Structural Biology in Antiviral Drug Discovery

Marcella Bassetto,1 Alberto Massarotti,2 Antonio Coluccia3 and Andrea Brancale*1

1School of Pharmacy & Pharmaceutical Sciences, Cardiff University, King Edward VII Avenue, Cardiff, CF10 3NB, UK.
2Dipartimento di Scienze del Farmaco, Università degli Studi del Piemonte Orientale A. Avogadro Largo Donegani 2, 28100 Novara, Italy.
3Istituto Pasteur Italia– Fondazione Cenci Bolognetti, Dipartimento di Chimica e Tecnologie del Farmaco, Sapienza Università di Roma, Piazzale Aldo Moro 5, I-00185 Roma, Italy.

Abstract
Structural biology has emerged during the last thirty years as a powerful tool for rational drug discovery. Crystal structures of biological targets alone and in complex with ligands and inhibitors provide essential insights into the mechanisms of actions of enzymes, their conformational changes upon ligand binding, the architectures and interactions of binding pockets. Structure-based methods such as crystallographic fragment screening represent nowadays invaluable instruments for the identification of new biologically active compounds. In this context, three-dimensional protein structures have played essential roles for the understanding of the activity and for the design of novel antiviral agents against several different viruses. In this review, the evolution in the resolution of viral structures is analysed, along with the role of crystal structures in the discovery and optimisation of new antivirals.

Keywords
X-ray crystallography, PDB, structure-based drug discovery, antiviral agents.

Abbreviations
DENV Dengue virus
EV71 Enterovirus 71
FOAMV Simian foamy virus
HCV Hepatitis C virus
HHV Human herpesvirus
HIV-1 Human immunodeficiency virus 1
MLV Murine leukemia virus
NV Norovirus
RSV-SRA Rous sarcoma virus
SARS-CoV Severe acute respiratory syndrome-related coronavirus, Human SARS coronavirus

1. Introduction
Knowledge of the three-dimensional structures of proteins has been long recognised as a powerful tool to accelerate drug discovery, providing information about target shape, hydrophobic and hydrophilic behaviours of macromolecules and interactions with substrates [1]. Since 1934, when the first X-ray diffraction structure was reported for pepsin, this method emerged as an invaluable source of detailed
and reliable information about protein structure, representing a significant step forward in comparison with previously used physical or chemical methods [2].

The idea that three-dimensional structural information could be useful in defining topographies of the complementary surfaces of ligands and their protein targets raised in the early 1980s, when scientists started to use this information to optimise potency and selectivity of lead compounds [3]. Structural details gathered from a target structure in the presence of unique ligands can provide fundamental insights on the geometric fit of these compounds into the binding site, on the binding of active conformations, on molecular electrostatic potentials and on hydrophobic interactions [4]. From 1980 to nowadays, the applicability of structural biology was extended to the assessment of target druggability, to the identification of hits by virtual screening with structure-based virtual screening methods, to target identification by structure-sequence homology recognition [5].

During the last three decades, the application of structural biology to drug discovery followed a classical path. Structure-based became extremely fashionable during the 1980s, as a consequence of the publication of the structures of the first important drug targets. Later on, due the limited number of protein structures available and to the cost and time required to set up the crystallization process, the use of these rational approaches dropped in favour of high-throughput screening methods and combinatorial chemistry [5]. The successful story of HIV protease inhibitors [6,7] and influenza antiviral drug Relenza [8] led to a renewed interest in target structure-driven drug discovery. Furthermore, the numerous advances in science and technology promoted a faster and less expensive application of structural biology, increasing the speed of macromolecular structure determination, increasing the resolution of new crystal structures and allowing smaller amounts of protein and fewer crystals to be required to solve a structure [9].

One of the most recent applications of structural biology is crystallographic fragment screening [1]. The low affinity for the target that characterises chemical fragments has made them unsuitable for classical high throughput screening, but the advances in high throughput NMR and crystallography have allowed the use of structural information of protein–fragment complexes, which provide reliable proof of binding pockets and hit binding mode, and give clear indications on how the fragment structures can be optimised into potent lead compounds [10]. These characteristics make fragments very attractive starting points for iterative medicinal chemistry optimization [1].

Several biologically active compounds discovered by structure-based design are now drugs in the market, confirming the crucial role played by structural biology in drug development [11], while protein structure determination has become one of the earliest and most crucial steps for many drug-discovery programs of pharmaceutical companies [1].

2. Viral structures in the Protein Data Bank

The Protein Data Bank (PDB) was established in 1971 as a curated archive that evolves with new developments in structural biology [12]. The holdings in the PDB continue to grow and the usage of PDB data is also growing. In 2015, 526 million downloads (ftp and website) of data occurred from the PDB sites, compared to 226 million downloads in 2008. Download statistics for the overall archive and for individual entries are available from the wwPDB website (http://www.wwpdb.org/stats/download ).
The first atomic viral structure was published less than 40 years ago [13], while now about 7,852 viral related structures are available in the PDB from about 475 different virus, more than 80% of them containing viruses only (Figure 1A-B). Most viral structures in the PDB (around 89%), have been determined using X-ray crystallography (Figure 1B). Since the establishment of this archive, the number of structure depositions has grown steadily. The average resolution of X-ray structures has remained constant at about 2.54 Å (Figure 1C). However, with the large volume of data available, there are now substantial numbers of structures determined to a high resolution (below 1.5 Å), including at least one viral structure [14]. At the same time, as more large macromolecular machines are being studied using X-ray methods, there are many examples of low resolution structures (above 3.5 Å) [15-17].

Figure 1. Statistical details of 7,852 PDB entries considered. RCSB Protein Data Bank was used to retrieve the viral PDB entries using “TAXONOMY is Viruses” as a search string. (A) Taxonomy, (B) detailed taxonomy, (C) experimental method, and (D) resolutions.

Over the years the number of viral PDB entries has increased at an increasingly faster rate (Figure 2A). Analysis of the taxonomy in the PDB shows that the most studied viruses are, respectively, HIV-1, Enterobacteria phage, Influenza A virus, HCV, Human herpesvirus, SARS-CoV, Dengue virus, Norwalk virus, Vaccinia virus, Enterovirus C, Enterovirus A, Influenza B virus, MLV, RSV-SRA and FOAMV (Figure 2B). This is most likely due to the important roles that these viruses play in biomedical research.

Figure 2. Growth in the numbers of viral PDB entries. The large increase shown from 1984 was due to the release of the tomato bushy stunt virus (PDB id: 2tbv) [13]. (A) Total number of entries, (B) details of viruses reported more than 50 times.
For each top virus a sequence cluster analysis (Table 1) was performed using a sequence identity of 50% as cut-off. More clusters than expected were obtained, as their number is bigger than the number of proteins of each viral proteome.

<table>
<thead>
<tr>
<th></th>
<th>HIV-1</th>
<th>Enterobacteria phage</th>
<th>Influenza A virus</th>
<th>HCV</th>
<th>Human herpesvirus</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDB entries</td>
<td>1617</td>
<td>1318</td>
<td>662</td>
<td>377</td>
<td>276</td>
</tr>
<tr>
<td>Clusters</td>
<td>108</td>
<td>242</td>
<td>47</td>
<td>32</td>
<td>78</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>SARS-CoV</th>
<th>Dengue virus</th>
<th>Norwalk virus</th>
<th>Vaccinia virus</th>
<th>Enterovirus C</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDB entries</td>
<td>170</td>
<td>127</td>
<td>110</td>
<td>110</td>
<td>84</td>
</tr>
<tr>
<td>Clusters</td>
<td>37</td>
<td>19</td>
<td>7</td>
<td>40</td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Enterovirus A</th>
<th>Influenza B virus</th>
<th>MLV</th>
<th>RSV-SRA</th>
<th>Simian foamy virus</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDB entries</td>
<td>58</td>
<td>54</td>
<td>53</td>
<td>52</td>
<td>51</td>
</tr>
<tr>
<td>Clusters</td>
<td>6</td>
<td>12</td>
<td>15</td>
<td>17</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 1. Diversity of PDB entries. All viral PDB entries were analysed using a sequence identity of 50% as cut-off to identify clusters.

Biomedical research on viruses does not involve only structural biology, as highlighted in Figure 3, where the general interest of scientists is compared with the entries of viruses in the PDB. A search in the PubMed database using each virus name revealed the number of original research papers published...
per virus. Surprisingly, the growth rate of publications in PubMed is not always correlated with the number of PDB entries.

Figure 3. Analysis of the top viruses in the PDB versus the corresponding publications in PubMed.

The number of ligand/viral PDB entries continues to increase; there are now more than 5,200 complexes with ligands, including different marketed drugs (Table 2).

<table>
<thead>
<tr>
<th>Name</th>
<th>Structure</th>
<th>PDB entry</th>
<th>Number of entries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aciclovir</td>
<td><img src="image1" alt="Aciclovir Structure" /></td>
<td>AC2</td>
<td>4</td>
</tr>
<tr>
<td>Amantadine</td>
<td><img src="image2" alt="Amantadine Structure" /></td>
<td>308</td>
<td>4</td>
</tr>
<tr>
<td>Name</td>
<td>Structure</td>
<td>PDB entry</td>
<td>Number of entries</td>
</tr>
<tr>
<td>-----------------</td>
<td>-----------</td>
<td>-----------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Amprenavir</td>
<td><img src="image" alt="Amprenavir Structure" /></td>
<td>478</td>
<td>18</td>
</tr>
<tr>
<td>Atazanavir</td>
<td><img src="image" alt="Atazanavir Structure" /></td>
<td>DR7</td>
<td>10</td>
</tr>
<tr>
<td>Darunavir</td>
<td><img src="image" alt="Darunavir Structure" /></td>
<td>017</td>
<td>49</td>
</tr>
<tr>
<td>Efavirenz (Sustiva)</td>
<td><img src="image" alt="Efavirenz Structure" /></td>
<td>EFZ</td>
<td>6</td>
</tr>
<tr>
<td>Foscarnet</td>
<td><img src="image" alt="Foscarnet Structure" /></td>
<td>PPF</td>
<td>7</td>
</tr>
<tr>
<td>Name</td>
<td>Structure</td>
<td>PDB entry</td>
<td>Number of entries</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------</td>
<td>-----------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Ganciclovir</td>
<td><img src="https://example.com/ganciclovir_structure.png" alt="Ganciclovir Structure" /></td>
<td>GA2</td>
<td>2</td>
</tr>
<tr>
<td>Idoxuridine</td>
<td><img src="https://example.com/idoxuridine_structure.png" alt="Idoxuridine Structure" /></td>
<td>ID2</td>
<td>1</td>
</tr>
<tr>
<td>Indinavir</td>
<td><img src="https://example.com/indinavir_structure.png" alt="Indinavir Structure" /></td>
<td>MK1</td>
<td>15</td>
</tr>
<tr>
<td>Lopinavir</td>
<td><img src="https://example.com/lopinavir_structure.png" alt="Lopinavir Structure" /></td>
<td>AB1</td>
<td>10</td>
</tr>
<tr>
<td>Nelfinavir</td>
<td><img src="https://example.com/nelfinavir_structure.png" alt="Nelfinavir Structure" /></td>
<td>1UN</td>
<td>10</td>
</tr>
<tr>
<td>Oseltamivir</td>
<td><img src="https://example.com/oseltamivir_structure.png" alt="Oseltamivir Structure" /></td>
<td>G39</td>
<td>28</td>
</tr>
<tr>
<td>Name</td>
<td>Structure</td>
<td>PDB entry</td>
<td>Number of entries</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------</td>
<td>-----------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Penciclovir</td>
<td><img src="image1" alt="Penciclovir Structure" /></td>
<td>PE2</td>
<td>2</td>
</tr>
<tr>
<td>Raltegravir</td>
<td><img src="image2" alt="Raltegravir Structure" /></td>
<td>RLT</td>
<td>4</td>
</tr>
<tr>
<td>Ribavirin</td>
<td><img src="image3" alt="Ribavirin Structure" /></td>
<td>RBV</td>
<td>2</td>
</tr>
<tr>
<td>Rimantadine</td>
<td><img src="image4" alt="Rimantadine Structure" /></td>
<td>RIM</td>
<td>2</td>
</tr>
<tr>
<td>Ritonavir</td>
<td><img src="image5" alt="Ritonavir Structure" /></td>
<td>RIT</td>
<td>13</td>
</tr>
<tr>
<td>Saquinavir</td>
<td><img src="image6" alt="Saquinavir Structure" /></td>
<td>ROC</td>
<td>25</td>
</tr>
</tbody>
</table>
### Table 2.
List of approved antiviral drugs currently available in the PDB in complex with their target.

<table>
<thead>
<tr>
<th>Name</th>
<th>Structure</th>
<th>PDB entry</th>
<th>Number of entries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simeprevir</td>
<td><img src="image" alt="Structure" /></td>
<td>30B</td>
<td>1</td>
</tr>
<tr>
<td>Sofosbuvir (diphosphate metabolite)</td>
<td><img src="image" alt="Structure" /></td>
<td>6GS</td>
<td>1</td>
</tr>
<tr>
<td>Tenofovir</td>
<td><img src="image" alt="Structure" /></td>
<td>TFO</td>
<td>1</td>
</tr>
<tr>
<td>Valaciclovir</td>
<td><img src="image" alt="Structure" /></td>
<td>TXC</td>
<td>1</td>
</tr>
<tr>
<td>Zanamivir</td>
<td><img src="image" alt="Structure" /></td>
<td>ZMR</td>
<td>20</td>
</tr>
</tbody>
</table>

#### Table 2. List of approved antiviral drugs currently available in the PDB in complex with their target.

3. **Recent applications of structural biology in antiviral research**

One of the most striking examples of successful application of structural biology for antiviral drug discovery is represented by HIV-1 protease and reverse-transcriptase inhibitors. HIV-1 was recognised as the responsible for the acquired immune deficiency syndrome (AIDS) in the early 80s [18], and the discovery of the first selective HIV-1 protease inhibitors is still one of the most popular examples of the use of X-ray crystallography in the development of a drug in clinical use. The HIV protease was validated as a potential drug target in 1985 [19,20], the first X-ray crystal structures of the enzyme began appearing in 1989 [21-24] and the first HIV protease inhibitor Saquinavir was licensed only six years later, followed by the approval of Ritonavir four months later [25]. Three fundamental steps led to Saquinavir discovery,
the first one being the classification of HIV protease as a member of the aspartate protease family, which comprises also pepsin and renin [26]. Homology with renin, already a target in the design of anti-hypertensive agents, suggested a potential approach for the development of selective inhibitors of this enzyme [26], and this research interest was reinforced by the resolution of the first crystal structures for HIV [21,24], along with the protease structure of the related Rous sarcoma virus [22]. Finally, useful information on protease inhibition by transition-state analogue inhibitors [27-29] guided a series of investigations on the minimum size required for a small molecule to inhibit this enzyme. Extensive structure-activity relationship studies and X-ray experiments directed to address this aspect resulted in the discovery of Saquinavir.

A key step in HIV-1 life cycle is reverse transcription (RT), therefore the RT/DNA polymerisation has been immediately considered as a prime drug target, with the first approved anti-AIDS drug being the nucleoside analogue AZT (zidovudine, ZDV) in 1987 [30].

The HIV reverse transcriptase (RT) is a heterodimer consisting of two polypeptide chains, p66 and p51 (Figure 4). The p66 chain contains an N-terminal polymerase domain and a C-terminal RNase H domain [31,32]. The subdomains in p66 are flexible and can rearrange to different conformational states, required to carry out the enzyme essential functions for the viral replication. Sequencing of the complete RT from clinical isolates have shown that mutations in the remote connection subdomain and the RNase H domain enhance resistance to both nucleoside (NRTIs) and non-nucleoside (NNRTIs) inhibitors [33,34], following indirect mechanisms that are not well understood.

Figure 4. HIV-1 RT structure (PDB ID: 5D3G), p66 and p55 subunits are depicted in light green and light orange, respectively.

Since the publication of the first crystal structure of the HIV reverse-transcriptase in complex with the non-nucleoside inhibitor Nevirapine in 1994 [35], the resolution of the structures of several conformations of the complex with different inhibitors has provided an extremely powerful tool for gaining essential insights into the mechanism of action of this enzyme, for understanding the importance of its flexibility and its different conformational states, for rationalising the occurrence of resistance, for elucidating the binding of nucleoside and non-nucleoside inhibitors and for the design and optimisation of new chemical agents targeting this protein.

Among the several examples available on how structural biology has been essential for the design and optimisation of new non-nucleoside inhibitors of the HIV-1 RT, the resolution of a crystal structure of
the enzyme in complex with the RNase inhibitor dihydroxy benzoyl naphthyl hydrazine in 2006 (DHBNH, Figure 5) has led to the discovery of a novel site of the protein, near both the polymerase active site and the NNRTI binding pocket [36]. Structure-based modifications on the DHBNH scaffold resulted in the identification of dual inhibitors of both the polymerase and the RNH activities of the HIV-1 RT (exemplified by compound 1 in Figure 5).

More recently, following the resolution of a crystal structure of HIV-1 RT containing the NNRTI TMC278 [37] in the DNA polymerase domain and α-hydroxytropolone manicol in the RNase H active site, the structure of manicol was rationally modified to obtain a new series of α-hydroxytropolones, which show antiviral activities at non-cytotoxic concentrations and occupy an additional site surrounding the DNA polymerase catalytic centre (compound 2 in Figure 5) [38].

In 2013, with the application of an X-ray crystallographic fragment screening methodology to evaluate the intrinsic flexibility of the RT for the discovery of new allosteric sites, seven new sites were identified within this protein [39]. Three of these sites (named the Knuckles, the NNRTI Adjacent and the Incoming Nucleotide binding sites) were proven inhibitory in an enzymatic assay, while the co-crystallised fragments (3a-c in Figure 5) were found to be novel scaffolds in comparison with previously reported RT inhibitors, thus providing the basis for the development of novel leads [39].

After the resolution of a crystal structure of HIV-1 RT in complex with potent pyrimidine-based NNRTI 4a, structure-based modifications on its chemical scaffold directed to the occupation of the entrance channel to the NNRTI binding site resulted in the identification of much more soluble analogues such as 4b (Figure 5), with which the solubility issues associated with 4a were significantly improved [40].

In 2011, the study of several reported structures of the enzyme in complex with Efavirenz and other non-nucleoside inhibitors, and the inspection of the ligand geometries required for the interaction with the enzyme revealed in these structures, guided the design of a new series of aryl-phospho-indoles as potent inhibitors of the enzyme (exemplified by compound 5a in Figure 5) [41]. Subsequent rational optimisation of the original lead 5a resulted in the identification of 5b, a nanomolar inhibitor of the Y181C/K103N double mutant (both mutations are clinically relevant) that reached phase IIb clinical trials [42].

In the same year, multiple RT crystal structures have been used for the in silico screening of a library of more than two million compounds using molecular docking methods, taking into account different protein conformations in order to overcome resistance [43]. One hit was found with 4.8 µM potency against WT HIV-1 (compound 6a in Figure 5). Computational analyses and rational modifications on the structure of 6a led to the discovery of catechol diether 6b, a 55-pM anti-HIV agent that retains nanomolar activity against the Y181C mutant [43].
Structural biology has often been extremely helpful for the identification or optimisation of novel antiviral agents also in the case of HCV, in particular for the discovery of novel inhibitors of the NS3-4a protein and the NS5b polymerase.

In 2012, a fragment-based screening of 176 fragments against the full length HCV NS3-NS4a genotype 1b enzyme led to the discovery of a new allosteric pocket at the interface between the protease and helicase domains [44]. Structure-based optimisation of a first hit found to bind this new pocket (compound 7a in Figure 6) guided the identification of potent inhibitor 7b (Figure 6), which also inhibits the viral replication with an EC_{50} value in the nanomolar range [44].

Crystallographic fragment-based screening methodologies have proven successful also in the identification of non-nucleoside inhibitors of the HCV NS5b polymerase. In 2008, a small bromo-aryl fragment (8a) was found to bind the thumb domain of the protein with an initial binding affinity in the millimolar range [45]. A series of structure-based optimization cycles on the fragment scaffold led to the identification of a family of structures with high affinity for the enzyme and low micromolar activities in the HCV replicon assay (compound 8b in Figure 6).

More recently, the application of a similar fragment-based approach guided the identification of sulfonamide fragment 9a (Figure 6), which binds the polymerase allosteric thumb pocket 2 [46]. The scaffold of this small fragment was the starting point for different structure-based modifications, which resulted in the identification of phenoxyantranilic acid sulfonamide derivative 9b as a 650-fold more potent inhibitor of the HCV NS5b polymerase [46]. Further structure-based optimisation attempts on the scaffold of 9b led to the discovery of derivative 9c, which is slightly more potent in inhibiting the enzyme and shows a much more potent inhibition of the viral replication in cell culture, with an EC_{50} < 100 nM [47].

**Figure 5:** Chemical structures of HIV-1 RT inhibitors discovered or optimised using structure-based methods.
Along with fragment-based screening methods, structure-based design and optimisation techniques have played an important role in the discovery of novel non-nucleoside inhibitors of the HCV NS5b polymerase. One of the first of these studies was reported in 2007, when the structure of the non-nucleoside inhibitor 10a, a hit identified by a biochemical HTS assay, was resolved in complex with the enzyme, revealing that it binds to an allosteric region between the thumb and palm domains [48]. Starting from the examination of this crystal structure, a series of rational modifications on the hit scaffold, directed to a better occupation of the allosteric pocket, resulted in the discovery of potent inhibitor 10b (Figure 6), with IC\textsubscript{50} < 17 nM and activities against the viral replication in cellular systems in the low micromolar range [48].

Other recent examples of structure-based lead optimisation include the identification of a novel quinazolinone chemotype as thumb pocket 2 allosteric inhibitor (compound 11a in Figure 6), which has been rationally designed following inspection of the enzyme crystal structures in complex with previously reported allosteric inhibitors [49]. After the crystal structure of the newly designed quinazolinone 11a in complex with the enzyme was resolved, further structure-based optimisation attempts aiming to improve the key interactions within the protein binding channel were carried out, resulting in the identification of derivative 11b, which shows improved potency in both the biochemical and the cellular antiviral assays (Figure 6) [49]. Finally, starting from the crystal structure of non-nucleoside inhibitor 12a in complex with the polymerase, synthetic efforts directed towards the optimisation of the interactions with different sub-pockets of the enzyme resulted in the identification of several new analogues with potent antiviral activities \textit{in vitro} against both genotype 1a and 1b, high metabolic stability and good oral bioavailability (represented by 12b in Figure 6) [50].

Another interesting case of application of structural biology to the identification of HCV NS5B non-nucleoside inhibitors has been reported in 2010, when the structure of compound 13a, an attractive hit deriving from a biochemical HTS assay of a large compound collection, was sequentially optimised using a combination of X-ray crystallography, NMR analyses and molecular modelling studies, along with binding-site resistance mutant experiments and photoaffinity labelling studies [51]. This approach resulted in the identification of different series of new analogues with significantly improved potency in both the enzymatic assay and a cellular antiviral assay (exemplified by compound 13b Figure 6).
As mentioned above, there are many cases in which structural biology has been key to the identification of novel and improved antivirals, involving many of the viruses for which structural information has become available in the last decades, and including examples of antiviral drugs on the market, such as the neuraminidase inhibitor Zanamivir for influenza A and B [8]. Another example of successful application of a structure-based approach has been recently reported for the identification of a new class of influenza endonuclease inhibitors (Figure 7) [52]. An engineered high-resolution crystal form of pandemic 2009 influenza polymerase acidic protein N-terminal endonuclease domain was used for the crystallographic fragment screening of 775 fragments, leading to the identification of hit fragment 14a, which showed a binding affinity to the enzyme in the range of 1000 µM and also revealed the presence of a third metal ion in the active site cleft, previously unknown. Different cycles of rational modifications on the hit scaffold were performed in order to maximise the interactions with the active site of the enzyme. This structure-based optimisation approach resulted in the identification of compound 14b, which shows an antiviral EC\textsubscript{50} of 11 µM (Figure 7).

Structural biology has been fundamental also to understand the mechanism of action and the importance of conformational flexibility of the Dengue Virus polymerase [53-55], and structure-based methods have been extremely useful for the discovery of non-nucleoside inhibitors of this enzyme. An X-ray-based screen of the Novartis fragment collection against DENV-3 polymerase resulted in the identification of
biphenyl acetic acid 15a (Figure 7) bound to a novel pocket in the palm subdomain of the protein [56]. Growing and optimisation of this first hit through crystallography and computer-aided structure-based design led to the development of 15b, an antiviral agent active against all four Dengue serotypes with EC\textsubscript{50} values in the low micromolar range.

![Chemical structures of influenza endonuclease inhibitors 14a-b and Dengue polymerase inhibitors 15a-b.](image)

Figure 7: Chemical structures of influenza endonuclease inhibitors 14a-b and Dengue polymerase inhibitors 15a-b.

Among other examples, structure-based methods have been particularly useful for the discovery of novel chemical agents against enterovirus 71 and norovirus. In 2014, the analysis of the crystal structures of the EV71 complete viral particle in complex with uncoating inhibitors (exemplified by derivative 16a in Figure 8), in combination with in silico docking-based methods, guided the design of an improved and potent inhibitor of the virus-induced cytopathic effect in cells (compound 16b), which inhibits the full range of EV71 subtypes [57]. More recently, the analysis of the crystal structure of the EV71 3C proteinase in complex with moderate inhibitor 17a, and the subsequent structure-based modification of its scaffold, guided the design of 17b, an improved inhibitor in both the enzymatic and a virus cell-based assay [58]. Finally, a structure-based fragment-wise design of new inhibitors of the same enzyme led to the discovery of the new chemical scaffold 18 as a potent inhibitor of both the enzyme activity and the virus-induced cytopathic effect in cells [59].

In the case of norovirus, an iterative process of structure-guided design and optimisation on dipeptidyl inhibitors of the viral 3C-like protease has guided the discovery of potent derivative 19, which displays in vivo efficacy in a murine model of norovirus infection [60]. For the study of the same viral target, different crystal structures in complex with peptidyl inhibitors have recently been used to design novel triazole-based macrocyclic inhibitors (represented by compound 20 in Figure 8), which also inhibit the viral replication in cells with EC\textsubscript{50} values in the low micromolar range [61]. Finally, structural biology has been essential also for the identification of novel inhibitors of norovirus RNA-dependent RNA-polymerase. A docking-based in silico search method on the polymerase structure using commercially available compounds led to the discovery of suramin and its analogues [62] and of PPNDS [63, 64] as novel inhibitors of the norovirus RdRp activity. The resolution of the crystal structure of PPNDS in complex with the enzyme has also indicated the possibility to target a new binding sub-pocket in the thumb domain of the enzyme for the design of new norovirus inhibitors [64].
4. Conclusions
Following the increasing number of viral target structures deposited in the PDB, structural biology has become a fundamental tool in antiviral research, not only providing essential insights into the mechanisms of actions of viral enzymes and their interactions with substrates and antiviral agents, but representing in several cases the very basis for the rational discovery and optimisation of new antivirals, as exemplified by the striking case of drugs in the market such as Saquinavir and Zanamivir. Used in combination with computational techniques, structure-based methods are often among the earliest and most important steps of drug-discovery campaigns of most pharmaceutical companies, and their fundamental application for the discovery of new antivirals can only be expected to increase over the next years, supported by the continuous evolution of associated technologies.

5. Highlights
Resolution of three-dimensional viral structures in the PDB statistically analysed.
Evolution of structure-based methods in antiviral drug-discovery overviewed.
Different examples of the role of structural biology in the discovery and optimisation of new antivirals examined.

Figure 8: Chemical structures of enterovirus 71 and norovirus inhibitors discovered with the aid of structural methods.
6. Author information

*Corresponding Author: Dr. Andrea Brancale,
School of Pharmacy & Pharmaceutical Sciences, Cardiff University, King Edward VII Avenue, Cardiff, CF10 3NB, UK.
Phone: +44 (0)29 2087 4485;
E-mail: brancalea@cardiff.ac.uk

Authors declare no competing financial interest.

8. References


This paper represents an excellent example of the use of structural biology in the discovery of new antivirals against HIV.


45. Antonysamy SS, Aubol B, Blaney J, Browner MF, Giannetti AM, Harris SF, Hébert N, Hendle J,


This paper represents a significant example of the use of structural biology for understanding the mechanism of action of viral enzymes and the importance of their conformational flexibility, and for discovering new inhibitors.

This paper represents an excellent example of the importance of structural biology for the discovery of novel antivirals.

This paper represents a significant example of the importance of structural biology for the rational optimisation of antiviral compounds.

This paper represents a significant example of the importance of structural biology for the rational optimisation of antiviral compounds.


