

RATIONAL DESIGN OF ARTIFICIAL ENZYMES

A thesis submitted for the degree of Doctor of Philosophy by

MAURIZIO MURONI

Supervisor: R. K. Allemann

School of Chemistry

Cardiff University

April 2011

UMI Number: U585432

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U585432

Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

DECLARATION

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed *Manish Kumar* (candidate) Date *01-04-2011*.....

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of PhD

Signed *Manish Kumar* (candidate) Date *01-04-2011*.....

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references.

Signed *Manish Kumar* (candidate) Date *01-04-2011*.....

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed *Manish Kumar* (candidate) Date *01-04-2011*.....

ABSTRACT

Despite the endeavours of many decades, the design of artificial enzymes remains challenging. The work presented here investigates two known molecules as scaffolds for the design of artificial enzymes; an 18 amino acids α helical peptide with two disulfide bridges – ‘Apoxaldie’ able to catalyse the decarboxylation of oxaloacetate, and the 86-amino acid colicin E9 immunity protein (Im9) with four α helices.

Apoxaldie was modified such that the active site lysines were substituted by 2,4-diaminobutyric acid in order to increase the proximity of the enzyme active site to the chiral environment of the α helix. The designed peptide (‘Apoxaldie-Dab’) was synthesized with two different strategies and the correct formation of disulfide bonds was achieved. However, Apoxaldie-Dab did not show the expected activity for the decarboxylation of oxaloacetate. Circular dichroism studies showed a 30% loss of α helicity upon introduction of 2,4-diaminobutyric acid into Apoxaldie which can explain the decrease in activity.

In the case of Im9, two series of mutants, constructed around histidine 10 and asparagine 78 respectively, were designed to introduce histidine-based active sites into hydrophobic clefts. Site directed mutagenesis, gene expression and variant protein purification were carried out for ten variant mutants together with the wild type. The secondary structure and thermal stability of each protein were studied and catalytic activities were examined by monitoring the hydrolysis of *p*-nitrophenol acetate via ultraviolet-visible spectroscopy. The Im9 variant ‘Im9-W74A/N78H’ demonstrated three times more activity compared to Im9, indicating the modification of IM9 to possess a histidine based active site increased activity as hypothesized through its rational design.

The initial work of applying directed evolution on 'Im9-W74A/N78H' was accomplished by constructing a phage display library. A transition state analogue was synthesised to test screening the expressed library. This method can be further developed to assist the design of Im9-based artificial enzymes.

ACKNOWLEDGMENTS

I would like to express my gratitude to my supervisor Professor Rudolf Allemann for giving me the opportunity to undertake this project and for his patience and guidance throughout my research. My special thanks to Dr Mahmoud Akhtar and Dr Joel Loveridge for proofreading this work.

I would like to thank Professor Gerald Richter and Dr James Redman for their support and guidance during *viva voce* examinations and Dr Anne-Sophie Rouziere for the help with the phage display. A special thanks to Dr Yuehua Cong for guiding me at the beginning of this adventure.

I would like to thank my parents, my brother and sister for their constant love and support. Thanks to my friends in Cardiff and Sassari: Vero, Niek, Non, Anabel, Paula, Tony, Soraya, Dirk, Hock, Juan, Claudio, Simone, Margherita, Luca.

TABLE OF CONTENTS

ABSTRACT	i
ACKNOWLEDGMENTS	iii
LIST OF ABBREVIATIONS	xi
1. INTRODUCTION	1
1.1 OVERVIEW	2
1.1.1 Structure of Enzymes	2
1.1.2 Cofactors	5
1.1.3 Henri-Michaelis-Menten Model	7
1.1.4 Enzyme Specificity	9
1.1.5 Rate Enhancement	11
1.2 ARTIFICIAL ENZYMES	16
1.2.1 Macrocycles as Scaffolds	17
1.2.2 Peptides as Scaffolds	19
1.2.3 Rational (Re)design of Existing Proteins	27
1.2.4 Directed Evolution	27
1.2.4.1 Selection and Screening Techniques	30
1.2.5 Exploiting Enzyme Promiscuity	31
1.2.5.1 Creating a New Function	34
1.2.6 Computational Methods	35
1.2.7 Catalytic Antibodies	41
1.3 SUMMARY	43
1.4 AIMS	44
2. MATERIALS AND METHODS	46
2.1 CHEMICALS	47

2.2	MICROBIAL MEDIA	47
2.2.1	Luria-Bertani (LB) medium	47
2.2.2	2 x YT medium	47
2.2.3	SOC medium	47
2.2.4	SB medium	48
2.2.5	LB agar	48
2.2.6	LB top agar	48
2.2.7	2 x YT agar	48
2.3	STERILE SOLUTIONS	49
2.3.1	Antibiotic stock solutions	49
2.3.2	Tris-Buffered saline (TBS)	49
2.3.3	1% (w/v) BSA in TBS	49
2.3.4	0.5% (v/v) Tween 20 in TBS	49
2.4	SOLUTIONS FOR COMPETENT CELL PREPARATION	50
2.4.1	Preparation of competent cells (with CaCl ₂)	50
2.4.2	Preparation of super competent cells (with RbCl)	50
2.4.3	Preparation of electro-competent cells	50
2.5	NON-STERILE SOLUTIONS	51
2.5.1	dNTPs	51
2.5.2	Ethidium bromide	51
2.5.3	DNA loading dye for agarose gels (10x)	51
2.5.4	50x TAE buffer stock for agarose gels	51
2.5.5	SDS stacking buffer	52
2.5.6	SDS resolving buffer	52
2.5.7	10% (w/v) ammonium persulfate	52
2.5.8	Electrode running buffer (10x) for SDS-PAGE	52
2.5.9	Protein loading dye (4x) for SDS-PAGE	52

2.5.10 Protein purification buffer A	53
2.5.11 Protein purification buffer B	53
2.5.12 Phosphate buffer	53
2.5.13 Protein dialysis buffer	53
2.6 DNA PURIFICATION	53
2.6.1 DNA miniprep buffers	54
2.6.2 Agarose gel DNA extraction	54
2.6.3 DNA from PCR purification miniprep	55
2.6.4 Nucleotide removal miniprep	55
2.7 <i>E. COLI</i> STRAINS	55
2.7.1 Cloning strains	55
2.7.2 Expression strains	56
2.7.3 Preparation of ultra-competent cells	56
2.7.4 Preparation of competent cells for protein expression	57
2.7.5 Preparation of electrocompetent cells	57
2.7.6 Calculation of the transformation efficiency	58
2.7.7 Transformation of cells with plasmids	58
2.7.8 Transformation of electrocompetent cells with DNA phagemid	58
2.8 MANIPULATION OF DNA	59
2.8.1 Determination DNA concentration	59
2.8.2 Site directed mutagenesis	59
2.8.3 Polymerase chain reaction (PCR)	62
2.8.4 Digestion of DNA with restriction endonucleases	63
2.8.5 DNA visualisation and purification	64
2.8.6 Sub-cloning of Im9-W74A/N78H(NcoI-SacI) gene in pHEN2 phagemid	64

2.8.7	Library ligation	65
2.8.8	Plasmid purification (QIAprep Spin Miniprep Kit)	65
2.8.9	DNA precipitation	66
2.8.10	DNA sequencing	66
2.9	PHAGE DISPLAY	66
2.9.1	Preparation of helper phage VCSM13	66
2.9.2	Library transformation	67
2.9.3	Panning	68
2.10	PROTEIN EXPRESSION	69
2.10.1	Test expression	69
2.10.2	Large scale expression	70
2.11	PROTEIN PURIFICATION	70
2.11.1	DEAE anion exchange chromatography	70
2.11.2	SourceQ™ anion exchange chromatography	71
2.11.3	Ni affinity chromatography	71
2.11.4	Amicon ultrafiltration	71
2.12	PROTEIN CHARACTERISATION	72
2.12.1	SDS polyacrylamide gel electrophoresis	72
2.12.2	MALDI-TOF	72
2.12.3	Determination of the concentration of Im9	73
2.12.4	Kinetics parameters for Im9	73
2.13	PEPTIDE SYNTHESSES	73
2.13.1	Standard synthesis	73
2.13.2	Oxidation and purification of apoxaldie	74
2.13.3	Orthogonal preparation of Apoxaldie-Dab	75
2.14	APOXALDIE AND APOXALDIE-DAB CHARACTERISATION	76
2.14.1	Determination of the concentration of apoxaldie-Dab	76

2.14.2	Determination of the concentration of apoxaldie	76
2.14.3	Kinetics parameters for Apoxaldie and Apoxaldie-Dab	77
2.15	CD SPECTROSCOPY	77
2.16	SYNTHESIS OF TRANSITION STATE ANALOGUE	78
3.	APOXALDIE-DAB	81
3.1	INTRODUCTION	82
3.1.1	Aims	86
3.2	RESULTS AND DISCUSSION	87
3.2.1	Design of a new apamin-based miniature enzyme	87
3.2.2	Synthesis and purification of Apoxaldie-Dab	88
3.2.3	Oxidation of Apoxaldie-Dab	90
3.2.4	Synthesis and purification of Apoxaldie-Dab with orthogonal protection	94
3.2.5	Synthesis and purification of Apoxaldie	97
3.2.6	Oxidation of Apoxaldie	97
3.2.7	Circular dichroism spectroscopy	98
3.2.8	Kinetics study	105
3.3	CONCLUSION	107
4.	IM9	109
4.1	INTRODUCTION	110
4.1.1	Artificial enzymes	110
4.1.2	Im9	110
4.1.3	p-Nitrophenol hydrolysis as a model reaction	113
4.1.4	Aims	115
4.2	RESULTS AND DISCUSSION	116
4.2.1	Expression of Im9 wild type	116

4.2.2	Purification of Im9	117
4.2.3	Characterisation of Im9	119
4.2.4	Design of the active site	120
4.3	DESIGN 1: IM9-Y10H/F15A SERIES	121
4.3.1	Test expression of Im9 mutants	123
4.3.2	Purification of Im9-Y10H/F15A mutants	123
4.3.3	Characterisation of Im9-Y10H	124
4.3.4	Characterisation of Im9-F15A	125
4.3.5	Characterisation of Im9-Y10H/F15A	127
4.4	DESIGN 2: IM9-W74A/N78H SERIES	130
4.4.1	Im9-W74A	131
4.4.2	Characterisation of Im9-N78H	132
4.4.3	Characterisation of Im9-W74A/N78H	134
4.4.4	Characterisation of Im9-L52E/W74A/N78H and Im9-L52D/W74A/N78H	135
4.4.5	Characterisation of Im9-D51A/L52D/W74A/N78H and Im9-D51A/L52E/W74A/N78H	138
4.5	KINETIC DATA	140
4.6	CONCLUSIONS	147
5.	DIRECTED EVOLUTION OF IM9-W74A/N78H	149
5.1	INTRODUCTION	150
5.1.1	Directed evolution	150
5.1.2	Phage display	150
5.1.3	Biopanning	153
5.1.4	Aim	158
5.2	RESULTS AND DISCUSSION	158
5.2.1	Design of the mutations	158

Table of contents

5.2.2	Sub-cloning Im9 into pHEN2	159
5.2.3	Random cassette mutagenesis	161
5.2.4	Library preparation	162
5.2.5	Panning	164
5.2.6	Test expression of Im9 mutants in pHEN2	169
5.3	CONCLUSIONS	171
6.	GENERAL CONCLUSION	173
6.1	APOXALDIE-DAB	174
6.2	IM9	175
6.3	PHAGE DISPLAY	177
7.	REFERENCES	179

LIST OF ABBREVIATIONS

A	adenosine
Abs	absorbance
AEE	L-Ala-D/L-Glu epimerase
AMP	ampicillin
Acm	acetamidomethyl
APS	ammonium persulfate
β -ME	β -mercaptoethanol
BES	<i>N,N</i> -bis(2-hydroxyethyl)-2-aminoethanesulfonic acid
BLAST	basic local alignment search tool
Boc	<i>tert</i> -butoxycarbonyl
BSA	bovine serum albumin
C	cytosine
c	concentration
CD	circular dichroism spectroscopy
cfu	colony forming unit
CHCA	α -cyano-4-hydroxycinnamic acid
CV	column volume
Dab	L-2,4-diaminobutyric acid
DEAE	diethylaminoethyl
dH ₂ O	deionised water
DHAP	dihydroxyacetone phosphate

List of abbreviations

DIEA	<i>N,N</i> -diisopropylethylamine
DIC	<i>N,N'</i> -diisopropylcarbodiimide
DMAP	4-dimethylamino pyridine
DMF	<i>N,N</i> -dimethylformamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dsDNA	double stranded DNA
DTNB	5,5'-dithio-bis-(2-nitrobenzoic acid)
DTT	dithiothreitol
E	enzyme
E _a	activation energy
EDTA	ethylenediaminetetraacetic acid
FASTA	FAST-all
FAST-N	fast nucleotide alignment
FAST-P	fast protein alignment
Fmoc	9-fluorenylmethyloxycarbonyl
FPLC	fast protein liquid chromatography
G	guanine
GAP	glyceraldehyde 3-phosphate
GSH	glutathione
HBTU	<i>O</i> -benzotriazol-1-yl- <i>N,N,N,N</i> -tetramethyluronium hexafluorophosphate

List of abbreviations

HOBT	hydroxybenzotriazole
IPTG	isopropyl- β -D-1-thiogalactopyranoside
KiPO ₄	potassium phosphate
l	pathlength
LB	Luria-Bertani growth media
l-LDH	L-lactate dehydrogenase
MALDI-TOF	matrix assisted laser desorption ionisation time of flight mass spectrometry
MOPS	3-(<i>N</i> -morpholino)propanesulfonic acid
MW	molecular weight
NAC	near attack conformation
NAD ⁺	oxidised nicotinamide adenine dinucleotide
NADH	reduced nicotinamide adenine dinucleotide
OAA	oxaloacetate
OD	optical density
OSBS	<i>O</i> -succinylbenzoate synthase
P	product
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PEG	polyethylene glycol
PMSF	phenylmethylsulfonylfluoride
PSI-BLAST	position specific iterative basic local alignment search tool

List of abbreviations

RBP	ribose binding protein
RNA	ribose nucleic acid
RP-HPLC	reverse phase high pressure/performance liquid chromatography
Rpm	revolutions per minute
S	substrate
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
ssDNA	single stranded DNA
T	thymine
TASP	template assembled synthetic proteins
TBS	tris buffered saline
tBu	tert-butyl
TCEP	tris(2-carboxyethyl)phosphine
TEMED	<i>N,N,N',N'</i> -tetramethylenediamine
TFA	trifluoroacetic acid
TFE	2,2,2-trifluoroethanol
THF	tetrahydrofuran
TIM	triose phosphate isomerase
TM	melting temperature
Tris	tris(hydroxymethyl)aminomethane
Trt	trityl
TS	transition state

List of abbreviations

U	uracil
UV/Vis	ultraviolet/visible spectroscopy
V_{\max}	maximal velocity
ΔG^\ddagger	free activation energy of the reaction
ΔH^\ddagger	activation enthalpy
ΔS^\ddagger	activation entropy
ϵ	extinction coefficient

Amino acid triple and single letter code

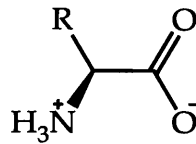
Amino acid	3 letter code	1 letter code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Iso	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

1. INTRODUCTION

1.1 Overview

1.1.1 Structure of Enzymes

Enzymes are proteins and therefore they are mostly made of α -L-amino acids. There are 20 natural α -amino acids that are universally found in proteins (Stryer 1988) and they share, at neutral pH, the general formula:



In relation to the nature of the R group, the amino acids can be classified in three large categories (Figure 1.1): polar-neutral, polar-charged and non-polar.

The amino acids are connected by amide (peptide) bonds and the order of the amino acids in the sequence of the protein defines its primary structure. Polypeptide chains can fold in organised structures *i.e.* secondary structure like α -helices (Figure 1.2), which are stabilised by intra-chain hydrogen bonds, and β -sheets (Figure 1.3), which are stabilised by hydrogen bonds between strands. The structured parts of the proteins are connected by structural elements like β -turns and loops.

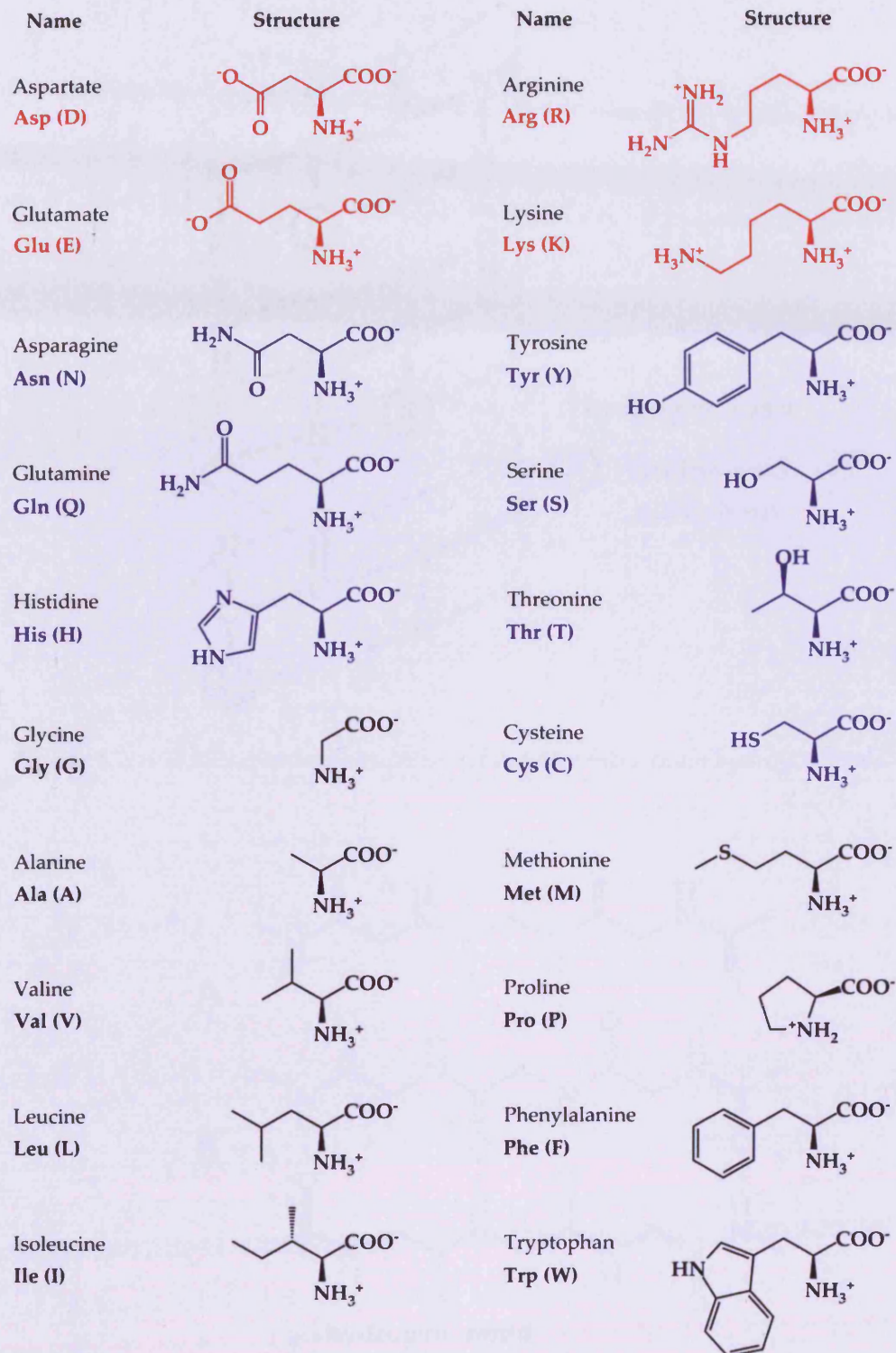


Figure 1.1: Amino Acids. Common amino acids can be classified into three large categories: non polar (in black); polar non-charged (in blue); and polar charged (in red).

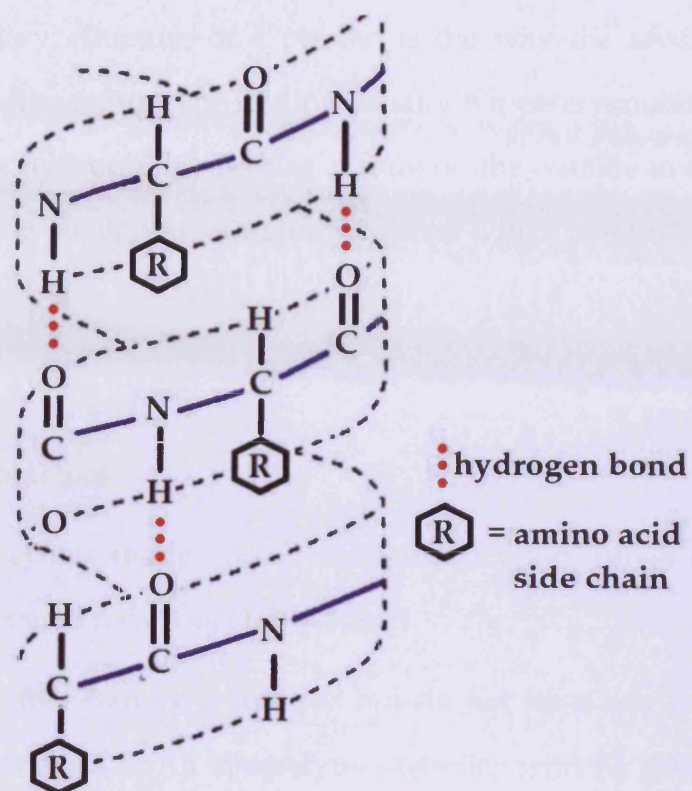


Figure 1.2: α -Helix secondary structure stabilised by intra-chain hydrogen bonds.

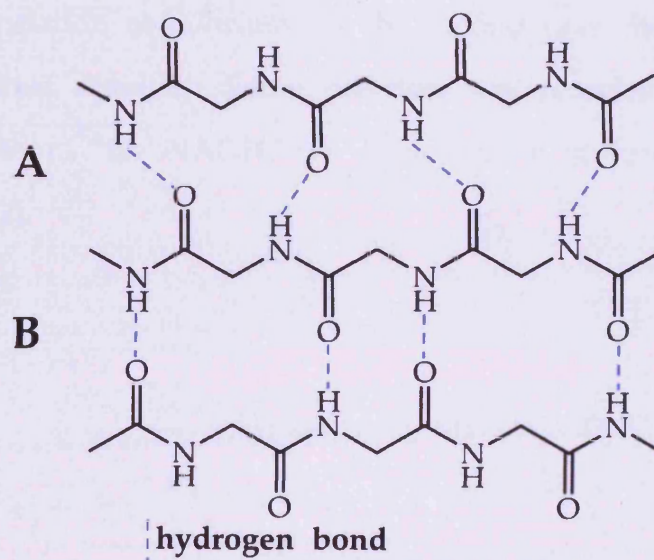


Figure 1.3: β -Sheet parallel (A) and antiparallel (B) secondary structure stabilised by hydrogen bonds between strands.

The tertiary structure of a protein is the way the secondary structure elements interact together. The folding usually happens around a hydrophobic core leaving the hydrophilic residues mainly on the outside in contact with the solvent, with the exception of some proteins which interact with biological membranes.

1.1.2 Cofactors

Many enzymes require small molecules to perform as catalysts. These molecules are termed cofactors (Table 1.2).

Enzymes that require a cofactor but do not have one bound are called *apoenzymes* or *apoproteins*. An apoenzyme together with its cofactor(s) is called the *holoenzyme*, which is the active form.

Cofactors can bind to the enzyme covalently or in association/dissociation equilibrium. In the second case their concentrations influence enzymes activities. Some cofactors are recycled in the catalytic reactions but others, like NADH, are consumed in stoichiometric amounts (Drauz *et al.* 2002).

Table 1.1: List of cofactors and their functions.

Compound	Function
Nicotinamide adenine dinucleotide (NADH)	Oxidation-reduction
Nicotinamide adenine dinucleotide phosphate (NADPH)	Oxidation-reduction
Flavin adenine dinucleotide (FAD)	Oxidation-reduction
Flavin mononucleotide	Oxidation-reduction
Haem	Electron transfer
CoenzymeA	Acyl group transfer
Adenosine triphosphate (ATP)	Phosphate/pyrophosphate transfer, adenylation
Pyridoxal phosphate	Transamination, amino acid decarboxylation, deamination, glycogenolysis, condensation reaction in heme synthesis
Thiamine pyrophosphate	Decarboxylation, transfer of C ₂ units
Biotin	Transfer of CO ₂
Tetrahydrofolic acid	Transfer of C ₁ groups
S-Adenosyl methionine	Methylation
Adenosyl-cobalamin	Isomerisation (hydrogen-shift)
Methyl-cobalamin	Methylation

1.1.3 Henri-Michaelis-Menten Model

Victor Henri published the first successful mathematical model for describing enzyme kinetics in 1903 (Henri 1903). In 1913 Leonor Michaelis and Maud Leonora Menten expanded Henri's earlier work and rederived the enzyme rate equation that today bears their names (Michaelis *et al.* 1913).

The Michaelis-Menten mechanism is based on the following features that characterise many enzymatic reactions:

- a) The initial rate of product formation is proportional, for a given initial concentration of the substrate ($[S]_0$), to the total concentration of the enzyme ($[E]_0$).
- b) For a given $[E]_0$ and for low $[S]$, the rate of product formation is proportional to $[S]_0$.
- c) For a given $[E]_0$ and high values of $[S]$ the rate of product formation becomes independent of $[S]$ reaching a maximum velocity v_{\max} .

In the mechanism the enzyme (E) forms a complex with the substrate (S). The complex ES can either release the unchanged S or the modified product P.



Michaelis and Menten assumed that ES is in thermodynamic equilibrium with E and S, an assumption that is not always true. Haldane and Briggs modified the mechanism introducing the steady state approximation (Briggs *et al.* 1925).

The rate of product formation is

$$v = k_b[ES] \quad 1.3$$

In the steady-state

$$\frac{d[ES]}{dt} = k_a[E][S] - k'_a[ES] - k_b[ES] = 0 \quad 1.4$$

It follows that

$$[ES] = \left(\frac{k_a}{k'_a + k_b} \right) [E][S] \quad 1.5$$

where [E] and [S] are the concentrations of unbound enzyme and substrate, respectively. Defining the Michaelis constant as

$$K_M = \frac{k'_a + k_b}{k_a} = \frac{[E][S]}{[ES]} \quad 1.6$$

K_M has the units of molar concentration. We can write

$$[E]_0 = [E] + [ES] \quad 1.7$$

Because the concentration of the substrate is usually in large excess compared to the enzyme we can write

$$[S] \approx [S]_0 \quad 1.8$$

and therefore

$$[ES] = \frac{[E]_0}{1 + K_M/[S]_0} \quad 1.9$$

Substituting this in equation 1.3 we obtain

$$v = \frac{k_b[E]_0}{1 + K_M/[S]_0} \quad 1.10$$

When

$$[S]_0 \ll K_M \quad 1.11$$

The rate is proportional to $[S]_0$

$$v = \frac{k_a}{K_M} [S]_0 [E]_0 \quad 1.12$$

When

$$[S]_0 \gg K_M \quad 1.13$$

the rate is independent of $[S]_0$ and reaches its maximum

$$v = v_{\max} = k_b [E]_0 \quad 1.14$$

Substituting the definition of v_{\max} in equation 1.10 we have

$$v = \frac{v_{\max}}{1 + K_M/[S]_0} \quad 1.15$$

k_b (k_{cat}), often called *turnover number*, is a first order rate constant related to the properties and reactions of the enzyme-substrate, enzyme-intermediate, and enzyme-product complexes. K_M is an apparent dissociation constant and can be seen as the dissociation constant of all complexes that the enzyme form with the different compound and it is equal to the substrate concentration at which $v = v_{\max}/2$ (Fersht 1984). The ratio k_{cat}/K_M can be used to compare the relative rates of an enzyme acting on different competitive substrates.

1.1.4 Enzyme Specificity

One of the outstanding characteristics of enzymes is the specificity they exhibit for the reactions they catalyse. There are different levels of specificities (DeMan 1999):

- a) Bond specificity: the enzyme acts on substrates that have similar structure and contain the same type of bond (*e.g.* amylase which catalyses the cleavage of the α -1,4-glycosidic bonds in starch, dextrin and glycogen).

- b) Group specificity: the enzyme acts on substrates that have the same type of bond and the same substructure surrounding it (*e.g.* pepsin which catalyses the hydrolysis of the central peptide bonds in which the amino group belongs to aromatic amino acids)
- c) Absolute specificity: the enzyme catalyses only one reaction (*e.g.* carbonic anhydrase which catalyses the reversible reaction between carbon dioxide hydration and bicarbonate dehydration).
- d) Stereochemical specificity: the enzyme acts only on a particular steric or optical isomer (*e.g.* D-amino acid oxidase which catalyses the oxidation of D-amino acids to the corresponding imino acids).
- e) Dual specificity: there are two types of dual specificity. In the first the enzyme acts on two substrates performing one kind of reaction (*e.g.* xanthine oxidase catalyse the oxidation of hypoxanthine to xanthine and the oxidation of this to uric acid). In the second type the enzyme performs two different reactions on the substrate (*e.g.* isocitrate dehydrogenase catalyses the conversion of isocitrate to α -ketoglutarate performing an oxidation followed by a decarboxylation (Stoddard *et al.* 1993)).

The specificity is connected with the geometry of the active site, which is designed to accommodate the substrate (Fischer's "lock and key", and Koshland's induced fit model) or, as it is now generally accepted, the transition state. Geometric complementarities, however, are not enough. Electrostatic, hydrophobic and hydrogen-bonding interactions are essential to the docking of the substrate in the active site (Fersht 1984).

1.1.5 Rate Enhancement

Enzymes are catalysts, so they speed up the rate of the reaction in which they are involved without undergoing any overall change and without modifying the equilibrium of the reaction. In the presence of enzymes, reactions that would otherwise take millions of years happen in a fraction of a second (Figure 1.4).

The “lock and key” model was based on the hypothesis that the active site of the enzyme is shaped in a way to accommodate exactly the substrate in its ground level. Pauling in his transition state theory suggested that enzymes are complementary to the structure of the activated complexes (transition state) of the reaction they catalyse and that the attraction of the enzyme for the transition state of the reaction would stabilise its energy, leading to a decrease in the activation energy of the reaction and an increase in the rate of the reaction (Pauling 1948) (Figure 1.5).

The higher affinity of inhibitors designed to mimic the transition state for the enzyme compared to the substrate supports Pauling’s theory, and can be used to gain useful indications of the particular mechanism by which a substrate is transformed by the enzyme that it inhibits. The structure of an effective inhibitor should reflect and confirm the mechanism on which its design was based.

Rate constant	Reaction	Half-time
$k = 10^{-15} \text{ sec}^{-1}$	glycine decarboxylation	1.1 billion years
	α -O-glycoside hydrolysis	12 million years
	phosphodiester hydrolysis	130,000 years
$k = 10^{-12} \text{ sec}^{-1}$	amino acid racemisation	6,000 years
	peptide hydrolysis	450 years
$k = 10^{-9} \text{ sec}^{-1}$	ribose phosphodiester hydrolysis	4 years
$k = 10^{-6} \text{ sec}^{-1}$	triosephosphate isomerization	2 days
$k = 10^{-3} \text{ sec}^{-1}$	peptide <i>cis-trans</i> isomerization	23 seconds
$k = 1 \text{ sec}^{-1}$	typical enzyme substrate complex formation	10 milliseconds
$k = 1 \text{ msec}^{-1}$		

Figure 1.4: Half-times of biological reactions proceeding spontaneously at 25 °C in neutral solution in the absence of a catalyst (Snider *et al.* 2000).

In certain cases it may be difficult to isolate reaction intermediates, and inhibitors can sometimes provide mechanistic information that is not easily accessible by other methods (Wolfenden 1976). In 1978 Schowen wrote: "...the entire and sole source of catalytic power is the stabilisation of the transition state; reactant-state interactions are by nature inhibitory and only waste catalytic power" (Gandour *et al.* 1978). Page, however, in 1987 listed 21 published theories of enzyme catalysis (Page *et al.* 1987) and Menger challenged

the transition state theory with the “split site” model in which the active site is subdivided into a region of binding and a region of reaction (Menger 1992). In 1995 Murphy replied to Menger’s observations showing again the centrality of the transition state stabilisation in the overall process of the catalytic activity of enzymes (Murphy 1995).

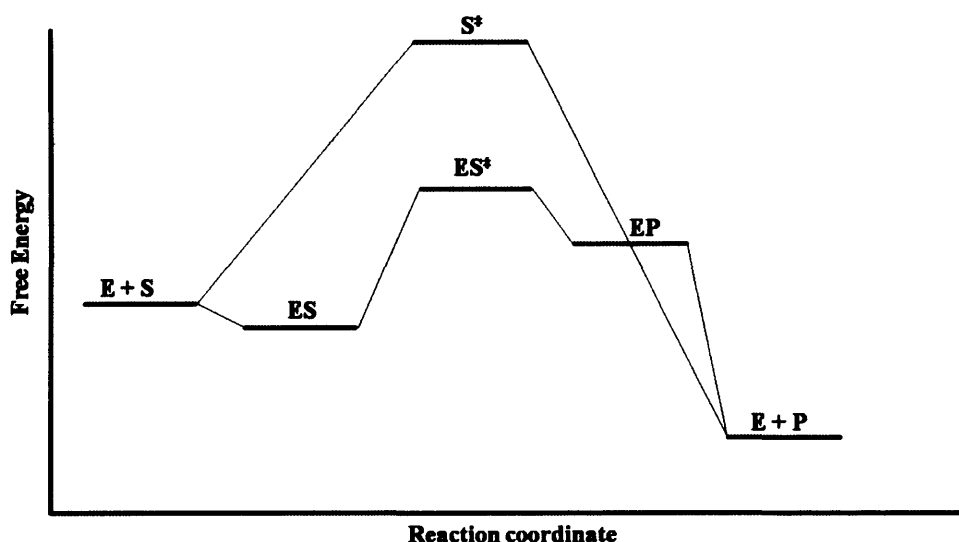


Figure 1.5: Energy level diagram of an enzyme-catalysed reaction and the corresponding non catalysed chemical reaction where E represent the free enzyme; S represent the free substrate; S^\ddagger represent the free transition state; ES represent the enzyme-substrate complex; ES^\ddagger represent the enzyme-transition state complex; EP represent the enzyme-product complex; P represent the free product states. Enzymes increase the rates of chemical reactions by stabilising the transition state of the reaction, hence lowering the activation energy barrier to product formation.

Bruice introduced the term near attack conformation (NAC) to define the required conformation for juxtaposed reactants to enter a transition state (TS) (Lightstone *et al.* 1996). The greater the mole fraction of reactant conformations that are present as NACs, the greater the rate constant. Rate constants for bond making and breaking in enzymatic reactions depend on, (i) the fraction of ES present as NACs, (ii) the change in solvation of reactant species within the

NAC, as compared to water, and (iii) electrostatic forces which can stabilize the TS (Bruice *et al.* 1998).

Lightstone *et al.* analysed a series of intramolecular reactions of dicarboxylic acid monoesters which yield five- and six-membered cyclic anhydride, as models for enzymatic reaction with the NAC method. They found that the rate constants for the reactions of the studied esters were directly dependent upon the mole fraction of the ground state of each ester present as NACs. When the ground state consists of only NACs, the rate enhancement is 10^8 (Lightstone *et al.* 1996). In a following report Lightstone *et al.* studied the transition state structures in anhydride formation from three esters (monoglutarate, succinate, and 3,6-endoxo- Δ^4 -tetrahydrophthalate esters) using *ab initio* calculation. The transition state structures for the three reactions were essentially identical. Applying the criteria of superimposable transition states and the lack of evidence for frequency changes in the transition states in the direction of increasing rate Lightstone *et al.* concluded that the structures of the transition states do not contribute to the different rate constants for the three esters analysed. Therefore any difference in activation energy from NAC to transition state must reside in small differences in the NAC structures for the three esters, a feature of the ground state (Lightstone *et al.* 1997).

Bruice suggested that a rational stance for the enzymologist is to consider driving forces for enzymatic reactions to arise from both ground state and transition state features. The free energy of reaction going from NAC to transition state can then be dominated by either entropy or enthalpy, depending upon the nature of the reaction. Electrostatic forces and the heterogeneous environment of the active site may provide up to another 10^8 fold rate enhancement due to stabilization of the transition state (Bruice *et al.* 1998).

The stabilisation of the transition state, and consequent rate enhancement, is accomplished by enzymes in different ways. These can be grouped into 5 categories (Copeland 2000):

- 1) Approximation (*i.e.* proximity) of reactants, which results in an increase of the effective concentration of the reactants.
- 2) Covalent catalysis (*i.e.* electrophilic and nucleophilic catalysis), with the formation of covalently bonded species, followed by the release of the product, the latter usually being the rate limiting step in the process.
- 3) General or specific acid-base catalysis, in which the catalytic residues stabilise the transition state *via* proton transfer.
- 4) Conformational distortion. In which the enzymes change conformation of the active site to force the substrate into a structure that resembles the transition state.
- 5) Preorganisation of the active site for transition state complementarity, in which the active site favours the formation of the transition state by removing the destabilising influence of the solvent.

The contribution of each mechanism to the overall stabilisation of the transition state is an intrinsic characteristic of each single enzyme.

Since any stabilisation of the transition state alone will lower the activation energy and any stabilisation of the ground state alone will increase

the activation energy (Fersht 1984), binding and stabilisation of the transition state have been the preferred targets in the design of artificial enzymes.

1.2 Artificial Enzymes

The crystallisation of urease by Sumner proved enzymes are proteins and gave the possibility to study the mechanism in details and to build models to mimic their actions (Sumner 1926).

Enzymes, with their high specificity and activity, can be regarded as perfect catalysts and it would be desirable to have an enzyme for those reactions not accessible to normal procedures, as in the functionalization of unactivated carbons remote from functional groups (Breslow 2005). Unfortunately, although several thousand enzymes have been characterised, this is still not realised. The design of new enzymes with the required characteristic for each different case is one of the ultimate goals of protein engineering (Ulmer 1983). The difficulty in the *ex novo* design of enzymes lays in the impossibility to predict the folding of protein from its primary structure, although great progress have been made in the prediction of small peptides folding (Venkatraman *et al.* 2001). To overcome above difficulty, several approaches have been used, like the use of pre-existing scaffolds (cyclodextrin and small peptides) or the (re)design of new active site onto existing enzymes (Trainor *et al.* 1981; Breslow *et al.* 1988; Ye *et al.* 1992; Johnsson *et al.* 1993; Broo *et al.* 1997; Broo *et al.* 1998; Nilsson *et al.* 2000; Taylor *et al.* 2001; Drauz *et al.* 2002; Taylor *et al.* 2002; Del Valle 2004; Weston *et al.* 2004).

1.2.1 Macrocycles as Scaffolds

Ronald Breslow coined the term Biomimetic Chemistry to describe chemistry which tries to imitate biological processes (Breslow 1972). Since 1956 Breslow used cyclodextrins as scaffolds on which to graft catalytic groups to imitate the general principles of enzymatic catalysis, with particular attention to the geometry of the enzyme-substrate complexes.

Cyclodextrins are cyclic oligosaccharides, composed of 5 or more α -D-glucopyranoside units linked α -1,4 as in amylose (Del Valle 2004). The cyclodextrin cavity mimics the active site of an enzyme, allowing substrate discrimination. Breslow attached a thiazolium group on β - and γ -cyclodextrin in order to catalyse the condensation of two benzaldehyde molecules to form benzoin. The β -cyclodextrin was too small to bind two benzaldehyde molecules and did not catalyse the reaction, but the γ -cyclodextrin increased the rate of the reaction 150-fold compared to the reaction without the cyclodextrin ring, and the benzoin was expelled out of the cyclodextrin cavity because of its larger size (Breslow *et al.* 1988). Breslow synthesised different bis-imidazole β -cyclodextrins (the difference was in the distance between imidazole rings, *i.e.* how many glucose units between them) and tested their catalytic activities in the hydrolysis of the cyclic phosphate of 4-t-butylcatechol. The three isomers with zero, one and two glucose residues between the imidazole rings, catalysed the reaction with one imidazole acting as base and the other protonated one as acid. The isomer with the higher activity was the one with the two imidazole rings right next to each other with the ImH^+ protonating the phosphate oxyanion to facilitate formation of a phosphorane intermediate (Figure 1.6) (Breslow *et al.* 1996a; Breslow *et al.* 1996b; Breslow *et al.* 1996c).

Cyclodextrins have been used to synthesise flavocyclodextrins as models for flavoenzymes (*i.e.* enzymes that bind the cofactor riboflavin and catalyse

several reactions like oxidation, hydroxylation, *etc.*). The model with flavin bound to the secondary side (the larger opening of the toroid like structure) of a β -cyclodextrin was able to catalyse the oxidation of 1-(1-naphthyl)methyldihydronicotinamide with saturation kinetics like a real enzyme (Ye *et al.* 1992).

Other macrocycles have been used to mimic the active site cavity of an enzyme. Cyclophanes (hydrocarbons consisting of an aromatic unit and an aliphatic chain that forms a bridge between two non-adjacent positions of the aromatic ring) were linked covalently to catalyse the benzoin condensation in different solvents. In methanol the equilibrium of the reaction was favourable to the formation of benzoin (Jimenez *et al.* 1989).

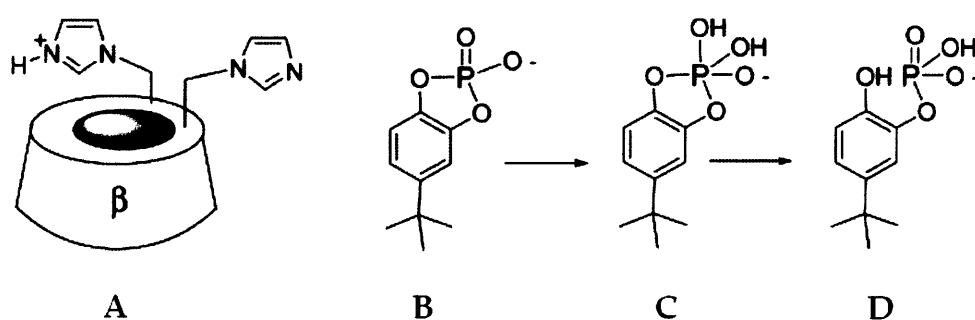


Figure 1.6: The bis-imidazole cyclodextrin (A) with the imidazole rings right next to each other catalysed the hydrolysis of cyclic phosphate of 4-t-butylcatechol (B) via a phosphorane intermediate (C) (Breslow 2005).

An octopus cyclophane was used to incorporate non-covalently bound pyridoxal-5'-phosphate into its cavity to construct an artificial holoenzyme (Figure 1.7). Different alkylamines were linked to the "tentacles" as substrates. The Schiff base formation was measured as a function of the different hydrophobicity of the alkylamines (Murakami *et al.* 1989).

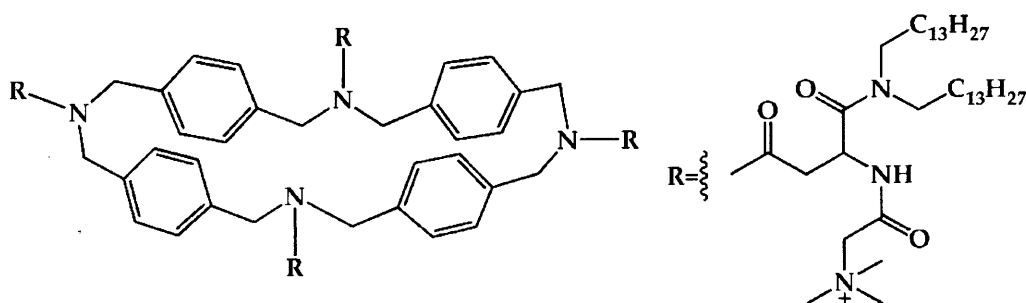


Figure 1.7: An octopus cyclophane made introducing multiple amphipathic components into a rigid macrocyclic skeleton. The R groups form the “tentacles” of the octopus and the alkylamine acting as substrate are bounded to them (Murakami *et al.* 1989).

1.2.2 Peptides as Scaffolds

The use of polypeptides in the design of new enzymes allows a greater synthetic flexibility than the methods previously discussed, especially when used in conjunction with genetic methods like mutagenesis. Changes to the gene sequence encoding a protein will result in modifications to that protein. The combination of the 20 proteinogenic amino acids gives the possibility to create a virtually infinite number of molecules (the number of possible combination for a simple 10 amino acid peptide is 1.024×10^{13}), although not all combinations are useful in the design of new enzymes. The polypeptide should fold in a specific way in order to create a catalytic arrangement of the side chains and this is the greatest difficulty in the *de novo* design of enzymes. The preference of each natural L-amino acid for the different secondary structure conformations (α -helix, β -sheet and random coil) has been calculated using known X-ray structure of proteins (Chou *et al.* 1974; Levitt 1978) statistical analysis (Munoz *et al.* 1994) thermodynamic analysis (O'Neil *et al.* 1990).

The α -helical fold has been extensively used in *de novo* design of folded peptides. Generally amino acid sequences that assume an α -helix conformation in a folded protein are not helical when excised from the protein. However the

use of appropriate strategies like incorporation of helix stabilising residues (Eisenberg *et al.* 1986; Betz *et al.* 1995; Monera *et al.* 1996; Doig 2008; Song *et al.* 2008; Huang *et al.*), the creation of salt bridges between residues separated by one α -helical turn (Goodman *et al.* 1989; Errington *et al.* 2005; Dzubiella 2008; Suveges *et al.* 2009; Walker *et al.* 2009; Sommese *et al.* 2010) allow the design of peptides that form α -helices.

Monomeric α -helices can interact to form coiled coil, a structure firstly described by Crick (Crick 1953) consisting of two to five α -helices wrapped around each other. Coiled coils can be left-handed or right handed. In the most common left handed coiled coil the number of residues per turn of the helix decrease from the typical 3.6 to 3.5. Each helix has a periodicity of seven (a heptad repeat) with a minimum of two repeats (Burkhard *et al.* 2000). This repeat is usually denoted **(a-b-c-d-e-f-g)_n** in one helix, and **(a'-b'-c'-d'-e'-f'-g')_n** in the other. In this model, **a** and **d** are typically nonpolar core residues found at the interface of the two helices, whereas **e** and **g** are solvent exposed, polar residues that give specificity between the two helices through electrostatic interactions (Mason *et al.* 2004). In the right handed coiled coil, the number of residues per turn of the helix increase to 3.67. The periodicity of the helix is eleven residues every three turn of the helix (undecatad repeat). The repeated unit is usually denoted with the letter from **a** to **k** with the residues. There are very few examples of right handed coiled coil in nature, one is tetrabrachion from *Staphylothermus marinus*, which forms a parallel tetramer (Stetefeld *et al.* 2000). Similar to coiled coil are α -helical bundles which most commonly contain four helices and can be grouped into two major types – those in which the helix crossing angles are all the same and near 20°, and those in which larger helix crossing angles occur (Kohn *et al.* 1998). One of the major differences between helical bundles and coiled coils is that a coiled coil contains

a narrow hydrophobic face (hydrophobic residues primarily at the **a** and **d** heptad positions), while bundles tend to contain a wider hydrophobic face, owing to a greater incidence of hydrophobicity at the **e** and **g** positions (Epanand 1993). α -Helical bundles have been used to mimic metallonucleases (Rossi *et al.* 2004), quinoproteins (redox components and enzymes in biological systems) (Li *et al.* 2006), polypeptides for control of nanoparticle assembly (Aili *et al.* 2007), protein models of radical enzymes (Dai *et al.* 2002), substrate-accessible carboxylate-bridged binuclear metal centre (Di Costanzo *et al.* 2001) and peptides with oxaloacetate activity (Johnsson *et al.* 1993; Taylor *et al.* 2001; Taylor *et al.* 2002) etc.

The use of β -sheet, the other common secondary structure configuration has been limited by experimental difficulties. The residues involved in the formation of β -sheets are at variable and often distant positions in the sequences. Also, the exposed amides at the edge of β -sheets can hydrogen-bond to other sheets, leading to insoluble aggregates (Bryson *et al.* 1995). Metal ion and disulfide bonds have been used to stabilise the β -sheet (Pessi *et al.* 1993; Quinn *et al.* 1994; Yan *et al.* 1994) and advances in computational methods allowed the design of small β -sheet motif (Kraemer-Pecore *et al.* 2003).

Gutte *et al.* synthesised one of the first artificial enzyme a 34 residue polypeptide, following the Chou-Fasman method (Chou *et al.* 1974), to bind the nucleotide sequence GAA (the anticodon of yeast tRNA^{Phe}) (Figure 1.8).

The polypeptide had tendency to polymerise, due to the presence of two cysteine residues, and the monomer and dimer were purified. The dimer showed a better ribonuclease activity compared to the monomer (Gutte *et al.* 1979).

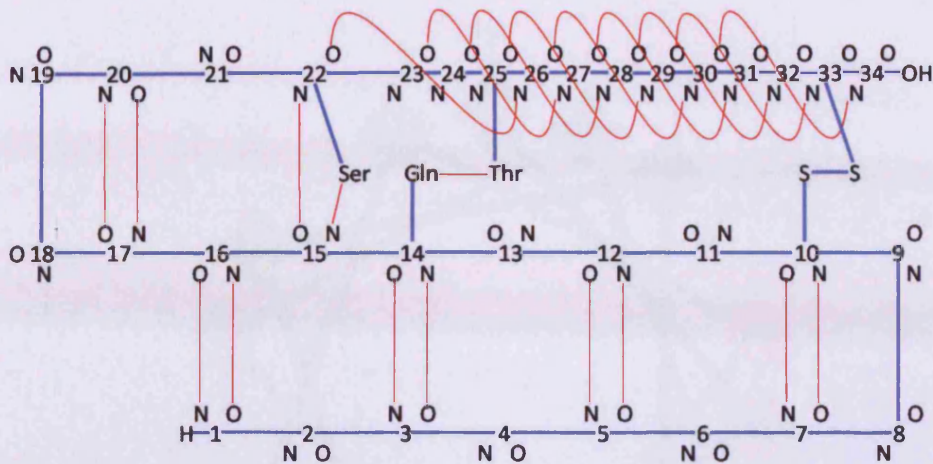


Figure 1.8: Proposed secondary structure of an artificial 34-residue DNA binding polypeptide. Hydrogen bonds are represented as thin red lines. In the proposed model the α -helix axis (from residues 23 to residues 33) lies above the hydrophobic side of the β -sheet (from residues 1 to residues 16). The amino acid symbols in boxes represent side chains involved in hydrogen bonds between different secondary structures (Gutte *et al.* 1979).

The first *de novo* rational designed functional peptides were the 14-residue amphipathic α -helices oxaldie 1 and oxaldie 2 (Figure 1.9) (Johnsson *et al.* 1993). The helicity of the two peptides increased with concentration. Oxaldie 1 and oxaldie 2 aggregate in solution to form helical bundles. The designed active site was based on amino groups – either the amino terminal in oxaldie 1 (the pK_a of which was depressed to 7.2 by its interaction with the helix dipole) or a lysine side chain in oxaldie 2 (pK_a depressed to 8.9 by electrostatic interactions with other lysine side chains). The peptides catalysed the decarboxylation of β -ketoacids through the formation of a Schiff base between the substrate and the active amino group (Johnsson *et al.* 1993).

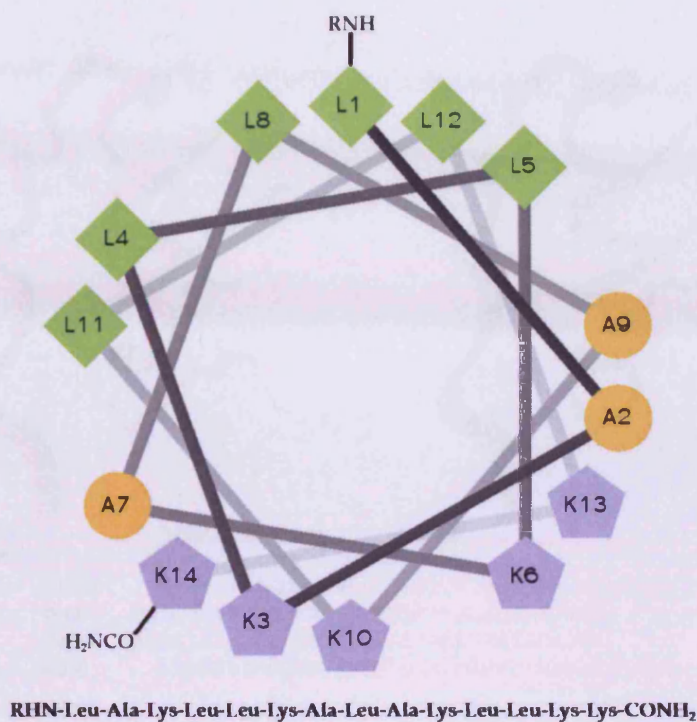


Figure 1.9: Sequence of oxaldie 1 and 2 (*R* is hydrogen for oxaldie 1 and acetyl for oxaldie 2). In the helix wheel representation the hydrophobic alanine residues are in orange circles, the hydrophobic leucine residues are in green squares and the positively charged lysine residues, acting as catalytic residues, in azure pentagons.

Based on oxaldie 2, Taylor *et al.* designed oxaldie 3 and oxaldie 4 (Taylor *et al.* 2001; Taylor *et al.* 2002). In this case the scaffolds were based on pancreatic polypeptides (avian for oxaldie 3 and bovine for oxaldie 4), which are 36 amino acid polypeptides with a proline rich helix (residues 1-8) and an α -helix (residues 14-31). Those peptides form dimers in solution (Figure 1.10) (Blundell *et al.* 1981). Lysine residues (4 in oxaldie 3 and 3 in oxaldie 4) were grafted on the α -helix portion of the peptides without disturbing the hydrophobic core or the dimer interface. Both peptides showed a higher activity compared to oxaldie 2 with K_M and k_{cat} independent from the concentration of the catalyst between 2 and 200 μ M (Taylor *et al.* 2001; Taylor *et al.* 2002).

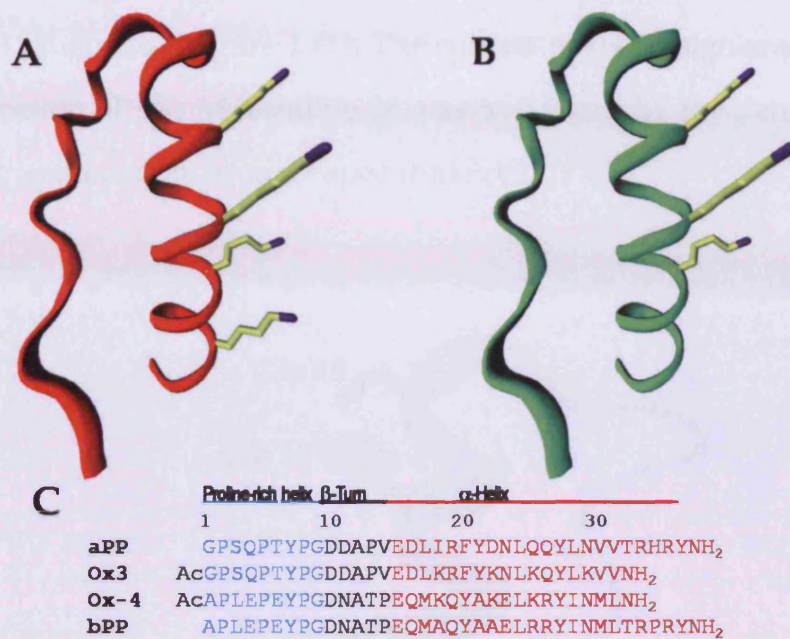


Figure 1.10: Structures of oxaldie-3 with four lysines (A) and oxaldie 4 with three lysines (B) acting as the active site, and sequence alignment of avian pancreatic polypeptides, bovine pancreatic polypeptides, Oxaldie-3 and Oxaldie-4 (C). The model is based on the X-ray structure of avian pancreatic polypeptides (Blundell et al. 1981).

Weston *et al.* used the 18 residues Apamin to design Apoxaldie, a monomeric miniature enzyme with high resistance to thermal and chemical denaturation. Apamin is a neurotoxin present in the bee venom. Two disulfide bonds stabilise a C-terminal α -helix on which solvent exposed face an active site based on lysine residues was grafted. The success in Apoxaldie design was confirmed by the lack of activity of reduced Apoxaldie in which the α -helix is not present. Apoxaldie inherited the thermal and chemical denaturation resistance of Apamin with a catalytic activity comparable to the best peptide-based oxaloacetate decarboxylase (Weston *et al.* 2004).

Nicoll *et al.* used bovine pancreatic peptide (the scaffold of oxaldie-4) to design Art-Est a miniature enzyme with *p*-nitrophenyl esterase activity. The catalytic moiety was made of histidine grafted on the solvent exposed face of

the C-terminal α -helix. The histidine side chains reactivity was modulated by arginine and glutamate (Figure 1.11). The success of the design was confirmed by the depression of pK_a of histidine 18 was to 5.5 and by the detection of the acyl-enzyme intermediate by mass spectrometry.

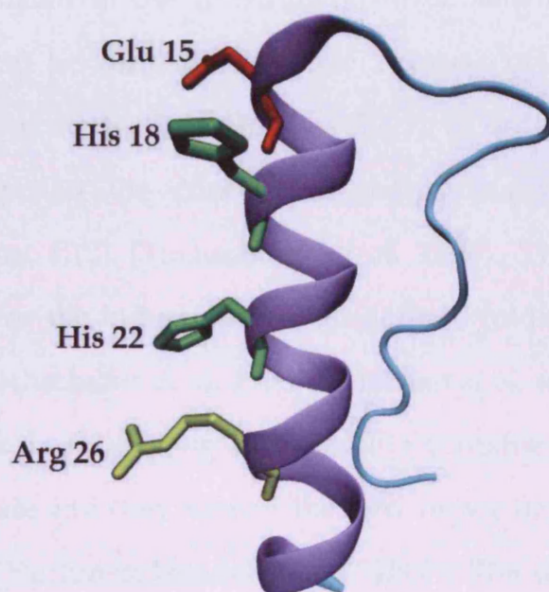


Figure 1.11: Structure of Art-Est. The active histidine 18 and histidine 22 (in red) were grafted on the solvent exposed surface of the α -helix (blue) of bovine pancreatic peptide. The pK_a of the active histidine was modulated by glutamate 15 and arginine 26. Model based on NMR study (Nicoll *et al.* 2004).

In 1994 Atassi *et al.* claimed the successful reproduction of the active site of trypsin and chymotrypsin in two cyclic peptides of 29 residues: TrPepz and ChPepz. The peptides, cyclised by disulfide bond, were constructed linking the residues present in trypsin and chymotrypsin active sites using glycine as spacer. TrPepz was able to catalyse the hydrolysis of *N*-tosyl-L-arginine with K_M similar to trypsin and k_{cat} almost half of the parent enzyme. ChPepz on the other hand was able to catalyse the hydrolysis of the ester group of *N*-benzoyl-L-tyrosine ester with K_M and k_{cat} values comparable to chymotrypsin. (Atassi *et*

al. 1993). Wells *et al.* and Corey *et al.* however in two independent studies were unable to reproduce Atassi results for TrPepz and ChPepz (Corey *et al.* 1994; Wells *et al.* 1994). A molecular dynamics study of TrPepz showed that the cyclopeptide was quite flexible and therefore lacking the spatial organisation of the catalytic residues found in trypsin and chymotrypsin (Marrone *et al.* 1994).

The folding problem in the construction of *de novo* designed artificial enzymes has proven to be difficult to tackle because of the extraordinary complexity of the folding mechanism (Breslow 2005). In order to overcome this difficulty Mutter proposed the concept of template assembled synthetic proteins (TASP) (Figure 1.12) (Tuchscherer *et al.* 1996). This approach uses topological templates for the induction of well-defined folding topologies (*e.g.* four helix bundles) (Tuchscherer *et al.* 1996). Sherman *et al.* used this approach to construct a four helix bundle using a cavitand (a container shaped molecule (Cram 1983)) as template and they named the new molecules “caviteins” from *cavitand* and *proteins* (Huttunen-Hennelly *et al.* 2007). The caviteins displayed some native-like properties like tertiary structures and a hydrophobic core (Huttunen-Hennelly *et al.* 2007; Seo *et al.* 2007).

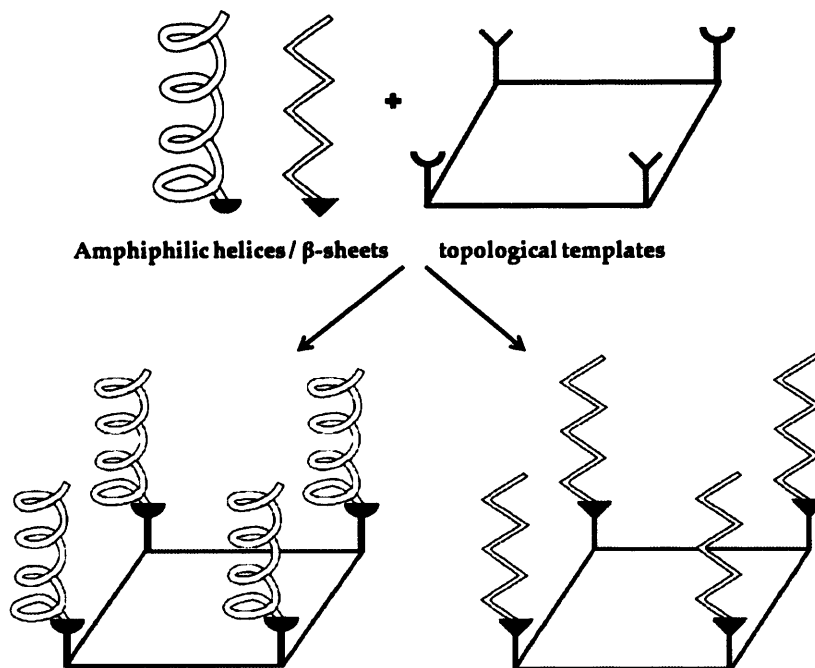


Figure 1.12: Topological templates have been as devices in order to induce and stabilise α -helix bundle or β -sheet TASP molecules to mimic some properties of natural proteins.

1.2.3 Rational (Re)design of Existing Proteins

Another approach to design artificial enzymes is to use existing proteins and modify them to gain new functions or improve the existing functions. This method can basically overcome the difficulty of predicting the tertiary structure of new designed proteins. The use of gene manipulation techniques and screening techniques make it possible to force a chosen protein to evolve in the desired direction.

1.2.4 Directed Evolution

Using methods mentioned above, few successes in rational design have been reported, although many have been attempted. When successful, the rate accelerations and catalytic efficiencies are almost always considerably lower than those seen for natural enzymes. This is because often the sites of mutation

which will improve or change the specificity of the enzyme are far away from the active site and their position is almost impossible to predict (Moore *et al.* 1996). The use of directed evolution can, however, help improve their efficiency. The procedure that “should allow a new type of evolutionary biomolecular engineering” (directed evolution) was suggested by Eigen and Gardiner (Eigen *et al.* 1984):

- 1) Produce a mutant spectrum of self-reproducing templates.
- 2) Separate and clone individual mutants.
- 3) Amplify clones.
- 4) Express clones.
- 5) Test for optimal phenotypes.
- 6) Identify optimal genotypes.
- 7) Return to 1 with a sample of optimal genotypes.

Directed evolution does not require any knowledge of the structure, function or even mechanism. Different strategies can be used in order to create the mutants. Random point mutagenesis of the whole gene is a common approach. The mutations are typically introduced by error prone PCR, mutator strains, or by treatment of the isolated DNA with UV light or chemical reagents. These can be simple non alkylating agents (like formaldehyde, hydroxylamine, methoxyamine, nitrous acid, bisulfite *etc.*), alkylating agents (like dialkyl sulfates, alkyl alkane sulfonates, dialkyl nitrosamine, *N*-nitrosoureas, *etc.*) or aromatic mutagens (like 2-acetylaminofluorene, *N*-hydroxy-1-naphthylamine, *N,N*-dimethyl-4-aminoazobenzene, benzo(α)pyrene, *etc.*) (Singer *et al.* 1982). An

appropriate error rate is essential for the success of this approach. Too low an error rate will produce not enough diversity and the screening is wasted by too many copies of the parent enzymes. If the error rate is too large, the number of positive mutants will be very low and the screening is wasted by the large number of inactive clones.

Error prone PCR uses non-optimal conditions for the reaction (*e.g.*, increasing the concentration of MgCl₂, adding MnCl₂ to the reaction mixture, increasing and unbalancing the concentrations of the four dNTPs, adding deoxyinosine triphosphate (dITP), increasing the concentration of *Taq* polymerase, or increasing the extension time and cycle numbers). The error rate of *Taq* DNA polymerase is, under normal reaction conditions, about 0.002% to 0.02% per nucleotide per replication, too low for efficient mutagenesis of small genes. With the modified reaction conditions the error rates can be as high as 2% per nucleotide position. Mutator strains introduce defined spectra of mutations at frequencies up to 100,000 times that in wild type bacteria due to defects in DNA replication and repair machinery. The main problem of the single point mutation approach lies in the structure of the genetic code which limits the accessible amino acid substitutions. Depending on the specific codons, only 24-40% of the possible amino acid changes are accessible by single base substitutions (Sirotkin 1986). Furthermore to obtain the best mutations (those that will change the physicochemical characteristic of the mutants) multiple substitution at a single codon are required (Miyazaki *et al.* 1999). Focused mutagenesis allows the library size to be reduced (a quintuple mutation gives a theoretical library size of only ~10⁶ mutants) but it is possible that important mutation sites are left out from the screening. Several kits are available to perform saturation mutagenesis on a single amino acid, such as the Altered Site®II (Promega, Madison, WI, USA), and QuikChange™ (Stratagene,

La Jolla, CA, USA) site-directed mutagenesis systems. When more amino acids are to be targeted and if they are all in one region of the gene, the usual approach is to use oligonucleotide-cassette mutagenesis, where a piece of the gene between two restriction sites is excised and a synthetic oligonucleotide with the random mutation is ligated in its place. If the sites chosen to be randomised are too far apart, other approaches like recursive PCR are used (Prodromou *et al.* 1992).

DNA shuffling is PCR without synthetic primers. In the process, a collection of mutants of the same gene, or related families of genes are first chopped up with enzymes. The DNA fragments then are separate into single-stranded templates by thermal denaturation. When the temperature is lowered, fragments that share complementary DNA regions will bind to each other and non-complementary regions will hang over the ends of the templates. PCR is then used to build new double-helical DNA with the complementary regions as primers creating a mixed structure or chimera. In the final step, PCR reassembles these chimeras into full-length, shuffled genes (Cohen 2001).

Once the appropriate random mutation technique is chosen a screening technique powerful enough to cover all the possible mutants is needed. The screening can be carried out *in vivo* (when the properties of the enzyme are related to the fitness of the cell) or *in vitro*.

1.2.4.1 Selection and Screening Techniques

Selection methods are designed to increase the percent of positive variation removing the non-positive mutants or allowing only those with the wanted characteristics to grow. Screening refers to a qualitative or quantitative assay of each single clone or few pooled clones of a mutant library.

The importance of the choice of an appropriate selection/screening technique cannot be overestimated. The conditions of the screening should be as similar as possible to those of the reaction, in which the enzyme is going to be used (same pH, temperature, solvent, *etc.*). Modified substrates are often used to simplify the analysis of the reaction with production of fluorescent or coloured compounds. In these cases however, the optimised enzyme may not be active when the 'real' substrate is used. This problem can be solved by performing a second screening against the real substrate or in conditions more similar to the real one (Moore *et al.* 1996). Another problem can arise when a stress is used in *in vivo* selection. The stress that is supposed to select for the desired characteristic can trigger an unpredictable response like activation of new genes (Patten *et al.* 1996).

Selection/screening methods can be divided in low throughput and high throughput. The first usually have a low level of error in the measurement but positive mutants can be missed because they are not sampled. High throughput methods have a higher error level in the measurement which means that positive mutant can be discarded. Microplate based screening can use fluorescence (Joo *et al.* 1999; Hardiman *et al.* 2010) or colour changes (Fox *et al.* 2007) to detect the mutation in the activity.

1.2.5 Exploiting Enzyme Promiscuity

The specificity of enzymes has been thought of as the cornerstone of catalysis however actively searching for promiscuous activities usually reveals one (Nobeli *et al.* 2009). Different classifications of promiscuity have been proposed. Copley defined four types of catalytic promiscuity: first, enzymes use different substrates to perform one chemical reaction; second, enzymes produce

different products due to imperfect control of the reactants and use of different sites in catalysis; third, enzymes use the same residues in the active site to perform different overall reactions; fourth, enzymes catalyse different transformations using different mechanisms (Copley 2003). Hult and Berglund defined three major types of enzymatic promiscuity that can be combined: first, condition promiscuity (catalysis in a variety of temperatures, pH, etc.); second, substrate promiscuity; and third, catalytic promiscuity (different chemical transformations performed) (Hult *et al.* 2007). Bornscheuer and Kazlauskas classify catalytic promiscuity according to whether the reactions involve different functional groups, different mechanisms or, more commonly, both (Bornscheuer *et al.* 2004). Nobeli *et al.* classify catalytic promiscuity in levels (promiscuity at the individual gene or transcript level; promiscuity at the individual protein level; and family promiscuity); manifestation (multiple substrates or partners; multiple chemical reactions); condition that triggered the promiscuity (differential expression; environmental conditions; concentration of ligand); and different mechanisms (post-translational modifications; multiple domains; oligomeric state; etc.). Almost any mechanism is available to any level, can be triggered by any condition and can have any effect (Nobeli *et al.* 2009).

Babtie *et al.* analysed the possible causes of enzyme promiscuity. Active site flexibility allows enzymes to accommodate different substrates. Hydrophobic binding, driven by desolvation, do not require a specific arrangement of functional groups and may be less discriminatory than electrostatic interactions or hydrogen bonding. The intrinsic reactivity of the active site due to the presence of multiple functional groups may be sufficient to trigger promiscuity as well as the presence of new functional group derived

from post translational modification. The presence of cofactors increasing the reactive potential of the active site can as well facilitate promiscuity.

An example of a promiscuous enzyme is chymotrypsin, which catalyses the hydrolysis of many different types of compounds like amides, esters, thiol esters, acid chlorides and anhydrides. All the substrates are thought to react *via* tetrahedral transition states or intermediates. The first step of the reaction consists of a nucleophile attack by a serine at a carbonyl carbon. Chymotrypsin also catalyses attack at a tetrahedral phosphoryl group, a reaction that proceeds *via* a trigonal bipyramidal intermediate. This alternative reaction involves attack on a different atom, cleavage of a different type of bonds and results in covalent modification of the enzyme (Zhao *et al.* 1994).

The concept that high selectivity is necessary for high activity has been challenged by BcPMH an enzyme from *Burkholderia caryophilli* PG2952 originally assigned as a phosphonate monoester hydrolase (van Loo *et al.* 2010). BcPMH belongs to the alkaline phosphatase super-family which members catalyse the hydrolysis of phosphate monoester and diester and sulphate monoester (Galperin *et al.* 1998). BcPMH is able to hydrolyse with high efficiency (from 10^7 to 10^{12} second-order rate acceleration) phosphonate monoester and diester, phosphate monoesters, phosphate triester, sulfate monoesters and sulfonate monoesters. For all the reactions except for phosphate triester BcPMH displays saturation kinetics. BcPMH presents almost all the characteristic listed by Babbitt *et al.* that can trigger promiscuity. Its active site has a diameter approximately twice that of the largest substrate and is able to accommodate all the six substrates. A highly reactive metal atom (iron or zinc) is coordinated in the active site. The reactive potential of the active site is further enhanced by the presence of a formylglycine (derived from a

posttranslational modification of a cysteine) which is thought to be involved in the hydrolysis of five of the six substrates (van Loo *et al.* 2010).

A single mutation is sometimes enough to increase the catalytic efficiency of the side reaction. An example is the protease papain which exhibits promiscuity by catalysing nucleophilic attack on nitriles. Its nitrile hydratase activity has been greatly improved by addition of a general acid residue. The single point mutation Gln19Glu increases the rate of multiple turnovers by more than 10^4 fold. It has been suggested that the introduced residues act as general acids that facilitate the successive additions of water required to convert the nitrile to its carboxylic acid and ammonia products (Dufour *et al.* 1995).

γ -Humulene synthase, a highly promiscuous sesquiterpene synthase was converted, through a series of mutations in the active site, to seven specific and active synthases that use different reaction pathways to produce specific and very different products. Saturation mutagenesis was performed for each chosen residue and the mutants were analysed treating each mutated residue as an independent variable. The mutations that increased the production of one particular product were then combined together to construct the seven specific synthases (Yoshikuni *et al.* 2006).

Promiscuity can be used as 'starting points' for both *in vivo* and *in vitro* evolution of new functions. The enhancement and exploitation of catalytic promiscuity has emerged as an important strategy for developing novel biocatalysts (Gerlt *et al.* 2009).

1.2.5.1 Creating a New Function

A more challenging task is to create a completely new function. A single substitution permits an L-Ala-D/L-Glu epimerase (AEE) to catalyse the *O*-succinylbenzoate synthase (OSBS) reaction (Figure 1.13). The progenitor is not

promiscuous for the OSBS reaction. The natural enzymes that catalyse the AEE and OSBS reactions share potential Lys acid/base catalysts on opposite faces of the active site at the ends of the second and sixth β -strands in the $(\beta/\alpha)_8$ -barrel (modified TIM barrel) domain. The mutant AEE-G-297-D showed OSBS activity while retaining a reduced level of the progenitor's activity (Schmidt *et al.* 2003).

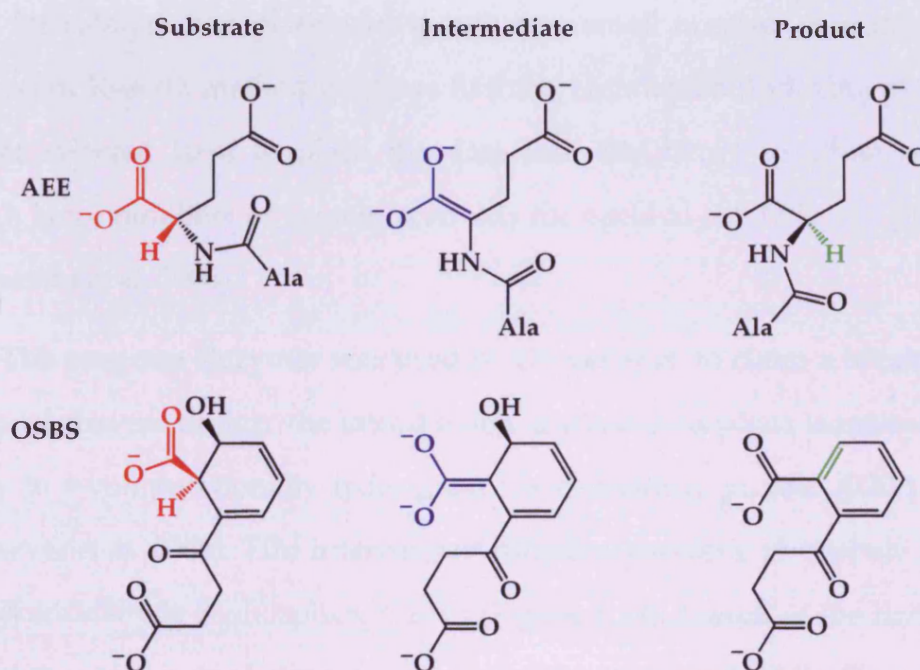


Figure 1.13: Reactions catalysed by the mutant AEE-G-297-D: 1,1-proton transfer (AEE) and β -elimination/dehydration (OSBS).

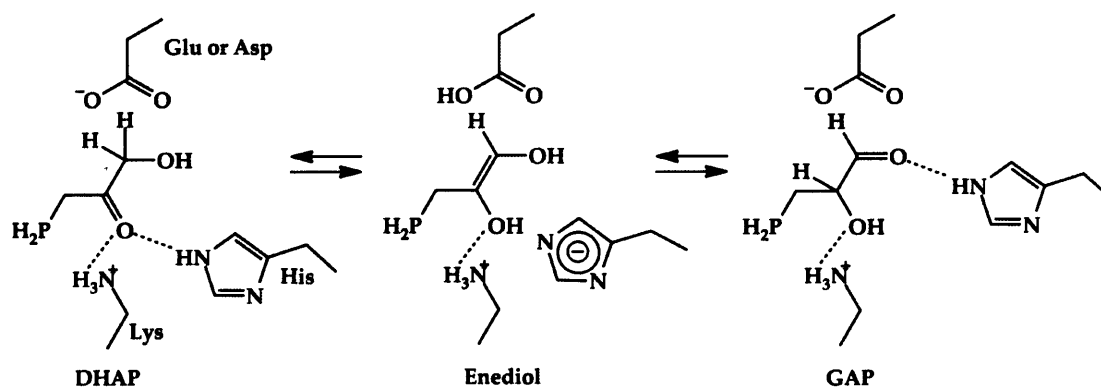
1.2.6 Computational Methods

The use of computational methods in *de novo* enzyme design is increasing in importance year by year with the increasing power of computers. Because of the complexity of the problem a simple energy minimisation of a single protein state seldom gives good results. Multi-objective searches are a better approach for designing specificity (to stabilise one or more states relative to others), improving binding affinity (to increase interaction while maintaining

folding stability), and designing *de novo* proteins (to avoid alternate structures and aggregation). Often enzyme design requires more detailed objectives than simply binding the transition state and coordination of key active-site functional groups (Lippow *et al.* 2007).

The first computational methods used in enzyme design, such as ORBIT (Dahiyki 1996) and Dezymer (Hellings *et al.* 1991), have primarily been used to search for catalytic site placement in one or a small number of scaffolds. The more recent Rosetta method employs hashing techniques (indexing algorithms that use selected keys to place the data into the array) to allow searching through large numbers of protein scaffolds for optimal catalytic site placement (Zanghellini *et al.* 2006).

The program Dezymer was used by Dwyer *et al.* to claim a breakthrough in rational enzyme design: the introduction of triose phosphate isomerase (TIM) activity in a computationally redesigned ribose-binding protein (RBP) from *E. coli* (Dwyer *et al.* 2004). TIM interconvert dihydroxyacetone phosphate (DHAP) and glyceraldehyde 3-phosphate (GAP) (Figure 1.14), therefore the first step in the design was the prediction of mutation that converted RBP into a receptor for DHAP using a combinatorial optimization algorithm (Looger *et al.* 2003).

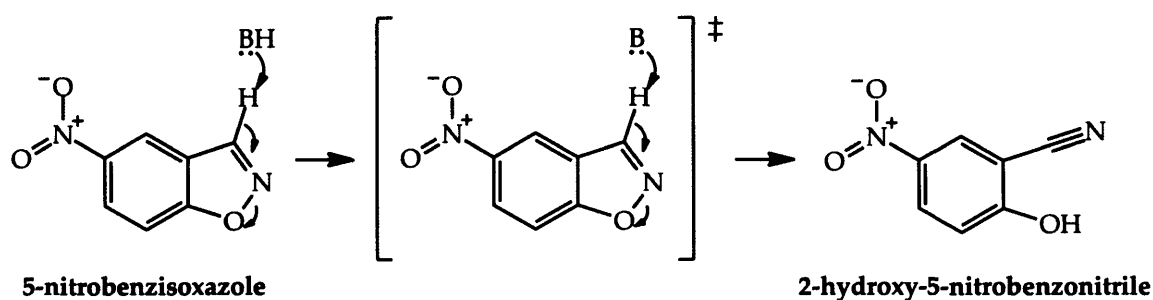


Scheme 1.1: TIM mechanism for the interconversion of DHAP and GAP via enediol

The design process to introduce catalytic activity was divided into three steps: a set of geometrical definition of key interactions contributing to catalysis was firstly generated; then a combinatorial search algorithm (Hellingsa *et al.* 1991) was used to identify positions where placement of catalytic residues and substrate simultaneously satisfies the defined geometrical constraints; finally a receptor design algorithm (Looger *et al.* 2003) was used to generate the remainder complementary surface around the placed substrate. One mutant (NovoTim1.0) out of the fourteen designed displayed TIM activity. A further series of mutation was designed to increase activity and thermal stability of NovoTim1.0 leading to NovoTim1.2. Complementation of a TIM-deficient strain by overexpressed NovoTims (1.0 and 1.2) was tested on gluconeogenic substrates (lactate and glycerol) in the presence and absence of the inducer isopropyl-D-thiogalactopyranoside (IPTG). NovoTims1.0 and 1.2 support IPTG-dependent growth on lactate, but not glycerol while NovoTim1.1 was not tested. The most active NovoTims showed a 10^5 - to 10^6 -fold increase in the Tim-catalysis over background reaction and an alanine-scanning mutagenesis indicates that all residues designed to be catalytically active contribute substantially to rate enhancement (Dwyer *et al.* 2004). However further analysis

on the NovoTims by J. Richard at State University of New York showed that the TIM activity registered by Dwyer *et al.* was due to a contamination of native TIM from the bacteria used to overexpress NovoTims leading to the retraction of the report (Dwyer *et al.* 2008). In the retraction Dwyer *et al.* acknowledge the finding of Richard and explained with the TIM contamination all the results obtained. J. F. Kirsch and J. Richard however pointed out in two letters in response to the retraction that several results published in the Dwyer report (lower K_m compared to the wild type TIM; decrease of activity after alanine mutation; the results of the *in vivo* experiment) could not be explained by simply contamination but only by really working NovoTims. The use of Dezymer to predict ligand binding has been questioned by Schreier *et al.* who tried to assemble and analyse five of the designed proteins that seemed to work best. Most of the analysed proteins were not suitable for structural studies due to instability and aggregation. However, they were able to solve the crystal structure of an arabinose binding protein designed to bind serotonin. The protein crystallised in presence of an excess of serotonin is in an open conformation with no serotonin bound, although the side-chain conformations in the empty binding pocket are very similar to the conformations predicted. During subsequent characterization using isothermal titration calorimetry, CD, and NMR spectroscopy, no indication of binding could be detected for any of the tested designed receptors, whereas wild-type proteins bound their ligands as expected. Schreier *et al.* concluded that although the computational prediction of side-chain conformations appears to be working, it does not necessarily confer binding as expected and that the computational design of ligand binding is not a solved problem and needs to be revisited (Schreier *et al.* 2009).

Recently the Rosetta algorithm has been used successfully in the design of two set of enzymes. In the first R othlisberger *et al.* designed enzymes that catalyse the Kemp elimination. The reaction, which needs a general base to remove a proton from a carbon, is known to be catalysed by bovine serum albumin and is used as model for proton transfer from carbon (Figure 1.15).



Scheme 1.2: Mechanism of Kemp elimination catalysed by a general base.

The first step was the design of two idealised active sites, which contained either an aspartate/glutamate or a histidine-aspartate couple as the catalytic base. Using quantum mechanical and classical methods functional groups were added to improve the transition state (TS) stabilisation. Next RosettaMatch (Zanghellini *et al.* 2006) was used to screen a large set of stable protein scaffolds with known X-ray structures to find backbone positions that could accommodate the designed active sites. RosettaDesign (Liu *et al.* 2006) was used to optimise the packing of residues on a given backbone by combining side-chain orientations for the various amino acids deposited in a rotamer library. RosettaMatch was used to check the quality of a design by means of a multi-parameter scoring on the basis of catalytic geometry and TS energy. On the basis of these scores, 59 designs were experimentally tested from the more than 10⁵ analysed. 8 of them, which contained between 10 and 20 residue exchanges, showed weak enzymatic activity with catalytic efficiencies

(k_{cat}/K_M) increased up to $2.5 \cdot 10^5$ fold. The X-ray structure of one of the variants was solved and the active site superimposes with a root mean square deviation of 0.95 Å with the calculated active site model. The back-mutation of the introduced catalytic bases in several designs abolished or strongly reduced catalytic activity suggesting that the new enzymes catalysed the Kemp elimination with the expected reaction mechanism. The catalytic activity was further improved by seven rounds of directed evolution and best variant showed a > 200-fold improvement in activity compared to the starting enzyme (Röthlisberger *et al.* 2008).

The second set of enzymes designed with the Rosetta method are able to catalyse a retro-aldol reaction, which requires the breaking of a carbon-carbon bond in a non-natural substrate (Jiang *et al.* 2008). A different method was used to address the higher complexity of the reaction with multiple intermediates and transition states. The various reaction intermediates and transition states were modelled in the context of a specific set of functional residues. The models were superimposed and four alternative active site motifs, all with a nucleophilic lysine and general acid/base for the proton transfer steps. Three dimensional models for each motif were generated varying the degrees of freedom of the composite TS, the orientation of catalytic site chains relative to the TS and the conformation of the side chains. RosettaMatch was then used to identify a suitable scaffold to host the candidate catalytic site and 72 candidates, based on 10 scaffolds, were experimentally tested. 32 designs showed a small catalytic activity. The X-ray structure of the best variant confirmed the active site design but showed a significant difference in the loops surrounding the active site. The designed enzymes displayed only modest catalytic activity, with one molecule of product generated in 2 hours. The low activity of the designed enzymes is principally due to the limited number of residues (maximum 4) that

can be designed simultaneously. In natural occurring enzymes group positioning and protonation states create a network of polar interactions and a more extended hydrogen-bond networks. Computational design of such extended polar networks is exceptionally challenging because of the difficulty of accurately computing the free energies of buried polar interactions (Jiang *et al.* 2008).

The reports show that it is possible to design enzymes using computational methods but the increase in the rates compared to the background reaction for the computational designed enzymes ($k_{\text{cat}}/K_{\text{uncut}} \approx 10^2$ - 10^4) are lower than the synzymes designed by Hollfelder *et al.* ($k_{\text{cat}}/K_{\text{uncut}} \approx 10^7$) (Hollfelder *et al.* 2001) and in the range of those of a set of catalytic antibody ($k_{\text{cat}}/K_{\text{uncut}} \approx 10^4$) (Thorn *et al.* 1995), and serum albumins ($k_{\text{cat}}/K_{\text{uncut}} \approx 10^3$ - 10^4) (Hollfelder *et al.* 1996). Although computational designs allow structure prediction close to atomic resolution they are not yet able to take account of all the interaction involved in a catalytic process. Several rounds of directed evolution for fine-tuning of structures and improvement of catalytic efficiency are still necessary (Damborsky *et al.* 2009).

1.2.7 Catalytic Antibodies

Pauling suggested that enzyme are designed to stabilise the transition state of the reaction and that the mechanism was similar to the binding of an antibody to its antigen (Pauling 1948). This suggestion led to the idea of using antibodies as catalysts. Antibodies (or immunoglobulins) are glycoproteins produced by vertebrates with the task to neutralise, upon binding, foreign bodies like virus, bacteria (the antigens) or simple molecules (haptens). Antibodies have specific and high affinity for the antigens that elicited their

synthesis, they recognise a specific group or cluster of amino acids on a large molecule called an antigenic determinant (epitope) (Stryer 1988). When a transition state analogue of a particular reaction is used as the hapten the resulting antibody is, in theory, able to catalyse that reaction and the hapten will be an inhibitor.

The technique was developed independently by Lerner and Schultz. Lerner used a tetrahedral phosphonate as the transition state analogue for the hydrolysis of carboxylic esters which proceed through a tetrahedral intermediate (Tramontano *et al.* 1986). Schultz also used a tetrahedral phosphonate as the transition state analogue but for the hydrolysis of carbonate esters (Pollack *et al.* 1986).

One of the most successful catalytic antibodies was developed by Barbas III in collaboration with Lerner (Wagner *et al.* 1989). The antibody 38C2 catalyses the aldol addition and the retro-aldol reaction and it is the first commercial available catalytic antibody. A β -diketone was used as hapten to elicit antibodies with a reactive lysine residue in the active site (Wagner *et al.* 1989). Another remarkable example is the antibody 1F7 which is able to catalyse the stereospecific Claisen rearrangement of (-)-chorismate into prephenate (Hilvert *et al.* 1988b). The abzyme 1F7 was raised against a transition state analogue inhibitor of natural occurring chorismate mutase and accelerates the mutase reaction by a factor of 2×10^2 over background (Hilvert *et al.* 1988a). When produced cytoplasmically at high concentrations in a chorismate mutase deficient yeast strain, the Fab fragment of this catalyst reconstitutes the shikimate biosynthetic pathway leading to the aromatic amino acids tyrosine and phenylalanine, conferring a substantial growth advantage under auxotrophic conditions (Tang *et al.* 1991).

Abzymes can be regarded as the most successful enzyme mimics so far described (Breslow 2005). However their activity is still far lower than the natural occurring enzymes. One of the reason of this lower activities can be traced in the affinity of antibody for their respective transition states, which is usually in the 10^{-9} M range while for enzymes can reach 10^{-24} M (Stewart *et al.* 1995). Another reason is the dynamic of the antibody/transition state interaction. Enzymes have evolved to recognise a series of structures that connect the substrate and product along the reaction coordinate; by contrast, catalytic antibodies are elicited to a single, specific structure. Consequently, the binding energy is applied by abzymes to preferential transition-state stabilisation whereas enzymes may also act to destabilise bound substrates (Stewart *et al.* 1995). The lower efficiency of catalytic antibody compared to natural enzymes may also be explained with the scaffold shared by all antibodies. Studies on catalytic antibodies show the recurrence of a basic hapten recognition motif (MacBeath *et al.* 1996; Charbonnier *et al.* 1997; Romesberg *et al.* 1998; Xu *et al.* 1999; Hilvert 2000). In favourable cases a nearly perfect shape complementarity of the binding pocket to its ligand can be achieved. In general, however, these frequently selected binding pockets may be poorly suited to particular catalytic tasks or may represent local minima from which it will be difficult to evolve further (Hilvert 2000).

1.3 Summary

Enzymes can be seen as the perfect catalysts with their specificity and huge rate enhancement. Artificial enzymes are designed to reproduce enzymes performance. The use of non-protein scaffolds on which to graft an active site with precise geometrically directed functionalisations can have a role in

chemical processes where harsh conditions are needed since they are usually more robust than proteins. *De novo* design of new enzymes is limited by the lack of understanding of the folding rules, therefore existing protein (often enzymes themselves) are generally used as starting points for the design of enzymes able to catalyse reactions not covered by the natural enzymes. Rational design and/or directed evolution are used for this purpose. Structural knowledge of the starting protein is a necessary prerequisite for rational design. When directed evolution is used, it is necessary to reduce the library size to match the performance of the screening method in use. Artificial enzymes design remains to be the fascinating yet extremely challenging field.

1.4 Aims

Different approaches have been used to create new enzymes. Rate accelerations comparable to catalytic antibodies have been achieved with miniature enzymes based on peptide scaffold. Their simple mechanism of action has allowed a detailed study of the reaction mechanism; however the exposure of the active site to the solvent limits their selectivity toward different substrates.

The overall aim of this project is to generate improved small artificial enzymes. Possibilities of achieving good catalytic activity and discrimination of the substrates, based on enzyme structure and on the geometry of the transition state will be examined.

Two different strategies are planned to achieve the design of artificial enzymes. In the first strategy the miniature oxaloacetate decarboxylase, Apoxaldie will be used as starting point. By shortening the side chains of the active lysine residues, the aim is to bring the active site closer to the peptide backbone,

in the hope that the chiral environment of the α -helix and the increased proximity of the amino group would help increase its selectivity and activity.

In the second strategy, the colicin DNase immunity protein Im9 would be used as a scaffold. Im9 has been extensively studied and it is easily produced. Its structure present several hydrophobic clefts, where active sites would be created using histidine. The designed active site will present a cavity, created by mutating a bulky aromatic amino acid into an alanine, on which the substrate would bind.

The rationally designed Im9-based artificial enzymes would then be submitted to directed evolution. A library of Im9 mutants would be displayed the on phage pIII protein and screened against a transition state analogue.

2. MATERIALS AND METHODS

2.1 Chemicals

All chemicals were purchased from Thermo Fisher Scientific or Sigma Aldrich unless otherwise stated. IPTG, ampicillin, carbenicillin and MOPS were purchased from Melford. All amino acids and peptide reagents were purchased from Novabiochem. Oligonucleotides were purchased from Operon Technologies (Germany).

2.2 Microbial Media

All media were prepared using deionised water, and were sterilised by autoclaving for 20 min at 121 °C prior to use (Sambrook *et al.* 2001).

2.2.1 Luria-Bertani (LB) medium

Bacto-Tryptone (1% w/v), NaCl (1% w/v) and yeast extract (0.5% w/v), pH adjusted to 7.0 with 5 M NaOH.

2.2.2 2 x YT medium

Bacto-Tryptone (1.6% w/v), yeast extract (1% w/v) and NaCl (0.5% w/v), pH adjusted to 7.0 with 5 M NaOH.

2.2.3 SOC medium

Bacto-Tryptone (2% w/v), Bacto Yeast Extract (0.5% w/v), 5 M NaCl (0.2% v/v), 1 M KCl (0.25% v/v) and 1 M MgCl₂ (1% v/v), pH adjusted to 7.0 with 5 M

NaOH. Prior to use 1 M MgSO₄ (1% v/v) and 1 M glucose (2% v/v), both filter-sterilised using a 0.2 µm syringe filter, were added to the autoclaved solution.

2.2.4 SB medium

MOPS (1% w/v), Bacto-Tryptone (3% w/v) and yeast extract (2% w/v), pH adjusted to 7.0 with 5 M NaOH.

2.2.5 LB agar

Bacto-Tryptone (1% w/v), NaCl (1% w/v), yeast extract (0.5% w/v) and agar (1.5% w/v), pH adjusted to 7.0 with 5 M NaOH. Following autoclaving, the solution was cooled to 37 °C, antibiotic was added where appropriate, and the agar poured to form plates.

2.2.6 LB top agar

Bacto-Tryptone (1% w/v), NaCl (1% w/v), yeast extract (0.5% w/v) and agar (0.75% w/v), pH adjusted to 7.0 with 5 M NaOH. Following autoclaving, the solution was cooled to 45 °C and poured on the top of LB agar plates.

2.2.7 2 x YT agar

Bacto-Tryptone (1.6% w/v), yeast extract (1% w/v), NaCl (0.5% w/v) and agar (1.5% w/v), pH adjusted to 7.0 with 5 M NaOH. Following autoclaving, the solution was cooled to 37 °C, antibiotic was added where appropriate, and the agar poured to form plates.

2.3 Sterile Solutions

All sterile solutions were prepared using deionised water and, unless stated otherwise, were filter-sterilised using a 0.2 µm syringe filter prior to storage.

2.3.1 Antibiotic stock solutions

Ampicillin (100 mg mL⁻¹), carbenicillin (100 mg mL⁻¹) and kanamycin (50 mg mL⁻¹) stocks were stored at -20 °C. Tetracycline (40 mg mL⁻¹) was sterilised as above, diluted with an equal volume of autoclaved, cooled glycerol, and stored at -20 °C.

2.3.2 Tris-Buffered saline (TBS)

Tris-HCl (0.5 mM) and NaCl (10 mM), pH adjusted to 7.0 with 5 M NaOH. Stored at room temperature.

2.3.3 1% (w/v) BSA in TBS

BSA (1% w/v) was dissolved in sterile TBS solution (Section 2.3.3). Stored at 5 °C.

2.3.4 0.5% (v/v) Tween 20 in TBS

Tween 20 (0.5% v/v) was dissolved in sterile TBS solution (Section 2.3.4). Stored at 5 °C.

2.4 Solutions For Competent Cell Preparation

All solutions were prepared using deionised water.

2.4.1 Preparation of competent cells (with CaCl₂)

CaCl₂ solution 1: CaCl₂ (0.1 M).

CaCl₂ solution 2: CaCl₂ (0.1 M) and glycerol (15% v/v).

Solutions were sterilised by autoclaving at 121 °C for 20 minutes and stored at 4 °C.

2.4.2 Preparation of super competent cells (with RbCl)

RbCl solution 1: RbCl (100 mM), CH₃COOK (30 mM), CaCl₂ (10 mM), MnCl₂ (50 mM) and glycerol (15% v/v), pH adjusted to 5.8 with dilute acetic acid.

RbCl solution 2: RbCl (10 mM), CaCl₂ (75 mM), 3-(*N*-morpholino)propanesulfonic acid (MOPS) (10 mM) and glycerol (15% v/v), pH adjusted to 7.0 with 1 M NaOH.

Solutions were filter-sterilised using a 0.2 µm syringe filter and stored at 5°C.

2.4.3 Preparation of electro-competent cells

10% (v/v) glycerol was sterilised by autoclaving for 20 minutes at 121 °C.

2.5 Non-Sterile Solutions

Except where otherwise stated, all non-sterile solutions were prepared using deionised water.

2.5.1 dNTPs

dNTP stock solutions (100 mM) were purchased from Bioline. These were diluted with ultra-pure water to a concentration of 10 mM each for site directed mutagenesis experiments or 25 mM each for PCR experiments, and stored at -20 °C.

2.5.2 Ethidium bromide

Ethidium bromide (25 mM) was stored in the dark at 5 °C. For gel staining, this stock was diluted to give a working concentration of 6 µM.

2.5.3 DNA loading dye for agarose gels (10x)

Bromophenol blue (2.5 mg mL⁻¹) and glycerol (30% w/v). Stored at room temperature. The dye was mixed with the DNA sample 1:10 immediately prior to use.

2.5.4 50x TAE buffer stock for agarose gels

EDTA (0.1 M), glacial acetic acid (5.7% v/v) and Tris base (2 M), pH adjusted to 8.5 with NaOH 5 M. Stored at room temperature. Diluted 1:50 immediately before use.

2.5.5 SDS stacking buffer

Tris base (0.5 M), pH adjusted to 6.8 with diluted hydrochloric acid.
Stored at 4 °C.

2.5.6 SDS resolving buffer

Tris base (1.5 M), pH adjusted to 8.8 with diluted hydrochloric acid.
Stored at 4 °C.

2.5.7 10% (w/v) ammonium persulfate

Ammonium persulfate (10% w/v). Stored at 4 °C.

2.5.8 Electrode running buffer (10x) for SDS-PAGE

Tris base (3.03% w/v), glycine (14.4% w/v) and SDS (1% w/v). Stored at room temperature. Diluted 1:10 prior to use.

2.5.9 Protein loading dye (4x) for SDS-PAGE

Glycerol (40%), bromophenol blue (0.08% w/v), SDS (8% w/v), β -mercaptoethanol (4%) and Tris-HCl (0.2 M), pH adjusted to 6.8 with diluted sodium hydroxide. Diluted 1:4 before use.

2.5.10 Protein purification buffer A

(EDTA) (1 mM), Tris-base (50 mM), β -mercaptoethanol (1 mM) and sodium chloride (10 mM), pH adjusted to 7.5 with diluted hydrochloric acid. The solution was filter-sterilised and degassed *in vacuo*. Stored at room temperature.

2.5.11 Protein purification buffer B

EDTA (1 mM), Tris-base (50 mM), β -mercaptoethanol (1 mM) and sodium chloride (1 M), pH adjusted to 7.5 with diluted hydrochloric acid. The solution was filter-sterilised and degassed *in vacuo*. Stored at room temperature.

2.5.12 Phosphate buffer

Potassium phosphate (10 mM), pH adjusted to 7.0 with diluted hydrochloric acid. The solution was degassed *in vacuo* and stored at room temperature.

2.5.13 Protein dialysis buffer

Protein purification buffer A was diluted 50-fold.

2.6 DNA Purification

All the buffers were stored at room temperature unless stated otherwise.

2.6.1 DNA miniprep buffers

Buffer P1 (suspension buffer): Tris-HCl (50 mM), EDTA (10 mM) RNase A (25 $\mu\text{g mL}^{-1}$), pH adjusted to 8.0 with diluted sodium hydroxide. Stored at 4 °C.

Buffer P2 (lysis buffer): NaOH (0.2 M) and SDS (1% w/v).

Buffer N3 (neutralisation and binding buffer): guanidine hydrochloride (4 M) and potassium acetate (0.5 M), pH adjusted to 4.2 with diluted hydrochloric acid.

Buffer PB (wash buffer): guanidine hydrochloride (8 M), ethanol (38% v/v), and Tris-HCl (20 mM), pH adjusted to 6.6 with diluted sodium hydroxide.

Buffer PE (wash buffer): NaCl (20 mM), ethanol (80% w/v) and Tris-HCl (2 mM), pH adjusted to 7.5 with diluted sodium hydroxide.

Buffer EB (elution buffer): Tris-HCl (10 mM), pH adjusted to 8.5 with diluted sodium hydroxide.

2.6.2 Agarose gel DNA extraction

Buffer QG (gel solubilising buffer): guanidine thiocyanate (5.5 M) and Tris-HCl (2 mM), pH adjusted to 6.6 with diluted sodium hydroxide.

Buffer PE (wash buffer): NaCl (20 mM), ethanol (80% w/v) and Tris-HCl (2 mM), pH adjusted to 7.5 with diluted sodium hydroxide.

Buffer EB (elution buffer): Tris-HCl (10 mM), pH adjusted to 6.6 with diluted sodium hydroxide.

2.6.3 DNA from PCR purification miniprep

Buffer PB (wash buffer): guanidine hydrochloride (8 M), ethanol (38% v/v) and Tris-HCl (20 mM), pH adjusted to 7.5 with diluted sodium hydroxide pH 6.6.

Buffer PE (wash buffer): NaCl (20 mM), ethanol (80% w/v) and Tris-HCl (2 mM), pH adjusted to 7.5 with diluted sodium hydroxide.

Buffer EB (elution buffer): Tris-HCl (10 mM), pH adjusted to 8.5 with diluted sodium hydroxide.

2.6.4 Nucleotide removal miniprep

Buffer PN (binding buffer): sodium perchlorate (16 M) and sodium acetate (0.5 M), pH adjusted to 4.2 with diluted hydrochloric acid.

Buffer PE (wash buffer): NaCl (20 mM), ethanol (80% w/v) and Tris-HCl (2 mM), pH adjusted to 7.5 with diluted sodium hydroxide.

Buffer EB (elution buffer): Tris-HCl (10 mM), pH adjusted to 8.5 with diluted sodium hydroxide.

2.7 *E. coli* Strains

2.7.1 Cloning strains

XL1-Blue ultracompetent cells (Stratagene, CA, USA) were used for cloning following site directed mutagenesis and ligation. The transformation

efficiency was $\geq 1 \times 10^9$ cfu/ μ g for super competent cells and $\geq 1 \times 10^{10}$ cfu/ μ g for electro competent cells

2.7.2 Expression strains

For expression of Im9 and all the mutants BL21(DE3) competent cells (Stratagene, CA, USA) were used. IPTG was used to induce protein expression.

2.7.3 Preparation of ultra-competent cells

XL1-Blue cells from a glycerol stock were streaked onto LB agar containing 20 μ g mL⁻¹ tetracycline, and incubated at 37 °C for 16 h. A negative control was performed using LB agar containing 20 μ g mL⁻¹ ampicillin.

100 mL of LB medium containing 20 μ g mL⁻¹ tetracycline was inoculated with a single colony from the agar plate and incubated at 37 °C overnight. This culture was subcultured (1:50 dilution) into 100 mL of LB medium containing 20 μ g mL⁻¹ tetracycline and incubated at 37 °C until an optical density of 0.6 (at 600 nm) was reached.

The culture was held on ice for 15 min and then harvested in a Thermo IEC 243 Centra CL3R centrifuge at 2750 *g*, 4 °C for 10 min. Cells were re-suspended in RbCl solution 1 (40% of the culture volume), held on ice for 15 min, and the centrifugation step repeated. The pellet was re-suspended in RbCl solution 2 (4% of the culture volume), held on ice for 10 min, and transferred to sterile Eppendorf tubes in 100 μ L aliquots. Cell suspensions were flash frozen in liquid nitrogen and stored at -80 °C. Transformation efficiency was calculated (Section 2.7.6) and expressed as cfu/ μ g DNA used.

2.7.4 Preparation of competent cells

BL21(DE3) cells from a glycerol stock were streaked onto non-selective LB agar and incubated at 37 °C for 16 h. A negative control was performed using LB agar containing 20 µg mL⁻¹ ampicillin.

100 mL of non-selective LB medium was inoculated with a single colony from the agar plate and incubated at 37 °C overnight. This culture was subcultured (1:50 dilution) into 100 mL of non-selective LB medium and incubated at 37 °C until an optical density of 0.6 (at 600 nm) was reached. Competent cells were then prepared as described in section 2.7.3.

2.7.5 Preparation of electrocompetent cells

XL1-Blue from a glycerol stock were streaked onto LB agar containing 20 µg mL⁻¹ of tetracycline and incubated at 37 °C for 16 h. A negative control was performed using LB agar containing 20 mg mL⁻¹ ampicillin.

15 mL of LB medium containing 20 µg mL⁻¹ tetracycline was inoculated with a single colony from the agar plate and incubated at 37 °C overnight. This culture was subcultured (1:200) into LB medium containing 20 µg mL⁻¹ tetracycline and incubated at 37 °C until an optical density of 0.7 (at 600 nm) was reached.

The culture was held on ice for 15 min and then harvested in a Thermo IEC 243 Centra CL3R centrifuge at 2750 g, 4 °C for 10 min. Cells were re-suspended in ice-cold sterile 10% (v/v) glycerol solution (50% of the culture volume) and the centrifugation step repeated. The re-suspension and centrifugation procedure was repeated two more times. The final pellet was resuspended in ice-cold sterile 10% (v/v) glycerol solution (2% of the culture

volume), and transferred to sterile Eppendorf tubes in 40 μL aliquots. Cell suspensions were flash frozen in liquid nitrogen and stored at $-80\text{ }^{\circ}\text{C}$. Transformation efficiency was calculated (Section 2.7.6) and expressed as cfu/ μg DNA used.

2.7.6 Calculation of the transformation efficiency

The transformation efficiency, expressed in colony formation units (cfu) per μg , of the competent cells was calculated with the following equation:

$$\text{Transformation efficiency (cfu}/\mu\text{g)} = \frac{\text{cfu on plate}}{\text{ng of DNA plated}} \times \frac{1 \times 10^3 \text{ ng}}{\mu\text{g}}$$

2.7.7 Transformation of cells with plasmids

Competent cells (Section 2.7.3 and 2.7.4) were thawed on ice and plasmid solution (1 μL) was added under sterile condition. After incubation for 30 min on ice the cells were heat shocked in a water bath at $42\text{ }^{\circ}\text{C}$ for 1 min then held on ice for at least 2 min. Non-selective LB medium (1 mL) was added and, after 1 h incubation at $37\text{ }^{\circ}\text{C}$, cells were harvested by centrifugation, resuspended in LB medium (100 μL), were spread onto LB agar containing $20\text{ }\mu\text{g mL}^{-1}$ ampicillin and incubated overnight at $37\text{ }^{\circ}\text{C}$. Plates were stored at $4\text{ }^{\circ}\text{C}$. A negative control was performed using sterile water in place of the plasmid solution.

2.7.8 Transformation of electrocompetent cells with DNA phagemid

Electrocompetent cells (Section 2.7.5) were thawed on ice and plasmid solution (1 μL) was added under sterile condition. The cells were transferred to

a pre-chilled sterile electroporation cuvette (0.1 cm gap), and pulsed once at 2500 V. Sterile SOC medium (960 μL , prewarmed to 37 °C) was immediately added, the solution was transferred to a sterile tube and incubated for 1 h at 37 °C with shaking at 200 rpm. 200 μL of the culture were spread onto LB agar containing 20 $\mu\text{g mL}^{-1}$ carbenicillin and incubated overnight at 37 °C. Plates were stored at 4 °C. A negative control was performed using sterile water in place of the plasmid solution.

2.8 Manipulation of DNA

2.8.1 Determination DNA concentration

Plasmid concentration was determined measuring the absorbance at 260 nm by UV spectroscopy on a UV-spectrophotometer (NanoDrop 1000, Thermo Scientific).

2.8.2 Site directed mutagenesis

DNA mutations were performed using the Stratagene® QuikChange® site directed mutagenesis kit following the manufacturer's instructions. The primers were designed to obtain the desired Im9 mutants (Table 2.1):

Table 2.1: Primers used to obtain the desired Im9 mutants by site directed mutagenesis

Mutant	Template	Primers
Im9-F15A	Im9	Fwd 5' -GTGATTATACAGAAGCTGAAGCTTTACAGCTTGTAACAAC-3' Bwd 5' -GTTGTTACAAGCTGTAAAGCTTCAGCTTCTGTATAATCAC-3'
Im9-Y10H	Im9	Fwd 5' -GCATAGCATTAGTGATCATACAGAAGCTGAAGC-3' Bwd 5' -GCTTCAGCTTCTGTATGATCACTAATGCTATGC-3'
Im9-Y10H-F15A	Im9-F15A	Fwd 5' -GCATAGCATTAGTGATCATACAGAAGCTGAAGC-3' Bwd 5' -GCTTCAGCTTCTGTATGATCACTAATGCTATGC-3'
Im9-W74A	Im9	Fwd 5' -GTAAAACAATGGAGAGCCGCTCACGGTAAGTCAGGATTTAAC-3' Bwd 5' -GTTAAATCCTGACTTACCGTGAGCGGCTCTCCATTGTTTTAC-3'
Im9-N78H	Im9	Fwd 5' -GTAAAACAAGCGAGAGCCGCTCACGGTAAGTCAGGATTTAAC-3' Bwd 5' -GTTAAATCCTGACTTACCGTGAGCGGCTCTCGCTTGTTTTAC-3'
Im9-L52E/W74A/N78H	Im9-W74A/N78H	Fwd 5' -GCACCCTAGTGGTAGTGATGAAATATATTACCCAAAAGAAGG-3' Bwd 5' -CCTTCTTTTGGGTAATATATTTTCATCACTACCACTAGGGTGC-3'
Im-L52D/W74A/N78H	Im9-W74A/N78H	Fwd 5' -GCACCCTAGTGGTAGTGATGATATATATTACCCAAAAGAAGG-3' Bwd 5' -CCTTCTTTTGGGTAATATATATCATCACTACCACTAGGGTGC-3'
Im9-D51A/L52E/W74A/N78H	Im9-L52E/W74A/N78H	Fwd 5' GAGCACCCCTAGTGGTAGTGCGGAAATATATTACCCAAAAG 3' Bwd 5' CTTTTGGGTAATATATTTCCGCACTACCACTAGGGTGCTC 3'
Im-D51A/L52D/W74A/N78H	Im9-L52D/W74A/N78H	Fwd 5' GAGCACCCCTAGTGGTAGTGCGGATATATATTACCCAAAAG 3' Bwd 5' CTTTTGGGTAATATATATCCGCACTACCACTAGGGTGCTC 3'

In order to sub-clone the Im9-W74A/N78H gene in the phagemid pHEN2 a *SacI* restriction site was engineered instead of the stop codon (TAG)

Im9-*SacI*

Fwd 5' -GTCAGGATTTAAACAGGGCGAGCTCATGAGTGCCGATGG
GGTACC-3'

Bwd 5' -GGTACCCCATCGGCACTCATGAGCTCGCCCTGTTTAAAT
CCTGAC-3'

In order to construct the library of random mutants a *KpnI* restriction site was engineered in the Im9-W74A-N78H(*SacI*)

Im9-*KpnI*

Fwd 5' -GATGACTCACCTTCAGGTACCGTAAACACAGTAAAACAAGC-3'

Bwd 5' -GCTTGTTTTACTGTGTTTACGGTACCTGAAGGTGAGTCATC-3'

Table 2.2: Reagents for site-directed mutagenesis.

Reagent	Volume (μL)
dNTPs (40 mM total, 10 mM each NTP)	1
Forwards primer (100 μM)	1
Backwards primer (100 μM)	1
Parent plasmid	1
10x <i>Pfu</i> buffer	2.5
<i>Pfu</i> polymerase	1
ddH ₂ O	17.5
Total volume	25

Table 2.3: Site directed mutagenesis program.

Mutagenesis Step		Temperature (°C)	Time (sec)
X 25	1 Initial denaturation	95	60
	2 Denaturation	95	30
	3 Annealing	55	60
	4 Primer extension	72	360
	5 Final extension	72	720

After the PCR reaction the template DNA was digested with *DpnI*.

2.8.3 Polymerase chain reaction (PCR)

The oligonucleotide to construct the random library was designed as follow and amplified using PCR.

5' -CCTCACCTTCAGGTACCGTAAACNNKVNSAAANNKVNSAGAGCC

NNKCACGGTNNKTCAGGATTCAAGCAGGGCGAGCTC-3'

N=G, C, A, T; K=G, T; V=A, C, G; S=G, C

Primers for PCR amplification of random sequence:

Fwd 5' -GCGGCATTAACCTCACCTTCAGGTACCGTAAAC-3'

Bwd 5' -GCGGGGGCAGCTGAGCTCGCCCTGCTTGAATCC-3'

Table 2.4: PCR reagents for the amplification of the random sequence.

Reagent	Volume (μL)
dNTPs (100 mM total, 25 mM each dNTP)	0.8
Forwards primer (100 μM)	1
Backwards primer (100 μM)	1
Random oligonucleotide (32 pM)	1
10x <i>Pfu</i> buffer	10
<i>Pfu</i> polymerase (2.5 U/ μL)	2
ddH ₂ O	84.2
Total volume	100

Table 2.5: PCR program for the amplification of the random sequence.

Step	T ($^{\circ}\text{C}$)	Time (s)
1 Initial denaturation	95	45
X 25	2 Denaturation	45
	3 Annealing	30
	4 Primer extension	30
	5 Final extension	300

2.8.4 Digestion of DNA with restriction endonucleases

For every endonuclease the manufacturer's instructions were followed.

Digestion reactions with *Nco*I (buffer: 50 mM Tris-HCl, 100 mM NaCl, 10 mM MgCl₂, 1 mM DTT, pH 7.9), *Sac*I (buffer: 10 mM Bis-Tris-Propane-HCl, 10 mM MgCl₂, 1 mM Dithiothreitol, pH 7.0) and *Kpn*I (buffer: 10 mM Bis-Tris-Propane-HCl, 10 mM MgCl₂, 1 mM Dithiothreitol, pH 7.0) were performed sequentially with a purification step, using agarose gel electrophoresis, before the following reaction. The reactions were carried at 37 $^{\circ}\text{C}$ for 4 h.

Digestion of the parent plasmid after site directed mutagenesis reactions was performed with *DpnI* (buffer: 20 mM Tris-acetate, 50 mM potassium acetate, 10 mM Magnesium acetate, 1 mM DTT) for 1 h at 37 °C.

2.8.5 DNA visualisation and purification

Agarose (0.5-1.5% w/v) was dissolved in 50 mL of TAE buffer (40 mM Tris acetate, 2 mM EDTA pH 8.0) by microwave heating and poured into a Mini Gel Kit Tray (CSB Scientific). Once cold the gel was covered in TAE buffer and the DNA samples (0.25% bromophenol blue, 30% glycerol) were loaded into wells. Gels were run at 60 A for 75 min, stained with ethidium bromide (0.2 µg ml⁻¹) and visualised by fluorescence of the intercalated ethidium bromide following excitation at 254 nm with a GeneFlash UV light box (Syngene, Cambridge, UK). Molecular weight ladders (1 kb for circular plasmid and 100 bp for small fragments) were used to identify the appropriate bands. When required, DNA fragments were excised from the agarose gel with a clean scalpel blade and stored in an eppendorf tube.

DNA was extracted from the gel according to the manufacturer's instructions using the QIAquick gel extraction kit (QIAGEN, Crawley, UK) and stored at -20 °C.

2.8.6 Sub-cloning of Im9-W74A/N78H(*NcoI-SacI*) gene in pHEN2 phagemid

Im9-W74A/N78H(*NcoI/SacI*) gene was excised from plasmid pET21d by sequential double digestion with *NcoI* and *SacI*. The reaction mixture was visualised on an agarose gel and the band with the gene was cut out and DNA

extraction was performed. The Im9-W74A/N78H(*NcoI/SacI*) gene was inserted with T4 ligase (New England Biolabs) in previously double digested and purified pHEN2 phagemid. The reaction was performed at room temperature for 30 minutes.

XL1-Blue ultracompetent cells were transformed with the ligation reaction (2 μ L) and with the double digested phagemid as a control.

2.8.7 Library ligation

Sequential double digestion with *KpnI* and *SacI* was performed on the DNA insert codifying for the random mutations (Section 2.8.3). The reaction mixture was visualised on agarose gel to confirm the digestion. The digested insert was purified with the nucleotide removal miniprep (Section 2.6.4) and ligated, with T4 ligase (NEB), in a equimolar amount of previously double digested and purified pHEN2 phagemid. The reaction was performed at room temperature for 30 min. The DNA was precipitated (Section 2.8.9) and stored overnight at -80 °C.

2.8.8 Plasmid purification (QIAprep Spin Miniprep Kit)

LB medium (5 mL) containing 50 μ g mL⁻¹ ampicillin was inoculated with a single colony of cells containing the desired plasmid. After overnight incubation at 37 °C with shaking at 150 rpm in an Innova® 43 shaker (New Brunswick Scientific, Hertfordshire, UK) cells were harvested at 13,500 rpm in an Eppendorf centrifuge 5415R. The QIAprep® Spin Miniprep Kit (QIAGEN, Crawley, UK) was used to purify the plasmid following the manufacturer's

instructions, using EconoSpin™ All-in-1 mini spin columns (Epoch Biolabs Inc, TX, USA).

2.8.9 DNA precipitation

DNA was precipitated by adding glycogen ($20 \mu\text{g mL}^{-1}$), 0.1 volume of 3 M sodium acetate (pH 5.2) and 2.2 volumes of ethanol. The DNA solution was stored at -80°C overnight before centrifugation.

2.8.10 DNA sequencing

DNA of plasmid constructs and PCR products was sequenced by Lark Technologies (Cogenics, Essex, UK) or the DNA Sequencing Facility at School of Bioscience, Cardiff University.

2.9 Phage Display

2.9.1 Preparation of helper phage VCSM13

SB medium (2 mL) containing $20 \mu\text{g mL}^{-1}$ tetracycline was inoculated with XL1-Blue cells (2 μL) (Section 2.7.5) and incubated for 1 h at 37°C with shaking at 250 rpm. 1 μL of a 10^{-6} dilution (in SB medium) of commercially obtained VCSM13 was added to the XL1-Blue culture and incubated for 15 min at room temperature. Liquefied LB top agar (3 mL) was added and the resulting solution was poured onto non-selective LB agar and incubated overnight at 37°C .

The following day, SB medium (10 mL) containing 20 $\mu\text{g mL}^{-1}$ tetracycline was inoculated with XL1-Blue cells (10 μL) and incubated for 1 h at 37 °C with shaking at 250 rpm. A single VCSM13 plaque was transferred to the cell culture. The infected culture was added to SB medium (500 mL) containing 20 $\mu\text{g mL}^{-1}$ tetracycline and 20 $\mu\text{g mL}^{-1}$ kanamycin. The flask was incubated overnight at 37 °C with shaking at 250 rpm.

The culture was centrifuged at 2500 g for 15 min. The supernatant solution was incubated at 70 °C for 20 minutes and centrifuged again at 2500 g for 15 min. The supernatant solution was stored at 4 °C.

The titre of the phage helper preparation was calculated and expressed in plaque-forming units (pfu) per mL.

2.9.2 Library transformation

For each transformation 300 μL of XL1-Blue electrocompetent cells (Section 2.7.5) were incubated on ice for 1 min with 1.4 μg of ligated library (Section 2.8.7). The cells were electroporated (Section 2.7.8). The cuvette was flushed with 5 mL of SOC medium (Section 2.2.3) at room temperature and the suspension was transferred to a 50 mL polypropylene tube and incubated shaking at 37 °C for 1 h. 10 mL of warm SB medium (Section 2.2.4), 3 μL of 100 mg/mL carbenicillin and 30 μL of 5 mg/mL tetracycline were added. 2 μL of the culture were plated to calculate the titer. The 15 mL culture was incubated for 1 h at 37 °C, 4.5 μL of 100 mg/mL carbenicillin were added and the culture was incubated for another h. 2 mL of VCSM13 helper phage (Section 2.9.1) were added and the solution was transferred to a 500 mL bottle prior the addition of 183 mL of warm SB medium, 92.5 μL of 100 mg/mL carbenicillin and 370 μL of 5 mg/mL tetracycline. The culture was incubated for 2 h at 37 °C. 280 μL of 5100

mg/mL kanamycin was added and the culture was incubated overnight at 37 °C.

The cells were harvested by centrifugation and the supernatant was transferred in a clean bottle. PEG-8000 (4% w/v) and sodium chloride (3% w/v) were added. The solution was stored on ice for 30 min, centrifuged at 15,000 g for 15 min, and the supernatant was discarded and the phage pellets were resuspended in 2 mL of Tris buffer (15 mM pH 8) with BSA (1% w/v). The solution was centrifuged at 10,000 g and the supernatant was passed through a 0.2 µm filter. The phage preparation was used for panning.

2.9.3 Panning

The fresh prepared library (Section 2.9.2) was incubated with 50 µL of Dynabeads with no transition state analogue bounded for 1.5 h in order to remove the non-specific binders. The phage solution was then incubated for 1.5 h with Dynabeads with the transition state analogue bound. In the meantime 2 x 2 mL of SB medium were inoculated with 2 µL of XL1-Blue electrocompetent cells and incubated shaking until an optical density of 1 at 600 nm. The phage solution was removed and the beads were washed 4 times with Tween 20 (0.05%) in TBS. The beads were then incubated with 50 µL of 100 mM glycine-HCl (pH 2.2) for 10 min at room temperature. The solution was then transferred to a microcentrifuge tube containing 3 µL of 2 M Tris base. The neutralised solution was added to one of the 2 mL XL1-Blue cultures and it was incubated for 15 min at room temperature. 6 mL of SB medium, 1.6 µL of 100 mg/mL carbenicillin and 12 µL of 5 mg/mL tetracycline, were added to the culture in a 50 mL polypropylene tube. 2 µL of this culture were plated on LB agar/carbenicillin plates for the output titering. For the input titering 50 µL of

the other 2 mL SB culture were inoculated with 1 μ L of a 10^{-6} dilution of the phage preparation and plated after 15 min incubation on LB agar/carbenicillin plates. The 8 mL culture was incubated at 37 °C for 2 h and ampicillin was added after the first hour. The culture was transferred to a 500 mL bottle and 1 mL of VCSM13 helper phage (Section 2.9.1) plus 91 mL of SB medium with carbenicillin and tetracycline were added. The culture was incubated overnight at 37 °C (kanamycin was added after the first 2 h).

The cells were harvested by centrifugation and the supernatant was transferred in a clean bottle. PEG-8000 (4% w/v) and sodium chloride (3% w/v) were added. The solution was stored on ice for 30 min, centrifuged at 15,000 g for 15 min, and the supernatant was discarded and the phage pellets were resuspended in 2 mL of Tris buffer (15 mM pH 8) with BSA (1% w/v). The solution was centrifuged at 10,000 g and the supernatant was passed through a 0.2 μ m filter. The phage preparation was used for the next panning round.

2.10 Protein Expression

2.10.1 Test expression

To optimise condition a small-scale expression was performed for every mutated protein.

LB medium (100 mL) was inoculated with 1 mL of an overnight culture and incubated at 37 or 30 °C to an OD_{600 nm} of 0.6. Expression was induced with IPTG and allowed to continue for 5 h at 37 °C or overnight at 16 °C. The expression was sampled every hour (37 °C) or every 4 h (16 °C) for SDS- PAGE analysis. All mutants gave best results at 16 °C.

2.10.2 Large scale expression

BL21(DE3) competent cells (Stratagene, CA, USA) were transformed with the required plasmid and incubated overnight on LB agar containing 50 $\mu\text{g mL}^{-1}$ ampicillin at 37 °C. LB medium (100 mL) containing 50 $\mu\text{g mL}^{-1}$ ampicillin was inoculated with a single colony and incubated overnight at 30 °C with shaking at 150 rpm in a Innova® 43 shaker. This starter culture was then subcultured (1:100 dilution) into LB medium (6 x 500 mL) containing 50 $\mu\text{g mL}^{-1}$ ampicillin and incubated at 30 °C, 150 rpm to an $\text{OD}_{600 \text{ nm}}$ of 0.6. Expression was induced with IPTG and the temperature was lowered to 16 °C. The expression was allowed to proceed overnight. Cells were harvested *via* centrifugation in a Sorvall RC5C Plus (Thermo Scientific, UK) centrifuge (rotor SLA-3000) at 16000 g for 12 minutes at 4 °C and stored at -20 °C.

2.11 Protein Purification

2.11.1 DEAE anion exchange chromatography

Frozen cells were thawed on ice, resuspended in protein purification buffer A, and lysed by sonication for 3 min. Cellular debris was removed via centrifugation in a Sorvall RC5C Plus centrifuge (rotor SS-34) at 34000 g for 20 min at 4 °C, and the supernatant solution was applied to a DEAE (diethylaminoethyl) column (70 mL). After equilibration with purification buffer A (3 column volumes) to remove unbound proteins, the desired protein was eluted with a linear salt gradient (NaCl 50-1000 mM) using purification buffer B over 6 column volumes. The fractions containing the protein, as judged by SDS-PAGE, were pooled and dialysed against protein dialysis buffer.

2.11.2 SourceQ™ anion exchange chromatography

The dialysed solution following DEAE purification was applied to a Tricorn™ (GE healthcare) column packed with SourceQ™ (GE healthcare) resin (8 mL). The column was washed with purification buffer A (3 column volumes) to remove unbound proteins. The desired protein was then eluted in purification buffer B with a linear salt gradient (NaCl 50-1000 mM) over 20 column volumes. The fractions containing the protein, as judged by SDS-PAGE, were pooled and dialysed against protein dialysis buffer.

2.11.3 Ni affinity chromatography

Frozen cells were thawed on ice, resuspended in 500 mM phosphate buffer (pH 7), and lysed by sonication for 3 min. Cellular debris was removed *via* centrifugation in a Sorvall RC5C Plus centrifuge (rotor SS-34) at 34000 g for 20 min at 4 °C, and the supernatant solution was applied to a HiTrap™ (GE Healthcare) column (1 mL) charged with NiSO₄ solution. The protein was eluted with phosphate buffer containing increasing concentrations of imidazole (10-200 mM). The fractions containing the protein, as judged by SDS-PAGE, were pooled and, dialysed against 500 mM phosphate buffer (pH 7).

2.11.4 Amicon ultrafiltration

High molecular weight impurities were removed by passing the solution through a YM30 membrane. The pure protein was collected in the flow-through, freeze dried, and stored at -20 °C.

2.12 Protein Characterisation

2.12.1 SDS polyacrylamide gel electrophoresis

SDS gels (5% stacking gel, 17% running gel) were produced following the Mini-PROTEAN™ 3 Cell (Bio-Rad) protocols. Protein loading dye was added to the samples, which were then heated to 90 °C for 1 min to denature proteins prior to loading onto the gel. Gels were run at 200 V for 55 min in 1 x electrode running buffer. Proteins were visualised by staining the gel (0.25% w/v Coomassie Brilliant Blue R250 in 40% v/v ethanol, 10% v/v glacial acetic acid) for 1 h and destaining it (40% v/v ethanol, 10% v/v glacial acetic acid) for 1 h.

2.12.2 MALDI-TOF

A solution of 10 mg mL⁻¹ sinapinic acid (for proteins) or 10 mg mL⁻¹ α -cyano-4-hydroxycinnamic acid (CHCA) (for peptides) in water:acetonitrile:trifluoroethanoic acid (40:60:0.1; v/v) was used as matrix. Sandwich sample preparation was used: 1 μ L of the matrix was loaded on the plate and allow to dry, then 1 μ L of sample solution was loaded on the top of the matrix and again allow to dry. Finally 1 μ L on the matrix was loaded on the top to complete the process. The samples were analysed with a MALDI-TOF Micro MX spectrophotometer (Waters, Manchester, UK) used in reflection mode, positive polarity, mass range 3,000-12,000 to identify the protein of interest (molecular weight of 9,000-10,000) and mass range 1,500-3,500 to identify Apoxaldie (molecular weight of 1,900).

2.12.3 Determination of the concentration of Im9

The concentration of Im9 (WT and mutants) was determined by UV spectroscopy measuring the absorbance at 280 nm. An extinction coefficient of $11400 \text{ M}^{-1} \text{ cm}^{-1}$ (Wallis *et al.* 1994) was used for Im9. The extinction coefficient for the mutants were calculated using the ProtParam tool (Gasteiger *et al.* 2005).

2.12.4 Kinetics parameters for Im9

Kinetic data were acquired on a Shimadzu UV-2101PC spectrophotometer equipped with a Shimadzu CPS thermocontroller, in 100 mM MOPS (pH 7) at 17 °C. The production of *p*-nitrophenol was monitored at 320 nm ($\epsilon_{320\text{nm}} = 7600 \text{ M}^{-1} \text{ cm}^{-1}$). An excess of protein (150–600 μM) was used to determinate the pseudo-first-order rate constants. Due to its low aqueous solubility, the substrate *p*-nitrophenol acetate was dissolved in acetonitrile to give a stock solution of 15 mM. This solution (3 μL) was added to temperature equilibrated protein solution (997 μL) in a 10 mm path-length cuvette to give a final substrate concentration of 30 μM . A linear regression analysis of the pseudo first order data determined experimentally gave the second order rate constants as a function of the protein concentration.

2.13 Peptide Syntheses

2.13.1 Standard synthesis

Apoxaldie and Apoxaldie-Dab were synthesised using a microwave assisted Liberty peptide synthesiser (CEM) and standard Fmoc chemistry (Jones

1992). The Fmoc and side-chain protected amino acids [Asn, Cys (Trt); Cys (Acm); Glu (OtBu); Thr (tBu); Dab, Lys (Boc)] were coupled to a Rink amide resin on a polystyrene support. The amino acids were activated *in situ* using *O*-benzotriazol-1-yl-*N,N,N,N*-tetramethyluroniumhexafluorophosphate, hydroxybenzotriazole and, *N,N*-diisopropylethylamine (HBTU, HOBt, DIEA). All peptides were acetylated with acetic anhydride on resin.

Cleavage from the resin and deprotection was performed with 10 mL trifluoroacetic acid (TFA):water:phenol:triisopropylsilane (88:5:5:2; v/v) per gram of resin for 2 h at room temperature. After filtration to remove the resin the solution was concentrated under reduced pressure and the peptide was precipitated with ice-cold diethyl ether (3 × 5 mL). The precipitate was dissolved in water:acetonitrile (50:50 v/v) and lyophilised to give yields in the range of 5-10 mg of peptide.

2.13.2 Oxidation and purification of Apoxaldie

Apoxaldie was reduced with two equivalents of solid phase supported *tris*(2-carboxyethyl)phosphine (TCEP) to ensure complete reduction in 5 mM potassium phosphate buffer (pH 8). The final peptide concentration was 0.3 mg ml⁻¹ to prevent oligomerisation. The TCEP resin was removed and oxidation was performed stirring vigorously overnight. Water was removed under reduced pressure with gentle heating. Apoxaldie was purified by semi-preparative reversed phase high pressure liquid chromatography (RP-HPLC) on a LUNA 10 μ C18 column (250 × 10 mm) using a linear gradient of 60-100% v/v aqueous acetonitrile (containing 0.05% TFA) over 40 min at 5 mL min⁻¹. RP-HPLC analysis using a linear gradient of 60-100% v/v aqueous acetonitrile

(containing 0.05% TFA) over 30 min at 2 mL min⁻¹ on an analytical LUNA 10 μ C18 column (250 x 4.6 mm) revealed the peptide was pure.

2.13.3 Orthogonal preparation of Apoxaldie-Dab

Orthogonally protected peptides were synthesised by Alta Bioscience (University of Birmingham, U.K.) using trityl (Trt) and S-acetamidomethyl aminoacetal (Acm) protection for the thiol groups (Trt for Cys 1 and Cys 11, Acm for Cys 3 and Cys 15). Standard protecting groups were used for all other amino acids and capping of the N-terminal amino group was performed with acetic anhydride.

The peptide was received fully deprotected except for Cys 3 and Cys 15.

The first disulfide bond was formed by dissolving the peptide in 5 mM potassium phosphate buffer (pH 8) at a concentration of 0.5 mg mL⁻¹. Oxidation occurred overnight at room temperature. After concentration under reduced pressure with gentle heating, the peptide was purified by RP-HPLC on a LUNA 10μ C18 column (250 x 21.2 mm) using a linear gradient of 60-100% v/v aqueous acetonitrile (containing 0.05% TFA) over 1 hour at 5 mL min⁻¹.

The Acm protecting groups were removed with silver triflate (10 peptide equivalents) in TFA (5 mL) for 1 h at 4 °C. TFA was removed under reduced pressure with gentle heating and the peptide precipitated and washed with ice-cold diethyl ether (3 x 15 mL). The peptides were purified by RP-HPLC on a LUNA 10μ C18 column (250 x 21.2 mm) using a linear gradient of 60-100% v/v aqueous acetonitrile (containing 0.05% TFA) over 40 minutes at 5 mL min⁻¹.

The second disulfide bond was then formed by air oxidation (see above) and the final product isolated by HPLC on a LUNA 10μ C18 column (250 x 21.2

mm) using a linear gradient of 60-100% v/v aqueous acetonitrile (containing 0.05% TFA) over 40 minutes at 5 mL min⁻¹. The identity of the products was confirmed by MALDI-TOF-MS.

2.14 Apoxaldie and Apoxaldie-Dab Characterisation

2.14.1 Determination of the concentration of Apoxaldie-Dab

The concentration of Apoxaldie-Dab was determined by quantitative amino acid analysis (Alta Bioscience).

2.14.2 Determination of the concentration of Apoxaldie

The concentration of Apoxaldie in its oxidised form was determined by UV spectroscopy using 5,5'-dithiobis(2-nitrobenzoic acid) (Ellman's reagent) after reduction of the samples with solid supported TCEP.

Oxidised glutathione (GSSG) was used to construct a calibration curve. GSSG (0.125-1 mM) and Apoxaldie (unknown concentration) solutions (200 µL each) were reduced with TCEP slurry (100 µL) for 1 h under nitrogen. The samples were then incubated with Ellman's reagent for 15 min following the manufacturer's instruction and the absorbance was measured at 412 nm. Linear regression of the data gave the calibration curve.

2.14.3 Kinetics parameters for Apoxaldie and Apoxaldie-Dab

Kinetic data were acquired on a Shimadzu UV-2101PC spectrophotometer equipped with a Shimadzu CPS thermocontroller. The rate of production of pyruvate from decarboxylation of oxaloacetate was determined at 25 °C in *N,N*-bis(2-hydroxyethyl)-2-aminoethanesulfonic acid (BES) (1 M, pH 7) containing NaCl (10 mM). An enzyme-coupled assay, with lactate dehydrogenase reducing pyruvate to lactic acid using NADH as cofactor, was used to record kinetic data for Apoxaldie and Apoxaldie-Dab. The oxidation of NADH (0.2 M) was monitored at 340 nm following the linear decrease in absorbance ($\epsilon_{\text{NADH}} = 6.23 \times 10^3$) over time.

The concentration of the peptides was held at 10 μM and the concentration of oxaloacetate was varied from 5 mM to 200 mM.

2.15 CD Spectroscopy

Circular dichroism (CD) spectroscopy was performed using a Chirascan™ (Applied Photophysics, UK). All experiments were carried out in degassed 10 mM potassium phosphate buffer (pH 7.0), using a 0.1 cm path-length cuvette. All the blank experiments were recorded with only buffer in the cuvette. A working concentration of 10 μM was used for Im9 at its mutants as well for Apoxaldie and Apoxaldie-Dab.

Spectra were recorded between 190 and 400 nm. Thermal unfolding experiments were performed between 5 and 90 °C in a 1 mm path cuvette with a temperature gradient of 0.5 °C min⁻¹. The signal (in mdeg) obtained from the CD experiments was converted to MRE using the following equation

$$\theta_{MRE} = \Theta \div (10 \times n \times c \times l)$$

where

Θ = CD signal in mdeg

n = number of backbone peptide bonds

c = concentration of the sample in mol dm⁻³

l = path-length of the cuvette in cm

2.16 Synthesis of Transition State Analogue

A mixture of ethyl 4-(diethoxyphosphoryl)butanoate **1** (2.65 g, 10.6 mmol) and concentrated HCl (60 mL) was stirred under reflux for 15 h. After removal of the solvent, the residues was washed with dioxane (10 mL) and concentrated under reduced pressure to dryness. This process was repeated twice. The solid was then washed with toluene (10 mL) and concentrate under reduced pressure. This process was repeated twice. The crude compound **2** was obtained as white solid (1.7 g, 10.1 mmol) and was suspended in SOCl₂ (3 mL, 40.4 mmol) and DMF (200 μ L) as a catalyst. The mixture was stirred at 55 °C for 15 h. After cooling, excess of SOCl₂ was removed under reduced pressure. The crude acyl chloride **3** was obtained as green oil (2.25 g).

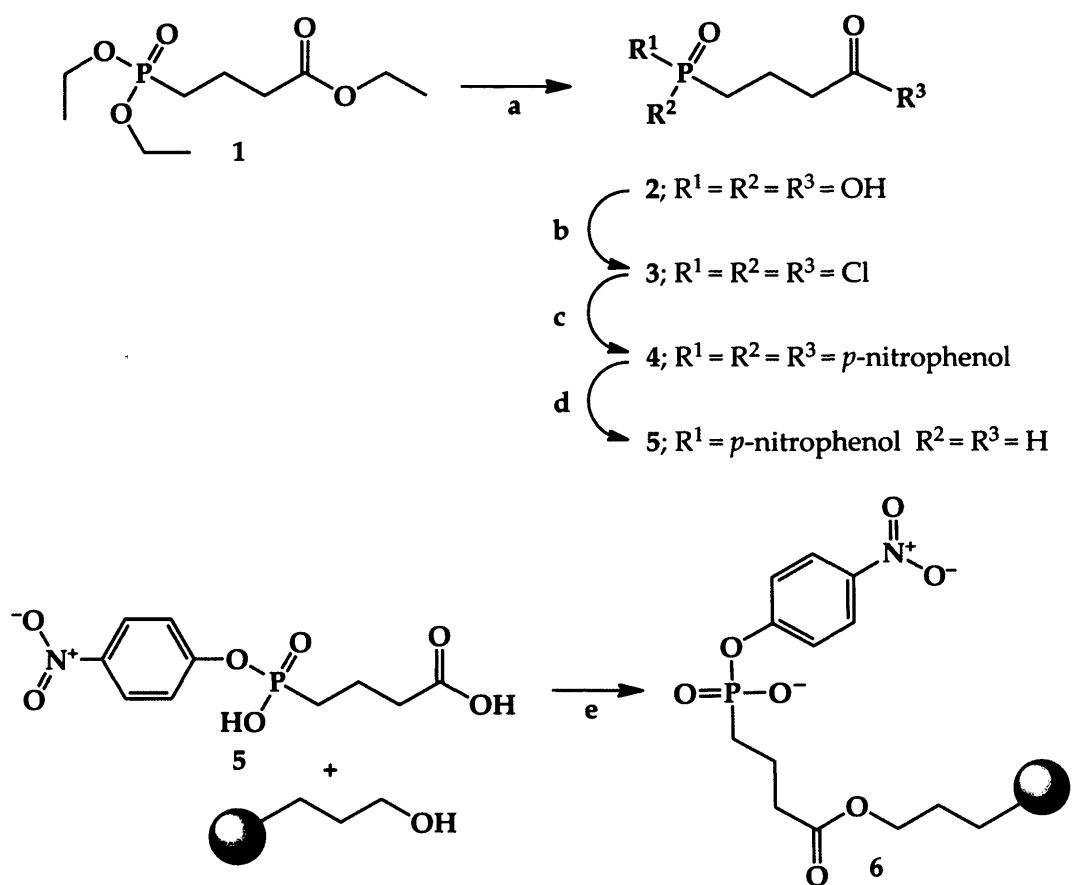
To *p*-nitrophenol (5.60 g, 40.4 mmol) in THF (40 mL) was added NaH (60% dispersion in paraffin liquid, 1.60 g (0.94 g), 40.4 mmol) with stirring at room temperature. Stirring was continued until evolution of H₂ ceased. To the mixture was added a solution of the crude trichloride **3** (1.7 g, 10.1 mmol) in THF (40 mL) gradually with vigorous stirring at room temperature. The mixture was stirred for 12 h. The reaction mixture was poured into an ice-

aqueous solution of HCl and then was extracted with ethyl acetate (3 x 50 mL). The combined organic layers were washed with brine, dried over MgSO₄ and concentrated under reduced pressure to dryness. Triester **4** was obtained as a brown powder (5.38 g, 25% from **1**).

The triester **4** (5.38 g, 10.1 mmol) was stirred vigorously in 0.5 M aqueous NaOH solution (200 mL) at room temperature for 48 h. After removal of unreacted impurities by filtration, the filtrate was acidified with 0.5 M aqueous HCl solution and concentrated under reduced pressure to dryness. The residue was purified by RP-HPLC (Luna: C-18 reverse phase column, CH₃CN:0.1% aqueous TFA = 20:80, 3.0 mL min⁻¹, 254 nm, retention time 14.6 min). The CH₃CN and TFA were removed under reduced pressure and the water by lyophilisation to give acid **5** as a white solid (231 mg, 8% from **1**). ¹H NMR (400MHz, D₂O) : δ 8.17 (d, *J*=9.0Hz, 2H), 7.17 (d, *J*=8.5Hz, 2H), 2.39 (m, 2H), 1.80~1.63 (m, 4H); ¹³C NMR (100 MHz, D₂O) δ 177.4, 158.1, 145.1, 126.2, 121.1, 34.7, 31.1, 27.8.

Tosyl-Dynabeads® (50 mg, 5-10 μmol) were stirred in CH₃CN (2.25 mL) and 10 μM aqueous NaOH (2.25 mL) at room temperature for 24 hours. The reaction was followed by analytical RP-HPLC monitoring the release of the tosyl group. The HO-Dynabeads were collected and washed with H₂O.

The beads (30 mg, 3-6 μmol) were mixed with acid **5** (2 mg, 6 μmol) in DMF (5 mL), *N,N'*-diisopropylcarbodiimide (1 μL, 7 μmol) and 4-dimethylaminopyridine (catalytic amount) were added to the mixture and stirred at room temperature for 24 h. The ester **6** was collected and washed with water (Scheme 2.1). The success of the reaction was confirmed by the presence of *p*-nitrophenol which was released by incubation of the beads with HCl.



Reagents and conditions: a) conc. HCL, reflux; b) SOCl_2 , DMF; c) *p*-nitrophenol, NaH, THF
 d) 0.5 M NaOH; e) DMF, DIC, DMAP.

Scheme 2.1: Synthesis of 4-(hydroxy(4-nitrophenoxy)phosphoryl) butanoate-Dynabeads®

3. APOXALDIE-DAB

3.1 Introduction

Small peptides without a definite secondary structure seldom display catalytic activity, so a simple array of functional groups is not enough to induce function. Multidisulfide peptides, with defined secondary structure are therefore of particular interest. The spacing between the cysteines plays a key role in the correct formation of the disulfide bonds avoiding formation of non-native forms (Loret *et al.* 1990; Olivera *et al.* 1991; Chau *et al.* 1992; Huyghues-Despointes *et al.* 1992; Ramalingam *et al.* 1993; Xu *et al.* 1994; Dauplais *et al.* 1995; Fainzilber *et al.* 1995; Price-Carter *et al.* 1996b; Price-Carter *et al.* 1996a; Landon *et al.* 1997; Blanc *et al.* 1998; Lecomte *et al.* 1998; Pegoraro *et al.* 1999; Savarin *et al.* 1999; Martin *et al.* 2000; Chin *et al.* 2001; Chagot *et al.* 2005; Fuller *et al.* 2005; Han *et al.* 2006; Tan *et al.* 2006; Kang *et al.* 2007; Zugasti-Cruz *et al.* 2008; Daly *et al.* 2009; Halai *et al.* 2009; Jacob *et al.*). Small rigid peptides can be used as scaffolds on which to graft active sites. The intrinsic stability of those peptides allows several mutations on the solvent exposed faces without disturbing their folding.

The 18-residue peptide apamin is a neurotoxin present in the venom of honeybees (Figure 3.1). Its neurotoxic activities consist in the selective blocking of the small conductance Ca^{2+} activated K^{+} (SKCa) channels (Fletcher *et al.* 2007). Two disulfide bonds make it extremely stable to thermal and chemical denaturation: apamin does not unfold completely even at 70 °C in 6 M guanidinium hydrochloride (Pease *et al.* 1988). Apamin has been used as scaffold to create hybrids with a broad range of functions (Pease *et al.* 1990; Brazil *et al.* 1997; Turner *et al.* 2003; Weston *et al.* 2004; Li *et al.* 2009).

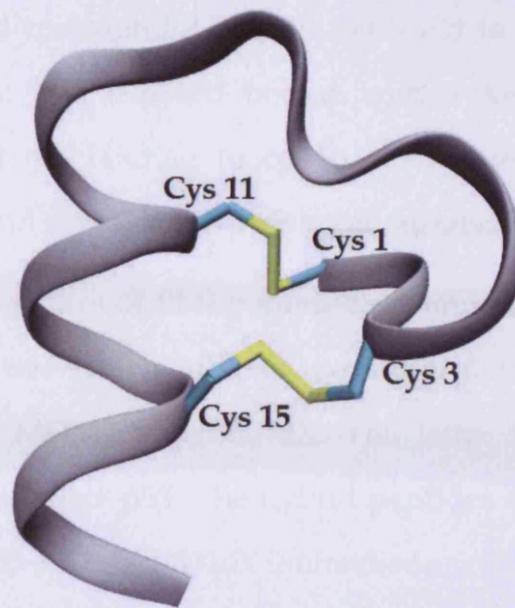


Figure 3.1: Molecular model of apamin based on NMR structure. The two disulfide bonds (Cys 1-Cys 11 and Cys 3-Cys 15) make the peptide extremely stable (Pease *et al.* 1988)

Pease *et al.* synthesised two hybrid apamin-S peptides. The S peptide consisting of the first 20 amino acid of RNase A and binds to S protein (with an association constant of 10^6 M^{-1}), the large fragment of RNase A from which it was cleaved, reactivating the nuclease activity (Potts *et al.* 1963). The two hybrids were constructed replacing the non-cysteine amino acids from the helical region of apamin with those from the S-peptide. Both hybrids folded into a structure essentially identical to that of the parent apamin and restored the nuclease activity to a level essentially equivalent to that found with the S peptide itself (Pease *et al.* 1990).

Brazil *et al.* synthesised three hybrid peptides encompassing a portion of apamin and a sequence from an amphipathic helix in the N-terminal region of bovine rhodanase in order to study the binding to molecular chaperone cpn60, which helps the protein fold (Brazil *et al.* 1997). The hybrids were designed to present either a hydrophobic or hydrophilic face of the amphipathic helix. The

peptide with the exposed hydrophilic face did not bind to cpn60, instead the one with a hydrophobic face exposed bound with a K_D of 106 μM . This observation showed that the binding to cpn60 is favoured by a secondary structure that organises and exposes a hydrophobic surface.

Li *et al.* grafted 4 residues of PMI peptide to apamin (Li *et al.* 2009). PMI is a 12 amino acids peptide which inhibits the interactions between protein tumour suppressor p53, MDM2 and MDMX. The latter two are oncogenic proteins that negatively regulate p53. The hybrid peptides demonstrated to be potent inhibitors of the p53-MDM2/MDMX interaction.

Turner *et al.* fused apamin to the DNA binding helix of MyoD creating the protein apa-MyoD (Turner *et al.* 2004). The disulfide stabilised helix from apamin was used to hold apa-MyoD's DNA recognition helix in a predominantly α -helical conformation. Consequently the DNA complexes with apa-MyoD showed an increased thermal stability (by 13 $^{\circ}\text{C}$) and a 20-fold increase in binding specificity relative to MyoD. The significant increase of DNA binding demonstrated by the oxidised apa-MyoD was due the presence of the disulfide bonds. The reduced apa-MyoD, on the other hand, showed only a 2-fold increase of binding due the lack of the disulfide bonds.

Weston *et al.* synthesised Apoxaldie, an apamin-derived miniature enzyme with oxaloacetate decarboxylase activity (Weston *et al.* 2004). The active site was designed with three Lys residues in positions 9, 13 and 16 on the solvent exposed face of the apamin α -helix (Figure 3.2).

The proximity of two other Lys had been reported to decrease the pK_a of an amino residue of Lys by two orders of magnitude in the Oxaldie peptides (Johnsson *et al.* 1993; Taylor *et al.* 2001; Taylor *et al.* 2002). The proposed mechanism of action was based on the amine-catalysed decarboxylation of β -

keto acids (as in acetoacetate decarboxylase) where a Schiff-base is formed after the attack of the amino group on the carbonyl group (Scheme 3.1).

Apoxaldie showed remarkable thermal stability, with no loss of secondary structure up to 75 °C. The peptide showed a remarkable chemical stability to denaturant with a little loss of helicity in presence of 6 M guanidinium chloride. The rate of decarboxylation was increased by almost four orders of magnitude compared to simple amines, while Apamin didn't show any measurable activity (Weston *et al.* 2004).

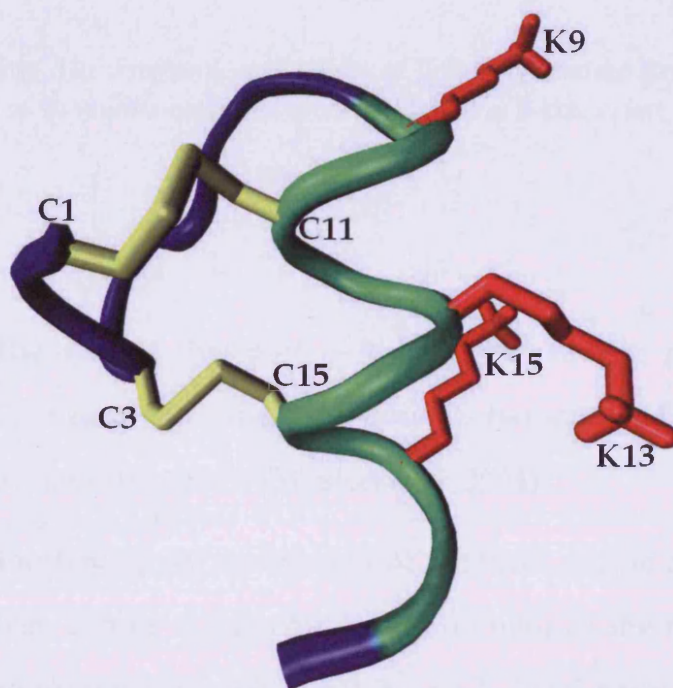
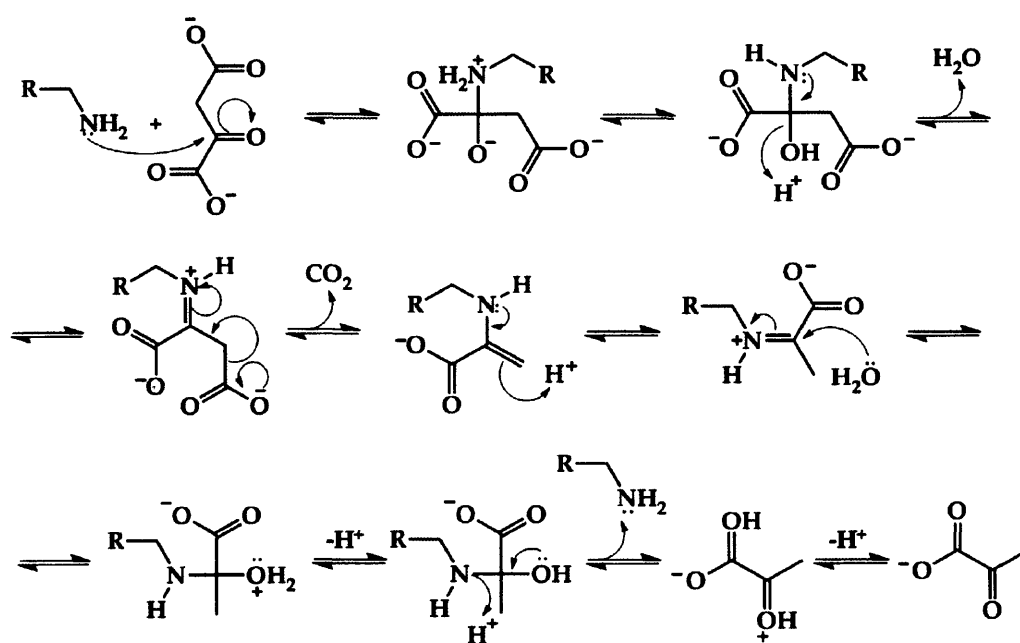


Figure 3.2: Molecular model of Apoxaldie-1 based on the NMR structure of apamin. Cysteine residues are shown in yellow, lysines in red, helical residues in green, and other residues in blue.



Scheme 3.1: Proposed mechanism of action of oxaldie peptides. The mechanism was based on the amine-catalysed decarboxylation of β -keto acids.

3.1.1 Aims

The aim of this part of the project was to produce a new miniature enzyme based on the previously-characterised miniature oxaloacetate decarboxylase Apoxaldie (Weston *et al.* 2004).

Shortening the side chains of the basic amino acids in the active site will bring them nearer to the chiral environment of the α -helix of apamin scaffold therefore increase the selectivity toward chiral substrates (Colonna *et al.* 1983; Takagi *et al.* 2000; Kelly *et al.* 2004; Blank *et al.* 2006; Maayan *et al.* 2009). At the same time a shorter side chains will bring the amino groups closer. This spatial proximity should further lower the basic amino acids' pK_a compared to Apoxaldie and consequently increase the catalytic activity.

3.2 Results and Discussion

3.2.1 Design of a new apamin-based miniature enzyme

In order to bring the active site of Apoxaldie closer to the chiral environment of the apamin scaffold and to restrict the flexibility of the side chain of Lys residues in the active-site, homologues with shorter side chains were considered (Figure 3.3).

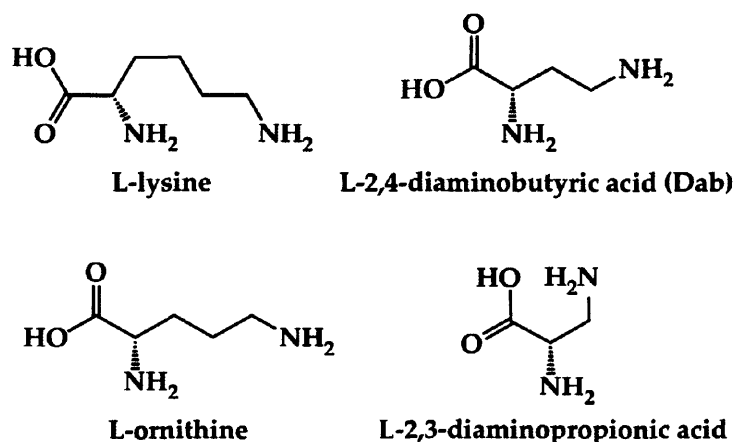


Figure 3.3: Structure of lysine and its homologues.

The possible candidates were L-2,3-diaminopropionic acid (three methylene groups fewer than lysine), L-2,4-diaminobutyric acid (Dab) (two methylene groups fewer than lysine), and L-ornithine (one methylene group fewer than lysine). The polar side chain of L-2,3-diaminopropionic acid can effectively form H-bonds with the CO or NH groups in the peptide backbone (Padmanabhan *et al.* 1996). This characteristic leads to the destabilisation of the α -helix. For this reason L-2,3-diaminopropionic acid was not considered. As L-ornithine has only one methylene group less than lysine, its physiochemical

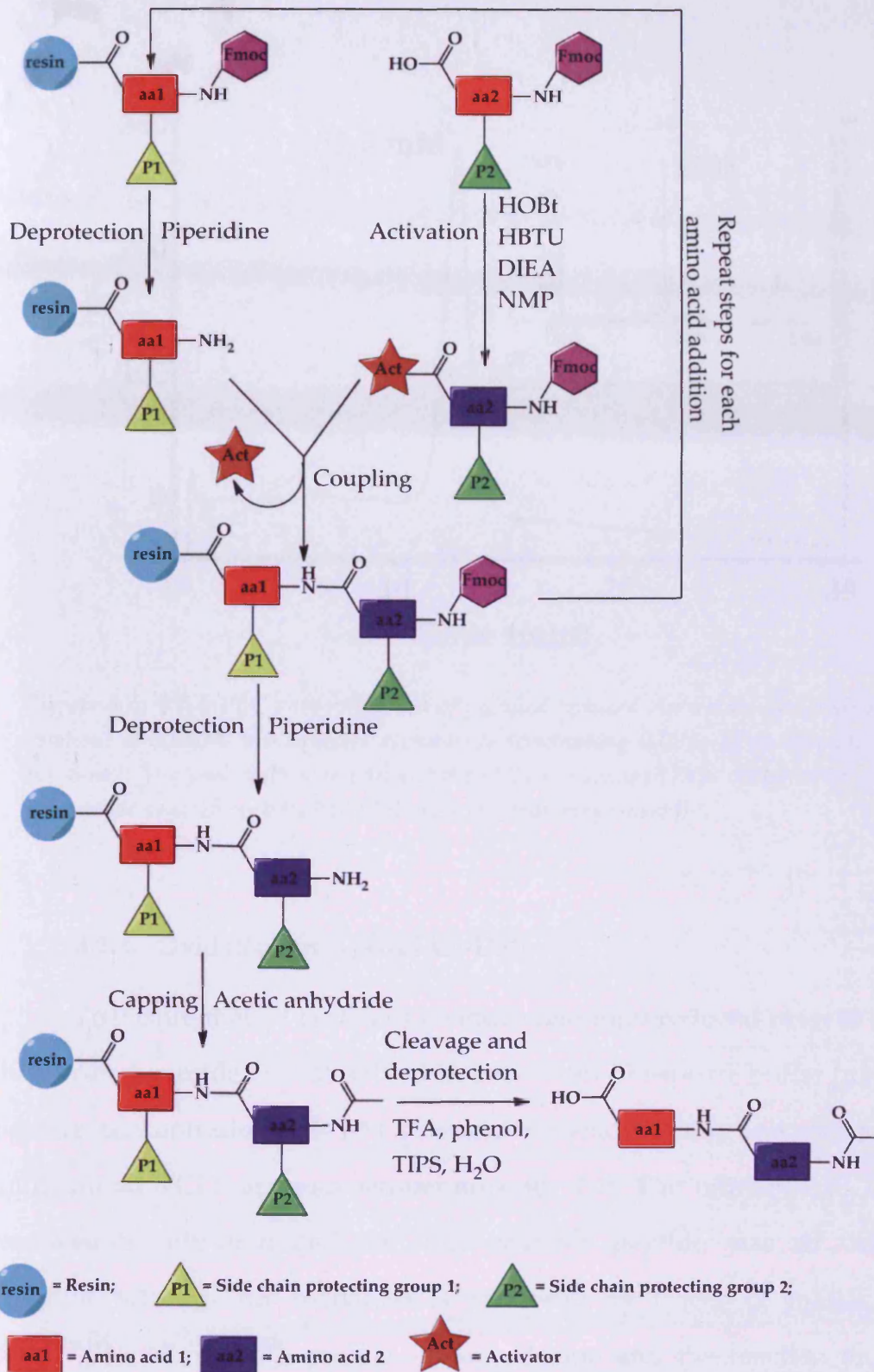
properties remain very similar to lysine. Dab's side chain is between L-2,3-Diaminopropionic acid and L-ornithine side chain length. The miniature enzyme was then designed using Dab and named as Apoxaldie-Dab.

3.2.2 Synthesis and purification of Apoxaldie-Dab

Apoxaldie-Dab was synthesised using standard Fmoc chemistry solid phase peptide synthesis (Jones 1992) using a microwave assisted peptide synthesiser (Scheme 3.2). The Fmoc and side-chain protected amino acids were coupled with Rink amide resin on polystyrene support. The amino acids were activated *in situ* using *O*-benzotriazol-1-yl-*N,N,N,N*-tetramethyluroniumhexafluorophosphate, hydroxybenzotriazole and, *N,N*-diisopropylethylamine (HBTU, HOBt, DIEA). The peptide was N-terminally acetylated on resin with acetic anhydride. Apoxaldie-Dab was then cleaved and fully deprotected in trifluoroacetic acid (TFA) in the presence of scavengers (phenol; triisopropylsilane; and water) to give the final product



The peptide was purified by reversed phase high pressure liquid chromatography (RP-HPLC) with a Luna C18 column (5 μm , 150 x 10 mm) using a linear gradient 60-90% aqueous acetonitrile containing 0.05% TFA over 30 min, flow rate 5 mL min⁻¹ and was detected by MALDI-TOF mass spectrometry, to have a mass of 1895 which is in agreement with the calculated value of 1894.88 (Figure 3.4). The final yield of the purified peptide was around 20%.



Scheme 3.2: Solid phase peptide synthesis using standard Fmoc chemistry.

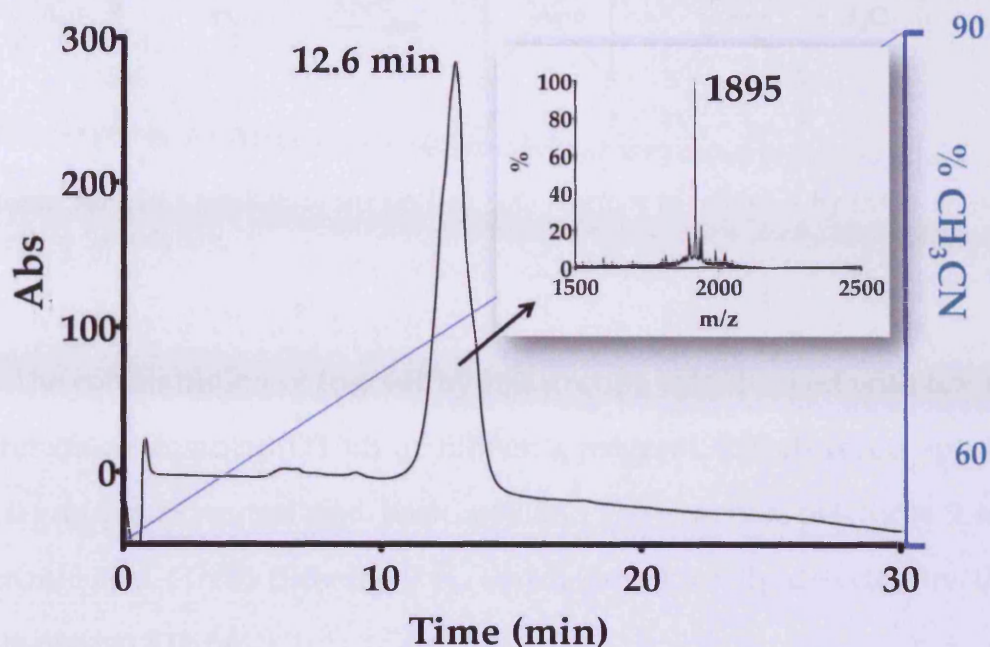
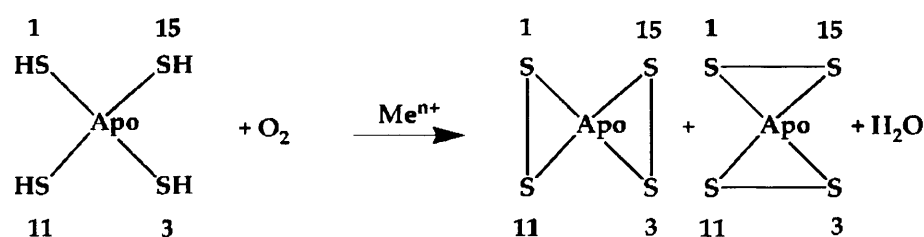


Figure 3.4: RP-HPLC chromatogram of purified reduced Apoxaldie-Dab using a linear gradient of 60-90% *v/v* aqueous acetonitrile (containing 0.05% TFA) over 30 min at 5 mL min⁻¹. The peak with a retention time of 12.6 minutes (73% acetonitrile) contained the peptide as confirmed by MALDI mass spectrometry (insert).

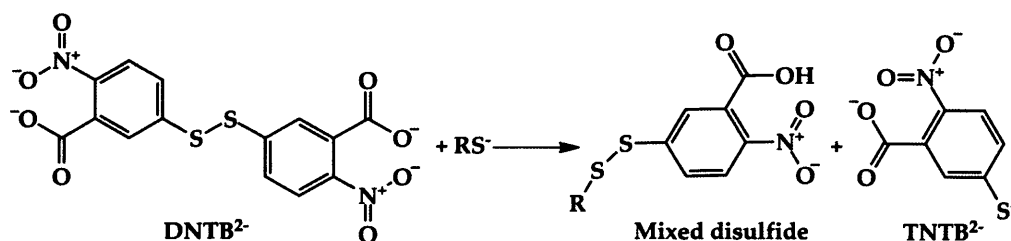
3.2.3 Oxidation of Apoxaldie-Dab

To ensure that all cysteines residue were fully reduced prior to oxidation, the purified peptide was dissolved in potassium phosphate buffer (pH 8.0, final peptide concentration 250 μ M to avoid polymerisation) and incubated with immobilised TCEP at room temperature for 1 h. The immobilised TCEP was removed by filtration and the fully reduced peptide was air oxidised by vigorous stirring. Air oxidation is catalysed by traces of metals, such as copper(II) and iron(III), present in the solution and the reaction requires the ionised thiolate species (Cecil *et al.* 1959) (Scheme 3.3).



Scheme 3.3: Air oxidation of the peptide. The reaction is catalysed by traces of metals present in the solution.

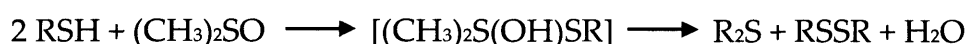
The concentration of free sulfhydryl groups was checked with 5,5'-dithio-bis-(2-nitrobenzoic acid) (DTNB or Ellman's reagent). DTNB reacts specifically to -SH groups at neutral and basic pH and the reaction produces 2-nitro-5-thiobenzoic acid (TNB) (Scheme 3.4), which can be easily detected by UV-VIS spectroscopy at 412 nm.



Scheme 3.4: Ellman's reagent reaction. The product of the reaction can be easily detected by UV-vis spectroscopy.

The timeline of oxidation of disulfide was checked by taking a sample of the peptide solution every 12 h and subjecting it to Ellman's reagent test. Typically a 250 μ L of peptide solution was mixed with a 50 μ L of DTNB 4 mg/mL stock solution in phosphate buffer pH 8.0. Total reaction volume was 2.5 mL. The reaction was incubated at room temperature for 15 min before the UV absorbance at 412 nm was measured. The UV readings indicated that even after 48 h there were still approximately 30% of the sulfhydryl groups

(concentration of $-SH = 300 \mu M$) present. Therefore DMSO was added to the reaction mixture to give a final concentration of 5% v/v to facilitate oxidation. DMSO as a mild oxidizing agent for thiols produces water and dimethylsulfide as by-products from the oxidation (Wallace 1964; Wallace *et al.* 1964).



DMSO is miscible with water and allows the oxidation reaction to be performed over a broad pH range. The reaction has been found to be up to forty times faster than simple air oxidation (Tam *et al.* 1991). In the case of Apoxaldie-Dab, upon addition of DMSO, the oxidation was complete in 2 h (concentration of sulfhydryl groups below Ellman's reagent detection limit) and the oxidised peptide was purified by RP-HPLC with a Luna C18 column (5 μm , 150 x 10 mm) using a linear gradient 60-90% aqueous acetonitrile containing 0.05% TFA over 30 min, flow rate 5 mL min⁻¹. The chromatogram shows the presence of 2 isomers (Figure 3.4). MALDI-TOF mass spectrometry confirmed both have the correct mass 1891, as the calculated mass for the oxidised peptide is 1891 (Figure 3.5).

The overlapping of the peaks made the complete separation of the two isomers impossible. In order to improve the separation a flatter gradient was used 70-80% aqueous acetonitrile and 0.05% TFA over 40 min but the separation did not improve. Another oxidation strategy, using reduced glutathione (GSH) mixed in a 10:1 ratio with its oxidized form (GSSG) as redox buffer, was then used. Thiols have been used to help renaturation of proteins; the presence of traces of metals help the reaction but the sulfhydryl-disulfide

exchange can be the only method of re-oxidation (Scheme 3.5) (Wetlaufer *et al.* 1970).

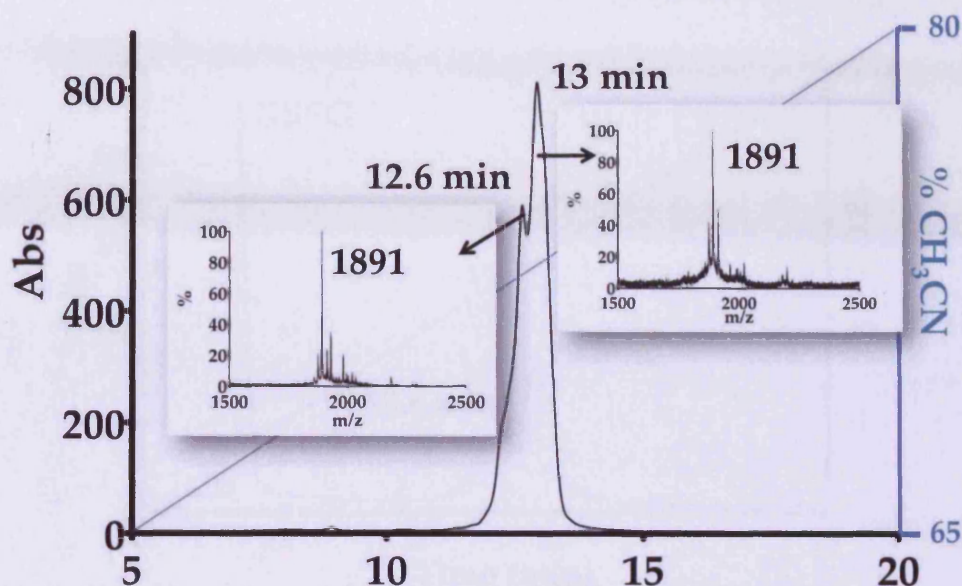
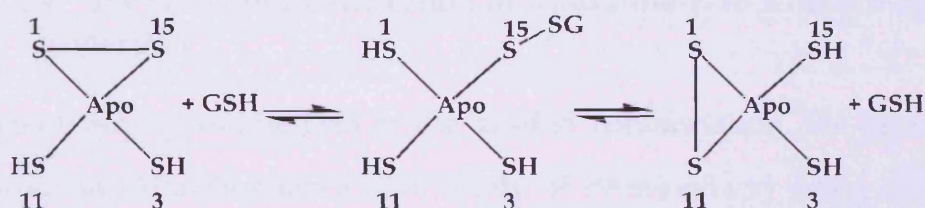


Figure 3.5: RP-HPLC trace of air oxidised Apoxaldie-Dab mixture recorded using a linear gradient of 60-90% *v/v* aqueous acetonitrile (containing 0.05% TFA) over 30 min at 5 mL min⁻¹. The two peaks contained the peptide in the oxidised form as confirmed by MALDI mass spectrometry (inserts).



Scheme 3.5: The use of GSH-GSSG redox buffer helps renaturation of the proteins.

The peptide (final concentration 250 μ M) was dissolved in phosphate buffer (pH 8.0) containing GSH/GSSG (4/0.4 mM) with constant stirring for 24 h. The oxidised peptide was purified by RP-HPLC with a Luna C18 column (5

μm , 150 x 10 mm) using the flatter gradient 70-80% aqueous acetonitrile and 0.05% TFA over 40 min. However no appreciable improvement in the separation was achieved (Figure 3.6).

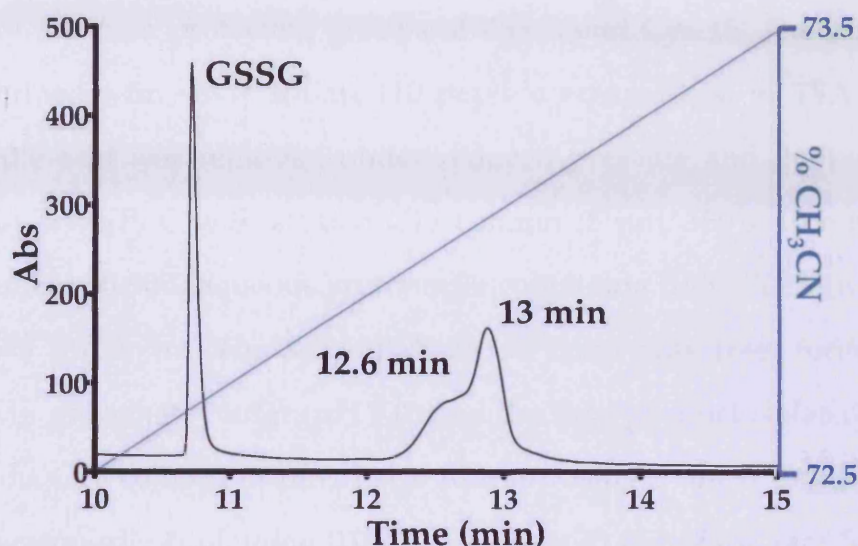


Figure 3.6: RP-HPLC trace after use of redox buffer using a linear gradient of 70-80% v/v aqueous acetonitrile (containing 0.05% TFA) over 40 min at 5 mL min⁻¹. The separation of the two isomers didn't improve.

3.2.4 Synthesis and purification of Apoxaldie-Dab with orthogonal protection

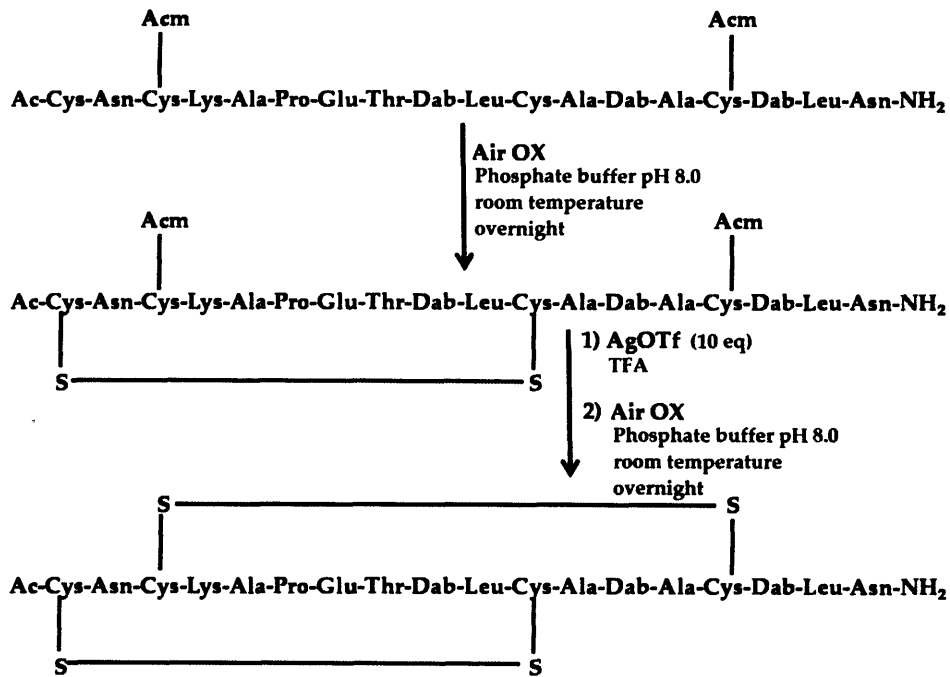
To obtain Apoxaldie-Dab in the desired conformation, the peptide was synthesized by Alta Bioscience (University of Birmingham) using orthogonal triphenylmethyl (Trt) (Cys 1 and 11) and acetamidomethyl (Acm) (Cys 3 and 15) protecting groups.

The peptide was received fully deprotected except for the Cys 3 and 15 which were protected with Acm (Scheme 3.6). The peptide (final concentration 250 μM) was dissolved in phosphate buffer (pH 8.0) and air oxidised to form the first disulfide bond between Cys 1 and Cys 11 overnight with constant

stirring. The completion of the reaction was monitored with Ellman's reagent (sulfhydryl concentration below detectable limit). The peptide was purified by RP-HPLC with a Luna C18 column (5 μm , 150 x 10 mm) using a linear gradient 60-90% aqueous acetonitrile containing 0.05% TFA over 30 min, flow rate 5 mL min^{-1} . The purified peptide with 1 disulfide bond was then subjected to the removal of the Ac protecting groups of Cys 3 and Cys 15. The deprotection was performed with silver triflate (10 peptide equivalents) in TFA. After the reaction, the acid was removed under reduced pressure and the peptide was purified by RP-HPLC with a Luna C18 column (5 μm , 150 x 10 mm) using a linear gradient 60-90% aqueous acetonitrile containing 0.05% TFA over 30 min, flow rate 5 mL min^{-1} . The second disulfide bond was then formed by air oxidation in phosphate buffer (pH 8.0) and the final product isolated by HPLC with a Luna C18 column (5 μm , 150 x 10 mm) using a linear gradient 60-90% aqueous acetonitrile containing 0.05% TFA over 30 min, flow rate 5 mL min^{-1} . The mass of the products was confirmed by MALDI-TOF-MS.

The purity of the purified peptide was assessed by RP-HPLC Luna C18 analytical column (5 μm , 150 x 5 mm) using a linear gradient (60-90% aqueous acetonitrile containing 0.05% TFA over 30 min, flow rate 1 mL min^{-1}). The peptide eluted as a single peak (Figure 3.7).





Scheme 3.6: Synthesis of Apoxaldie-Dab using AcM protecting group for Cys 3 and 15.

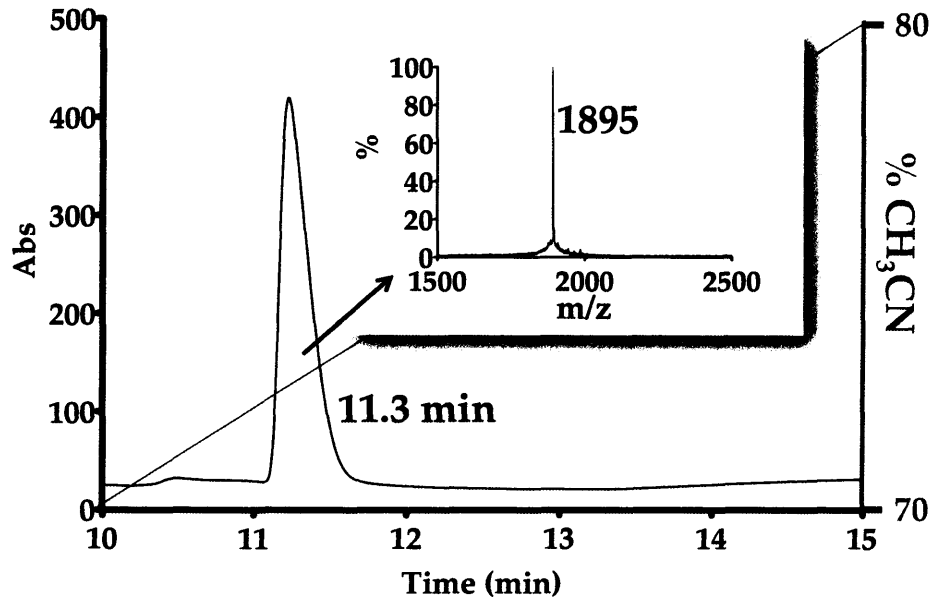


Figure 3.7: RP-HPLC trace of Apoxaldie-Dab on an analytical column using a linear gradient (60-90% *v/v* aqueous acetonitrile containing 0.05% TFA over 30 min). The peptide eluted as single peak.

3.2.5 Synthesis and purification of Apoxaldie

Apoxaldie was synthesised as a positive control using standard Fmoc chemistry (Jones 1992) using a microwave assisted peptide synthesiser (Scheme 3.2). The peptide was N-terminally acetylated on resin with acetic anhydride and C-terminally amidated to give the final sequence



The peptide was purified by RP-HPLC with a Luna C18 column (5 μm , 150 x 10 mm) using a linear gradient 60-90% aqueous acetonitrile containing 0.05% TFA over 30 min, flow rate 5 mL min⁻¹ and was found by MALDI-TOF mass spectrometry to have a mass of 1938 Da, which is in agreement with the calculated value of 1937 Da.

3.2.6 Oxidation of Apoxaldie

The purified peptide (0.5 mM) was treated with immobilised TCEP to assure that the cysteines were present in the reduced form. The fully reduced peptide was then air oxidised at pH 8.0. Testing with Ellman's reagent indicated that the reaction was complete after 24 h and the oxidised peptide was purified by RP-HPLC with a Luna C18 column (5 μm , 150 x 10 mm) using a linear gradient 60-90% aqueous acetonitrile containing 0.05% TFA over 30 min, flow rate 5 mL min⁻¹. The HPLC chromatogram showed the presence of a single peak and the mass of the eluted, 1933 Da, for the oxidised form was confirmed by MALDI-TOF mass spectrometry (Figure 3.8).

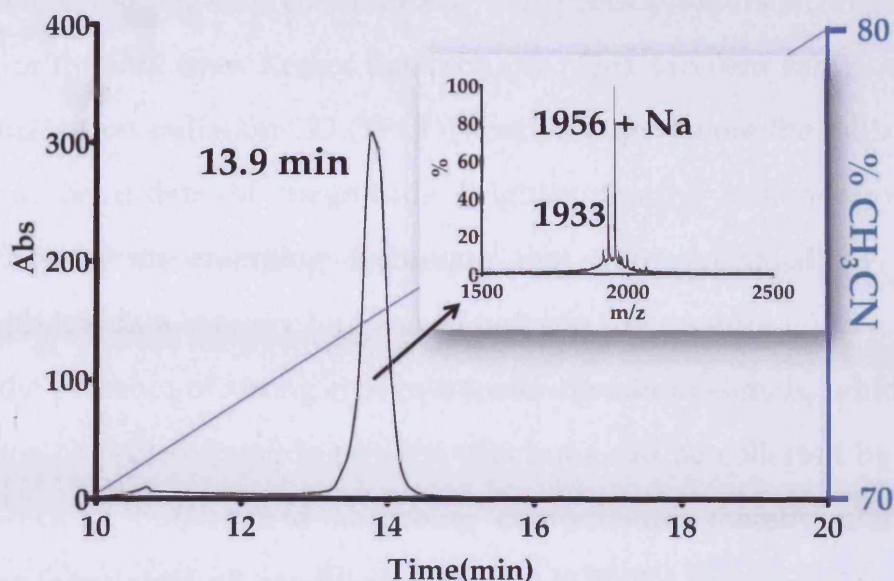


Figure 3.8: RP-HPLC trace using a linear gradient (60-90% *v/v* aqueous acetonitrile containing 0.05% TFA over 30 min) of Apoxaldie in its fully oxidised form. The peptide eluted as a single peak and its mass confirmed by MALDI mass spectrometry (insert).

3.2.7 Circular dichroism spectroscopy

Circular dichroism spectroscopy measures the difference in absorption for left and right circularly polarized components of plane polarized radiation by optically active molecules. The intensity of the CD signal is zero unless the chromophore is optically active either intrinsically or because it is in an asymmetric environment (Kelly *et al.* 1997). Polypeptides and proteins form structures that are intrinsically chiral due to the exclusive presence of L-amino acids and CD spectroscopy has been used extensively to study their secondary structure.

The far UV region from 250 to 180 nm (transitions $n \rightarrow \pi^*$ (210 nm) and $\pi \rightarrow \pi^*$ (*ca.* 190 nm) in the amide bond) carries most of the information on the secondary structure (Figure 3.9), particularly the region below 190 nm. Unfortunately in biological systems it is difficult to acquire data with high

quality below 190 nm with conventional CD spectrophotometer due to loss of intensity of the flux from Xenon light sources. This problem can be overcome using synchrotron radiation CD (SRCD) spectroscopy where the radiation light source can be orders of magnitude brighter than a Xenon lamp. SRCD spectroscopy is an emerging technique that allows acquisition of lower wavelength UV data. At very low wavelengths in the vacuum UV region below 190 nm, the presence of strong charge-transfer transition signals, which contain information of protein/peptide tertiary structures can be collected by SRCD in the presence of buffers and absorbing components, thereby more closely mimicking 'physiological' conditions (Miles *et al.* 2006).

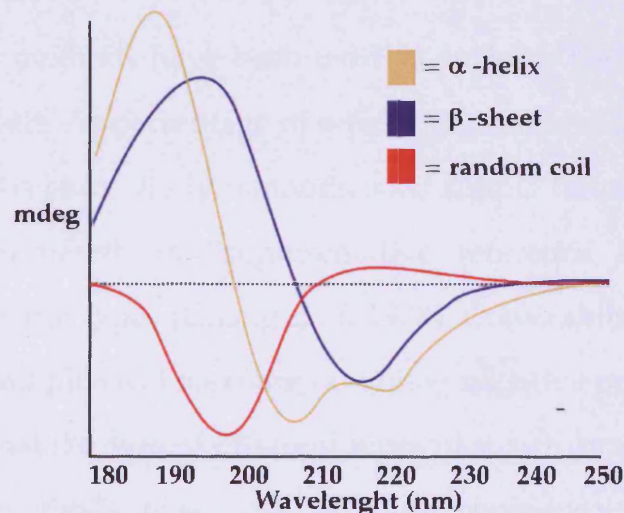


Figure 3.9: CD spectra of different protein secondary structures. In yellow α -helix; in blue β -sheet; and in red random coil. Picture from <http://www.proteinchemist.com/cd/cdspec.html>.

The CD of a protein can be measured as the mean ellipticity per residue, θ , and the helicity (fraction of residues in a helix) is often estimate from the mean ellipticity per residue at a wavelength of 222 nm $[\theta]_{222}$ as helical structures the ellipticity has a negative maximum at 222 nm. These simple estimates of

helicity assume that each residue in a helix contributes equally to $[\theta]_{222}$ (Yang *et al.* 1986). Theoretical calculation have suggested that the ellipticity, $[\theta]_{222}$, is sensitive to both the length of the helix and the conformation of the residues in the helix (Gans *et al.* 1991; Manning *et al.* 1991). The length dependence makes $[\theta]_{222}$ a linear function of the number of residues in a helix and the number of helices. Thus for the same number of residues in helices $[\theta]_{222}$ may differ depending on the number of helices that these residues are distributed over. This obscures the simple proportionality between helix content and CD signal at 222 nm (Hirst *et al.* 1994). Manning *et al.* calculated the CD spectrum of helices with residues in different conformation and found that $[\theta]_{222}$ may vary by up to a factor of three for helices ranging from hydrophilic ($\phi = -67^\circ$, $\psi = -44^\circ$) to hydrophobic ($\phi = -48^\circ$, $\psi = -57^\circ$) (Manning *et al.* 1991).

Different methods have been used to perform deconvolution of the CD spectra to calculate the percentage of α -helix, β -sheet and random coil present in the secondary structure. Early methods used simple linear and non-linear least square analyses based on “representative reference spectra” of different secondary structural types (Chang *et al.* 1978). Constraints were used to ensure that the result had physical meaning (avoiding negative percentages of one type structure) and that the sum of different types of structure was one (Wallace *et al.* 1987). Recent methods make use of more sophisticated methods such as parameterised fits, self-consistency and neural networks and have been used routinely for more than twenty years now (Whitmore *et al.* 2008). Generally, deconvolution gives the best results for α -helices because of their regular structure and the stronger CD signals produced in comparison to β -sheets and random coils. Another important factor for the success of the analysis is the quality of the reference spectra (how well they resemble the structure analysed *e.g.* the use of spectra of peptides as references for the deconvolution of the

spectrum of a peptide will give more reliable results than the use of spectra of molten globular proteins) (Whitmore *et al.* 2008). The program CDNN, the algorithm of which is based on a back-propagating neural network model, is frequently used to evaluate CD spectra from globular proteins (Böhm 1997). Poschner *et al.* implemented the program with a series of 20 artificial reference spectra of peptides (calculated with PEPFIT (Reed *et al.* 1997)) whose secondary structure is composed of different percentages of α -helix, β -sheet (parallel and antiparallel), β -turn and random coil (Poschner *et al.* 2007). This allows the program CDNN-PEPFIT, used in this work, to evaluate more precisely the secondary structure of small peptides.

The CD spectrum of Apoxaldie-Dab (Figure 3.10) shows a minimum at 205 nm and a shoulder at 222 nm as seen for apamin (Miroshnikov *et al.* 1978) and Apoxaldie (Weston *et al.* 2004). The deconvolution of the spectrum however shows that the percent of α -helix is smaller in Apoxaldie-Dab than in Apoxaldie (Table 3.1). Weston *et al.* detected a similar reduction in helix content in Apoxaldie compared to Apamin.

The proposed reason of the decreased helicity was either a distortion of the C-terminal end of Apoxaldie or the presence of 3_{10} -helix. The first hypothesis was supported by molecular dynamic simulations which suggested an unravelling of the last three N-terminal residues (Weston *et al.* 2004). The distortion of the end of the helix should be more important in Apoxaldie-Dab due the lower helical propensity of Dab compared to Lys (Hatano *et al.* 1970; Padmanabhan *et al.* 1996). The difficulties in the oxidation reaction (see above) suggest that a similar distortion is present at the beginning of the helix and this can prevent the peptide from folding into the native Apamin conformation. The mixture of isomers produced using the earlier synthesis and step-oxidation strategy (see above) also showed the presence of α -helical secondary structural

elements (Figure 3.9, non-native fold), but the CD spectrum is clearly different to that of the purified correct isomer produced using the step-oxidation strategy.

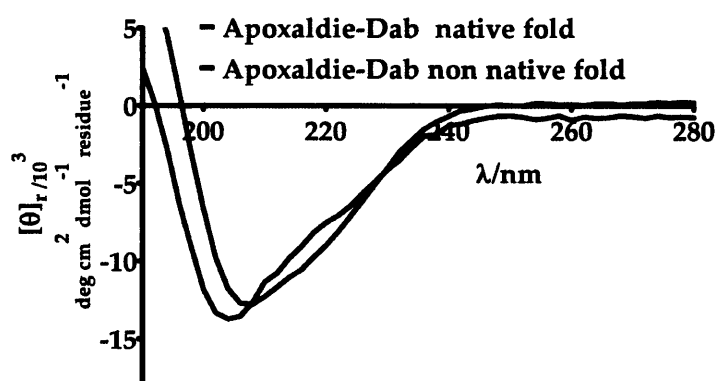


Figure 3.10: CD spectra of oxidised Apoxaldie-Dab isomers mixture (25 μM) from the direct oxidation strategy in potassium phosphate buffer (10 mM, pH 7). The black line indicated the native oxidised form of Apoxaldie-Dab and the blue line the non native oxidised form.

In order to increase the helicity of the peptide, which is essential for the right orientation of the amino side chain groups (Weston *et al.* 2004) and therefore the catalytic efficiency, 2,2,2-trifluoroethanol was used. Trifluoroethanol is a co-solvent known to increase the helix propensity of peptides and proteins (Goodman *et al.* 1969). Two mechanisms have been proposed to explain this effect: a) direct binding of trifluoroethanol (Jasanoff *et al.* 1994); b) stabilisation of the H-bonds in the helix *via* weakening of the hydrogen bonds between water and peptide backbones in the coil form (Cammers-Goodwin *et al.* 1996). The second mechanism has been supported by molecular calculations which suggested that the stabilizing effect of TFE is induced by the preferential aggregation of TFE molecules around the peptides. This coating displaces water, thereby removing alternative hydrogen-bonding

elements (Figure 3.9, non-native fold), but the CD spectrum is clearly different to that of the purified correct isomer produced using the step-oxidation strategy.

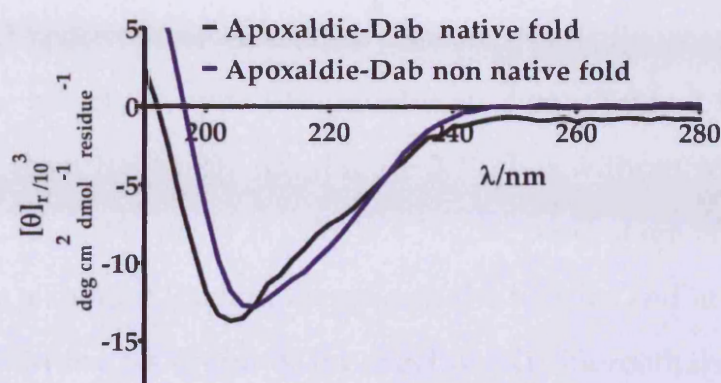


Figure 3.10: CD spectra of oxidised Apoxaldie-Dab isomers mixture (25 μ M) from the direct oxidation strategy in potassium phosphate buffer (10 mM, pH 7). The black line indicated the native oxidised form of Apoxaldie-Dab and the blue line the non native oxidised form.

In order to increase the helicity of the peptide, which is essential for the right orientation of the amino side chain groups (Weston *et al.* 2004) and therefore the catalytic efficiency, 2,2,2-trifluoroethanol was used. Trifluoroethanol is a co-solvent known to increase the helix propensity of peptides and proteins (Goodman *et al.* 1969). Two mechanisms have been proposed to explain this effect: a) direct binding of trifluoroethanol (Jasanoff *et al.* 1994); b) stabilisation of the H-bonds in the helix *via* weakening of the hydrogen bonds between water and peptide backbones in the coil form (Cammers-Goodwin *et al.* 1996). The second mechanism has been supported by molecular calculations which suggested that the stabilizing effect of TFE is induced by the preferential aggregation of TFE molecules around the peptides. This coating displaces water, thereby removing alternative hydrogen-bonding

partners and providing a low dielectric environment that favours the formation of intrapeptide hydrogen bonds. Because TFE interacts only weakly with nonpolar residues, hydrophobic interactions within the peptides are not disrupted. As a consequence, TFE promotes stability rather than inducing denaturation (Roccatano *et al.* 2002).

The CD spectrum of Apoxaldie-Dab in 15% trifluoroethanol shows an increase in the helicity content (deconvolution data) (Table 3.1) with a shift of the minimum from 205 to 207 nm (Figure 3.11) but without reaching the same level as Apoxaldie (Weston *et al.* 2004). A higher percentage of trifluoroethanol usually shows a smaller fraction increase in the helicity and at the same time a drastic change in the condition of the reaction. Trifluoroethanol (pK_a 12.4) is a stronger acid than water (pK_a 15.7) with a lower dielectric constant ($\epsilon_{TFE} = 26.1$, $\epsilon_{water} = 76.4$ for water) and it has been hypothesized that the condition with TFE become more acidic than the apparent pH (Cammers-Goodwin *et al.* 1996). For this reason no further addition of trifluoroethanol was conducted.

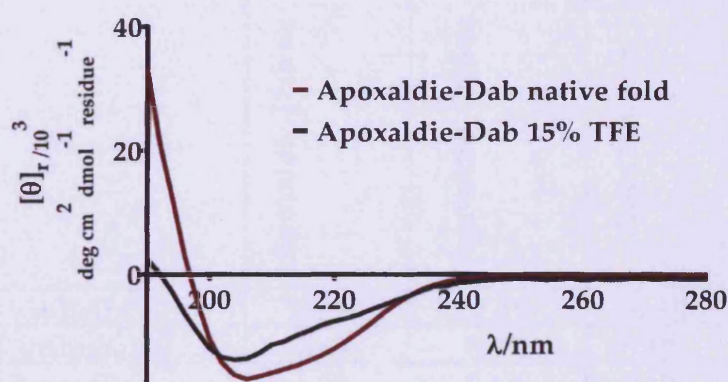


Figure 3.11: CD spectra of Apoxaldie-Dab native fold and in presence of TFE in potassium phosphate buffer (10 mM, pH 7). The black line indicated the oxidised Apoxaldie-Dab (25 μ M) and the red line the same peptide in presence of TFA (15%).

The spectrum of Apoxaldie shows a minimum at 206 nm and a shoulder at 222 nm (Figure 3.12). The deconvolution of the spectra estimated an α -helical content of 38% in good agreement with the published 35% (Weston *et al.* 2004).

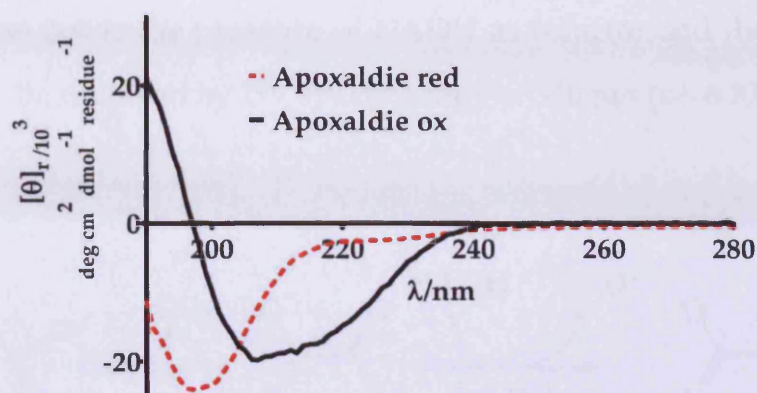


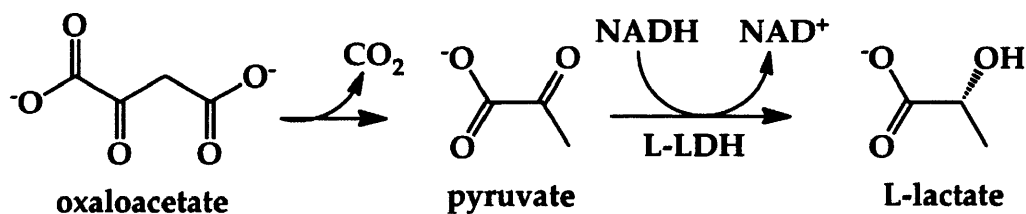
Figure 3.12: CD spectrum of Apoxaldie (25 μ M) in potassium phosphate buffer (10 mM, pH 7). The black line indicated the oxidised Apoxaldie and the red dotted line the reduced form.

Table 3.1: Deconvolution of the CD spectra using the program CDNN-PEPFIT.
*Apamine value from NMR data (Pease *et al.* 1988)

	Apoxaldie-Dab ox	Apoxaldie -Dab ox 15% TFA	Apoxaldie ox	Apamine*
α -helix	26.3%	32.4%	36.5%	~50%
β -antiparallel	2.6%	6.3%	6.1%	-
β -parallel	0.0%	0.1%	0.1%	-
β -turn	1.0%	0.6%	0.8%	-
Random coil	74.9%	63.0%	57.4%	~50%
Total Sum	104.8%	102.4%	100.9%	

3.2.8 Kinetics study

An enzyme-coupled assay was used to study the peptide-catalysed decarboxylation of oxaloacetate (Johnsson *et al.* 1993; Taylor *et al.* 2001; Weston *et al.* 2004). The assay uses lactate dehydrogenase to convert the pyruvate produced by the decarboxylation reaction into lactate (Scheme 3.7). Lactate dehydrogenase needs the presence of NADH as cofactor and the consumption of NADH can be followed by UV spectroscopy at 340 nm ($\epsilon = 6.23 \times 10^3 \text{ M}^{-1} \text{ cm}^{-1}$).



Scheme 3.7: The conversion of NADH to NAD⁺, during the reduction of pyruvate to lactate, was used to monitor the decarboxylation of oxaloacetate to form pyruvate and carbon dioxide.

The decarboxylation of oxaloacetate was followed at 25 °C in *N,N*-bis(2-hydroxyethyl)-2-aminoethanesulfonic acid (BES) (pH 7). The concentration of the peptides was kept constant (10 μM) whereas the concentration of oxaloacetate varied from 5 mM to 200 mM.

The catalytic activity of Apoxaldie-Dab was however not able to be detected using this assay. A ten-fold increase in its concentration did not improve the readout of the assay. This can be due to the low α -helical content in Apoxaldie-Dab structure compare to Apoxaldie, 26% against 36% (Table 3.1). A helicity of 26% means that the α -helix is only 5 residues in length, probably in the region between the two disulfide bonds from Cys11 to Cys 15. In this case

Dab9 and Dab16 side chains will be not aligned with the side chain of Dab13 (as they are in Apamin Figure 3.13). As a consequence the pK_a of the side chain amino groups of Dab residues cannot be modulated by their mutual interaction.

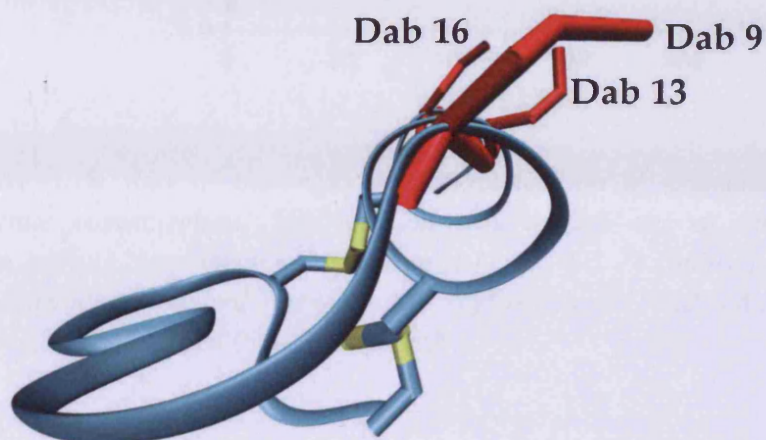


Figure 3.13: Alignment of Dab residues in Apoxaldie-Dab in a model based on Apamin crystal structure.

The catalytic reaction using Apoxaldie-Dab was then tested in the presence of 15% 2,2,2-trifluoroethanol (to improve the helicity (Figure 3.9 and Table 3.1)) in BES (pH 7). While the peptide concentration constant (10 μ M), oxaloacetate concentration was increased from 5 mM to 200 mM. Again no appreciable activity was detected, probably because the denaturation of lactate dehydrogenase when the concentration of TFA is higher than 5%, consequently the catalytic reaction didn't take place (Allemann 1989).

As a positive control the experiment was repeated with Apoxaldie using the same conditions. The peptide showed saturation kinetics with $k_{cat} = 0.08 \pm 0.02 \text{ s}^{-1}$ and $K_M = 32 \pm 8$ (Figure 3.14) in agreement with the published data ($k_{cat} = 0.07 \pm 0.02$ and $K_M = 30 \pm 5$ (Weston *et al.* 2004).

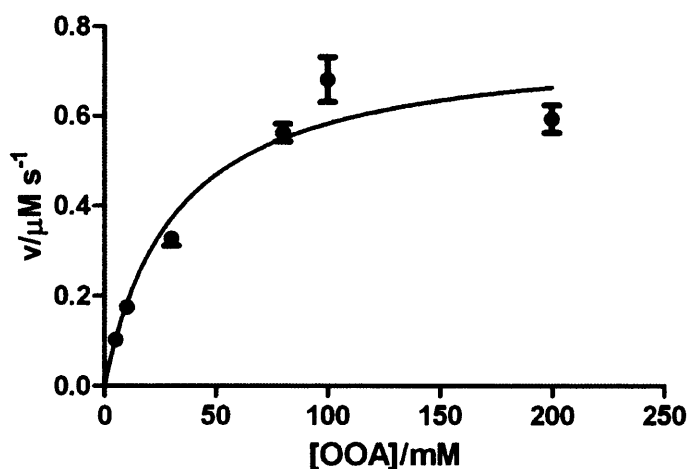


Figure 3.14: Rate of the catalysed decarboxylation of oxaloacetate as a function of substrate concentration. The reaction was carried out at 293 K in *N,N*-bis-(2-hydroxyethyl)-2-aminoethanesulfonic acid (1 M, pH 7) containing 10 mM NaCl. The concentration of oxidised Apoxaldie was kept constant (10 μ M) and the concentration of oxaloacetate was varied (5 mM to 200 mM).

3.3 Conclusion

A new putative oxaloacetate decarboxylase, Apoxaldie-Dab, was designed by replacing the three catalytic lysine residues of Apoxaldie with L-2,4-diaminobutyric acid (Dab) residues. Production of Apoxaldie-Dab was more complicated than that of the parent Apoxaldie due to the formation of two isomers during the oxidation reaction required to form the two disulfide bridges between four cysteines. The desired isomer was obtained using orthogonally protected cysteine pairs. The formation of incorrect disulfide bridges was therefore successfully prevented allowing the production of Apoxaldie-Dab. However, Apoxaldie-Dab showed no appreciable catalytic activity, whereas the parent Apoxaldie was an efficient oxaloacetate decarboxylase as described previously [Section 3.5.2 and (Weston *et al.* 2004)].

The helix breaking activity of Dab is likely to be one of the main causes for the lack of activity of Apoxaldie-Dab. The distortion of the helix increased as

the helicity decreased, shown by the deconvolution of the CD spectra that Apoxaldie contains approximately 36% helix whereas Apoxaldie-Dab contains only 26%. The two cysteine residues in the helix are one turn of helix apart (Figure 3.1) and this makes its C-terminal region subject to distortion. Such distortion, which is also present in Apoxaldie, although to less extent (Weston *et al.* 2004), may prevent the interaction between the amino group of the Dab residues necessary to modulate their pK_a therefore increase their activity in the decarboxylation reaction. If the pK_a is not modulated as desired, the protonated side-chains of the Dab residues will be unable to carry out their intended catalytic functions. In order to understand better the conformation assumed by Apoxaldie-Dab, a molecular modelling study in conjunction with NMR and X-ray spectroscopy will be necessary. The use of ornithine instead of Dab may solve the problem of strong helix distortion when apamin is used as a scaffold.

4. IM9

4.1 Introduction

4.1.1 Artificial enzymes

The design of new enzymes able to catalyse a chosen reaction is one of the ultimate goals of protein engineering. The task is particularly challenging due to the difficulty of predicting the folded structure of a new designed protein from its primary structure (Section 1.7). The use of natural proteins as “blue prints” is a way to overcome this problem. Small peptides have been used as scaffolds on which to graft an active site with good results (Section 1.11), but the exposure of the catalytic residues to the solvent limits the specificity of the reactions. Catalytic antibodies as artificial enzymes, have given good results but the difficulties in developing improved transition-state analogues and in the refinements of immunization and screening protocols limit the possibility of further improvement of their activities (Section 1.17).

The use of small proteins as scaffolds has several advantages. Proteins are easily expressed and purified and they are big enough to accommodate an active site in a cavity shielded by the solvent.

4.1.2 Im9

Im9 is a monomeric 86-amino acid protein (MW 9580) (Figure 4.1) which inhibits the DNase activity of colicin E9 (James *et al.* 1993). Im9 amino acids sequence is:

```
MELKHSISDYTEAEFLQLVTTICNADTSSEEELVKLVTH
FEEMTEHPSGSDLIYYPKEGDDDSPSGIVNTVKQWRAA
NGKSGFKQG
```

Its structure consists of a distorted four-helix bundle without a prosthetic group or disulfide bond.

Im9 has been intensively studied as a model for protein folding. It folds, like the majority of the small proteins (< 100 amino acid residues) (Cranz-Mileva *et al.* 2005), with a two state mechanism at pH 7 without an intermediate (Ferguson *et al.* 1999).

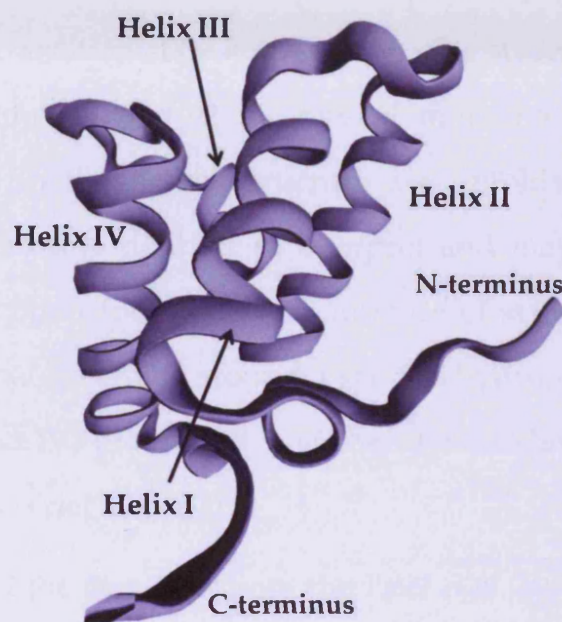


Figure 4.1: Ribbon diagram of the structure of Im9 illustrating its four-helical fold (Osborne *et al.* 1996).

Wallis *et al.* performed an alanine scan mutagenesis of 34 residues: C23; N24; D26; T27; S28; S29; E30; E31; E32; L33; V34; K35; L36; V37; T38; E41; E42; T44; E45; H46; P47; S48; G49; S50; D51; L52; I53; Y54; Y55; P56; D60; S63; V68; N69. These residues were chosen on the basis of a NMR analysis of the binding reaction with E9 DNase: the residues whose amides showed large or intermediate changes in chemical shift were mutated to alanine. The scan revealed that the mutants, with the exception of P47 and Y54, did not have

significantly altered solvent accessibility. The buried residues showed a greater effect on stability than the solvent exposed residues (Wallis *et al.* 1998).

Friel *et al.* used Φ analysis to study the effect of several mutations on the folding of Im9 (Friel *et al.* 2003). In Φ analysis, mutations are designed to delete or alter existing weak interactions. The parameter Φ ($= \Delta\Delta G^\ddagger/\Delta\Delta G_0$, where $\Delta\Delta G^\ddagger$ is the change of free energy of activation and $\Delta\Delta G_0$ the change in free energy of folding on mutation) is a measure of the average extent of structure formation at the mutated site on a scale of 0 to 1: $\Phi = 0$ implies that at the site of mutation the structure in the transition state and in the native structure are folded in the same way; $\Phi = 1$ implies that at the site of mutation the structure in the transition state and in the native structure are unfolded in the same way. Fractional values are more difficult to interpret and may imply that there is partial non-covalent bond formation or a mixture of states (Fersht *et al.* 1992). The study showed that Im9 folds around a specific hydrophobic core with three of the helix, (I, II, and IV) preformed, and that the rate-limiting transition state is robust to mutation (Friel *et al.* 2003).

Paci *et al.* used the Φ values from the Friel *et al.* study as restraints in an *in silico* study to assess the importance of each residue in the folding mechanism of Im9 (Paci *et al.* 2004). They used B analysis to determine which amino acid residues make the most important contribution to generating the small-world network (a network in which the connection topology, in this case between amino acid residues, is half way between completely regular or completely random (Watts *et al.* 1998)) of interaction that characterise a protein structure, with B_k defined as the number of pairs (i, j) of vertices such that the shortest path between i and j passes through k , normalised by the total number of pairs (Freeman 1977). Thus B measures the centrality of a residue in the network of interactions that characterise a protein structure and provides a correction to

the use of the number of contacts for describing the structural relevance of a residue; *i.e.* the key residues are not necessarily the residues with the largest number of contacts (Vendruscolo *et al.* 2002). In their study Paci *et al.* found that the residues F15, V19, I67, V71, K72, and R75 are important in determining the network of interactions that stabilise the transition state of Im9. The results of the B analysis were used to back calculate the Φ values for all the residues of Im9. The experimental values of Φ used as constrain in the B analysis and the calculated data were essentially identical. Several non-mutated residues (S6, I7, S8, Y10, Y54, Y55, W74, and R75) showed to be important in the stabilisation of the transition state.

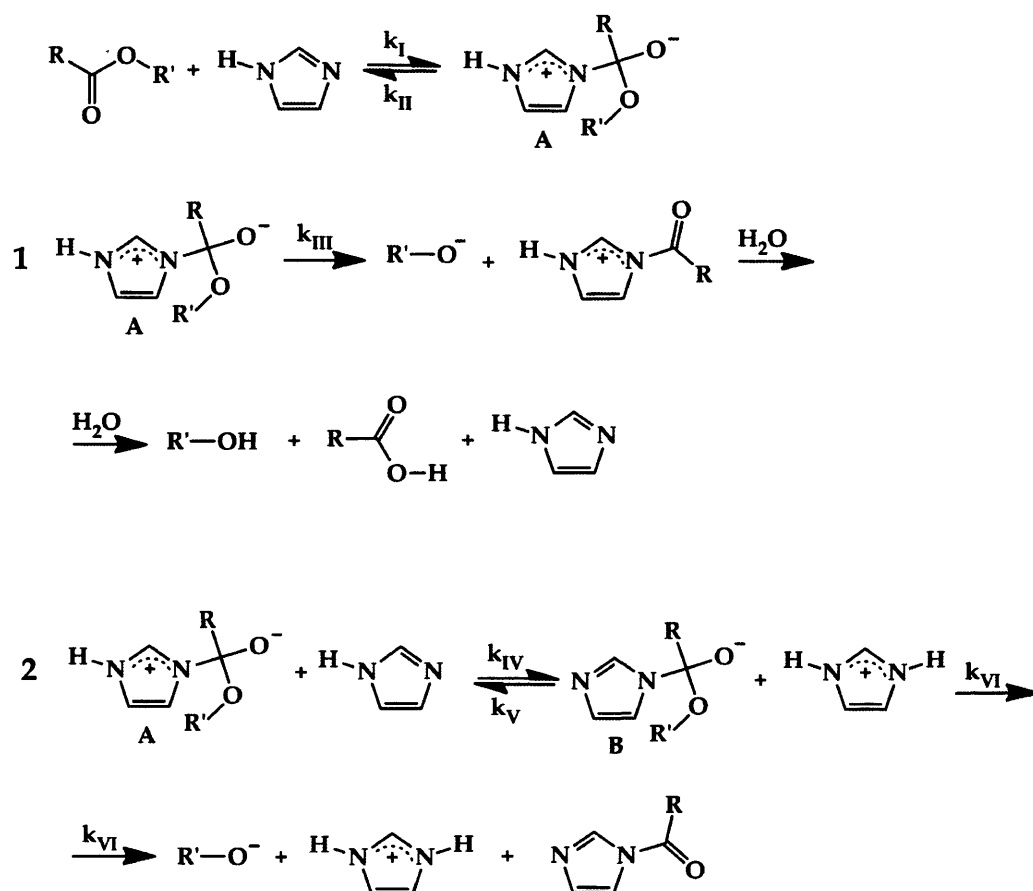
4.1.3 *p*-Nitrophenol hydrolysis as a model reaction

The hydrolysis of *p*-nitrophenyl acetate (PNPA) into *p*-nitrophenol (PNP) and acetate catalysed by a nucleophilic histidine was selected as the target reaction because it can easily be followed by UV-VIS spectroscopy and has a low activation barrier at neutral pH in aqueous solution.

The aminolysis of carboxylic acid esters has been extensively studied. In the past, several partly contradictory expressions of the reaction mechanism were proposed (Bender *et al.* 1957; Sacher *et al.* 1964; Menger 1966; Shawali *et al.* 1967; Anderson *et al.* 1969; Menger *et al.* 1969; Nakamizo 1969; Satchell *et al.* 1969; Satchell *et al.* 1970). Today the generally accepted mechanism of ester aminolysis consists of two parallel reaction channels (Scheme 4.1) (Schmeer *et al.* 1999).

The contributions of the two paths to the overall reaction are determined by the difference of the basicities of the imidazole and the leaving group in the respective solvent. Reaction channel 2 becomes detectable when the basicity of

the leaving group (*p*-nitrophenolate) is at least three units greater than the basicity of the attacking nucleophile. In water, the difference of the basicities is nearly zero and the formation of the intermediate A is therefore the rate-determining process ($k_{III} > k_I$ and $k_{III} > k_{II}$), and the resulting N-acetylimidazole from path 1 is hydrolysed in a fast reaction.



Scheme 4.1: Generally accepted mechanism of ester aminolysis.

In non-aqueous solvents (or in the hydrophobic pocket of an enzyme active site), the basicity of the imidazole is much lower than that of the *p*-nitrophenolate ion. A general base can, however, increase it by withdrawing the proton from the heterocycle. Thus, the decomposition of the intermediate A *via*

both reaction paths becomes rate determining: k_{III} , k_{IV} , and k_V are lower than all other rate constants (Schmeer *et al.* 1999).

In natural proteases nucleophile-catalysed hydrolysis is usually performed by a catalytic serine or cysteine. Histidine usually acts as a general acid or base, but can act as a nucleophile as in phosphate transfer reactions (Fersht 1984). However, a key requirement for a successful nucleophilic catalysis is a higher relative instability of the acylated enzyme intermediate relative to the substrate in buffer in order to increase the rate of product formation over the uncatalysed buffer rate (Fersht 1984; Bolon *et al.* 2001). An acyl-histidine intermediate would be less stable than an acyl-serine or an acyl-cysteine intermediate (Bolon *et al.* 2001), therefore, histidine was chosen as nucleophile in the design of the active site.

4.1.4 Aims

The aim of this part of the project was to produce an artificial enzyme using the colicin DNase immunity protein Im9. Its robustness to mutation and the fact that it can easily be expressed and purified made Im9 an ideal candidate. The protein will be used as a scaffold on which to graft an active site in the hydrophobic cleft between helices, using histidine residues as nucleophiles for the catalytic reaction.

In the design strategy, a bulky aromatic amino acid positioned in a hydrophobic cleft, would be mutated to alanine in order to create a cavity for substrate binding (Figure 4.2). Then a second amino acid would be mutated to histidine. The second active-site histidine would be provided by the native protein sequence. Mutations of nearby residues to tune the pKa values of the two histidines were also considered.

The new miniature enzymes would be characterised using mass spectrometry and CD spectroscopy. The kinetic activity of the designed miniature enzymes will be measured following the hydrolysis of *p*-nitrophenyl acetate by UV spectroscopy.

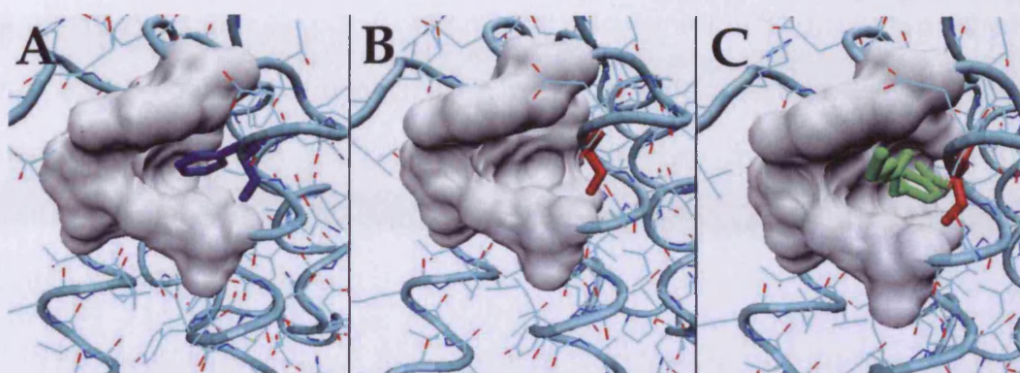


Figure 4.2: The mutation of a bulky aromatic amino acid (Phenylalanine in blue) situated in a hydrophobic cleft (A) to a small amino acid (Alanine in red) will create a cavity (B) for substrate binding (*p*-nitrophenyl acetate in green) (C).

4.2 Results and Discussion

4.2.1 Expression of Im9 wild type

The Im9 gene, sub-cloned into pET21d was obtained from Dr. Nicoll, previous member of the group. BL21(DE3) cells were used as the expression host. For large scale expression cells were grown in LB medium at 37 °C to an OD_{600nm} of 0.6 (typically 4 h after inoculation). Gene expression was induced with IPTG (0.5 mM) and cells were harvested 4 h after induction.

4.2.2 Purification of Im9

The published Im9 purification protocol involves a three-step procedure, an initial ammonium sulfate precipitation followed by DEAE ion-exchange chromatography and size exclusion chromatography (Wallis *et al.* 1992). To avoid the time-consuming size exclusion purification, a method involving two steps of anion exchange purification was developed. The first step involved a weak anion-exchange resin (DEAE) for group separation, and the second step applied a strong anion resin (SourceQ[®]) to realise the final purification. An ultra-filtration step was also included. This method was used to purify all the mutants as well.

The protein expressed overnight at 16 °C (Section 2.10.2) was present in the supernatant (Figure 4.3), confirmed by MALDI-TOF mass spectrometry. The supernatant was applied to a DEAE (anion exchange) column and eluted with a salt gradient (from 10 mM NaCl to 400 mM over 40 min and from 400 mM to 1 M NaCl over 10 min with a constant flow rate of 8 mL min⁻¹) (Figure 4.4).

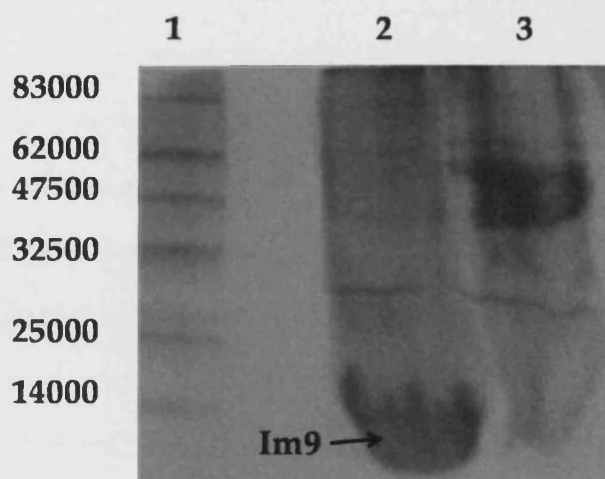


Figure 4.3: SDS-polyacrylamide gel electrophoresis showing cell lysate. Lane 1: broad range protein marker; lane 2: supernatant solution (containing the protein); lane 3: cell debris pellet.

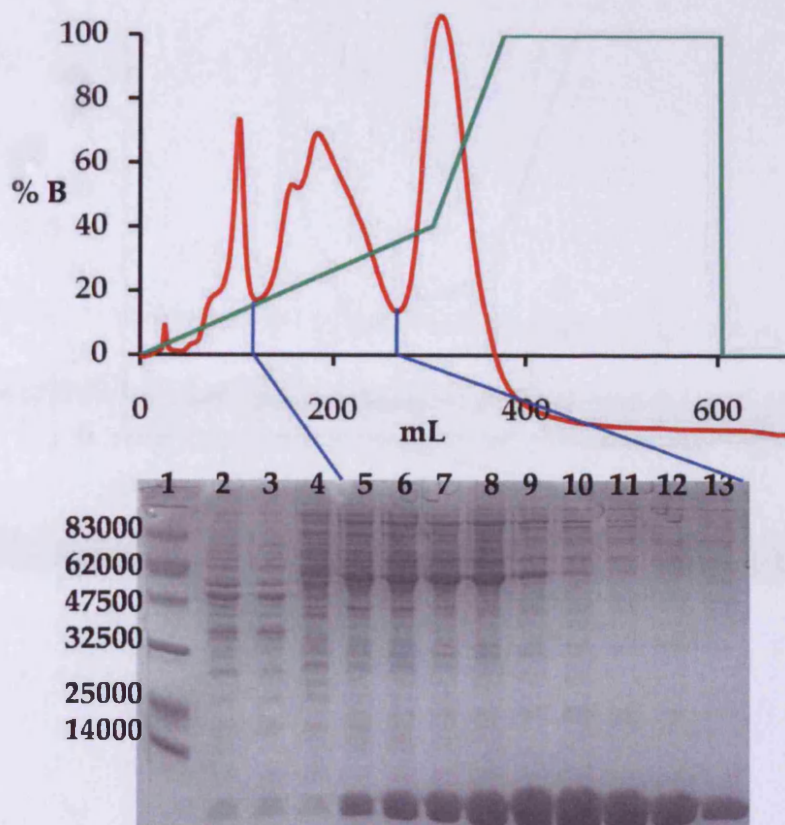


Figure 4.4: Top: Chromatogram of DEAE purification of Im9. The protein was eluted with a salt gradient (from 10 mM NaCl to 400 mM over 40 min and from 400 mM to 1 M NaCl over 10 min with a constant flow rate of 8 mL min⁻¹). The absorbance at 280 nm (red) and NaCl concentration (green) are shown. The peak containing Im9 is indicated between the blue bars. Bottom: 12% SDS-PAGE following DEAE purification; lane 1: broad range protein marker; lane 2-4 impurity; lane 5-13 fractions containing Im9.

The fractions containing Im9 were determined using SDS-polyacrylamide gel electrophoresis (Figure 4.4). Those fractions were then pooled and dialysed to remove the salt. The resulting solution was then applied onto a SourceQ[®] column and eluted with a salt gradient (from 10 mM NaCl to 400 mM over 12.5 min and from 400 mM to 1 M NaCl over 2.5 min with a constant flow rate of 8 mL min⁻¹) (Figure 4.5). SDS-PAGE was performed to identify the fractions containing Im9 (Figure 4.5). The purified protein was freeze-dried and stored at -20 °C.

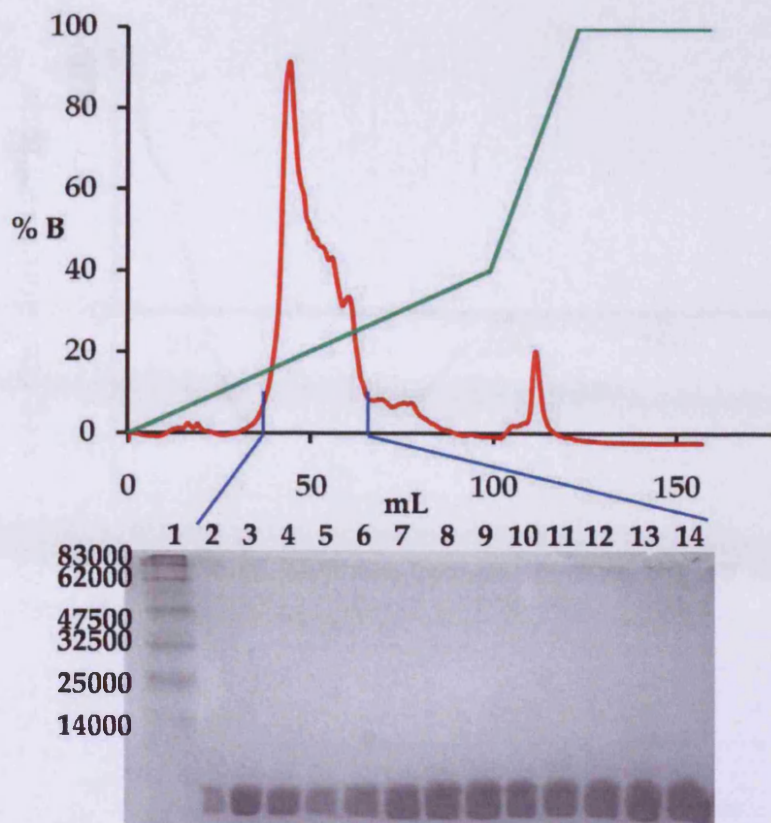


Figure 4.5: Top: Chromatogram of ResourceQ™ purification of Im9. The protein was eluted with a salt gradient (from 10 mM NaCl to 400 mM over 12.5 min and from 400 mM to 1 M NaCl over 2.5 min with a constant flow rate of 8 mL min⁻¹). The absorbance at 280 nm (red) and NaCl concentration (green) are shown. The peak containing Im9 is indicated between the blue bars. Bottom: 12% SDS-PAGE following ResourceQ purification; lane 1: broad range protein marker; lane 2-14 fractions containing Im9.

4.2.3 Characterisation of Im9

The mass of purified Im9WT was analysed *via* MALDI mass spectrometry giving a mass of 9583 (calculated mass 9582.5). The secondary structure and thermal stability of the protein were analysed using CD spectroscopy (Figure 4.6) and the spectra were deconvoluted with the program CDNN (Poschner *et al.* 2007). The protein showed, as expected, around 48% α -helical structure with a melting temperature of 56 °C (Figure 4.7) in good agreement with the literature data of 57 °C (Wallis *et al.* 1992).

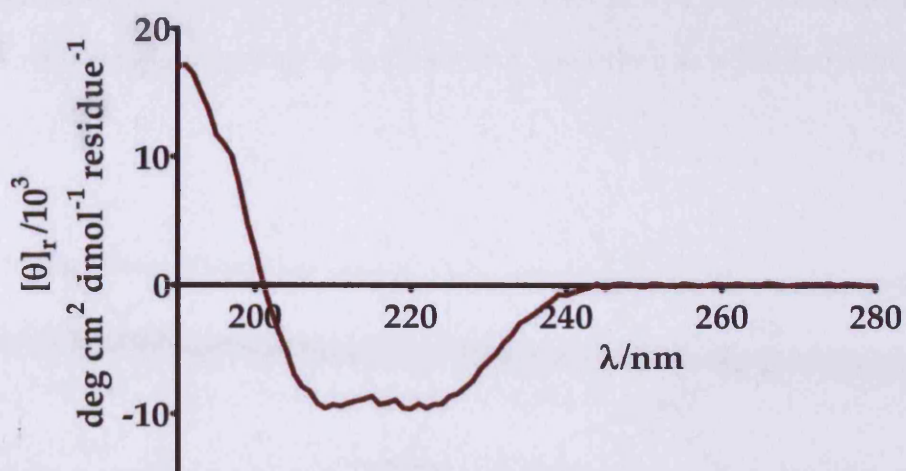


Figure 4.6: CD spectra of Im9. Spectrum recorded at 20 °C in 10 mM potassium phosphate buffer (pH 7.0).

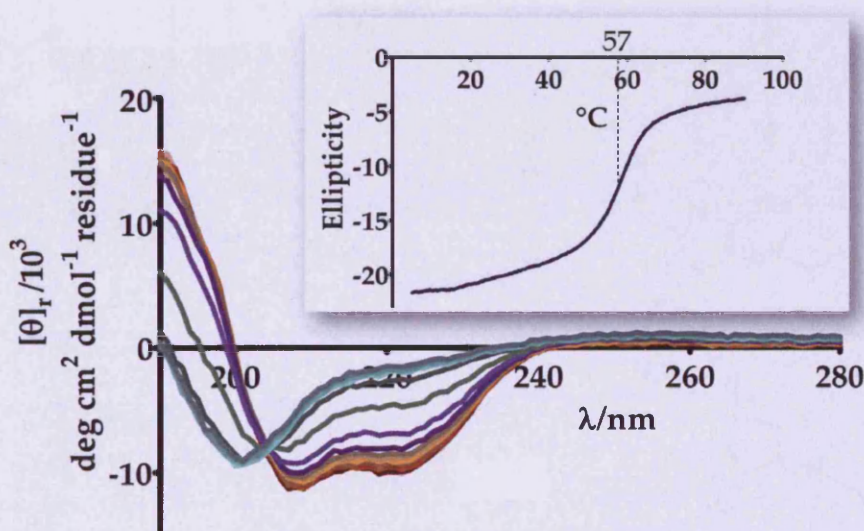
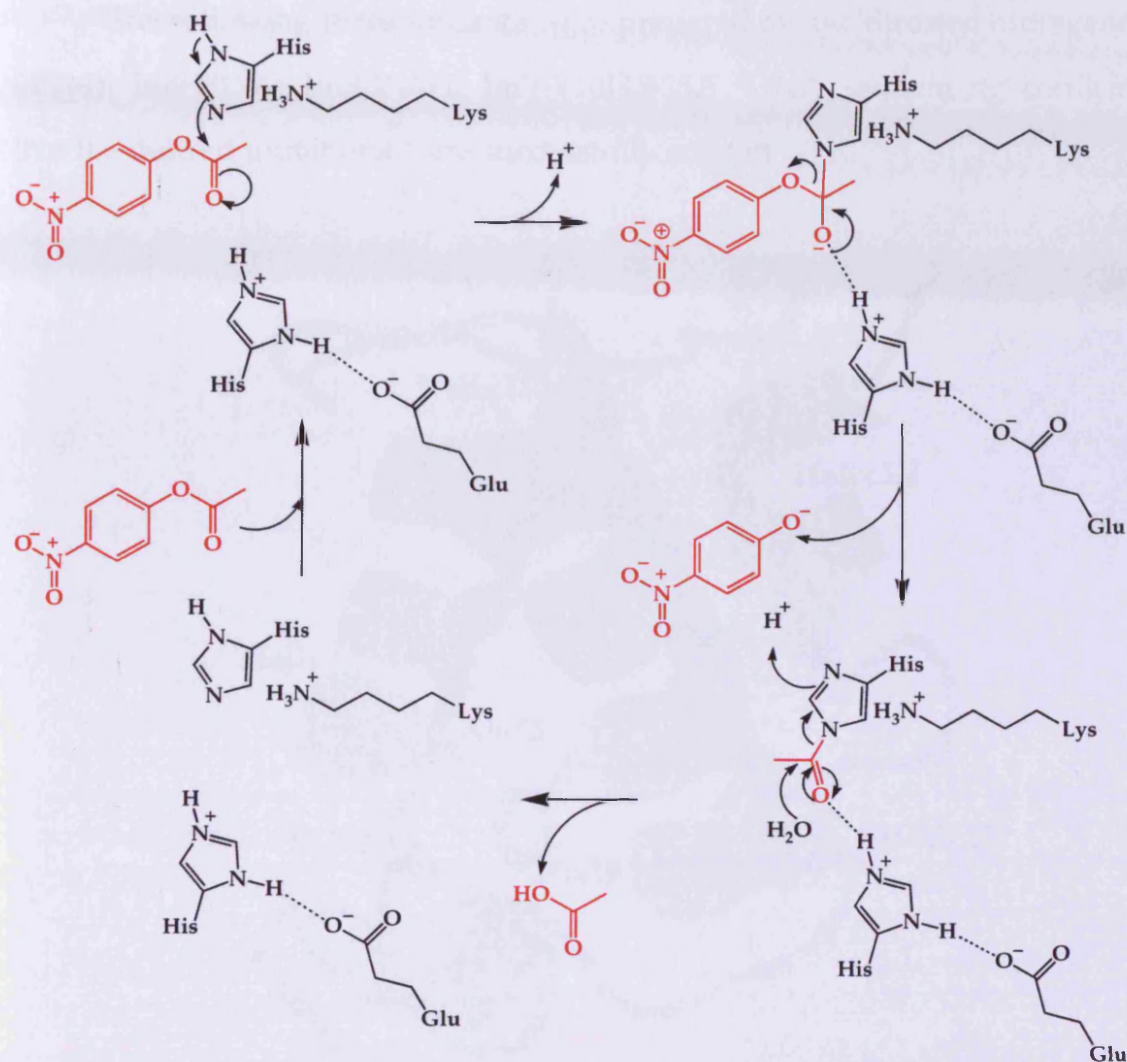


Figure 4.7: Effect of the temperature on Im9 structure across the UV range and followed at 222 nm (insert). Spectra recorded in 10 mM potassium phosphate buffer (pH 7.0).

4.2.4 Design of the active site

The NMR-derived structure of Im9 was used to find a relatively solvent shielded cavity in which to design an active site (Osborne *et al.* 1996). A semi-hydrophobic site might enhance the binding of *p*-nitrophenyl esters. A

mechanism-based design was used to sketch an active site based around two histidine residues, one acting as an acid and the other as a nucleophile (Scheme 4.2).



Scheme 4.2: Proposed mechanism of the hydrolysis of *p*-nitrophenyl acetate used to sketch the active site.

4.3 Design 1: Im9-Y10H/F15A series

In the first design, phenylalanine 15, situated at the beginning of helix I, was mutated to alanine to create the active site cavity, while the replacement of

tyrosine 10 with histidine generated a possible general acid within hydrogen bonding distance of the carbonyl of the substrate helping to stabilise the negative charge formed after nucleophilic attack of native histidine 39 (Figure 4.8).

The following three mutants were prepared by site directed mutagenesis (SDM): Im9-F15A; Im9-Y10H; Im9-Y10H/F15A. DNA sequencing confirmed that the desired mutations were successfully created.

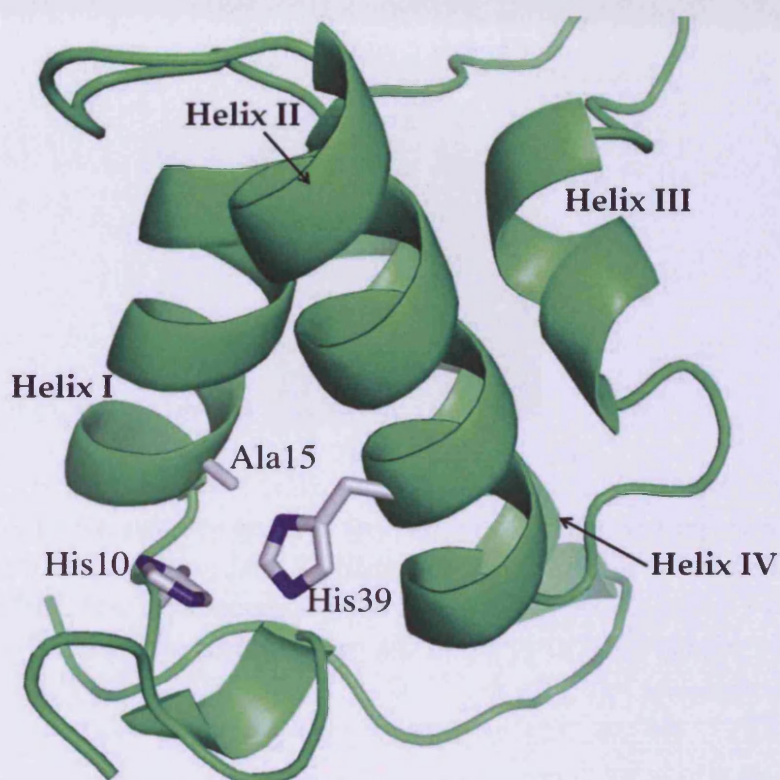


Figure 4.8: Molecular model of Im9-Y10H/F15A based on the NMR structure of Im9 (Osborne et al. 1996).

The protein was produced overnight in good yield in soluble form with no presence of inclusion body (Figure 4.9). For the purification, cells were harvested 16 h after induction.

4.3.1 Test expression of Im9 mutants

Although wild-type Im9 expressed well at 37 °C (Section 4.2.1), the expression of Im9 mutants at 37 °C led to the formation of inclusion bodies (Figure 4.9). A low temperature protein expression protocol was developed to tackle this problem.

BL21(DE3) cells were transformed with the mutated plasmid and grown in LB medium at 30 °C to an OD_{600nm} of 0.6 (typically after 7 h). The temperature was lowered to 16 °C and gene expression was induced with IPTG (0.5 mM).

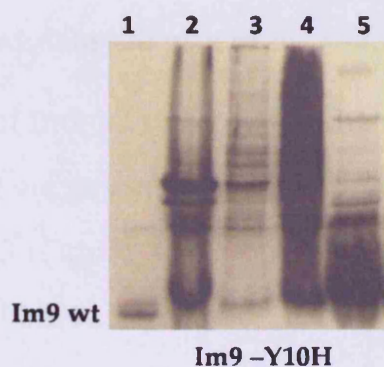


Figure 4.9: SDS-PAGE showing lysate of Im9 mutants at different temperatures; lane 1: purified Im9 wt used as marker; lane 2 pellets from expression at 37 °C with inclusion bodies of Im9-Y10H; lane 3 supernatant from expression at 37 °C; lane 4 2 pellets from expression at 16 °C; lane 5 supernatant from expression at 16 °C containing the majority of Im9-Y10H.

4.3.2 Purification of Im9-Y10H/F15A mutants

Im9 mutants were purified using the same method for the wild type protein (Section 4.2.2). Following purification, MALDI mass spectrometry indicated masses of 9555 for Im9-Y10H (calculated mass 9556.5), 9505 for Im9-F15A (calculated mass 9506.4), and 9483 for Im9-Y10H/F15A (calculated mass 9480.4).

4.3.3 Characterisation of Im9-Y10H

Residue Y10 is located in an unstructured loop before helix I. A molecular dynamic simulation suggested that Y10 is more stable in the transition state ensembles for Im9 folding than in the native form suggesting that it can play a role in the stabilisation of the transition state together with Y54 and R75 (Paci *et al.* 2004). Morton *et al.* investigated further the importance of non-native interactions in the correct folding of Im9 but without elucidating the role of Y10 (Morton *et al.* 2007). Analysis of the three-dimensional structure suggests a possible hydrogen bond between Y10 and H39 which may stabilise the tertiary structure. Mutation to histidine is expected to maintain the interaction, helping to modulate the pKa of the mutated histidine.

The CD spectrum of Im9-Y10H shows a lower helicity than the wild type, and deconvolution result suggests Y10H has approximate 27% α -helix, almost half that of wild type Im9 (Figure 4.10). There is shift of the minima from 222 and 208 nm to 226 and 205 nm. The melting temperature also decreased to 41 °C, 17 °C lower than the wild type (Figure 4.11).

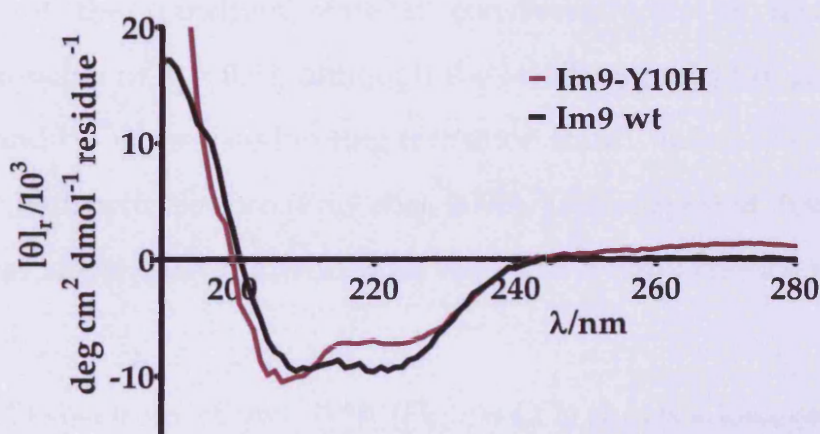


Figure 4.10: CD spectra of Im9-Y10H (red) and Im9 wt (black) recorded at 20 °C. Spectra recorded in 10 mM potassium phosphate buffer (pH 7.0).

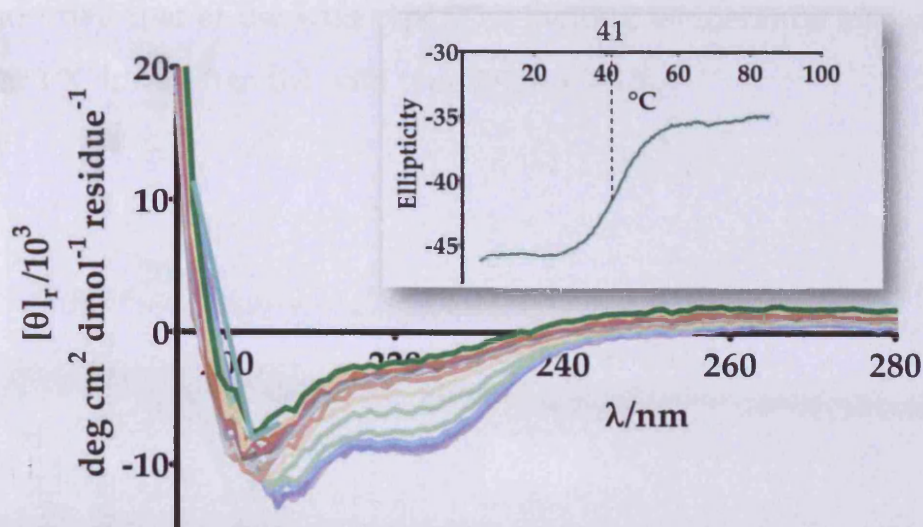


Figure 4.11: Effect of the temperature on Im9-Y10H structure across the UV range and followed at 222 nm (insert). Spectra recorded in 10 mM potassium phosphate buffer (pH 7.0).

4.3.4 Characterisation of Im9-F15A

F15 is located in the beginning of helix I and has been recognised as an important residue in determining the network of interactions that stabilise the transition state. Along with residues V19, I67, V71, K72 and R75, F15 helps the packing of helices I and IV (Paci *et al.* 2004). The importance of F15 in the stabilisation of the transition state is consistent with its relatively high experimental value of $\Phi = 0.58$, although the stabilising effect of residues from helices I, II and IV on the rate-limiting transition state is much lower than those in the native hydrophobic core (Friel *et al.* 2003). This suggested that the mutant would be less stable than the wild type with less α -helix content and a lower melting point.

The CD spectrum of Im9-F15A (Figure 4.12) shows a lower helicity than the wild type, and deconvolution gives 27% α -helix, similar to that of Im9-Y10H

and almost half that of the wild type. The melting temperature also decreased to 26 °C, 30 °C lower than the wild type (Figure 4.13).

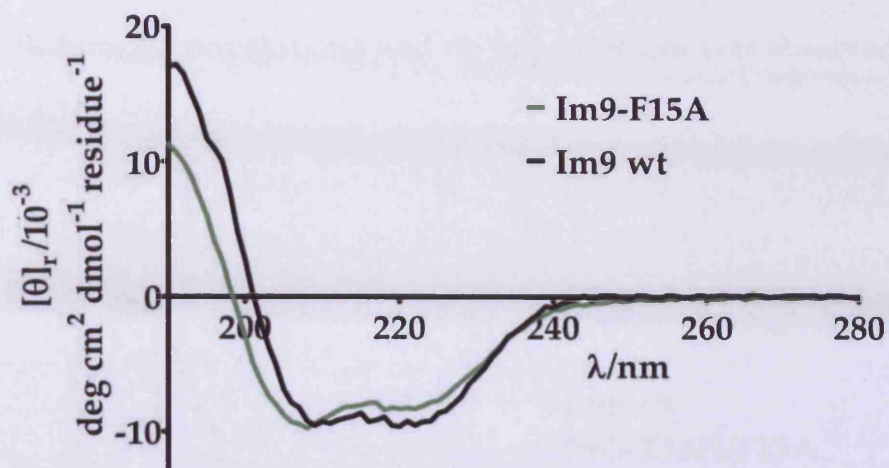


Figure 4.12: CD spectra of Im9-F15A (green) and Im9 wt (black) recorded at 20 °C. Spectra recorded in 10 mM potassium phosphate buffer (pH 7.0).

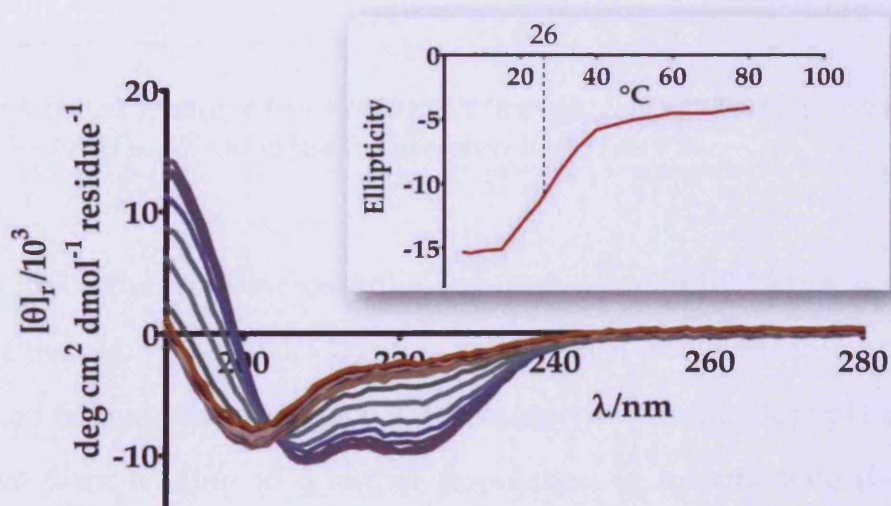


Figure 4.13: Right: effect of the temperature on Im9-F15A structure across the UV range and followed at 222 nm (insert). Spectra recorded in 10 mM potassium phosphate buffer (pH 7.0).

4.3.5 Characterisation of Im9-Y10H/F15A

This mutant was expected to be less stable than Im9-Y10H and Im9-F15A due to the combination of the destabilising effect of the two mutations. The CD spectrum of Im9-Y10H/F15A shows very little α -helix secondary structure (around 7% from deconvolution), and no improvement was observed even at 5 °C (Figure 4.14).

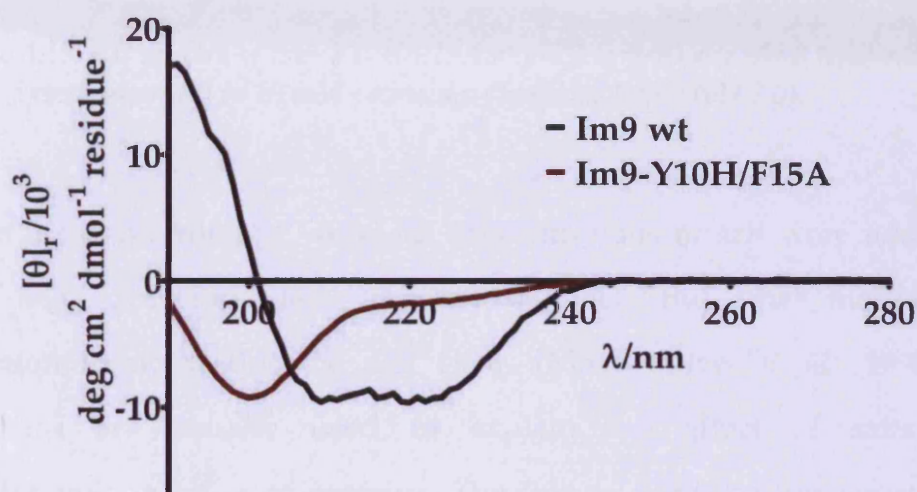


Figure 4.14: CD spectra of Im9-Y10H/F15A (red) and Im9 wt (black) recorded at 20 °C. Spectra recorded in 10 mM potassium phosphate buffer (pH 7.0).

A first attempt to increase the stability of Im9-Y10H/F15A was made by changing the pH. Im9-Y10H/F15A has a theoretical isoelectric point of 4.53 with a predicted net negative charge of 9. It has been shown that low pH destabilises the native form leading to a higher population of intermediate (Gorski *et al.* 2001), therefore the CD spectrum was measured at pH 8. The spectrum showed no improvement of α -helix content comparing to the one at pH 7 (Figure 4.15).

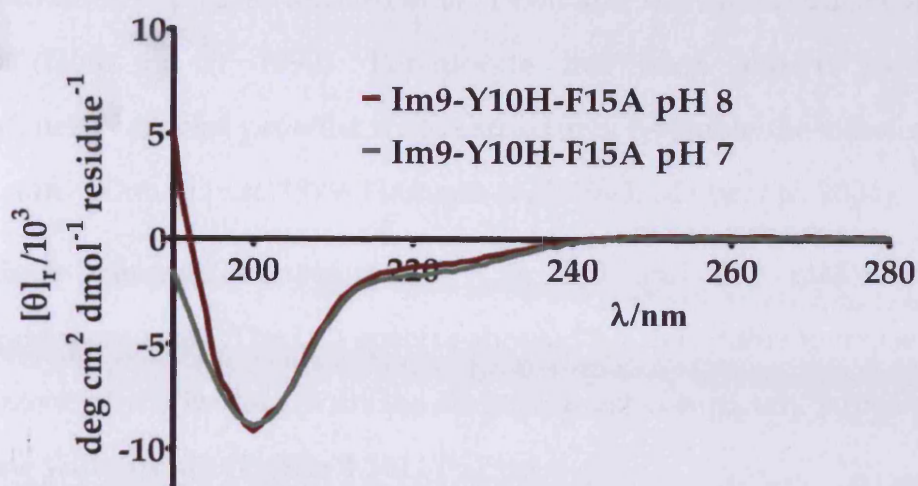


Figure 4.15: CD spectra of Im9-Y10H/F15A at pH 7 (green) and pH 8 (red) recorded at 20 °C. Spectra recorded in 10 mM potassium phosphate buffer (pH 7.0).

In a second attempt, different concentrations of salt were added to the buffer. Salt can modulate protein stability but the mechanism of stabilisation/destabilisation is not clear (Makhatadze *et al.* 1998). Three mechanisms are usually used to explain the effect of salts on the thermodynamic stability of proteins. The first is a Hofmeister effect with the increase in salt concentration likely leading to an increase in the apparent hydrophobic effect (and the salting-out of the non-polar core) due to changes in the solvent structure (Leberman *et al.* 1995). The second theory, suggested by computational studies results (Gilson *et al.* 1989) and by experimental data (Robinson *et al.* 1993), is the Debye-Hückel screening of unfavourable electrostatic interactions between parallel helix macro dipoles. The Debye-Hückel screening effect is nonspecific in nature, so both favourable and unfavourable interactions will be screened by mobile ions. The third theory is an anion binding model like the one for the salt interaction with nucleic acids (Goto *et al.* 1990). In order to obtain good CD data, sodium perchlorate, which is “invisible” in the far UV, was used. Perchlorate stabilises the α -helix

conformation of peptides (Masaru *et al.* 1997) and the acid-induced A state of proteins (Goto *et al.* 1990). Perchlorate has been shown to populate intermediates of several proteins whose structures resemble the classical molten globule state (Ahmad *et al.* 1979; Hamada *et al.* 1993; Maity *et al.* 2004).

Three different concentrations (50, 100 and 250 mM) of sodium perchlorate were used. The CD spectra showed no detectable increase in the α -helix content of the mutant with the traces almost completely superimposable on the one without salt (Figure 4.16)

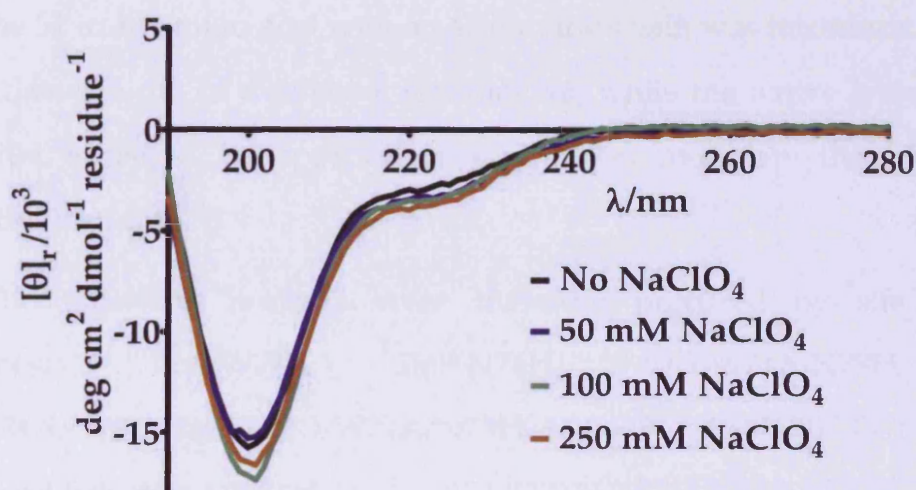


Figure 4.16: CD spectra of Im9-Y10H/F15A recorded at 20 °C with different concentrations of NaClO₄: no salt (black); 50mM (blue); 100 mM (green); 250 mM (yellow). Spectra recorded in 10 mM potassium phosphate buffer (pH 7.0).

The almost complete lack of α -helix in Im9-Y10H/F15A is quite surprising. Paci *et al.* suggested that helices I and IV are not only substantially formed in the rate-limiting transition state, but their hydrophobic packing is also essentially native-like (Paci *et al.* 2004). The mutation F15A although destabilising the tertiary structure of the protein, should not influence the formation of helix I, as suggested by 27% α -helix present in the Im9-F15A

mutant. A possible explanation for the lack of helicity of Im9-Y10H/F15A is that the destabilisation caused by mutated histidine (Y10H) of the hydrophobic core around which the helices I and IV fold led to destabilisation of all helices.

4.4 Design 2: Im9-W74A/N78H series

In the second set of design, mutations around tryptophan 74 and asparagine 78 were thoroughly investigated by generating seven mutants. Tryptophan 74 was mutated to alanine to create the cavity and asparagine 78 was mutated to histidine to act as nucleophile (Figure 4.17). A further mutation of leucine 52 to an amino acid with an acidic side-chain was introduced in order to modulate the pK_a of the native histidine 46, while the native lysine 80 was considered to be in an appropriate position to modulate the pK_a of the introduced histidine 78.

The following mutants were therefore prepared by site directed mutagenesis: Im9-W74A; Im9-N78H; Im9-W74A/N78H, Im9-L52E/W74A/N78H, Im9-L52D/W74A/N78H, Im9-D51A/L52D/W74A/N78H and Im9-D51A/L52E/W74A/N78H.

DNA sequencing confirmed that all the desired mutations were successful with no unwanted mutations. Im9 mutants were purified as described for the wild type protein (Section 4.2.2). Following purification, MALDI mass spectrometry indicated masses of 9468 for Im9-W74A (calculated mass 9467.3), 9605 for Im9-N78H (calculated mass 9605.5), 9491 for Im9-W74A/N78H (calculated mass 9490.4), 9507 for Im9-L52E/W74A/N78H (calculated mass 9506.3), 9492 for Im9-L52D/W74A/N78H (calculated mass 9492.3), 9462 for Im9-D51A/L52E/W74A/N78H (calculated mass 9462.3) and 9449 for Im9-D51A/L52D/W74A/N78H (calculated mass 9448.3).

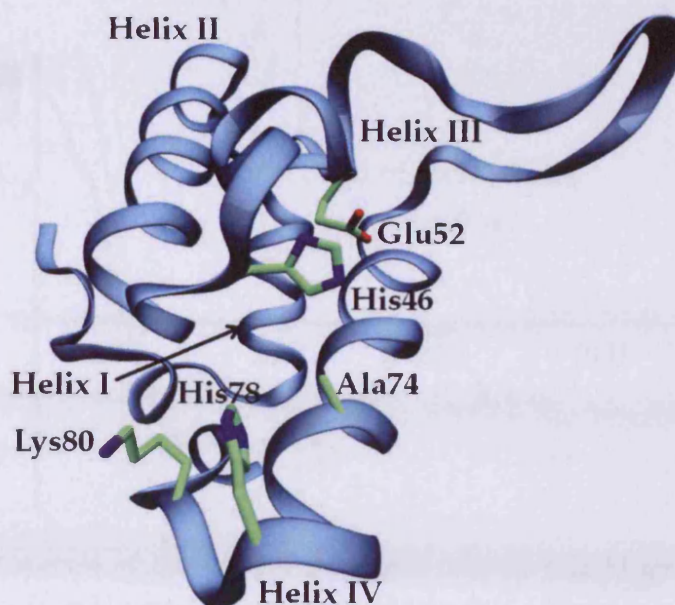


Figure 4.17: Molecular model of Im9-L52D/W74A/N78H based on the NMR structure of Im9 (Osborne *et al.* 1996).

4.4.1 Im9-W74A

W74 is located in the C-terminal part of helix IV and has been used to study the kinetics of Im9 protein folding (Friel *et al.* 2003) because its fluorescence in the native fold is quenched by interaction with histidine 46, the imidazole ring of which forms a stacking interaction with the indole ring (Wallis *et al.* 1998). It has been shown that W74 interacts with Y54 and Y55 in the transition state ensemble and the interaction persisted even in presence of 6 M urea, an interaction that is not present in the native fold of the protein and neither in the denatured state (Paci *et al.* 2004). This suggests a role of W74 in the folding of the protein.

The CD spectrum shows a change in the intensity of the two minima, with an increase of the signal at 222 nm and a decrease at 208 nm (Figure 4.18). The deconvolution of the spectrum indicated 34.8% α -helix. The mutant has a melting point of 37 °C (Figure 4.19).

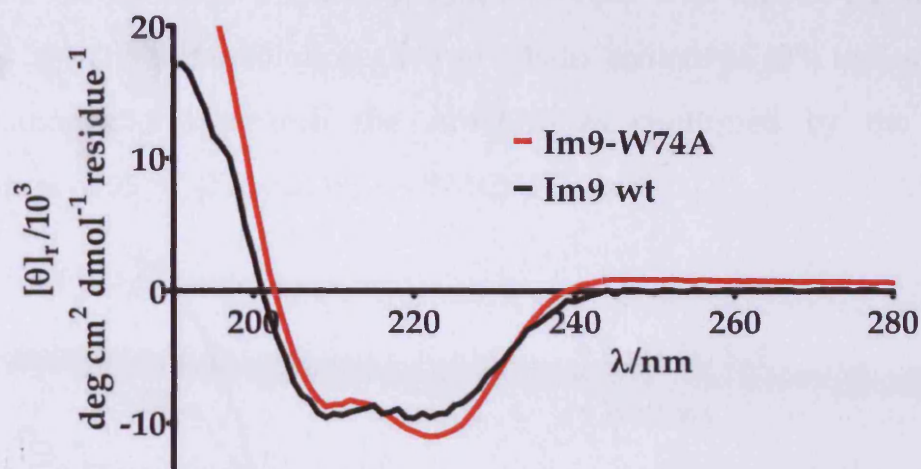


Figure 4.18: CD spectra of Im9-W74A (red) and Im9 wt (black) recorded at 20 °C. Spectra recorded in 10 mM potassium phosphate buffer (pH 7.0).

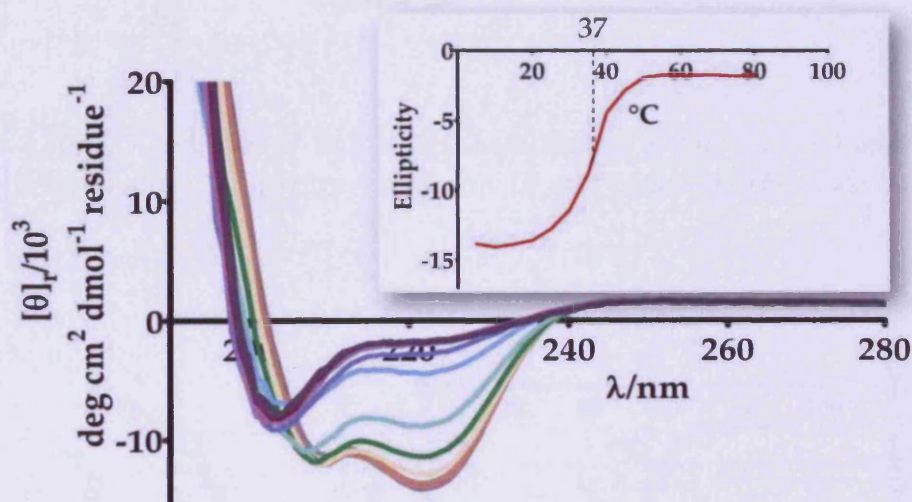


Figure 4.19: Effect of the temperature on Im9-W74A structure across the UV range and followed at 222 nm (insert). Spectra recorded in 10 mM potassium phosphate buffer (pH 7.0).

4.4.2 Characterisation of Im9-N78H

N78 is the last residue in the C-terminal part of helix IV. It doesn't have any interaction with other residues in the protein so it was expected that the mutation to histidine wouldn't disturb the folding of the protein.

The CD spectrum is almost superimposable with that of the wild type (Figure 4.20). The deconvolution gave an α -helix content of 49% indicating that the mutation didn't perturb the structure as confirmed by the melting temperature of 55 °C (T_m wild type = 57 °C) (Figure 4.21).

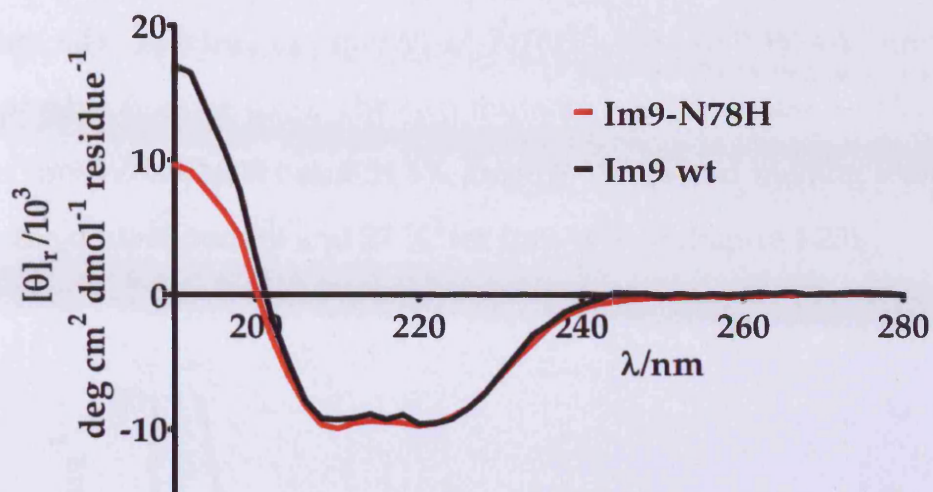


Figure 4.20: CD spectrum of Im9-N78H. Left: spectra of Im9 wt (red) and Im9 wt (black) recorded at 20 °C. Spectra recorded in 10 mM potassium phosphate buffer (pH 7.0).

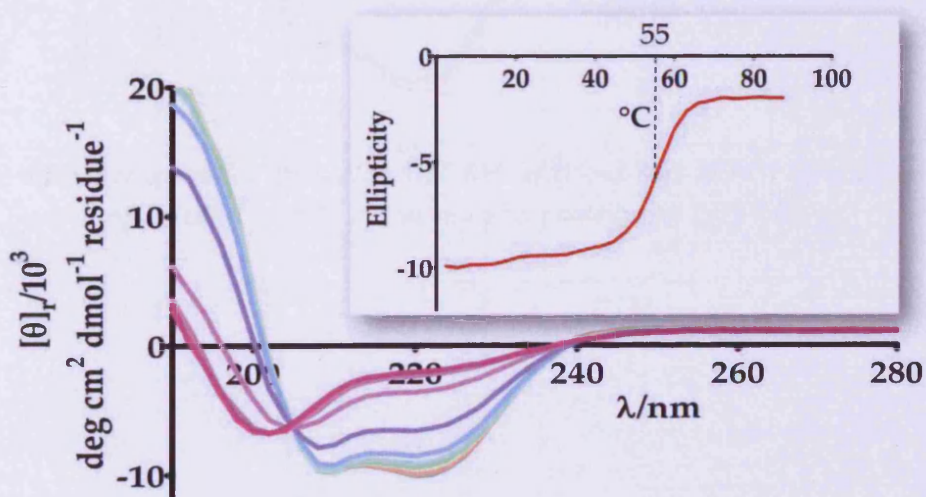


Figure 4.21: Effect of the temperature on Im9-N78H structure across the UV range and followed at 222 nm (insert). Spectra recorded in 10 mM potassium phosphate buffer (pH 7.0).

4.4.3 Characterisation of Im9-W74A/N78H

This mutant was expected to have the same characteristics as Im9-W74A due to the fact that Im9-N78H shares similarity with the wild type thermal stability wise and secondary structure wise.

The CD spectra of Im9W74A/N78H and Im9-W74A are almost superimposable (Figure 4.22). The two mutants have a similar α -helix content, 34.4% for Im9-W74A/N78H and 34.8% for Im9-W74A and melting temperature 35 °C for the double mutant and 37 °C for Im9-W74A) (Figure 4.23).

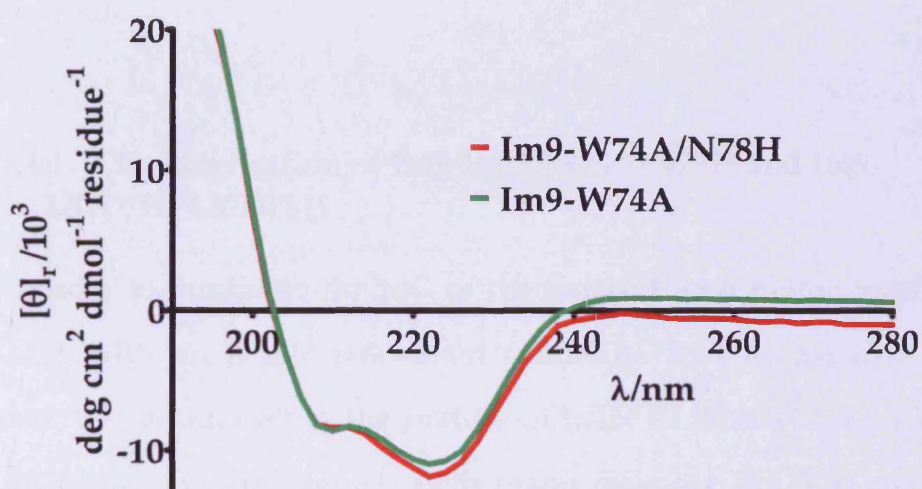


Figure 4.22: CD spectra of Im9-W74A/N78H (red) and Im9-W74A (green) recorded at 20 °C. Spectra recorded in 10 mM potassium phosphate buffer (pH 7.0).

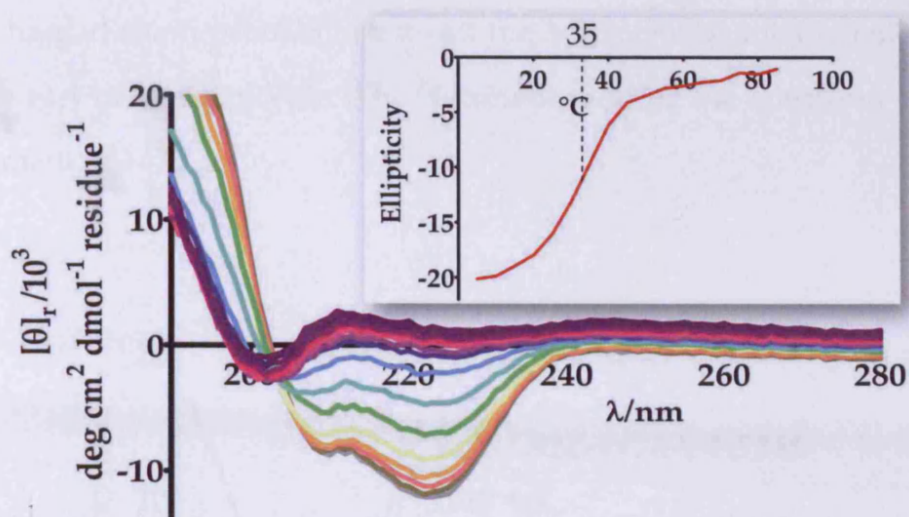


Figure 4.23: Effect of the temperature on Im9-W74A/N78H structure across the UV range and followed at 222 nm (insert). Spectra recorded in 10 mM potassium phosphate buffer (pH 7.0).

4.4.4 Characterisation of Im9-L52E/W74A/N78H and Im9-L52D/W74A/N78H

In order to modulate the pK_a of the native H46 a mutation of L52 to an amino acid with an acidic side-chain (glutamic acid or aspartic acid) was performed. L52 is situated in the middle of helix III which plays a key role in the hydrophobic core. It interacts with many residues: I67, T70 and V71 from helix IV; H46, S48 and G49 from the loop between helices II and III; F40 from helix II.

The CD spectrum of Im9-L52E/W74A/N78H shows a loss of secondary structure compared with the wild type, with a shift of the minimum to 206 nm, a decrease of the minimum at 222 nm (Figure 4.24), which more resembles the denaturated state of Im9 wild type. The thermal denaturation of the protein didn't show a melting point but a progressive loss of the secondary structure (Figure 4.25). Friel *et al.* showed that the mutation L52A does not destabilise the rate-limiting transition state of Im9 (Friel *et al.* 2003) but the introduction of a

longer charged chain probably destroys the hydrophobic interaction of helix III with the rest of the molecule. The deconvolution of the spectrum gave an α -helix content of 14%.

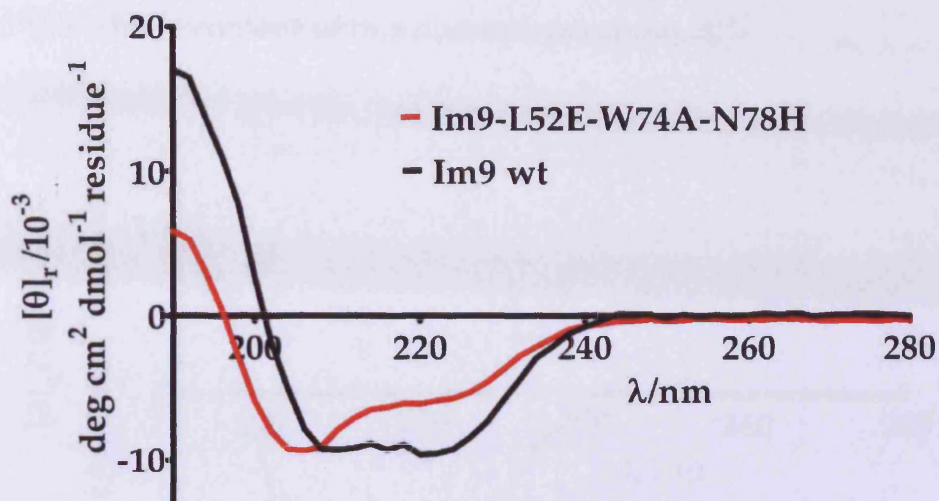


Figure 4.24: CD spectra of *Im9-L52E/W74A/N78H* (red) and *Im9 wt* (black) recorded at 20 °C. Spectra recorded in 10 mM potassium phosphate buffer (pH 7.0).

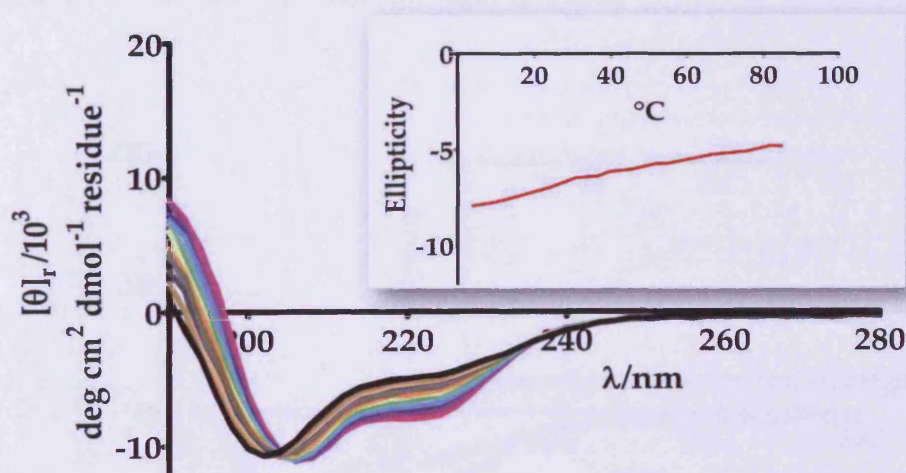


Figure 4.25: Effect of the temperature on *Im9-L52E/W74A/N78H* structure across the UV range and followed at 222 nm (insert). Spectra recorded in 10 mM potassium phosphate buffer (pH 7.0).

Im9-L52D/W74A/N78H shows minima at 208 nm and 222 nm in its CD spectrum (Figure 4.26), as in Im9-W74A/N78H. Although more structured than the previous mutant with glutamic acid, probably because of the shorter side-chain, the mutant was far less stable than Im9-W74A/N78H, with complete loss of secondary structure at 26 °C (Figure 4.27). The deconvolution of the spectrum indicated the α -helix content of this mutant was about 20%.

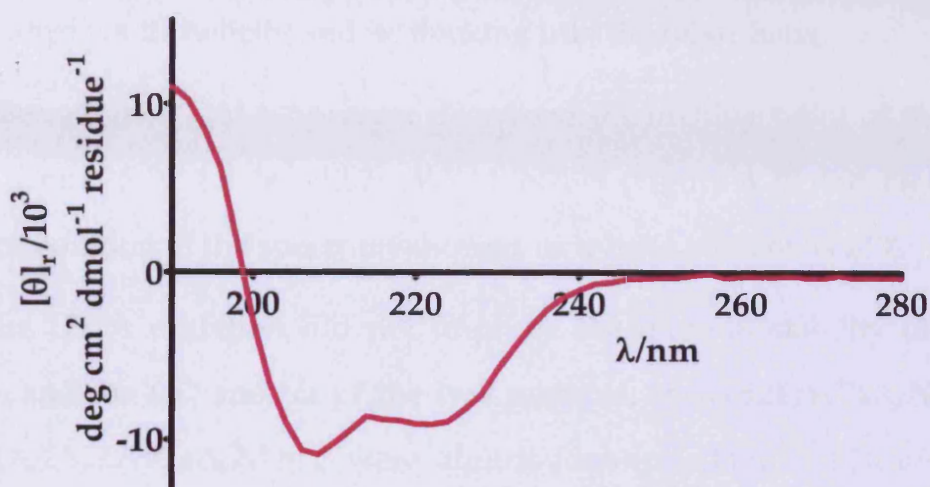


Figure 4.26: CD spectrum of Im9-L52D/W74A/N78H recorded at 20 °C. Spectrum recorded in 10 mM potassium phosphate buffer (pH 7.0).

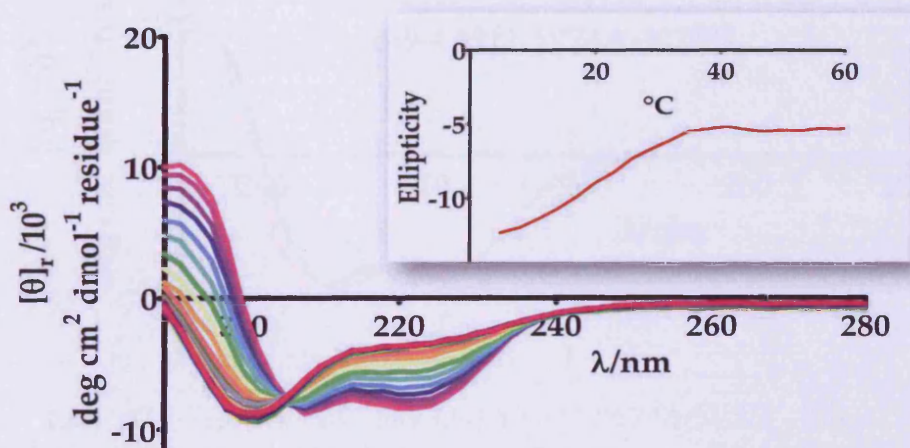


Figure 4.27: Effect of the temperature on Im9-L52D/W74A/N78H structure across the UV range and followed at 222 nm (insert). Spectra recorded in 10 mM potassium phosphate buffer (pH 7.0).

4.4.5 Characterisation of Im9-D51A/L52D/W74A/N78H and Im9-D51A/L52E/W74A/N78H

The mutations L52D and L52E, located in the middle of helix III, appeared to cause further destabilisation to the mutant Im9-W74A/N78H. This observation confirmed the importance of helix III in the hydrophobic core of the native state of Im9 for the protein structure integrity. Residue D51 which is the first amino acid in the N-terminal part of the helix III, was mutated to alanine in order to improve its helicity and its docking into the other helix.

The mutation D51A however decreased the melting point of the mutant to 20 °C from 26 °C for Im9-L52D/W74A/N78H (Figure 4.28 and Figure 4.29). The deconvolution of the spectrum showed its α -helix content as 21%

The D51A mutation did not improve the thermal stability of the two mutants, and the CD spectra of the two mutants, Im9-L52E/W74A/N78H and Im9-D51A/L52E/W74A/N78H, were almost identical (Figure 4.30 and Figure 4.31).

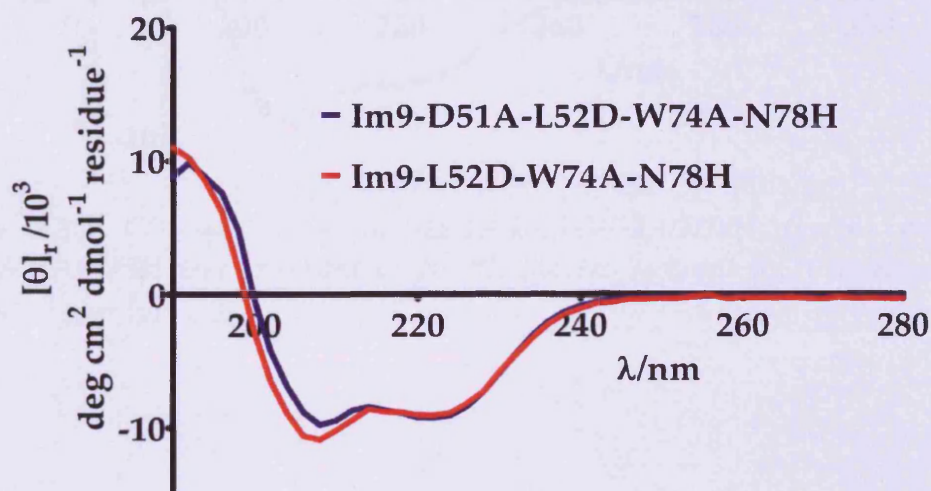


Figure 4.28: CD spectra of Im9-D51A/L52D/W74A/N78H (blue) and Im9-L52D/W74A/N78H (red) recorded at 20 °C. Spectra recorded in 10 mM potassium phosphate buffer (pH 7.0).

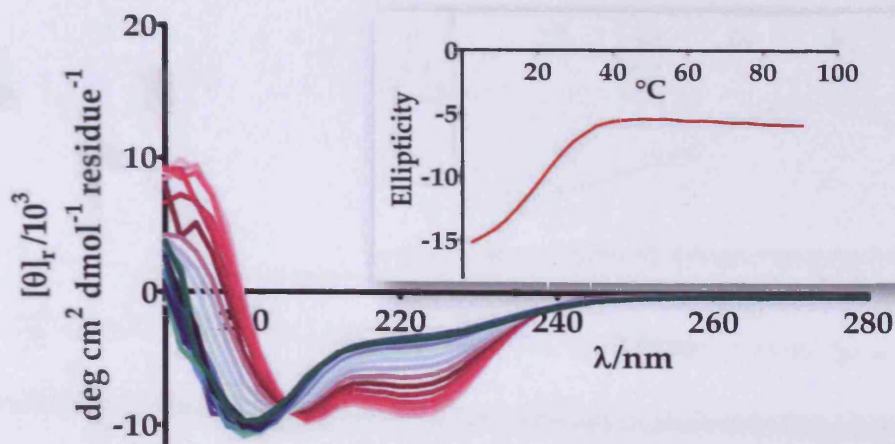


Figure 4.29: Effect of the temperature on Im9-D51A/L52D/W74A/N78H structure across the UV range and followed at 222 nm (insert). Spectra recorded in 10 mM potassium phosphate buffer (pH 7.0).

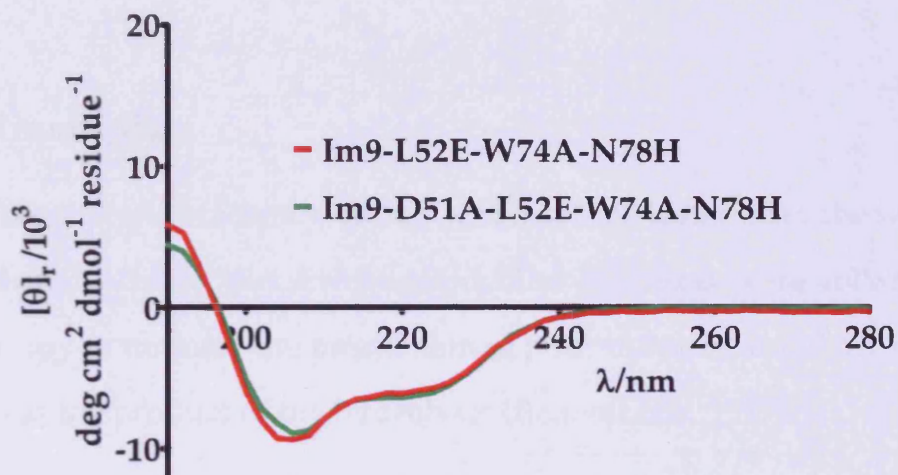


Figure 4.30: CD spectra of Im9-D51A/L52E/W74A/N78H (green) and Im9-L52E/W74A/N78H (red) recorded at 20 °C. Spectra recorded in 10 mM potassium phosphate buffer (pH 7.0).

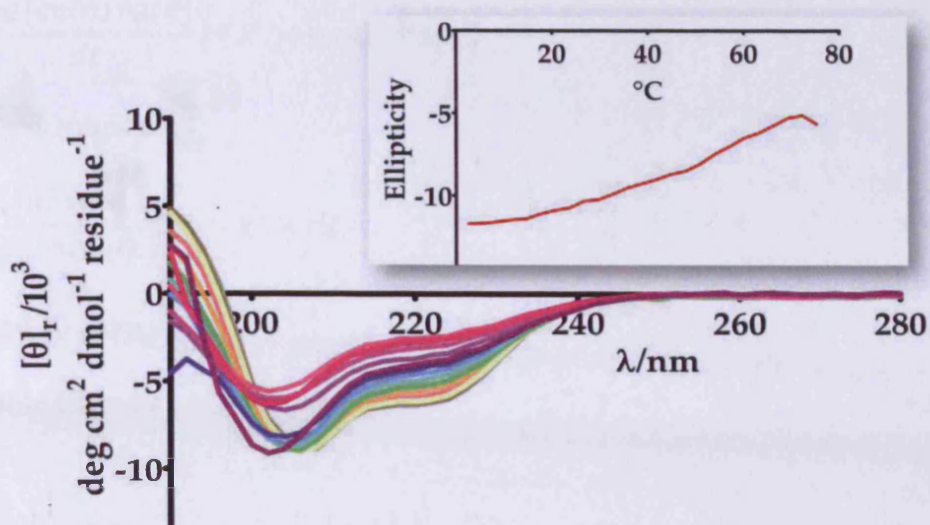


Figure 4.31: Effect of the temperature on Im9-D51A/L52E/W74A/N78H structure across the UV range and followed at 222 nm (insert). Spectra recorded in 10 mM potassium phosphate buffer (pH 7.0).

4.5 Kinetic data

The catalytic efficiency of Im9 wt and its 10 mutants in the hydrolysis of *p*-nitrophenyl acetate was investigated. The reactions were followed by UV spectroscopy to monitor the production of *p*-nitrophenol at 320 nm ($\epsilon_{320\text{nm}} = 7600 \text{ M}^{-1} \text{ cm}^{-1}$) as the product of the hydrolysis (Bennett *et al.* 1977).

The reaction rates were then determined under pseudo-first order conditions assuming that in the presence of a great excess of catalyst compared to substrate the concentration of the former will remain constant:

$$\frac{-d[\text{substrate}]}{dt} = k \times [\text{substrate}] \times [\text{protein}] \quad 4.1$$

When

$$[\text{substrate}] \ll [\text{protein}] \quad k' = k \times [\text{protein}] \quad 4.2$$

then

$$\frac{-d[\text{substrate}]}{dt} = k' \times [\text{substrate}] \quad 4.3$$

rearranged to give

$$\frac{d[\text{substrate}]}{[\text{substrate}]} = -k' \times dt \quad 4.4$$

integrated to give

$$\ln[\text{substrate}] = -k' \times t + C \quad 4.5$$

for $t = 0$

$$\ln[\text{substrate}]_0 = C \quad 4.6$$

substitute in answer

$$\ln[\text{substrate}] = -k' \times t + \ln[\text{substrate}]_0 \quad 4.7$$

Which is the mathematical presentation of a straight line. The slope of the line (k') gives the pseudo first-order rate constant of the reaction for a given catalyst concentration. Substitution of k' and catalyst concentration in equation 2 gives the pseudo- first order rate (Atkins). The reaction was first tested using the 4-methyl imidazole, which mimics the histidine side-chain, catalysed reaction (pH 7), as a comparison (Figure 4.32 and Figure 4.33).

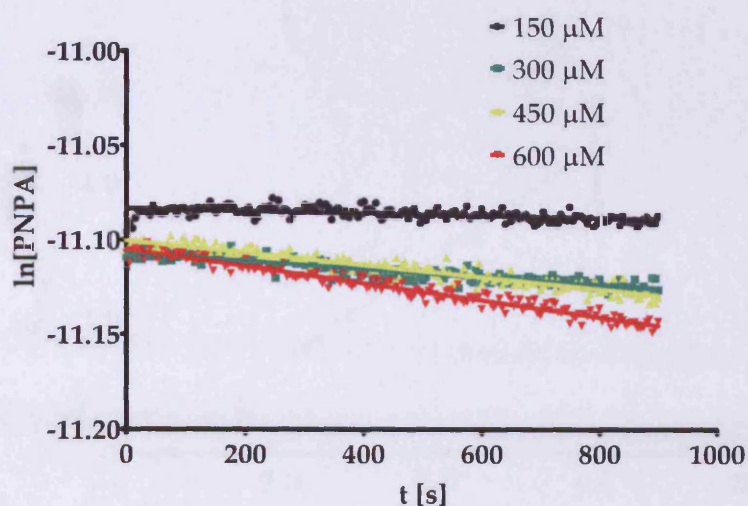


Figure 4.32: Reaction of 4-methylimidazole (150–600 μM) with *p*-nitrophenyl acetate (15 μM) at pH 7, 20 $^{\circ}\text{C}$, as followed by the production of *p*-nitrophenol. [*p*-nitrophenyl acetate] was plotted on a log scale to give pseudo first order rate constants at different catalyst concentrations.

As expected, the reaction rates increased with increasing 4-methylimidazole concentrations. The apparent second-order rate constant for the 4-methylimidazole catalysed hydrolysis of *p*-nitrophenyl acetate (pH 7) is $0.080 \pm 0.015 \text{ M}^{-1} \text{ s}^{-1}$, consistent with the published value of $0.11 \text{ M}^{-1} \text{ s}^{-1}$ (Bolon *et al.* 2001).

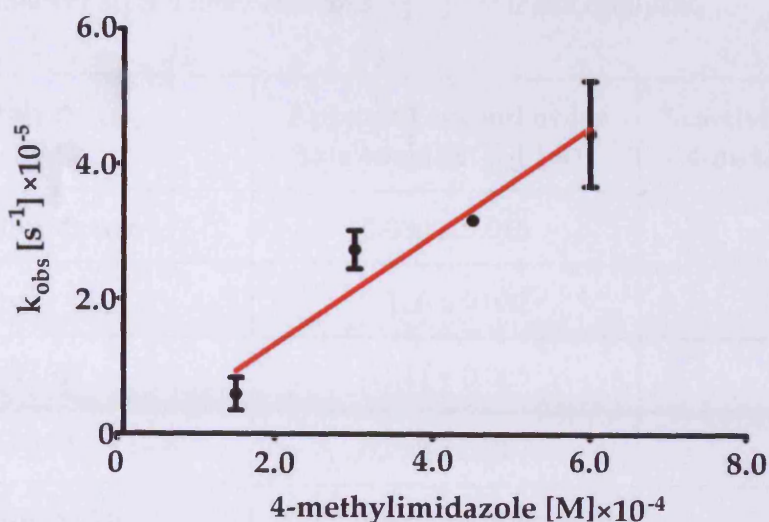


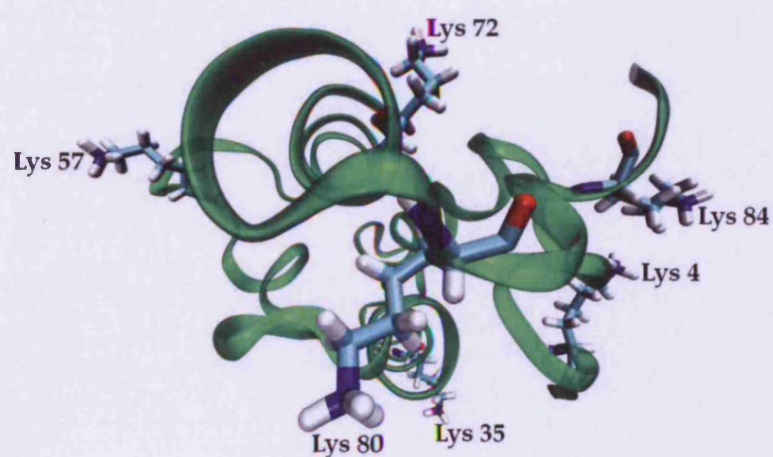
Figure 4.33: Pseudo first order rate constant for the 4-methylimidazole (150-600 μM) catalysed hydrolysis of *p*-nitrophenyl acetate (15 μM) at pH 7, 20 °C over a range of catalyst concentrations. The gradient of the line gives the apparent second order rate constant.

The catalytic efficiency of Im9 and its 10 mutants to catalyse the hydrolysis *p*-nitrophenyl acetate were tested under the same conditions (Table 4.1).

Im9 and the mutants Im9-N78H and Im9-W74A showed activities similar to 4-methylimidazole. The mutant Im9-W74A/N78H increased the rate of the reaction three-fold. The further mutations, L52D and D51A, however showed a decrease in the rate of catalysis compared to Im9-W74A/N78H, probably due to the loss in stability caused by the mutations. The mutants Im9-N78H and Im9-Y10H/F15A, with the latter lacking any secondary structure, have the same number of histidine residues as Im9-W74A/N78H. Their lower activities show that the increase in the rate of Im9-W74A/N78H is not just a simple result of the additional histidine residues in the protein, but also related to the designed active site. The presence of several lysine residues on the surface of the protein can explain the activity of Im9 itself (Figure 4.34).

Table 4.1: Apparent second order rate constant of the tested catalysts.

Catalyst	Apparent second order rate constant [M ⁻¹ s ⁻¹]	% activity relatively to 4-methylimidazole
4-methylimidazole	0.080 ± 0.015	100%
Im9	0.120 ± 0.032	150%
Im9-Y10H	0.041 ± 0.005	50%
Im9-F15A	0.098 ± 0.007	122%
Im9-Y10H/F15A	0.130 ± 0.006	162%
Im9-N78H	0.130 ± 0.012	162%
Im9-W74A	0.090 ± 0.005	112%
Im9-W74A/N78H	0.364 ± 0.032	455%
Im9-L52D/W74A/N78H	0.234 ± 0.027	292%
Im9-D51A/L52D/W74A/N78H	0.083 ± 0.005	103%

**Figure 4.34:** Cartoon representation of the NMR structure of Im9 (Osborne et al. 1996), with the surface lysine residues shown.

In order to trap the acetylated intermediates of Im9-W74A/N78H, Im9-W74A and Im9-N78H, the proteins were incubated with a five-fold excess of *p*-nitrophenyl acetate. The reaction mixtures were analysed by MALDI mass spectrometry every 20 min for 1 h. The three proteins showed a similar profile with two additional peaks, M+42 (acetylation) and M+84 mass units (diacetylation) detected from all of them (Figure 4.35). The accumulation of acetyl residues in Im9-W74A and Im9-N78H shows that these are nonspecific sites of acetylation on the proteins, probably on the lysine residues which are known to be irreversibly acetylated by *p*-nitrophenyl acetate (De Caro *et al.* 1988). Im9-W74A/N78H showed a faster accumulation of acetyl groups compared with the other two, which can be explained by the higher activity of this mutant compared to the other two.

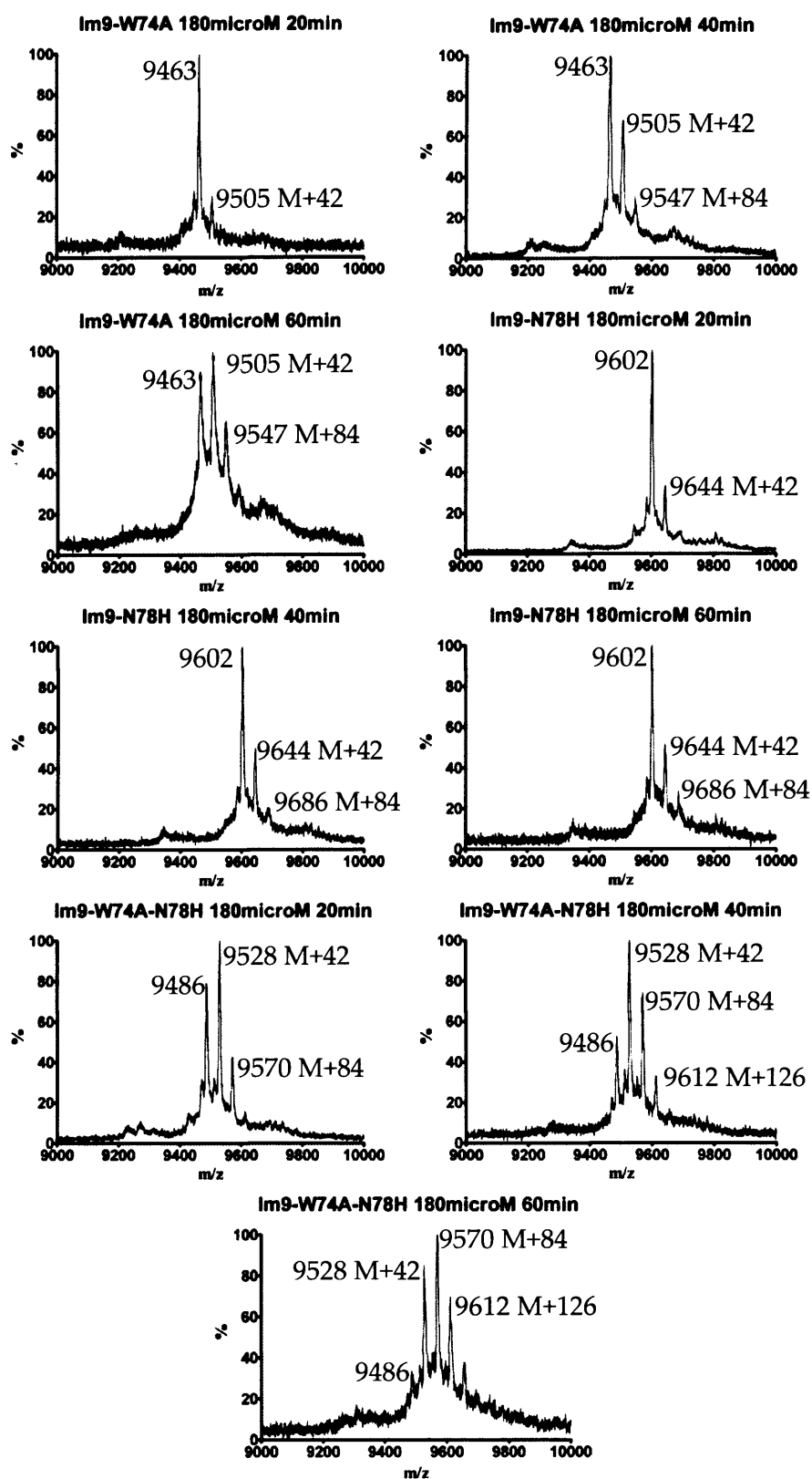


Figure 4.35: Mass spectra of Im9-W74A, Im9-N78H, Im9-W74A/N78H incubated with five-fold excess of p-nitrophenyl acetate. The presence of peaks M+42 (acetylation) and M+84 (diacetylation) indicate the presence of nonspecific sites of acetylation.

4.6 Conclusions

Im9 was chosen to be the scaffold for the novel design of catalyst for the hydrolysis of *p*-nitrophenyl acetate. The designed active site was based on histidine residues. Site-directed mutagenesis was used to generate the sequences of 2 series, Im9-Y10H/F15A and Im9-W74A/N78H, to give a total of ten mutants. Together with the wild type Im9, 11 proteins were expressed and purified. MALDI-TOF measurements confirmed the mass for all of them. CD experiments were conducted to investigate the protein and mutants secondary structures and thermal stabilities. Kinetic data using *p*-nitrophenyl acetate as substrate were collected to study the artificial enzymes' activities. As proof of concept, artificial enzymes based on Im9 were designed and obtained.

Among the two series of mutants tested, Im9-Y10H/F15A did not demonstrate enough stability; either secondary structure consequently was not able to accelerate the rate of *p*-nitrophenol acetate hydrolysis any more than wild-type Im9. Whereas Im9-W74A/N78H retained good stability and increased the rate of the reaction three-fold compared to 4-methylimidazole and wild-type Im9. Additional mutations designed to improve the catalytic efficiency however destabilised the protein secondary structures.

Detection of acetylated proteins suggests that, rather than complete catalysis, a single turnover event is observed with consequent inactivation of the active site. This is suggested by the faster acetylation detected with the more reactive Im9-W74A/N78H when compared to the other mutants.

Interestingly diacetylation of Im9-W74A/N78H, Im9-W74A and Im9-N78H were also observed. This suggests the presence of nonspecific acetylation sites, probably solvent-exposed lysines. Indeed Im9-W74A/N78H with 34% α -helix, comparing to the 48% for wild type, may have more residues exposed to

solvent. However the three fold rate increase displayed by the Im9-W74A/N78H cannot be explained only by solvent exposed lysines being able to remove acetyl group from the substrate, because the completely unstructured Im9-Y10H/F15A did not show a similar effect. Therefore the increase catalytic activity demonstrated by Im9-W74A/N78H can be closely related to its active site in which the native histidine 46 acts in synergy with the mutated histidine 78.

A single turnover event would mean that a further modification of the active site is necessary to allow water to enter and hydrolyse the acetylated histidine. The increased instability exhibited by Im9 mutants with the increased number of altered amino acids indicates the structural stability, which is connected with the tolerance to mutation (Besenmatter *et al.* 2006), of the protein needs to be improved. The introduction of disulfide bonds to stabilise the structure can be a solution, however due to the strict geometrical requirement, this is not a straightforward solution. Furthermore the disulfide stabilises both the native and the unfolded state of a protein and therefore can lead to destabilisation of the protein (Zavodszky *et al.* 2001). Another solution can be the construction of a chimeric protein with a sterically constrained and immunodominant antibody loop region as done for Im7 which led to an increase in the melting temperature by 10 °C (Juraja *et al.* 2006).

5. DIRECTED EVOLUTION OF IM9-W74A/N78H

5.1 Introduction

5.1.1 Directed evolution

Rational design seldom produces enzymes that can rival natural enzymes in terms of rate enhancement. To increase their catalytic efficiency a directed evolution step is usually included in the design process (Breslow 2005). The number of mutants generated by randomisation of the protein sequence increases exponentially with the number of amino acids involved: a total of 20^n (where n is the number of randomised amino acids) possible mutants are obtained when a complete degeneration of the genetic code is used. The high number of mutants to be screened requires a powerful selection method and phage display was chosen because of its relative simplicity (the selection method can be described as affinity chromatography) compared to *in vivo* selection (Farinas *et al.* 2001).

5.1.2 Phage display

The phage display technique was first introduced by Smith who showed that DNA fragments can be inserted into filamentous phage gene III to create a fusion protein with the inserted sequence in the middle. The "fusion phages" retained their infectivity and could be enriched more than 1000-fold over ordinary phage by affinity for antibody directed against the foreign sequence (Smith 1985).

Filamentous bacteriophages are a group of viruses in which a circular DNA (+) single-stranded genome is encapsulated in a long protein capsid cylinder. The viruses enter the host cells and use the biological machinery of the

host to replicate its DNA and express the proteins that constitute the capsid. The phage is assembled at the moment of its extrusion from the cell. The main consequence of the infection is an approximately 50% increase in the replication time of the host (Barbas III *et al.* 2004).

The capsid is formed by 5 proteins (Figure 5.1). The major capsid protein VIII (pVIII) is the constituent of the cylinder and is mainly α -helical (Marvin *et al.* 1994). The minor protein III (pIII), responsible for the infection process, is made up of three domains separated by glycine rich regions (Marvin 1998).

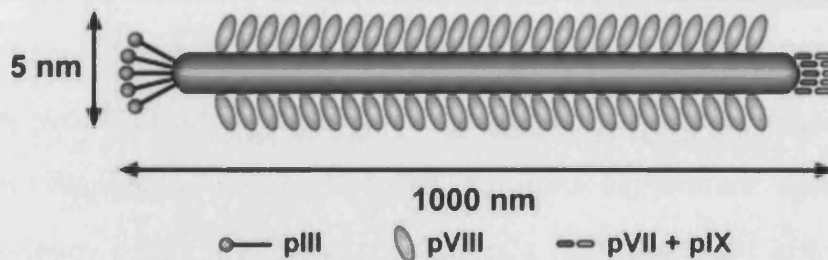


Figure 5.1: Structure of a typical filamentous phage virion (Krumpe *et al.* 2006).

The replication process starts with the synthesis of the complementary DNA (-) strand by bacterial enzymes to produce a supercoiled double-stranded circular DNA. One of the first proteins produced from the newly synthesised circular DNA, pII, nicks the (+) strand and a replication cycle, carried out by bacterial enzymes, starts using the resulting 3'-hydroxyl terminus as a primer and the DNA (-) strand as template. The displaced open single DNA (+) strand is then cyclised and converted to a double strand by bacterial enzymes. The process continues until a critical concentration of protein V (pV) is reached, then the pV dimer inhibits double strand production and only the DNA (+) is then replicated. This switch initiates the assembly process which leads to the

extrusion of the single DNA (+) strand encapsulated in the coating proteins (Barbas III *et al.* 2004).

All five coating proteins have been used to create a fusion with the protein or peptide to be displayed, but pIII and pVIII fusions are most common (Brakmann *et al.* 2002). Only small peptides can be displayed in fusion with pVIII (Petrenko *et al.* 1996) because larger peptides prevent the correct packing of the capsid (Endemann *et al.* 1995) and the extrusion of the phage due to steric hindrance (Marciano *et al.* 1999). To display a large protein in pVIII a hybrid system is usually used with the chimeric gene placed in a plasmid (Greenwood *et al.* 1991). When bacteria transformed with this plasmid are infected with wild type phage, the chimeric pVIII is used as well as the pVIII wild type in the coating process leading to a hybrid phage with multiple copies of the fusion protein displayed at the phage surface. Polyvalent display leads to multivalency binding events and confers a high apparent affinity (avidity) on weak-binding virions. When the initial binding affinity of the WT protein is low ($K_i > 1 \mu\text{M}$) a polyvalent display strategy may be advantageous because it allows a far greater representation of rare library variants (Zani *et al.*).

pIII is the capsid protein most often used to create fusion proteins. The disadvantage is that only 5 molecules can be displayed (the number of pIII molecules in the capsid) but the advantage is that even large proteins can be displayed without compromising the packing of the capsid (Barbas III *et al.* 2004). Sometimes a large protein displayed on pIII can decrease the infectivity (Smith 1985) but this can be overcome by using a phagemid, a plasmid that contains a phage origin of replication in addition to its plasmid origin of replication and that can be packed in the phage coat (Barbas III *et al.* 2004). The wild type pIII is produced by a helper phage, which encodes for the production of the coating proteins for the phagemid and for the helper phage itself. A

hybrid phage will then be assembled with a mixture of wild type pIII and the chimeric pIII (Dunn 1996). Phagemids are easier to handle than phages but the number of foreign proteins displayed in each phage is difficult to control and sometime much fewer than one copy per phage is displayed (Brakmann *et al.* 2002).

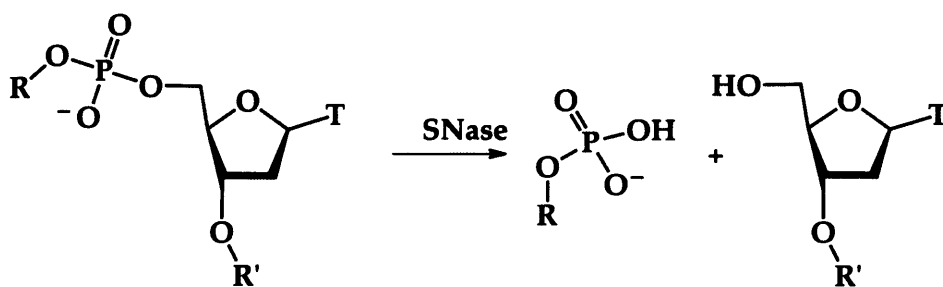
Phage display was invented to perform affinity screening of protein fragments encoded by the associated complementary DNA fragment (Barbas III *et al.* 2004). In the first phage selection experiment a polyclonal antibody against the *EcoRI* endonuclease was used to coat a polystyrene dish and to selectively bind phage that displayed a fragment this protein fused to pIII, from a pool containing a large excess of phage without insert. The phages with the insert that still bound to the plate after several washes were then eluted by denaturation of the immobilised antibodies in acidic conditions. The eluted phage were after neutralisation used to infect *E. coli* cells, which amplified the enriched pool of antibody-binding phage (Smith 1985).

5.1.3 Biopanning

All the methods that have been used to screen phage display libraries have the same objective: to find ligands for the screening molecule which is usually immobilised (Brakmann *et al.* 2002). The phage population can be divided into four major groups. Firstly, nonbinding phage, which represent the vast majority of the library and are removed by repeated washing. Tween 20 helps to remove this population lowering the background binding (Smith *et al.* 1993). Secondly, plate binders, which contain a particular sequence WXXWXXXW that binds polyurethane/polystyrene magnetic particles even in the presence of surfactants (Gebhardt *et al.* 1996). This population can be

removed by performing a pre-screening with all the components except the screening molecule. The last two populations are binders of the screening molecule, with low affinity binders more abundant than high affinity binders.

When selecting for catalytic activity the screening molecule should not only bind the desired phage population but should select for mutants able to catalyse the reaction for which the enzyme is designed. Several strategies have been used. One is the use of an inhibitor of the reaction that resembles the product or the substrate. Substrate analogues were used to screen a phage displayed library of mutants of staphylococcal nuclease (SNase), a Ca^{2+} dependent phosphodiesterase which cleaves DNA with preference for thymidine on the 5' side of the cleavage site (Scheme 5.1). The best mutant screened against a thymidine substrate analogue (Figure 5.2A) was almost as active as the wild type; those screened against a guanosine substrate analogue (Figure 5.2B) were ten time less active but displayed a change in specificity (Light *et al.* 1995).



Scheme 5.1: Hydrolysis of the DNA substrate by SNase.

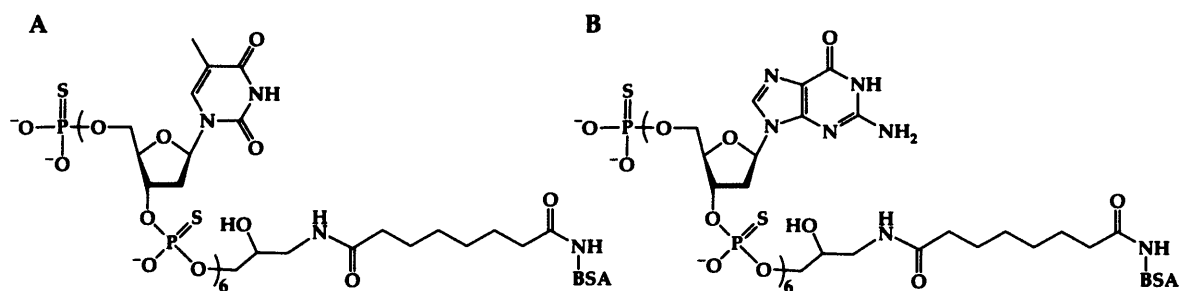
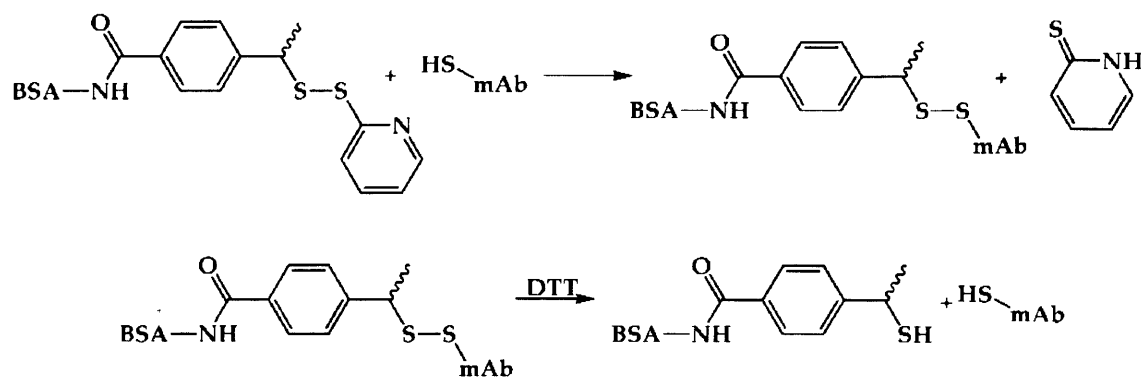


Figure 5.2: Thymidine-based (A) and guanosine-based (B) phosphorothionate substrate analogues used in panning selections. The substrate analogues were attached to BSA to facilitate immobilisation for panning (Light *et al.* 1995).

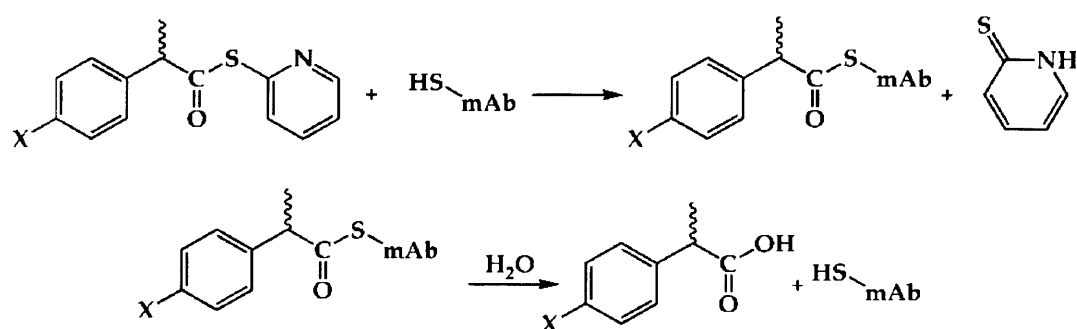
Molecules that are able to bind covalently with a side chain of a residue in the active site have been used to screen for active enzymes or abzymes (monoclonal antibody with catalytic activity). A screening molecule with a reactive cysteine, BSA-conjugated α -phenethyl pyridyl disulfide was used to bind phage-antibodies (Scheme 5.2). The substrate analogues were attached to BSA to facilitate immobilisation for panning. Two out of ten of the selected phage-antibodies contained an unpaired cysteine, one was able to catalyse the hydrolysis of a thioester whose electrophilic carbonyl occupies the position of the reactive sulphur during selection (Scheme 5.3) (Janda *et al.* 1994).

The recognition of the transition state theory has led to the design of transition-state analogues (TSA) as potent inhibitors of enzymes and to their use as haptens to induce the immune system into generating antibodies endowed with catalytic activity, named abzymes or catalytic antibodies. A library of carbonic anhydrase mutants, a metalloenzyme that catalyses the hydration of CO_2 , was screened for zinc-ion binding against an immobilised sulphonamide, a transition-state analogue inhibitor (Figure 5.3). Selected mutants had different grades of affinity for zinc, from equal to 100-fold lower

than the wild type, and 80% had CO₂ hydrase activity close to that of the wild type (Hunt *et al.* 1997).



Scheme 5.2: Mechanism of covalent binding of a BSA-conjugated disulfide to the active site of a monoclonal antibody.



X=H or HCONH

Scheme 5.3: One of the selected antibodies catalyses the hydrolysis of a thioester (Janda *et al.* 1994).

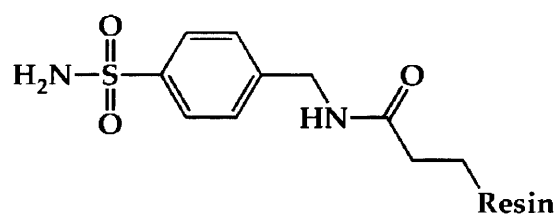


Figure 5.3: *p*-Aminomethyl benzenesulfonamide, a transition-state analogue, used to select mutants of carbonic anhydrase (Hunt *et al.* 1997).

Lipolytic abzymes were selected from a phage displayed library of mutants against a transition-state analogue of lipases/esterases (Figure 5.4). Four mutants, able to bind the TSA, were selected and their catalytic activity was measured (Leong *et al.* 2007).

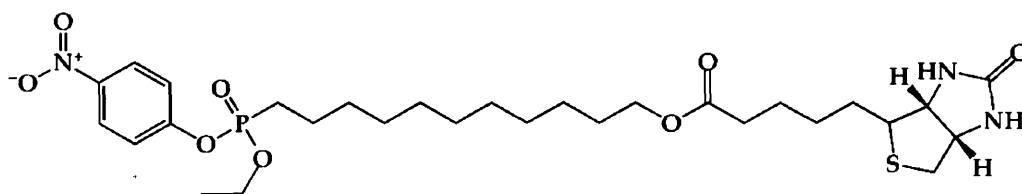


Figure 5.4: TSA of lipase, 5-(2-oxo-hexahydro-thieno[3,4-d]imidazol-6-yl)-pentanoic acid 11-[ethoxy-(4-nitro-phenoxy)-phosphoryl]-undecylester (Leong *et al.* 2007).

The success of the biopanning experiment can be followed by determining the number of phage eluted from the immobilised target protein versus a control without binding molecule by titring the recovered phage on a antibiotic medium (Rudgers *et al.* 1999). After each round of binding and enrichment, the number of phage eluted from the target protein should increase while the number of phage eluted from the control should stay approximately the same. This method is good for monitoring the success of a panning experiment, but it does not always hold true if the displayed peptides bind weakly to the target molecule. If after three rounds of binding and enrichment there is no increase in the number of phage eluted from the target protein versus the control, it is useful to check the sequence of the recovered phages. A comparison of the target protein and control sequence data should indicate if a specific type of sequence is being selected (Arnold *et al.* 2003).

5.1.4 Aim

This part of thesis was the preliminary work to design and construct a library of Im9 mutants using directed evolution and phage display technologies. The possibility of screening the library would also be studied using a transition state analogue of the hydrolysis of *p*-nitrophenol esters by biopanning, which would also involve the synthesis of this analogue.

5.2 Results and Discussion

5.2.1 Design of the mutations

As Im9-W74A/N78H showed a three-fold increase in the rate of hydrolysis of *p*-nitrophenyl acetate compared to wild type Im9 (Section 4.2.17), this mutant was chosen as a starting point for directed evolution to further improve the esterase activity. The NMR structure of wild type Im9 was used to analyse the region around the designed active site (Osborne *et al.* 1996). To improve the binding of the substrate and to allow the use of random cassette mutagenesis (Section 5.3.2), mutations were concentrated around the cavity created by the mutated alanine 74. Six positions were chosen: Thr 70; Val 71; Gln 73; Ala 74; Ala 77; and Lys 80 (Figure 5.5). Ala 74 was included to allow the formation of a cavity with a different shape.

For the randomisation of the gene in four of the six targeted positions (Thr 70; Gln 73; Ala 77; and Lys 80) the NNK codon doping strategy was used (where N = adenine, cytosine, guanine or thymine and K = guanine or thymine) which codes for all the twenty amino acids and a stop codon. For the remaining two targeted positions (Val 71 and Ala 74) the VNS codon doping strategy was

used (where V = adenine, cytosine, or guanine and S = cytosine or guanine) which codes for 16 amino acids with Trp, Phe, and Tyr as well as Cys and all stop codons excluded. The aromatic amino acids were excluded for these mutations because they will not create a cavity in position 74 and to avoid steric clash with the helix II. The number of possible different mutants, not including those containing stop codons, with this design is 4×10^7 ($20^4 \times 16^2$) and the number of possible codon variants is 6×10^8 .

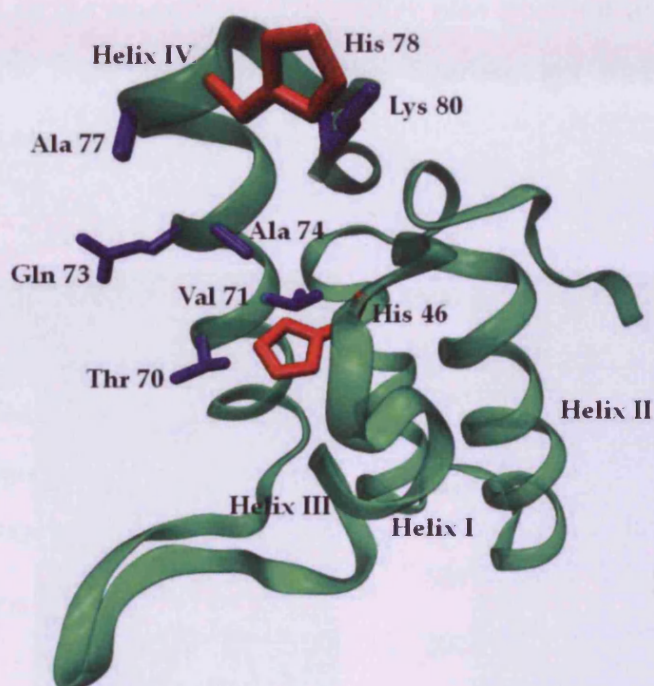


Figure 5.5: Cartoon representation of the NMR structure of Im9, with residues chosen for randomisation shown as blue sticks and catalytic histidine residues as red sticks.

5.2.2 Sub-cloning Im9 into pHEN2

The pET21 vector containing the gene encoding Im9-W74A/N78H (Section 4.2.14) was used as a starting point for the sub-cloning of this gene into a pHEN2 phagemid (gift from Dr. James Redman) to create a fusion with the N-

terminus of the pIII minor coating protein. pHEN2 encodes a hexa-histidine tag downstream from the inserted gene, which is then followed by an amber stop codon. This allows production of the fused protein without the pIII in non amber suppressor cells like *E. coli* BL21.

First, a *SacI* restriction site was inserted at the 3' terminus of the gene using site directed mutagenesis to allow transfer from the plasmid pET21 to the phagemid pHEN2. The Im9 gene was then excised from the pET21 vector using two consecutive digestion reactions, first with *NcoI* and then with *SacI*. pHEN2 was also digested in the same way. The DNA was purified after each reaction and the fragments were visualised using agarose gel electrophoresis and ethidium bromide staining (Figure 5.6).

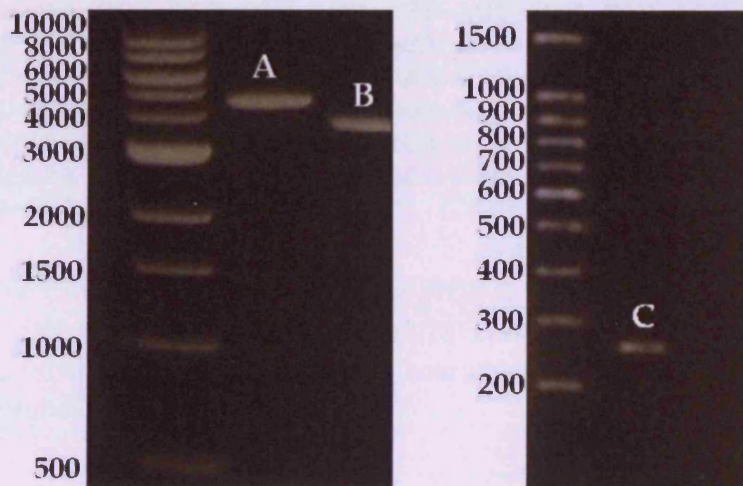
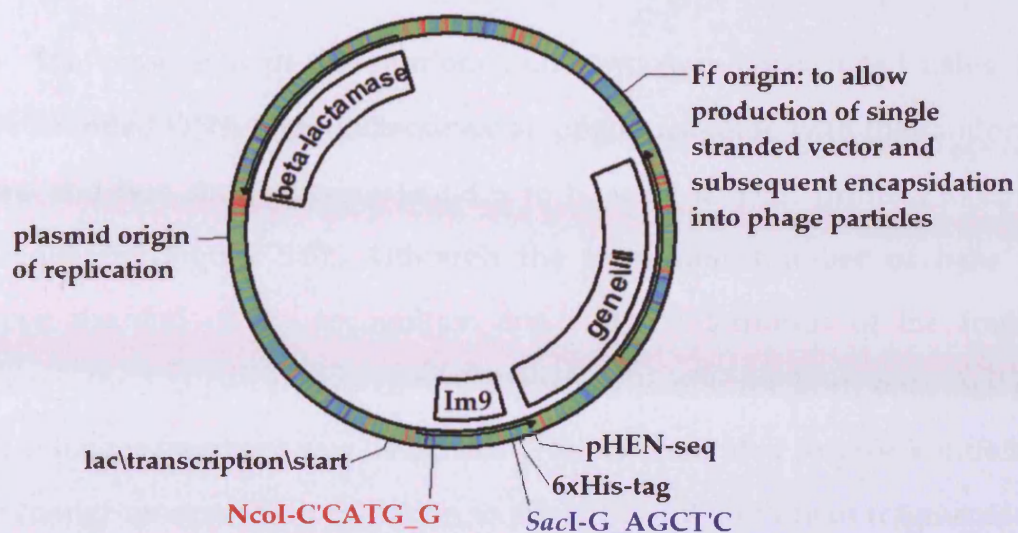


Figure 5.6: Left: 1kb DNA marker, doubly digested with *NcoI* and *SacI* (A) and intact (B) phagemid. Right: 100 bp DNA marker and excised Im9 gene (C). Numbers indicate base pairs per fragment in the marker lanes.

The linearised pHEN2 vector and the Im9 gene were ligated using T4 DNA Ligase. The reaction mixture was transformed into *E. coli* XL1-Blue

competent cells and the closed phagemid (Figure 5.7) was extracted and purified. The sequence was analysed to confirm the insertion.



CC-ATG-GAA-CTG-AAG-CAT-AGC-ATT-AGT-GAT-TAT-ACA-GAA-GCT-GAA-TTT-TTA-CAG-CTT-GTA-ACA-ACA-ATT-TGT-AAT-GCG-GAC-ACT-TCC-AGT-GAA-GAA-GAA-CTG-GTT-AAA-TTG-GTT-ACA-CAC-TTT-GAG-GAA-ATG-ACT-GAG-CAC-CCT-AGT-GGT-AGT-GAT-TTA-ATA-TAT-TAC-CCA-AAA-GAA-GGT-GAT-GAT-GAC-TCA-CCT-TCA-GGT-ACC-GTA-AAC-ACA-GTA-AAA-CAA-GCG-AGA-GCC-GCT-CAC-GGT-AAG-TCA-GGA-TTT-AAA-CAG-GGC-GAG-CTC

Figure 5.7 Top: Circular representation of the pHEN2 phagemid with the inserted Im9 gene. Bottom: DNA sequence of Im9-W74A/N78H, showing the NcoI restriction site (red), SacI restriction site (blue) and KpnI restriction site (green) used for random cassette mutagenesis.

5.2.3 Random cassette mutagenesis

As the designed random mutations were concentrated in a restricted area, the cassette mutagenesis technique was chosen (Botstein *et al.* 1985). This method consists of excising a DNA fragment and replacing it with an oligonucleotide carrying the desired mutations. Using site directed mutagenesis, isoleucine 67 (codon ATT), at the beginning of helix IV, was

mutated to threonine (codon ACC) in order to engineer a new restriction site, *KpnI*, in the middle of the *Im9*-W74A/N78 gene. The new *KpnI* restriction site and the *SacI* site at the end of the gene were used as termini for the fragment carrying the mutations.

The cassette with the random mutations was constructed using three single stranded DNA oligonucleotides: an oligonucleotide with the randomised codons and two short oligonucleotides to be used as PCR primers to amplify the sequence (Figure 5.8). Although the minimum number of base pairs between the end of the recognition site and the terminus of the fragment necessary for 100 percent successful digestion is one for both enzymes (NEB 2009), a longer fragment was designed. This was intended to give a sufficiently large change on restriction digestion to allow the cut and uncut fragments to be distinguished on an agarose gel.

```

GCG GCA TTA ACC TCA CCT TCA GGT ACC GTA AAC
          CC TCA CCT TCA GGT ACC GTA AAC NNK VNS
AAA NNK VNS AGA GCC NNK CAC GGT NNK TCA GGA TTC AAG
CAG GGC GAG CTC                                AAG TTC
GTC CCG CTC GAG TCG ACG GGG GCG

```

Figure 5.8: The single stranded oligonucleotide with the degenerate codons (black), with the forward (red) and reverse complementary (blue) primer for PCR.

5.2.4 Library preparation

The vector and the random cassette (after amplification and purification) were prepared using two consecutive digestion reactions, first with *KpnI* and then with *SacI*. The DNA was purified after each reaction and the fragments were visualised following agarose gel electrophoresis (Figure 5.9).

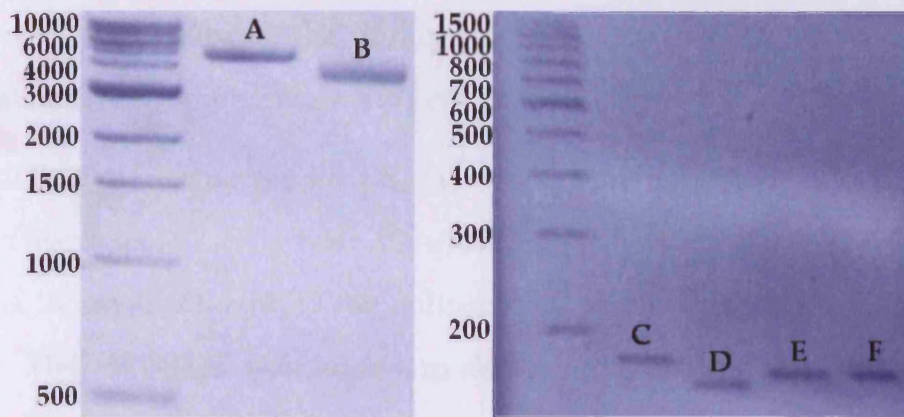


Figure 5.9: Left: 1kb DNA marker, double digested (A) and intact (B) phagemid pHEN2. Right: 100 bp DNA marker, uncut random insertion (C), double digested random insertion (D), random insertion single digested with *KpnI* (E) and random insertion single digested with *SacI* (F).

In order to determine the amount of DNA necessary to give a library complexity big enough to cover the all the possible mutants, and to find the optimal ligation conditions, a series of test ligation reactions was performed changing the ratio of vector and insert (1:4; 2:4; 1:1; 2:1; 4:1). A back self-ligation was performed with the double digested phagemid alone to measure the background reaction. The reaction products were transformed after heat inactivation of the T4 DNA ligase, by electroporation into *E. coli* XL1-Blue. The heat inactivation of the ligase is necessary to avoid inhibition of the transformation (Ymer 1991). The transformed cultures were plated and the number of transformants per μg of vector DNA in the more successful reaction (1:4) was calculated to be equal to 4.11×10^7 . The absence of colonies from the back self-ligation reaction indicated the absence of undigested or singly digested phagemid. The total amount of DNA required to cover the number of mutants was calculated by dividing the number of possible mutants (6.04×10^8 , Section 5.2.1) by the transformant per μg (4.11×10^7) and was equal to 15.1 μg of double digested vector. The web program GLUE (Firth *et al.* 2005), used to

analyse the complexity of the library, gave a completeness of 96% with an expected number of different amino acid of 3.9×10^7 .

10 ligation reactions with 1.51 μg vector were performed and *E. coli* XL1-Blue electrocompetent cells were transformed with the DNA from the ligation reactions. A small amount of the culture was plated to titer the transformed bacteria. Helper phage was added to the culture to start the production of phage and after overnight incubation the phage pellets were precipitated, resuspended in TBS and filtered. The freshly prepared phage library was used for panning.

5.2.5 Panning

The transition state analogues normally used for the selection of enzymes able to catalyse *p*-nitrophenyl esters are phosphates, phosphonates, phosphoramidates, and phosphinates and their esters, which are expected to mimic the charge delocalisation and tetrahedral geometry of the transition states (Figure 5.10) (Tantillo *et al.* 1999).

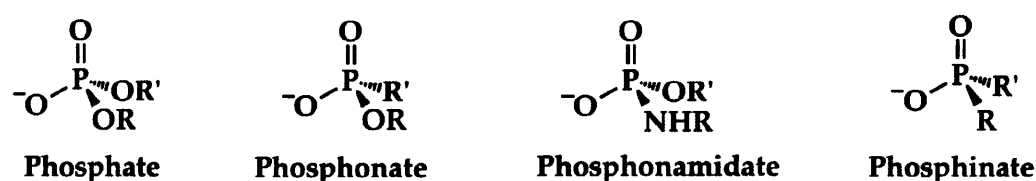


Figure 5.10: Transition-state analogues for *p*-nitrophenyl esters.

Several antibodies which catalyse the hydrolysis of aryl esters have been designed using such transition state analogues (Figure 5.11) (Golinelli-Pimpaneau *et al.* 1994; Wilmore *et al.* 1994; Zhou *et al.* 1994; Mu *et al.* 1997;

Tawfik *et al.* 1997; Wedemayer *et al.* 1997; Stephens *et al.* 1998; Odenbaugh *et al.* 2000; Gul *et al.* 2003; Sonkaria *et al.* 2004; Boucher *et al.* 2007; Zhang *et al.* 2007).

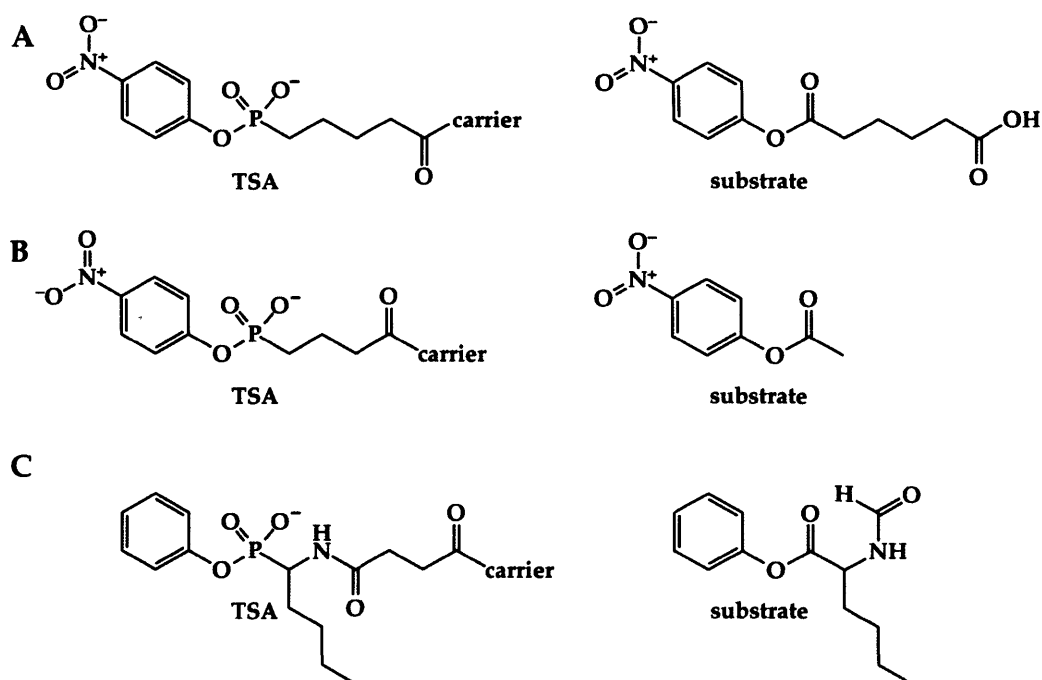
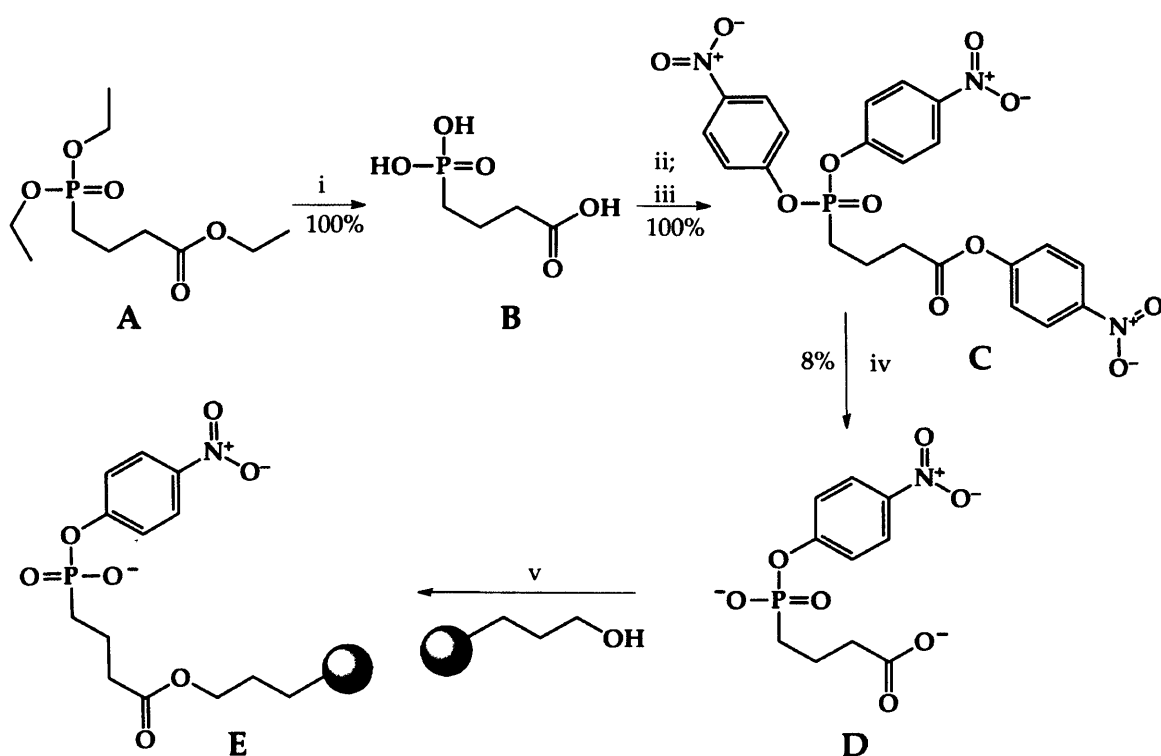


Figure 5.11: Transition state analogue and substrate of the abzymes 48G7 (A) (Wedemayer *et al.* 1997); CNJ206 (B) (Golinelli-Pimpaneau *et al.* 1994); 17E8 (C) (Zhou *et al.* 1994).

For the panning the phosphonate transition state analogue 4-[hydroxy(4-nitrophenoxy)phosphoryl]butanoic acid was synthesised and coupled to magnetic beads (Dynabeads[®]) (Scheme 5.4). The phosphonate 4-[hydroxy(4-nitrophenoxy)phosphoryl]butanoic acid has been successfully used to raise catalytic antibodies able to catalyse the hydrolysis of p-nitrophenyl esters (Tawfik *et al.* 1990)



Reagents and conditions: i) HCl conc; ii) SOCl_2 , DMF; iii) $p\text{-NO}_2\text{C}_6\text{H}_4\text{OH}$; iv) NaOH; v) DMF, DIC, DMAP

Scheme 5.4: Synthesis of the transition state analogue coupled with magnetic beads.

A mixture of ethyl 4-(diethoxyphosphoryl)butanoate **A** (10.6 mmol) and concentrated HCl (60 mL) was stirred under reflux for 15 h. The solvent was removed under reduced pressure and the residue was washed with dioxane (10 mL) and concentrated under reduced pressure to dryness. This procedure was repeated twice. The solid was then washed with toluene (10 mL) and dried under reduced pressure. This procedure was again repeated twice. The 4-phosphonobutanoic acid **B** was obtained as white solid (10.1 mmol) and was suspended in SOCl_2 (40.4 mmol) and DMF (200 μL) as a catalyst. The mixture was stirred at 55 $^\circ\text{C}$ for 15 h. After cooling, excess of SOCl_2 was removed under reduced pressure and crude acyl chloride was obtained as green oil (2.25 g).

Sodium *p*-nitrophenolate was prepared by mixing *p*-nitrophenol (40.4 mmol) dissolved in 40 mL of THF with NaH dispersed in paraffin liquid (40.4 mmol) with constant stirring at room temperature. Stirring was continued until evolution of H₂ ceased. A solution of the crude acyl chloride (10.1 mmol) in 40 mL of THF was gradually added with vigorous stirring to the *p*-nitrophenolate mixture at room temperature. The reaction mixture was stirred for 12 h and then poured into an ice-aqueous solution of HCl. The triester 4-nitrophenyl 4-(bis(4-nitrophenoxy)phosphoryl)butanoate **C** was extracted with ethyl acetate (3 x 50 mL). The combined organic layers were washed with brine, dried over MgSO₄ and concentrated under reduced pressure to dryness.

The triester **C** obtained as a brown powder (10.1 mmol) was stirred vigorously in 0.5 M aqueous NaOH solution (200 mL) at room temperature for 48 h. After removal of unreacted impurities by filtration, the filtrate was acidified with 0.5 M aqueous HCl solution and concentrated under reduced pressure to dryness. The residue was purified by RP-HPLC Luna C18 (25 x 10 mm, 5 μm) isocratic with 20% acetonitrile in 0.01% aqueous TFA solution (3.0 mL min⁻¹, retention time 14.6 min). The acetonitrile and TFA were removed under reduced pressure and the water was removed by lyophilisation to give 4-(hydroxy(4-nitrophenoxy)phosphoryl)butanoic acid **D** as a white solid with a final yield of 8% from **A**.

Tosyl-Dynabeads® (50 mg, 5-10 μmol) were stirred in CH₃CN (2.25 mL) and 10 μM aqueous NaOH (2.25 mL) at room temperature for 24 hours. The reaction was followed by analytical RP-HPLC Luna C18 (15 x 4.6 mm, 5 μm) isocratic with 20% acetonitrile in 0.01% aqueous TFA solution (1.0 mL min⁻¹) monitoring the release of the tosyl group (retention time 17 min). The HO-Dynabeads were collected and washed with H₂O.

The beads (30 mg, 3-6 μmol) were coupled with acid **D** (2 mg, 6 μmol) in DMF (5 mL) using *N,N'*-diisopropylcarbodiimide (1 μL , 7 μmol) and 4-dimethylaminopyridine (catalytic amount). The mixture was stirred at room temperature for 24 h. The ester **E** was collected and washed with water. The success of the reaction was tested by hydrolysing the final product with hydrochloric acid and checking the production of 4-nitrophenol by UV spectroscopy. Due to the variable substitution of the beads, a quantitative measure of the yield of the reaction was not possible.

The freshly prepared library was pooled and divided into three groups for panning with the phagemid containing the Im9-W74A/N78H gene as the control. The library was first pre-screened against Dynabeads[®] without the TSA to remove non-specific binders. The second round of screening used Dynabeads[®] with TSA bound. Unbound phages were removed with detergent (Tween 20). Bound phages were eluted under acidic conditions and after neutralisation were used to inoculate a new culture. Three rounds of panning were performed and the phages were titered after each round (Table 5.1).

Although there was a good enrichment in the second round, the third round gave no colonies in two of the three phage groups. Because there was no enrichment in the third round of panning, it was decided to sequence the DNA extracted from the phages. Three colonies were selected from each successful round of panning for a total of twenty-one mutants. Twelve of the twenty-one sequencing reaction failed. The nine successful sequences did not show enrichment for any particular sequence (Table 5.2).

Table 5.1: Panning results.

Groups	In/out 1 st round	In/out 2 nd round	In/out 3 rd round
1	$3.11 \cdot 10^{-7}$	$1.78 \cdot 10^{-5}$	No colonies
2	$2.05 \cdot 10^{-8}$	$1.12 \cdot 10^{-5}$	$1.25 \cdot 10^{-6}$
3	$1.05 \cdot 10^{-7}$	$9.11 \cdot 10^{-6}$	No colonies
Wild type	$1.51 \cdot 10^{-8}$	$1.26 \cdot 10^{-8}$	$1.63 \cdot 10^{-8}$

Table 5.2: List of mutations present in sequenced mutants.

position mutant	70	71	73	74	77	80
Im9- W74A/N78H	Thr	Val	Gln	Ala	Ala	Lys
group 1 round 1	Val	Leu	Tyr	Asn	Gln	Arg
group 1 round 1	Val	Ala	Ala	Arg	Gly	Phe
group 2 round 1	Met	Val	Cys	Ser	Trp	Thr
group 3 round 1	Asp	Thr	Ser	Ile	Val	Val
group 1 round 2	Pro	Val	Tyr	Leu	Asn	Ala
group 2 round 2	Arg	Ala	Asp	Gly	Ile	Val
group 3 round 2	Trp	Asp	Leu	Ala	Arg	Phe
group 3 round 2	Ser	His	Pro	Thr	Cys	Ser
group 1 round 3	Gly	Ile	Asn	Val	Stop	Ser

5.2.6 Test expression of Im9 mutants in pHEN2

The phagemid pHEN2 was constructed to allow the expression of the fused protein without the need to subclone. An amber stop codon (TAG) is situated between the pIII protein and the fused protein Im9. When amber suppressor strains like XL1-Blue are transformed with the phagemid translation

proceeds through the amber codon and the two proteins are expressed as one. When non amber suppressor cells are transformed with the phagemid the fused protein is expressed as a separate entity. A His-tag is present downstream from the protein sequence to facilitate the purification of the protein.

Im9 mutants were produced in BL21(DE3) cells (non amber suppressor). The protein of interest, present in the supernatant (Figure 5.12), was applied to a nickel column and eluted with an imidazole step gradient (from 20 mM; 40 mM; 60 mM NaCl to 1 M NaCl).

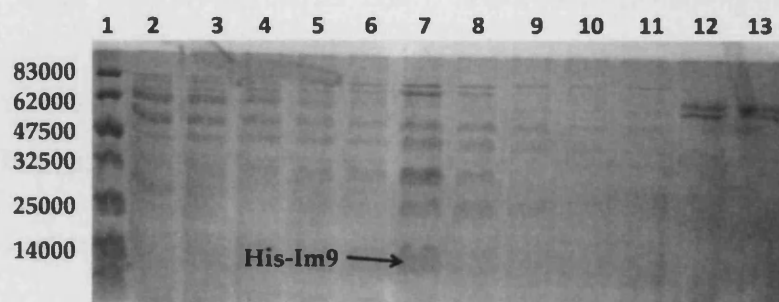


Figure 5.12: 12% SDS polyacrylamide gel from purification of His-Im9 with a Ni-column. Lane 1: broad range protein marker; lane 2- 5: fractions eluted with 20 mM imidazole; lane 6- 9; fractions eluted with 40 mM imidazole containing the His-Im9; lane 10- 13: fractions eluted with 60 mM imidazole.

The fractions 6-8 containing the His-Im9 mutant, identified by SDS-PAGE (Figure 5.13), were collected and Amicon ultra-filtration (MWCO 30 kDa) was used to remove the larger proteins leading to a pure product (Figure 5.13) as confirmed by MALDI mass spectrometry 12978 (calculated 12979.2).

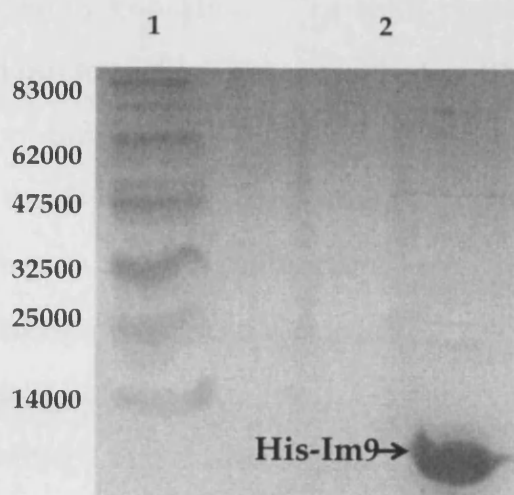


Figure 5.13: 12% SDS polyacrylamide gel from Amicon ultrafiltration (MWCO 30 kDa). Lane 1: broad range protein marker; lane 2: flow through containing the His-Im9.

5.3 Conclusions

The construction of a phage display library of Im9-W74A/N78H was accomplished. First step was the design of two restriction sites into Im9W74A/N78H gene to allow the subcloning to take place. Second step was the successful subcloning of Im9 mutant gene into a pHEN2 phagemid. This step was then followed by applying the cassette mutagenesis to introduce random mutations. The introduced random mutations were located around the active site of Im9. The success of library construction was confirmed by the absence of self-ligation of the open vector in the control ligation.

A test expression and purification of His-Im9 in pHEN2 phagemid demonstrates the possibility of expression and analysis of eventually successfully screened mutants without the necessity of a time consuming subcloning procedure. A transition state analogue was synthesised and coupled to magnetic beads for the screening of the library. The screening however did not lead to an enrichment of any particular sequence. One of the main reasons

could be non-optimal screening conditions. The solid-phase panning typically has a high capturing yield but low discrimination (Barbas III *et al.* 2004) and this can explain the presence of binder with bulky group in position 74 (see Table 5.2). The positive note is that binding mutants can be isolated. To improve the panning results a screening in solution should be included: with the “right” concentration of the screening molecule the capture of the clones with higher affinity will be maximised (Barbas III *et al.* 2004). For this purpose the screening molecule needs to be modified. The transition state analogue can be bound to biotin and, after capturing the binders from the solution, the complexes will be captured on immobilised streptavidin. The combination of the two screening methods, first a solid-phase screening and then a solution screening, would hopefully lead to a selection of only high affinity binders. Another reason for the poor results of the panning can be the library itself therefore different designs should be tried including a randomisation of the residues that are around the native histidine 46 (see Figure 5.6).

6. GENERAL CONCLUSION

This thesis focused on the ever-challenging topic of artificial enzyme design. A peptide, Apoxaldie and a protein, Im9 were chosen to initiate the work. Apoxaldie-Dab and 10 Im9 mutants were designed and prepared. Their secondary structures were characterised by CD and the catalytic reactions were studied. Two Im9 mutants showed improved activities. The work extended to the pioneering study of Im9 phage displayed library construction. The library together with the transition state analogue were successfully prepared. The mechanism of the designed artificial enzymes was discussed.

6.1 Apoxaldie-Dab

Apoxaldie-Dab, a new putative oxaloacetate decarboxylase based on Apoxaldie, was designed and synthesised mutating Lysine residues in the active site of Apoxaldie with Dab residues in order to bring the active amino groups nearer to the chiral environment of the α -helix backbone. The designed 18 aa peptide has two pair of cysteines and the formation of the native disulfide bonds has proved to be crucial for the peptide structure. Two synthetic strategies were explored: the former involved the synthesis of Apoxaldie-Dab primary sequence with 4 free cysteines following by direct oxidation; the latter strategy applied the synthesised peptide with two cysteine residues protected by AcM and the other two cysteines protected by Trt. The use of orthogonal protecting groups allowed to deprotect and to oxidise the two pair of cysteines sequentially. This method overcame the difficulty in the formation of the native disulfide bridges. The obtained peptide was then studied in the decarboxylation reaction using oxaloacetate as the substrate.

Apoxaldie-Dab's catalytic activity however was not observed in the assay. This can be explained by the helix breaking property of Dab, which

probably prevents the modulation of the pK_a of the amino residues, which is necessary to achieve a catalytic activity by destabilising the active site secondary structure. The deconvolution of the CD spectra supports this hypothesis, showing a decrease of helicity in Apoxaldie-Dab (26%) compared to Apoxaldie (36%).

The synthesis of Apoxaldie derived peptides using ornithine to substitute the lysine residues in the active site is suggested for the future work. As the experimental results with Apoxaldie-Dab indicating the helicity of the peptide could be the index of the activity, using ornithine residue, which has the intermediate side chain length and helix propensity between lysine and Dab, may solve the helix distortion problem with Dab, consequently lead to an active artificial peptide.

6.2 Im9

In order to design a catalyst for the hydrolysis of *p*-nitrophenyl acetate the colicin DNase inhibitor Im9 was used. It was chosen because of its known structure and well-studied folding mechanism. A histidine based active site was designed in a hydrophobic cleft in order to shield it from the solvent and achieve substrate selectivity.

Two series of 10 mutants were then expressed, purified and characterised. Together with Im9 wild type, the 11 proteins secondary structures and thermal stabilities were examined by CD and their catalytic properties were studied by UV spectroscopy following the hydrolysis of *p*-nitrophenol.

The three Im9-Y10H/F15A series mutants showed comparable activity to the wild type with the two single mutants Y10H and F15A slightly unstable and the double mutant Y10H/F15A almost completely unstructured even at 5 °C.

In the second series, the single mutant N78H had similar stability and catalytic activity as the WT, the other single mutant W74A was slightly less stable but had the same activity. The double mutant Im9-W74A/N78H retained the secondary structure and good stability with a three-fold increase in the hydrolysis of *p*-nitrophenyl acetate compared to Im9 wild type. Encouraged by this result, attempts to further increase the double mutant efficiency were carried out by tuning the pK_a of the active histidine residues *via* mutations of nearby amino acid residues. Two mutants with 3 mutation sites and two with 4 mutation sites were generated. Im9-L52D/W74A/N78H showed a doubled activity with half of the α -helicity comparing to WT. The rest of the mutants however demonstrated similar catalytic activity with great loss of the helicity. They are thermally unstable comparing to the WT.

The increased catalytic activity of Im9-W74A/N78H compared to the unstructured Im9-Y10H/F15A, which contains the same number of histidine residues, indicates that the spatial arrangement of the histidines plays an important role in the reaction. An attempt to trap the acetylated intermediate by monitoring the reaction using MALDI-MS showed an accumulation of acetyl residues on the protein. This reveals that other residues on the surface of the protein are able to remove an acetyl group from the substrate. Furthermore the increased acetyl accumulation for Im9-W74A/N78H, which was the most active among the tested mutants, may suggest that a single turnover event is happening with the consequent inactivation of the active site. If this is the case, further mutation will be necessary, in order to open a channel toward the active site, to allow water molecules to hydrolyse the acetyl-histidine intermediate.

This additional mutation however would probably further destabilise the mutants, as shown by the set of mutants attempting to tune the pK_a . Therefore, prior to test, the intrinsic stability of Im9 needs to be improved.

6.3 Phage Display

As Im9-W74A/N78H showed increased catalytic activity, the work of applying directed evolution and phage display to facilitate the design of artificial enzyme was also initiated aiming to construct a phage display library of Im9-W74A/N78H. This was achieved by first designing two restriction sites into Im9W74A/N78H gene to allow the subcloning to take place, then subcloning of Im9 mutant gene into a pHEN2 phagemid, following by applying the cassette mutagenesis using designed inserts. The introduced random mutations were located around the active site of Im9. The success of library construction was confirmed by the absence of self-ligation of the open vector in the control ligation. The phagemid pHEN2 allows production of the displayed proteins in non amber suppressor strains. The displayed proteins are expressed with a His-tag that facilitates purification; therefore the selected mutants can be tested without the necessity of subcloning the gene in another vector. A test expression of His-tagged Im9-W74A/N78H was successfully performed and the protein was purified in a two-step procedure.

A transition state analogue phosphonate was also synthesised and immobilised to magnetic beads to facilitate solid phase the screening of the library. Test screening of the constructed library against the synthesised solid phase immobilised transition state analogue, although leading to no enrichment in any particular sequence, shows that binding mutants can be easily isolated.

This suggests that the detection of mutants with high affinity for the transition state analogue could be possible by adjusting the screening conditions.

For future work, solution screening is suggested to maximise the selection of high affinity binders as the solid phase screening has a high capturing yield but low discrimination toward high affinity binders. Different designed libraries should also be tested to increase the possibility of finding an active mutant.

7. REFERENCES

- Ahmad, F. and C. C. Bigelow (1979). "The denaturation of ribonuclease A by combinations of urea and salt denaturants." *J. Mol. Biol.* **131**: 607.
- Aili, D., K. Enander, L. Baltzer and B. Liedberg (2007). "Synthetic de novo designed polypeptides for control of nanoparticle assembly and biosensing." *Biochem. Soc. Trans.* **35**: 532.
- Allemann, R. K. (1989). Evolutionary guidance as a tool in organic chemistry. Ph. D. thesis, Zurich, Swiss Federal Institute of Technology.
- Anderson, H., C. W. Su and J. W. Watson (1969). "Aminolysis reactions. I. Mechanism of aminolysis and amidinolysis of *p*-nitrophenyl acetate in chlorobenzene." *J. Am. Chem. Soc.* **91**: 482.
- Arnold, F. H. and G. Georgiou (2003). Directed enzyme evolution: screening and selection methods. Springer, Düsseldorf.
- Atassi, M. Z. and T. Manshour (1993). "Design of peptide enzymes (pepzymes): surface-simulation synthetic peptides that mimic the chymotrypsin and trypsin active sites exhibit the activity and specificity of the respective enzyme." *Proc. Natl. Acad. Sci. USA* **90**: 8282.
- Atkins, P. W. (2006). Physical Chemistry 8th ed. Oxford University Press, Oxford, Melbourne, Tokyo.
- Barbas III, C. F., D. R. Burton, J. K. Scott and G. J. Silverman (2004). Phage display: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
- Bender, M. L. and B. W. Turnquest (1957). "The Imidazole-catalyzed Hydrolysis of *p*-Nitrophenyl Acetate 1." *J. Am. Chem. Soc.* **79**: 1652.

- Bennett, J. and E. A. Wren (1977). "The interaction of *p*-nitrophenyl carbamate with urease." *Biochim. Biophys. Acta* **482**: 421.
- Besenmatter, W., P. Kast and D. Hilvert (2006). "Relative tolerance of mesostable and thermostable protein homologs to extensive mutation." *Proteins: Struct. Funct. Bioinform.* **66**: 500.
- Betz, S., R. Fairman, K. O'Neil, J. Lear and W. Degrado (1995). "Design of two-stranded and three-stranded coiled-coil peptides." *Philos. Trans. R. Soc. Lond., Ser. B: Biol. Sci.* **348**: 81.
- Blanc, E., R. Romi-Lebrun, O. Bornet, T. Nakajima and H. Darbon (1998). "Solution structure of two new toxins from the venom of the Chinese scorpion *Buthus martensi Karsch* blockers of potassium channels." *Biochemistry* **37**: 12412.
- Blank, J. T. and S. J. Miller (2006). "Studies of folded peptide-based catalysts for asymmetric organic synthesis." *Biopolymers (Peptide Sci.)* **84**: 38.
- Blundell, T. L., J. E. Pitts, I. J. Tickle, S. P. Wood and C. W. Wu (1981). "X-ray analysis (1.4-Å resolution) of avian pancreatic polypeptide: Small globular protein hormone." *Proc. Natl. Acad. Sci. USA* **78**: 4175.
- Böhm, G. (1997). CDNN CD Spectra Deconvolution Software, Version.
- Bolon, D. N. and S. L. Mayo (2001). "Enzyme-like proteins by computational design." *Proc. Natl. Acad. Sci. USA* **98**: 14274.
- Bornscheuer, U. T. and R. J. Kazlauskas (2004). "Catalytic promiscuity in biocatalysis: using old enzymes to form new bonds and follow new pathways." *Angew. Chem. Int. Ed.* **43**: 6032.
- Botstein, D. and D. Shortle (1985). "Strategies and applications of in vitro mutagenesis." *Science* **229**: 1193.

- Boucher, G., B. Said, E. L. Ostler, M. Resmini, K. Brocklehurst and G. Gallacher (2007). "Evidence that the mechanism of antibody-catalysed hydrolysis of arylcarbamates can be determined by the structure of the immunogen used to elicit the catalytic antibody." *Biochem. J.* **401**: 721.
- Brakmann, S. and K. Johnsson (2002). Directed Molecular Evolution of Proteins: or How to Improve Enzymes for Biocatalysis. Wiley, New York.
- Brazil, B. T., J. L. Cleland, R. S. McDowell, N. J. Skelton, K. Paris and P. M. Horowitz (1997). "Model peptide studies demonstrate that amphipathic secondary structures can be recognized by the chaperonin GroEL (cpn60)." *J. Biol. Chem.* **272**: 5105.
- Breslow, R. (1972). "Centenary lecture. Biomimetic Chemistry." *Chem. Soc. Rev.* **1**: 553.
- Breslow, R. (2005). Artificial enzymes. Wiley, Weinheim.
- Breslow, R. and W. H. Chapman (1996a). "On the mechanism of action of ribonuclease A: relevance of enzymatic studies with a *p*-nitrophenylphosphate ester and a thiophosphate ester." *Proc. Natl. Acad. Sci. USA* **93**: 10018.
- Breslow, R., S. D. Dong, Y. Webb and R. Xu (1996b). "Further studies on the buffer-catalyzed cleavage and isomerization of uridyluridine. Medium and ionic strength effects on catalysis by morpholine, imidazole, and acetate buffers help clarify the mechanisms involved and their relationship to the mechanism used by the enzyme ribonuclease and by a ribonuclease mimic." *J. Am. Chem. Soc.* **118**: 6588.
- Breslow, R. and E. Kool (1988). "A α -cyclodextrin thiazolium salt holoenzyme mimic for the benzoin condensation." *Tetrahedron Lett.* **29**: 1635.

- Breslow, R. and C. Schmuck (1996c). "Goodness of fit in complexes between substrates and ribonuclease mimics: Effects on binding, catalytic rate constants, and regiochemistry." *J. Am. Chem. Soc.* **118**: 6601.
- Briggs, G. E. and J. B. S. Haldane (1925). "A note on the kinetics of enzyme action." *Biochem. J.* **19**: 338.
- Broo, K. S., L. Brive, P. Ahlberg and L. Baltzer (1997). "Catalysis of hydrolysis and transesterification reactions of *p*-nitrophenyl esters by a designed helix-loop-helix dimer." *J. Am. Chem. Soc.* **119**: 11362.
- Broo, K. S., H. Nilsson, J. Nilsson and L. Baltzer (1998). "Substrate recognition and saturation kinetics in de novo designed histidine-based four-helix bundle catalysts." *J. Am. Chem. Soc.* **120**: 10287.
- Bruice, T. C. and F. C. Lightstone (1998). "Ground State and Transition State Contributions to the Rates of Intramolecular and Enzymatic Reactions." *Acc. Chem. Res.* **32**: 127.
- Bryson, J. W., S. F. Betz, H. S. Lu, D. J. Suich, H. X. Zhou, K. T. O'Neil and W. F. DeGrado (1995). "Protein design: a hierarchic approach." *Science* **270**: 935.
- Burkhard, P., M. Meier and A. Lustig (2000). "Design of a minimal protein oligomerization domain by a structural approach." *Protein Sci.* **9**: 2294.
- Cammers-Goodwin, A., T. J. Allen, S. L. Oslick, K. F. McClure, J. H. Lee and D. S. Kemp (1996). "Mechanism of stabilization of helical conformations of polypeptides by water containing trifluoroethanol." *J. Am. Chem. Soc.* **118**: 3082.
- Cecil, R. and J. R. McPhee (1959). "The sulfur chemistry of proteins." *Adv. Protein Chem.* **14**: 255.

- Chagot, B., C. Pimentel, L. Dai, J. Pil, J. Tytgat, T. Nakajima, G. Corzo, H. Darbon and G. Ferrat (2005). "An unusual fold for potassium channel blockers: NMR structure of three toxins from the scorpion *Opisthacanthus madagascariensis*." *Biochem. J.* **388**: 263.
- Chang, C. T., C. S. Wu and J. T. Yang (1978). "Circular dichroic analysis of protein conformation: inclusion of the beta-turns." *Anal. Biochem.* **91**: 13.
- Charbonnier, J. B., B. Golinelli-Pimpaneau, B. Gigant, D. S. Tawfik, R. Chap, D. G. Schindler, S. H. Kim, B. S. Green, Z. Eshhar and M. Knossow (1997). "Structural convergence in the active sites of a family of catalytic antibodies." *Science* **275**: 1140.
- Chau, M. H. and J. W. Nelson (1992). "Cooperative disulfide bond formation in apamin." *Biochemistry* **31**: 4445.
- Chin, J. W. and A. Schepartz (2001). "Concerted evolution of structure and function in a miniature protein." *J. Am. Chem. Soc.* **123**: 2929.
- Chou, P. Y. and G. D. Fasman (1974). "Conformational parameters for amino acids in α -helical, β -sheet, and random coil regions calculated from proteins." *Biochemistry* **13**: 211.
- Cohen, J. (2001). "How DNA Shuffling Works." *Science* **293**: 237.
- Colonna, S., H. Molinari, S. Banfi, S. Juliá, J. Masana and A. Alvarez (1983). "Synthetic enzymes--4 : Highly enantioselective epoxidation by means of polyaminoacids in a triphase system: influence of structural variations within the catalysts." *Tetrahedron* **39**: 1635.
- Copeland, R. A. (2000). Enzymes: a practical introduction to structure, mechanism, and data analysis. Wiley-VCH, New York.

- Copley, S. D. (2003). "Enzymes with extra talents: moonlighting functions and catalytic promiscuity." *Curr. Opin. Chem. Biol.* **7**: 265.
- Corey, D. R. and M. A. Phillips (1994). "Cyclic peptides as proteases: a reevaluation." *Proc. Natl. Acad. Sci. USA* **91**: 4106.
- Cram, D. J. (1983). "Cavitands: organic hosts with enforced cavities." *Science* **219**: 1177.
- Cranz-Mileva, S., C. T. Friel and S. E. Radford (2005). "Helix stability and hydrophobicity in the folding mechanism of the bacterial immunity protein Im9." *Protein Eng. Des. Sel.* **18**: 41.
- Crick, F. H. C. (1953). "The packing of α -helices: simple coiled-coils." *Acta Crystallogr., Sect D: Biol. Crystallogr.* **6**: 689.
- Dahiyki, B. I. (1996). "Protein design automation." *Protein Sci.* **5**: 895.
- Dai, Q. H., C. Tommos, E. J. Fuentes, M. R. Blomberg, P. L. Dutton and A. J. Wand (2002). "Structure of a de novo designed protein model of radical enzymes." *J. Am. Chem. Soc.* **124**: 10952.
- Daly, N. L. and D. J. Craik (2009). "Structural studies of conotoxins." *IUBMB Life* **61**: 144.
- Damborsky, J. and J. Brezovsky (2009). "Computational tools for designing and engineering biocatalysts." *Curr. Opin. Chem. Biol.* **13**: 26.
- Dauplais, M., B. Gilquin, L. D. Possani, G. Gurrola-Briones, C. Roumestand and A. Menez (1995). "Determination of the three-dimensional solution structure of noxiustoxin: analysis of structural differences with related short-chain scorpion toxins." *Biochemistry* **34**: 16563.

- De Caro, J. D., M. P. Chautan, P. Rouimi and M. Rovero (1988). "Acetylation of Lys-373 in porcine pancreatic lipase after reaction of the enzyme or its C-terminal with *p*-nitrophenyl acetate." *Biochimie* **70**: 1785.
- Del Valle, E. M. (2004). "Cyclodextrins and their uses: a review." *Process Biochem.* **39**: 1033.
- DeMan, J. M. (1999). Principles of food chemistry. Aspen Pub, Westport.
- Di Costanzo, L., H. Wade, S. Geremia, L. Randaccio, V. Pavone, W. F. DeGrado and A. Lombardi (2001). "Toward the de novo design of a catalytically active helix bundle: a substrate-accessible carboxylate-bridged dinuclear metal center." *J. Am. Chem. Soc.* **123**: 12749.
- Doig, A. J. (2008). "Stability and design of alpha-helical peptides." *Prog. Mol. Biol. Transl. Sci.* **83**: 1.
- Drauz, K., H. Waldmann and S. M. Roberts (2002). Enzyme catalysis in organic synthesis: a comprehensive handbook. Wiley-VCH Weinheim.
- Dufour, E., A. C. Storer and R. Menard (1995). "Engineering nitrile hydratase activity into a cysteine protease by a single mutation." *Biochemistry* **34**: 16382.
- Dunn, I. S. (1996). "Phage display of proteins." *Curr. Opin. Biotechnol.* **7**: 547.
- Dwyer, M. A., L. L. Looger and H. W. Hellinga (2004). "Computational design of a biologically active enzyme." *Science* **304**: 1967.
- Dwyer, M. A., L. L. Looger and H. W. Hellinga (2008). "Retraction." *Science* **319**: 569.
- Dzubiella, J. (2008). "Salt-specific stability and denaturation of a short salt-bridge-forming alpha-helix." *J. Am. Chem. Soc.* **130**: 14000.

- Eigen, M. and W. Gardiner (1984). "Evolutionary molecular engineering based on RNA replication." *Pure Appl. Chem.* **56**: 967.
- Eisenberg, D., W. Wilcox, S. M. Eshita, P. M. Pryciak, S. P. Ho and W. F. DeGrado (1986). "The design, synthesis, and crystallization of an alpha-helical peptide." *Proteins* **1**: 16.
- Endemann, H. and P. Model (1995). "Location of Filamentous Phage Minor Coat Proteins in Phage and in Infected Cells." *J. Mol. Biol.* **250**: 496.
- Epanand, R. M. (1993). The Amphipathic Helix. CRC Press, Boca Raton.
- Errington, N. and A. J. Doig (2005). "A phosphoserine-lysine salt bridge within an alpha-helical peptide, the strongest alpha-helix side-chain interaction measured to date." *Biochemistry* **44**: 7553.
- Fainzilber, M., T. Nakamura, A. Gaathon, J. C. Lodder, K. S. Kits, A. L. Burlingame and E. Zlotkin (1995). "A new cysteine framework in sodium channel blocking conotoxins." *Biochemistry* **34**: 8649.
- Farinas, E. T., T. Bulter and F. H. Arnold (2001). "Directed enzyme evolution." *Curr. Opin. Biotechnol.* **12**: 545.
- Ferguson, N., A. P. Capaldi, R. James, C. Kleanthous and S. E. Radford (1999). "Rapid folding with and without populated intermediates in the homologous four-helix proteins Im7 and Im9." *J. Mol. Biol.* **286**: 1597.
- Fersht, A. (1984). Enzyme structure and mechanism. W. H. Freeman and Company, New York.
- Fersht, A. R., A. Matouschek and L. Serrano (1992). "The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding." *J. Mol. Biol.* **224**: 771.

- Firth, A. E. and W. M. Patrick (2005). "Statistics of protein library construction." *Bioinformatics* **21**: 3314.
- Fletcher, D. I., C. R. Ganellin, A. Piergentili, P. M. Dunn and D. H. Jenkinson (2007). "Synthesis and pharmacological testing of polyaminoquinolines as blockers of the apamin-sensitive Ca²⁺-activated K⁺ channel (SKCa)." *Biorg. Med. Chem.* **15**: 5457.
- Fox, R. J., S. C. Davis, E. C. Mundorff, L. M. Newman, V. Gavrilovic, S. K. Ma, L. M. Chung, C. Ching, S. Tam and S. Muley (2007). "Improving catalytic function by ProSAR-driven enzyme evolution." *Nat. Biotechnol.* **25**: 338.
- Freeman, L. C. (1977). "A set of measures of centrality based on betweenness." *Sociometry* **40**: 35.
- Friel, C. T., A. P. Capaldi and S. E. Radford (2003). "Structural analysis of the rate-limiting transition states in the folding of lm7 and lm9: Similarities and differences in the folding of homologous proteins." *J. Mol. Biol.* **326**: 293.
- Fuller, E., B. R. Green, P. Catlin, O. Buczek, J. S. Nielsen, B. M. Olivera and G. Bulaj (2005). "Oxidative folding of conotoxins sharing an identical disulfide bridging framework." *FEBS J.* **272**: 1727.
- Galperin, M. Y., A. Bairoch and E. V. Koonin (1998). "A superfamily of metalloenzymes unifies phosphopentomutase and cofactor-independent phosphoglycerate mutase with alkaline phosphatases and sulfatases." *Protein Sci.* **7**: 1829.
- Gandour, R. D. and R. L. Schowen (1978). Transition states of biochemical processes. Plenum Press, New York.

- Gans, P. J., P. C. Lyu, M. C. Manning, R. W. Woody and N. R. Kallenbach (1991). "The helix-coil transition in heterogeneous peptides with specific side-chain interactions: theory and comparison with CD spectral data." *Biopolymers* **31**: 1605.
- Gasteiger, E., C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel and A. Bairoch (2005). Protein identification and analysis tools on the ExPASy server.
- Gebhardt, K., V. Lauvrak, E. Babaie, V. Eijsink and B. H. Lindqvist (1996). "Adhesive peptides selected by phage display: characterization, applications and similarities with fibrinogen." *Pept. Res.* **9**: 269.
- Gerlt, J. A. and P. C. Babbitt (2009). "Enzyme (re) design: lessons from natural evolution and computation." *Curr. Opin. Chem. Biol.* **13**: 10.
- Gilson, M. K. and B. Honig (1989). "Destabilization of an alpha-helix-bundle protein by helix dipoles." *Proc. Natl. Acad. Sci. USA* **86**: 1524.
- Golinelli-Pimpaneau, B., B. Gigant, T. Bizebard, J. Navaza, P. Saludjian, R. Zemel, D. S. Tawfik, Z. Eshhar, B. S. Green and M. Knossow (1994). "Crystal structure of a catalytic antibody Fab with esterase-like activity." *Structure* **2**: 175.
- Goodman, E. M. and P. S. Kim (1989). "Folding of a peptide corresponding to the alpha-helix in bovine pancreatic trypsin inhibitor." *Biochemistry* **28**: 4343.
- Goodman, M., A. S. Verdini, C. Toniolo, W. D. Phillips and F. A. Bovey (1969). "Sensitive criteria for the critical size for helix formation in oligopeptides." *Proc. Natl. Acad. Sci. USA* **64**: 444.

- Gorski, S. A., A. P. Capaldi, C. Kleanthous and S. E. Radford (2001). "Acidic conditions stabilise intermediates populated during the folding of Im7 and Im9." *J. Mol. Biol.* **312**: 849.
- Goto, Y., N. Takahashi and A. L. Fink (1990). "Mechanism of acid-induced folding of proteins." *Biochemistry* **29**: 3480.
- Greenwood, J., A. E. Willis and R. N. Perham (1991). "Multiple display of foreign peptides on a filamentous bacteriophage. Peptides from Plasmodium falciparum circumsporozoite protein as antigens." *J. Mol. Biol.* **220**: 821.
- Gul, S., S. Sonkaria, S. Pinitglang, J. Florez-Alvarez, S. Hussain, E. W. Thomas, E. L. Ostler, G. Gallacher, M. Resmini and K. Brocklehurst (2003). "Improvement in hydrolytic antibody activity by change in haptenic structure from phosphate to phosphonate with retention of a common leaving-group determinant: evidence for the 'flexibility' hypothesis." *Biochem. J.* **376**: 813.
- Gutte, B., M. Däumigen and E. Wittschieber (1979). "Design, synthesis and characterisation of a 34-residue polypeptide that interacts with nucleic acids." *Nature* **281**: 650.
- Halai, R. and D. J. Craik (2009). "Conotoxins: natural product drug leads." *Nat. Prod. Rep.* **26**: 526.
- Hamada, D., M. Hoshino, M. Kataoka, A. L. Fink and Y. Goto (1993). "Intermediate conformational states of apocytochrome c." *Biochemistry* **32**: 10351.
- Han, Y. H., Q. Wang, H. Jiang, L. Liu, C. Xiao, D. D. Yuan, X. X. Shao, Q. Y. Dai, J. S. Cheng and C. W. Chi (2006). "Characterization of novel M-superfamily conotoxins with new disulfide linkage." *FEBS J.* **273**: 4972.

- Hardiman, E., M. Gibbs, R. Reeves and P. Bergquist (2010). "Directed Evolution of a Thermophilic beta-glucosidase for Cellulosic Bioethanol Production." *Appl. Biochem. Biotechnol.* **161**: 301.
- Hatano, M. and M. Yoneyama (1970). "Side-chain effect on the helix stability of poly-alpha-amino acids." *J. Am. Chem. Soc.* **92**: 1392.
- Hellinga, H. W. and F. M. Richards (1991). "Construction of new ligand binding sites in proteins of known structure. I. Computer-aided modeling of sites with pre-defined geometry." *J. Mol. Biol.* **222**: 763.
- Henri, V. (1903). Lois générales de l'action des diastases. Librairie Scientifique A. Hermann.
- Hilvert, D. (2000). "Critical analysis of antibody catalysis." *Annu. Rev. Biochem.* **69**: 751.
- Hilvert, D., S. H. Carpenter, K. D. Nared and M. T. Auditor (1988a). "Catalysis of concerted reactions by antibodies: the Claisen rearrangement." *Proc. Natl. Acad. Sci. USA* **85**: 4953.
- Hilvert, D. and K. D. Nared (1988b). "Stereospecific Claisen rearrangement catalyzed by an antibody." *J. Am. Chem. Soc.* **110**: 5593.
- Hirst, J. D. and C. L. Brooks 3rd (1994). "Helicity, circular dichroism and molecular dynamics of proteins." *J. Mol. Biol.* **243**: 173.
- Hollfelder, F., A. J. Kirby and D. S. Tawfik (1996). "Off-the-shelf proteins that rival tailor-made antibodies as catalysts." *Nature* **383**: 60.
- Hollfelder, F., A. J. Kirby and D. S. Tawfik (2001). "On the magnitude and specificity of medium effects in enzyme-like catalysts for proton transfer." *J. Org. Chem.* **66**: 5866.

- Huang, J. F., Y. M. Xu, D. M. Hao, Y. B. Huang, Y. Liu and Y. X. Chen (2010). "Structure-guided de novo design of α -helical antimicrobial peptide with enhanced specificity." *Pure Appl. Chem.* **82**: 243.
- Hult, K. and P. Berglund (2007). "Enzyme promiscuity: mechanism and applications." *Trends Biotechnol.* **25**: 231.
- Hunt, J. A. and C. A. Fierke (1997). "Selection of Carbonic Anhydrase Variants Displayed on Phage." *J. Biol. Chem.* **272**: 20364.
- Huttunen-Hennelly, H. E. K. and J. C. Sherman (2007). "The design, synthesis, and characterization of the first cavitand-based de novo hetero-template-assembled synthetic proteins (Hetero-TASPs)." *OBC* **5**: 3637.
- Huyghues-Despointes, B. M. and J. W. Nelson (1992). "Stabilities of disulfide bond intermediates in the folding of apamin." *Biochemistry* **31**: 1476.
- Jacob, R. B. and O. M. McDougal (2010). "The M-superfamily of conotoxins: a review." *Cell. Mol. Life Sci.* **67**: 17.
- James, R., R. Wallis, G. R. Moore and C. Kleanthous (1993). "Molecular analysis of the interaction between the DNase Colicin E9 and its inhibitor protein Im9." *J. Cell. Biochem.*: 175.
- Janda, K. D., C. H. Lo, T. Li, C. F. Barbas, P. Wirsching and R. A. Lerner (1994). "Direct selection for a catalytic mechanism from combinatorial antibody libraries." *Proc. Natl. Acad. Sci. USA* **91**: 2532.
- Jasanoff, A. and A. R. Fersht (1994). "Quantitative determination of helical propensities from trifluoroethanol titration curves." *Biochemistry* **33**: 2129.
- Jiang, L., E. A. Althoff, F. R. Clemente, L. Doyle, D. Rothlisberger, A. Zanghellini, J. L. Gallaher, J. L. Betker, F. Tanaka and C. F. Barbas Iii

- (2008). "De novo computational design of retro-aldol enzymes." *Science* **319**: 1387.
- Jimenez, L. and F. Diederich (1989). "Catalytic cyclophanes: A highly efficient model for pyruvate oxidase." *Tetrahedron Lett.* **30**: 2759.
- Johnsson, K., R. K. Allemann, H. Widmer and S. A. Benner (1993). "Synthesis, structure and activity of artificial, rationally designed catalytic polypeptides." *Nature* **365**: 530.
- Jones, J. (1992). Amino Acid and Peptide Synthesis. University Press, Oxford.
- Joo, H., Z. Lin and F. H. Arnold (1999). "Laboratory evolution of peroxide-mediated cytochrome P450 hydroxylation." *Nature* **399**: 670.
- Juraja, S. M., T. D. Mulhern, P. J. Hudson, M. K. Hattarki, J. A. Carmichael and S. D. Nuttall (2006). "Engineering of the Escherichia coli Im7 immunity protein as a loop display scaffold." *Protein Eng. Des. Sel.* **19**: 231.
- Kang, T. S., Z. Radic, T. T. Talley, S. D. Jois, P. Taylor and R. M. Kini (2007). "Protein folding determinants: structural features determining alternative disulfide pairing in alpha- and chi/lambda-conotoxins." *Biochemistry* **46**: 3338.
- Kelly, D. R., T. T. T. Bui, E. Caroff, A. F. Drake and S. M. Roberts (2004). "Structure and catalytic activity of some soluble polyethylene glycol-peptide conjugates." *Tetrahedron Lett.* **45**: 3885.
- Kelly, S. M. and N. C. Price (1997). "The application of circular dichroism to studies of protein folding and unfolding." *Biochim. Biophys. Acta* **1338**: 161.

- Kohn, W. D. and R. S. Hodges (1998). "De novo design of [alpha]-helical coiled coils and bundles: models for the development of protein-design principles." *Trends Biotechnol.* **16**: 379.
- Kraemer-Pecore, C. M., J. T. Lecomte and J. R. Desjarlais (2003). "A de novo redesign of the WW domain." *Protein Sci.* **12**: 2194.
- Krumpe, L. R. H. and T. Mori (2006). "The use of phage-displayed peptide libraries to develop tumor-targeting drugs." *Int. J. Pept. Protein Res.* **12**: 79.
- Landon, C., P. Sodano, B. Cornet, J. M. Bonmatin, C. Kopeyan, H. Rochat, F. Vovelle and M. Ptak (1997). "Refined solution structure of the anti-mammal and anti-insect LqqIII scorpion toxin: comparison with other scorpion toxins." *Proteins* **28**: 360.
- Leberman, R. and A. K. Soper (1995). "Effect of high salt concentrations on water structure." *Nature* **378**: 364.
- Lecomte, C., J. M. Sabatier, J. Van Rietschoten and H. Rochat (1998). "Synthetic peptides as tools to investigate the structure and pharmacology of potassium channel-acting short-chain scorpion toxins." *Biochimie* **80**: 151.
- Leong, M. K., C. Chen, K. C. Shar and D. Shiuan (2007). "Selection and characterization of lipase abzyme from phage displayed antibody libraries." *Biochem. Biophys. Res. Commun.* **361**: 567.
- Levitt, M. (1978). "Conformational preferences of amino acids in globular proteins." *Biochemistry* **17**: 4277.
- Li, C., M. Pazgier, M. Liu, W. Y. Lu and W. Lu (2009). "Apamin as a template for structure-based rational design of potent peptide activators of p53." *Angew. Chem. Int. Ed.* **48**: 8712.

- Li, W. W., P. Hellwig, M. Ritter and W. Haehnel (2006). "De novo design, synthesis, and characterization of quinoproteins." *Chemistry* **12**: 7236.
- Light, J. and R. A. Lerner (1995). "Random mutagenesis of staphylococcal nuclease and phage display selection." *Biorg. Med. Chem.* **3**: 955.
- Lightstone, F. C. and T. C. Bruice (1996). "Ground State Conformations and Entropic and Enthalpic Factors in the Efficiency of Intramolecular and Enzymatic Reactions. 1. Cyclic Anhydride Formation by Substituted Glutarates, Succinate, and 3,6-Endoxo- Δ^4 -tetrahydrophthalate Monophenyl Esters." *J. Am. Chem. Soc.* **118**: 2595.
- Lightstone, F. C. and T. C. Bruice (1997). "Separation of Ground State and Transition State Effects in Intramolecular and Enzymatic Reactions. 2. A Theoretical Study of the Formation of Transition States in Cyclic Anhydride Formation." *J. Am. Chem. Soc.* **119**: 9103.
- Lippow, S. M. and B. Tidor (2007). "Progress in computational protein design." *Curr. Opin. Biotechnol.* **18**: 305.
- Liu, Y. and B. Kuhlman (2006). "RosettaDesign server for protein design." *Nucleic Acids Res.* **34**: W235.
- Looger, L. L., M. A. Dwyer, J. J. Smith and H. W. Hellinga (2003). "Computational design of receptor and sensor proteins with novel functions." *Nature* **423**: 185.
- Loret, E. P., P. Mansuelle, H. Rochat and C. Granier (1990). "Neurotoxins active on insects: amino acid sequences, chemical modifications, and secondary structure estimation by circular dichroism of toxins from the scorpion *Androctonus australis Hector*." *Biochemistry* **29**: 1492.

- Maayan, G., M. D. Ward and K. Kirshenbaum (2009). "Folded biomimetic oligomers for enantioselective catalysis." *Proc. Natl. Acad. Sci. USA* **106**: 13679.
- MacBeath, G. and D. Hilvert (1996). "Hydrolytic antibodies: variations on a theme." *Chem. Biol.* **3**: 433.
- Maity, H., M. Maity and S. Walter Englander (2004). "How cytochrome c folds, and why: Submolecular foldon units and their stepwise sequential stabilization." *J. Mol. Biol.* **343**: 223.
- Makhatadze, G. I., M. M. Lopez, J. M. Richardson Iii and S. T. Thomas (1998). "Anion binding to the ubiquitin molecule." *Protein Sci.* **7**: 689.
- Manning, M. C. and R. W. Woody (1991). "Theoretical CD studies of polypeptide helices: examination of important electronic and geometric factors." *Biopolymers* **31**: 569.
- Marciano, D. K., M. Russel and S. M. Simon (1999). "An aqueous channel for filamentous phage export." *Science* **284**: 1516.
- Marrone, T. J. and J. A. McCammon (1994). "Pepzyme Dynamics and Conformation: A Molecular Dynamics Study in Water." *J. Am. Chem. Soc.* **116**: 6987.
- Martin, L., P. Barthe, O. Combes, C. Roumestand and C. Vita (2000). "Engineering novel bioactive mini-proteins on natural scaffolds." *Tetrahedron* **56**: 9451.
- Marvin, D. A. (1998). "Filamentous phage structure, infection and assembly." *Curr. Opin. Struct. Biol.* **8**: 150.
- Marvin, D. A., R. D. Hale, C. Nave and M. Helmer-Citterich (1994). "Molecular models and structural comparisons of native and mutant class I

- filamentous bacteriophages Ff (fd, f1, M13), If1 and IKe." *J. Mol. Biol.* **235**: 260.
- Masaru, H., G. Yuji, Y. Noboru and Y. Susumu (1997). "Design and characterization of the anion-sensitive coiled-coil peptide." *Protein Sci.* **6**: 1396.
- Mason, J. M. and K. M. Arndt (2004). "Coiled coil domains: stability, specificity, and biological implications." *ChemBioChem* **5**: 170.
- Menger, F. M. (1966). "The Aminolysis and Amidinolysis of *p*-Nitrophenyl Acetate in Chlorobenzene. A Facile Bifunctional Reactivity." *J. Am. Chem. Soc.* **88**: 3081.
- Menger, F. M. (1992). "Analysis of ground-state and transition-state effects in enzyme catalysis." *Biochemistry* **31**: 5368.
- Menger, F. M. and J. H. Smith (1969). "Aminolysis of phenyl salicylate in acetonitrile. A study of intramolecular catalysis in an aprotic solvent." *J. Am. Chem. Soc.* **91**: 5346.
- Michaelis, L. and M. L. Menten (1913). "Kinetics of invertase action." *Biochem. Z* **49**: 333.
- Miles, A. J. and B. A. Wallace (2006). "Synchrotron radiation circular dichroism spectroscopy of proteins and applications in structural and functional genomics." *Chem. Soc. Rev.* **35**: 39.
- Miroshnikov, A. I., E. G. Elyakova, A. B. Kudelin and L. B. Senyavina (1978). "A study of the physicochemical characteristics of the neurotoxin apamin from the venom of the honeybee, *Apis mellifica* ." *Bioorg. Khim* **4**: 1022.

- Miyazaki, K. and F. H. Arnold (1999). "Exploring nonnatural evolutionary pathways by saturation mutagenesis: rapid improvement of protein function." *J. Mol. Evol.* **49**: 716.
- Monera, O. D., F. D. Sonnichsen, L. Hicks, C. M. Kay and R. S. Hodges (1996). "The relative positions of alanine residues in the hydrophobic core control the formation of two-stranded or four-stranded alpha-helical coiled-coils." *Protein Eng.* **9**: 353.
- Moore, J. C. and F. H. Arnold (1996). "Directed evolution of a para-nitrobenzyl esterase for aqueous-organic solvents." *Nat. Biotechnol.* **14**: 458.
- Morton, V. L., C. T. Friel, L. R. Allen, E. Paci and S. E. Radford (2007). "The effect of increasing the stability of non-native interactions on the folding landscape of the bacterial immunity protein Im9." *J. Mol. Biol.* **371**: 554.
- Mu, Y. Q. and R. A. Gibbs (1997). "Design and synthesis of chiral and racemic phosphonate-based haptens for the induction of aldolase catalytic antibodies." *Bioorg. Med. Chem.* **5**: 1327.
- Munoz, V. and L. Serrano (1994). "Intrinsic secondary structure propensities of the amino acids, using statistical phi-psi matrices: comparison with experimental scales." *Proteins* **20**: 301.
- Murakami, Y., J. I. Kikuchi and O. Hayashida (1989). "Molecular recognition of hydrophobic ammonium substrates by a cationic octopus cyclophane bearing noncovalently bound pyridoxal-5-phosphate: A vitamin B 6-dependent holoenzyme model." *J. Incl. Phenom. Macrocycl. Chem.* **7**: 91.
- Murphy, D. J. (1995). "Revisiting ground-state and transition-state effects, the split-site model, and the "fundamental position" of enzyme catalysis." *Biochemistry* **34**: 4507.

- Nakamizo, N. (1969). "Catalysis in Peptide Synthesis with Active Esters. II. Effects of Concentrated Carboxylic Acids on the Aminolysis of Benzyloxycarbonyl-L-phenylalanine *p*-Nitrophenyl Ester in Dioxane." *Bull. Chem. Soc. Jpn.* **42**: 1078.
- NEB. (2009). "Cleavage Close to the End of DNA Fragments (linearized vector)." from http://www.neb.com/nebecomm/tech_reference/restriction_enzymes/cleavage_linearized_vector.asp.
- Nicoll, A. J. and R. K. Allemann (2004). "Nucleophilic and general acid catalysis at physiological pH by a designed miniature esterase." *OBC* **2**: 2175.
- Nilsson, J. and L. Baltzer (2000). "Reactive-site design in folded-polypeptide catalysts--the leaving group pKa of reactive esters sets the stage for cooperativity in nucleophilic and general-acid catalysis." *Chem. Eur. J.* **6**: 2214.
- Nobeli, I., A. D. Favia and J. M. Thornton (2009). "Protein promiscuity and its implications for biotechnology." *Nat. Biotechnol.* **27**: 157.
- O'Neil, K. T. and W. F. DeGrado (1990). "A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids." *Science* **250**: 646.
- Odenbaugh, A. L., E. D. Helms and B. L. Iverson (2000). "An investigation of antibody acyl hydrolysis catalysis using a large set of related haptens." *Bioorg. Med. Chem.* **8**: 413.
- Olivera, B. M., J. Rivier, J. K. Scott, D. R. Hillyard and L. J. Cruz (1991). "Conotoxins." *J. Biol. Chem.* **266**: 22067.

- Osborne, M. J., A. L. Breeze, L. Y. Lian, A. Reilly, R. James, C. Kleanthous and G. R. Moore (1996). "Three-dimensional solution structure and C-13 nuclear magnetic resonance assignments of the colicin E9 immunity protein Im9." *Biochemistry* **35**: 9505.
- Paci, E., C. T. Friel, K. Lindorff-Larsen, S. E. Radford, M. Karplus and M. Vendruscolo (2004). "Comparison of the transition state ensembles for folding of Im7 and Im9 determined using all-atom molecular dynamics simulations with phi value restraints." *Proteins: Struct. Funct. Bioinform.* **54**: 513.
- Padmanabhan, S., E. J. York, J. M. Stewart and R. L. Baldwin (1996). "Helix propensities of basic amino acids increase with the length of the side-chain." *J. Mol. Biol.* **257**: 726.
- Page, M. I. and A. Williams (1987). Enzyme mechanisms. Royal Society of Chemistry London.
- Patten, P. A., T. Sonoda and M. M. Davis (1996). "Directed evolution studies with combinatorial libraries of T4 lysozyme mutants." *Mol. Divers.* **1**: 97.
- Pauling, L. (1948). "Nature of forces between large molecules of biological interest." *Nature* **161**: 707.
- Pease, J. H., R. W. Storrs and D. E. Wemmer (1990). "Folding and activity of hybrid sequence, disulfide-stabilized peptides." *Proc. Natl. Acad. Sci. USA* **87**: 5643.
- Pease, J. H. B. and D. E. Wemmer (1988). "Solution structure of apamin determined by nuclear magnetic resonance and distance geometry." *Biochemistry* **27**: 8491.

- Pegoraro, S., S. Fiori, J. Cramer, S. Rudolph-Bohner and L. Moroder (1999). "The disulfide-coupled folding pathway of apamin as derived from diselenide-quenched analogs and intermediates." *Protein Sci.* **8**: 1605.
- Pessi, A., E. Bianchi, A. Cramer, S. Venturini, A. Tramontano and M. Sollazzo (1993). "A designed metal-binding protein with a novel fold." *Nature* **362**: 367.
- Petrenko, V. A., G. P. Smith, X. Gong and T. Quinn (1996). "A library of organic landscapes on filamentous phage." *Protein Eng. Des. Sel.* **9**: 797.
- Pollack, S. J., J. W. Jacobs and P. G. Schultz (1986). "Selective chemical catalysis by an antibody." *Science* **234**: 1570.
- Poschner, B. C., J. Reed, D. Langosch and M. W. Hofmann (2007). "An automated application for deconvolution of circular dichroism spectra of small peptides." *Anal. Biochem.* **363**: 306.
- Potts, J. T., Jr., D. M. Young and C. B. Anfinsen (1963). "Reconstitution of fully active RNase S by carboxypeptidase-degraded RNase S-peptide." *J. Biol. Chem.* **238**: 2593.
- Price-Carter, M., W. R. Gray and D. P. Goldenberg (1996a). "Folding of omega-conotoxins. 1. Efficient disulfide-coupled folding of mature sequences in vitro." *Biochemistry* **35**: 15537.
- Price-Carter, M., W. R. Gray and D. P. Goldenberg (1996b). "Folding of omega-conotoxins. 2. Influence of precursor sequences and protein disulfide isomerase." *Biochemistry* **35**: 15547.
- Prodromou, C. and L. H. Pearl (1992). "Recursive PCR: a novel technique for total gene synthesis." *Protein Eng. Des. Sel.* **5**: 827.

- Quinn, T. P., N. B. Tweedy, R. W. Williams, J. S. Richardson and D. C. Richardson (1994). "Betadoublet: de novo design, synthesis, and characterization of a beta-sandwich protein." *Proc. Natl. Acad. Sci. USA* **91**: 8747.
- Ramalingam, K. and G. H. Snyder (1993). "Selective disulfide formation in truncated apamin and sarafotoxin." *Biochemistry* **32**: 11155.
- Reed, J. and T. A. Reed (1997). "A set of constructed type spectra for the practical estimation of peptide secondary structure from circular dichroism." *Anal. Biochem.* **254**: 36.
- Robinson, C. R. and S. G. Sligar (1993). "Electrostatic stabilization in four-helix bundle proteins." *Protein Sci.* **2**: 826.
- Roccatano, D., G. Colombo, M. Fioroni and A. E. Mark (2002). "Mechanism by which 2,2,2-trifluoroethanol/water mixtures stabilize secondary-structure formation in peptides: a molecular dynamics study." *Proc. Natl. Acad. Sci. USA* **99**: 12179.
- Romesberg, F. E., B. Spiller, P. G. Schultz and R. C. Stevens (1998). "Immunological origins of binding and catalysis in a Diels-Alderase antibody." *Science* **279**: 1929.
- Rossi, P., P. Tecilla, L. Baltzer and P. Scrimin (2004). "De novo metallonucleases based on helix-loop-helix motifs." *Chem. Eur. J.* **10**: 4163.
- Röthlisberger, D., O. Khersonsky, A. M. Wollacott, L. Jiang, J. DeChancie, J. Betker, J. L. Gallaher, E. A. Althoff, A. Zanghellini and O. Dym (2008). "Kemp elimination catalysts by computational enzyme design." *Nature* **453**: 190.

- Rudgers, G. W. and T. Palzkill (1999). "Identification of residues in β -lactamase critical for binding β -lactamase inhibitory protein." *J. Biol. Chem.* **274**: 6963.
- Sacher, E. and K. J. Laidler (1964). "Kinetics of the catalysed hydrolysis of *p*-nitrophenyl acetate " *Can. J. Chem.* **42**: 2404.
- Sambrook, J. and D. W. Russell (2001). Molecular cloning: a laboratory manual. CSHL Press, New York.
- Satchell, D. P. N. and Secemski, II (1969). "Acylation. Part XXIX. The mechanism of ester aminolysis in non-hydroxylic media and the effect of nitrogen-containing leaving groups." *J. Chem. Soc. Perk. Trans. 2* **1969**: 130.
- Satchell, D. P. N. and Secemski, II (1970). "Acylation. Part XXXI. Preferred transition states for ester aminolysis in non-hydroxylic solvents." *J. Chem. Soc. Perk. Trans. 2* **1970**: 1013.
- Savarin, P., R. Romi-Lebrun, S. Zinn-Justin, B. Lebrun, T. Nakajima, B. Gilquin and A. Menez (1999). "Structural and functional consequences of the presence of a fourth disulfide bridge in the scorpion short toxins: solution structure of the potassium channel inhibitor HsTX1." *Protein Sci.* **8**: 2672.
- Schmeer, G., C. Six and J. Steinkirchner (1999). "Investigations on Substituent and Solvent Effects of Solvolysis Reactions. VIII. The Influence of Water and Nonaqueous Solvents on the Imidazolysis of 4-Nitrophenyl Acetate." *J. Solution Chem.* **28**: 211.
- Schmidt, D. M. Z., E. C. Mundorff, M. Dojka, E. Bermudez, J. E. Ness, S. Govindarajan, P. C. Babbitt, J. Minshull and J. A. Gerlt (2003). "Evolutionary Potential of ([β]/[α]) 8-Barrels: Functional

- Promiscuity Produced by Single Substitutions in the Enolase Superfamily†." *Biochemistry* **42**: 8387.
- Schreier, B., C. Stumpp, S. Wiesner and B. Hocker (2009). "Computational design of ligand binding is not a solved problem." *Proc. Natl. Acad. Sci. USA* **106**: 18491.
- Seo, E. S., W. R. P. Scott, S. K. Straus and J. C. Sherman (2007). "Optimal Attachment Position and Linker Length Promote Native-like Character of Cavitand-Based Template-Assembled Synthetic Proteins (TASPs)." *Chem. Eur. J.* **13**: 3596.
- Shawali, A. and S. S. Biechler (1967). "Aminolysis of esters. I. Kinetics and mechanism in anhydrous doxane." *J. Am. Chem. Soc.* **89**: 3020.
- Singer, B. and J. T. Kusmierek (1982). "Chemical mutagenesis." *Annu. Rev. Biochem.* **51**: 655.
- Sirotkin, K. (1986). "Advantages to mutagenesis techniques generating populations containing the complete spectrum of single codon changes." *J. Theor. Biol.* **123**: 261.
- Smith, G. P. (1985). "Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface." *Science* **228**: 1315.
- Smith, G. P. and J. K. Scott (1993). "Libraries of peptides and proteins displayed on filamentous phage." *Methods Enzymol.* **217**: 228.
- Snider, M. J. and R. Wolfenden (2000). "The rate of spontaneous decarboxylation of amino acids." *J. Am. Chem. Soc.* **122**: 11507.
- Sommese, R. F., S. Sivaramakrishnan, R. L. Baldwin and J. A. Spudich (2010). "Helicity of short E-R/K peptides." *Protein Sci.* **10**: 2001.

- Song, K., J. M. Stewart, R. M. Fesinmeyer, N. H. Andersen and C. Simmerling (2008). "Structural insights for designed alanine-rich helices: comparing NMR helicity measures and conformational ensembles from molecular dynamics simulation." *Biopolymers* **89**: 747.
- Sonkaria, S., G. Boucher, J. Flórez-Álvarez, B. Said, S. Hussain, E. L. Ostler, S. Gul, E. W. Thomas, M. Resmini and G. Gallacher (2004). "Evidence for 'lock and key' character in an anti-phosphonate hydrolytic antibody catalytic site augmented by non-reaction centre recognition: variation in substrate selectivity between an anti-phosphonate antibody, an anti-phosphate antibody and two hydrolytic enzymes." *Biochem. J.* **381**: 125.
- Stephens, D. B., R. E. Thomas, J. F. Stanton and B. L. Iverson (1998). "Polyclonal antibody catalytic variability." *Biochem J* **332 (Pt 1)**: 127.
- Stetefeld, J., M. Jenny, T. Schulthess, R. Landwehr, J. Engel and R. A. Kammerer (2000). "Crystal structure of a naturally occurring parallel right-handed coiled coil tetramer." *Nat. Struct. Biol.* **7**: 772.
- Stewart, J. D. and S. J. Benkovic (1995). "Transition-state stabilization as a measure of the efficiency of antibody catalysis." *Nature* **375**: 388.
- Stoddard, B. L. and D. E. Koshland Jr (1993). "Structure of isocitrate dehydrogenase with. alpha.-ketoglutarate at 2.7- Å resolution: conformational changes induced by decarboxylation of isocitrate." *Biochemistry* **32**: 9317.
- Stryer, L. (1988). Biochemistry (3rd edn). Feeman and Company, New York.
- Sumner, J. B. (1926). "The isolation and crystallization of the enzyme urease. Preliminary paper." *J. Biol. Chem.* **69**: 435.

- Suveges, D., Z. Gaspari, G. Toth and L. Nyitray (2009). "Charged single alpha-helix: a versatile protein structural motif." *Proteins* **74**: 905.
- Takagi, R., A. Shiraki, T. Manabe, S. Kojima and K. Ohkata (2000). "The Juliá-Colonna Type Asymmetric Epoxidation Reaction Catalyzed by Soluble Oligo-L-leucines Containing an α -Aminoisobutyric Acid Residue: Importance of Helical Structure of the Catalyst on Asymmetric Induction." *Chem. Lett.* **29**: 366.
- Tam, J. P., C. R. Wu, W. Liu and J. W. Zhang (1991). "Disulfide bond formation in peptides by dimethyl sulfoxide. Scope and applications." *J. Am. Chem. Soc.* **113**: 6657.
- Tan, P. T., A. Veeramani, K. N. Srinivasan, S. Ranganathan and V. Brusica (2006). "SCORPION2: a database for structure-function analysis of scorpion toxins." *Toxicon* **47**: 356.
- Tang, Y., J. B. Hicks and D. Hilvert (1991). "In vivo catalysis of a metabolically essential reaction by an antibody." *Proc. Natl. Acad. Sci. USA* **88**: 8784.
- Tantillo, D. J. and K. N. Houk (1999). "Fidelity in Hapten Design: How Analogous Are Phosphonate Haptens to the Transition States for Alkaline Hydrolyses of Aryl Esters?" *J. Org. Chem.* **64**: 3066.
- Tawfik, D. S., A. B. Lindner, R. Chap, Z. Eshhar and B. S. Green (1997). "Efficient and selective p-nitrophenyl-ester-hydrolyzing antibodies elicited by a p-nitrobenzyl phosphonate hapten." *Eur. J. Biochem.* **244**: 619.
- Tawfik, D. S., R. R. Zemel, R. Arad-Yellin, B. S. Green and Z. Eshhar (1990). "Simple method for selecting catalytic monoclonal antibodies that exhibit turnover and specificity." *Biochemistry* **29**: 9916.

- Taylor, S. E., T. J. Rutherford and R. K. Allemann (2001). "Design, synthesis and characterisation of a peptide with oxaloacetate decarboxylase activity." *Bioorg. Med. Chem. Lett.* **11**: 2631.
- Taylor, S. E., T. J. Rutherford and R. K. Allemann (2002). "Design of a folded, conformationally stable oxaloacetate decarboxylase." *J. Chem. Soc. Perk. Trans. 2* **2002**: 751.
- Thorn, S. N., R. G. Daniels, M. T. Auditor and D. Hilvert (1995). "Large rate accelerations in antibody catalysis by strategic use of haptenic charge." *Nature* **373**: 228.
- Trainor, G. L. and R. Breslow (1981). "High acylation rates and enantioselectivity with cyclodextrin complexes of rigid substrates." *J. Am. Chem. Soc.* **103**: 154.
- Tramontano, A., K. D. Janda and R. A. Lerner (1986). "Catalytic antibodies." *Science* **234**: 1566.
- Tuchscherer, G. and M. Mutter (1996). "Template assisted protein de novo design." *Pure Appl. Chem.* **68**: 2153.
- Turner, E. C., C. H. Cureton, C. J. Weston, O. S. Smart and R. K. Allemann (2004). "Controlling the DNA binding specificity of bHLH proteins through intramolecular interactions." *Chem. Biol.* **11**: 69.
- Turner, E. C., C. J. Weston, O. S. Smart and R. K. Allemann (2003). "Apamin-induced stabilization of the basic helix-loop-helix domain of MyoD." *Biophys. J.* **84**: 366A.
- Ulmer, K. M. (1983). "Protein engineering." *Science* **219**: 666.

- van Loo, B., S. Jonas, A. C. Babbie, A. Benjdia, O. Berteau, M. Hyvonen and F. Hollfelder (2010). "An efficient, multiply promiscuous hydrolase in the alkaline phosphatase superfamily." *Proc. Natl. Acad. Sci. USA* **107**: 2740.
- Vendruscolo, M., N. V. Dokholyan, E. Paci and M. Karplus (2002). "Small-world view of the amino acids that play a key role in protein folding." *Phys. Rev. E: Stat. Phys., Plasmas, Fluids*, **65**: 61910.
- Venkatraman, J., S. C. Shankaramma and P. Balaram (2001). "Design of folded peptides." *Chem. Rev.* **101**: 3131.
- Wagner, J., R. A. Lerner and C. F. Barbas Iii (1989). "Efficient aldolase catalytic antibodies that use the enamine mechanism of natural enzymes." *Science* **25**: 19.
- Walker, K. D. and T. P. Causgrove (2009). "Contribution of arginine-glutamate salt bridges to helix stability." *J. Mol. Model.* **15**: 1213.
- Wallace, B. A. and C. L. Teeters (1987). "Differential absorption flattening optical effects are significant in the circular dichroism spectra of large membrane fragments." *Biochemistry* **26**: 65.
- Wallace, T. J. (1964). "Reactions of thiols with sulfoxides. I. Scope of the reaction and synthetic applications." *J. Am. Chem. Soc.* **86**: 2018.
- Wallace, T. J. and J. J. Mahon (1964). "Reactions of Thiols with Sulfoxides. II. Kinetics and Mechanistic Implications¹." *J. Am. Chem. Soc.* **86**: 4099.
- Wallis, R., K. Y. Leung, M. J. Osborne, R. James, G. R. Moore and C. Kleanthous (1998). "Specificity in protein-protein recognition: conserved Im9 residues are the major determinants of stability in the Colicin E9 DNase-Im9 complex." *Biochemistry* **37**: 476.

- Wallis, R., A. Reilly, K. Barnes, C. Abell, D. G. Campbell, G. R. Moore, R. James and C. Kleanthous (1994). "Tandem overproduction and characterisation of the nuclease domain of colicin E9 and its cognate inhibitor protein Im9." *Eur. J. Biochem.* **220**: 447.
- Wallis, R., A. Reilly, A. Rowe, G. R. Moore, R. James and C. Kleanthous (1992). "In vivo and In vitro Characterization of Overproduced Colicin E9 Immunity Protein." *Eur. J. Biochem.* **207**: 687.
- Watts, D. J. and S. H. Strogatz (1998). "Collective dynamics of 'small-world' networks." *Nature* **393**: 440.
- Wedemayer, G. J., L. H. Wang, P. A. Patten, P. G. Schultz and R. C. Stevens (1997). "Crystal structures of the free and liganded form of an esterolytic catalytic antibody1." *J. Mol. Biol.* **268**: 390.
- Wells, J. A., W. J. Fairbrother, J. Otlewski, M. Laskowski, Jr. and J. Burnier (1994). "A reinvestigation of a synthetic peptide (TrPepz) designed to mimic trypsin." *Proc. Natl. Acad. Sci. USA* **91**: 4110.
- Weston, C. J., C. H. Cureton, M. J. Calvert, O. S. Smart and R. K. Allemann (2004). "A stable miniature protein with oxaloacetate decarboxylase activity." *ChemBioChem* **5**: 1075.
- Wetlaufer, D. B. and V. P. Saxena (1970). "Formation of three-dimensional structure in proteins. I. Rapid nonenzymic reactivation of reduced lysozyme." *Biochemistry* **9**: 5015.
- Whitmore, L. and B. A. Wallace (2008). "Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases." *Biopolymers* **89**: 392.

- Wilmore, B. H. and B. L. Iverson (1994). "Phosphate versus phosphorothioate haptens for the production of catalytic polyclonal antibodies." *J. Am. Chem. Soc.* **116**: 2181.
- Wolfenden, R. (1976). "Transition state analog inhibitors and enzyme catalysis." *Annu. Rev. Biophys. Bioeng.* **5**: 271.
- Xu, J., Q. Deng, J. Chen, K. N. Houk, J. Bartek, D. Hilvert and I. A. Wilson (1999). "Evolution of shape complementarity and catalytic efficiency from a primordial antibody template." *Science* **286**: 2345.
- Xu, X. and J. W. Nelson (1994). "One-disulfide intermediates of apamin exhibit native-like structure." *Biochemistry* **33**: 5253.
- Yan, Y. and B. W. Erickson (1994). "Engineering of betabellin 14D: disulfide-induced folding of a beta-sheet protein." *Protein Sci.* **3**: 1069.
- Yang, J. T., C. S. Wu and H. M. Martinez (1986). "Calculation of protein conformation from circular dichroism." *Methods Enzymol.* **130**: 208.
- Ye, H., D. Rong, W. Tong and V. T. D'Souza (1992). "Artificial redox enzymes. Part 3. Structure and properties." *J. Chem. Soc. Perkin Trans. 2* **1992**: 2071.
- Ymer, S. (1991). "Heat inactivation of DNA ligase prior to electroporation increases transformation efficiency." *Nucleic Acids Res.* **19**: 6960.
- Yoshikuni, Y., T. E. Ferrin and J. D. Keasling (2006). "Designed divergent evolution of enzyme function." *Nature* **440**: 1078.
- Zanghellini, A., L. Jiang, A. M. Wollacott, G. Cheng, J. Meiler, E. A. Althoff, D. Röthlisberger and D. Baker (2006). "New algorithms and an in silico benchmark for computational enzyme design." *Protein Sci.* **15**: 2785.
- Zani, M. L. and T. Moreau (2010). "Phage display as a powerful tool to engineer protease inhibitors." *Biochimie* **92**: 1689.

- Zavodszky, M., C. W. Chen, J. K. Huang, M. Zolkiewski, L. Wen and R. Krishnamoorthi (2001). "Disulfide bond effects on protein stability: Designed variants of Cucurbita maxima trypsin inhibitor-V." *Protein Sci.* **10**: 149.
- Zhang, D., S. Li, W. Li and Y. Chen (2007). "Biomimic recognition and catalysis by an imprinted catalysts: a rational design of molecular self-assembly toward predetermined high specificity." *Catal. Lett.* **115**: 169.
- Zhao, Q., I. M. Kovach, A. Bencsura and A. Papathanassiou (1994). "Enantioselective and Reversible Inhibition of Trypsin and alpha.-Chymotrypsin by Phosphonate Esters." *Biochemistry* **33**: 8128.
- Zhou, G. W., J. Guo, W. Huang, R. J. Fletterick and T. S. Scanlan (1994). "Crystal structure of a catalytic antibody with a serine protease active site." *Science* **265**: 1059.
- Zugasti-Cruz, A., M. B. Aguilar, A. Falcon, B. M. Olivera and E. P. Heimer de la Cotera (2008). "Two new 4-Cys conotoxins (framework 14) of the vermivorous snail *Conus austini* from the Gulf of Mexico with activity in the central nervous system of mice." *Peptides* **29**: 179.

