NMR (500mHz, CDCl$_3$) δ9.78 (s, 1H), δ2.52 (a) d, J=12.5Hz, J=11.4Hz, J=1.0Hz and J=0.9Hz, 2H), δ2.42 (b) d, J=12.5Hz, J=1.1Hz, J=7.1Hz and J=1.1Hz, 2H), δ2.32 (a) d, J=11.4Hz, J=1.1Hz, J=12.5Hz, J=12.4Hz and J=4.0Hz, 2H), δ2.14 (b) d, J=1.0Hz, J=7.1Hz, J=12.5Hz, J=3.5Hz and J=2.9Hz, 2H), δ2.63 (a) d, J=12.4Hz, J=3.5Hz and J=12.5Hz, 2H), δ2.67 (b) d, J=0.9Hz, J=1.1Hz, J=4.0Hz, J=2.9Hz and J=12.5Hz, 2H), δ9.85 (s, 1H), δ7.53 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.45 (d, J=7.5Hz, J=7.5Hz and J=1.5Hz, 1H), δ7.20 (d, J=1.5Hz , J=7.5Hz and J=7.5Hz, 1H), δ7.42 (d, J=1.5Hz and J=7.5Hz, 1H), δ7.67 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.33 (d, J=7.5Hz, J=7.5Hz and J=1.5Hz, 1H), δ7.48 (d, J=1.5Hz, J=7.5Hz and J=7.5Hz, 1H), δ7.79 (d, J=1.5Hz and J=7.5Hz, 1H), δ4.14 (s, 3H) and δ3.99 (s, 3H) ppm.

^{13}C NMR (CDCl3) δ: 169.1 (Cf=O), 175.1 (Ca=O), 140.3 (Ce, Ar), 119.6 (Cc, Ar), 130.5 (CdH, Ar), 120.9 (CbH, Ar), 135.6 (CgH, Ar), 120.8 (ChH, Ar), 37.5 (CiH2), 21.3 (ClH2).
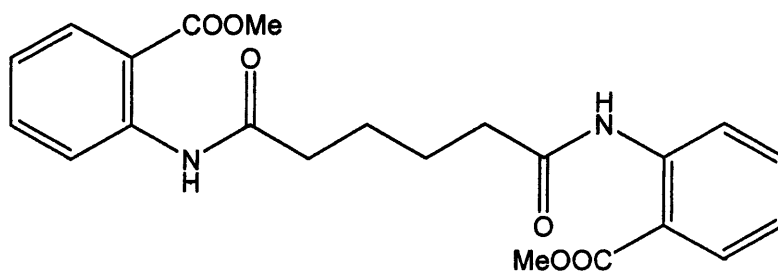
Synthesis of Hexanedioic acid bis-[(2-methyl ester formoyl-phenyl)-amide]

To a solution of methyl 2-aminobenzoate (0.43mL, 3.3mmol, 2eq) in anhydrous THF (20mL) and *N,N* diisopropylethylamine (0.60mL, 3.4mmol, 2.1eq) was added adipoyl chloride (0.24mL, 1.6mmol, 1eq).

The reaction was left stirring overnight at room temperature under a nitrogen atmosphere; the green solution collected was concentrated *in vacuo* to leave a green solid that was crystallized with absolute MeOH (4.84% yield).

^{1}H NMR (500mHz, CDCl$_3$) δ7.75 (s, 1H), δ2.08 (a) d, J=12.5Hz, J=6.7Hz, J=1.1Hz, J=1.0Hz and J=0.5Hz, 2H), δ2.97 (b) d, J=12.5Hz, J=1.0Hz, J=12.0Hz and J=0.8Hz, 2H), δ1.88 (a) d, J=6.7Hz, J=1.0Hz, J=12.5Hz, J=4.7Hz, J=2.6Hz, J=1.1Hz and J=0.8Hz, 2H), δ2.24 (b) d, J=1.1Hz, J=12.0Hz, J=12.5Hz, J=2.6Hz, J=12.6Hz and J=0.5Hz, 2H), δ1.82 (a) d, J=1.0Hz, J=0.8Hz, J=4.7Hz, J=2.6Hz, J=12.5Hz, J=6.2Hz and J=1.0Hz, 2H), δ2.27 (b) d, J=0.5Hz, J=2.6Hz, J=12.6Hz, J=12.5Hz, J=1.2Hz and J=12.3Hz, 2H), δ2.18 (a) d, J=1.1Hz, J=0.5Hz, J=6.2Hz and J=12.5Hz, 2H), δ2.67 (b) d, J=0.8Hz, J=1.0Hz, J=12.3Hz and J=12.5Hz, 2H), δ10.33 (s, 1H), δ7.35 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.34 (d, J=7.5Hz, J=7.5Hz and J=1.5Hz, 1H), δ7.43 (d, J=1.5Hz , J=7.5Hz and J=7.5Hz, 1H), δ7.54 (d, J=1.5Hz and J=7.5Hz, 1H), δ7.90 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.34 (d, J=7.5Hz, J=7.5Hz and J=1.5Hz, 1H), δ7.45 (d, J=1.5Hz, J=7.5Hz and J=7.5Hz, 1H), δ7.60 (d, J=1.5Hz and J=7.5Hz, 1H), δ4.09 (s, 3H), and δ3.85 (s, 3H) ppm.

^{13}C NMR (CDCl3) δ: 169.1 (Cf=O), 175.1 (Ca=O), 140.3 (Ce, Ar), 117.6 (Cc, Ar), 130.5 (CdH, Ar), 120.9 (CbH, Ar), 135.6 (CgH, Ar), 120.8 (ChH, Ar), 36.5 (CiH2), 25.3 (ClH2).
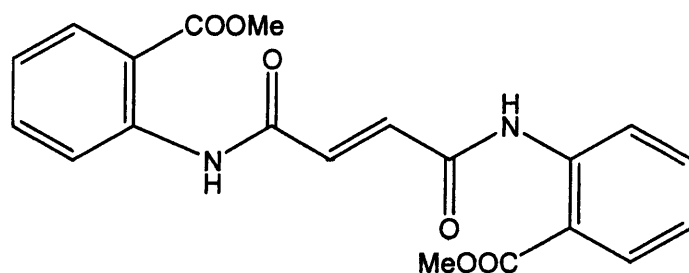
Synthesis of But-2-enedioic acid bis-[(2-methyl ester formoyl-phenyl)-amide]

To a solution of methyl 2-aminobenzoate (0.43mL, 3.3mmol, 2eq) in anhydrous THF (20mL) and *N,N* diisopropylethylamine (0.60mL, 3.4mmol, 2.1eq) was added fumaryl chloride (0.18mL, 1.6mmol, 1eq). A sudden change of colour (from a brown to a green solution) was observed and further anhydrous THF (15mL) was added to quench the reaction mixture.

The reaction was left stirring overnight at room temperature under a nitrogen atmosphere; the green solution collected was concentrated *in vacuo* to leave a green solid that was crystallized with absolute MeOH (85.17% yield).

^{1}H NMR (500mHz, CDCl$_3$) δ14.23 (s, 1H), δ6.94 (d, J=15.5Hz, 1H), δ6.81 (d, J=15.5Hz, 1H), δ13.46 (s, 1H), δ7.91 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.43 (d, J=7.5Hz, J=7.5Hz and J=1.5Hz, 1H), δ7.59 (d, J=1.5Hz , J=7.5Hz and J=7.5Hz, 1H), δ8.11 (d, J=1.5Hz and J=7.5Hz, 1H), δ7.69 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.34 (d, J=7.5Hz, J=7.5Hz and J=1.5Hz, 1H), δ7.53 (d, J=1.5Hz, J=7.5Hz and J=7.5Hz, 1H), δ8.05 (d, J=1.5Hz and J=7.5Hz, 1H), δ3.99 (s, 1H) and δ3.86 (s, 1H) ppm.

^{13}C NMR (CDCl3) δ: 169.1 (Cf=O), 1626.2 (Ca=O), 139.8 (Ce, Ar), 118.6 (Cc, Ar), 130.5 (CdH, Ar), 120.9 (CbH, Ar), 134.6 (CgH, Ar), 131.7 (ChH, Ar), 129.0 (CiH2).
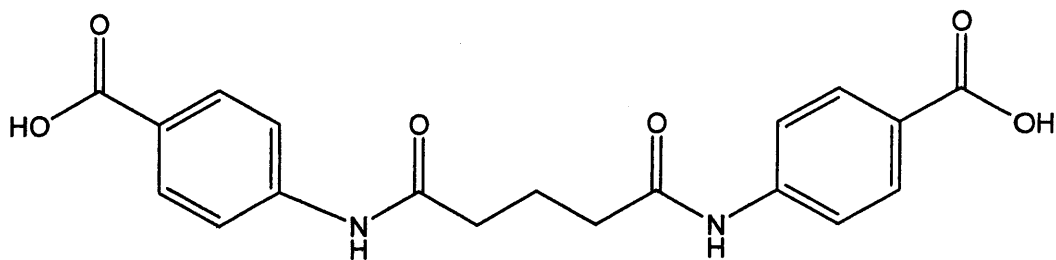
Synthesis of Pentanedioic acid bis-[(4-hydroxy formoyl-phenyl)-amide]

Sodium hydroxide monohydrate (11,5 mg, 0,29 mmol, 6eq) was added to a solution of pentanedioic acid bis-[(4-methyl ester formoyl-phenyl)-amide] (19.2 mg, 0.048 mmol, 1eq) in THF (0.50 mL) and water (0.50 mL). Then the reaction mixture was stirred at room temperature for four days. After dilution with water, the mixture was acidified with 2 N HCl until pH 2 was reached. The acid was extracted with ethyl acetate, washed with brine, and dried. Removal of the solvent by concentration *in vacuo* provided a pure compound (98.21% yield).

^{1}H NMR (500mHz, CDCl$_3$) δ7.15 (s, 1H), δ2.57 (a) d, J=12.5Hz, J=9.6Hz, J=3.0Hz and J=0.5Hz, 2H), δ2.66 (b) d, J=12.5Hz, J=3.2Hz, J=6.5Hz, J=0.9Hz and J=0.6Hz, 2H), δ2.30 (a) d, J=9.6Hz, J=3.2Hz, J=12.5Hz, J=7.5Hz and J=2.4Hz, 2H), δ2.12 (b) d, J=3.0Hz, J=6.5Hz, J=12.5Hz, J=2.5Hz and J=9.4Hz, 2H), δ2.62 (a) d, J=0.5Hz, J=0.9Hz and J=7.5Hz, J=2.5Hz and J=12.5Hz, 2H), δ2.54 (b) d, J=0.6Hz, J=2.4Hz, J=9.4Hz and J=12.5Hz, 2H), δ6.84 (s, 1H), δ7.02 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.55 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.55 (d, J=1.5Hz and J=7.5Hz, 1H), δ7.02 (d, J=1.5Hz and J=7.5Hz, 1H), δ7.09 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.14 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.14 (d, J=1.5Hz and J=7.5Hz, 1H), δ7.09 (d, J=1.5Hz and J=7.5Hz, 1H), δ20.42 (s, 3H), and δ16.42 (s, 3H) ppm.

^{13}C NMR (CDCl3) δ: 168.4 (Cf=O), 173.9 (Ca=O), 144.1 (Ce, Ar), 126.0 (Cc, Ar), 131.0 (CdH, Ar), 118.7 (CbH, Ar), 38.1 (CgH2), 22.4 (ChH2).
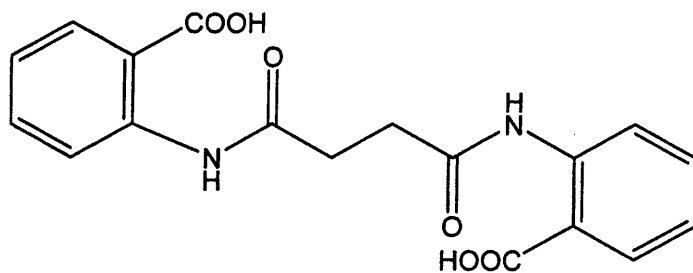
Synthesis of *N, N'*- bis -(2-hydroxy formoyl-phenyl)-succinamide

Sodium hydroxide monohydrate (60 mg, 1.48 mmol, 6eq) was added to a solution of *N, N'*- bis -(2-methyl ester formoyl-phenyl)-succinamide (95 mg, 0.25 mmol, 1eq) in THF (10 mL) and water (10 mL). Then the reaction mixture was stirred at room temperature for 48 hours. After dilution with water, the mixture was acidified with 1 N HCl until pH 2 was reached. The acid was extracted with ethyl acetate, washed with brine, and dried. Removal of the solvent by concentration *in vacuo* and precipitation from DCM to get rid of colour impurities provided a pure compound (31.13% yield).

^{1}H NMR (500mHz, CDCl$_3$) δ12.18 (s, 1H), δ2.88 (a) d, J=12.5Hz, J=11.1Hz and J=1.3Hz, 2H), δ3.03 (b) d, J=12.5Hz, J=1.3Hz and J=6.2Hz, 2H), δ3.25 (a) d, J=11.1Hz, J=1.3Hz and J=12.5Hz, 2H), δ3.02 (b) d, J=1.3Hz, J=6.2Hz and J=12.5Hz, 2H), δ11.02 (s, 1H), δ8.10 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.50 (d, J=7.5Hz, J=7.5Hz and J=1.5Hz, 1H), δ7.68 (d, J=1.5Hz, J=7.5Hz and J=7.5Hz, 1H), δ8.37 (d, J=1.5Hz and J=7.5Hz, 1H), δ7.01 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.38 (d, J=7.5Hz, J=7.5Hz and J=1.5Hz, 1H), δ7.38 (d, J=1.5Hz, J=7.5Hz and J=7.5Hz, 1H), δ6.63 (d, J=1.5Hz and J=7.5Hz, 1H), δ12.17 (s, 3H) and δ22.40 (s, 3H) ppm.

^{13}C NMR (CDCl3) δ: 171.4 (Cf=O), 175.0 (Ca=O), 141.7 (Ce, Ar), 124.0 (Cc, Ar), 132.1 (CdH, Ar), 120.3 (CbH, Ar), 135.6 (CgH, Ar), 120.8(ChH, Ar), 32.1 (CiH2).
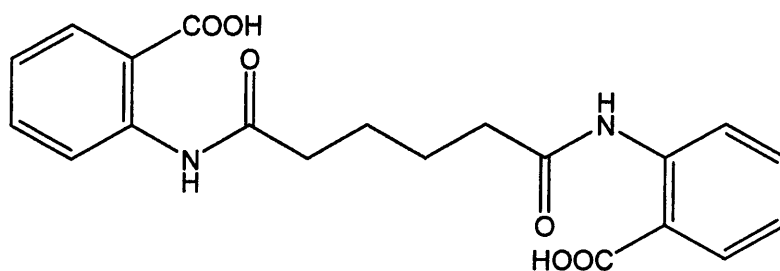
Synthesis of Hexanedioic acid bis-[(2-hydroxy formoyl-phenyl)-amide

Sodium hydroxide monohydrate (60 mg, 1.47 mmol, 6eq) was added to a solution of hexanedioic acid bis-[(2-methyl ester formoyl-phenyl)-amide] (101 mg, 0.245 mmol, 1eq) in THF (10 mL) and water (10 mL). Then the reaction mixture was stirred at room temperature for 48 hours. After dilution with water, the mixture was acidified with 1 N HCl until pH 2 was reached. The acid was extracted with ethyl acetate, washed with brine, and dried. Removal of the solvent by concentration *in vacuo* and precipitation from DCM to get rid of colour impurities provided a pure compound (32.17% yield).

^{1}H NMR (500mHz, CDCl$_3$) δ10.68 (s, 1H), δ2.26 (a) d, J=12.5Hz, J=6.2Hz, J=4.8Hz and J=0.6Hz, 2H), δ3.08 (b) d, J=12.5Hz, J=5.2Hz, J=8.0Hz and J=0.5Hz, 2H), δ1.55 (a) d, J=6.2Hz, J=5.2Hz, J=12.5Hz, J=9.0Hz, J=3.4Hz and J=0.5Hz, 2H), δ1.21 (b) d, J=4.8Hz, J=8.0Hz, J=12.5Hz, J=3.6Hz, J=7.1Hz and J=0.6Hz, 2H), δ1.26 (a) d, J=9.0Hz, J=3.6Hz, J=12.5Hz, J=9.9Hz and J=3.4Hz, 2H), δ1.82 (b) d, J=0.6Hz, J=0.5Hz, J=3.4Hz, J=7.1Hz, J=12.5Hz, J=3.2Hz and J=6.0Hz, 2H), δ2.75 (a) d, J=9.9Hz, J=3.2Hz and J=12.5Hz, 2H), δ2.73 (b) d, J=0.5Hz, J=0.6Hz, J=3.4Hz, J=6.0Hz and J=12.5Hz, 2H), δ8.08 (s, 1H), δ7.70 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.26 (d, J=7.5Hz, J=7.5Hz and J=1.5Hz, 1H), δ7.31 (d, J=1.5Hz , J=7.5Hz and J=7.5Hz, 1H), δ7.65 (d, J=1.5Hz and J=7.5Hz, 1H), δ7.22 (d, J=7.5Hz and J=1.5Hz, 1H), δ7.47 (d, J=7.5Hz, J=7.5Hz and J=1.5Hz, 1H), δ7.52 (d, J=1.5Hz, J=7.5Hz and J=7.5Hz, 1H), δ6.75 (d, J=1.5Hz and J=7.5Hz, 1H), δ22.15 (s, 3H) and δ23.27 (s, 3H) ppm.

^{13}C NMR (CDCl3) δ: 171.4 (Cf=O), 175.1 (Ca=O), 142.4 (Ce, Ar), 124.0 (Cc, Ar), 132.1 (CdH, Ar), 120.3 (CbH, Ar), 135.6 (CgH, Ar), 120.8(ChH, Ar), 36.5 (CiH2), 25.3 (ClH2).
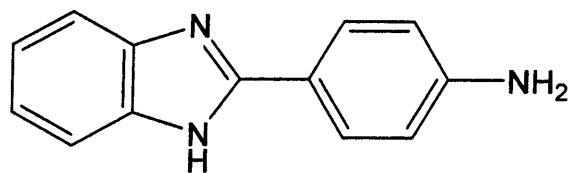
Synthesis of 4-(1*H*-benzo[d] imidazol-2-yl)benzenamine

A mixture of *o*-phenylenediamine (2.0 g, 18.5 mmol, 1eq) and aminobenzoic acid (2.43 g, 17.72 mmol, 1.06 eq) in polyphosphoric acid (27 g, 13.5eq) was heated with stirring in an oil bath at 180°C for 5 hours, cooled to room temperature and poured into 500 mL of water. The resulting green precipitate was filtered off, stirred with aqueous 10% sodium hydroxide (200 mL), and filtered off affording a pure compound (43.25% yield).[II]

II. M.F. Brana et al., Synthesis of benzimidazo-substituted 3-quinolinecarboxylic acids as antibacterial agents, *J.Heterocyclic Chem.*, 1990, Vol.27, p.1177-1180.
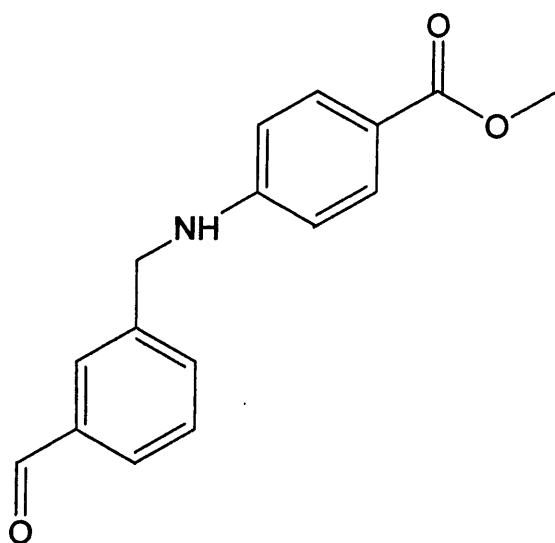
Synthesis of methyl 4-[(3-formylbenzyl)amino]benzoate

Isophthalaldehyde (100 mg, 0.746 mmol, 1eq) was added to a round-bottomed flask containing THF (5mL) which was being stirred under argon. Diethyl1,4-dihydro-2,6-dimethyl-3,5-pyridinedecarboxylate (189 mg, 0.746 mmol, 1eq) was then added to the same flask together with scandium triflate (7.3 mg, 0.0149 mmol, 0.02eq) as catalyst. Methyl-4-aminobenzoate (113 mg, 0.746 mmol, 1eq) was dissolved in 20 mL THF and added dropwise every 5 minutes over the first hour of the four-hour reaction. The reaction mixture was then extracted with ethyl acetate, washed initially with 5% Na_2CO_3 and subsequently with brine before drying over $MgSO_4$ and evaporating *in vacuo* to give a pale yellow solid. Purification of the final compound was carried out by column chromatography (50/50 DEE/CHCl$_3$) (12.34% yield).

^{1}H NMR (500mHz, CDCl$_3$) δ7.61 (d, J=7.5Hz, J=1.5Hz and J=1.5Hz, 1H), δ7.56 (d, J=7.5Hz and J=7.5Hz 1H), δ7.77 (d, J=1.5Hz , J=7.5Hz and J=1.5Hz, 1H), δ7.83 (d, J=1.5Hz and J=1.5Hz, 1H), δ4.66 (a) d, J=12.5Hz, 2H), δ4.65 b) d, J=12.5Hz, 2H), δ3.87 (s, 1H), δ7.76 (d, J=1.5Hz and J=7.5Hz, 1H), δ7.76 (d, J=1.5Hz and J=7.5Hz, 1H), δ6.93 (d, J=7.5Hz and J=1.5Hz, 1H), δ6.93 (d, J=7.5Hz and J=1.5Hz, 1H), δ9.89 (s, 1H) and δ3.92 (s, 3H) ppm.

^{13}C NMR (CDCl3) δ: 165.9 (Cf=O), 190.7 (Ca=O), 141.4 (Ce, Ar), 137.0 (Cc, Ar), 134.3 (CdH, Ar), 128.5 (CbH, Ar), 130.2 (CgH, Ar), 127.2 (ChH, Ar), 153.5 (CiH, Ar), 124.7 (CjH, Ar), 132.1 (CkH, Ar), 132.1(ClH, Ar), 111.4 (CmH, Ar), 111.4 (CnH, Ar), 47.9 (CoH2), 51.9 (CpH3).

References

References

Abstract

1. N. Claude Cohen, *Guidebook on Molecular Modelling in Drug Design*, Academic Press.
2. H.- D. Holtje, W. Sippl, D. Rognan, G. Folkers, *Molecular modelling – basic principles and applications*, 2nd Edition
3. Burger. *Molecular Modelling in Drug Design. Medicinal Chemistry & Drug Design*, 2003, Vol.1, 79, 123, pp. 113-115.
4. Andrew Leach. *The use of molecular modelling and chemoinformatics to discover and design new molecules. Molecular Modelling – principles and applications*. Second edition. Chapter 12, pp. 641-719.
5. M.A.Jordan and L.Wilson, Microtubules as target for anticancer drugs, *Nature*, April 2004, Vol.4, pp. 253-265.
6. 6) G. De Martino, G. La Regina, A. Coluccia, M.C. Edler, **M. Chiara Barbera**, A. Brancale, E. Wilcox, E. Hamel, M. Artico, and R. Silvestri, Arylthioindoles, potent inhibitors of tubulin polymerization, *Journal of Medicinal Chemistry*, **2004**, 47, 6120-6123.
7. 7) G. De Martino, M.C. Edler, G. La Regina, A. Coluccia, **M. Chiara Barbera**, D. Barrow, R. I. Nicholson, G. Chiosis, A. Brancale, E. Hamel, M. Artico, and Romano Silvestri, New Arylthioindoles: Potent Inhibitors of Tubulin Polymerization. 2. Structure-Activity Relationships and Molecular Modeling Studies, *J. Med. Chem.*, ASAP Article 10.1021/jm050809s S0022-2623(05)00809-5 (**Web Release Date:** January 19, 2006).
8. Francis V. Chiari, Unscrambling hepatitis C virus–host interactions, *Nature Insight*, Vol.**436**, pp.930-932, 18 August 2005.
9. G. Krauss, *Biochemistry of signal transduction and regulation*, 3rd Edition, Wiley-Vch.
10. John K. Heath, *Principles of cell proliferation*.

Chapter 1

1. G.Patrick, *Instant notes – Medicinal Chemistry*.
2. D.A. Williams and T.L. Lemke, *Foye's principle of medicinal chemistry*, 5th Edition.
3. F.D. King, *Medicinal chemistry – principles and practice*, 2nd Edition.
4. G. Thomas, *Fundamentals of medicinal chemistry*, 2004.
5. M.A. Jordan and L. Wilson, Microtubules as a target for anticancer drugs, *Nature Reviews, Cancer*, Vol.4, pp. 253-265, April 2004.
6. G. De Martino, G. La Regina, A. Coluccia, M.C. Edler, **M. Chiara Barbera**, A. Brancale, E. Wilcox, E. Hamel, M. Artico, and R. Silvestri, Arylthioindoles, potent inhibitors of tubulin polymerization, *Journal of Medicinal Chemistry*, **2004**, 47, 6120-6123.
7. G. De Martino, M.C. Edler, G. La Regina, A. Coluccia, **M. Chiara Barbera**, D. Barrow, R. I. Nicholson, G. Chiosis, A. Brancale, E. Hamel, M. Artico, and Romano Silvestri, New Arylthioindoles: Potent Inhibitors of Tubulin Polymerization. 2. Structure-Activity Relationships and Molecular Modeling Studies, *J. Med. Chem.*, ASAP Article 10.1021/jm050809s S0022-2623(05)00809-5 (**Web Release Date:** January 19, 2006).
8. Francis V. Chiari, Unscrambling hepatitis C virus–host interactions, *Nature Insight*, Vol.**436**, pp.930-932, 18 August 2005.
9. G. Krauss, *Biochemistry of signal transduction and regulation*, 3rd Edition, Wiley-Vch.
10. John K. Heath, *Principles of cell proliferation*.
11. P. Ehrlich, Uber den jetzigen standeer chemotherapie, *Chem. Ber.*, Vol.42,p. 17, 1909.
12. E. Fischer, Einfluss der configuartion auf die wirkung der enzyme, *Chem. Ber.*, Vol.27,p. 2985, 1894.
13. T. Fujita, The extrathermodynamic approach to drug design, *Comprehensive Medicinal Chemistry*, Vol.4, pp. 497-560, Pergamon, Elmsford, NY, 1990.
14. J.P. Snider, Computer assisted drug design. I. Conditions in the 1980s, *Med. Res. Rev.*, Vol.11(6), pp. 64-662, 1991.
15. U. Burkert and N.L. Allinger, *Molecular Mechanics*, ACS Monograph 177, American Chemical Society, Washington D.C. 1982.
16. W.H. Press et al., *Numerical recipes in C.*, Cambridge University Press, Cambridge 1988.
17. T. Schlick, *Optimization methods in computational chemistry*, Reviews in computational chemistry, Vol.3, K.B. Lipkowitz and D.B. Boyd (Eds.), VCH, New York, pp. 1-71, 1992.
18. E.L. Eliel et al., *Conformational analysis*, Wiley-Interscience, New York, 1965.
19. A.K. Ghose et al., *J.Med. Chem.*, Vol.32, pp. 746-756, 1989.
20. W.L. Jörgensen, *Science*, Vol.254, pp. 954-963, 1991.
21. N. Metropolis et al., *J. Chem. Phys.*, Vol.32, pp. 1087-1092, 1953.

22. R.A. Leach, *A survey of methods for searching the conformational space of small and medium-sized molecules*, Reviews in computational chemistry, Vol.2 K.B. Lipkowitz and D.B. Boyd (Eds.), VCH, New York, pp. 1-47, 1991.
23. C. Lemmen and T. Lengauer, *J. Comput-Aided Mol. Design*, Vol.11, pp.215-232, 2000.
24. J.S. Mason et al., *J. Curr. Pharm. Design*, Vol.7, pp. 567-597, 2001.
25. A.C. Good and J.S. Mason, *Three-dimensional structure database search*, Reviews in computational chemistry, Vol.7 K.B. Lipkowitz and D.B. Boyd (Eds.), VCH, New York, pp. 73-95, 1995.
26. Jaroslaw Polanski, Self-organizing neural network for modelling 3D QSAR of colchinoids, *Acta Biochimica Polonica*, 2000, Vol.47, No.1, pp. 37-45.
27. A.Bairoch and B. Boeckmann, *Nucleic Acids Res.*, Vol.22, pp.3578-3580, 1994.
28. F.C. Bernstein et al., *J. Mol. Biol.*, Vol.112, pp. 535-542, 1977.
29. M.S. Johnson et al., *Crit. rev. Biochem.Mol. Biol.*, Vol.29, pp. 193-316, 1994.
30. Sali et al., *TIBS.*, Vol.15, pp. 235-240, 1990.
31. T.A. Jones and S. Thirup, *EMBO J..*, Vol.5, pp. 819-822, 1986.
32. M.J. Dudek and H.A. Sheraga, *J. Comput. Chem.*, Vol.11, pp. 121-151, 1990.
33. R. Doolittle, *Methods Enzymol.*, Vol.183, pp.736-772, 1990.
34. S.H. Bryant, *Proteins Struc. Func. Gen.,*, Vol.26, pp. 172-185, 1996.
35. M. Wilmanns and D. Eisenberg, *Protein Eng.*, Vol.8, pp. 626-635, 1995.
36. Kolinski et al., *Proteins Struc. Func. Gen.,*Vol.44, pp. 133-149, 2001.
37. Clothia and A.M. Lesk, *EMBO J.*, Vol.5, pp. 823-826, 1986.
38. W.R. Parson, *Methods in Enzymology.*, Vol.183, pp. 63-98, 1990.
39. S.F. Altshult et al., *J. Mol. Biol.*, Vol.215, pp. 403-410, 1990.
40. M.F. Perutz et al., *J. Mol. Biol.*, Vol.13, pp. 669-678, 1965.
41. J. Devereux et al., *Nucleic Acids Res.*, Vol.12, pp. 387-395, 1984.
42. S. Mosimann et al., *Proteins Struc. Func. Gen.*, Vol.23, pp. 301-317, 1995.
43. J. Moult, *Curr. Opinion Biotechnology*, Vol.10, pp.583-588, 1999.

[i] Cummins, A.-C. Computer-Aided Design & Synthesis of Novel Antiviral Agents Thesis for Degree of Master of Pharmacy (MPharm honours). (2003) *Welsh School of Pharmacy Cardiff University*.

Chapter 2

1. M.A.Jordan and L.Wilson, Microtubules as target for anticancer drugs, *Nature*, April 2004, Vol.4, pp. 253-265.
2. K.Yamamoto et al.,Phase I study of E7010, *Cancer Chem. Pharmacol.*,1998, Vol.42, pp. 124-134.
3. G.Bacher et al., New small molecule tubulin inhibitors, *Pure Appl. Chem.*, 2001, Vol.73, pp.1459-1464.
4. G. De Martino, G. La Regina, A. Coluccia, M.C. Edler, **M. Chiara Barbera**, A. Brancale, E. Wilcox, E. Hamel, M. Artico, and R. Silvestri, Arylthioindoles, potent inhibitors of tubulin polymerization, *Journal of Medicinal Chemistry*, **2004**, 47, 6120-6123.
5. G. De Martino, M.C. Edler, G. La Regina, A. Coluccia, **M. Chiara Barbera**, D. Barrow, R. I. Nicholson, G. Chiosis, A. Brancale, E. Hamel, M. Artico, and Romano Silvestri, New Arylthioindoles: Potent Inhibitors of Tubulin Polymerization. 2. Structure-Activity Relationships and Molecular Modeling Studies, *J. Med. Chem.*, ASAP Article 10.1021/jm050809s S0022-2623(05)00809-5 **(Web Release Date:** January 19, 2006).
6. Luduena, R. F. Multiple forms of tubulin: different gene products and covalent modifications, *Int. Rev. Cytology*, Vol.178, pp. 207–275, 1998.
7. Verdier-Pinard, P. et al. Direct analysis of tubulin expression in cancer cell lines by electrospray ionization mass spectrometry, *Biochemistry*, Vol.42, pp.12019–12027, 2003.
8. Ligon, L. A., Shelly, S. S., Tokito, M. & Holzbaur, E. L. The microtubule plus-end proteins EB1 and dynactin have differential effects on microtubule polymerization, *Mol. Biol. Cell*, Vol.14, pp. 1405–1417, 2003.
9. Galmarini, C. M. et al. Drug resistance associated with loss of p53 involves extensive alterations in microtubule composition and dynamics, *Br. J. Cancer*, Vol. 88, pp. 1793–1799, 2003.
10. Giodini, A. et al. Regulation of microtubule stability and mitotic progression by surviving, *Cancer Res.*, Vol.62, pp. 2462–2467, 2002.
11. Cassimeris, L. The oncoprotein 18/stathmin family of microtubule destabilizers, *Curr. Opin. Cell Biol.*, Vol.14, 18–24, 2002.
12. Spittle, C., Charrasse, S., Larroque, C. & Cassimeris, L. The interaction of TOGp with microtubules and tubulin, *J. Biol.Chem.*, Vol.275, pp. 20748–20753, 2000.
13. Maney, T., Wagenbach, M. & Wordeman, L. Molecular dissection of the microtubule depolymerizing activity of mitotic centromere-associated kinesin, *J. Biol. Chem.*, Vol.276, pp. 34753–34758, 2001.
14. M. A. Jordan and L. Wilson, Microtubules as a target for anticancer drugs, *Nature Reviews, Cancer*, Vol.4, pp.253-265, April 2004.
15. Wordeman, L. & Mitchison, T. J. in *Microtubules* (eds Hyams, J. S. & Lloyd, C. W.), pp.287–302 (Wiley–Liss, New York, 1994).
16. Wilson, L. & Jordan, M. A. in *Microtubules* (eds Hyams, J. S. & Lloyd, C. W.), pp. 59–84 (Wiley–Liss, New York, 1994).
17. McIntosh, J. R. in *Microtubules* (eds Hyams, J. S. & Lloyd, C. W.), pp. 413–434 (Wiley–Liss, New York, 1994).

18. Waterman-Storer, C. & Salmon, E. D. Microtubule dynamics: treadmilling comes around again, *Curr. Biol.*, Vol. 7, pp. 369–372, 1997.
19. Mitchison, T. J. & Kirschner, M. Dynamic instability of microtubule growth, *Nature*, Vol.312, pp. 237–242, 1984.
20. Margolis, R. L. & Wilson, L. Opposite end assembly and disassembly of microtubules at steady state *in vitro*, *Cell*, Vol. 13, pp. 1–8, 1978.
21. Margolis, R. L. & Wilson, L. Microtubule treadmilling: what goes around comes around, *Bioessays*, Vol.20, pp. 830–836, 1998.
22. Rodionov, V. 1. & Borisy, G. G. Microtubule treadmilling *in vivo*, *Science*, Vol.275, pp. 215–218, 1997.
23. Shaw, S. L., Kamyar, R. & Ehrhardt, D. W. Sustained microtubule treadmilling in *Arabidopsis* cortical arrays, *Science*, Vol.300, pp. 1715–1718, 2003.
24. Panda, D., Miller, H. P. & Wilson, L. Rapid treadmilling of MAP-free brain microtubules *in vitro* and its suppression by tau, *Proc. Natl Acad. Sci. USA*, Vol.96, pp. 12459–12464, 1999.
25. Wilson, L., Panda, D. & Jordan, M. A. Modulation of microtubule dynamics by drugs: a paradigm for the actions of cellular regulators, *Cell Struct. Funct.*, Vol.24, pp. 329–335, 1999.
26. Wilson, L. & Jordan, M. A. Microtubule dynamics: taking aim at a moving target, *Chem. Biol.*, Vol.2, pp. 569–573, 1995.
27. Lodish, H. *et al. Molecular Cell Biology* (W. H. Freeman, NewYork, 1999).
28. Panda, D., Miller, H. & Wilson, L. Determination of the size and chemical nature of the stabilizing cap at microtubule ends using modulators of polymerization dynamics, *Biochemistry*, Vol.41, pp. 1609–1617, 2002.
29. Caplow, M. & Fee, L. Concerning the chemical nature of tubulin subunits that cap and stabilize microtubules, *Biochemistry*, Vol.42, pp. 2122–2126, 2003.
30. Wittmann, T., Bokoch, G. & Waterman-Storer, C. Regulation of leading edge microtubule and actin dynamics downstream of Rac1, *J. Cell Biol.*, Vol.161, pp. 845–851, 2003.
31. Alli, E., Bash-Babula, J., Yang, J.-M. & Hait, W. N. Effect of stathmin on the sensitivity to antimicrotubule drugs in human breast cancer, *Cancer Res.*, Vol.62, pp. 6864–6869, 2002.
32. McNally, F. Microtubule dynamics: new surprises from an old MAP, *Curr. Biol.*, Vol.13, pp. 597–599, 2003.
33. Ohi, R., Coughlin, M. L., Lane, W. S. & Mitchison, T. J. An inner centromere protein that stimulates the microtubule depolymerizing activity of a KinI kinesin, *Dev. Cell*, Vol.5, pp. 309–321, 2003.
34. Carvalho, P., Tirnauer, J. & Pellman, D. Surfing on microtubule ends, *Trends Cell Biol.*, Vol.13, pp. 229–237, 2003.
35. Hergovich, A., Lisztwan, J., Barry, R., Ballschmieter, P. & Krek, W. Regulation of microtubule stability by the von Hippel–Lindau tumour suppressor protein pVHL, *Nature Cell Biol.*, Vol.5, pp. 64–70, 2003.
36. Komarova, Y., Akhmanova, A., Kojima, S., Galjart, N. & Borisy, G. Cytoplasmic linker proteins promote microtubule rescue *in vivo*, *J. Cell Biol.*, Vol.159, pp. 589–599, 2002.
37. Mitchison, T. J. Microtubule dynamics and kinetochore function in mitosis, *Annu. Rev. Cell Biol.*, Vol.4, pp. 527–549, 1988.
38. Saxton, W. M. et al. Tubulin dynamics in cultured mammalian cells, *J. Cell Biol.*, Vol.99, pp. 2175–2186, 1984.
39. Rusan, N. M., Fagerstrom, C. J., Yvon, A.-M. C. & Wadsworth, P. Cell cycle-dependent changes in microtubule dynamics in living cells expressing green fluorescent protein–α tubulin, *Mol. Biol. Cell*, Vol.12, pp. 971–980, 2001.
40. Zhou, J. et al. Brominated derivatives of noscapine are potent microtubule-interfering agents that perturb mitosis and inhibit cell proliferation, *Mol. Pharmacol.*, Vol.63, pp. 799–807, 2003.
41. Hamel, E. & Covell, D. G. Antimitotic peptides and depsipeptides, *Curr. Med. Chem. Anti-Canc. Agents*, Vol.2, pp. 19–53, 2002.
42. Hoffman, J. C. & Vaughn, K. C. Mitotic disrupter herbicides act by a single mechanism but vary in efficacy, *Protoplasma*, Vol.179, pp. 16–25, 1994.
43. Lacey, E. & Gill, J. H. Biochemistry of benzimidazole resistance, *Acta Trop.*, Vol.56, pp. 245–262, 1994.
44. Lobert, S., Ingram, J. & Correia, J. Additivity of dilantin and vinblastine inhibitory effects on microtubule assembly, *Cancer Res.*, Vol.59, pp. 4816–4822, 1999.
45. Cann, J. R. & Hinman, N. D. Interaction of chlorpromazine with brain microtubule subunit protein, *Molec. Pharmacol.*, Vol.11, pp. 256–267, 1975.
46. Boder, G. B., Paul, D. C. & Williams, D. C. Chlorpromazine inhibits mitosis of mammalian cells, *Eur. J. Cell Biol.*, Vol.31, pp. 349–353, 1983.
47. Jordan, M. A. Mechanism of action of antitumor drugs that interact with microtubules and tubulin, *Curr. Med. Chem. Anti-Canc. Agents*, Vol.2, pp. 1–17, 2002.
48. Jimenez-Barbero, J., Amat-Guerri, F. & Snyder, J. P. The solid state, solution and tubulin-bound conformations of agents that promote microtubule stabilization, *Curr. Med. Chem. Anti-Canc. Agents*, Vol.2, pp. 91–122, 2002.
49. Johnson, I. S., Wright, H. F. & Svoboda, G. H. Experimental basis for clinical evaluation of anti-tumor principles derived from *Vinca rosea* Linn, *J. Lab. Clin. Med.*, Vol.54, pp. 830–837, 1959.
50. Noble, R. L., Beer, C. T. & Cutts, J. H. Further biological activities of vincaleukoblastine: an alkaloid isolated from *Vinca rosea* (L.), *Biochem. Pharmacol.*, Vol.1, pp. 347–348, 1958.
51. Gidding, C. E., Kellie, S. J., Kamps, W. A. & de Graaf, S. S. Vincristine revisited, *Crit. Rev. Oncol. Hematol.*, Vol.29, pp. 267–287, 1999.
52. Wani, M. C., Taylor, H. L., Wall, M. E., Coggon, P. & McPhail, A. T., *J. Am. Chem. Soc.*, Vol.93, pp. 2325–2327, 1971.
53. Schiff, P. B., Fant, J. & Horwitz, S. B. Promotion of microtubule assembly *in vitro* by taxol, *Nature*, Vol.277, pp. 665–667, 1979.
54. Horwitz, S. B. How to make taxol from scratch, *Nature*, Vol. 367, pp. 593–594, 1994.
55. Von Hoff, D. D. The taxoids: same roots, different drugs, *Semin. Oncol.*, Vol.24 (4 Suppl. 13), S13-3-S13-10 (1997).

56. Markman, M. Managing taxane toxicities, *Support Care Cancer*, Vol.11, pp. 144–147, 2003.

57. Nogales, E., Wolf, S. G., Khan, I. A., Luduena, R. F. & Downing, K. A. Structure of tubulin at 6. 5Å and location of the taxol-binding site, *Nature*, Vol.375, pp. 424–427, 1995.

58. Nogales, E. Structural insights into microtubule function, *Annu. Rev. Biophys. Biomol. Struct.*, Vol.30, pp. 397–420, 2001.

59. Pryor, D. E. et al. The microtubule stabilizing agent laulimalide does not bind in the taxoid site, kills cells resistant to paclitaxel and epothilones, and may not require its epoxide moiety for activity, *Biochemistry*, Vol.41, pp. 9109–9115, 2002.

60. Hastie, S. B. Interactions of colchicine with tubulin, *Pharmacol. Ther.*, Vol.512, pp. 377–401, 1991.

61. Skoufias, D. & Wilson, L. Mechanism of inhibition of microtubule polymerization by colchicine: inhibitory potencies of unliganded colchicine and tubulin–colchicines complexes, *Biochemistry*, Vol.31, pp. 738–746, 1992.

62. Tozer, G. M., Kanthou, C., Parkins, C. S. & Hill, S. A. The biology of the combretastatins as tumour vascular targeting agents, *Int. J. Exp. Pathol.*, Vol.83, pp. 21–38, 2002.

63. Prise, V. E., Honess, D. J., Stratford, M. R., Wilson, J. & Tozer, G. M. The vascular response of tumor and normal tissues in the rat to the vascular targeting agent, combretastatin A-4-phosphate, at clinically relevant doses, *Int. J. Oncol.*, Vol.21, pp. 717–726, 2002.

64. a) Ana B. S. Maya et al., Further naphthylcombretastatins. An investigation on the role of the naphthalene moiety, *J. Med. Chem.*, Vol.48, pp. 556-568, 2005. b) D. Simoni et al., Heterocyclic and phenyl double-bond-locked combretastatin analogues possessing potent apoptosis-inducing activity in HL60 and in MDR cell lines, *J. Med. Chem.*, Vol.48, pp. 723-736., 2005.

65. Dumontet, C. & Sikic, B. Mechanisms of action of and resistance to antitubulin agents: microtubule dynamics, drug transport, and cell death, *J. Clin. Oncol.*, Vol.17, pp. 1061–1070, 1999.

66. Ambudkar, S. V., Kimchi-Sarfaty, C., Sauna, Z. E. & Gottesman, M. M. P-glycoprotein: from genomics to mechanism, *Oncogene*, Vol.22, pp. 7468–7485, 2003.

67. Safa, A. R. Identification and characterization of the binding sites of P-glycoprotein for multidrug resistance-related drugs and modulators, *Curr. Med. Chem. Anti-Canc. Agents*, Vol.4, pp. 1–17, 2004.

68. Thomas, H. & Coley, H. M. Overcoming multidrug resistance in cancer: an update on the clinical strategy of inhibiting P-glycoprotein, *Cancer Control*, Vol.10, pp. 159–165, 2003.

69. Geney, R., Ungureanu, M., Li, D. & Ojima, I. Overcoming multidrug resistance in taxane chemotherapy, *Clin. Chem. Lab. Med.*, Vol.40, pp. 918–925, 2002.

70. Orr, G. A., Verdier-Pinard, P., McDaid, H. & Horwitz, S. B. Mechanisms of taxol resistance related to microtubules, *Oncogene*, Vol.22, pp. 7280–7295, 2003.

71. Kavallaris, M. et al. Multiple microtubule alterations are associated with *Vinca* alkaloid resistance in human leukaemia cells, *Cancer Res.*, Vol.61, pp. 5803–5809, 2001.

72. Minotti, A. M., Barlow, S. B. & Cabral, F. Resistance to antimitotic drugs in Chinese hamster ovary cells correlated with changes in the level of polymerized tubulin, *J. Biol. Chem.*, Vol.266, pp. 3987–3994, 1991.

73. James, S. W., Silflow, C. D., Stroom, P. & Lefebvre, P. A. A mutation in the α1-tubulin gene of *Chlamydomonas reinhardtii* confers resistance to anti-microtubule herbicides, *J. Cell Sci.*, Vol.106, pp. 209–218, 1993.

74. Lee, W.-P. Purification and characterization of tubulin from parental and vincristine-resistant HOB1 lymphoma cells, *Arch. Biochem. Biophys.*, Vol.319, pp. 498–503, 1995.

75. Ohta, S. et al. Characterization of a taxol-resistant human small-cell lung cancer cell line, *Jpn. J. Cancer Res.*, Vol.85, pp. 290–297, 1994.

76. Laing, N. M. et al. Amplification of the ATP-binding cassette 2 transporter gene is functionally linked with enhanced efflux of estramustine in ovarian carcinoma cells, *Cancer Res.*, Vol.58, pp. 1332–1337, 1998.

77. R. J Owellen et al., The binding of vincristine, vinblastine and colchicine to tubulin, *Biochem. Biophys. Res. Commun.*, Vol.47, pp. 685-691, 1972.

78. K. Pohle et al., On the tumor growth-inhibiting action of cochicine, *Arch. Geschwulstforsh.*, Vol.25, pp. 17-20,1965.

79. M. A Jordan, Mechanism of action of antitumor drugs that interact with microtubules and tubulin, *Curr. Med. Chem. – Anticancer Agents*, Vol.1, pp. 1-17, 2002.

80. P.B. Schiff et al., Promotion of microtubule assembly in vitro by taxol, *Nature*, Vol.277, pp. 665-667,1979.

81. G. Hoefle et al., Epothilone A and B – novel 16-membered macrolides with cytotoxic activity. Isolation, crystal structure and conformation in solution, *Angew. Chem. Int. Ed. Engl.*, Vol.35, pp. 1567-1569, 1996.

82. D. G. Corley et al., Laulimalide. New potent cytotoxic macrolides from a marine sponge and a nudibranch predator, *J. Org. Chem.*, Vol.53, pp. 3644-3646, 1998.

83. (a) E. Ter Haar et al., Discodermolide, cytotoxyc marine agent that stabilizes microtubules more potently than Taxol, *Biochemistry*, Vol.35, pp. 243-250, 1996. (b) I. Paterson et al., A pratical synthesisi of (+)-discodermolide and analogues: fragment union by complex aldol reactions, *J. Am. Chem. Soc.*, Vol.123, pp. 9535-9544, 2001.

84. E. Hamel and D.G. Covell, Antimitotic Peptides and Depsipeptides, *Curr. Med. Chem. – Anti-Cancer Ag.*, Vol.2, pp. 19-53, 2002.

85. T. Beckers and S. Mahboobi, Natural, semisynthetic and synthetic microtubule inhibitors for cancer therapy, and references therein, *Drugs Fut.*, Vol.28, pp. 767-785, 2003.

86. A. Duflos et al., Novel aspects of natural and modified Vinca alkaloids, *Curr. Med. Chem. – Anti-Cancer Agents*, Vol.2, pp. 55-70, 2002.

87. A. Kruczynski and B.T. Hill, Vinflumine, the latest vinca alkaloid in clinicla development. A review of its preclinical anticancer properties, *Crit. Rev. Oncol. Hematol.*, Vol.40, pp. 159-173, 2001.

88. J.A. Mcintyre and J. Castaner, Vinflumine: antimitotic vinca alkaloid, *Drugs Fut.*, Vol.29, pp. 574-580, 2004.

89. S. Zhao et al., Biological activity of Tryprostatins and their diastereoisomers on human carcinoma cell lines, *J. Med. Chem.*, Vol.45, pp. 1559-1562, 2002.

90. T. Usui et al., Tryprostatin A, a specific and novel inhibitor of microtubule assembly, *Biochemical J.*, Vol.**333**, pp. 543-563, 1998.

91. H. Morita et al., Antimitotic activity of moroidin, a bicyclic peptide from the seeds of *Celosia argentea, Bioorg. & Med. Chem. Lett.*, Vol.**10**, pp. 469-471, 2000.

92. J. Kobayashi et al., Celogentins A-C, new antimitotic bicyclic peptides from the seeds of *Celosia argentea, J. Org. Chem.*, Vol.**66**, pp. 6626-6633, 2001.

93. A. Zask et al., Synthesis and biologiacl activity of analogues of the antimicrotubule agent N,b,b-trimethyl-L-phenylalanyl-N^1-[(1S,2E)-3-carboxy-1-isopropylbut-2-enyl]N^1,3-dimethyl-L-valinamide (HTI-286), *J. Med. Chem.*, Vol.**47**, pp. 4774-4786, 2004.

94. N. Lindquist et al., *J. Am. Chem. Soc.*, Vol.**113**, pp. 2303-2304, 1991.

95. G.M. Cragg and D.J. Newman, A tale of two tumor targets: topoisoimerase I and tubulin. The Wall and Wani contribution to cancer chemotherapy, *J. Nat. Prod.*, Vol.**67**, pp. 232-244, 2004.

96. (a) E. Hamel and C.M. Lin, Interactions of combretastatin, a new plant-derived antimitotic agent with tubulin, *Biochem. Pharmacol.*, Vol.**32**, pp. 3864-3867, 1983. (b) C.M. Lin et al., Antimitotic natural products combretatstatin A-4 and combretatstatin A-2. Studies on the mechanism of their inhibition of the binding of colchicine tu tubulin, *Biochemistry*, Vol.**28**, pp. 6984-6991, 1998.

97. B. Shan et al., Selective, covalent modification of beta-tubulin residue Cys-239 by T138067, an antitumor agent with in vivo efficacy against multidrug-resistant tumors, *Proc. Nat. Acad. Sci. USA*, Vol.**96**, pp. 5686-5891,1999.

98. Mithu Banerjee et al., Sulfonamide drugs binding to the colchicine site of tubulin: thermodynamic analysis of the drug-tubulin interactions by isothermal titration calorimetry, *J. Med. Chem.*, Vol.**48**, pp. 547-555, 2005.

99. R.J. D'Amato et al., 2-Methoxyestradiol, an endogenous mammalian metabolite, inhibits tubulin polymerization by interacting at the colchicine site, *Proc. Nat. Acad. Sci. USA*, Vol.**91**, pp. 3964-3968, 1994.

100. S. J. Van Belle et al., Phase I trial of erbulozole (R55104), *Anticancer Res.*, Vol.**13**, pp. 2389-2391, 1993.

101. T. Haga et al., Preparation and formulation of (pyrimidinyloxyphenyl)benzoylureas as antitumor agents, Eur. Pat. Appl. 1991, EP 413977.

102. Mounetou, E.; Legault, J.; Lacroix, J.; C-Gaudreault, R. Antimitotic antitumor agents: synthesis, structure-activity relationships, and biological characterization of N-aryl-N'-(2-chloroethyl)ureas as new selective alkylating agents, *J. Med. Chem.*, Vol.**44**, pp. 694-702, 2001.

103. H. Yoshino et al., Preparation of sulfonamide and sulfonic ester derivatives each having tricyclic heterocyclic ring as antitumor agents, PCT Int. Appl. 1995, WO 9503279.

104. (a) V. Lisowski et al., Design, synthesis and antiproliferative activity of tripentones: a new series of antitubulin agents, *Bioorg. & Med. Chem. Lett.*, Vol.**16**, pp. 2205-2208, 2001. (b) V. Lisowski et al., Design, synthesis, and evaluation of novel thienopyrrolizinones as antitubulin agents, *J. Med. Chem.*, Vol.**47**, pp. 1448-1464, 2004.

105. K. Yoshimatsu et al., Mechanism of action of E-7010, an orally active sulfonamide antitumor agents: inhibition of mitosis by binding to the colchicine site of tubulin, *Cancer Res.*, Vol.**57**, pp. 3208-3213, 1997.

106. Gautam Bhattacharya et al., Synthesis and antitubulin activity of N1- and N4-ubstituted 3,5-dinitro sulfanilamides against african trypanosomes and *Leishmania*, *.J. Med. Chem.*, Vol.**47**, pp. 1823-1832, 2004.

107. Semi Kim et al., Novel diarylsulfonylurea derivatives as potent antimitotic agents, *Bioorganic & Medicinal Chemistry Letters*, Vol.**14**, pp. 6075–6078, 2004.

108. Yasser A. Elnakadya et al., Disorazol A1, a highly effective antimitotic agent acting on tubulin polymerization and inducing apoptosis in mammalian cells, *Biochemical Pharmacology*, Vol.**67**, pp. 927–935, 2004.

109. James T. Anderson et al., Identification of novel and improved antimitotic agents derived from noscapine, *J. Med. Chem.*, Vol.**48**, pp. 7096-7098, 2005.

110. Ya-Yun Lai et al., Synthesis and biological relationships of 3',6-substituted 2-phenyl-4-quinolone-3-carboxylic acid derivatives as antimitotic agents, *Bioorganic & Medicinal Chemistry*, Vol.**13**, pp. 265–275, 2005.

111. R. Gastpar et al., Methoxy-substituted 3-formyl-2-phenylindoles inhibit tubulin polymerization, *J. Med. Chem.*, Vol.**41**, pp. 4965-4972, 1998.

112. M. Goldbrunner et al., Inhibition of tubulin polymerization by 5,6-5,6-dihydroindolo[2,1-a]isoquinoline derivatives, *J. Med. Chem.*, Vol.**40**, pp. 3524-3533, 1997.

113. M. Medarde et al., Synthesis and antineoplastic activity of combretastatin analogues, heterocombretastatins, *Eur. J. Med. Chem.*, Vol.**33**, pp. 71-77, 1998.

114. C.M. Lin et al., *Mol. Pharmacol.*, Vol.**34**, pp. 200-208, 1988.

115. M. Medarde et al., Synthesis and pharmacological activity of diarylindole derivatives, cytotoxic agents based on combretastatins. *Bioorg. & Med. Chem. Lett.*, Vol.**9**, pp. 2303-2308, 1999.

116. E. Haar et al., *Bioorg. Med. Chem.*, Vol.**4**, pp. 1659-1671, 1996.

117. Q. Shi et al., *Current Pharm. Design*, Vol.**4**, pp. 219-248, 1998.

118. S. Mahboobi et al., Synthetis 2-aroylindole derivatives as a new class of potent tubulin-inhibitory, antimitotic agents, *J. Med. Chem.*, Vol.**44**, pp. 4535-4553, 2001.

119. T. Beckers et al., 2-Aroylindoles, a novel class of potent, orally active small molecule tubulin inhibitors, *Cancer Res.*, Vol.**62**, pp. 3113-3119, 2002.

120. M. Knaack et al., Synthesis and characterization of the biologically active 2-[1-(4-chlorobenzyl)-1*H*-indol-3-yl]-2-oxo-*N*-pyridin-4-yl acetamide, *Eur. J. Org. Chem.*, Vol.**20**, pp. 3843-3847, 2001.

121. B. Nickel et al., Indol-3-yl-glyoxylsäure-derivate mit antitumorwirkung. Patent WO 51224, 1999.

122. G. Bacher et al., T. D-24851, a novel synthetic microtubule inhibitor, exerts curative antitumoral activity in vitro, shows efficacy toward multidrug-resistant tumor cells, and lacks neutoxicity, *Cancer Res.*, Vol.**61**, pp. 392-399, 2001.

123. B.L. Flynn et al., One-pot synthesis of benzo[b]furan and indole inhibitors of tubulin polymerization, *J. Med. Chem.*, Vol.**45**, pp. 2670-2673, 2002.

124. A. Arcadi et al., A versatile approach to 2,3-disubstituted indoles through the palladium-catalyzed cyclization of o-alkynylfluoroacetanilides with vinyl triflates and aryl halides, *Tetrahedron*, Vol.**33**, pp. 3915-3918, 1992.

125. A. Arcadi et al., 2-Substituted-3-acylindoles through the palladium-catalyzed carbonilative cyclization of 2-alkynyltrifuoroacetanilides with aryl halides anf vinyl triflates, *Tetrahedron*, Vol.50, pp. 437-452, 1994.

126. a) J.-P. Liou et al., Coincise synthesis and structure-activity relationships of combretastatins A-4 analogues, 1-aroylindoles and 3-aroylindoles, a novel class of potent antitubulin agents, *J. Med. Chem.*, Vl.47, pp. 4247-4257, 2004. b) J.-P. Liou et al., Synthesis and structure-activity relationships of 3-aminobenzophenones as antimitotic agents, *J. Med. Chem.*, Vol.47, pp. 2897-2905, 2004.

127. C.-C Kuo et al., BPR0L075, a novel synthetic indole compound with antimitotic activity in human cancer cells, exerts effective antitumoral acitivity in vivo, *Cancer Res.*, Vol.64, pp. 4621-4628, 2004.

128. G. De Martino, G. La Regina, A. Coluccia, M.C. Edler, **M. Chiara Barbera**, A. Brancale, E. Wilcox, E. Hamel, M. Artico, and R. Silvestri, Arylthioindoles, potent inhibitors of tubulin polymerization, *J. Med. Chem.*, Vol.47, pp. 6120-6123, 2004.

129. B.L. Flynn et al., The synthesis an tubulin binding activity of thiophene-based analogues of combretastatin A-4, *Bioorg. & Med. Chem. Lett.*, Vol.11, pp. 2341-2343, 2001.

130. B.L. Flynn et al., One-pot synthesis of benzo[b]furan and indole inhibitors of tubulin polymerization, *J. Med. Chem.*, Vol.45, pp. 2670-2673, 2002.

131. K. Yamamoto et al., Phase I study of E7010, *Cancer Chem. Pharmacol.*, Vol.42, pp. 124-134, 1998.

132. J.G. Atkinson et al., A new synthesis of 3-arylthioindoles, *Synthesis*, pp. 480-481, 1998.

133. G. De Martino, M.C. Edler, G. La Regina, A. Coluccia, **M. Chiara Barbera**, D. Barrow, R. I. Nicholson, G. Chiosis, A. Brancale, E. Hamel, M. Artico, and Romano Silvestri, New Arylthioindoles: Potent Inhibitors of Tubulin Polymerization. 2. Structure-Activity Relationships and Molecular Modeling Studies, *J. Med. Chem.*, **ASAP Article** 10.1021/jm050809s S0022-2623(05)00809-5 (**Web Release Date:** January 19, 2006).

134. R.B.G. Ravelli et al., Insight into tubulin regulation from a complex with colchicine and a stathmin-like domain, *Nature*, Vol.428, pp. 198-202, 2002.

135. G. De Martino, G. La Regina, A. Coluccia, M.C. Edler, **M. Chiara Barbera**, A. Brancale, E. Wilcox, E. Hamel, M. Artico, and R. Silvestri, Arylthioindoles, potent inhibitors of tubulin polymerization, *Journal of Medicinal Chemistry*, **2004**, 47, 6120-6123.

136. Flynn, B. L.; Flynn, G. P.; Hamel, E.; Jung, M. K. The synthesis and tubulin binding activity of thiophene-based analogues of combretastatin A-4. *Bioorg. & Med. Chem. Lett.* **2001**, *11*, 2341-2343.

137. Flynn, B. L., Hamel, E.; Jung, M. K. One-pot synthesis of benzo[b]furan and indole inhibitors of tubulin polymerization. *J. Med. Chem.* **2002**, *45*, 2670-2673.

138. Yamamoto, K.; Noda, K.; Yoshimura, A.; Kukuoka, M.; Furuse, K.; Niitani, H. Phase I study of E7010. *Cancer Chem. Pharmacol.* **1998**, *42*, 124-134.

139. G. De Martino, M.C. Edler, G. La Regina, A. Coluccia, **M. Chiara Barbera**, D. Barrow, R. I. Nicholson, G. Chiosis, A. Brancale, E. Hamel, M. Artico, and Romano Silvestri, New Arylthioindoles: Potent Inhibitors of Tubulin Polymerization. 2. Structure-Activity Relationships and Molecular Modeling Studies, *Journal of Medicinal Chemistry*, **ASAP Article** 10.1021/jm050809s S0022-2623(05)00809-5 (**Web Release Date:** January 19, 2006).

140. Ravelli, R. B. G.; Gigant, B.; Curmi, P. A.; Jourdain, I.; Lachkar, S.; Sobel, A.; Knossow, M. Insight into tubulin regulation from a complex with colchicine and a stathmin-like domain *Nature* **2004**, *428*, 198-202.

141. Tripos SYBYL 7.0; Tripos Inc., 1699 South Hanley Rd, St. Louis, Missouri 63144, USA..http://www. tripos.com.

142. Molecular Operating Environment (MOE). Chemical Computing Group, Inc. Montreal, Quebec, Canada. http://www.chemcomp.com

143. Jaroslaw Polanski, Self-organizing neural network for modeling 3D QSAR of colchinoids, *Acta Biochimica Polonica*, 2000, Vol.47, No.1, pp. 37-45.

144. Code "more_dock.svl" obtained from SLV Exchange website http://svl.chemcomp.com., Chemical Computing Group, Inc., Montreal, Canada.

145. Silvestri, R.; De Martino, G.; Artico, M.; Massa, S.; Marceddu, T.; Loi, A. G.; Musiu, C.; La Colla, C. Indolyl Aryl Sulfones (IASs). Part 1: SAR Studies and in vitro anti-HIV-1 activity against wt RT and related mutants. *J. Med. Chem.* **2003**, *46*, 2482-2493.

146. http://www.baxter-oncology.com/english/proects/index.htmL.

147. R. Gastpar, E. von Angere et al., Methoxy-substituted 3-formyl-2-phenylindoles inhibit tubulin polymerization, *Journal of Medicinal Chemistry*, 1998, Vol. 41, pp. 4965-4972.

148. M. Medarde et al., Synthesis and antineoplastic activity of combretastatin analogue: heterocombretastatin, *European Journal of Medicinal Chemistry*, 1998, Vol. 33, pp. 71-77.

149. B. L. Flynn et al., One-pot synthesis of benzo[b]furan and indole inhibitors of tubulin polymerization, *Journal of Medicinal Chemistry*, 2002, Vol. 45, pp. 2670-2673.

150. R. Silvestri et al., Novel indolyl aryl sulfones active against HIV-1 carrying NNRTI resistance mutations: synthesis and SAR studies, *Journal of Medicinal Chemistry*, 2003, Vol. 46, pp. 2482-2493.

151. Gay P. Fagan et al., Indoline analogues of idazoxan: potent α2-antagonists and α1-agonists, *Journal of Medicinal Chemistry*, 1988, Vol.31, pp. 944-948.

152. D.C. Beshore and C.J. Dinsmore, Preparation of ethyl 5-iodo-1*H*-indole-2-carboxylate, *Synthetic Communications®*, 2003, Vol.33, No.14, pp. 2423-2427.

Chapter 3

1. Hoofnagle, J. H. Course and outcome of hepatitis C. *Hepatology*, Vol.36, S21–S29 (2002).

2. Alter, H. J. & Seeff, L. B. Recovery, persistence, and sequelae in hepatitis C virus infection: a perspective on long-term outcome. *Semin. Liver Dis.*, Vol.20, pp. 17–35 (2000).

3. Alter, M. J. Epidemiology of hepatitis C. *Eur. J. Gastroenterol. Hepatol.*, Vol. **8**, pp. 319–323 (1996).
4. Alter, M. J. Epidemiology of hepatitis C. *Hepatology*, Vol.**26**, pp. 62S–65S (1997).
5. Alter, M. J. Hepatitis C virus infection in the United States. *J. Hepatol.*, Vol.**31** (Suppl. 1), pp. 88–91 (1999).
6. Alter, M. J. Prevention of spread of hepatitis C. *Hepatology*, Vol. **36**, S93–S98 (2002).
7. David G. Bowen and Christopher M. Walker, Adaptive immune responses in acute and chronic hepatitis C virus infection, *Nature Insight*, Vol.**436**, pp. 946-952, 18 August 2005.
8. Choo, Q. L. et al. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science*, Vol.**244**, pp. 359–362 (1989).
9. Bartenschlager, R. Hepatitis C virus replicons: potential role for drug development. *Nature Rev. Drug Discov.*, Vol.**1**, pp. 911–916 (2002).
10. Zhong, J. et al. Robust hepatitis C virus infection *in vitro*. *Proc. Natl Acad. Sci. USA*, Vol.**102**,pp. 9294–9299 (2005).
11. Lindenbach, B. D. *et al.* Complete replication of hepatitis C virus in cell culture. *Science*, Vol.**309**, pp. 623–626 (2005).
12. Wakita, T. et al. Production of infectious hepatitis C virus in tissue culture from a cloned viral genome. *Nature Med.*, Vol.**11**, pp. 791–796 (2005).
13. Wieland, S. F. & Chisari, F. V. Stealth and cunning: hepatitis B and hepatitis C. *J. Virol.*, Vol.**79**, pp. 9369–9380 (2005).
14. Neumann, A. U. et al. Hepatitis C viral dynamics *in vivo* and the antiviral efficacy of interferon-α therapy. *Science*, Vol. **282**, pp. 103–107 (1998).
15. Lanford, R. E. & Bigger, C. Advances in model systems for hepatitis C virus research. *Virology*, Vol. **293**, pp. 1–9 (2002).
16. Mondelli, M. U. et al. Hypervariable region 1 of hepatitis C virus: immunological decoy or biologically relevant domain? *Antiviral Res.*, Vol.**52**, pp. 153–159 (2001).
17. Bowen, D. G. & Walker, C. M. Mutational escape from CD8+ T cell immunity: HCV evolution, from chimpanzees to man. *J. Exp. Med.*, Vol.**201**, pp. 1709–1714 (2005).
18. Feld, J.J. and Hoofnagle, J.H., Mechanism of action of interferon and ribavirin in treatment of hepatitis C, Nature Insight, Vol.**436**, pp. 967-972, 18 August 2005.
19. McHutchison, J. G. & Fried, M. W. Current therapy for hepatitis C: pegylated interferon and ribavirin. *Clin. Liver Dis.*, Vol., **7**, pp. 149–161 (2003).
20. Pawlotsky, J. M. Mechanisms of antiviral treatment efficacy and failure in chronic hepatitis C. *Antiviral Re.*, Vol.**59**, pp. 1–11 (2003).
21. Strader, D. B., Wright, T., Thomas, D. L. & Seeff, L. B. Diagnosis, management, and treatment of hepatitis C. *Hepatology*, 3 Vol.**9**, pp. 1147–1171 (2004).
22. Lamarre, D. et al. An NS3 protease inhibitor with antiviral effects in humans infected with hepatitis C virus. *Nature*, Vol.**426**, pp. 186–189 (2003).
23. Nevens, F. et al. A pilot study of therapeutic vaccination with envelope protein E1 in 35 patients with chronic hepatitis C. *Hepatology*, Vol.**38**, pp. 1289–1296 (2003).
24. Houghton, M. and Abrignani, S., Prospects for a vaccine against the hepatitis C virus, *Nature Insight*, Vol.**436**, pp. 961-966, 18 August 2005.
25. De Francesco, R. and Migliaccio, G., Challenges and successes in developing new therapies for hepatitis C, *Nature Insight*, Vol.**436**, pp. 953-960, 18 August 2005.
26. Choo, Q.-L. et al. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science*, Vol.**244**, pp. 359–362 (1989).
27. Feinstone, S. M., Kapikian, A. Z., Purcell, R. H., Alter, H. J. & Holland, P. V. Transfusionassociated hepatitis not due to viral hepatitis type A or B. *N. Engl. J. Med.*, Vol.**292**, pp.767–770 (1975).
28. Spahn, C. M. et al. Hepatitis C virus IRES RNA-induced changes in the conformation of the 40S ribosomal subunit. *Science*, Vol.**291**, pp. 1959–1962 (2001).
29. Ji, H., Fraser, C. S., Yu, Y., Leary, J. & Doudna, J. A. Coordinated assembly of human translation initiation complexes by the hepatitis C virus internal ribosome entry site RNA. *Proc. Natl Acad. Sci. USA*, Vol.**101**, pp. 16990–16995 (2004).
30. Otto, G. A. & Puglisi, J. D. The pathway of HCV IRES-mediated translation initiation. *Cell*, Vol.**119**, pp. 369–380 (2004).
31. Griffin, S. D. et al. The p7 protein of hepatitis C virus forms an ion channel that is blocked by the antiviral drug, Amantadine. *FEBS Lett.*, Vol.**535**, pp. 34–38 (2003).
32. Pavlovic, D. et al. The hepatitis C virus p7 protein forms an ion channel that is inhibited by long-alkyl-chain iminosugar derivatives. *Proc. Natl Acad. Sci. USA*, Vol.**100**, pp. 6104–6108 (2003).
33. Branch, A. D., Stump, D. D., Gutierrez, J. A., Eng, F. & Walewski, J. L. The hepatitis C virus alternate reading frame (ARF) and its family of novel products: the alternate reading frame protein/F-protein, the double-frameshift protein, and others. *Semin. Liver Dis.*, Vol.**25**, pp. 105–117 (2005).
34. Penin, F., Dubuisson, J., Rey, F. A., Moradpour, D. & Pawlotsky, J. M. Structural biology of hepatitis C virus. *Hepatology*, Vol. **39**, pp. 5–19 (2004).
35. Failla, C., Tomei, L. & De Francesco, R. Both NS3 and NS4A are required for proteolytic processing of hepatitis C virus nonstructural proteins. *J. Virol.*, Vol.**68**, pp. 3753–3760 (1994).
36. Bartenschlager, R., Ahlborn-Laake, L., Mous, J. & Jacobsen, H. Kinetic and structural analyses of hepatitis C virus polyprotein processing. *J. Virol.*, Vol.**68**, pp. 5045–5055 (1994).
37. Lin, C., Thomson, J. A. & Rice, C. M. A central region in the hepatitis C virus NS4A protein allows formation of an active NS3–NS4A serine proteinase complex *in vivo* and *in vitro*. *J. Virol.*, Vol.**69**, pp. 4373–4380 (1995).
38. De Francesco, R. and Migliaccio, G., Challenges and successes in developing new therapies for hepatitis C, *Nature Insight*, Vol.**436**, pp. 953-960, 18 August 2005.
39. Tai, C.-L., Chi, W.-K., Chen, D.-S. & Hwang, L.-H. The helicase activity associated with hepatitis C virus nonstructural protein 3 (NS3). *J. Virol.*, Vol.**70**, pp. 8477–8484 (1996).
40. Levin, M. K., Gurjar, M. & Patel, S. S. A Brownian motor mechanism of translocation and strand separation by hepatitis C virus helicase. *Nature Struct. Mol. Biol.*, Vol.**12**, pp. 429–435 (2005).

41. Serebrov, V. & Pyle, A. M. Periodic cycles of RNA unwinding and pausing by hepatitis C virus NS3 helicase. *Nature*, Vol.430, pp. 476–480 (2004).

42. Frick, D. N., Rypma, R. S., Lam, A. M. & Gu, B. The nonstructural protein 3 protease/helicase requires an intact protease domain to unwind duplex RNA efficiently. *J. Biol. Chem.*, Vol.279, pp. 1269–1280 (2004).

43. Pang, P. S., Jankowsky, E., Planet, P. J. & Pyle, A. M. The hepatitis C viral NS3 protein is a processive DNA helicase with cofactor enhanced RNA unwinding. *EMBO J.*, Vol.21, pp. 1168–1176 (2002).

44. Xu, T. et al., Structure of the Dengue Virus Helicase/Nucleoside Triphosphatase Catalytic Domain at a Resolution of 2.4 Å, *Journal of Virology*, Vol. 79, No. 16, pp. 10278–10288, Aug. 2005.

45. He, Y. & Katze, M. G. To interfere and to anti-interfere: the interplay between hepatitis C virus and interferon. *Viral Immunol.*, Vol.15, pp. 95–119 (2002).

46. Macdonald, A. & Harris, M. Hepatitis C virus NS5A: tales of a promiscuous protein. *J. Gen. Virol.*, Vol.85, pp. 2485–2502 (2004).

47. Tellinghuisen, T. L. & Rice, C. M. Interaction between hepatitis C virus proteins and host cell factors. *Curr. Opin. Microbiol.*, Vol. 5, pp. 419–427 (2002).

48. Katze, M. G. et al. Ser2194 is a highly conserved major phosphorylation site of the hepatitis C virus nonstructural protein NS5A. *Virology*, Vol.278, pp. 501–513 (2000).

49. Reed, K. E. & Rice, C. M. Identification of the major phosphorylation site of the hepatitis C virus H strain NS5A protein as serine 2321. *J. Biol. Chem.*, Vol.274, pp. 28011–28018 (1999).

50. Coito, C., Diamond, D. L., Neddermann, P., Korth, M. J. & Katze, M. G. High-throughput screening of the yeast kinome: identification of human serine/threonine protein kinases that phosphorylate the hepatitis C virus NS5A protein. *J. Virol.*, Vol.78, pp. 3502–3513 (2004).

51. Ide, Y., Tanimoto, A., Sasaguri, Y. & Padmanabhan, R. Hepatitis C virus NS5A protein is phosphorylated *in vitro* by a stably bound protein kinase from HeLa cells and by cAMPdependent protein kinase A-alpha catalytic subunit. *Gene*, Vol.201, pp. 151–158 (1997).

52. Kim, J., Lee, D. & Choe, J. Hepatitis C virus NS5A protein is phosphorylated by casein kinase II. *Biochem. Biophys. Res. Commun.*, Vol.257, pp. 777–781 (1999).

53. Reed, K. E., Xu, J. & Rice, C. M. Phosphorylation of the hepatitis C virus NS5A protein *in vitro* and *in vivo*: properties of the NS5A-associated kinase. *J. Virol.*, Vol.71, pp. 7187–7197 (1997).

54. Brass, V. et al. An amino-terminal amphipathic alpha-helix mediates membrane association of the hepatitis C virus nonstructural protein 5A. *J. Biol. Chem.*, Vol.277, pp. 8130–8139 (2002).

55. Tellinghuisen, T. L., Marcotrigiano, J., Gorbalenya, A. E. & Rice, C. M. The NS5A protein of hepatitis C virus is a zinc metalloprotein. *J. Biol. Chem.*, Vol.279, pp. 48576–48587 (2004).

56. Tellinghuisen, T. L., Marcotrigiano, J. & Rice, C. M. Structure of the zinc-binding domain of an essential replicase component of hepatitis C virus. *Nature*, Vol.435, pp. 374–379 (2005).

57. Hong, Z. et al. A novel mechanism to ensure terminal initiation by hepatitis C virus NS5B polymerase. *Virology*, Vol.285, pp. 6–11 (2001).

58. Butcher, S. J., Grimes, J. M., Makeyev, E. V., Bamford, D. H. & Stuart, D. I. A mechanism for initiating RNA-dependent RNA polymerization. *Nature*, Vol.410, pp. 235–240 (2001).

59. van Dijk, A. A., Makeyev, E. V. & Bamford, D. H. Initiation of viral RNA-dependent RNA polymerization. *J. Gen. Virol.*, Vol.85, pp. 1077–1093 (2004).

60. Bressanelli, S., Tomei, L., Rey, F. A. & De Francesco, R. Structural analysis of the hepatitis C virus RNA polymerase in complex with ribonucleotides. *J. Virol.*, Vol.76, pp. 3482–3492 (2002).

61. Schmidt-Mende, J. et al. Determinants for membrane association of the hepatitis C virus RNA-dependent RNA polymerase. *J. Biol. Chem.*, Vol.276, pp. 44052–44063 (2001).

62. Wang, Q. M. et al. Oligomerization and cooperative RNA synthesis activity of hepatitis C virus RNA-dependent RNA polymerase. *J. Virol.*, Vol.76, pp. 3865–3872 (2002).

63. Michael Gale Jr and Eileen M. Foy, Evasion of intracellular host defence by hepatitis C virus, *Nature Insight*, Vol.436, pp. 939-945, 18 August 2005.

64. Blight, K. J., Kolykhalov, A. A. & Rice, C. M. Efficient initiation of HCV RNA replication in cell culture. *Science*, Vol.290, pp. 1972–1974 (2000).

65. Lohmann, V. et al. Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science*, Vol.285, pp. 110–113 (1999).

66. Brown, R. S. Jr & Gaglio, P. J. Scope of worldwide hepatitis C problem. *Liver Transpl.*, Vol.9, S10–S13 (2003).

67. Wasley, A. & Alter, M. J. Epidemiology of hepatitis C: geographic differences and temporal trends. *Semin. Liver Dis.*, Vol.20, pp. 1–16 (2000).

68. Feld, J.J. and Hoofnagle, J.H., Mechanism of action of interferon and ribavirin in treatment of hepatitis C, Nature Insight, Vol.436, pp. 967-972, 18 August 2005.

69. R. De Francesco and G. Migliaccio, Challenges and successes in developing new therapies for hepatitis C, *Nature Insight*, Vol.436, pp. 953-960, 18 August 2005.

70. De Francesco, R. and Migliaccio, G., Challenges and successes in developing new therapies for hepatitis C, *Nature Insight*, Vol.436, pp. 953-960, 18 August 2005.

71. Lindenbach, Brett D. and Rice, Charles M., Unravelling hepatitis C virus replication from genome to function, *Nature Insight*, Vol. 436, pp.933 – 938, 18 August 2005

72. Gale, Michael Jr and. Foy, Eileen M, Evasion of intracellular host defence by hepatitis C virus, *Nature Insight*, Vol.436, pp. 939-945, 18 August 2005.

73. Urbani, A. et al. Substrate specificity of the hepatitis C virus serine protease NS3. *J. Biol. Chem.*, Vol.272, pp. 9204–9209 (1997).

74. Yan, Y. et al. Complex of NS3 protease and NS4A peptide of BK strain hepatitis C virus: a 2.2 Å resolution structure in a hexagonal crystal form. *Protein Sci.*, Vol.7, pp. 837–847 (1998).

75. Kim, J. L. et al. Crystal structure of the hepatitis C virus NS3 protease domain complexed with a synthetic NS4A cofactor peptide. *Cell,* Vol.87, pp. 343–355 (1996).

76. Love, R. A. et al. The crystal structure of hepatitis C virus NS3 proteinase reveals a trypsinlike fold and a structural zinc binding site. *Cell,* Vol.87, pp. 331–342 (1996).

77. Llinas-Brunet, M. *et al.* Peptide-based inhibitors of the hepatitis C virus serine protease. *Bioorg. Med. Chem. Lett.,* Vol.8, pp. 1713–1718 (1998).

78. Steinkuhler, C. et al. Product inhibition of the hepatitis C virus NS3 protease. *Biochemistry,* Vol.37, pp. 8899–8905 (1998).

79. Lamarre, D. et al. An NS3 protease inhibitor with antiviral effects in humans infected with hepatitis C virus. *Nature,* Vol.426, pp. 186–189 (2003).

80. Llina`s-Brunet, M., Structure-Activity Study on a Novel Series of Macrocyclic Inhibitors of the Hepatitis C Virus NS3 Protease Leading to the Discovery of BILN 2061, *J. Med. Chem.,* Vol.47, pp. 1605-1608, 2004.

81. Llinas-Brunet, M. et al. Highly potent and selective peptide-based inhibitors of the hepatitis C virus serine protease: towards smaller inhibitors. *Bioorg. Med. Chem. Lett.,* Vol.10, pp. 2267–2270 (2000).

82. Goudreau, N. et al. NMR structural characterization of peptide inhibitors bound to the Hepatitis C virus NS3 protease: design of a new P2 substituent. *J. Med. Chem.,* Vol.47, pp. 123–132 (2004).

83. Tsantrizos, Y. S. et al. Macrocyclic inhibitors of the NS3 protease as potential therapeutic agents of hepatitis C virus infection. *Angew. Chem. Int. Ed. Engl.,* Vol.42, pp. 1356–1360 (2003).

84. Reiser, M. et al. Antiviral efficacy of NS3-serine protease inhibitor BILN-2061 in patients with chronic genotype 2 and 3 hepatitis C. *Hepatology,* Vol.41, pp. 832–835 (2005).

85. Lu, L. et al. Mutations conferring resistance to a potent hepatitis C virus serine protease inhibitor *in vitro*. *Antimicrob. Agents Chemother.,* Vol.48, pp. 2260–2266 (2004).

86. Lin, C. et al. *In vitro* resistance studies of hepatitis C virus serine protease inhibitors, VX-950 and BILN 2061: structural analysis indicates different resistance mechanisms. *J. Biol. Chem.,* Vol.279, pp. 17508–17514 (2004).

87. Chen, S.-H. et al. P1 and P1'-optimization of [3,4]-bicycloproline P2 incorporated tetrapeptidyl α-ketoamide based HCV protease inhibitors. *Lett. Drug Des. Disc.,* Vol.2, pp. 118–123 (2005).

88. Vertex Pharmaceuticals reports that oral hepatitis C protease inhibitor VX-950 dramatically reduces viral levels in phase Ib clinical study. http://www.vrtx.com/Pressreleases2005/pr051705.html (2005).

89. Lemon, S. M., Yi, M. & Li, K. Strong reasons make strong actions. The antiviral efficacy of NS3/4A protease inhibitors. *Hepatology,* Vol.41, pp. 671–673 (2005).

90. Lin, C. et al. *In vitro* resistance mutations against VX-950 and BILN 2061, two protease inhibitor clinical candidates: single-resistance, cross-resistance and fitness. *Hepatology,* Vol.40 (Suppl. 1), 404A (2004).

91. Chen, K.X. et al., Design, Synthesis, and Biological Activity of *m*-Tyrosine-Based 16- and 17-Membered Macrocyclic Inhibitors of Hepatitis C Virus NS3 Serine Protease *J. Med. Chem.,* Vol.48, pp. 6229-6235, 2005

92. Chen, K.X. et al., Potent 7-Hydroxy-1,2,3,4-tetrahydroisoquinoline-3-carboxylic Acid-Based Macrocyclic Inhibitors of Hepatitis C Virus NS3 Protease, *J. Med. Chem.,* Vol.49, pp. 567-574, 2006.

93. Behrens, S. E., Tomei, L. & De Francesco, R. Identification and properties of the RNAdependent RNA polymerase of hepatitis C virus. *EMBO J.,* Vol.15, pp. 12–22 (1996).

94. Ago, H. et al. Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. *Struct. Fold. Des.,* Vol.7, pp. 1417–1426. (1999).

95. Bressanelli, S. et al. Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. *Proc. Natl Acad. Sci. USA,* Vol.96, pp. 13034–13049. (1999).

96. Lesburg, C. A. et al. Crystal structure of the RNA-dependent RNA polymerase from hepatitis C virus reveals a fully encircled active site. *Nature Struct. Biol.,* Vol.6, pp. 937–943. (1999).

97. Biswal, B. K. et al. Crystal structures of the RNA dependent RNA polymerase genotype 2a of hepatitis C virus reveal two conformations and suggest mechanisms of inhibition by nonnucleoside inhibitors. *J. Biol. Chem.,* Vol.280, pp. 18202–18210 (2005).

98. Beaulieu, P. L. & Tsantrizos, Y. S. Inhibitors of the HCV NS5B polymerase: new hope for the treatment of hepatitis C infections. *Curr. Opin. Investig. Drugs,* Vol.5, 838–850 (2004).

99. Afdhal, N. et al. Final phase I/II trial results for NM283, a new polymerase inhibitor for hepatitis C: antiviral efficacy and tolerance in patients with HCV-1 infection, including previous interferon failures. http://www.idenix.com/products/datapres_nm283/AfdhalAASLD04_10-04.pdf (2004).

100. Afdal, N. et al. Enhanced antiviral efficacy for Valopicitabine (NM283) plus peg-interferon in hepatitis C patients with HCV genotype-1 infection: Results of a phase IIa multicenter trial. http://www.idenix.com/products/datapres_nm283/AfdhalEASL2005_4-05.pdf (2005).

101. Carroll, S.et al. Susceptibility of different genotypes of hepatitis C virus to inhibition by nucleoside and non nucleoside inhibitors. *Antiviral Res.,* Vol.62, A83 (2004).

102. Ludmerer, S. W. et al. Replication fitness and NS5B drug sensitivity of diverse hepatitis C virus isolates characterized by using a transient replication assay. *Antimicrob. Agents Chemother.,* Vol.49, pp. 2059–2069 (2005).

103. Migliaccio, G. et al. Characterization of resistance to non-obligate chain-terminating ribonucleoside analogs that inhibit hepatitis C virus replication *in vitro*. *J. Biol. Chem.,* Vol.278, pp. 49164–49170 (2003).

104. Olsen, D. B. et al. A 7-deaza-adenosine analog is a potent and selective inhibitor of hepatitis C virus replication with excellent pharmacokinetic properties. *Antimicrob .Agents Chemother.,* Vol.48, pp. 3944–3953 (2004).

105. Clark, J.L. et al., Design, Synthesis, and Antiviral Activity of 2'-Deoxy-2'-fluoro-2'-*C*-methylcytidine, a Potent Inhibitor of Hepatitis C Virus Replication, *J. Med. Chem.,* Vol. 48, pp. 5504-5508, 2005.

106. Yung-hyo Koh, Y. et al., Design, Synthesis, and Antiviral Activity of Adenosine 5'-Phosphonate Analogues as Chain Terminators against Hepatitis C Virus, *Journal of Medicinal Chemistry*

107. Hashimoto, H., Mizutani, K. & Yoshida, A. in *WO 0014 7883* (Japan Tobacco Inc., Published International Patent Application, 2001).

108. La Plante, S. R. et al. Binding mode determination of benzimidazole inhibitors of the hepatitis C virus RNA polymerase by a structure and dynamics strategy. *Angew Chem. Int. Ed. Engl.*, Vol.**43**, pp. 4306–4311 (2004).

109. Tomei, L. et al. Mechanism of action and antiviral activity of benzimidazole-based allosteric inhibitors of the hepatitis C virus RNA-dependent RNA polymerase. *J. Virol.*, Vol.**77**, pp. 13225–13231 (2003).

110. Di Marco, S. et al. Interdomain communication in hepatitis C virus polymerase abolished by small-molecule inhibitors bound to a novel allosteric site. *J. Biol. Chem.* published online 13 June 2005.

111. Lu, H. in WO 2005/000308 (Rigel Pharmaceuticals, USA. Published International Patent Application, 2005).

112. Poor bioavailability results in insignificant clinical effects for Rigel R803 in phase I/II HCV trial. http://www.rigel.com/rigel/pr_1101094254 (2004).

113. ViroPharma announces data from HCV-086 proof of concept study. http://phx.corporateir.net/phoenix.zhtml?c=92320&p=irol-researchNewsArticle&ID=684145&highlight (2005).

114. Chan, L. et al. Discovery of thiophene-2-carboxylic acids as potent inhibitors of HCV NS5B polymerase and HCV subgenomic RNA replication. Part 2: tertiary amides. *Bioorg. Med. Chem. Lett.*, Vol.**14**, pp. 797–800 (2004).

115. Chan, L. et al. Discovery of thiophene-2-carboxylic acids as potent inhibitors of HCV NS5B polymerase and HCV subgenomic RNA replication. Part 1: Sulfonamides. *Bioorg. Med. Chem. Lett.*, Vol.**14**, pp. 793–796 (2004).

116. Wang, M. et al. Non-nucleoside analogue inhibitors bind to an allosteric site on HCV NS5B polymerase. Crystal structures and mechanism of inhibition. *J. Biol. Chem.*, Vol.**278**, pp. 9489–9495 (2003).

117. Love, R. A. et al. Crystallographic identification of a noncompetitive inhibitor binding site on the hepatitis C virus NS5B RNA polymerase enzyme. *J. Virol.*, Vol.**77**, pp. 7575–7581 (2003).

118. Dhanak, D. et al. Identification and biological characterization of heterocyclic inhibitors of the hepatitis C virus RNA-dependent RNA polymerase. *J. Biol. Chem.*, Vol.**277**, pp. 38322–38327 (2002).

119. Gu, B. et al. Arresting initiation of hepatitis C virus RNA synthesis using heterocyclic derivatives. *J. Biol. Chem.*, Vol.**278**, pp. 16602–16607 (2003).

120. Tomei, L. et al. Characterization of the inhibition of hepatitis C virus RNA replication by nonnucleosides. *J. Virol.*, Vol.**78**, pp. 938–946 (2004).

121. Harper, S. et al., Potent Inhibitors of Subgenomic Hepatitis C Virus RNA Replication through Optimization of Indole-*N*-Acetamide Allosteric Inhibitors of the Viral NS5B Polymerase, *J. Med. Chem.*, Vol.**48**, pp. 4547-4557, 2005

122. Harper, S. et al., Development and Preliminary Optimization of Indole-*N*-Acetamide Inhibitors of Hepatitis C Virus NS5B Polymerase, *J. Med. Chem., Letters*, Vol.**48**, pp. 1314-1317, 2005.

123. Powers, J.P. et al., SAR and Mode of Action of Novel Non-Nucleoside Inhibitors of Hepatitis C NS5b RNA Polymerase *J. Med. Chem*

124. Di Santo, R. et al., Simple but Highly Effective Three-Dimensional Chemical-Feature-Based Pharmacophore Model for Diketo Acid Derivatives as Hepatitis C Virus RNA-Dependent RNA Polymerase Inhibitors, *J. Med. Chem.*, Vol.**48**, pp.6304-6314, 2005

125. Tedesco, R. et al., 3-(1,1-Dioxo-2*H*-(1,2,4)-benzothiadiazin-3-yl)-4-hydroxy-2(1*H*)-quinolinones, Potent Inhibitors of Hepatitis C Virus RNA-Dependent RNA Polymerase, *J. Med. Chem*

126. Braasch, D. A. *et al.* Biodistribution of phosphodiester and phosphorothioate siRNA. *Bioorg. Med. Chem. Lett.*, Vol.**14**, pp. 1139–1143 (2004).

127. Foster, G. R. Past, present, and future hepatitis C treatments. *Semin. Liver Dis.*, Vol.**24** (Suppl. 2), pp. 97–104 (2004).

128. Kronke, J. et al. Alternative approaches for efficient inhibition of hepatitis C virus RNA replication by small interfering RNAs. *J. Virol.*, Vol.**78**, pp. 3436–3446 (2004).

129. Kapadia, S. B., Brideau-Andersen, A. & Chisari, F. V. Interference of hepatitis C virus RNA replication by short interfering RNAs. *Proc. Natl Acad Sci USA*, Vol.**100**, pp. 2014–2018 (2003).

130. Yokota, T. et al. Inhibition of intracellular hepatitis C virus replication by synthetic and vector-derived small interfering RNAs. *EMBO Rep.*, Vol.**4**, pp. 602–608 (2003).

131. Wilson, J. A. et al. RNA interference blocks gene expression and RNA synthesis from hepatitis C replicons propagated in human liver cells. *Proc. Natl Acad. Sci. USA*, Vol.**100**, pp. 2783–2788 (2003).

132. Randall, G., Grakoui, A. & Rice, C. M. Clearance of replicating hepatitis C virus replicon RNAs in cell culture by small interfering RNAs. *Proc. Natl Acad. Sci. USA*, Vol.**100**, pp. 235–240 (2003).

133. Han, J. et al. Inhibition of HCV replication *in vivo* by nuclease-resistant siRNAs that are targeted to the liver. Presented at 11th International Symp. Hepatitis C Virus and Related Viruses (Heidelberg, Germany 2004).

134. Benitec announces clinical candidate for treatment of hepatitis C. http://www.benitec.com/PRDownloads/Hepatitis%20C%20Clinical%20Candidate%20 050905%20.pdf (2005).

135. O'Neill, L. A. TLRs: Professor Mechnikov, sit on your hat. *Trends Immunol.*, Vol.**25**, pp. 687–693 (2004).

136. Takeda, K., Kaisho, T. & Akira, S. Toll-like receptors. *Annu. Rev. Immunol.*, Vol.**21**, pp. 335–376 (2003).

137. Hahn, Y. S. Subversion of immune responses by hepatitis C virus: immunomodulatory strategies beyond evasion? *Curr. Opin. Immunol.*, Vol.**15**, pp. 443–449 (2003).

138. Schetter, C. & Vollmer, J. Toll-like receptors involved in the response to microbial pathogens: development of agonists for toll-like receptor 9. *Curr. Opin. Drug Discov. Dev.*, Vol.**7**, pp. 204–210 (2004).

139. Coley reports results from phase I studies of ActilonTM for hepatitis C. http://www.coleypharma.com/coley/pr_1105025921 (2005).

140. Lee, J. et al. Molecular basis for the immunostimulatory activity of guanine nucleoside analogs: activation of Toll-like receptor 7. *Proc. Natl Acad. Sci. USA*, Vol.**100**, pp. 6646–6651 (2003).

141. Anadys Pharmaceuticals announces selection of ANA975 as a development candidate for front-line treatment of chronic hepatitis C. http://phx.corporateir. net/phoenix.zhtml?c=148908&p=irol-newsArticle&ID=575761&highlight (2004).

142. Kim, J. L., Morgenstern, K. A., Griffith, J. P., Dwyer, M. D., Thomson, J. A., Morcko, M. A., Lin, C., Caron, P. R. (1998) Hepatitis C virus NS3 helicase domain with a bound oligonucleotide: the crystal structure provides insight into the mode of

unwinding. *Structure,* Vol.6, pp. 89-100 [WWW]
<http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6VSR-4CJB9BG-9S-
H&_cdi=6269&_user=2203889&_orig=browse&_coverDate=01%2F15%2F1998&_sk=999939998&view=c&wchp=dGL
bVlz-zSkzk&md5=26023b712ee7d1ff4a02062d9dc365fd&ie=/sdarticle.pdf> [accessed March 2005].

143. Phoon, C. Ng, P. Ting, A. Yeo, S. Sim, M. 2001. Biological Evaluation of Hepatitis C Virus Helicase Inhibitors. *Bioorganic & Medicinal Chemistry Letters,* Vol.11, pp. 1647–1650.
144. Diana, G. Bailey, T. 1997. Compounds, compositions and methods for treatment of hepatitis C, *United States Patent* No. 5,633,388.
145. Mayfield, E. 2004. Unpublished Work.
146. Matusan, A. Pryor, M. Davidson, A. Wright, P. 2001. Mutagenesis of the *Dengue Virus Type 2* NS3 Protein within and outside Helicase Motifs: Effects on Enzyme Activity and Virus Replication, *Journal of Virology,* pp. 9633–9643.
147. Vlachakis, D. 2004. Unpublished Work.
148. *Journal of Medicinal Chemistry,* 2004, Vol.47, No.15, p.3895.
149. M.F. Brana et al., Synthesis of benzimidazo-substituted 3-quinolinecarboxylic acids as antibacterial agents, *J.Heterocyclic Chem.,* 1990, Vol.27, p.1177-1180.
150. Horscroft, N. Lai, V. Cheney, W. Yao, N. Wu, JZ. Hong, Z. Zhong, W., Replicon cell culture system as a valuable tool in antiviral drug discovery against hepatitis C virus. *Antiviral Chemistry & Chemotherapy*; 16: 1–12, 2005.
151. Molecular Operating Environment (MOE). Chemical Computing Group, Inc. Montreal, Quebec, Canada. http://www.chemcomp.com
152. 152Kim, J. L. Morgenstern et al., Hepatitis C virus NS3 RNA helicase domain with a boundoligonucleotide: the crystal structure provides insights into the mode of unwinding, Structure, Vol.6, pp. 89, 1998.
153. Tripos SYBYL 7.0; Tripos Inc., 1699 South Hanley Rd, St. Louis, Missouri 63144, USA..http://www. tripos.com
154. P. Bianco, Hepatitis C NS3 helicase unwinds RNA in leaps and bounds, The Lancet, Vol.364, pp 1385-1387, 2004.
155. Itoh, T., Nagate, K., Kurihara, A., Miyazaki, M., Ohsawa, A. Reductive amination of aldehydes and ketones by a Hantzsch dihydropyridine using scandium triflate as a catalyst. (2002) *Tetrahedrom Lett,* 43 3105-3108
156. a) Smith, J. G. Ho, I. The Effect of Electronegative Substituents on the Reductive Dimerization of Schiff Bases. Formation of Vicinal Dianions. (1973) *J Org Chem,* 38 2776-2779 (b) Brown, F. J., Bernstein, P. R., Cronk, L. A., Dosset, D. L., Hebbel, K. C., Madskuie Jr., T. P., Shapiro, H. S., Vacek, E. P., Yee, Y. K., Willard, A. K., Krell, R. D., Snyder, D. W. Hydroxyacetophenone-Derived Antagonists of the Peptidoleukotrienes. (1989) *J Med Chem,* 32 807-826 (c) Baures, P. W., Oza, V. B., Peterson, S. A., Kelly, J. W. Synthesis and Evaluation of Inhibitors of Transthyretin Amyloid Formation Based on the Non-steroidal Anti-inflammatory Drug, Flufenamic Acid. (1999) *Bioorg Med Chem,* 7 1339-1347.
157. S. Sellarajah et al., Synthesis of Analogues of Congo Red and Evaluation of Their Anti-Prion Activity, *J. Med. Chem.,* 2004, Vol. 47, pp.5515-5534_.
158. List, B. (2000) The Direct Catalytic Asymmetric Three Component Mannich Reaction. *Journal of the American Chemical Society.* 122; 9336-9337.
159. Notz, W., Sakthivel, K., Bui, T., et al. Amine-Catalysed Direct Asymmetric Mannich-Type Reactions. *Tetrahedron Letters* 42; 199-201.
i. Cummins, A.-C. Computer-Aided Design & Synthesis of Novel Antiviral Agents Thesis for Degree of Master of Pharmacy (MPharm honours). (2003) *Welsh School of Pharmacy Cardiff University.*

Chapter 4

1. Ulf R. Rapp, Apoptosis in tumor cells, *Cell,* Vol. 87-120 (1996)
2. C.B. Thompson et al., Pathways of apoptosis in lymphocyte development, homeostasis and disease, *Cell,* 109 Suppl, S97-107.
3. John K. Heath, *Principles of cell proliferation,* Chapters 1, 8, 9.
4. Motoshi Suzuki et al., Structure of Bax: Coregulation of Dimer Formation and Intracellular Localization, *Cell,* Vol. 103, 645–654, November 10, 2000.
5. S. Cori et al., The Bcl-2 family regulators of the cellular life-or-death switch, *Nat. Re. Cancer,* 2002, Vol. 2, pp. 647-656.
6. Mark G.Hinds, Martin Lackmann et al., The structure of Bcl-w reveals a role for the C-terminal residues in modulating biological activity, *The EMBO Journal,* 2003, Vol.22, No.7, 1497-1507.
7. Motoshi Suzuki et al., Structure of Bax:coregulation of dimmer formation and intracellular localization, *Cell,* November 10, 2000, Vol.103, *645-654.*
8. Michael Sattler, Heng Liang et al., Structure of Bcl-xL –Bak peptide complex: recognition between regulators of apoptosis, *Science,* 14 February 1997, Vol. 275, 983-986.
9. Andrew M. Petros et al., Structural biology of the Bcl-2 family of proteins, *Biochimica et Biophysica Acta,* Vol.1644, pp. 83–94, 2004.
10. Anna Schinzel et al., Bcl-2 family members: intracellular targeting, membrane-insertion, and changes in subcellular localization, Biochimica et Biophysica Acta, Vol. 1644, pp. 95–105, 2004.
11. Xinqi Liu et al.,The structure of a Bcl-xL –Bim fragment complex; implications for Bim function, *Immunity,* September, 2003, Vol.19, 341-352.
12. Ameeta Kelekar and Craig B. Thompson. Bcl-2-family proteins: the role of the BH3 domain in apoptosis. *Cell Biology,* August , 1998, Vol.8, 324-330.
13. Jerry M. Adams and Suzanne Cory, The Bcl-2 Protein Family: Arbiters of Cell Survival, *Science,* Vol.281, pp.1322-1326, 28 August 1998.

14. S. Mazel et al., Regulation of Cell Division Cycle Progression by *bcl-2* Expression: A Potential Mechanism for Inhibition of Programmed Cell Death, *J. Exp. Med.*, Vol.**183**, pp. 2219-2226, May 1996.

15. Tohru Ohmori et al., Apoptosis of lung cancer cells caused by some anti-cancer agents (MMC, CPT-11 and ADM) is inhibited by Bcl-2, *Biophysical and biomedical communications*, Vol.**192**, No.1, pp. 30-36, 15 April 1993.

16. Andrew M. Petros et al., Discovery of a Potent Inhibitor of the Antiapoptotic Protein Bcl-xL from NMR and Parallel Synthesis, *J. Med. Chem.*, Vol.**49**, pp. 656-663, 2006.

17. Atsushi Suzuki and Yumi Tsutomi, Bcl-2 accelerates the neuronal differentiation: new evidence approaching to the biofunction of Bcl-2 in the neuronal system, *Brain Research*, Vol. **801**, pp. 59-66, 1998.

18. Cory, S. et al., The Bcl-2 family: Roles in cell survival and oncogenesis. *Oncogene*, Vol. **22**, pp. 8590-8607, 2003

19. Kitrkin, V. et al., The role of Bcl-2 family members in tumorigenesis, *Biochim. Biophys. Acta*, Vol. **1644**, pp. 229-249, 2004.

20. Juin, P. et al., Shooting at survivors: Bcl-2 family members as drug targets for cancer, *Biochim. Biophys. Acta*, Vol.**1644**, pp. 251-260, 2004.

21. Xingming Deng et al., Mono- and multisite phosphorylation enhances Bcl2's antiapoptotic function and inhibition of cell cycle entry functions, *PNAS*, Vol.**101**, No.1, pp. 153-158, 6 January, 2004.

22. Brian S.Chang et al., Identification of a novel regulatory domain in Bcl-xL and Bcl-2, *The EMBO Journal*, Vol.**16**, No.5, pp.968-977, 1997.

23. Shujie Wang et al., Loss of the Bcl-2 Phosphorylation Loop Domain Increases Resistance of Human Leukemia Cells (U937) to Paclitaxel-Mediated Mitochondrial Dysfunction and Apoptosis, *Biochemical and Biophysical Research Communications*, Vol. **259**, pp. 67-72, 1999.

24. Diane J. Rodi et al., Screening of a Library of Phage-displayed Peptides Identifies Human Bcl-2 as a Taxol-binding Protein, *J. Mol. Biol.*, Vol.**285**, pp.197-203, 1999.

25. R. Kroning and A. Lichtenstein, taxol can induce phosphorylation of Bcl-2 in multiple myeloma cells and potentiate dexamethasone-induced apoptosis, *Leukemia Research*, Vol.**22**, No.3, pp. 275-286, 1998.

26. Jan-Lun Wang et al. Structure-based discovery of an organic compound that binds Bcl-2 protein and induces apoptosis of tumor cells, *PNAS*, June 20, 2000, Vol. **97**, *No.13, 7124-7128*.

27. Degterev A. et al. Identification of small-molecule inhibitors of interaction between the BH3 domain and Bcl-xL. *Natural Cell Biology*, 2001, **3**, 173-182.

28. Lugovskoy A. et al. A novel approach for characterizing protein ligand complexes: molecular basis for specificity of small-molecule Bcl-2 inhibitors. *J.Am.Chem.Soc.*, 2002, **124**, 1234-1240.

29. Shie-Pon Tzung et al.. Antimycin A mimics a cell-death-inducing Bcl-2 homology domain 3. *Nature Cell Biology*, February, 2001, Vol.**3**, 183-191.

30. Istvan J. Enyedy et al. Discovery of small-molecule inhibitors of Bcl-2 through structure-based computer screening. *Journal of Medicinal Chemistry*, 2001, Vol.**44**, No. 25, 4313-4324.

31. Olaf Kutzki et al. Development of a potent Bcl-xL antagonist based on α-helix mimicry. *Journal of American Chemical Society*, 2002, Vol.**124**, 11838-11839.

32. Maryanne C.S. Herzig et al. Apoptosis induction by the dual action DNA- and protein-reactive antitumor drug irofulven is largely Bcl-2 independent. *Biochemical Pharmacology*, 2003, **65**, 503-513.

33. Barbara Becattini et al., Rational Design and Real Time, In-Cell Detection of the Proapoptotic Activity of a Novel Compound Targeting Bcl-XL, *Chemistry & Biology*, Vol. **11**, pp. 389-395, March, 2004.

34. Garbiñe Roy et al., Mechanistic Aspects of the Induction of Apoptosis by Lauryl Gallate in the Murine B-Cell Lymphoma Line Wehi 231, *Archives of Biochemistry and Biophysics*, Vol. **383**, No. 2, pp. 206-214, November 15, 2000.

35. Marilisa Leone et al., Cancer Prevention by Tea Polyphenols Is Linked to Their Direct Inhibition of Antiapoptotic Bcl-2-Family Proteins, *Advances in Brief, Cancer Research*, **63**, pp. 8118-8121, December 1, 2003.

36. Tripos SYBYL 7.0; Tripos Inc., 1699 South Hanley Rd, St. Louis, MO 63144, USA. http://www.tripos.com.

37. Molecular Operating Environment (MOE 2004.03). Chemical Computing Group, Inc. Montreal, Quebec, Canada. http://www.chemcomp.com. 37a. Code "more_dock.svl" obtained from SLV Exchange website http://svl.chemcomp.com., Chemical Computing Group, Inc., Montreal, Canada 37b Code "scoring.svl" obtained from SLV Exchange website http:// svl.chemcomp.com., Chemical Computing Group, Inc., Montreal, Canada.

38. a) R. A. Laskowski et al., PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Cryst.*, Vol.**26**, pp. 283-291, 1993. b) A L Morris et al., Stereochemical quality of protein structure coordinates,*Proteins*,Vol.**12**,pp.345-364,1992.
http://www.biochem.ucl.ac.uk/~roman/procheck/procheck.html

39. Paul Bamborough and Fred E. Cohen, Modelling protein-ligand complexes, *Current opinion in structural biology*, pp.1-8, April 1996.

40. Ajay and Mark A. Murcko, Computational Methods to predict binding free energy in ligand-receptor complexes, *Journal of Medicinal Chemistry*, Vol.**38**, No.26, pp. 4953-4967, 1995.

41. Smith and Williams, *Introduction to the Principles of Drug Design and Action*, 3rd Edition, Chapters 3, 9.

42. Renxiao wang et al., LigBuilder: a multi-purpose program for structure-based drug design, *Journal of Molecular Modelling*, Vol.**6**, pp. 498-516, 2000.

43. G.Moont, G.R. Smith et al., 3D-Dock- incorporating FTDock (version 2.0), RPScore, and Multidock, pp. 1-20, March 2001.

44. Thomas R. Gadek and John B.Nicholas, Small molecule antagonist of proteins, *Biochemical Pharmacology*, Vol.65, pp. 1-8, 2003.
45. Xingming Deng et al., Mono- and multisite phosphorylation enhances Bcl-2's antiapoptotic function and inhibition of cell cycle entry functions, *PNAS*, Vol.101, No.1, pp. 153-158, January 6, 2004.
46. Diane J. Rodi et al., Screening of a library of phage–displayed peptides identifies human Bcl-2 as taxol-binding protein, *Journal of Molecular Biology*, Vol.285, pp.197-203, 1999.
47. a) H. Bekker, H.J.C. Berendsen, E.J. Dijkstra, S. Achterop, R. van Drunen, D. van der Spoel, A. Sijbers, H. Keegstra, B. Reitsma and M.K.R. Renardus: *Gromacs: A parallel computer for molecular dynamics simulations.* In Physics Computing 92 (Singapore, 1993). R.A. de Groot and J. Nadrchal, eds. . World Scientific. b) H.J.C. Berendsen, D. van der Spoel and R. van Drunen: *GROMACS: A message-passing parallel molecular dynamics implementation* Comp. Phys. Comm., **91**, pp. 43-56, (1995). c) E. Lindahl, B. Hess and D. van der Spoel: *GROMACS 3.0: A package for molecular simulation and trajectory analysis.* J. Mol. Mod., 7, pp. 306-317 (2001). d) D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. C. Berendsen: *GROMACS: Fast, Flexible and Free*, J. Comp. Chem., **26**, pp. 1701-1718 (2005). e) http://www.gromacs.org/
48. Wei Hu, John J Kavanagh, Anticancer therapy targeting the apoptotic pathway, *Lancet Oncol*, 2003; **4**: 721–29.
49. http://www.rcsb.org/pdb/searchlite.htmL
50. Xingming Deng et al., Mono- and multisite phosphorylation enhances Bcl-2's antiapoptotic function and inhibition of cell cycle entry functions, *PNAS*, January 6, 2004, Vol. **101**, No.1,153-158.

Chapter 5

1. Andrei M. Petro set al., Defining the p53 DNA-binding domain/Bcl-xL-binding interface using NMR, *FEBS Letters*, Vol. **559**, pp. 171-174, 2004.
2. Mihara, M. et al., *Mol. Cell*, Vol. **11**, pp. 577-590, 2003.
3. Paul H. Kussie, Svetlana Gorina et al., Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain, *Science*, 8 November 1996, Vol. **274**, Issue 5289, pp. 948-953.
4. Lyubomir T. Vassilev, Binh T. Vu et al., In vivo activation of the p53 pathway by small-molecule antagonists of MDM2, *Science*, 6 February 2004 Vol.**303**.
5. Sara J. Duncan, Sabine Gru Schow et al., Isolation and structure elucidation of Chlorofusin, a novel p53-MDM2 antagonist from a Fusarium sp., *J. Am. Chem. Soc.*, 2001, **123**, pp. 554-560.
6. Jianhua Zhao, Mijuan Wangb et al., The initial evaluation of non-peptidic small-molecule HDM2 inhibitors based on p53–HDM2 complex structure, *Cancer Letters*, (2002), Vol. **182**, pp. 69–77.
7. Jerry E. Chipuk, Tomomi Kuwana et al., Direct activation of Bax by p53 mediates mitochondrial membrane permeabilization and apoptosis, *Science*, 13 February 2004, Vol. **303**.
8. J.I-Ju Leu, Patrick Dumont et al., Mitochondrial p53 activates Bak and causes disruption of a Bak–Mcl1 complex, *Nature Cell Biology*, May 2004, Vol. **6**, No.5.
9. Thomas G. Hofmann, Andreas Möller et al., Regulation of p53 activity by its interaction with homeo-domain-interacting protein kinase-2, *Nature Cell Biology*, January 2002, Vol. **4**.
10. http://www.rcsb.org/pdb/searchlite.html
11. Molecular Operating Environment (MOE). Chemical Computing Group, Inc. Montreal, Quebec, Canada. http://www.chemcomp.com
12. Tripos SYBYL 7.0; Tripos Inc., 1699 South Hanley Rd, St. Louis, Missouri 63144, USA..http://www.tripos.com.
13. Lyubomir T. Vassilev, Binh T. Vu et al., In vivo activation of the p53 pathway by small-molecule antagonists of MDM2, *Science*, 6 February 2004 Vol.**303**.

Chapter 6

1. Molecular Operating Environment (MOE). Chemical Computing Group, Inc. Montreal, Quebec, Canada. http://www.chemcomp.com
2. Tripos SYBYL 7.0; Tripos Inc., 1699 South Hanley Rd, St. Louis, Missouri 63144, USA..http://www.tripos.com.
3. G. De Martino, G. La Regina, A. Coluccia, M.C. Edler, **M. Chiara Barbera**, A. Brancale, E. Wilcox, E.
4. Hamel, M. Artico, and R. Silvestri, Arylthioindoles, potent inhibitors of tubulin polymerization, *Journal of Medicinal Chemistry*, **2004**, 47, 6120-6123.
5. G. De Martino, M.C. Edler, G. La Regina, A. Coluccia, **M. Chiara Barbera**, D. Barrow, R. I. Nicholson, G. Chiosis, A. Brancale, E. Hamel, M. Artico, and Romano Silvestri, New Arylthioindoles: Potent Inhibitors of Tubulin Polymerization. 2. Structure-Activity Relationships and Molecular Modeling Studies, *Journal of Medicinal Chemistry*, ASAP Article 10.1021/jm050809s S0022-2623(05)00809-5 (Web Release Date: January 19, 2006).
6. Ravelli, R. B. G.; Gigant, B.; Curmi, P. A.; Jourdain, I.; Lachkar, S.; Sobel, A.; Knossow, M. Insight into tubulin regulation from a complex with colchicine and a stathmin-like domain. *Nature* **2004**, *428*, 198-202.
7. Code "scoring.svl" obtained from SLV Exchange website http://svl.chemcomp.com., Chemical Computing Group, Inc., Montreal, Canada.

Bibliography

Bibliography

Chapter 1

G.Patrick , *Instant notes – Medicinal Chemistry*.

D.A. Williams and T.L. Lemke, *Foye's principle of medicinal chemistry*, 5th Edition.

F.D. King, *Medicinal chemistry – principles and practice*, 2nd Edition.

G. Thomas, *Fundamentals of medicinal chemistry*, 2004.

G. Krauss, *Biochemistry of signal transduction and regulation*, 3rd Edition, Wiley-Vch.

John K. Heath, *Principles of cell proliferation*.

F. Gualtieri, *New trends in synthetic medicinal chemistry*,Volume 7.

Gisbert Schneider and Uli Fechner, Computer-based *De Novo* design of drug-like molecules, *Nature Reviews, Drug Discovery*, Vol.4, pp. 649-663, August 2005.

Florence L. Stahura et al., Methods for compound selection focused on hits and application in drug discovery, *Journal of Molecular Graphics and Modelling*, Vol.20, pp. 439–446, 2002.

N. Claude Cohen, *Guidebook on Molecular Modelling in Drug Design*, Academic Press.

H.- D. Holtje, W. Sippl, D. Rognan, G. Folkers, *Molecular modelling – basic principles and applications*, 2nd Edition

Santosh Putta et al., Conformation mining: an algorithm for finding biologically relevant conformations, *J. Med. Chem.*, Vol.48, pp. 3313-3318, 2005.

Burger. *Molecular Modelling in Drug Design. Medicinal Chemistry & Drug Design*, 2003, Vol.1, 79, 123, pp. 113-115.

Andrew Leach. *The use of molecular modelling and chemoinformatics to discover and design new molecules.*

Molecular Modelling – principles and applications. Second edition. Chapter 12, pp. 641-719.

Smith and Williams, *Introduction to the Principles of Drug Design and Action*, 3rd Edition, Chapter 3, 52-80, Chapter 9, pp. 333-385.

Jianfeng Pei et al., Improving the quality of 3D-QSAR by using Flexible-Ligand receptor models, *J. Chem. Inf. Model.*

Jeffrey J. Sutherland et al., A comparison of methods for modelling Quantitative Structure-Activity Relationships, *J. Med. Chem.*, Vol.47, pp. 5541-5554, 2004.

Casey E. Bohl et al., A Ligand-Based approach to identify Quantitative Structure-Activity Relationships for the androgen receptor, *J. Med. Chem.*, Vol.47, pp.3765-3776, 2004.

Jaroslaw Polanski et al., Modelling robust QSAR, *J. Chem. Inf. Model.*

Rajarshi Guha and Peter C. Jurs, Determining the validity of a QSAR model - a classification approach, *J. Chem. Inf. Model.*, Vol. 45, pp. 65-73, 2005.

Sabcho Dimitrov et al., A stepwise approach for defining the applicability domain of SAR and QSAR Models, *J. Chem. Inf. Model.*, Vol. 45, pp. 839-849, 2005.

Ling Xue and Jürgen Bajorath, Molecular descriptors for effective classification of biologically active compounds based on principal component analysis identified by a genetic algorithm, *J. Chem. Inf. Comput. Sci.*, Vol.40, pp. 801-809, 2000.

A. Yasri and D. Hartsough, Toward an optimal procedure for variable selection and QSAR model building, *J. Chem. Inf. Comput. Sci.*, Vol.41, pp. 1218-1227, 2001.

Hetal Mishra et al., Three-dimensional quantitative structure-activity relationship and comparative molecular field analysis of dipeptide hydroxamic acid *Helicobacter pylori* urease inhibitors, *Antimicrobial agents and chemotherapy*, Vol. 46, No. 8, pp. 2613–2618, August 2002.

Alexander Böcker et al., A hierarchical clustering approach for large compound libraries, *J. Chem. Inf. Model.*, Vol.**45**, pp. 807-815, 2005.

Mary Pat Beavers and Xin Chen, Structure-based combinatorial library design: methodologies and applications, *Journal of Molecular Graphics and Modelling*, Vol.**20**, pp. 463–468, 2002.

Andreas Evers and Thomas Klabunde, Structure-based drug discovery using GPCR Homology Modelling: successful virtual screening for antagonists of the alpha1A adrenergic receptor, *J. Med. Chem.*, Vol.**48**, pp. 1088-1097, 2005.

Paul Bamborough and Fred E. Cohen, Modelling protein-ligand complexes, *Current opinion in structural biology*, April 1996, pp. 1-8.

Ajay and Mark A. Murcko, Computational Methods to predict binding free energy in ligand-receptor complexes, *Journal of Medicinal Chemistry*, 1995, Vol. **38**, No.26, pp. 4953-4967.

Graham R Smith and Michael J.E. Sternberg, Prediction of protein-protein interactions by docking methods, *Current Opinion in structural Biology*, 2002, 12, pp. 28-35.

Henry A. Gabb, Michael J.E.Sternberg et al., Modelling protein docking using shape complementarity, electrostatics and biochemal information, Journal of Molecular Biology, 1997, Vol. **272**, pp. 106-120.

Vladimir Maiorov and Robert P. Sheridan, Enhanced virtual screening by combined use of two docking methods: getting the most on a limited budget, *J. Chem. Inf. Model., Vol.45, pp.* 1017-1023, 2005.

Maxwell D. et al., Comparison of automated docking programs as virtual screening tools, *J. Med. Chem.*, Vol.**48**, pp. 962-976, 2005.

Jinn-Moon Yang et al., Consensus scoring criteria for improving enrichment in virtual screening, *J. Chem. Inf. Model.*

J. Christian Baber et al.,The use of consensus scoring in ligand-based virtual screening, *J. Chem. Inf. Model.*
Angelo Vedani et al., Combining protein modelling and 6D-QSAR. Simulating the binding of structurally diverse ligands to the estrogen receptor, *J. Med. Chem., Vol.48, pp.* 3700-3703, 2005.

Gregory L. Warren et al., A Critical assessment of docking programs and scoring functions, *J.Med.Chem.*

Jean-Christophe Mozziconacci et al., Optimization and validation of a docking-scoring protocol; application to virtual screening for COX-2 inhibitors, *J. Med. Chem.*, Vol.**48**, pp. 1055-1068, 2005.

Ian L. Alberts et al., Receptor flexibility in *De Novo* ligand design and docking.

Yukihisa S. Watanabe et al., Free energy landscapes of small peptides in an implicit solvent model determined by force-biased multicanonical molecular dynamics simulation, *Chemical Physics Letters*, Vol. **400**, pp. 258–263, 2004.

Lillian T. Chong et al., Dimerization of the p53 oligomerization domain: identification of a folding nucleus by Molecular Dynamics Simulations, *J. Mol. Biol.*, Vol.**345**, pp. 869–878, 2005.

Chapter 2

Lubert Stryer, *Biochemistry,* 4th Edition.

H.Campbell, *Biology*, 5th Edition

H.P. Rang, M.M. Dale, J.M. Ritter, *Pharmacology*, 5th edition, Chapter 50.

Goodman, *Louis S. Goodman & Gilman's the pharmacological basis of therapeutics*, 9th Edition, Chapter 51.

David A.Williams and Thomas L.Lemke, *Foye's principles of Medicinal Chemistry*, 5th edition.

M.C. Wani et al., The isolation and structure of taxol, a novel antileukemic and antitumor agent from Taxus brevifolia, *J. Am. Chem. Soc.*, 1971, Vol.**93**, pp.2325-2327.

E.K. Rowinsky et al., Sequences of taxol and cisplatin: a phase I and a pharmacological study, *J. Clin., Oncol.*, 1991, Vol.**9**, pp.1692-1703.

K.C. Nicolaou et al., Total synthesis of taxol, *Nature,* 1994, Vol.**47**, pp. 632-634.

Chapter 3

Francis V. Chiari,· Unscrambling hepatitis C virus–host interactions, *Nature Insight*, Vol.**436**, pp.930-932, 18 August 2005.

Ursula Weiss, Hepatitis C, *Nature Insight*, Vol.**436**, Issue no. 7053, 18 August 2005.

Darius Moradpouret al., Hepatitis C: molecular virology and antiviral targets, *TRENDS in Molecular Medicine*, Vol.**8**, No.10, pp. 476-482, October 2002.

Jonathan M Caruthers and David B McKay, Helicase structure and mechanism, *Current Opinion in Structural Biology*, Vol.12, pp. 123–133, 2002.

Peter Borowski et al.,, Nucleotide triphosphatase/helicase of hepatitis C virus as a target for antiviral therapy, *Antiviral Research*, Vol.**55**, pp. 397-412, 2002.

Christopher P. Gordon and Paul A. Keller, Control of Hepatitis C: A Medicinal Chemistry Perspective, *J.Med.Chem.*, Vol.**48**, No.1, pp. 1-20, 13 January, 2005.

Seng-Lai Tanet al., Hepatitis C therapeutics: current status and emerging strategies, *Nature Reviews, Drug Discovery*, Vol.1, pp. 867-881, November 2002.

Hugo R. Rosen and David R. Gretch, Hepatitis C virus: current understanding and prospects for future therapies, *Molecular Medicine Today*, Vol.5, pp.393-399, September 1999.

Brett D. Lindenbach and Charles M. Rice, Unravelling hepatitis C virus replication from genome to function, *Nature Insight*, Vol. 436, pp.933 – 938, 18 August 2005.

R. De Francesco and G. Migliaccio, Challenges and successes in developing new therapies for hepatitis C, *Nature Insight*, Vol.**436**, pp. 953-960, 18 August 2005.

Chapter 4

M.D. Jacobson and N. McCarthy, *Apoptosis, the molecular biology of programmed cell death*, Oxoford University Press, 2002, Chapters 1, 5, 9, 11.

G. Krauss, *Biochemistry of signal transduction and regulation*, 3rd Edition, Wiley-Vch, Chapters 6, 14, 15.

John K. Heath, *Principles of cell proliferation*, Chapters 1, 8, 9.

C.B. Thompson et al., Pathways of apoptosis in lymphocyte development, homeostasis and disease, *Cell*, 109 Suppl, S97-107.

Chapter 5

Peter L. Toogood. Inhibition of protein-protein association by small molecules: approaches and progress. *Journal of Medicinal Chemistry*, 2002, Vol.45, No.8, 1543-1558.

Ziwei Huang. The chemical biology of apoptosis: exploring protein-protein interactions and the life and death of cells with small molecules. *Chemistry & Biology*, October, 2002, Vol.9, 1059-1072.

M.R.Arkin and J.A. Wells. Small molecule inhibitors of protein-protein interactions: progressing towards the dream, *Nature*, Vol. 3, April 2004, pp. 301-317.

M.D. Jacobson and N. McCarthy, *Apoptosis, the molecular biology of programmed cell death*, Oxoford University Press, 2002, Chapters 1, 5, 9, 11.

G. Krauss, *Biochemistry of signal transduction and regulation*, 3rd Edition, Wiley-Vch, Chapters 6, 14, 15.
John K. Heath, *Principles of cell proliferation*, Chapters 1, 8, 9.

Glossary of
Abbreviations

Glossary of Abbreviations

AcOEt	Ethyl acetate
Arg	Arginine
BOP	Benzotriazol-1-yloxytris(dimethylamino)-phosphonium hexafluorophosphate
CHCl$_3$	Chloroform
DCC	Cyclohexanamine,N,N'-methanetetraylbis
DCM	Dichloromethane
DEE	Diethyl ether
DENV	Dengue virus
DHF	Dengue haemorrhagic fever
DIPEA	N,N'-diisopropylethylamine
DMAP	4-Dimethylaminopyridine
DMF	Dimethyl formammide
DMSO	Dimethylsulphoxide
DNA	Deoxyribonucleic acid
dsDNA	Double stranded deoxyribonucleic acid
DSS	Dengue shock syndrome
EC50	50% Effective concentration
HCl	Hydrochloric acid
HCV	Hepatitis C virus
IC50	50% Inhibition concentration
IRES	Internal ribosomal entry site
IV	Intravenous
LogP	Partition coefficient (Oil/Water)
M	Molar
MOPS	4-Morpholinepropanesulfonic acid
NMR	Nuclear magnetic resonance
RdRp	RNA-dependant RNA-polymerase
RNA	Ribonucleic acid
ssDNA	Single stranded deoxyribonucleic acid
THF	Tetrahydrofuran
TLC	Thin layer chromatography

Arylthioindoles, Potent Inhibitors of Tubulin Polymerization

Gabriella De Martino,[†] Giuseppe La Regina,[†]
Antonio Coluccia,[†] Michael C. Edler,[‡]
Maria Chiara Barbera,[§] Andrea Brancale,[§]
Elizabeth Wilcox,[‡] Ernest Hamel,[‡]
Marino Artico,[†] and Romano Silvestri*,[†]

Dipartimento di Studi Farmaceutici, Università di Roma
"La Sapienza", Piazzale Aldo Moro 5, I-00185 Roma, Italy,
Welsh School of Pharmacy, Cardiff University,
King Edward VII Avenue, Cardiff, CF10 3XF, U.K., and
Screening Technologies Branch, Developmental Therapeutics
Program, Division of Cancer Treatment and Diagnosis,
National Cancer Institute at Frederick, National Institutes
of Health, Frederick, Maryland 21702

Abstract: Several arylthioindoles had excellent activity as inhibitors both of tubulin polymerization and of the growth of MCF-7 human breast carcinoma cells. Methyl 3-[(3,4,5-trimethoxyphenyl)thio]-5-methoxy-1H-indole-2-carboxylate (21), the most potent derivative, showed $IC_{50} = 2.0 \, \mu M$, 1.6 times more active than colchicine and about as active as combretastatin A-4 (CSA4). Compound 21 inhibited the growth of the MCF-7 cells at $IC_{50} = 13$ nM. Colchicine and CSA4 had 13 nM and 17 nM IC_{50} values, respectively, with these cells.

Microtubules are involved in a wide number of cellular functions, such as motility, division, shape maintenance, and intracellular transport. The major protein component found in microtubules is tubulin. Interference with microtubule assembly, either by inhibition of tubulin polymerization or by blocking microtubule disassembly, leads to an increase in the number of cells in metaphase arrest. Inhibition of microtubule function using tubulin targeting agents is a validated approach to anticancer therapy.[1–5]

Many well-described antimitotic agents are derived from natural sources. Colchicine (1) and *Vinca* alkaloids, the first tubulin binding agents discovered, cause destabilization of microtubules and, as a consequence, induce the cell to undergo apoptosis.[6] On the other hand paclitaxel, a standard antitumor agent originally extracted from *Taxus brevifolia*, binds to a different site within tubulin leading to the stabilization of microtubules.[7] Unfortunately, clinical use of these compounds is restricted by toxicity, drug resistance, complex formulations, and limited bioavailability.

Various synthetic molecules have also been reported as tubulin inhibitors. The majority of them inhibit the binding of colchicine to tubulin and thereby act as destabilizing agents. In recent years about 30 natural, semisynthetic, or synthetic tubulin binding agents, acting by either stabilizing or destabilizing mechanisms, have been under preclinical/clinical investigation.[2]

* Corresponding author. Phone: +39 06 4991 3800. Fax: +39 06 491 491. E-mail: romano.silvestri@uniroma1.it.
† Università di Roma.
‡ National Cancer Institute.
§ Cardiff University.

Chart 1

colchicine (1)

3a R=H,
3b R=OCH$_3$

combretastatin A-4 (6)
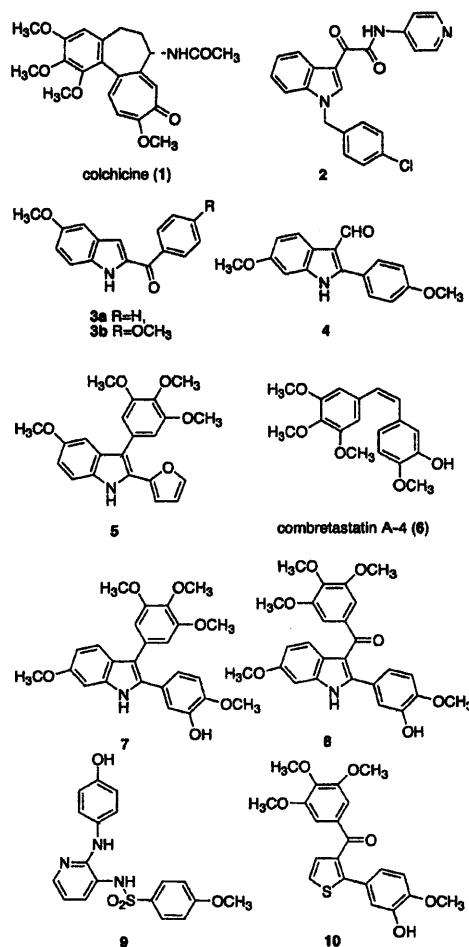
The indole nucleus is the core structure of a great number of tubulin polymerization inhibitors.[2,8] The indolyl-3-glyoxamide D-24851 (2) and the 2-aroylindoles D-64131 (3a) and D-68144 (3b) were discovered by Baxter Oncology. Compounds 3 are highly active against various tumors including those resistant to paclitaxel[9] (Chart 1).

Several 2-phenylindoles were designed by von Angerer as simple analogues of 12-formyl-5,6-dihydroindolo[2,1-a]isoquinoline. Among them, indole 4 completely blocked microtubule assemby at a concentration of 40 μM.[10] On the basis of the structure of the natural product combretastatin A-4 (CSA4, 6), some 2,3-diarylindoles, known as heterocombretastatins, were prepared (e.g., 5) by Medarde.[11] Flynn et al. reported the tubulin polymerization inhibitory activity of 2,3-diarylindole 7 and 2-aryl-3-arylcarbonylindole 8.[12]

Sulfur containing compounds, such as sulfonamide E-7010 (9),[13] thiophene 10,[14] and benzothiophene[12] derivatives, proved effective inhibitors of tubulin polymerization. To our knowledge there have been no reports on the inhibition of tubulin polymerization by arylthio/sulfonylindoles.

Table 2. Inhibition of Tubulin Polymerization, Colchicine Binding and Growth of MCF-7 Human Breast Carcinoma Cells by Compounds 11−21

compd	tubulin[a] $IC_{50} \pm SD$ (μM)	colchicine binding[b] (% $\pm$ SD)	MCF-7[c] $IC_{50} \pm SD$ (nM)
11	15 ± 0.7	13 ± 5	>2500
12	8.3 ± 0.6	21± 7	>2500
13	>40	5.1 ± 1	>2500
14	2.9 ± 0.1	81 ± 1	25 ± 1
15	4.4 ± 0.3	19 ± 7	>1250
16	4.4 ± 0.2	33 ± 0.6	>1000
17	>40	3.6 ± 8	>2500
18	2.5 ± 0.3	76 ± 0.2	42 ±10
19	>40	1.6 ± 2	>2500
20	6.2 ± 0.3	31 ± 8	>2500
21	2.0 ± 0.2	93 ± 0.8 (68 ± 3)[d]	13 ± 3
1 (colchicine)	3.2 ± 0.4		13 ± 3
4[e]	nd[g]	nd	160
6 (CSA4)	2.2 ± 0.2	100 ± 0.9 (88 ± 0.4)[d]	17 ± 10
7[f]	4.1 ± 0.6	28 ± 8	370 ± 2
8[f]	1.6	54 ± 0.7	45 ± 5

[a] Inhibition of tubulin polymerization. Tubulin was at 10 μM.[13] [b] Inhibition of [^{3}H]colchicine binding. Tubulin was at 1 μM; both [^{3}H]colchicine and inhibitor were at 5 μM.[14] [c] Inhibition of growth of MCF-7 human breast carcinoma cells.[14] [d] Inhibition of [^{3}H]colchicine binding. [^{3}H]Colchicine was at 5 μM; both tubulin and inhibitor were at 1 μM. [e] Reference 6. [f] Reference 8. [g] No data.

inhibitor of tubulin polymerization (IC_{50} = 2.5 μM), whereas inhibition of the growth of MCF-7 cells (IC_{50} = 42 nM) was decreased almost 2-fold. In contrast, the presence of a methoxy group at position 5 of the indole, giving methyl 3-[(3,4,5-trimethoxyphenyl)thio]-5-methoxy-1*H*-indole-2-carboxylate (**21**), enhanced inhibitory potency in both the biochemical and cytological assays. Compound **21** (IC_{50} = 2.0 μM) was 1.6 times more active than colchicine and about as active as CSA4 as an inhibitor of tubulin polymerization, while as an inhibitor of the growth of MCF-7 cells (IC_{50} = 13 nM) it was as potent as colchicine and CSA4. Among the derivatives tested, this compound was the most active.

It is interesting to note that, independently of the presence/nature of the substituent at position 5 of the indole, the 3,4,5-trimethoxyphenylthio moiety was crucial for effective inhibition of the growth of MCF-7 cells (compare **12** with **14**, **17** with **18**, and **20** with **21**).

In contrast, inhibition of tubulin polymerization was less affected by the substituent at position 5 (compare **12** with **14** and **20** with **21**). Surprisingly, methyl 3-[(4-methoxyphenyl)thio]-5-chloro-1*H*-indole-2-carboxylate (**17**) was quite inactive in both assays.

To investigate the possible binding mode of these novel compounds, we carried out docking studies of the arylthioindole **21** in the colchicine binding site of tubulin, which was obtained from the recently reported 3D structure of tubulin cocrystallized with a colchicine analogue: *N*-deacetyl-*N*-(2-mercaptoacetyl)colchicine (DAMA-colchicine).[19]

Compound **21** was docked using a modified version of the docking tool of MOE,[20] which implements a genetic algorithm based search method.[21] The resulting protein/ligand complex was then minimized, revealing a very interesting and efficient binding model for the inhibitor to tubulin.

Figure 1 clearly shows how the trimethoxy ring is well situated in proximity to Cys241 (residue numbers as in
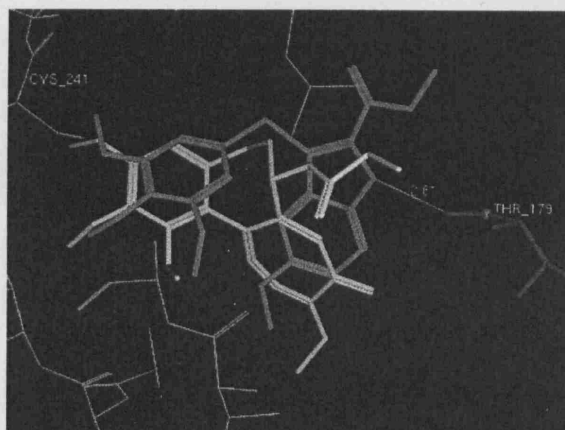
Figure 1. Proposed binding of **21**. DAMA-colchicine is represented in yellow, compound **21** in green. In red are represented the residues of the X-ray structure of tubulin, in cyan the same residues after minimization of tubulin with bound **21**. Residue numbers are those used by Ravelli et al.[19]

ref 19), adopting an orientation very similar to that of the corresponding ring in the DAMA-colchicine of the crystallized structure (see also Figure 2 in the Supporting Information). The methoxy substituent of the indole is also very close to the corresponding group on ring C of colchicine, leading to a very similar general binding of the two inhibitors. Furthermore, the indole moiety establishes one hydrogen bond between the NH and the backbone of Thr179 (shown in blue). Several other hydrophobic contacts (not shown for the sake of clarity) stabilize further the binding of **21** to the protein. These observations are consistent with the highly efficient inhibition of [^{3}H]colchicine binding that occurs with compound **21** (Table 2).

Further studies are currently underway to verify if this model can give a rational justification of the SAR obtained from the biological data.

In conclusion, we synthesized arylthioindoles, a novel class of inhibitors of tubulin polymerization. Some compounds had excellent activity in both inhibition of tubulin polymerization and inhibition of the growth of MCF-7 human breast carcinoma cells. Methyl 3-[(3,4,5-trimethoxyphenyl)thio]-5-methoxy-1*H*-indole-2-carboxylate (**21**), the most potent derivative, showed IC_{50} = 2.0 μM, 1.6 times more active than colchicine and about as active as CSA4 as an inhibitor of tubulin assembly.

This compound inhibited the growth of MCF-7 human breast carcinoma cells with IC_{50} = 13 nM, essentially equivalent to the values obtained with colchicine and CSA4.

A SAR study led to identification of several crucial structural requirements needed to enhance the effectiveness of the arylthioindoles. In particular, we found required determinants to be (i) the presence of a methoxy/ethoxycarbonyl group at position 2 of the indole nucleus; (ii) the 3,4,5-trimethoxyphenylthio group at position 3 of the indole; (iii) the sulfur atom in the sulfide oxidation state; (iv) the substituent at position 5 of the indole.

The last seems especially important for inhibition of the growth of MCF-7 cells. Furthermore, initial molecular modeling studies have revealed a possible binding mode for these novel inhibitors that could support the

Therefore, here we describe the synthesis of some novel 3-arylthio/sulfonylindole-2-carboxylates (12−21). Several were found to be highly potent inhibitors of tubulin polymerization.

3-Phenylthio-1*H*-indole (11) was synthesized by reacting indole (22) with 1,1'-diphenyldisulfide (29) in the presence of NaH in anhydrous DMF, following the procedure reported by Atkinson.[15] By this method was prepared in poor yield methyl 3-phenylthio-1*H*-indole-2-carboxylate (12) starting from methyl indole-2-carboxylate (23).

To improve the yield, compound 12 was synthesized by reacting ester 23 with *N*-phenyl-thiosuccinimide (30) in the presence of boron trifluoride diethyl etherate, as we previously reported for compound 15 starting from 24.[16]

By the same procedure were prepared methyl 3-[(4-methoxyphenyl)thio]-5-chloro-1*H*-indole-2-carboxylate (17) starting from methyl 5-chloro-1*H*-indole-2-carboxylate (25) and *N*-[(4-methoxyphenyl)thio]succinimide (31), and methyl 3-(phenylthio)-5-methoxy-1*H*-indole-2-carboxylate (20) starting from methyl 5-methoxyindole-2-carboxylate (26) and 30. Methyl 3-[(3,4,5-trimethoxyphenyl)thio]-1*H*-indole-2-carboxylate (14) and methyl 3-[(3,4,5-trimethoxyphenyl)thio]-5-chloro-1*H*-indole-2-carboxylate (18) were prepared by heating at 50 °C indole-2-carboxylic acid (27) or its 5-chloro derivative 28 with the 1,1'-(3,4,5-trimethoxyphenyl)disulfide (32) in the presence of NaH. The crude acids were then transformed into the corresponding methyl esters by reaction with trimethylsilyldiazomethane (TMSDM) at room temperature for 30 min. 5-Methoxyindole 21 was prepared following a two step procedure involving the addition of 3,4,5-trimethoxylthiophenol (33) to a solution of *N*-chlorosuccinimide (NCS) at −78 °C, and then this mixture was treated with acid 26 while the reaction temperature was warmed to 0 °C within 1 h. Oxidation of sulfur derivatives 12, 15, and 18 with 3-chloroperoxybenzoic acid (MCPBA) furnished the corresponding sulfones 13, 16, and 19, respectively (Scheme 1).

Inhibition of tubulin polymerization, colchicine binding,[17] and the growth of MCF-7 human breast carcinoma cells[18] by the novel indoles 11−21 in comparison with the effects of the reference compounds colchicine (1) and CSA4 (6) are shown in Table 2.

The compound we initially evaluated was 3-phenylthio-1*H*-indole (11), which inhibited tubulin polymerization with an IC$_{50}$ value of 15 μM, ca. 5 and 7 times inferior to the reference compounds 1 and 6, respectively.

We therefore attempted the introduction of a methoxycarbonyl function at position 2 of the indole ring. This chemical modification produced methyl 3-(phenylthio)-1*H*-indole-2-carboxylate (12), which was about twice as potent as 11 (IC$_{50}$ = 8.2 μM).

A further step was the oxidation of the sulfur atom of 12 to the sulfone to produce methyl 3-(phenylsulfonyl)-1*H*-indole-2-carboxylate (13). This compound was found to be inactive at the highest concentration tested.

However, a different behavior was observed for the analogous pair 15/16. Replacement of the 2-methoxycarbonyl group of 12 with an ethoxycarbonyl group (15) doubled the inhibitory potency of the parent compound, and the inhibitory activity of 15 was retained by the

Scheme 1[a]

compd	R$_1$	R$_2$	R$_3$	R$_4$	R$_5$	S/SO$_2$
11	H	H	H	H	H	S
12	COOCH$_3$	H	H	H	H	S
13	COOCH$_3$	H	H	H	H	SO$_2$
14	COOCH$_3$	OCH$_3$	OCH$_3$	OCH$_3$	H	S
15	COOC$_2$H$_5$	H	H	H	H	S
16	COOC$_2$H$_5$	H	H	H	H	SO$_2$
17	COOCH$_3$	H	OCH$_3$	H	Cl	S
18	COOCH$_3$	OCH$_3$	OCH$_3$	OCH$_3$	Cl	S
19	COOCH$_3$	OCH$_3$	OCH$_3$	OCH$_3$	Cl	SO$_2$
20	COOCH$_3$	H	H	H	OCH$_3$	S
21	COOCH$_3$	OCH$_3$	OCH$_3$	OCH$_3$	OCH$_3$	S

Table 1. Structure of Compounds 11−21
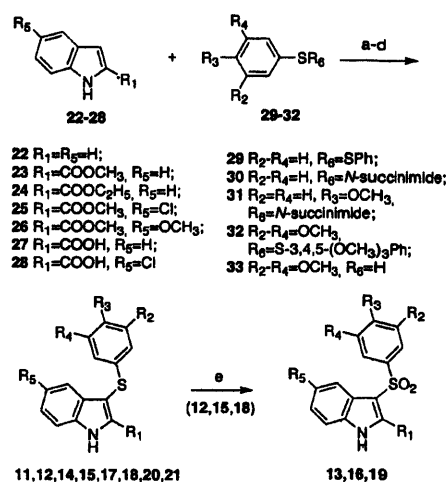
[a] Reagents and reaction conditions: (a, 11) ArSSAr, NaH, anhydrous DMF, rt, 2 h; (b, 12, 15, 17, 20) *N*-(ArS)succinimide, BF$_3$·Et$_2$O, anhydrous CH$_2$Cl$_2$, rt, 1.5 h, then 45 °C, 2 h; (c, 14, 18) (i) ArSSAr, NaH, anhydrous DMF, 50 °C, overnight, anhydrous nitrogen stream; (ii) TMSDM, CH$_3$OH/CH$_2$Cl$_2$, rt, 30 min; (d, 21) (i) ArSH, NCS, CH$_2$Cl$_2$, −78 °C; (ii) 26, −78 to 0 °C, 1 h; (e, 13, 16, 19) MCPBA (2.5 equiv), CHCl$_3$, rt, 1 h.

sulfone 16 (compare 12 with 15 and 16). Despite appreciable inhibition of tubulin polymerization displayed by indoles 11, 12, 15, and 16, these compounds were unable to inhibit the growth of MCF-7 human breast carcinoma cells.

At this point we decided to introduce a trimethoxy substitution pattern, a common structural feature of colchicine, CSA4, and other tubulin inhibitors (see Chart 1). Replacement of the phenylthio of 12 with the 3,4,5-trimethoxyphenylthio moiety furnished methyl 3-[(3,4,5-trimethoxyphenyl)thio]-1*H*-indole-2-carboxylate (14), which showed IC$_{50}$ = 2.9 μM, 2.8 times superior to that of 12 and comparable with those of colchicine (IC$_{50}$ = 3.2 μM) and CSA4 (IC$_{50}$ = 2.2 μM). Most importantly, this compound inhibited the growth of the MCF-7 cells by 50% at 25 nM, a concentration only 2 or 1.4 times higher than observed with colchicine (IC$_{50}$ = 13 nM) or CSA4 (IC$_{50}$ = 17 nM), respectively.

We also evaluated the effects of substituents at position 5 of the indole ring. To this end we chose two substituents with opposite properties, the electron withdrawing chlorine atom and the electron donating methoxy group. Introduction of a chlorine atom at position 5 of the indole of 14 gave methyl 3-[(3,4,5-trimethoxyphenyl)thio]-5-chloro-1*H*-indole-2-carboxylate (18), which only marginally increased potency as an

observed biological data. These findings have led us to continue our studies on novel arylthioindoles in order to investigate further SAR rules and mechanism of action and to evaluate their activity against a wider range of tumor cell lines.

Supporting Information Available: Experimental procedures for synthesis and biological evaluation of 11–21 and for molecular modeling. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Lin, M. C.; Ho, H. H.; Pettit, G. R.; Hamel, E. Antimitotic natural products combretastatin A-4 and combretastatin A-2: studies on the mechanism of their inhibition of the binding of colchicine to tubulin. *Biochemistry* **1989**, *28*, 6984–6991.

(2) Beckers, T.; Mahboobi, S. Natural, semisynthetic and synthetic microtubule inhibitors for cancer therapy. *Drugs Future* **2003**, *28*, 767–785 and references therein.

(3) Li, Q.; Sham, H. L. Discovery and development of antimitotic agents that inhibit tubulin polymerization for treatment of cancer. *Expert Opin. Ther. Pat.* **2002**, *12*, 1663–1702.

(4) Prinz, H. Recent advances in the field of tubulin polymerization inhibitors. *Expert Rev. Anticancer Ther.* **2002**, *2*, 695–708.

(5) Checchi, P. M.; Netlles, J. H.; Zhou, J.; Snyder, P.; Joshi, H. C. Microtubule-interacting drugs for cancer treatment. *Trends Pharmacol. Sci.* **2003**, *24*, 361–365.

(6) Pohle, K.; Matthies, E.; Peters, J. E. On the tumor growth-inhibiting action of colchicine. *Arch. Geschwulstforsh.* **1965**, *25*, 17–20.

(7) Schiff, P. B.; Fant J.; Horwitz, S. B. Promotion of microtubule assembly in vitro by Taxol. *Nature* **1979**, *277*, 665–667.

(8) Bacher, G.; Beckers, T.; Emig, P.; Klenner, T.; Kutshert, B.; Nickel, B. New small-molecule tubulin inhibitors. *Pure Appl. Chem.* **2001**, *73*, 1459–1464.

(9) http://www.baxter-oncology.com/english/projects/index.html.

(10) Gastpar, R.; Goldbrunner, M.; Marko, D.; von Angere, E. Methoxy-substituted 3-formyl-2-phenylindoles inhibit tubulin polymerizazion. *J. Med. Chem.* **1998**, *41*, 4965–4972.

(11) Medarde, M.; Ramos, A.; Caballero, E.; Peláz-Lamamié de Clairac, R.; López, J. L.; Garciá Grávalos, D. San Felicia, A. Synthesis and antineoplastic activity of combretastatin analogue: heterocombretastatin. *Eur. J. Med. Chem.* **1998**, *33*, 71–77.

(12) Flynn, B. L.; Hamel, E.; Jung, M. K. One-pot synthesis of benzo[b]furan and indole inhibitors of tubulin polymerization. *J. Med. Chem.* **2002**, *45*, 2670–2673.

(13) Yamamoto, K.; Noda, K.; Yoshimura, A.; Kukuoka, M.; Furuse, K.; Niitani, H. Phase I study of E7010. *Cancer Chem. Pharmacol.* **1998**, *42*, 124–134.

(14) Flynn, B. L.; Flynn, G. P.; Hamel, E.; Jung, M. K. The synthesis and tubulin binding activity of thiophene-based analogues of combretastatin A-4. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2341–2343.

(15) Atkinson, J. G.; Hamel, P.; Girard, Y. A. A new synthesis of 3-arylthioindoles. *Synthesis* **1988**, 480–481.

(16) Silvestri, R.; De Martino, G.; La Regina, G.; Artico, M.; Massa, S.; Vargiu, L.; Mura, M.; Loi, A. G.; Marceddu, T.; La Colla, P. Novel indolyl aryl sulfones active against HIV-1 carrying NNRTI resistance mutations: synthesis and SAR studies. *J. Med. Chem.* **2003**, *46*, 2482–2493.

(17) Hamel, E. Evaluation of antimitotic agents by quantitative comparisons of their effects on the polymerization of purified tubulin. *Cell Biochem. Biophys.* **2003**, *38*, 1–21.

(18) Verdier-Pinard, P.; Lai, J.-Y.; Yoo, H.-D.; Yu, J.; Marquez, B.; Nagle, D. G.; Nambu, M.; White, J. D.; Falck, J. R.; Gerwick, W. H.; Day, B. W.; Hamel, E. Structure-activity analysis of the interaction of curacin A, the potent colchicine site antimitotic agent, with tubulin and effects of analogs on the growth of MCF-7 breast cancer cells. *Mol. Pharmacol.* **1998**, *53*, 62–76.

(19) Ravelli, R. B. G.; Gigant, B.; Curmi, P. A.; Jourdain, I.; Lachkar, S.; Sobel, A.; Knossow, M. Insight into tubulin regulation from a complex with colchicine and a stathmin-like domain. *Nature* **2004**, *428*, 198–202.

(20) Molecular Operating Environment (MOE 2004.03). Chemical Computing Group, Inc., Montreal, Quebec, Canada. http://www.chemcomp.com.

(21) Code "more_dock.svl" obtained from SLV Exchange website http://svl.chemcomp.com., Chemical Computing Group, Inc., Montreal, Canada.

New Arylthioindoles: Potent Inhibitors of Tubulin Polymerization. 2. Structure–Activity Relationships and Molecular Modeling Studies

Gabriella De Martino,[†] Michael C. Edler,[§] Giuseppe La Regina,[†] Antonio Coluccia,[†] Maria Chiara Barbera,[‡] Denise Barrow,[‖] Robert I. Nicholson,[‖] Gabriela Chiosis,[⊥] Andrea Brancale,[‡] Ernest Hamel,[§] Marino Artico,[†] and Romano Silvestri*,[†]

Istituto Pasteur - Fondazione Cenci Bolognetti, Dipartimento di Studi Farmaceutici, Università di Roma "La Sapienza", Piazzale Aldo Moro 5, I-00185 Roma, Italy, Welsh School of Pharmacy, Cardiff University, King Edward VII Avenue, Cardiff, CF10 3XF, UK, Program in Cell Biology, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, New York 10021, Toxicology and Pharmacology Branch, Developmental Therapeutics Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute at Frederick, National Institutes of Health, Frederick, Maryland 21702, and Tenovus Centre for Cancer Research, Welsh School of Pharmacy, Cardiff University, King Edward VII Avenue, Cardiff, CF10 3XF, UK

Received August 15, 2005

Arylthioindoles (ATIs) that possess a 3-methoxyphenylthio or a 3,5-dimethoxyphenylthio moiety at position 2 of the indole ring were effective tubulin assembly inhibitors, but weak inhibitors of MCF-7 cell growth. ATIs bearing a 3-(3,4,5-trimethoxyphenyl)thio moiety were potent tubulin polymerization inhibitors, with IC_{50}s in the 2.0 (35) to 4.5 (37) μM range. They also inhibited MCF-7 cell growth at nanomolar concentrations. The 3,4,5-trimethoxy substituted ATIs showed potencies comparable to those of the reference compounds colchicine and combretastatin A-4 in both tubulin assembly and cell growth inhibition assays. Dynamics simulation studies correlate well with the observed experimental data. Furthermore, from careful analysis of the biological and in silico data, we can now hypothesize a basic pharmacophore for this class of compounds.

Introduction

Anticancer therapy based on microtubule-targeting agents is receiving growing attention. Tubulin heterodimers, the major component of microtubules, are the molecular target of these agents. Inhibition of tubulin polymerization or interfering with microtubule disassembly disrupts several cellular functions, including cell motility and mitosis.[1–3] The former group of agents includes such compounds as colchicine (1), combretastatin A-4 (2a; CSA4), and *Catharanthus* alkaloids, such as vinblastine, vincristine, and vinorelbine. The microtubule stabilizing drugs include the clinically important taxoids paclitaxel and docetaxel, as well as more recently discovered natural products, such as the epothilones, discodermolide, and eleutherobin. As a group, drugs that bind to either unpolymerized tubulin or tubulin polymer interfere both with cell division and interphase functions that require a normal microtubule cytoskeleton. Cells treated with these agents invariably undergo apoptosis.[4]

Since clinically useful antitubulin drugs all have problems with toxicity, drug resistance, and bioavailability, there is a continuing effort to find new compounds that might be safer or more effective.[2,5] Several compounds are currently in ongoing clinical trials, for example various epothilones and taxoids, representing the stabilizer class, and the destabilizer combretastatin A-4 phosphate (2b), ZD 6126[6] (3), the semisynthetic vinca alkaloid vinflunine,[7] the naturally occurring peptide dolastatin 10,[8] the sulfonamide E7010,[9] and T138067/T900607.[3]

Several indoles were found to be effective as inhibitors of tubulin assembly.[10–13] We recently discovered arylthioindoles (ATIs) as a new class of potent inhibitors of tubulin assembly

Chart 1

colchicine (1)

R=H, combretastatin A-4 (CSA4) (2a);
R=OP(O)(ONa)₂ or OP(O)(OH)(ONa),
combretastatin A-4 phosphate (2b)

ZD 6126 (3)

Arylthioindoles (ATIs) (4),
R₁=ester, R₂-R₆=H,Cl,OMe
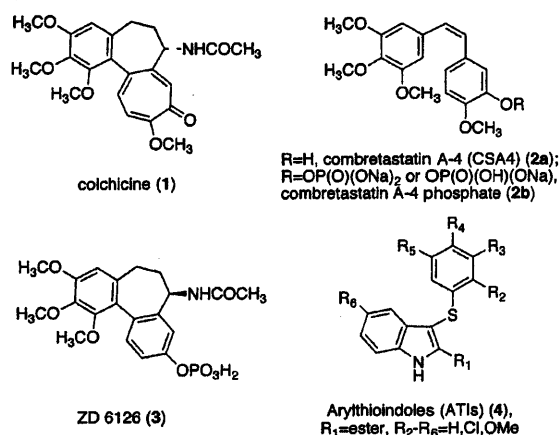
and the growth of MCF-7 human breast carcinoma cells[14] (for general formula of ATIs, see structure 4). The mechanism of action was through binding at the colchicine site on β-tubulin. In our preliminary work, structure–activity relationship (SAR) analysis highlighted four determinant structural requirements: (A) the ester function at position 2 of the indole; (B) the 3-arylthio group; (C) the sulfur atom bridge; (D) the substituent at position 5 of the indole, especially for the inhibition of MCF-7 cell growth.[14] Pursuing our research project, here we describe the synthesis, biological evaluation, and molecular modeling studies of new ATI derivatives (Chart 2).

Chemistry

Synthesis and structures of new ATI derivatives are depicted in Scheme 1 and Table 1, respectively. Compounds 12, 13, 14, 17, 19, 28, 31–34, and 38 were synthesized by treating appropriate indole-2-carboxylates with *N*-arylthiosuccinimides in the presence of boron trifluoride diethyl etherate in anhydrous

* Corresponding author. Phone +39 06 4991 3800, Fax +39 06 491 491, E-mail: romano.silvestri@uniroma1.it.
† Università di Roma.
‡ Cardiff University.
§ National Cancer Institute.
⊥ Memorial Sloan-Kettering Cancer Center.
‖ Tenovus Centre for Cancer Research.

Chart 2. SAR Study of ATIs
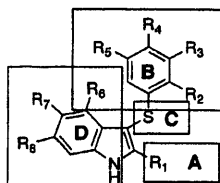
CH₂Cl₂, as we previously reported.[15] Compounds **16, 27, 36,** and **41** were obtained following a two-step procedure involving addition of 3,4,5-trimethoxylthiophenol to a solution of *N*-chlorosuccinimide (NCS) at −78 °C, and this mixture was treated with the indole-2-carboxylate while the reaction temperature was warmed to 0 °C over 1 h. Compound **25** was prepared by treatment of indole-2-carboxylic acid with 1,1′-(3,5-dimethoxyphenyl)disulfide in the presence of NaH. The crude acid was transformed into the corresponding methyl ester by reaction with trimethylsilyldiazomethane (TMSDM) at room temperature for 30 min. In the case of compounds **37** and **40**, the crude acids were transformed into the corresponding methyl esters by refluxing in methanol in the presence of thionyl chloride for 48 h under a stream of anhydrous argon. This procedure was also used for the preparation of esters **5**−**11** starting from 3-phenylthio-1*H*-indole-2-carboxylic acid[15] and appropriate alcohols. Oxidation of sulfur derivatives **17, 19, 21, 29,** and **38** with 3-chloroperoxybenzoic acid (MCPBA) furnished the corresponding sulfones **18, 20, 22, 30,** and **39**, respectively.

Results and Discussion

Table 2 summarizes the biological data for inhibition of tubulin polymerization, colchicine binding to tubulin (more active compounds only), and growth of MCF-7 human breast carcinoma cells by the indoles **5**−**41** in comparison with data obtained with the reference compounds colchicine (**1**) and CSA4 (**2a**). Biological evaluation of selected ATIs against small cell lung cancer (SCLC) carcinoma cells is also reported.

Previously, we found a 2- to 4-fold increase in the inhibitory effects of ATI derivatives on tubulin polymerization by introducing either a methoxycarbonyl (IC₅₀ = 8.2 μM) or an ethoxycarbonyl (IC₅₀ = 4.4 μM) function at position 2 of the indole ring of 3-phenylthio-1*H*-indole.[14] We therefore evaluated the effect of additional esters with either a linear or a branched alkoxy chain (compounds **5**−**11**) at this position. Only the isopropyl ester of 3-(phenylthio)-1*H*-indole-2-carboxylic acid (**6**) had any activity (IC₅₀ = 18 μM). All other tested compounds were inactive in both the tubulin polymerization and MCF-7 cell assays.

The tubulin assembly inhibitory activity of ATIs bearing a single methoxy group on the 3-arylthio moiety was strongly dependent on the position of the methoxy on the phenyl ring. Invariably, within the cohort of compounds synthesized, the greatest activity occurred when the methoxy group was located meta, the least when located para, to the bridging atom (compare **13** with **12** and **14, 19** with **17** and **21, 33** with **32** and **34**).

This inhibitory effect was unpredictably affected by substituents at position 5 of the indole ring (compare the ethyl esters **12** and **32** and also the methyl ester **17**, the ethyl esters **13** and **33** and also the methyl ester **19**).

Notably, of the compounds with a single methoxy group on the phenyl ring, potent inhibitory activity against tubulin assembly was observed with compounds **13, 17, 19,** and **33**, and all except **17** were reasonably effective inhibitors of MCF-7

cell growth (IC₅₀'s in 0.15 to 0.33 μM range). Thus, for example, methyl 3-[(3-methoxyphenyl)thio]-5-methoxy-1*H*-indole-2-carboxylate (**33**, IC₅₀ = 3.1 μM) was as active as colchicine as an inhibitor of tubulin assembly, and methyl 3-[(3-methoxyphenyl)-thio]-5-chloro-1*H*-indole-2-carboxylate (**19**, IC₅₀ = 1.8 μM) was 1.8 times more potent than colchicine and slightly superior to CSA4. These potent tubulin assembly inhibitors, however, were more than 10-fold less inhibitory than the reference compounds **1** (colchicine) and **2a** (CSA4) as inhibitors of MCF-7 cell growth.

Introduction of a second methoxy group at position 5 of the arylthio moiety of **19** produced methyl 3-[(3,5-dimethoxyphenyl)thio]-5-chloro-1*H*-indole-2-carboxylate (**25**). Compound **25** differed little from **19** in its inhibitory effect on tubulin assembly, but **25** was almost 10-fold more cytotoxic. Worthy of note, replacement of the 3-(3,5-dimethoxyphenyl)thio group of **25** with a 3-(3,5-dimethylphenyl)thio moiety (**23**) produced an inactive compound. The ethyl ester analogue of **23**, compound **24**, was also essentially inactive.

ATI derivatives **15, 16, 26, 27, 35, 36, 37, 40,** and **41** bear a 3-(3,4,5-trimethoxyphenyl)thio moiety. The first seven were all highly active as inhibitors of tubulin polymerization, with IC₅₀ values in the 2.0 (**35**) to 4.5 (**37**) μM range. These values were comparable to those of the reference compounds colchicine (IC₅₀ = 3.2 μM) and CSA4 (IC₅₀ = 2.2 μM). The 3,4,5-trimethoxy substitution of the phenylthio group was found to be of crucial importance for potent MCF-7 cell growth inhibition. In fact, with the sole exception of the 3,5-dimethoxy derivative **25**, only compounds bearing the 3-(3,4,5-trimethoxyphenyl)thio moiety had IC₅₀ values of less than 50 nM against the MCF-7 cells, with the most potent ATI being compound **35** (compare **15** (IC₅₀ = 25 nM), **16** (IC₅₀ = 40 nM), **26** (IC₅₀ = 42 nM), and **35** (IC₅₀ = 13 nM)). With the MCF-7 cells, methyl esters were about 2 to 3 times more active than the corresponding ethyl esters (compare **15** with **16, 26** with **27,** and **35** with **36**). A similar difference was not observed in effects on tubulin assembly. We also obtained some data regarding effects of susbtituents at position 5 of the indole ring. Order of activity, at least for inhibition of cell growth, was H₃CO (**35**) > H (**15**) > Cl (**26**) > NO₂ (**37**), thus suggesting a correlation between electron-donating effect and cytotoxicity.

Changing the position of the methoxy group on the indole from position 5 to 4 or 6 led to loss of activity in inhibiting tubulin assembly (compare **29** with **31** and **28**; all three compounds were inactive against the MCF-7 cells). Similarly, introduction of an additional methoxy group at position 6 of **35** or **36** to give compounds **40** and **41** resulted in a dramatic loss of activity. Finally, as observed previously,[14] oxidation of the sulfur atom to the sulfone always resulted in a large loss of activity (compare **17** with **18, 19** with **20, 29** with **30,** and **38** with **39**).

Biological data against the SCLC cells showed good agreement with the data for the MCF-7 cells, although, in this case, compound **25** appeared to be the most active agent in the series.

Due to structural similarly of the ATIs with the Hsp90 inhibitor PU3,[16] we evaluated compound **26** for binding to Hsp90. We found that **26** did not compete with geldanamycin for binding to Hsp90α, and consequently it did not induce the degradation of Her2. As would be expected for an agent with antitubulin activity, however, compound **26** caused MDA-MB-468 breast cancer cells to accumulate in mitotic arrest, as measured with the Tg3 antibody (Tg3 recognizes p-nucleolin, which is highly expressed during mitosis).

Scheme 1[a]
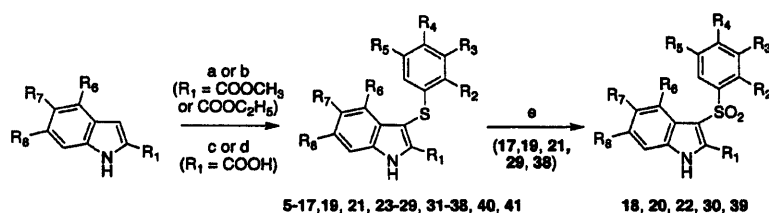
R_1 = COOMe, COOEt, COO-*n*-Pr, COO-*i*-Pr, COO-*n*-Bu, COO-*i*-Bu, COO-*s*-Bu, COO-*t*-Bu, COOCH$_2$Ph;
R_2-R_8 and R_8 = H, OMe; R_7 = H, Cl, OMe.

[a] Reagents and reaction conditions: (a) (**12, 13, 14, 17, 19, 28, 31**−**34** and **38**) *N*-[(R_2−R_5)PhS]succinimide, BF$_3$.Et$_2$O, anhydrous CH$_2$Cl$_2$, r.t., 1.5 h, then 45 °C, 2 h; (b) (**16, 27, 36,** and **41**) (i) 3,4,5-(MeO)$_3$PhSH, NCS, CH$_2$Cl$_2$, −78 °C to 0 °C, 1 h; (c) (**25**) (i) [3,5-(MeO)$_2$PhS]$_2$, NaH, anhydrous DMF, 50 °C, overnight, anhydrous argon stream; (ii) TMSDM, CH$_3$OH/CH$_2$Cl$_2$, r.t., 30 min; (d) (**5**−**11, 37,** and **40**) (i) [(R_2−R_5)PhS]$_2$, NaH, anhydrous DMF, 50 °C, overnight, anhydrous argon stream; (ii), SOCl$_2$, ROH, reflux, 48 h, anhydrous argon stream; (e) (**18, 20, 22, 30,** and **39**) MCPBA (2.5 eq), CHCl$_3$, r.t., 1 h.

Table 1. Structure of Arythioindoles **5**−**41**

compd	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	X
5	COO-*n*-Pr	H	H	H	H	H	H	H	S
6	COO-*i*-Pr	H	H	H	H	H	H	H	S
7	COO-*n*-Bu	H	H	H	H	H	H	H	S
8	COO-*s*-Bu	H	H	H	H	H	H	H	S
9	COO-*t*-Bu	H	H	H	H	H	H	H	S
10	COO-*i*-Pe	H	H	H	H	H	H	H	S
11	COOCH$_2$Ph	H	H	H	H	H	H	H	S
12	COOEt	OMe	H	H	H	H	H	H	S
13	COOEt	H	OMe	H	H	H	H	H	S
14	COOEt	H	H	OMe	H	H	H	H	S
15[a]	COOMe	H	OMe	OMe	OMe	H	H	H	S
16	COOEt	H	OMe	OMe	OMe	H	H	H	S
17	COOMe	OMe	H	H	H	H	Cl	H	S
18	COOMe	OMe	H	H	H	H	Cl	H	SO$_2$
19	COOMe	H	OMe	H	H	H	Cl	H	S
20	COOMe	H	OMe	H	H	H	Cl	H	SO$_2$
21[a]	COOMe	H	H	OMe	H	H	Cl	H	S
22	COOMe	H	H	OMe	H	H	Cl	H	SO$_2$
23[b]	COOMe	H	Me	H	Me	H	Cl	H	S
24[b]	COOEt	H	Me	H	Me	H	Cl	H	S
25	COOMe	H	OMe	H	OMe	H	Cl	H	S
26[a]	COOMe	H	OMe	OMe	OMe	H	Cl	H	S
27	COOEt	H	OMe	OMe	OMe	H	Cl	H	S
28	COOMe	H	H	H	H	H	H	OMe	S
29[a]	COOMe	H	H	H	H	H	OMe	H	S
30	COOMe	H	H	H	H	H	OMe	H	SO$_2$
31	COOMe	H	H	H	H	OMe	H	H	S
32	COOEt	OMe	H	H	H	H	OMe	H	S
33	COOEt	H	OMe	H	H	H	OMe	H	S
34	COOEt	H	H	OMe	H	H	OMe	H	S
35[a]	COOMe	H	OMe	OMe	OMe	H	OMe	H	S
36	COOEt	H	OMe	OMe	OMe	H	OMe	H	S
37	COOMe	H	OMe	OMe	OMe	H	NO$_2$	H	S
38	COOEt	H	H	H	H	H	OMe	OMe	S
39	COOEt	H	H	H	H	H	OMe	OMe	SO$_2$
40	COOMe	H	OMe	OMe	OMe	H	OMe	OMe	S
41	COOEt	H	OMe	OMe	OMe	H	OMe	OMe	S

[a] Ref 14. [b] Ref 15.

Table 2. Inhibition of Tubulin Polymerization, Growth of MCF-7 Human Breast Carcinoma Cells, Colchicine Binding, and Growth of SCLC Cells by Compounds **5**−**41**

compd	tubulin[a] IC$_{50}$ ± SD (μM)	MCF-7[b] IC$_{50}$ ± SD (nM)	inhibition colchicine binding[c] (% ± SD)	SCLC[d] IC$_{50}$ ± SD (nM)
5	>40	>2.5	-	-
6	18 ± 2	>2.5	-	-
7	>40	>2.5	-	-
8	>40	>2.5	-	-
9	>40	>2.5	-	-
10	>40	>2.5	-	-
11	>40	>2.5	-	-
12	12 ± 2	>2.5	-	-
13	2.9 ± 0.3	150 ± 90	48 ± 7	585 ± 50
14	>40	>2.5	-	>10000
15[e]	2.9 ± 0.1	25 ± 1	74 ± 2	52 ± 3
16	2.9 ± 0.2	40 ± 2	51 ± 3	84 ± 5
17	4.2 ± 0.6	1300 ± 400	34 ± 6	7000 ± 300
18	>40	>2.5	-	-
19	1.8 ± 0.2	330 ± 40	53 ± 4	3200 ± 800
20	31 ± 4	>2.5	-	>10000
21[e]	>40	>2.5	-	>10000
22	>40	>2.5	-	-
23[e]	>40	>2.5	-	-
24[e]	>40	1200 ± 200	-	-
25	2.2 ± 0.2	34 ± 10	82 ± 2	18 ± 1
26[e]	2.5 ± 0.3	42 ± 10	57 ± 2	216 ± 17
27	2.2 ± 0.2	110 ± 20	53 ± 6	93 ± 10
28	>40	>2.5	-	-
29[e]	6.2 ± 0.3	>2.5	30 ± 6	-
30	>40	>2.5	-	-
31	>40	>2.5	-	-
32	16 ± 0.5	350 ± 60	-	2200 ± 200
33	3.1 ± 0.2	280 ± 100	39 ± 3	584 ± 40
34	>40	>2.5	-	>10000
35[e]	2.0 ± 0.2	13 ± 3	90 ± 1	47 ± 2
36	2.4 ± 0.2	46 ± 3	71 ± 2	-
37	4.5 ± 0.1	120 ± 40	32 ± 2	-
38	14 ± 1	>2.5	-	-
39	>40	>2.5	-	-
40	>40	1600 ± 400	-	-
41	22 ± 0.7	1000 ± 200	-	-
Col.[e]	3.2 ± 0.4	13 ± 3	-	-
CSA4[e]	2.2 ± 0.2	17 ± 10	97 ± 0.5	-

[a] Inhibition of tubulin polymerization. [b] Inhibition of growth of MCF-7 human breast carcinoma cells. [c] Inhibition of [³H]colchicine binding. [d] Inhibition of growth of SCLC cells. [e] Data from ref 14.

Molecular Modeling

To investigate the structural basis of the SAR that emerged from the biological results, we carried out docking studies on the entire series of compounds reported in this paper, using the FlexX module included in SYBYL.[17] The results were evaluated using the built in scoring function as well as the functions included in the CScore module, with the aim of finding a good correlation between the experimental results and the in silico predictions. Some docking results were very similar to the proposed binding mode reported recently,[14] but, unfortunately, none of the scoring results correlated with the biological data (data not shown). In particular, the scoring results for the *p*-methoxyphenyl analogues were often similar to or better than the corresponding results for the trimethoxyphenyl compounds.

However, FlexX was able to identify an alternate possible binding conformation for this class of compounds compared
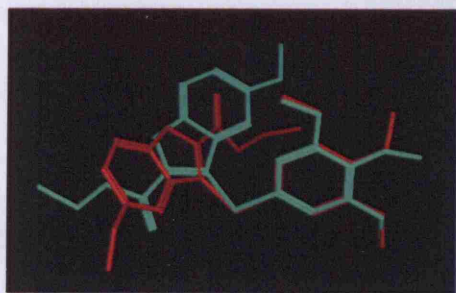
Figure 1. Two possible binding modes of **36**. In green the previously reported conformation (old pose); in red the current proposed conformation (new pose).
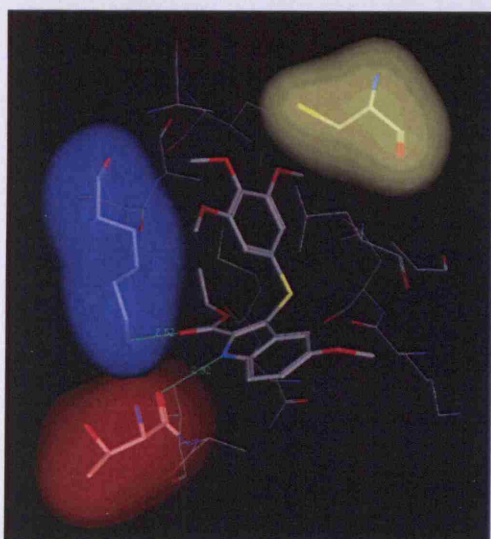
Figure 2. Proposed binding mode for **36**. Lys352 in blue, Thr179 in red, Cys241 in yellow.

Figure 3. Proposed binding mode for colchicine. Leu248 and Leu255 in gray, Cys241 in yellow, Thr179 and Lys352 in stick representation.

with the one reported previously (Figure 1). In this new pose, the indole moiety still forms a hydrogen bond with Thr179 (residue number based on the crystal structure used), but now the ester group is positioned deep in the binding pocket. In addition, the carbonyl of the ATI ester group forms another hydrogen bond with Lys352 (Figure 2).

In comparing these two possible binding modes, the scoring functions employed were not able to clearly indicate which would be the preferred conformation. To solve this problem, we carried out molecular dynamics simulations (MD) on the tubulin–ligand complexes to evaluate the potential energy of the binding of the two different poses of **36** in the colchicine binding site. The calculations were performed using the MOE (Molecular Operating Environment)[18] software package, allowing free movement of the residues within a 15 Å distance of the ligand, keeping the rest of the protein fixed. This site was soaked in water, and the simulation was run in the NVT (number of particles, volume, and temperature of the system are kept constant during the simulation) environment, for a total of 600 ps at 300 K. The states generated during the first 100 ps were not considered in evaluating the results.

The interaction potential energy values calculated by the dynamics simulations revealed a difference of almost 20 kcal/mol between the two conformations (average values: −64.7 kcal/mol for the new pose, −44.9 kcal/mol for the old pose), suggesting that **36** would bind with the indole moiety oriented as in Figure 2.

These results encouraged us to extend our molecular dynamics studies to other compounds and, in particular, to the *p*-, *m*-, and *o*-methoxyphenyl analogues (**32**, **33**, **34**). We also performed the simulation under the same conditions on colchicine to have a better understanding of its binding mode with tubulin. For the latter study, we modified the structure of the *N*-deacetyl-*N*-(2-mercaptoacetyl)colchicine cocrystallized with tubulin, and the resulting complex was energy minimized before the molecular dynamics experiment was performed.

Colchicine established one hydrogen bond between the carbonyl group of ring C and Lys352 and another between the NH group and Thr179 (Figure 3), the same residues involved in the binding of **36** (Figure 2). During the simulation, the amide moiety changed its orientation, losing the hydrogen bond with Thr179 and forming another one between the amide carbonyl group and the hydroxyl group of Ser178. The trimethoxy ring A remains placed in a hydrophobic pocket where, in particular, two leucine residues (Leu248 and Leu255) interact strongly with the aromatic ring. Cys241, a key residue for the binding and biological activity of many colchicine analogues,[19] is in close proximity to the C-3 methoxy group of ring A. The average value of the interaction potential energy for colchicine was −71.6 kcal/mol.

The para-substituted arylthioindole **34**, on the other hand, does not establish a hydrogen bond with Thr179 (Figure 4). The interaction with Lys352 remains stable during the simulation time, while the aromatic ring does not establish any contact with Cys241. The energy value obtained for binding, −49.8 kcal/mol, is in agreement with the poor activity observed in the biological experiments. This is ∼15 kcal/mol higher than the value obtained for **36**, indicating a reduced binding affinity for **34** relative to **36**.

The results obtained for the ortho- and meta-substituted arylthioindoles (**32** and **33**) also correlated well with the experimental data. Both compounds established hydrogen bonds with Thr179 and Lys352 through the indole ring. The phenyl ring of compound **33** assumes a similar orientation as the trimethoxy ring of **36**, but, in the case of compound **32**, the *o*-methoxy aryl moiety is almost orthogonal to the corresponding aromatic ring of **36** (Figure not shown). The energy values obtained for **32** and **33** are in agreement with these observations and with the experimental results (Table 3).

Figure 4. Proposed binding mode for **34**. Lys352 in blue, Thr179 in red, Cys241 in yellow.

Table 3. Comparison between Average Potential Energy of Interaction (U_{ab}) and Inhibition of Tubulin Polymerization for Compounds **32**, **33**, **34**, and **36**

compd	U_{ab} (kcal/mol)	tubulin IC$_{50}$ ± SD (μM)
32	−57.6	16 ± 0.5
33	−60.1	3.1 ± 0.2
34	−49.8	>40
36	−64.7	2.4 ± 0.2

Conclusions

ATI derivatives that possess a 3-methoxyphenylthio or a 3,5-dimethoxyphenylthio moiety at position 2 of the indole ring were potent tubulin inhibitors but weak inhibitors of MCF-7 cell growth. On the other hand, ATIs bearing the 3-(3,4,5-trimethoxyphenyl)thio moiety were not only potent tubulin inhibitors, with potencies comparable to those of the reference compounds colchicine (**1**) and CSA4 (**2a**), but also potent inhibitors of cell growth. Conversely, the 4-methoxyarylthio substitution was detrimental for inhibition of both tubulin polymerization and MCF-7 cell growth.

The results obtained from the dynamics simulation correlate well with the observed experimental data. The agreement between the SAR findings and the modeling studies increase our confidence in the binding model we have obtained, and these results may enable us to design more selective and potent analogues.

Experimental Section

Chemistry. Melting points (mp) were determined on a Büchi 510 apparatus and are uncorrected. Infrared spectra (IR) were run on a SpectrumOne FT spectrophotometer. Band position and absorption ranges are given in cm^{-1}. Proton nuclear magnetic resonance (^{1}H NMR) spectra were recorded on Bruker 200 and 400 MHz FT spectrometers in the indicated solvent. Chemical shifts are expressed in δ units (ppm) from tetramethylsilane. Column chromatography was performed on columns packed with alumina from Merck (70−230 mesh) or silica gel from Merck (70−230 mesh). Aluminum oxide TLC cards from Fluka (aluminum oxide precoated aluminum cards with fluorescent indicator at 254 nm) and silica gel TLC cards from Fluka (silica gel precoated aluminum cards with fluorescent indicator at 254 nm) were used for thin-layer chromatography (TLC). Developed plates were visualized by a Spectroline ENF 260C/F UV apparatus. Organic solutions were dried over anhydrous sodium sulfate. Concentration and evaporation of the solvent after reaction or extraction was carried out on a Büchi Rotavapor rotary evaporator operating at reduced pressure. Elemental analyses were found within ±0.4% of the theoretical values.

Method A. General Procedure for the Synthesis of Compounds 12, 13, 14, 17, 19, 28, 31−34, and 38. Example. Ethyl 3-[(2-Methoxyphenyl)thio]-1*H*-indole-2-carboxylate (12). Boron trifluoride diethyl etherate (0.067 g, 0.06 mL, 0.00047 mol) was added to a mixture of ethyl 1*H*-indole-2-carboxylate (0.53 g, 0.0028 mol), *N*-(2-methoxyphenylthio)succinimide (0.57 g, 0.0024 mol), and anhydrous dichloromethane (20 mL) under a dry argon atmosphere. After stirring at room temperature for 2 h, boron trifluoride diethyl etherate (0.27 g, 0.24 mL, 0.002 mol) was added, and the reaction was heated at 45 °C for 2 h. After cooling, the reaction mixture was diluted with chloroform and brine while being shaken. The organic layer was separated, washed with a saturated solution of sodium hydrogen carbonate and with brine, and dried. The solvent was evaporated to give a residue that was purified by silica gel column chromatography (ethyl acetate−*n*-hexane 1:3 as eluent) to afford **12**, yield 26%, mp 164−166 °C (from ethanol). Anal. Calcd. (C$_{18}$H$_{17}$NO$_3$S (327.40)) C, H, N, S.

Ethyl 3-[(3-methoxyphenyl)thio]-1*H*-indole-2-carboxylate (13) was prepared as **12** using *N*-(3-methoxyphenylthio)succinimide. Yield 23%, mp 101−104 °C (from ethanol).[20]

Ethyl 3-[(4-methoxyphenyl)thio]-1*H*-indole-2-carboxylate (14) was synthesized as **12** using *N*-(4-methoxyphenylthio)succinimide. Yield 15%, mp 118−120 °C (from ethanol). Anal. Calcd. (C$_{18}$H$_{17}$NO$_3$S (327.40)) C, H, N, S.

Methyl 5-chloro-3-[(2-methoxyphenyl)thio]-1*H*-indole-2-carboxylate (17) was synthesized as **12** using methyl 5-chloro-1*H*-indole-2-carboxylate and *N*-(2-methoxyphenylthio)succinimide. Yield 17%, mp 222−226 °C (from ethanol). Anal. Calcd. (C$_{17}$H$_{14}$ClNO$_3$S (347.82)) C, H, Cl, N, S.

Methyl 5-chloro-3-[(3-methoxyphenyl)thio]-1*H*-indole-2-carboxylate (19) was synthesized as **12** using methyl 5-chloro-1*H*-indole-2-carboxylate and *N*-(3-methoxyphenylthio)succinimide. Yield 17%, mp 160−163 °C (from ethanol). Anal. Calcd. (C$_{17}$H$_{14}$ClNO$_3$S (347.82)) C, H, Cl, N, S.

Methyl 6-methoxy-3-(phenylthio)-1*H*-indole-2-carboxylate (28) was synthesized as **12** using methyl 6-methoxy-1*H*-indole-2-carboxylate and *N*-(phenylthio)succinimide. Yield 5%, mp 120−124 °C (from ethanol). Anal. Calcd. (C$_{17}$H$_{15}$NO$_3$S (313.38)) C, H, N, S.

Methyl 4-methoxy-3-(phenylthio)-1*H*-indole-2-carboxylate (31) was synthesized as **12**, using methyl 4-methoxy-1*H*-indole-2-carboxylate and *N*-(phenylthio)succinimide. Yield 40%, mp 115−118 °C (from ethanol). Anal. Calcd. (C$_{17}$H$_{15}$NO$_3$S (313.38)) C, H, N, S.

Ethyl 5-methoxy-3-[(2-methoxyphenyl)thio]-1*H*-indole-2-carboxylate (32) was synthesized as **12** using ethyl 5-methoxy-1*H*-indole-2-carboxylate and *N*-(2-methoxyphenylthio)succinimide. Yield 35%, mp 158 °C (from ethanol). Anal. Calcd. (C$_{19}$H$_{19}$NO$_4$S (357.43)) C, H, N, S.

Ethyl 5-methoxy-3-[(3-methoxyphenyl)thio]-1*H*-indole-2-carboxylate (33) was synthesized as **12** using ethyl 5-methoxy-1*H*-indole-2-carboxylate and *N*-(3-methoxyphenylthio)succinimide. Yield 31%, mp 104−107 °C (from ethanol). Anal. Calcd. (C$_{19}$H$_{19}$NO$_4$S (357.43)) C, H, N, S.

Ethyl 5-methoxy-3-[(4-methoxyphenyl)thio]-1*H*-indole-2-carboxylate (34) was synthesized as **12** using ethyl 5-methoxy-1*H*-indole-2-carboxylate and *N*-(4-methoxyphenylthio)succinimide. Yield 54%, mp 140−145 °C (from ethanol). Anal. Calcd. (C$_{19}$H$_{19}$NO$_4$S (357.43)) C, H, N, S.

Ethyl 5,6-dimethoxy-3-(phenylthio)-1*H*-indole-2-carboxylate (38) was synthesized as **12** using ethyl 5,6-dimethoxy-1*H*-indole-2-carboxylate and *N*-(phenylthio)succinimide. Yield 35%, mp 137−139 °C (from ethanol). Anal. Calcd. (C$_{19}$H$_{19}$NO$_4$S (357.42)) C, H, N, S.

Method B. General Procedure for the Synthesis of Compounds 16, 27, 36, and 41. Example. Ethyl 3-[(3,4,5-Trimethoxyphenyl)thio)]-1*H*-indole-2-carboxylate (16). 3,4,5-Trimethoxybenzenethiol[21] (1.87 g, 0.0093 mol) was added to a solution of NCS (1.26 g, 0.0094 mol) in anhydrous dichloromethane (74 mL) at −78 °C. The reaction was warmed to 0 °C over 15 min, and then a solution of ethyl 1*H*-indole-2-carboxylate (1.5 g, 0.008 mol) in the

same solvent (15 mL) was added. The reaction mixture was stirred at 0 °C for 1 h and concentrated under reduced pressure. The residue was suspended in water (150 mL) and extracted with ethyl acetate, and the organic layer was washed with brine and dried. Removal of the solvent furnished a crude product that was purified by silica gel column chromatography (ethyl acetate—n-hexane 1:1 as eluent) to give **16**, yield 1%, mp 90—95 °C (from ethanol). Anal. Calcd. ($C_{20}H_{21}NO_5S$ (387.46)) C, H, N, S.

Ethyl 5-chloro-3-[(3,4,5-trimethoxyphenyl)thio)]-1H-indole-2-carboxylate (27) was synthesized as **16** using ethyl 5-chloro-1H-indole-2-carboxylate. Yield 6%, mp 110—120 °C (from ethanol). Anal. Calcd. ($C_{20}H_{20}ClNO_5S$ (421.90)) C, H, Cl, N, S.

Ethyl 5-methoxy-3-[(3,4,5-trimethoxyphenyl)thio]-1H-indole-2-carboxylate (36) was synthesized as **16** using ethyl 5-methoxy-1H-indole-2-carboxylate. Yield 4%, mp 123—128 °C (from ethanol). Anal. Calcd. ($C_{21}H_{23}NO_6S$ (417.48)) C, H, N, S.

Ethyl 5,6-dimethoxy-3-[(3,4,5-trimethoxyphenyl)thio]-1H-indole-2-carboxylate (41) was synthesized as **16** using ethyl 5,6-dimethoxy-1H-indole-2-carboxylate. Yield 10%, mp 140—146 °C (from ethanol). Anal. Calcd. ($C_{22}H_{25}NO_7S$ (447.51)) C, H, N, S.

Method C. Synthesis of Methyl 5-Chloro-3-[(3,5-dimethoxyphenyl)thio]-1H-indole-2-carboxylate (25). 5-Chloro-1H-indole-2-carboxylic acid (0.16 g, 0.0008 mol) was added by portions to a mixture of sodium hydride (60% dispersion in mineral oil, 0.094 g, 0.0023 mol) in anhydrous DMF (4 mL) under a nitrogen stream at 0 °C. After 15 min, 1,1'-(3,5-dimethoxyphenyl)disulfide (0.35 g, 0.001 mol) was added portionwise, then the reaction was heated at 50 °C overnight under a nitrogen atmosphere. After cooling, the mixture was poured onto crushed ice, made acidic with 2 N HCl and extracted with ethyl acetate. The organic layer was separated, washed with brine, and dried. Removal of the solvent gave a residue that was triturated with cyclohexane and filtered to give 5-chloro-3-[(3,5-dimethoxyphenyl)thio)]-1H-indole-2-carboxylic acid, which was used as the crude product. The acid was dissolved in dichloromethane (12 mL) and methanol (3 mL) and treated with TMSDM (1.02 mL of a 2 N solution in hexane, 0.002 mol) while stirring at room temperature for 30 min. After reduction to a small volume, the suspension was treated with ethyl acetate (10 mL × 5), and the solvent was evaporated. The crude product was purified by silica gel column chromatography (ethyl acetate—n-hexane 1:5 as eluent) to give **25**, yield 10%, mp 160—165 °C (from ethanol). Anal. Calcd. ($C_{18}H_{16}ClNO_4S$ (377.84)) C, H, Cl, N, S.

Method D. General Procedure for the Synthesis of Compounds 5—11, 37, and 40. Propyl 3-(Phenylthio)-1H-indole-2-carboxylate (5). Methyl 1H-indole-2-carboxylate was obtained by esterification of the corresponding acid using TMSDM, yield 88%, mp 145—148 °C (from ethanol). Lit.,[22] mp 149—151 °C (from ethyl acetate—n-hexane). Methyl 3-(phenylthio)-1H-indole-2-carboxylate was prepared by reaction of methyl 1H-indole-2-carboxylate with N-(phenylthio)succinimide in the presence of boron trifluoride diethyl etherate as above. The crude residue was purified by silica gel column chromatography (ethyl acetate—n-hexane 1:3 as eluent), yield 55%, mp 175—180 °C (from ethanol). Lit.,[23] mp 178—180 °C (from diethyl ether—n-hexane). Lithium hydroxide monohydrate (8.46 g, 0.20 mol) was added to a suspension of methyl 3-(phenylthio)-1H-indole-2-carboxylate (18.52 g, 0.065 mol) in THF (250 mL) and water (250 mL). The mixture was stirred at room-temperature overnight. After reduction to a small volume, it was made acidic (pH = 2) with 1 N HCl and extracted with ethyl acetate. The organic layer was washed with brine, dried, and evaporated to afford pure 3-(phenylthio)-1H-indole-2-carboxylic acid, yield 94%, mp 161—162 °C (from glacial acetic acid). Lit.,[23] 160—162 °C (from glacial acetic acid). To a cooled suspension of 3-(phenylthio)-1H-indole-2-carboxylic acid (1.00 g, 0.0037 mol) in 1-propanol (6.50 mL) was added dropwise thionyl chloride (0.33 mL) under an anhydrous argon atmosphere. The reaction mixture was heated at 70 °C for 5 h. After cooling, the suspension was filtered to afford **5**, yield 52%, mp 132—135 °C (from ethanol). Anal. Calcd. ($C_{18}H_{17}NO_2S$ (311.40)) C, H, N, S.

Isopropyl 3-(phenylthio)-1H-indole-2-carboxylate (6) was synthesized as **5** using *iso*-propyl alcohol. Yield 50%, mp 146 °C (from ethanol). Anal. Calcd. ($C_{18}H_{17}NO_2S$ (311.40)) C, H, N, S.

n-Butyl 3-(phenylthio)-1H-indole-2-carboxylate (7) was synthesized as **5** using n-butyl alcohol. Yield 25%, mp 138 °C (from ethanol). Anal. Calcd. ($C_{19}H_{19}NO_2S$ (325.43)) C, H, N, S.

sec-Butyl 3-(phenylthio)-1H-indole-2-carboxylate (8) was synthesized as **5** using sec-butyl alcohol. Yield 20%, mp 145 °C (from ethanol). Anal. Calcd. ($C_{19}H_{19}NO_2S$ (325.43)) C, H, N, S.

$tert$-Butyl 3-(phenylthio)-1H-indole-2-carboxylate (9) was synthesized as **5** using $tert$-butyl alcohol. Yield 12%, mp 143—147 °C (from ethanol). Anal. Calcd. ($C_{19}H_{19}NO_2S$ (325.43)) C, H, N, S.

iso-Pentyl 3-(phenylthio)-1H-indole-2-carboxylate (10) was synthesized as **5** using iso-pentyl alcohol. Yield 21%, mp 124—128 °C (from ethanol). Anal. Calcd. ($C_{20}H_{21}NO_2S$ (339.46)) C, H, N, S.

Benzyl 3-(phenylthio)-1H-indole-2-carboxylate (11) was synthesized as **5** using benzyl alcohol. Yield 20%, mp 135—138 °C (from ethanol). Anal. Calcd. ($C_{22}H_{17}NO_2S$ (359.45)) C, H, N, S.

Methyl 5-nitro-3-[(3,4,5-trimethoxyphenyl)thio]-1H-indole-2-carboxylate (37) was synthesized as **5** using 5-nitro-1H-indole-2-carboxylic acid and 1,1'-(3,4,5-trimethoxyphenyl)disulfide. Transformation of 5-nitro-3-[(3,4,5-trimethoxyphenyl)thio]-1H-indole-2-carboxylic acid into **37** was achieved by heating at reflux in methanol in the presence of thionyl chloride. The crude product was purified by silica gel column chromatography (ethyl acetate—n-hexane 1:2 as eluent), yield 3%, mp 194—198 °C (from ethanol). Anal. Calcd. ($C_{19}H_{18}N_2O_7S$ (418.43)) C, H, N, S.

Methyl 5,6-dimethoxy-3-[(3,4,5-trimethoxyphenyl)thio]-1H-indole-2-carboxylate (40) was synthesized as **5** using 5,6-dimethoxy-1H-indole-2-carboxylic acid. Yield 8%, mp 164—168 °C (from ethanol). Anal. Calcd. ($C_{21}H_{23}NO_7S$ (433.48)) C, H, N, S.

Methyl 5-chloro-3-[(2-methoxyphenyl)sulfonyl]-1H-indole-2-carboxylate (18). MCPMA (0.16 g, 0.0009 mol) was added to an ice-cold solution of **17** (0.13 g, 0.0004 mol) in chloroform (6 mL). The reaction mixture was stirred at room temperature for 2 h, poured on crushed ice, and extracted with chloroform. The organic solution was shaken with a saturated solution of sodium hydrogen carbonate and with brine and dried. After concentration to a small volume, the solution was purified by silica gel column chromatography (ethyl acetate—n-hexane 1:3 as eluent) to afford **18**, yield 50%, mp 250 °C (from ethanol). Anal. Calcd. ($C_{17}H_{14}ClNO_5S$ (379.82)) C, H, Cl, N, S.

Methyl 5-chloro-3-[(3-methoxyphenyl)sulfonyl]-1H-indole-2-carboxylate (20) was synthesized as **18**, starting from **19**. Yield 92%, mp 170—173 °C (from ethanol). Anal. Calcd. ($C_{17}H_{14}ClNO_5S$ (379.82)) C, H, Cl, N, S.

Methyl 5-chloro-3-[(4-methoxyphenyl)sulfonyl]-1H-indole-2-carboxylate (22) was synthesized as **18**, starting from **21**. Yield 29%, mp 230—232 °C (from ethanol). Anal. Calcd. ($C_{17}H_{14}ClNO_5S$ (379.82)) C, H, Cl, N, S.

Methyl 5-methoxy-3-(phenylsulfonyl)-1H-indole-2-carboxylate (30) was synthesized as **18**, starting from **29**. Yield 28%, mp 166—179 °C (from ethanol). Anal. Calcd. ($C_{17}H_{15}NO_5S$ (345.37)) C, H, N, S.

Ethyl 5,6-dimethoxy-3-(phenylsulfonyl)-1H-indole-2-carboxylate (39). Was synthesized as **18**, starting from **38**. Yield 47%, mp 264—273 °C (from ethanol). Anal. Calcd. ($C_{19}H_{19}NO_6S$ (389.42)) C, H, N, S.

N-(2-Methoxyphenylthio)succinimide. 2-Methoxythiophenol (2.80 g, 0.02 mol) was added dropwise to an ice-cold mixture of NCS (3.34 g, 0.025 mol) and anhydrous dichloromethane (30 mL) under an argon atmosphere. After 1 h, additional NCS (0.4 g, 0.003 mol) was added, and the reaction mixture was stirred for 2.5 h. Triethylamine (2.83 g, 3.9 mL, 0.028 mol) was added while stirring for 15 min, and then dichloromethane (20 mL) and 1 N HCl (10 mL) were added. After shaking, the organic layer was dried, concentrated to a small volume, and passed through a Celite column. After evaporation of the solvent, the residue was triturated with diethyl ether to give N-(2-methoxyphenylthio)succinimide, yield

60%, mp 165—168 °C, (from ethanol), Lit.,[24] mp 173 °C. Anal. Calcd. ($C_{11}H_{11}NO_3S$ (237.27)).

N-(3-Methoxyphenylthio)succinimide was synthesized as N-(2-methoxyphenylthio)succinimide using 3-methoxythiophenol. Yield 60%, mp 147—149 °C, Lit.,[25] mp 149—150 °C.

N-(4-Methoxyphenylthio)succinimide was synthesized as N-(2-methoxyphenylthio)succinimide using 4-methoxythiophenol. Yield 60%, mp 135—140 °C, Lit.[26] mp 100—106 °C. Anal. Calcd. ($C_{11}H_{11}$-NO_3S (237.27)).

1,1′-(3,5-Dimethoxyphenyl)disulfide. A mixture of 3,5-dimethoxyaniline (5.00 g, 0.033 mol), concentrated HCl (5.9 mL), and crushed ice (8 g) was prepared and cooled to −5 °C. To this mixture was added dropwise a solution of sodium nitrite (2.25 g, 0.033 mol) in water (12 mL) over 45 min. The reaction mixture was stirred at 0 °C for an additional 15 min, at which time it was added to a solution of potassium ethyl xantate (10.46 g, 0.065 mol) in water (31 mL) at 65 °C. After the mixture was stirred for 30 min at 65 °C, it was extracted with ethyl acetate. The organic layer was washed with brine, dried, and evaporated to give 3,5-dimethoxybenzenethiol (2.86 g, yield 51%), which was used without further purification. A solution of iodine (6.10 g, 0.024 mol) in ethanol 96% (15 mL) was added dropwise to a solution of 3,5-dimethoxybenzenethiol (2.86 g, 0.017 mol) in the same solvent (5 mL) until a persistent color was observed. After stirring for 10 min, the reaction mixture was extracted with ethyl acetate. The organic layer was washed with a saturated solution of sodium thiosulfate and with brine and dried. Evaporation of the solvent gave 1,1′-(3,5-dimethoxyphenyl)-disulfide, yield 52%, which was used without purification.

Biology. Tubulin Assembly. The reaction mixtures contained 0.8 M monosodium glutamate (pH 6.6 with HCl in 2 M stock solution), 10 μM tubulin, and varying concentrations of drug. Following a 15 min preincubation at 30 °C, samples were chilled on ice, GTP to 0.4 mM was added, and turbidity development was followed at 350 nm in a temperature controlled recording spectrophotometer for 20 min at 30 °C. Extent of reaction was measured. Full experimental details were previously reported.[27]

Colchicine Binding Assay. The reaction mixtures contained 1.0 μM tubulin, 5.0 μM [³H]colchicine, and 5.0 μM inhibitor and were incubated 10 min at 37 °C. Complete details were described previously.[28]

MCF-7 Cell Growth. The above paper[28] can also be referenced for methodology of MCF-7 cell growth.

SCLC Cell Growth. Experiments were carried out following a previously reported methodology.[29]

Hsp90 Competition Assay. Experiments were carried out as previously reported.[30] Briefly: Fluorescence polarization measurements were performed on an Analyst AD instrument (Molecular Devices, Sunnyvale, CA). Measurements were taken in black 96-well microtiter plates (Corning #3650). The assay buffer contained 20 mM potassium HEPES pH 7.3, 50 mM KCl, 5 mM $MgCl_2$, 20 mM Na_2MoO_4, and 0.01% NP40. Before each use, 0.1 mg/mL bovine gamma globulin (Panvera Corporation, Madison, WI) and 2 mM dithiothreitol (Fisher Biotech, Fair Lawn, NJ) were added to the buffer. GM-BODIPY (fluorescently labeled geldanamycin) was synthesized as previously reported[31] and was dissolved in dimethyl sulfoxide to form a 10 μM solution. Recombinant Hsp90α was purchased from Stressgen Bioreagents (cat. No. SPP-776), (Victoria, BC, Canada). For the competition studies, each 96-well contained 5 nM fluorescent GM-BODIPY, 30 nM Hsp90α and potential inhibitor (initial stock in dimethyl sulfoxide) in a final volume of 100 μL. The plate was left on a shaker at 4 °C for 24 h, and the fluorescence polarization values in mP were recorded. EC_{50} values were determined as the competitor concentrations at which 50% of the fluorescent GM-BODIPY was displaced.

Her2 Assay. The SKBr3 breast cancer cells were plated in black, clear-bottom microtiter plates (Corning #3603) at 3000 cells/well in growth medium (100 μL) and allowed to attach for 24 h at 37 °C and in a 5% CO_2 atmosphere. Growth medium (100 μL) with drug or vehicle (dimethyl sulfoxide) was carefully added to the wells, and the plates were placed at 37 °C in the 5% CO_2 atmosphere. Following a 24 h incubation with drugs, wells were washed with ice-cold Tris buffered saline containing 0.1% Tween 20 (TBST) (200 μL). A house vacuum source attached to an eight-channel aspirator was used to remove the liquid from the plates. Methanol (100 μL at −20 °C) was added to each well, and the plate was placed at 4 °C for 10 min. The methanol was removed by washing with TBST (2 × 200 μL). After further incubation at r.t. for 2 h with SuperBlock (Pierce 37535) (200 μL), anti-Her-2 (c-erbB-2) antibody (Zymed Laboratories #28—004) (100 μL, 1:200 in SuperBlock) was placed in each well. The plate was incubated overnight at 4 °C. For control wells, 1:200 dilution of a normal rabbit IgG (Santa Cruz #SC-2027) in Superblock was used. Each well was washed with TBST (2 × 200 μL) and incubated at r.t. for 2 h with an anti-rabbit HRP-linked antibody (Sigma, A-0545) (100 μL, 1:2000 in SuperBlock). Unreacted antibody was removed by washing with TBST (3 × 200 μL), and the ECL Western blotting reagent (Amersham #RPN2106) (100 μL) was added. The plate was immediately read in an Analyst AD plate reader (Molecular Devices). Each well was scanned for 0.1 s. Readings from wells containing only control IgG and the corresponding HRP-linked secondary antibody were set as background and subtracted from all measured values. Luminescence readings from drug-treated cells versus untreated cells were quantified and plotted against drug concentration to give the EC_{50} values as the concentration of drug that caused 50% decrease in luminescence.

Antimitotic Assay. Black, clear-bottom microtiter 96-well plates (Corning Costar #3603) were used to accommodate experimental cultures. MDA-MB-468 cells were seeded in each well at 8000 cells per well in growth medium (100 μL) and allowed to attach overnight at 37 °C in a 5% CO_2 atmosphere. Growth medium (100 μL) with drug or vehicle (dimethyl sulfoxide) was gently added to the wells, and the plates were incubated at 37 °C in a 5% CO_2 atmosphere for 24 h. Wells were washed with ice-cold TBST (2 × 200 μL). A house vacuum source attached to an eight-channel aspirator was used to remove the liquid from the 96-well plates. Ice-cold methanol (100 μL) was added to each well, and the plate was placed at 4 °C for 5 min. Methanol was removed by suction, and plates were washed with ice-cold TBST (2 × 200 μL). Plates were further incubated with SuperBlock blocking buffer (Pierce #37535) (200 μL) for 2 h at r.t. The Tg-3 antibody (gift of Dr. Davies, Albert Einstein College of Medicine) diluted 1:200 in SuperBlock was placed in each well (100 μL), except the control column that was treated with control antibody (Mouse IgM, NeoMarkers, NC-1030-P). After 72 h, wells were washed with ice-cold TBST (2 × 200 μL). The secondary antibody (goat Anti-Mouse IgM, Southern Biotech #1020—05) was placed in each well at 1:2000 dilution in SuperBlock and incubated on a shaker at r.t. for 2 h. Unreacted antibody was removed by washing the plates with ice-cold TBST (3 × 200 μL) for 5 min on a shaker. The ECL Western Blotting Detection Reagents 1 and 2 in 1:1 mixture (100 μL) was placed in each well, and the plates were read immediately in an Analyst AD plate reader (Molecular Devices). Luminescence readings were imported into SOFTmax PRO 4.3.1. Antimitotic activity was defined as a concentration dependent increase in luminescence readings in compound-treated wells as compared to vehicle-treated wells. The antibody (Tg-3), originally described as a marker of Alzheimer's disease, is highly specific for mitotic cells, Tg-3 immunofluorescence being > 50-fold more intense in mitotic cells than in interphase cells.

Molecular Modeling. All molecular modeling studies were performed on a RM Innovator with Pentium IV 2.8 GHz processor, running Linux Fedora Core 3 using Molecular Operating Environment (MOE) 2004.03[17] and the FlexX module in SYBYL 7.0.[16] The tubulin structure was downloaded from the PDB data bank (http://www.rcsb.org/pdb/index.html PDB code: 1SA0).[32] All the minimizations were performed with MOE until RMSD gradient of 0.05 kcal mol^{-1} $Å^{-1}$ was reached with the MMFF94x force field. The partial charges were automatically calculated. Docking experiments were carried out using the FlexX docking program of SYBYL 7.0 using the default settings. The output of FlexX docking was visualized in MOE, and the scoring.svl script[33] was used to identify interaction types between ligand and protein. Molecular dynamics

was performed with MOE using the NVT environment for 600 ps and constant temperature of 300 K using the MMFF94x force field with a time step of 2 fs. Residues within 15 Å of the ligand were allowed to move freely, keeping the rest of the protein fixed. The binding site was soaked in a water sphere of 25 Å radius from the sulfur atom of the ligand, and the total charge of the system included in the water droplet did not require any adjustment. The water molecules were energy minimized keeping the coordinates of the protein—ligand complex fixed before the MD simulation. A distance restraint of 25 Å with a weight of 100 between the oxygen atoms of the water molecules and the sulfur atom of the ligand was also applied.

Supporting Information Available: ^{1}H NMR and IR spectral data, and elemental analyses of new compounds **5—12, 14, 16—20, 22, 25, 27, 28, 30—34,** and **36—41** are available free of charge via the Internet at http://pubs.acs.org.

References

(1) Lin, C. M.; Ho, H. H.; Pettit, G. R.; Hamel, E. Antimitotic natural products combretastatin A-4 and combretastatin A-2.: studies on the mechanism of their inhibition of the binding to colchicine to tubulin. *Biochemistry* **1989**, *28*, 6984—6991.

(2) Beckers, T.; Mahboobi, S. Natural, semisynthetic and synthetic microtubule inhibitors for cancer therapy. *Drugs Future* **2003**, *28*, 767—785.

(3) Medina, J. C.; Houze, J.; Clark, D. L.; Schwendner, S.; Beckmann, H.; Shan, B. Selective irreversible tubulin binders with efficacy against multi-drug resistant tumor cells. *Abstracts of Papers*, 222nd American Chemical Society National Meeting, Chicago, August 26—30, 2001; ACS: Washington, DC.

(4) (a) Pellegrini, F.; Budman D. R. Review: tubulin function, action of antitubulin drugs, and new drug development. *Cancer Invest.* **2005**, *23*, 264—273. (b) Iyer, S.; Chaplin, D. J.; Rosenthal, D. S.; Boulares, A. H.; Li, L.-Y.; Smulson, M. E. Induction of apoptosis in proliferating human endothelial cells by the tumor-specific antiangiogenesis agent combretastatin A-4. *Cancer Res.* **1998**, *58*, 4510—4514.

(5) Sridhare, M.; Macapinlac, M. J.; Goel, S.; Verdier-Pinard, D.; Fojo, T.; Rothenberg, M.; Colevas, D. The clinical development of new mitotic inhibitors that stabilize the microtubule. *Anticancer Drugs* **2004**, *15*, 553—555.

(6) Soltau, J.; Drevs, D. ZD-6126 (AstraZeneca). *IDrugs* **2004**, *7*, 380—387.

(7) Mcintyre, J. A.; Castaner, J. Vinflunine: antimitotic vinca alkaloid. *Drugs Future* **2004**, *29*, 574—580.

(8) (1) Bai, R.; Pettit, G. R.; Hamel, E. Binding of dolastatin 10 to tubulin at a distinct site for peptide antimitotic agents near the exchangeable nucleotide and vinca alkaloid sites. *J. Biol. Chem.* **1990**, *265*, 17141—17149. (2) Hamel, E.; Covell, D. G. Antimitotic peptides and depsipeptides. *Curr. Med. Chem. — Anticancer Agents* **2002**, *2*, 19—53.

(9) Owa, T.; Yokoi, A.; Yamazaki, K.; Yoshimatsu, K.; Yamori, T.; Nagasu, T. Array based structure and gene expression relationship study of antitumor sulfonamides including *N*-[2-[(4-hydroxyphenyl)-amino]-4-methoxybenzene sulfonamide and *N*-(3-chloro-7-indolyl)-1,4-benzenedisulfonamide. *J. Med. Chem.* **2002**, *45*, 4913—4922.

(10) http://www.baxter-oncology.com/english/projects/index.html.

(11) Gastpar, R.; Goldbrunner, M.; Marko, D.; von Angerer, E. Methoxy-substituted 3-formyl-2-phenylindoles inhibit tubulin polymerization. *J. Med. Chem.* **1998**, *41*, 4965—4972.

(12) Medarde, M.; Ramos, A.; Caballero, E.; Peláz-Lamamié de Clairac, R.; López, J. L.; García Grávalos, D.; San Feliciano, A. Synthesis and antineoplastic activity of combretastatin analogues: heterocombretastatins. *Eur. J. Med. Chem.* **1998**, *33*, 71—77.

(13) Flynn, B. L.; Hamel, E.; Jung, M. K. One-pot synthesis of benzo-[*b*]furan and indole inhibitors of tubulin polymerization. *J. Med. Chem.* **2002**, *45*, 2670—2673.

(14) De Martino, G.; La Regina, G.; Coluccia, A.; Edler, M. C.; Barbera, M. C.; Brancale, A.; Wilcox, E.; Hamel, E.; Artico, M.; Silvestri, R. Arylthioindoles, potent inhibitors of tubulin polymerization. *J. Med. Chem.* **2004**, *47*, 6120—6123.

(15) Silvestri, R.; De Martino, G.; Artico, M.; Massa, S.; Marceddu, T.; Loi, A. G.; Musiu, C.; La Colla, C. Indolyl Aryl Sulfones (IASs). Part 1: SAR Studies and in vitro anti-HIV-1 activity against wt RT and related mutants. *J. Med. Chem.* **2003**, *46*, 2482—2493.

(16) Chiosis, G.; Lucas, B.; Shtil, A.; Huezo, H.; Rosen, N. Development of a purine-scaffold novel class of Hsp90 binders that inhibit the proliferation of cancer cells and induce the degradation of Her2 tyrosine kinase. *Bioorg. Med. Chem.* **2002**, *10*, 3555—3564.

(17) Tripos SYBYL 7.0; Tripos Inc., 1699 South Hanley Rd, St. Louis, MO 63144, USA. http://www.tripos.com.

(18) Molecular Operating Environment (MOE 2004.03). Chemical Computing Group, Inc. Montreal, Quebec, Canada. http://www.chem-comp.com.

(19) Bai, R.; Covell, D. G.; Pei, X. F.; Ewell, J. B.; Nguyen, N. Y.; Brossi, A. Hamel, E. Mapping the binding site of colchicinoids on *β*-tubulin: 2-chloroacety-2-demethylthiocolchicine covalently reacts predominantly with cysteine 239 and secondarily with cysteine 354. *J. Biol. Chem.* **2000**, *275*, 40443—40452.

(20) For spectral data and elemental analysis of **13**, Schlosser, K. M.; Krasutsky, A. P.; Hamilton, H. W.; Reed, J. E.; Sexton, K. A highly efficient procedure for 3-sulfenylation of indole-2-carboxylates. *Org. Lett.* **2004**, *6*, 819—821.

(21) Offer, J.; Boddy, C. N. C.; Dawson, P. E. Extending Synthetic Access to Proteins with a Removable Acyl Transfer Auxiliary. *J. Am. Chem. Soc.* **2002**, *124*, 4642—4646.

(22) Boger, D. L.; Nishi, T. Diastereoselective Dieckmann condensation suitable for introduction of the duocarmycin A C6 center: development of a divergent strategy for the total synthesis of duocarmycins A and SA. *Bioorg., Med. Chem.* **1995**, *3*, 67—77.

(23) Atkinson, J. G.; Hamel, P.; Girard, Y. A new synthesis of 3-arylth-ioindoles. *Synthesis* **1988**, *6*, 480—481.

(24) Arcoria, A.; Scarlata, G. Reactivity of methoxythioanisoles with *N*-bromosuccinimide and with bromine. *Ann. Chim.* **1964**, *54*, 139—155.

(25) Savarin, C.; Srogl, J.; Liebeskind, L. S. A Mild, Nonbasic Synthesis of Thioethers. The Copper-Catalyzed Coupling of Boronic Acids with *N*-Thio(alkyl, aryl, heteroaryl)imides. *Org. Lett.* **2002**, *4*, 4309—4312.

(26) Ondetti, M. A.; Krapcho, J. Mercaptoacyl derivatives of substituted prolines. US Patent 4316906, 1982.

(27) Hamel, E. Evaluation of antimitotic agents by quantitative comparisons of their effects on the polymerization of purified tubulin. *Cell Biochem. Biophys.* **2003**, *38*, 1—21.

(28) Verdier-Pinard, P.; Lai, J.-Y.; Yoo, H.-D.; Yu, J.; Marquez, B.; Nagle, D. G.; Nambu, M.; White, J. D.; Falck, J. R.; Gerwick, W. H.; Day, B. W.; Hamel, E. Structure—activity analysis of the interaction of curacin A, the potent colchicine site antimitotic agent, with tubulin and effects of analogs on the growth of MCF-7 breast cancer cells. *Mol. Pharmacol.* **1998**, *53*, 62—76.

(29) Jones, H. E.; Goddard, L.; Gee, J. M. W.; Hiscox, S.; Rubini, M.; Barrow, D.; Knowlden, J. M.; Williams, S.; Wakeling, A. E.; Nicholson, R. I. Insulin- like growth factor-1 receptor signalling and acquired resistance to gefitnib (ZD1839; Iressa) in human breast and prostate cancer cells. *Endocr. Relat. Cancer* **2004**, *11*, 793—814.

(30) Llauger, L.; He, H.; Kim, J.; Aguirre, J.; Rosen, N.; Peters, U.; Davies, P.; Chiosis, G. 8-Arylsulfanyl and 8-arylsulfoxyl adenine derivatives as inhibitors of the heat shock protein 90. *J. Med. Chem.* **2005**, *48*, 2892—2905.

(31) Llauger, L.; Felts, S.; Huezo, H.; Rosen, N.; Chiosis, G. Synthesis of novel fluorescent probes for the molecular chaperone Hsp90. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 3975—3978.

(32) Ravelli, R. B. G.; Gigant, B.; Curmi, P. A.; Jourdain, I.; Lachkar, S.; Sobel, A.; Knossow, M. Insight into tubulin regulation from a complex with colchicine and a stathmin-like domain. *Nature* **2004**, *428*, 198—202.

(33) Code "scoring.svl" obtained from SLV Exchange website http://svl.chemcomp.com., Chemical Computing Group, Inc., Montreal, Canada.

Homology Model of Human Retinoic Acid Metabolising Enzyme Cytochrome P450 26A1 (CYP26A1): Active Site Architecture and Ligand Binding

MOHAMED SAYED GOMAA, SOOK WAH YEE, CERI ELIZABETH MILBOURNE, MARIA CHIARA BARBERA, CLAIRE SIMONS, & ANDREA BRANCALE

Medicinal Chemistry, Welsh School of Pharmacy, Cardiff University, King Edward VII Avenue, Cardiff CF10 3XF, U.K

Abstract

Homology models of cytochrome P450 RA1 (CYP26A1) were constructed using three human P450 structures, CYP2C8, CYP2C9 and CYP3A4 as templates for the model building. Using MOE software the lowest energy CYP26A1 model was then assessed for stereochemical quality and side chain environment. Further active site optimisation of the CYP26A1 model built using the CYP3A4 template was performed by molecular dynamics to generate a final CYP26A1 model. The natural substrate, all-*trans*-retinoic acid (atRA), and inhibitor R115866, were docked into the model allowing further validation of the active site architecture. Using the docking studies structurally and functionally important residues were identified with subsequent characterisation of secondary structure. Multiple hydrophobic interactions, including the side chains of TRP112, PHE299, PHE222, PHE84, PHE374 and PRO371, are important for binding of atRA and R115866. Additional hydrogen bonding interactions were noted as follows: atRA – C=O of the atRA carboxylate group and ARG86; R115866 – benzothiazole nitrogen and the backbone NH of SER115.

Keywords: *CYP26A1, homology model, MOE, all-*trans*-retinoic acid (atRA), docking studies, inhibitor interactions, inhibition, R115866*

Introduction

Retinoic acid (RA) regulates genes involved in cell proliferation, differentiation and apoptosis [1], and exerts activity by binding to transcription-regulatory factors in the cell nucleus known as RAR (retinoic acid receptor) and RXR (retinoid X receptor), each having subtypes α, β and δ [2]. RAR receptors are activated by all-*trans*- (atRA) and 9-*cis*- (9-*cis*-RA) isomers of RA, whereas RXR receptors are only activated by 9-*cis*-RA. Upon RA binding the activated receptor transcriptionally regulates its target genes by binding to specific response elements in promoter regions of RA target genes (Figure 1) [3,4].

Retinoic acid is synthesised from vitamin A (retinol), which is oxidised through retinal by dehydrogenases in the cytoplasm of target cells, in low yields, to all-*trans*-retinoic acid (Figure 1). atRA is subsequently metabolised by human liver and intestine cytochrome P450s to the inactive 4-hydroxy-RA and thence by dehydrogenases to the partially active 4-keto-RA and inactive polar metabolites (Figure 1) [5].

The P450s responsible for 4-hydroxylation of atRA in the human liver are CYP2C8 as a major contributor as well as 3A7, 3A5, 3A4, 2C9 and 1A1 [6]. However, in living tissues, atRA administration induces another RA-metabolising enzyme, CYP26 [7], which recognises only atRA as its substrate, and the expression of this isozyme can be induced by atRA both *in vitro* and *in vivo* [8]. Three members of the CYP26 family have
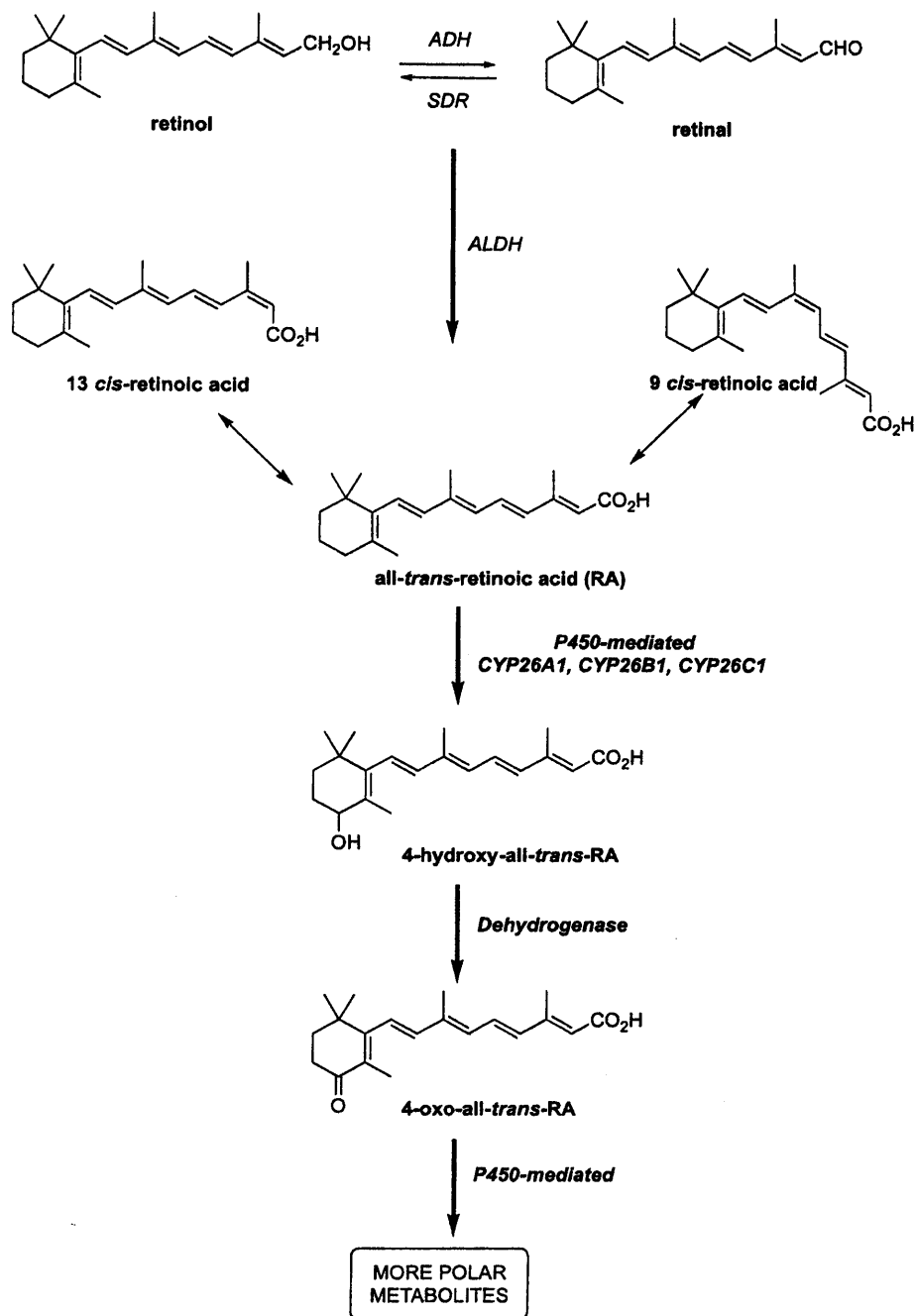
Figure 1. Alcohol dehydrogenases (ADH) and short-chain dehydrogenase/reductase catalyse the oxidation of retinol to retinaldehyde, which is subsequently oxidised by aldehyde dehydrogenases (ALDH) to retinoic acid. Retinoic acid, all-*trans* (atRA) and 9-*cis* (9cRA) isomers are further metabolised by cytochrome P450 enzymes (including CYP26A1, B1 and C1) to inactive polar metabolites resulting in retinoid excretion.

now been identified: CYP26A1 [7] and CYP26B1 [9], which metabolise atRA in the embryo and adult and, more recently [10], CYP26C1 that may have a role in the metabolism/isomerisation of both all-*trans* and 9-*cis* isomers of RA.

Retinoids in general have been used for some time in the treatment of psoriasis, cystic acne, cutaneous malignancies due to hyperkeratinisation as well as in the treatment of photo-damaged skin [11,12]. Retinoic acid has been used in a number of clinical situations, especially oncology and dermatology, atRA has also been shown to improve the efficacy of other treatments such as radiation, cisplatin and interferon therapies [13,14].

Although the amino acid sequence of CYP26A1 has been characterised no crystal structure is available. A homology model for human CYP26A1 would provide valuable information regarding active site architecture and allow docking studies to analyse ligand binding interactions, the information obtained from these studies would assist with rationale drug design of retinoid mimetics and CYP26A1 inhibitors. Importantly, crystal structures for three human cytochrome P450 enzymes, CYP2C8 [15], CYP2C9 [16] and CYP3A4 [17] have recently become available allowing the use of a human crystal template for the homology building.

Materials and methods

Computational approaches

All molecular modelling studies were performed on a RM Innovator Pentium IV 2.4 GHz running either Linux Fedora Core 3 or Windows XP using Molecular Operating Environment (MOE) 2004.03 [18] and FlexX module in SYBYL 7.0 [19] molecular modelling software.

Molecular dynamics simulations were performed with GROMACS 3.2 [20,21] and the Gromacs force field in a NVT (canonical) environment. Individual ligand/protein complexes obtained from the docking results were soaked in a triclinic water box and minimised using a steepest descent algorithm to remove unfavorable van der Waals contacts. The system was then equilibrated via a 20 ps MD simulation at 300°K with restrained ligand/protein complex atoms. Finally, a 800 ps simulation was performed at 300°K with a time step of 2 fs and hydrogen atoms constrained with a LINCS algorithm. Visualisation of the dynamics trajectories was performed with the VMD software package, version 1.8.3 [22].

All the minimisations were performed with MOE until RMSD gradient of 0.05 Kcal $mol^{-1}Å^{-1}$ with the forcefield specified and the partial charges were automatically calculated.

Homology searching

The protein sequence of human CYP26A1 was obtained from the ExPASy server (O43174) [23]. Homologous proteins with known crystal structures were found by performing a PSI-BLAST search (comparison matrix, BLOSUM62; E-threshold, 10), using the ExPASy server, aligning the query sequence (CYP26A1) against sequences in the Protein Data Bank (PDB) [24]. The alignment parameters and thresholds used for screening candidate homologues were used with their default values and BLOSUM62 comparison matrix.

Multiple sequence and structure alignment

MOE-Align, which implements a modified version of the alignment methodology originally introduced by Needleman and Wunsch [25], was employed. The sequence alignment was performed with an alignment constraint between the haem cysteine residue of the query sequence (CYP26A1) and the corresponding haem cysteine residue of the templates. All the default settings in the MOE-Align panel were used for the sequence alignment.

D Model Building

Homology models were built using MOE-Homology using a single template approach with CHARMM22 forcefield. The three human P450 crystal structures CYP2C8, CYP2C9 and CY3A4 were selected as the templates for constructing CYP26A1 models. Ten intermediate models were generated and the final model was taken as the Cartesian average of all the intermediate models. The haem was positioned using the same coordinates as in the template and the complex model was energy minimised.

Model Validation

Stereochemical quality of the polypeptide backbone and side chains was evaluated using Ramachandran

Table I. P450 sequences identified by PSI-BLAST search using CYP26A1 as the query sequence.

Protein	PDB code	PSI-BLAST score[a]	Sequence identity[b]	% Sequence identity	Chain length	E-value
CYP51	1EA1-A	275	115/447	25	455	1.3e-22
CYP102A1	1JPZ-A	253	113/432	25	473	5.7e-20
CYP3A4	1TQN-A	249	111/437	26	486	1.8e-19
CYP2C9	1R90-A	171	52/203	25	477	3.8e-13
CYP2B4	1PO5-A	164	71/285	24	476	6.9e-16
CYP2C8	1PQ2-A	156	67/277	24	476	1.1e-11
CYP2C5	1DT6-A	153	53/219	24	473	4.5e-12
CYP108	1CPT	153	42/133	31	412	4.0e-10

[a] The PSI-BLAST score for an alignment is calculated by summing the scores for each aligned position and the scores for gaps. [b] (Number of identical residues)/(length of sequence fragment identified by PSI-BLAST).

```
                   | membrane anchor region            |
         CYP2C9    -----------MDSLVVLVLCLSCLLLLSLWRQSSGRG----KLPPGPTPLPVIGNILQI 45
         CYP2C8    -----------MEPFVVLVLCLSFMLLFSLWRQSCRRR----KLPPGPTPLPIIGNMLQI 45
         CYP3A4    ----ALIPDLAMETWLLLAVSLVLLYLYGTHSHGLFKK----LGIPGPTPLPFLGNILSY 52
         CYP26A1   MGLPALLASALCTFVLPLLLFLAAIKLWDLYCVSGRDRSCALPLPPGTMGFPFFGETLQM 60
                      : * : *   : *  .     .          **.   :*.:*: *.


                                                          | SRS1
         CYP2C9    GIKDISKSLTNLSKVYGPVFTLYFGLKPIVVLHGYEAVKEALIDLG-EEFSGRGIFPLAE 104
         CYP2C8    DVKDICKSFTNFSKVYGPVFTVYFGMNPIVVFHGYEAVKEALIDNG-EEFSGRGNSPISQ 104
         CYP3A4    -HKGFCMFDMECHKKYGKVWGFYDGQQPVLAITDPDMIKTVLVKECYSVFTNRR--PFGP 109
         CYP26A1   VLQRR-KFLQMKRRKYGFIYKTHLFGRPTVRVMGADNVRRILLGDD---RLVSVHWPASV 116
                   :         : ** ::   :    .* :  . : :: *:         * .


                   SRS1             |
         CYP2C9    RANRGFGIVFSNGKKWKEIRRFSLMTLRNFGMGKRSIEDRVQEEARCLVEELRKTKAS-- 162
         CYP2C8    RITKGLGIISSNGKRWKEIRRFSLTTLRNFGMGKRSIEDRVQEEAHCLVEELRKTKAS-- 162
         CYP3A4    VGFMKSAISIAEDEEWKRLR--SLLSPTFTSGKLKEMVPIIAQYGDVLVRNLRREAETGK 167
         CYP26A1   RTILGSGCLSNLHDSSHKQRK-KVIMRAFSREALECYVPVITEEVGSSLEQWLSCGERGL 175
                        .       . :. *  .:           : :          :.:


                            |      SRS2        |
         CYP2C9    PCDPTFILGCAPCNVICSIIFHKRFDYKDQQFLNLMEKLNENIKI-LSSPWIQICNNFSP 221
         CYP2C8    PCDPTFILGCAPCNVICSVVFQKRFDYKDQNFLTLMKRFNENFRI-LNSPWIQVCNNFPL 221
         CYP3A4    PVTLKDVFGAYSMDVITSTSFGVNIDSLNNPQDPFVENTKKLLRFDFLDPFFLSITVFPF 227
         CYP26A1   LVYPEVKRLMFRIAMRILLGCEPQLAGDGDSEQQLVEAFEEMTRNLFSLPIDVPFSGLYR 235
                        :         .:    .:      :::   ::   :   :  *    . :


                        |SRS3|
         CYP2C9    IIDYFPGTHNKLL-KNVAFMKSYILEKVKEHQESMDMNNPQDFIDCFLMKMEKEKHNQPS 280
         CYP2C8    LIDCFPGTHNKVL-KNVALTRSYIREKVKEHQASLDVNNPRDFIDCFLIKMEQEKDNQKS 280
         CYP3A4    LIPILEVLNICVFPREVTNFLRKSVKRMKESRLEDTQKHRVDFLQLMIDSQNSKETESHK 287
         CYP26A1   GMKARNLIH--------ARIEQNIRAKICGLRASEAGQGCKDALQLLIEHS----WERGE 283
                        :            :      ::   : .     :    * :: ::       : .


                        |SRS4 I-helix|
         CYP2C9    EFTIESLENTAVDLFGAGTETTSTTLRYALLLLLLKHPEVTAKVQEEIER-----VIGRNR 335
         CYP2C8    EFNIENLVGTVADLFVAGTETTSTTLRYGLLLLLLKHPEVTAKVQEEIDH-----VIGRHR 335
         CYP3A4    ALSDLELVAQSIIFIFAGYETTSSVLSFIMYELATHPDVQQKLQEEIDA-----VLPNKA 342
         CYP26A1   RLDMQALKQSSTELLFGGHETTASAATSLITYLGLYPHVLQKVREELKSKGLLCKSNQDN 343
                   :    *        :: .* ***::.     : *  :*.*  *::**:.        ..


                          |ETLR |   | SRS5 |
         CYP2C9    SPCMQDRSHMPYTDAVVHEVQRYIDLLPTSLPHAVTCDIKFRNYLIPKGTTILISLTSVL 395
         CYP2C8    SPCMQDRSHMPYTDAVVHEIQRYSDLVPTGVPHAVTTDTKFRNYLIPKGTTIMALLTSVL 395
         CYP3A4    PPTYDTVLQMEYLDMVVNETLRLFPIAMR-LERVCKKDVEINGMFIPKGWVVMIPSYALH 401
         CYP26A1   KLDMEILEQLKYIGCVIKETLRLNPPVPGGFRVALKT-FELNGYQIPKGWNVIYSICDTH 402
                   :     :: *  . *::*  *          .   ..   ::..  ****   ::


                          |PERF|                 |  haem  |
         CYP2C9    HDNKEFPNPEMFDPHHFLDEGGNFKKSKYFMPFSAGKRICVGEALAGMELFLFLTSILQN 455
         CYP2C8    HDDKEFPNPNIFDPGHFLDKNGNFKKSDYFMPFSAGKRICAGEGLARMELFLFLTTILQN 455
         CYP3A4    RDPKYWTEPEKFLPERFSKKNKDNIDPYIYTPFGSGPRNCIGMRFALMNMKLALIRVLQN 461
         CYP26A1   DVAEIFTNKEEFNPDRFMLPHPEDASRFSFIPFGGGLRSCVGKEFAKILLKIFTVELARH 462
                   : :.: : * * :*      :  . **..* * * *  :*  : :     : ::


                          | SRS6   |
         CYP2C9    FNLKSLVDPKNLDTTPVVNGFASVPPFYQLCFIPV------ 490
         CYP2C8    FNLKSVDDLKNLNTTAVTKGIVSLPPSYQICFIPV------ 490
         CYP3A4    FSFKPCKETQIPLKLSLGGLLQPEKPVVLKVESRDGTVSGA 502
         CYP26A1   CDWQLLNGPPTMKTSPTVYPVDNLPARFTHFHGEI------ 497
                   .:     ..  .  .
```

Figure 2. CYP26A1 sequence and ClustalW (1.82) alignment with CYP2C9, CYP2C8 and CYP3A4. The colour coding of amino acid type: red, small + hydrophobic (AVFPMILWY); blue, acid (DE); magenta, basic (RHK); green, hydroxyl + amine + basic (STYHCNGQ). Substrate recognition sites (SRS) and conserved P450 domains (ETLR, PERF and Haem binding domain) are indicated. "*" means that the residues are identical, ":" means that conserved substitutions have been observed, "." means that semi-conserved substitutions are observed.

plots obtained from the RAMPAGE server [26]. Amino acid environment was evaluated using Verify3D [27] and Errat plots [28]. Verify 3D assesses environment of the sidechain based on the solvent accessibility of the sidechain and the fraction of the sidechain covered by polar atoms. Errat assesses the distribution of different types of atoms with respect to one another in the protein model. Validation data for the templates CYP2C8, CYP2C9 and CYP3A4 was used as the baseline to assess the respective models.

Table II. Validation results for the lowest energy CYP26A1 models (produced from the three different templates) and the 3D structural templates.

	Ramachandran plot[a] (%)	Errat (%)	Verify 3D[b] (total score)
Model			
CYP26A1-CYP3A4 model	81.4	84.5	121
CYP26A1-CYP2C8 model	82.3	84.5	137
CYP26A1-CYP2C9 model	82.4	82.3	126
3D Template (Resolution)			
CYP3A4 (2.05 Å)	91.2	93.7	178
CYP2C8 (2.70 Å)	92.3	89.6	183
CYP2C9 (2.00 Å)	93.8	90.8	161

Validation results for the final CYP26A1 models built from the 1jpzB template. [a] Percentage of residues with ϕ, ψ conformation in the 'most favoured' regions of the Ramachandran plot. [b] The total Verify 3D score summed over all of the residues.

Docking

Ligands were docked within the active site of the homology model using the FlexX docking programme of SYBYL, performed with the default values. The active site was defined by all the amino acid residues within a 6.5 Å distance from TRP112, VAL116, THR304, VAL370 and GLY373, including the haem in a heteroatom file.

The output of FlexX docking was visualised in MOE and the scoring.svl script [29] was used to identify interaction types between ligand and protein.

Results and discussion

Comparative modelling methods use structural templates that have the highest sequence homology with the target protein. Homologous proteins were identified by scanning the protein sequence of CYP26A1 [30], obtained from the ExPASy server [23], against 3D structures deposited in the Protein Data Bank (PDB) [24] using PSI-BLAST [31]. The search returned amino acid sequences of different P450s isolated from different species, including the three human P450s recently deposited in the PDB. The percent sequence identity, chain length and E-value for the homologous P450 sequences are shown in Table I. CYP3A4 (1TQN), CYP2C8 (1PQ2) and CYP2C9 (1R90) were promising templates with percent sequence identity ranging from 24-26% and good E-values.

All three sequences have a similar chain length to the query sequence CYP26A1 and are high resolution structures. Structural alignment of the CYP26A1 with the three human P450 templates was performed using ClustalW [32] on the EBI server [33].

Alignment allowed the assignment of substrate recognition sites and P450 binding motifs (ETLR, PERF and haem binding domains) (Figure 2).
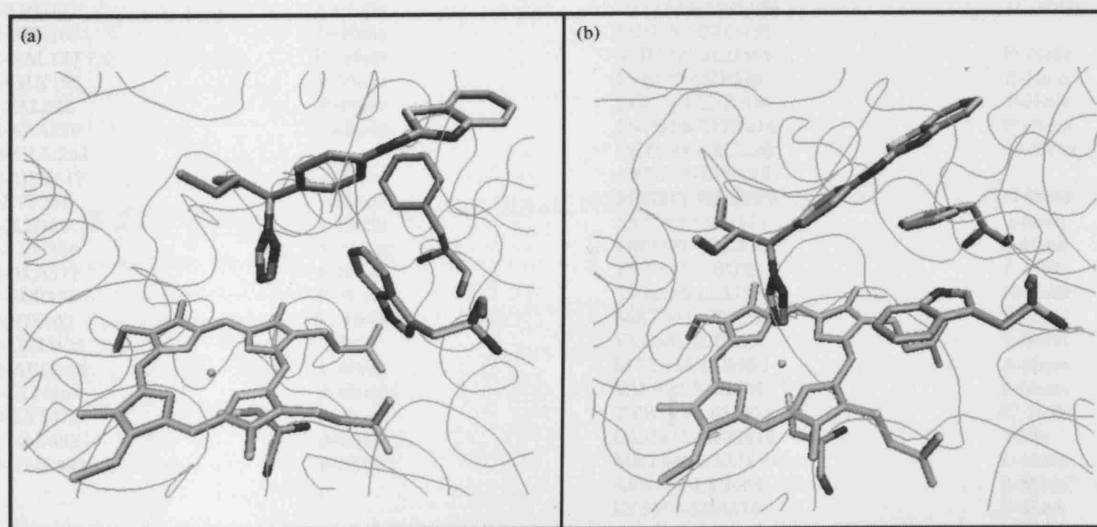

Figure 3.   CYP26A1 model active site (a) before and (b) after active site optimisation with the (*S*)-R115866 bound inhibitor.
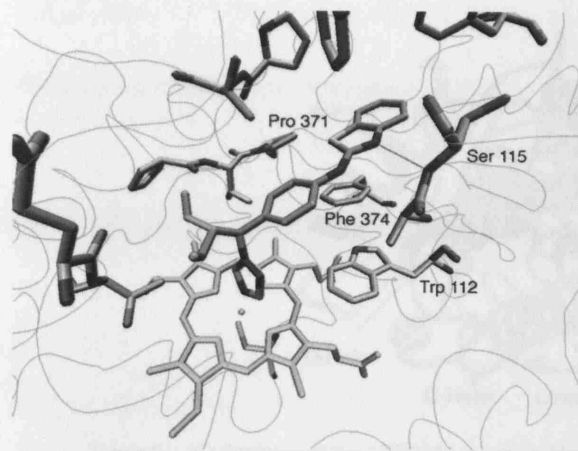
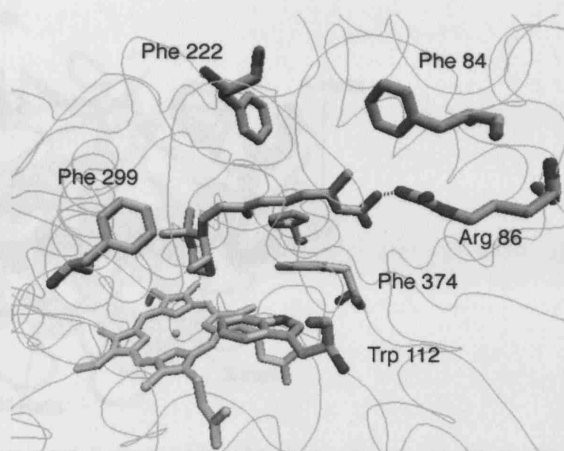Figure 4. Interactions between (S)-R115866 and the CYP26A1 active site.

Figure 5. atRA docked in CYP26A1 active site.

Models, based on each template using a single template alignment approach, were generated using Molecular Operating Environment (MOE) [18] software as described in the Experimental section. To evaluate the quality of the modelled structures, the lowest energy model generated from each template was subjected to a number of checks. Stereochemical quality was assessed using Ramachandran plots [34], using the Cambridge RAMPAGE server [26], and amino acid environment was assessed using Verify3D

Table III. **3.** Comparison of CYP26A1 model and CYP3A4 template secondary structure.

| CYP26A1 Model		CYP3A4 Template	
Residues	Secondary structure	Residues	Secondary structure
LEU58-MET60	Helix	LEU32-LEU36	Helix
ARG64-TYR75	A-Helix	PHE57-TYR68	A-Helix
TYR79-LEU83	β-Sheet	LYS70-GLY77	β-Sheet
ARG86-VAL91	β-Sheet	GLN79-THR85	β-Sheet
ALA94-LEU101	}B-Helix	PRO87-LEU94	B-Helix
ASP105-LEU107		MET114-LYS115	B'-Helix
SER132-ARG142	C-Helix	GLU124-ARG128	}C-Helix
ARG146-SER169	D-Helix	LEU132-PRO135	
ARG173-VAL177	D'-Helix	SER139-GLU165	D-Helix
PHE186-GLY195	E-Helix	LYS173-SER186	E-Helix
ASP227-VAL228	F'-Helix	PRO202-LYS208	F-Helix
GLY232-ALA239	}G-Helix	PRO218-THR224	F'-Helix
GLU248-GLU262		PHE228-LEU236	}G-Helix
GLN288-GLY317	I-Helix	ARG243-LYS257	
HIS321-LYS331	J-Helix	PHE271-GLN279	H-Helix
ASP346-ILE349	J'-Helix	ASP292-LEU321	I-Helix
VAL359-LEU366	K-Helix	PRO325-VAL338	J-Helix
GLY373-ALA377	β-Sheet	TYR347-LEU351	J'-Helix
TRP392-SER397	β-Sheet	TYR355-LEU366	K-Helix
CYS399-HIS402	K'-Helix	MET371-CYS377	β-Sheet
PRO422-GLU425	Helix	VAL381-ILE383	β-Sheet
GLU446-ARG461	L-Helix	MET386-ILE388	β-Sheet
ASP464-GLN466	β-Sheet	VAL392-SER398	β-Sheet
THR473-LYS475	β-Sheet	TYR399-LEU401	K'-Helix
TYR481-VAL483	β-Sheet	GLU417-ARG418	Helix
ARG489-THR491	β-Sheet	MET445-LEU460	L-Helix
		ASN462-LYS466	β-Sheet
		LYS476-SER478	β-Sheet
		GLN484-GLU486	β-Sheet
		VAL489-ARG496	β-Sheet
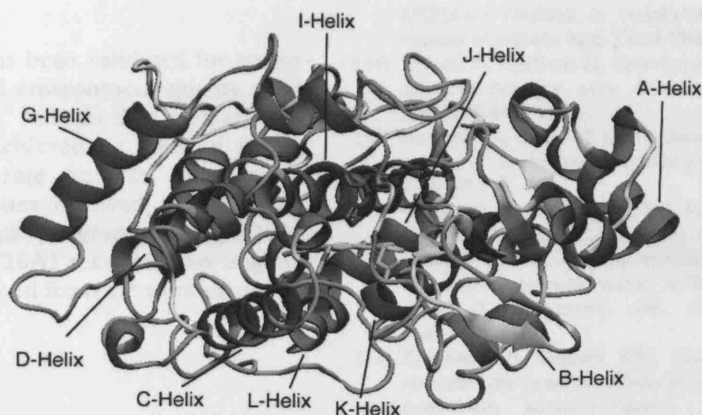
Figure 6.   3D Structure of the CYP26A1 model. α-Helices are in red, β-sheets are in yellow, loops are cyan and coils are grey.

[27] and Errat [28], using the UCLA-DOE server [35] (Table II).

Validation results would suggest that all three models performed equally well in terms of mainchain stereochemistry and amino acid environment, therefore substrate and ligand docking was required for further validation of the active site architecture.

Ligand Docking

atRA was found to dock satisfactorily in the active site of all three models, with atRA orientated in a position that favours 4-hydroxylation at C-4. However, only the CYP26A1 model built using the CYP3A4 template was able to accommodate the inhibitor ligand (S)-R115866 [8] in an orientation that would allow coordination between the nitrogen of the triazole heterocycle and the haem iron transition metal. However, docking studies had indicated that TRP112 was partially obstructing coordination between the triazole nitrogen of the ligand R115866 and the transition metal resulting in a distance of 5.50 Å (Figure 3a). Docking studies do not take into account protein flexibility therefore molecular dynamics were performed on the active site containing (S)-R115866 resulting in an optimised active site architecture. This resulted in the TRP112 residue positioned in a more favourable conformation with respect to ligand binding (Figure 3b) with the triazole nitrogen perpendicular to the haem iron at a distance of 2.6 Å.

Furthermore, (S)-R115688 establishes a hydrogen bond between the benzothiazole nitrogen and the NH of SER115, as well as several hydrophobic interactions with the side chains of TRP112, PHE374, PHE84, PHE299, VAL116, PRO371 and other residues (Figure 4).

The natural substrate atRA was docked with the optimised active site. Figure 5 shows the putative active site of CYP26A1 containing the bound substrate orientated for 4-hydroxylation with the C4 atom positioned above the haem iron at a distance of 5.3 Å, this distance would accommodate a water molecule between the 4-position of atRA and the haem iron. atRA interacts with amino acid residues at the active site by multiple hydrophobic interactions, including the side chains of TRP112, PHE299, PHE222, PHE84, PHE374 and PRO371. Hydrogen bonding interactions between the carbonyl group of atRA and ARG86 hold the molecule within the hydrophobic tunnel.

The secondary structure of the CYP26A1 model was determined using Swiss-PdbViewer 3.7 [36] with identification of α-helices, β-sheets, coils and loops as shown in Table III. The secondary structure is derived from the template structure CYP3A4 therefore further comparison with the template allows identification of specific helices and β-sheets. Studies have shown that the active site region (Figure 2) of the P450s is conserved across the entire family [37] and the middle section of the I-helix (Figure 2) is also well conserved despite low overall sequence homology [38]. However, although similarities do exist across the family, there are flexible regions in each P450 which have significantly different amino acid composition, such as the B'-helix, which may or may not be present, and F-G loop regions.

A good overlap between the model and template structure was found in particular for the A-, E- G-, I-, J-, J'- and K-helices. The model differs from the template in that the B'-, F- and H-helices are not identified in the model. As seen in Figure 6, the large I-helix lies above the haem and is connected to the J-helix through a loop. The active site cysteine 442 lies below the haem within a coil and it is also worth noting that all the residues involved in the binding of atRA and (S)-R115688, are located in the substrate recognition sites (SRSs)(Figure 2).

Conclusions

The CYP26A1 model has been validated for stereo-chemical and amino acid environment quality using appropriate programmes, with further validation of active site architecture achieved by docking studies with the natural substrate and an inhibitor to determine the key residues involved in the ligand binding. A drug design program of CYP26A1 inhibitors using this CYP26A1 model is now ongoing based on the results obtained from the work presented here.

Acknowledgements

References

[1] Sporn MB, Roberts AB, Goodman DS, editors. The Retinoids: Biology Chemistry and Medicine New York: Academic Press; 1984. p 424.

[2] Sonneveld E, Van den Brink CE, Van der Leede BJ, Schulkes RK, Petkovich M, Van der Burg B, Van der Saag PT. Human retinoic acid (RA) 4-hydroxylase (CYP26) is highly specific for all-trans-RA and can be induced through RA receptors in human breast and carcinoma cells. Cell Growth Differ 1998;9:629–637.

[3] Johnson A, Chandraratna RAS. Novel retinoids with receptor selectivity and functional selectivity. Br J Dermatol (Suppl) 1999;140:12–17.

[4] Torma H, Rollman O, Vahlquist A. The vitamin A metabolism and expression of retinoid-binding proteins differ in HaCaT cells and normal human keratinocytes. Arch Dermatol Res 1999;291:339–345.

[5] Smith HJ, Nicholls PJ, Simons C, LeLain R. Inhibitors of steroidogenesis as agents for the treatment of hormone-dependent cancers. Exp Opin Ther Patents 2001;11:789–824.

[6] Marill J, Cresteil T, Lanotte M, Chabot GG. Identification of human cytochrome P450s involved in the formation of all-trans-retinoic acid principal metabolites. Mol Pharmacol 2000;58:1341–1348.

[7] Ray WJ, Bain G, Yao M, Gottlieb DI. CYP26 a novel mammalian cytochrome P450 is induced by retinoic acid and defines a new family. J Biol Chem 1997;272:18702–18708.

[8] Stoppie P, Borgers M, Borghgraef P, Dillen L, Goossens J, Sanz G, Szel H, Van Hove C, Van Nyen G, Nobels G, Vanden Bossche H, Venet M, Willemsens G, Van Wauwe J. R115866 inhibits all-trans-retinoic acid metabolism and exerts retinoidal effects in rodent. J Pharmacol Exp Ther 2000;293:304–312.

[9] MacLean G, Abu-Abed S, Dolle P, Tahayato A, Chambon P, Petkovich M. Cloning of a novel retinoic-acid metabolizing cytochrome P450 Cyp26B1 and comparative expression analysis with Cyp26A1 during early murine development. Mech Develop 2001;107:195–201.

[10] Taimi M, Helvig C, Wisniewski J, Ramshaw H, White J, Amad M, Korczak B, Petkovich M. A novel human cytochrome P450 CYP26C1 involved in metabolism of 9-cis and all-trans isomers of retinoic acid. J Biol Chem 2004;279:77–81.

[11] Ahmad N, Mukhtar H. Cytochrome P450: A target for drug development for skin diseases. J Investig Dermatol 2004;123:417–425.

[12] Brecher AR, Orlow SJ. Oral retinoid therapy for dermatologic conditions in children and adolescents. J Am Acad Dermatol 2003;49:171–182.

[13] Weiss GR, Liu PY, Alberts DS, Peng YM, Fisher E, Xu MJ, Scudder SA, Baker LH, Moore DF, Lippman SM. 13-cis-Retinoic acid or all-trans-retinoic acid plus inteferon-alpha in recurrent cervical cancer: A Southwest Oncology Group phase II randomized trial. Gynecol Oncol 1998;71:386–390.

[14] Pettersson F, Colston KW, Dalgleish AG. Retinoic acid enhances the cytotoxic effects of gemcitabine and cisplatin in pancreatic adenocarcinoma cells. Pancreas 2001;23:273–279.

[15] Schoch GA, Yano JK, Wester MR, Griffin KJ, Stout CD, Johnson EF. Structure of human microsomal cytochrome P450 2C8 Evidence for a peripheral fatty acid binding site. J Biol Chem 2004;279:9497–9503.

[16] Williams PA, Cosme J, Ward A, Angove HC, Matak-Vinkovic D, Jhoti H. Crystal structure of human cytochrome P450 2C9 with bound warfarin. Nature 2003;424:464–468.

[17] Yano JK, Wester MR, Schoch GA, Griffin KJ, Stout D, Johnson EF. The structure of human microsomal cytochrome P450 3A4 determined by X-ray crystallography to 2.05-Å resolution. J Biol Chem 2004;279:38091–38094.

[18] Molecular Operating Environment. (MOE) Chemical Computing Group Inc Montreal Quebec Canada http://www.chemcomp.com 2004.03.

[19] Tripos SYBYL 7.0, Tripos Inc. South Hanley Rd St Louis Missouri 63144 USA http://www.tripos.com 1699.

[20] Berendsen HJC, van der Spoel D, van Drunen R. GROMACS: A message-passing parallel molecular dynamics implementation. Comp Phys Comm 1995;91:43–56.

[21] Lindahl E, Hess B, van der Spoel D. GROMACS 3.0: A package for molecular simulation and trajectory analysis. J Mol Mod 2001;7:306–317.

[22] Humphrey W, Dalke A, Schulten K. VMD - Visual Molecular Dynamics. J Mol Graph 1996;14:33–38.

[23] The ExPASy (Expert Protein Analysis System)., proteomics server of the *Swiss Institute of Bioinformatics* (SIB) http://ca.expasy.org

[24] R CSB Protein Data Bank (PDB)., http://www.rcsb.org/pdb

[25] Needleman SB, Wunsch CD. A general method applicable to search for similarities in amino acid sequence of 2 proteins. J Mol Biol 1970;48:443–453.

[26] RAMPAGE Server., http://ravenbioccam.ac.uk/rampage.php

[27] Bowie JU, Luthy R, Eisenberg D. A method to identify protein sequences that fold into a known three-dimensional structure. Science 1991;253:164–170.

[28] Colovos C, Yeates TO. Verification of protein structures: Patterns of nonbonded atomic interactions. Protein Science 1993;2:1511–1519.

[29] Code "scoring svl" obtained from SVL Exchange website., http://svl.chemcomp.com Chemical Computing Group Inc Montreal Canada.

[30] White JA, Beckett-Jones B, Guo Y-D, Dilworth FJ, Bonasoro J, Jones G, Petkovich M. cDNA cloning of human retinoic acid-metabolizing enzyme (hP450RAI) identifies a novel family of cytochromes P450. J Biol Chem 1997;272:18538–18541.

[31] Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res 1997;25:3389–3402.

[32] Higgins D, Thompson J, Gibson T, Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673–4680.

[33] ClustalW WWW Service at the European Bioinformatics Institute., http://www.ebi.ac.uk/clustalw

[34] Lovell SS, Davis IW, Arendall III WB, e Bakker PIW, Word JM, Prisant MG, Richardson JS, Richardson DC. Structure validation by Cα geometry: Φ, Ψ and Cβ deviation. Proteins: Structure Function & Genetics 2002;50:437–450.

[35] UCLA-DOE Institute for Genomics & Proteomics Server., -mbiucla.edu/services">http://www.doe -mbiucla.edu/Services.

[36] Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modelling. Electrophoresis (http://wwwexpasyorg/spdbv/) 1997;18: 2714–2723.

[37] Hasemann CA, Kurumbail SS, Boddupalli SS, Peterson JA, Deisenhofer J. Structure and function of cytochrome P450: A comparative analysis of three crystal structures. Structure 1995;:41–62.

[38] Podust LM, Poulos TL, Waterman MR. Crystal structure of cytochrome P450 14(-sterol demethylase (CYP51) from Mycobacterium tuberculosis in complex with azole inhibitors. PNAS 2001;98:3068–3073.

Abstracts

- Gabriella De Martino, Giuseppe La Regina, **Maria Chiara Barbera**, Andrea Brancale, Ernest Hamel, Marino Artico, Romano Silvestri. Potenti Inibitori della Tubulina a Struttura Indolica. *XVII Convegno Nazionale della Divisione di Chimica Farmaceutica della Società Chimica Italiana*, (*XVII National Conference of the Pharmaceutical Chemistry Department of the Italian Chemical Society*). Acts of the conference P-60, pg 126, 6 - 10 September 2004, Pisa (Italy).

- Gabriella De Martino, Giuseppe La Regina, **Maria Chiara Barbera**, Andrea Brancale, Ernest Hamel, Marino Artico, Romano Silvestri. Arylthioindoles, Potent Inhibitors of Tubulin Polymerization. Istituto Pasteur – Fondazione Cenci Bolognetti. Giornate Scientifiche Della Fondazione. Molecular Recognition (Pasteur Institute – Cenci Bolognetti Foundation. Molecular Recognition) . Rome (Italy), 2-3 December 2004.

- G. De Martino, G. La Regina, **M. C. Barbera**, A. Brancale, E. Hamel, M. Artico, R. Silvestri. New Agents Highly Active in Cancer Therapy. Indoles, Potent Inhibitors of Tubulin Polymerization. Conferenza sulla Ricerca Scientifica Facoltà di Farmacia (Scientific Research Conference – Faculty of Pharmacy), *Dalle molecole agli organismi (From molecules to organisms)*, Summary, P44, Rome (Italy) 9-10 December 2004.

- De Martino, G.; La Regina, G.; Coluccia, A.; Edler, M. C.; **Barbera, M.C.**; Brancale, A.; Wilcox, E.; Hamel, E.; Artico, M.; Silvestri, R. Arylthioindoles, Potent Inhibitors of Tubulin Polymerization. 20th Internationational Congress Heterocyclic Chemistry, Palermo (Italy) 31 July – 05 August 2005.

- Gabriella De Martino, Giuseppe La Regina, Antonio Coluccia, Francesco Piscitelli, Michael C. Edler, **Maria Chiara Barbera**, Andrea Brancale, Ernest Hamel, Marino Artico and Romano Silvestri. Arylthioindoles, Potent Inhibitors of Tubulin Polymerization. Second Joint Italian – Swiss Meeting on Medicinal Chemistry, Modena (Italy) 12-16 September 2005.