Structural and biophysical characterization of photoswitchable peptides and their complexes

by

Piotr Wysoczański

A thesis submitted to

Cardiff University

for the degree of

DOCTOR OF PHILOSOPHY

School of Chemistry

Cardiff University

2012

Acknowledgements

I would like to express my gratitude to Professor Rudolf K. Allemann for granting me the chance to join his team and suggesting fruitful projects. I also recognize his help, encouragement and patience during the course of my doctorate studies. I would like to express my gratitude to my collaborators Dr Christopher Williams, Dr Mathew Crump at the University of Bristol bioNMR lab, Andrea Brancali, Antonio Ricci at the School of Pharmacy and Sara B. Whittaker at the Henry Wellcome Building for NMR of University of Birmingham. I would like to express my gratitude to all members of Prof. Rudolf K. Allemann, Dr. Daffyd Jones, Prof. Matthias Bochtler and Dr. Eric Tippmann labs that have actively or passively helped in my PhD I hold an eternal gratefulness for Dr Robert Mart; his knowledge, self-learning skills and intelligence were often invaluable. I also acknowledge Dr Mahmoud Akhtar for his help with the revisions of my thesis and my friends Dr Oliver Ladeback, Dr Thomas Fricke, Matt Cyrankiewicz, Ola Rajska, Amang Li and Sarah Adams for financial and moral support. Many thanks to my girlfriend Toni.

Abstract

Networks of biomacromolecular interaction form the core of internal cellular logic where a wide range of input signals get analysed and results in a genetically programmed output. This description holds not only for healthy cells but is true also for those virally infected when a virus implements new biomacromolecular elements and modifies existing networks. Those hostile interactions are crucial for the HIV virus replication cycle. Due to the scale-free nature of biomacromolecular interaction networks they are prone to malfunction if a hub (node with the highest interactions) is damaged. Perhaps the best example of such a situation can be observed in the apoptotic machinery. Overexpression, deletion or mutations (depending on the context) in apoptotic proteins can lead to various diseases, most prominently cancer. Therefore current research efforts concentrate on designing inhibitors of malfunctioning interactions. Occasionally interaction between the two binding partners relies on a well-defined secondary structures element. This allows to engineer this element and use it to bias the given interaction equilibrium towards dissociation. One such approach concentrates on α-helices that are crosslinked with azobenzene derivatives. This permits not only to increase the helical content of a peptide, which translates to increased affinity (potency), but transplants a photo-controllability. Photoswitchable peptides created in this way can be controlled externally; their affinity towards targets can be effectively photo-modulated.

The first chapter of the thesis concentrates on the development and evaluation of photo-controllable peptides derived from the arginine rich region of the HIV type I Rev protein. Rev, which is an early product generated from the HIV genome, binds to RRE containing mRNA and removes it from the nucleus, allowing expression of unspliced or singly spliced mRNA into different protein products. As a consequence the RRE-Rev interaction controls the infection cycle of the HIV type 1, as these later products are necessary for virus particle assembly. The α-helical peptide derived from Rev was shown to retain most of the affinity to RRE. Rev-RRE serves as a prototypical protein-RNA interaction; a model for the development of the first photo-controllable peptide that targets RNA. This research presents two approaches for the peptide production: chemical synthesis and biosynthesis (overexpression of the Rev peptide gene). Initially the water soluble azobenzene derivative was used bearing two negatively charged sulphonate substituents on the aromatic rings. Surprisingly, the first generation of the peptides had substantially decreased affinity to RRE

as compared to the wild type Rev peptide. By applying computational chemistry techniques a possible rationalisation was found: sulphonates interact strongly with crucial arginines on the peptide making them unavailable to engage the phosphate backbone; additionally the peptide was forced into non-helical conformation prior to binding. This hypothesis was validated by crosslinking Rev peptides with azobenzene derivative deprived of the sulphonates. The second generation of Rev peptides binds RRE with high affinity and the interaction in question can be controlled effectively by light.

The second chapter of the thesis focuses on the structural characterization of a photo-controllable Bak peptide bound to Bcl-x_L – a prominent member of the Bcl-2 family of proteins. The Bcl-2 family of proteins includes the major regulators and effectors of the intrinsic apoptosis pathway. Cancers are frequently formed when activation of the apoptosis mechanism is compromised either by misregulated expression of prosurvival family members or, more frequently, by damage to the regulatory pathways that trigger intrinsic apoptosis. Short peptides derived from the pro-apoptotic members of the Bcl-2 family can activate mechanisms that ultimately lead to cell death. The recent development of photocontrolled peptides that are able to change their conformation and activity upon irradiation with an external light source has provided new tools to target cells for apoptosis induction with temporal and spatial control. In this thesis the first NMR solution structure of a photoswitchable peptide derived from the proapoptotic protein Bak in complex with the antiapoptotic protein Bcl-x_L is reported. This structure provides insight into the molecular mechanism, by which the increased affinity of such photopeptides compared to their native forms, is achieved, and offers a rationale for the large differences in the binding affinities between the helical and nonhelical states.

Index

Table of Figures

Table of Tables

Table of Equations

List of Abbreviations

A	Adenine
ABL	Abelson murine leukaemia viral oncogene homolog 1
AIF	Apoptosis inducing factor
AlwN1	*Acinetobacter iwoffii* N type II site-specific deoxyribonuclease
Apaf-1	Apoptosis protease activating factor 1
APO-1	Apoptosis antigen 1
Apo1L	Apo1 ligand
APS	Ammonium persulphate
ATP	Adenosine triphosphate
Bad	Bcl-2-associated death promoter
Bak	Bcl-2 homologous antagonist/killer
Bax	Bcl-2–associated X protein

Bcl-2	B-cell lymphoma-2
Bcl-2A1	Bcl-2-related protein A1
Bcl-b	Apoptosis regulator Bcl-B
Bcl-x_L	B-cell lymphoma extra large
BCR	breakpoint cluster region gene
BH	Bcl-2 homology
Bid	BH3-interacting domain death agonist
Bim	Bcl2-interacting mediator of cell death
Boc	Butoxycarbonyl
Bok	Bcl-2 ovarian killer
C	Cytosine
C. elegans	*Caenorhabditis elegans*
CARD	Caspase activation and recruitment domain
CD	Circular dichroism
CD4+	Cluster of differentiation 4 plus
CD95	Cluster of differentiation 95
ChR1	Channelrhodopsin-1
CNBr	Cyanogen bromide
COSY	Correlation spectroscopy
Crm1	Chromosomal region maintenance protein 1
Cy3	Cyanine 3

DCM	Dichloromethane
DD	Death domain
DED	Death effector domain
DIEA	N,N-Diisopropylethylamine
DISC	Death inducing signalling complex
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DR	Death receptor
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EMSA	Electrophoretic mobility shift assay
EndoG	Endonuclease G
F2f	F2 filtered
FADD	Fas-associated DD adapter
FAM	Fluoresceinamide
Fas	Apoptosis stimulating fragment
FasL	Fas ligand
FKHR	Forkhead in rhabdomyosarcoma

FLIP	FLICE inhibitory protein
Fmoc	Fluorenylmethylcarbamoyl
FPLC	Fast protein liquid chromatography
FRET	Fluorescence resonance energy transfer
Fyn	Src family tyrosine-protein kinase
G	Guanine
Gag	Group Antigens
Gdm:HCl	Guanidine hydrochloride
gp120	Glycoprotein 120
gp41	Glycoprotein 41
HBS	Hydrogen bond surrogate derived α-helices
HBTU	O-Benzotriazole-N,N,N',N'-tetramethyl-uronium-hexafluoro-phosphate
HDM2	Human double minute protein 2
HDMX	Human double minute protein X
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
HIV	Human immunodeficiency virus
HOBt	Hydroxybenzotriazole
HPLC	High performance liquid chromatography
Hrk	Harakiri protein
Hs	homoserine
HSQC	Heteronuclear single quantum coherence

IAP	Inhibitor of apoptosis
IC_{50}	Half maximum inhibitory concentration
I	Inosine
IPTG	Isopropyl-β-D-thiogalactopyranoside
K_D	Dissociation constant
K_i	Inhibition constant
KSI	Ketosteroid isomerase
LB	Luria-Bertani
LTR	Long terminal repeats
MAC	Mitochondrial apoptotic channels
MALDI	Matrix assisted laser desorption and ionization
MALDI-TOF	Matrix assisted laser desorption/ionisation - time of flight
Mcl-1	Myeloid cell leukaemia sequence 1
MD	Molecular dynamics
MDM2	Murine double minute protein 2
MDMX	Murine double minute protein X
MeCN	Acetonitrile
miRNA	microRNA
MM	Molecular mechanics
mRNA	Messenger ribonucleic acid
MS	Mass spectroscopy

MscL	Mechanosensitive channel of large conductance
Mtd	Matador
mTOR	Mammalian target of rapamycin
MWCO	Molecular weight cut off
NCCD	Nomenclature Committee on Cell Death
NEB	New England Biolabs
nef	Negative regulatory factor
NES	Nuclear export sequence
NiNTA	Nickel-nitrilotriacetic acid
NLS	Nuclear localization sequence
NMP	N-Methylpyrrolidinone
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
NOESY	Nuclear Overhauser effect spectroscopy
NOXA	Noxious stress induced protein
OD	Optical density
ORF	Open reading frame
p17	HIV protein 17
p53	Tumour suppressor protein p53
p7	HIV protein 7
PAGE	Polyacrylamide gel electrophoresis

PAX3	Paired box 3
Pbf	2,2,4,6,7- Pentamethyldihydrobenzofurane
PCD	Programmed cell death
PCR	Polymerase chain reaction
PDB	Protein Data Bank
pI	Isoelectric point
PLAD	Preligand binding assembly domain
PMSF	Phenylmethylsulphonyl fluoride
Pol	Polyprotein
PPI	Protein-protein interactions
Pr	Phototropin form red
Prf	Phototropin form far-red
PUMA	p53 Unregulated modulator of apoptosis
QM	Quantum mechanics
R.E.D.	RESP and ESP charge derive
RESP	Restrained electrostatic potential
Rev	Regulator of viron
RMSD	Root mean square deviation
RNA	Ribonucleic acid
ROCK I	Rho effector protein I
RRE	Rev response element

RT	Reverse transcriptase
SDS	Sodium dodecyl sulphate
SEM	Scanning electron microscopy
SH3	SRC homology 3 domain
SIV	Simmian immunodeficiency virus
SMAC	Second mitochondria-derived activator of caspases
SPR	Surface plasmon resonance
T	Thymidine
TAE	Tris-acetate-EDTA buffer
Tar	Trans-activation-responsive region
TAT	Transactivator of transcription
TB	Terrific broth
TBA	Tris-borane-acetate buffer
TBE	Tris-borane-EDTA buffer
tBid	Truncated Bid
tBu	Tert-butyl
TCEP	Tris(2-chloroethyl) phosphate
TEMED	N,N,N′,N′-Tetramethylethan-1,2-diamine
TFA	Trifluoroacetic acid
TFE	2,2,2-Trifluoroethanol
THF	Tetra-hydrofuran

TIS	Triisopropylsilane
TM	Transmembrane region
TNF	Tumor necrosis factor
TNFR	TNF receptor
TOCSY	Total correlation spectroscopy
TOF	Time of flight
TRAIL	Tumour necrosis factor related apoptosis inducing ligand
TRAILR	TRAIL receptor
Tris	Tris(hydroxymethyl)-aminomethane
tRNA	Transfer ribonucleic acid
Trt	Trityl
U	Uridine
UTP	Uridine triphosphate
UV	Ultraviolet
vif	Virulence factor
Vis	Visual
vpr	Viral protein R
vpu	Viral protein U
wt	Wild type
XIAP	X-linked inhibitor of apoptosis
XL-H	4,4'-Bis(chloroacetamido)azobenzene

XL-SO_3H	3,3'-Bis(sulfonato)-4,4'-bis(chloroacetamido)azobenzene

1 INTRODUCTION

1.1 Protein Structure

1.1.1 Amino acids and primary structure of proteins and peptides

Proteins and peptides are macromolecular polymers composed of amino acids. Definitions of proteins and peptides vary but are loosely based on the polymer chain length, so that short chains, between 2 and 30 amino acids, are considered to be peptides. Polypeptides that are longer and more sophisticated in higher order structures tend to be called proteins. Apart from a relatively small group of short peptides, all proteins are biosynthesized by ribosomes that use a strand of mRNA, transcribed from a DNA sequence, as a template that encodes a protein sequence in triplets of nucleobases together with complementary tRNA molecules loaded with corresponding amino acids. Jointly these molecules translate the encoded sequence into a one-dimensional protein sequence. Naturally occurring proteins and peptides are composed of 22 proteinogenic amino acids: alanine (Ala, A), arginine (Arg, R), aspartate (Asp, D), asparagine (Asn, N) cysteine (Cys, C), glycine (Gly, G), glutamate (Glu, E), glutamine (Gln, Q), isoleucine (Ile, I), histidine (His, H), leucine (Leu, L), lysine (Lys, K), phenylalanine (Phe, F), methionine (Met, M), proline (Pro, P), pyrrolysine (Pyl, O), selenocysteine (Sec, U), serine (Ser, S), threonine (Thr, T), tryptophan (Trp, W), tyrosine (Tyr, Y), valine (Val, V). Although many other amino acids are biosynthesized and sometimes incorporated into polypeptides by non-ribosomal mechanism, only the aforementioned 22 are encoded by triplets of mRNA nucleobases (N.B. pyrrolysine and selenocysteine have additional requirements to be incorporated).[1-3] Amino acids share a common scaffold of an amino group and a carboxylic acid group bound to a central alpha carbon (Cα). It is the side chain of amino acids, also connected to the alpha carbon, which gives them their unique properties. Amino acids can be categorized into groups of similar size, hydrophobicity, charge or aromaticity based on the properties of their side chains. Amino acids in peptides or proteins may also be selectively post-translationally modified by attachment of a variety of chemical groups including lipids, sugars, nucleic acids, hydroxyls, methyl groups or other products of a secondary metabolism. Amino acids within a protein are linked by a peptide bond, which due to delocalisation, has a partial double bond character and a tendency to remain planar (Figure 1). It is the sequence of amino acids connected by these

peptide bonds that defines the primary structure of protein and peptides. As the peptide bond is asymmetric, the polypeptide strand also has directionality with distinct N-, and C-termini.

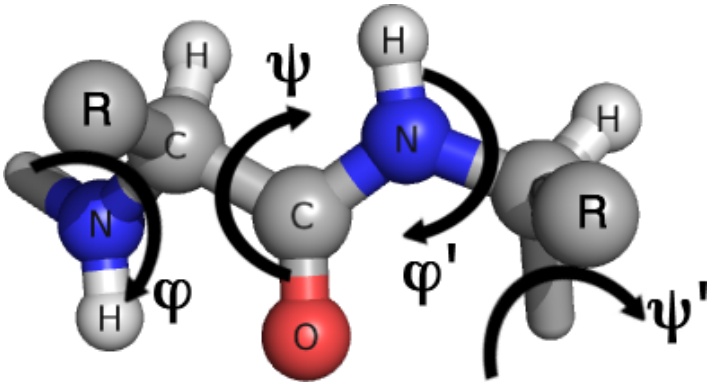

Figure 1: A model of two peptide bonds with ϕ (ϕ') and Ψ (Ψ') angles indicated.

1.1.2 Secondary structure

The one dimensional primary structure is driven, predominantly by the formation of hydrogen bonds between peptide bonds, to fold into few privileged three-dimensional secondary structures. The lowest energy arrangements of φ and ψ angles (which are correlated to hydrogen bonds patterns and steric effects) also contribute to secondary structure formation. The most frequently adopted secondary structures are the β-sheet and α-helix (Figure 2). The β-strands are an extended helical arrangement of a peptide chain where 2 amino acids form each turn (2 amino acids per pitch). Because the peptide chain has directionality, β-strands can associate into two sorts of β–sheets with the relative direction of the strands parallel or anti-parallel. In a parallel β-sheet φ/ψ angles are close to -139°/-135° and in an antiparallel -119/+113°.[4] In the α-helical arrangement, polypeptide chains are twisted into a characteristic right-handed helix with hydrogen bonds occurring between amino acids 4 residues apart giving a helix pitch of 3.6 amino acids (Figure 2); φ and ψ angles of such alpha helix cluster around the values of -57°/-47°, respectively (left-handed arrangements also occur, but relatively rarely).[4]

Depending on the environment around an α-helix, they can be predominantly hydrophobic (e.g. the core of a globular protein, or spanning a lipid bilayer), hydrophilic (e.g. helices that contact other macromolecules *via* electrostatic interactions) or amphipathic with hydrophobic and hydrophilic regions either perpendicular or parallel to the direction of the helix (e.g. localized at the interface between hydrophilic and hydrophobic environment). The helical structure is usually reinforced by other elements of the overall protein structure, as the

hydrogen bonds alone do not provide enough energy to stabilise an isolated α-helix. α-Helices deprived of their scaffold tend to be unfolded in aqueous solutions; however the structure depends on the additives present in a solution and its *p*H.

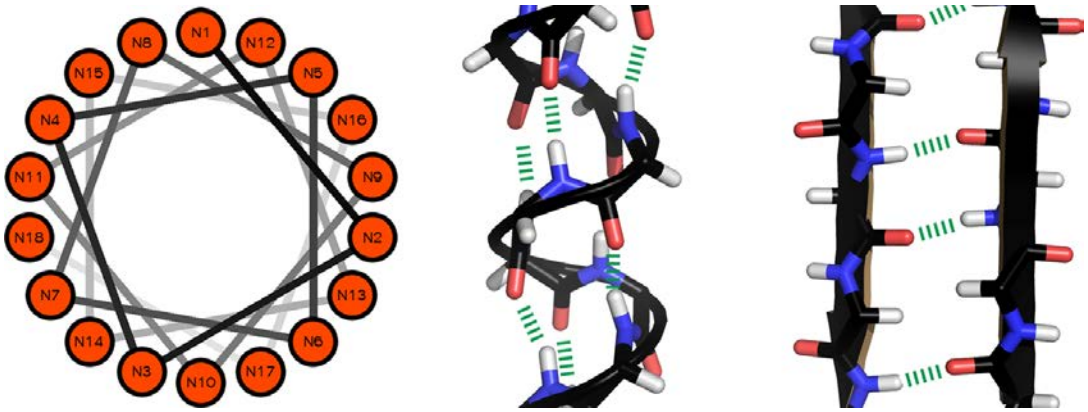

Figure 2: Example of protein secondary structure: a helical projection of a right-handed α-helix (on the right), model of a right-handed α-helix presented along the long axis (in the middle) and an antiparallel β-sheet (on the left). Hydrogen bonds are indicated in green; the intensity of the gray lines corresponds to the depth.

Due to the presence of so many hydrogen bonds aligned in the same direction, helices are dipolar with the negative pole at the C-terminus. That increases the energetic cost of assembling the helix and is very often compensated for by oppositely charged residues at respective ends, i.e. arginines at the C-terminus. α-Helical peptides are usually synthesized with a C-terminal amide cap that removes the negative charge of the carboxy acid terminus and N-terminal acetyl cap that removes the positive one of the N-terminus. In addition to α-helices, there are four other helical type secondary structures: π helix (φ/ψ −57/−70°), 3_{10} helix (-49/-26°) and two types of secondary structure adopted only by polyprolines: poly(Pro) I (φ/ψ −83/−158°) poly(Pro) II (φ/ψ −78/−149°).[4] Two different geometries of turns are also classified: a four-residue β-turn and a three-residue gamma turn.[4]

Amino acids can be divided into groups based on their propensity to form and stabilize different types of secondary structure; both statistical methods and direct biochemical investigation provide similar results. Glutamate, methionine, alanine appear often in α-helices, whereas valine, isoleucine and tyrosine are typical of β−sheets. Proline and glycine have tendency to break helices and are often found in parts of protein structure that change direction. Loops are long stretches of polypeptide sequences that do not appear to have any of the mentioned structures. Often they are flexible and probe many conformations, but occasionally loops are very stable but just do not fall into any canonical category of torsion angles.

1.1.3 Tertiary, quaternary structure and protein domains

Tertiary structure refers to the positions of all atoms within the given protein. Some proteins can be additionally subdivided into domains; elements of the protein structure, which can fold independently. The driving force for this larger scale folding and association is predominantly the hydrophobic effect; as the protein folds it places hydrophobic regions together excluding many ordered water molecules (water clusters around hydrophobic areas) and making the process entropically favourable despite the increased order of the protein. Hydrogen bonds, van der Waals interactions (dipole-dipole, dipole-induced dipole, and induced dipole/induced-dipole) and electrostatic interactions (salt bridges) can also promote folding, as can the presence of disulfide bridges between cysteines. Folded proteins can further interact with each other to form stable protein complexes, constituting the quaternary structure of proteins. Protein interactions will be discussed further later. Although it was once believed that proteins form rigid and well defined structures, many examples are now known of proteins that are natively unfolded and even more often only a part of the protein structure is well defined, with the remainder being natively unfolded. The advantage of this arrangement is that inherent structural flexibility allows them to adapt to many conditions and interact with many different proteins.

1.2 Protein interactions

1.2.1 Overview of protein interaction

There are four main categories of macromolecular interactions that a protein can be involved in: protein-protein (PPI), protein-DNA, protein-RNA and protein-polysaccharides (sugar trees) interactions. Proteins can also interact with many different types of small molecules including but not limited to polysaccharides, lipids, amino acids or molecules that are hybrids of aforementioned both in terms of covalent attachment and non-covalent complexes. The scope of this thesis is confined to protein-protein and protein-RNA interactions.

1.2.2 Protein-protein interactions

The average protein in the yeast proteome interacts with five other proteins, which leads to a total of 25,000 to 35,000 protein-protein interactions.[5] This number is estimated to be even higher for *Homo sapiens* at around 650,000 total interactions.[6] Such large numbers

underscore the crucial importance of protein-protein interactions which underpin almost all cellular functions. The average surface area of two interacting proteins is approximately 1600 Å^2.[2,7] Many such interfaces involve elements of secondary structure but are generally flat and were long considered poor targets for small molecule drugs, which are better suited to targeting deep pockets than shallow protein-protein interfaces.[8] Although such interfaces encompass many hydrophilic and hydrophobic residues, only certain percentage of amino acids forming the interface contribute significantly to the binding energy;[9] these are called "hot-spots", residues which upon mutation decrease the binding affinity the most.[9] It has also been observed that the larger the area of the interface, the more hydrophilic residues it contains; this effect corresponds to the necessity of dissociated units to remain in solution long enough to find their binding partner.[10] Not all protein interfaces are flat and dispersed over large areas, in a number of cases only a certain key element of the secondary structure is responsible for the majority of the binding energy. Bioinformatic methods reveal that 62% of protein-protein complexes structures deposited in protein data bank (PDB) contain an α–helix as one of the elements of the interface.[11] Of those 16% (10% of the total) contain key hot-spot residues within an α-helix and could be targeted by helix mimics.[11] Broadly, those interfaces may be divided into two categories; those with binding clefts where the distance between flanking hotspot residues is less than 7 Å and those with extended interfaces where it is between 7 and 30 Å.[11] In principle, α-helix protein interactions could be approximated by a peptide corresponding to the given α-helix and its binding protein, however such a peptide deprived of its structural scaffold is unfolded or adopts a different conformation in solution. The associated entropic cost of rearrangement upon binding of such peptide may significantly decrease its binding affinity. In the case of intrinsically (natively) unfolded proteins this difference might be less pronounced.

1.2.2.1 Targeting helix-binding clefts

Considerable efforts in recent years have led to the development of a number of different approaches to target helix-binding clefts and disrupt protein-protein interactions. Four major strategies have been adopted: the development of "classic" small molecule inhibitors, mimicking α–helical conformation through modular non-peptide scaffolds, using β–amino acid peptides and finally, modifying existing α-helical peptides to reinforce their structure and increase binding and/or stability or transplanting the binding epitope onto a stable, small protein.[12]

1.2.2.2 Classic small molecule inhibitors targeting helix-binding clefts

A limited number of successful small molecule inhibitors of protein-protein interactions have been reported so far. The list of somewhat successful inhibitors of protein-protein interactions is too long to be contained in this thesis. The engagement of various elements of protein structure in the binding interface reflects (at least to the first approximation) the statistical distribution within Protein Data Bank; this implies that only a subset of compounds within that group targets an α-helix binding cleft. This thesis will be only concerned with that type of protein-protein interactions.

Typical targets of those molecules are well-characterized members of the B cell lymphoma 2 (Bcl-2) family and human double minute 2/X (HDM2/X) protein; both of which are well-known oncogenes, making inhibitors promising targets for a cancer therapy. The two acylsulfonamide compounds W1191542 and ABT-737 target respectively B cell lymphoma extra-large (Bcl-x_L) and Bcl-2 proteins (Figure 3) with nanomolar dissociation constants.

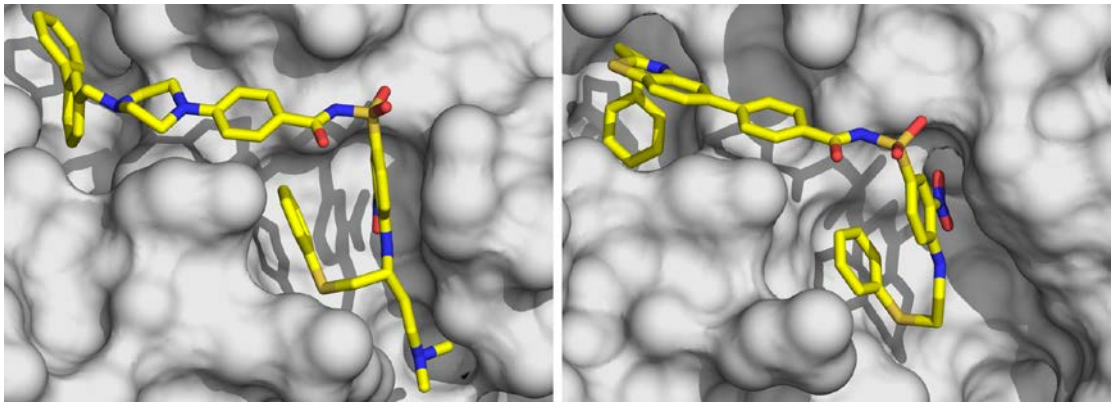

Figure 3: Surface representation of: Bcl-x_L with the W1191542 inhibitor (left) and Bcl-2 with ABT-737 inhibitor (right); PDB entries: 3INQ, 1YSW.

ABT-737 was shown to induce regression of solid tumours (Figure 3). Both compounds mimic the interactions present between the Bcl-2 homology domain 3 only (BH3-only) proteins as do others that aim to mimic the periodic structure of α–helix in a modular fashion.[13-15]

Interesting PPI occurs between eukaryotic translation initiation factor EIF4E and EIF4G. This complex is directly involved in the recognition of 5′, 7- methylguanosine mRNA cap and mRNA translocation to the ribosome for subsequent translation. The *EIF4E* gene was found to be recurrently overexpressed in cancer cells.[16] EIF4G was crystallized with a 14 amino acids long peptide derived from EIF4G (Figure 4).[17] The bound peptide is predominantly α-helical, but five amino acids run perpendicularly to the helix; one of the

amino acids form this secondary structure uncategorized stretch belongs to the group that most significantly contributes to the binding energy.

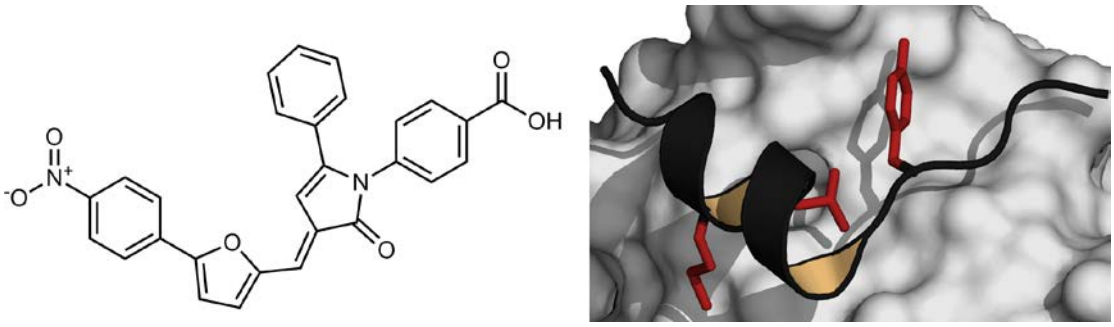

Figure 4: Structure of the 4E1RCat inhibitor of the EIF4E/EIF4G interaction (left); model of the crystal structure between EIF4E and its binding peptide (4E-BP1). Key residues are shown in read. PDB entries: 1EJ4.

This structure was used to design an inhibitor with K_d of 16 μM.[18] This inhibitor was shown to alter chemoresistance in a mouse model.[18] Unfortunately no structure of EIF4E with any of the inhibitors bound exists but the interaction site with the mRNA cap can also be targeted and these structures are available.[19] The mode of binding was confirmed by NMR techniques and the inhibitor binds in the same pocket as the wild type peptide.

The only other proteins whose structure with the α-helical part of the binding partner (a peptide) was solved is Z interacting protein A (ZipA) and the 17 amino acids long peptide derived from filamenting temperature-sensitive mutant Z (FtsZ);[20] both of which are *E. coli* cell division proteins. Although, similarly as in the case of EIF4E/peptide interaction α-helix is accompanied by even longer structurally uncategorized binding segment, all residues contributing significantly to the binding energy cluster in the middle of the α-helix. This PPI was deemed undruggable according to some researchers.[21] Nevertheless a series of inhibitors were synthesised the best of which exhibited K_i of 12 μM.[22-24] Structure of this compound bound to ZipA was solved by X-ray techniques (Figure 5).[23]

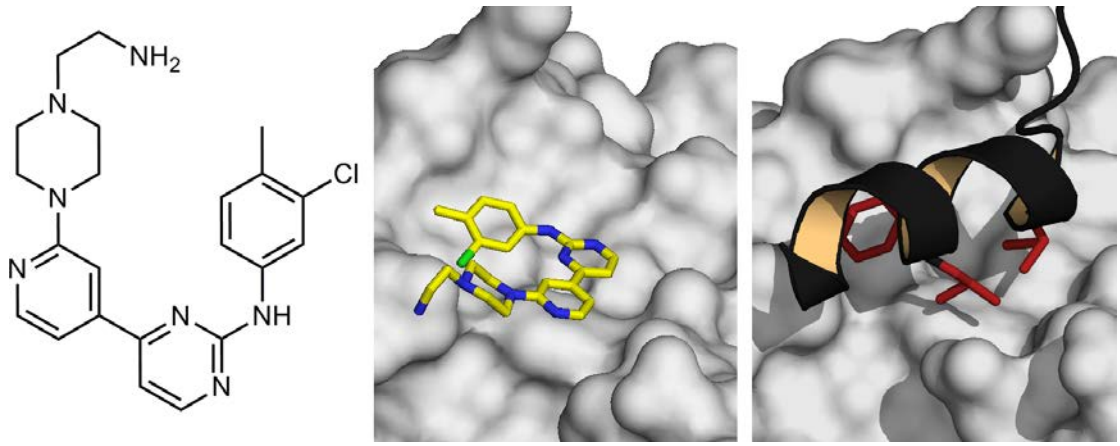

Figure 5: Structure pyridylpyrimidine inhibitor of the ZipA/FtsZ interaction (left); the inhibitor bound to ZipA; ZipA/FtsZ structure. Key residues are shown in read. PDB entries: 1YTF, 1F47.

At this point a distinction between classic small molecules and peptidomimetics has to be made. Given a simple interaction between an α-helical peptide and a target protein or a potent enough inhibitor and the target protein it is highly possible that the structure of the inhibitor would mimic the α-helical peptide. This rule might be broken in situations were the interaction in question is more complex (larger surface area and various elements of protein structure involved) or a significant conformational shift occurs upon binding. In the former situation, the inhibitor might mimic an epitope rather than a distinct secondary structure element. In the latter, the conformation of the inhibitor/protein could be trapped somewhere along the conformational shift coordinate due to some additional, sufficiently deep energetic minima present. Therefore, the difference between classic small molecule inhibitors and an α-helical peptidomimetic lies in the concept rather than the final outcome. Small molecules are developed to fit the binding pocket with disregard to the natural binder that is to be displaced, whereas peptidomimetics are designed specifically to mimic hotspot residues and their respective position in an α-helix (or broader in any structural element of a protein). Moreover peptidomimetics structures can be subdivided into scaffold (an equivalent of a protein backbone) and substituents (corresponding to protein amino acid sidechains). In the end peptidomimetics are going to be subdivided into modular, semi-modular and “custom-made”. Custom-made peptidomimetics relay on a scaffold that can’t be easily extended and can’t be straightforwardly subdivided into modules that could be connected by simple synthetic routes or wouldn’t be stable in reality.

Different scaffolds have been developed in order to replicate a small helical section of p53 as bound to HDM2/X. Targeting the p53 HDM2/X interaction is ostensibly simple as just three residues contribute to the majority of the binding affinity with the remainder acting as a

scaffold. One compound resulting from this approach, a pyrrolopyrimidine-based inhibitor, has an affinity similar to its parent peptide, as strategically positioned chemical substituents mimic the three hot-spot residues: Phe19, Trp23 and Leu26 (Figure 6).[25] p53-HDM2/X is a crucial interaction that, amongst other roles, controls activation of the intrinsic apoptotic pathway in response to a range of different stress making it an important therapeutic target.

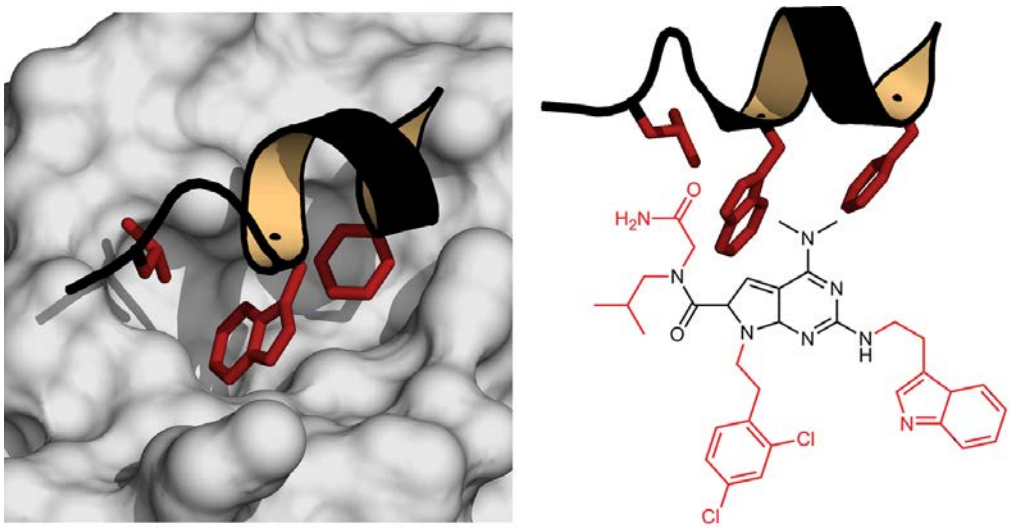

Figure 6: MDMX (murine HDMX analog) bound to p53 peptide (left); p53 compared to pyrrolopyrimidine-based inhibitor (right). Key residues of p53 and corresponding substituents of the inhibitor are shown in red

The most recent inhibitor for which a model of a crystal structure exists is based on piperidinone scaffold.[26] It was shown to be a potent binder of HDM2 with K_d of 0.4 nM. Moreover, in a mouse model it was able to suppress the tumour growth. Similarly as in the previous case, this inhibitor protrudes three substituents into corresponding binding pockets on HDM2.[26] The structure of the inhibitor can be subdivided into scaffold and "sidechains" (Figure 7) and although this design resembles a peptidomimetic it was created by screening a library of compounds based on previous successful inhibitors rather than the p53 peptide.

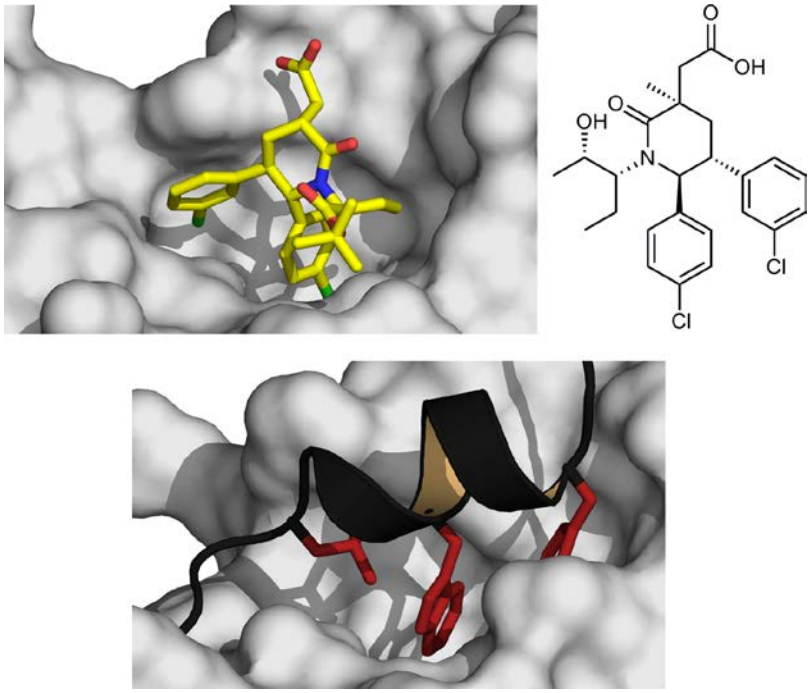

Figure 7: Structure compound 9 inhibitor of the HDM2 interaction (upper right); the inhibitor bound to HDM2 (upper left); HDM2/p53 peptide structure (down). Key residues are shown in read. PDB entries: 1YCR, 4ERF.

1.2.2.3 Modular peptide mimics inhibitors targeting helix-binding clefts

As mentioned previously, modular approaches to constructing α–helical mimics can be taken by providing a repeating scaffold that can be extended for different lengths of helices. The most widely used group of those mimetics is derived from a *para*-oligophenyl scaffold.[27,28] The "backbone" is functionalized appropriately in the *ortho* or *meta* positions to mimic *i*, *i*+4 or *i*, *i*+7, etc. side chain placements on an α–helix. Hamilton and co-workers developed this approach and applied it to target Bcl-x_L/Bak interaction.[29] The most successful from this series of inhibitors displaced Bak wild type peptide with K_i of 114 nM validating the method (Figure 8).[29] The same scaffold was shown to disrupt the HIV glycoprotein (gp41) helical hydrophobic core formation[30] and p53/HDM2 interaction.[31,32] The biggest issue with this type of scaffold and "sidechains" is the solubility of those compounds that limits possible applications.

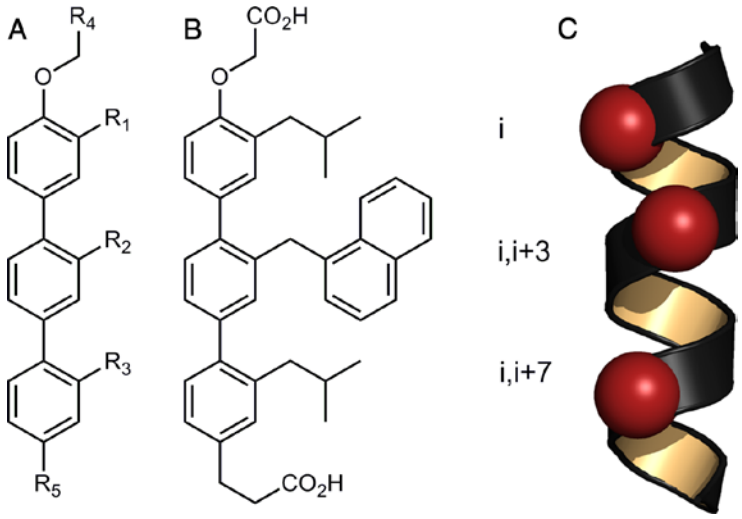

Figure 8: The general *para*-oligophenyl scaffold (A); the same scaffold with a set of "sidechains" that was shown to bind Bcl-x_L with 114 nM affinity (B); idealized α-helix with sidechains positions that are mimicked. Please not that the above structures do not represent the minimal energy conformation.

Similar approaches use amide bonds between pyridine rings and phenyl (oligobenzamides, oligopyridylamides). The former scaffold is additionally reinforced (has reduced degrees of freedom) by the extensive network of hydrogen bonds between amide hydrogen, carbonyl oxygen (Figure 9). This in theory reduces the entropic cost upon binding. Oligopyridylamides with appropriate "sidechains" were utilized to disrupt the Bcl-x_L/Bak interaction with an order of magnitude worse K_i as compared to the oligophenyl based scaffolds (*vide supra*).[33]

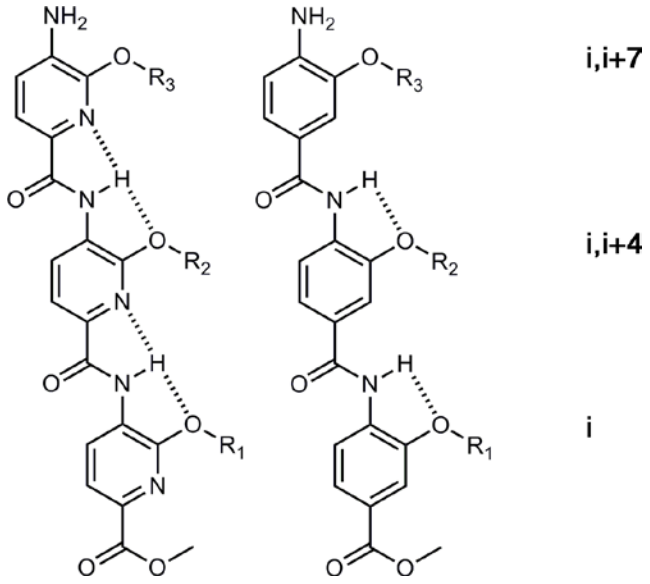

Figure 9: Oligopyridylamides (left) and oligobenzamides (right) peptidomimetics scaffolds. Please not that the above structures do not represent the minimal energy conformation.

More recent investigation revealed that a simpler scaffold of oligobenzamides, which lacks the bifurcated hydrogen bonds of oligopyridylamides allows easier access to conformations

more favourable for engaging Bcl-x_L as well as exposes more hydrophobic surface to complement the binding grove.[34] Also, both of those scaffolds are more water soluble.

The modular and universal designs of repeating identical units outline above was further developed into scaffolds that break this simplicity. Combinatorial and rational approaches used all scaffolds outlined above as starting points. Both modification of terminal or middle units was tested and lead to results (semi-modular peptidomimetics) ranging from a mild improvement to a mild decrease of the affinity (Figure 10).[35-37] This indicates a trend to combine "universal and minimalistic" peptidomimetics with classic small molecule design approaches.

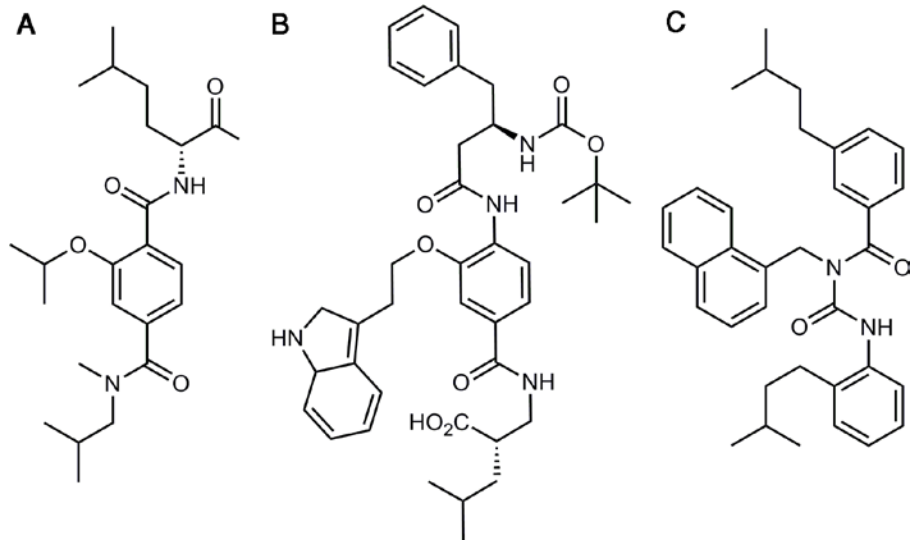

Figure 10: A sample of semi-modular peptidomimetics design to target Bcl-xL, K_i=0.78 µM (A); HDM2, K_i=8 µM (B) and again Bcl-x_L, K_i=2.4 µM

The great number of iterative and novel designs of peptidomimetic scaffolds does not allow for a full overview of the field. The conclusion can be made that to the best of my knowledge none of the modular or semi-modular peptidomimetics had higher affinity to a given protein as compared to "custom-made" peptidomimetics or classic small molecule binders. This does not mean that modular peptidomimetics do not have potential to become truly universal and potent, but it indicates that this generally simple and clever concept needs more development.

One of the missed points in this field is lack of any research that would systematically screen scaffolds once a small set of possibly optimal sidechains is found. In a number of important cases crystal structure of the complexes are available and sidechains can be readily identified. The reason why screening for optimal sidechains is the routine nowadays lies in the scaffolds themselves, because scaffolds are not ideal and this needs to be compensated by

sidechains. This could provide solutions to three issues haunting some of the peptidomimetics: affinity as contributed by the position of the sidechain, solubility and affinity contributions from the backbone. Of course, the rich data acquired so far allows for somewhat similar in effect retro-analysis but it wouldn't replace the experimental scaffolds screening.

1.2.2.4 Beta and gamma a amino acids peptide inhibitors targeting helix-binding clefts

An interesting group of strategies targeting helix-binding clefts involves the use of β- or γ-amino acids (Figure 11). Those singly or doubly homologated amino acids to come existence after insertion of CH_2 groups at any position between NH_2 and C=O without altering the sidechains.[38] Two positions for β- and three for γ-amino acids are possible for sidechain extension (β^2, β^3 and γ^2, γ^3, γ^4). Moreover occasionally amino acids are synthesized with more than one sidechain either on the same carbon (e.g. $\beta^{2,2}$, $\gamma^{3,3}$) or on different ones (e.g. $\beta^{2,3}$, $\gamma^{3,4}$), or a combination of those (e.g. $\beta^{2,2,3}$, $\gamma^{3,3,4}$). Additionally, some of the recurring types of amino acids are cyclised. Those peptides are nowadays rarely constructed solely from the un-natural amino acids, but rather β- or γ-amino acids are intertwined with regular α-amino acids.

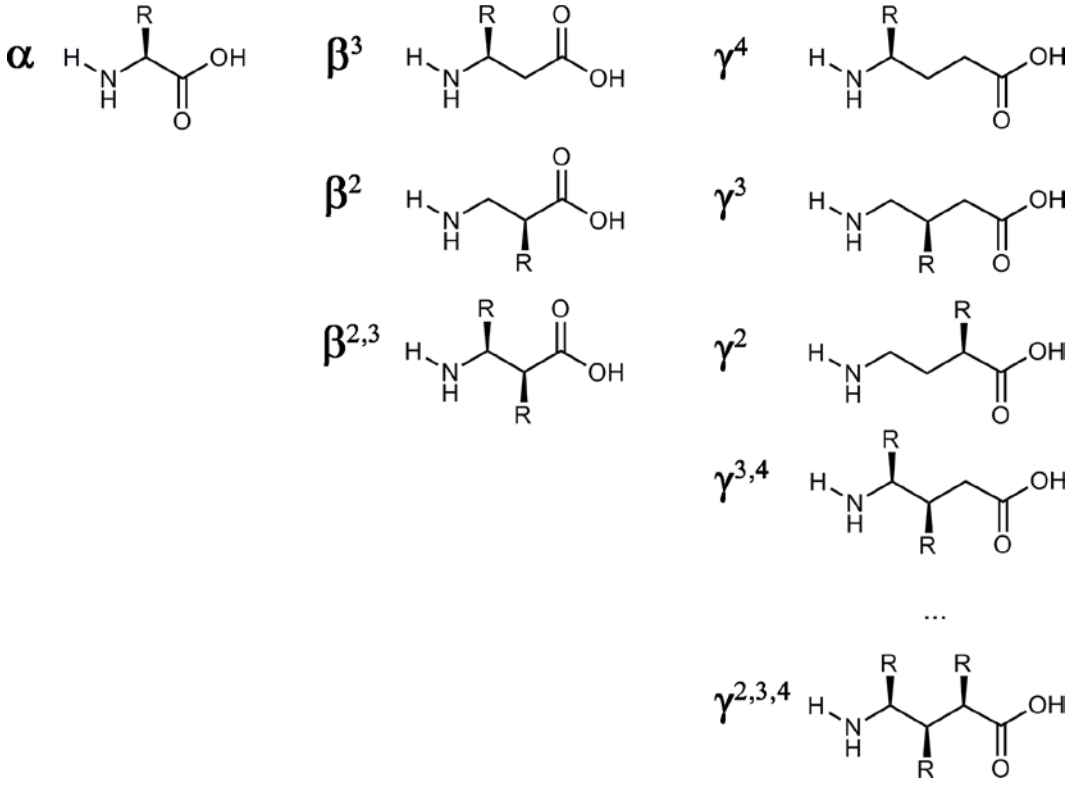

Figure 11: Examples of α, β and γ amino acids used in the construction of peptides.

The main advantage of constructing helix mimics from β/γ-amino acids over naturally occurring α-amino acids is their stability towards enzymatic degradation.[39] The extensive studies of proteolytic stability of all β- or γ-peptides (including stability against 20S proteasome) revealed that all peptides examined were stable for at least 48 h as compared to the 15 min stability of a regular poly-α-alanine peptide.[39] α/β/γ –amino acid (in any combination) containing peptides have been studied extensively and structural paradigms have been established for their design.[40-43] An especially interesting pattern has been recently examined: αγααβα (hexad) which is isoatomic to αααααααα (heptad).[44] This patter was shown to form a regular α-helix structure.[44]

Some of those peptides have been successfully used as inhibitors of protein-protein interaction and they retain the affinity of the sequences they are derived from (K_i of 2 nM).[45] Bim peptide was mimicked by a peptide that has its one half composed entirely of α-amino acids and the other half of the αβ mix (Figure 12). All β-amino acids in the sequence have their sequence cyclised and serve as a scaffold.

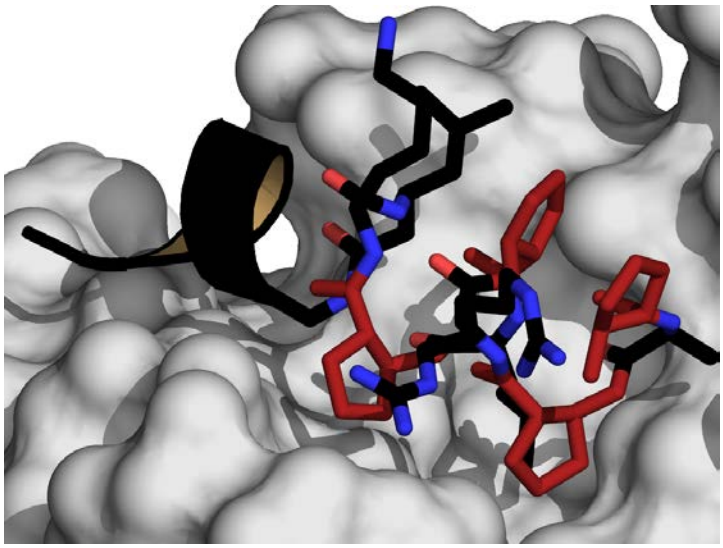

Figure 12: Structural model of α/β amino acids peptide bound to Bcl-2; βamino acids are indicated in red. PDB entries: 3FDM.

It was also possible to create a peptide mimicking the Bim peptide using αααβ pattern that had similar affinity to Bcl-x_L as the wild type α-peptide (50 nM vs. 23 nM).[46] The crystal structure reveals that β-amino acids are positioned not only on the outer face of the helix, but also contribute to the binding interface (Figure 13).

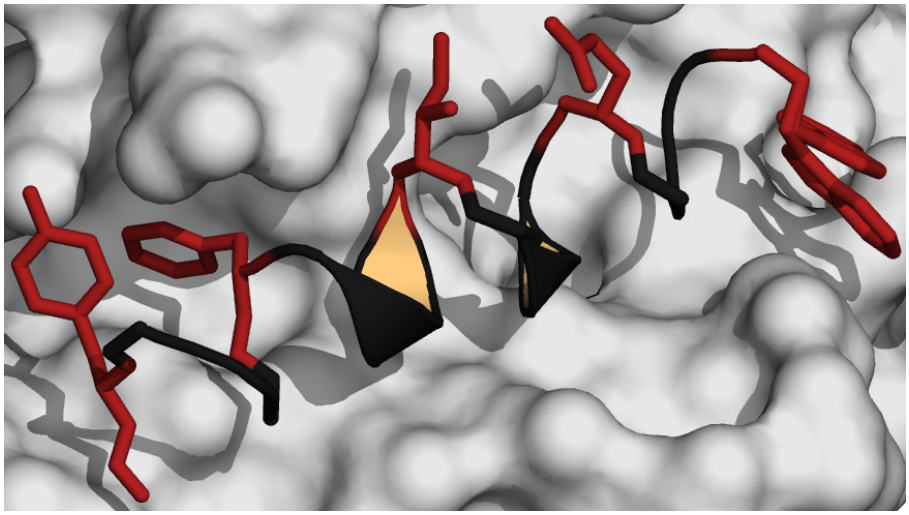

Figure 13: The crystal structure model of ααaβ peptide bound to Bcl-x_L; β−amino acids are indicated in red. PDB entries: 4A1W

1.2.2.5 Modified peptide inhibitors targeting helix-binding clefts

Introduction of bridging or crosslinking elements into a peptide structure can stabilize their α-helical structure and make them more resistant to a proteolytic degradation. The most straightforward and readily available crosslinking elements are cysteines. Unfortunately crosslinking two L-cysteines, even with optimal spacing *i,i+3*, is not ideal for the α-helix nucleation as sidechains are not pointing towards each other and the length of the disulfide bond is not sufficient to deal with this problem one can introduce D-cysteine at the *i* position while keeping L-cysteine at *i,+3* (S-S distance is 6.7 Å in an idealized α-helix); although strictly speaking this bridging forces formation of a helical but non-α-helical conformation it still can be utilized for α-helical binding clefts (Figure 14).[47,48] The estrogen receptor α is a transcription factor that upon binding of its small molecule activator undergoes a conformational shift that allows it to bind a coactivator peptide; only then it becomes fully activated and can promotes specific genes expression. The so called NR motif (LXXLL) is known to be almost exclusively responsible for the binding of the coactivator peptide. It was possible to engineer the peptide by introducing D- and L-cysteines to stabilize the structure (PERM 1) and increase binding affinity to 25 nM (Figure 14).[48] Interestingly, this potent binder appears to have almost no helical character in solution as judged by CD spectroscopy.[48] This observation is congruent with the effect that *i,i+3* spaced, disulfide bond forming D- and L-cysteine has on the conformation (*vide supra*).[47]

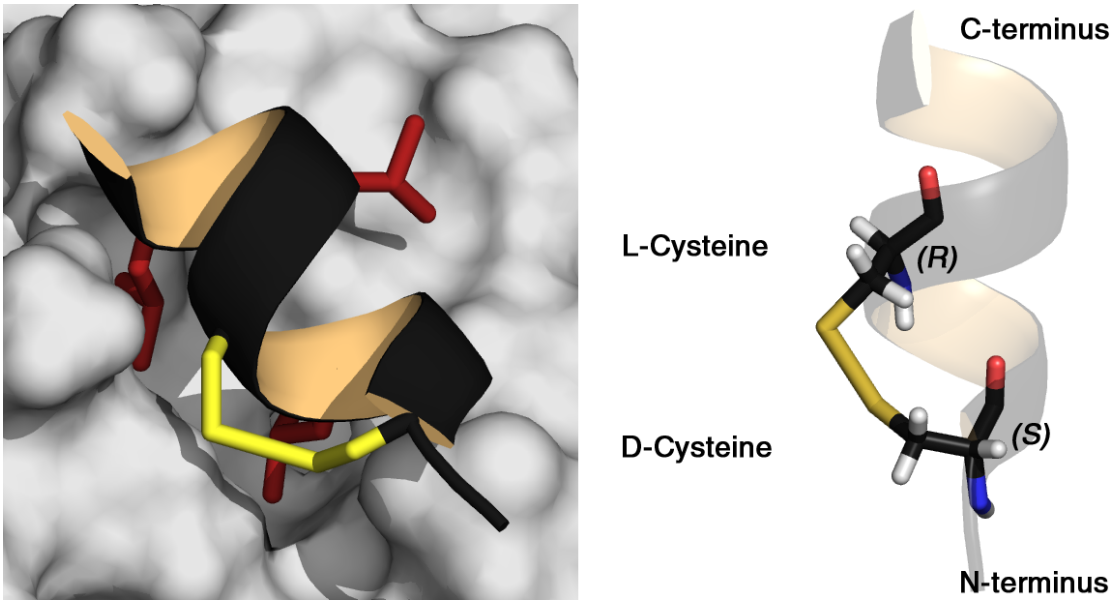

Figure 14: X-ray structure model of PERM 1 peptide bound to estrogen receptor α (left); close-up on the disulfide bond within the PERM 1 peptide. Key residues are shown in read. PDB entries: 1PCG

The reversible character of true disulfide bridges often makes them unsuitable for *in vivo* applications as the intracellular environment has a reducing potential,[48] but constructing a covalent bridge as a stable analogue lowers the entropy of the free peptide considerably. Within this strategy there is a broad range of linkages that can be used which may include synthesising a small molecule crosslinker or appropriately modified/extended amino acids to incorporate a bridging element whose length would match the calculated distances in the idealized α-helix. Usually, residues that are modified or crosslinked are spaced in the peptide sequence according to the multiples of the helix pitch (3.6 residues per pitch), which forms so called *i*, *i*+4, *i*, *i*+7; *i*, *i*+11 configurations.

Though covalent modifications are the method usually chosen, metal chelation approaches have also been reported.[49] p53/HDM2 interaction served as a prove of principle target. Two rhodium atoms are chelated by two glutamate carboxylic sidechains in *i, i+4* spacing and two molecules of acetic acid effectively crosslinking the p53 peptide (Figure 15). The problem of multiple carboxylic acid residues in p53 peptide (both part of sidechains and C-terminus) was solved by orthogonal protection (sidechains) and preparative chromatography (C-terminus). The peptide affinity to HDM2 was not improved; K_d of 66 vs. 77 nM for the wild type and the crosslinked peptide respectively. Nevertheless, as the dirhodium centre can serve as a catalyst of carbon-hydrogen insertion reactions which opens up interesting possibilities for directing the catalyses via protein-protein interaction.

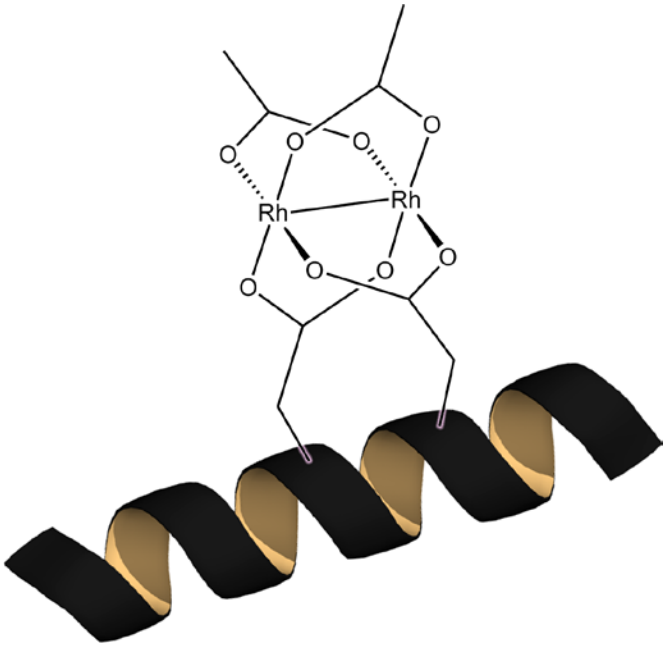

Figure 15: An idealized α-helix crosslinked *via* dirhodium chelation.

Arguably the most elegant embodiment of this method creates an additional amide bond parallel to the helix. To achieve this pyrrolopyrimidine-based orthogonal amino acid sidechain protection can be employed during chemical peptide synthesis to selectively form a number of peptide bonds between Glu/Asp and Lys. Non-proteinogenic amino acids may also be incorporated (ornithine instead of lysine) to create an optimal crosslink length (Figure 16). This methodology has been used to create stable peptides capable of binding their targets improving affinities for their targets compared to non-crosslinked equivalents.[48,50,51] Indeed, the smallest ever stable alpha helix reported was created using this strategy.[52] Particularly successful were inhibitors of Human respiratory syncytial virus (RSV) cell fusion. The main disadvantage of this method is that simple amides may still be proteolysed *in vivo*. RSV F glycoprotein is responsible for the recognition of host cell receptors and fusion with the host cell membrane. The interaction is associated with a major rearrangement of the fusion protein. Peptides derived from this protein is α-helical segments could in principle block the transition to a 6 α-helix bundle (three HR-C and three HR-N). The synthetic doubly constrained peptide derived from the HR-C α-helix (Figure 16) was able to inhibit the cell fusion event at IC of 190 pM, which was a four orders of magnitude improvement over the wild type peptide; it is essential to mention that the constrained peptide had 1-naphthylalanine residue at N-term, whilst this was absent in the control peptide.[50]

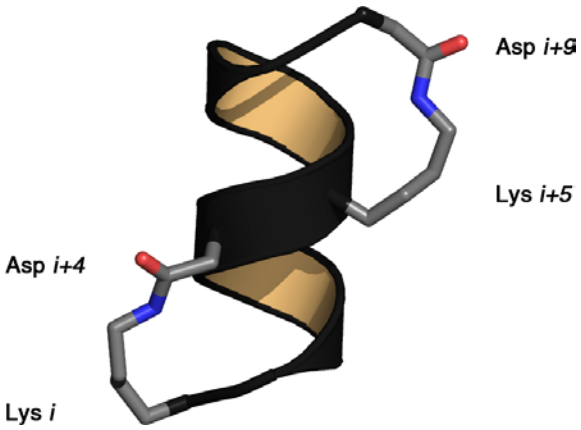

Figure 16: Model of doubly constrained α-helix inhibitor of RSV cell fusion. The peptide crosslinked in two places by forming lactam rings between lysine and aspartate residues (red) in *i*, *i*+4 configuration.

A similar method uses a short hydrocarbon bridge to stabilise α–helices.[53] A peptide is chemically synthesized with two non-natural amino acids, each bearing an olefin sidechain, which is subsequently closed by a metathesis reaction catalysed by Grubbs' first generation catalyst to form a ring (Figure 17).[54] In the literature this form of crosslinking is referred to as a staple. Although other functionally and chemically equivalent methods exist, hydrocarbon stapling has been reported as additionally conferring the sought after property of cell permeability, making stapled peptides promising therapeutics.[55] This strategy has been shown to effectively enhance properties of a number of medically and biologically relevant peptides.[56-59] The nearly archetypical PPI inhibitor target, the p53/HDM2 interaction, was perhaps the most successful complex to be disrupted by stapled peptides. The best inhibitor bound p53 with K_d of 0.92 nM, which is nearly 500 fold better than the wild type peptide and still better than any of the "natural" p53 peptides with improved affinities (*vide supra*).[60] Unfortunately the peptide was unable to cross the cell membrane possibly due to accumulation of negative charges. Further engineering made the peptide positively charged and the staple position was altered, thus decreasing the affinity to 55 nM but making the peptide cell permeable and proapoptotic (active *in vivo*).[60] The recently solved crystal structure of the aforementioned peptide bound to HDM2 hints that the staple not only stabilizes the helix and helps the peptide to cross the cell membrane, but also makes positive contributions to the binding affinity (Figure 17).[61] ERα and coactivator peptide interaction succumbed to the stapled peptides too. The stapled peptide sequence (SP-2) most directly comparable to the wild type was a seven fold better binder.[58] Addition of Glu and Lys at the N-term and repositioning of the staple lead to further increase and the peptide (SP-6) had a K_d of 75 nM (2500 nM for the wild type). The crystal structure of the ERα bound to SP-2 was

solved (Figure 17).[58] The most interesting feature of this complex is the correspondence between Leu 693 of the wild type (LXX**L**L) absent in the SP-2 sequence and the staple. It's clear that part of the staple acts as the missing leucine.[58]

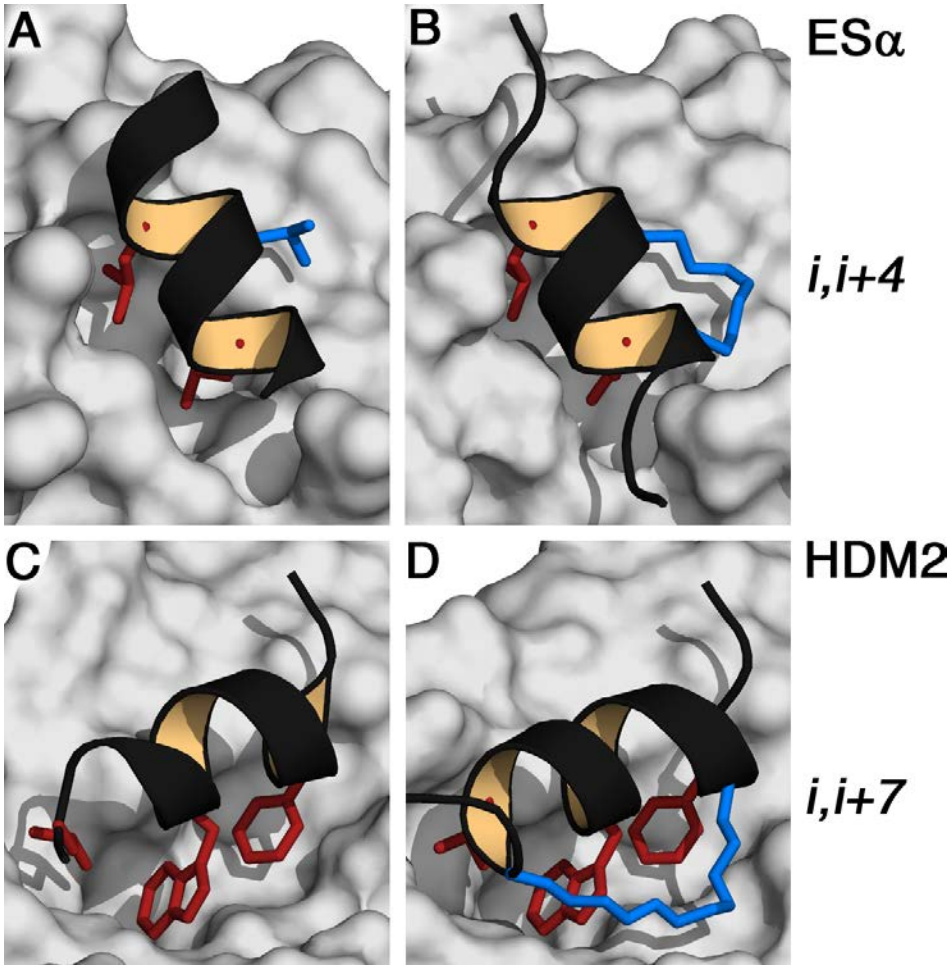

Figure 17: The wild type coactivator peptide bound to ESα (A); stapled coactivator peptide bound to ESα (B); the wild type p53 peptide bound to HDM2 (C); stapled p53 peptide bound to HDM2 (D). Key residues are shown in read; additionally Leu 693 is shown in blue (LXXLL); hydrocarbon staples are shown in blue. PDB entries: 3G03, 3V3B, 2YJA, 2QGT.

Many groups reported similar chemical approaches to crosslinking using azide click reactions or photoinduced cycloaddition schemes.[62,63] A somewhat different group of crosslinking strategies employ more elaborate photoswitchable chemical groups to stabilize or destabilize α-helix formation in response to light. These will be discussed in more detail later.

The formation of hydrogen bond surrogates (HBS) as another interesting technique to stabilise α–helical structure has been described in the literature (Figure 18). In this approach, part of the pattern of internal hydrogen bonds in the helix is mimicked by a short hydrocarbon chain to produce a peptide that is forced to adopt α-helical conformation around the modified hydrogen bond, nucleating overall helix formation.[64,65] This strategy allows for a creation of

peptides that are free to interact from any of their faces, unlike other crosslinkers which can obscure a considerable area of a face of the helix, thus imposing a compromise between crosslinker position, steric clashes and hot-spot residues mutation in any designed binders. Also, it was shown that HBS makes peptides more resistant to proteolysis (e.g. Bak HBS peptide is 30 fold more resistant to trypsin than Bak wild type peptide).[66] Interestingly, HBS peptides tend not to have dramatically increased affinity to their respective targets, although the increase is helicity between HBS and wild type peptide is often dramatic.[66,67] HBS Bak peptide affinity to Bcl-x_L is in fact decreased by two fold from 154 to 325 nM.[66] The affinity of a HBS peptide derived from Hypoxia Inducible Factor 1 to its native target CBP/p300 is two fold better as compared to the wild type peptide (825 vs. 420 nM) but *in vitro* potency to stop the transcription activated by aforementioned complex is nearly six time better for the HBS peptide.[67]

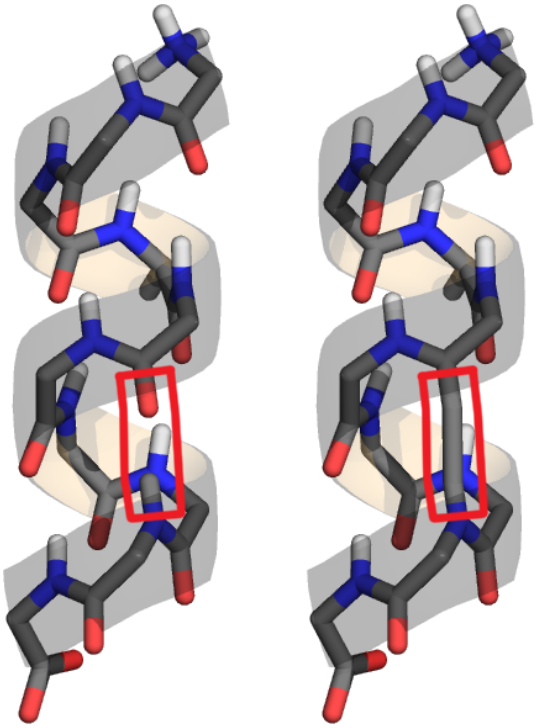

Figure 18: Representation of a standard α-helix (on the left) versus hydrogen bond surrogate α-helix (on the right). The hydrogen bond replaced by carbon-carbon bond in indicated with a red frame.

1.2.2.6 Miniature protein targeting helix-binding clefts

The miniature protein approach capitalizes upon the high stability of some known small proteins to reinforce binding helices (Figure 19). Naturally evolved structures that tolerate certain degrees of mutations are used to both transplant the binding epitope onto the helical element of a protein and retain its conformation. Usually peptides are derived from

avian pancreatic protein, Trp cage proteins, zinc finger proteins or the Z domain of staphylococcal protein A.[12]

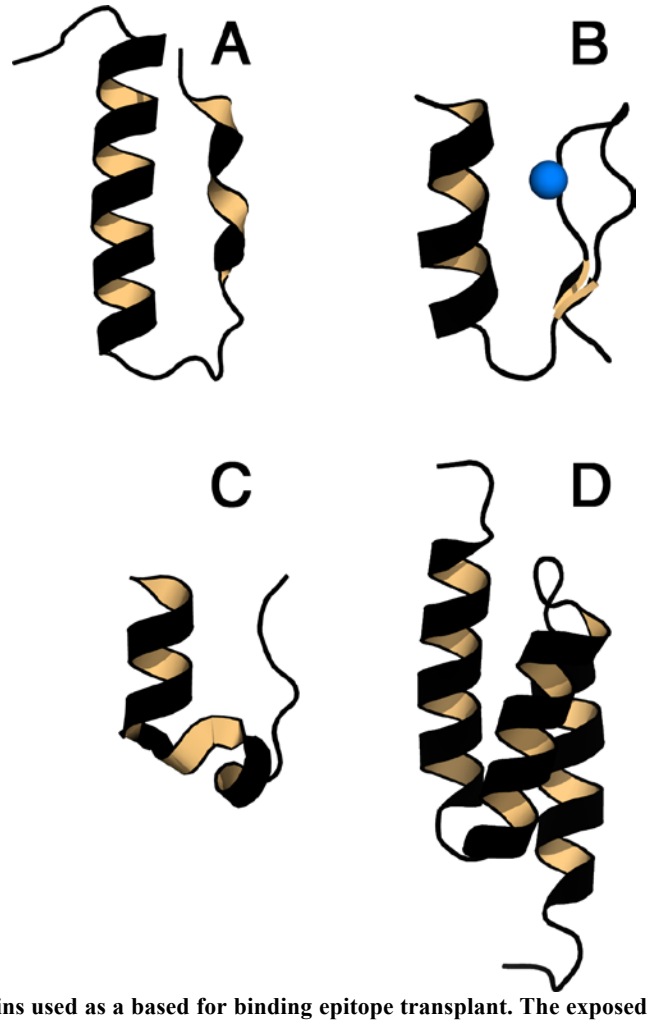

Figure 19: Miniature proteins used as a based for binding epitope transplant. The exposed side of the helix capable of accepting the mutation is always on the left side of each image; (A) avian pancreatic peptide, (B) zinc finger domain, (C) Trp-cage, (D) Z domain of staphylococcal protein A. PDB entries: 1PPT, 1L2Y, 1A1L, 2B88.

The interaction between Kinase inducible activation domain (KID) of cAMP-response element-binding protein (CREB) and KIX domain of CREB binding protein (CBP) is an unusual one (Figure 20).[68] Not only the helix-binding cleft is very shallow, but phosporylation of Ser 133 of KID is necessary for the binding as it provides crucial hydrogen bonding (the typical situation being a phosphorylation that abolishes binding).[68] Although KIX binds as two perpendicular α-helices only the C-terminal helix contributes significantly to the binding affinity. It was possible to graft the binding epitope of KIX onto avian pancreatic peptide, while retaining its native properties (Figure 20).[69] This was done via rational positioning of the crucial residues and using a phage display to provide a selection pressure over a library of randomized at 5 positions residues on the antiparallel extended helix. The first generations of peptides were able to bind KID, but still were dependent on the phosporylation of KID Ser

133; the later generations were able to bind KID regardless of phosporylation with affinity only 2 fold lower (1.5 μM) than the wild type KIX peptide to the phosporylated KID.[69]

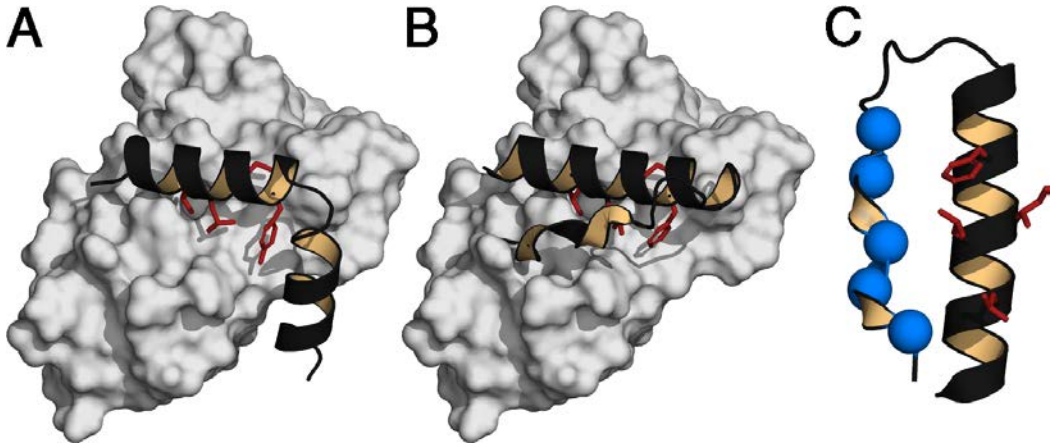

Figure 20: Solution structure models of KID-KIX complex (A); avian pancreatic overlaid with the KIX peptide in the KID-KIX complex (B); avian pancreatic peptide with four crucial hydrophobic residues grafted from KIX and randomized positions for phage display selection (C); the crucial residues of KIX are in red and randomized position are in blue. PDB entries: 1KDX, 1PPT.

p53/HDM2 and p53/HDMX interaction was disrupted using the binding epitope grafting approach. Relaying on the stability of venomous protein from scorpion's and bee's venom (apamin and BmBKTtx1 respectively) a series of inhibitors was developed (Figure 21).[70,71] Apamin served as a starting point for a peptide (stingin 1) that bound HDM2 with K_d of 25 nM and HDMX with 11 nM (respectively 3 and 8.5 nM for wild types).[70] The structure was of the stingin 1 bound to HDM2 was solved by means of X-ray crystallography (Figure 21).[70]

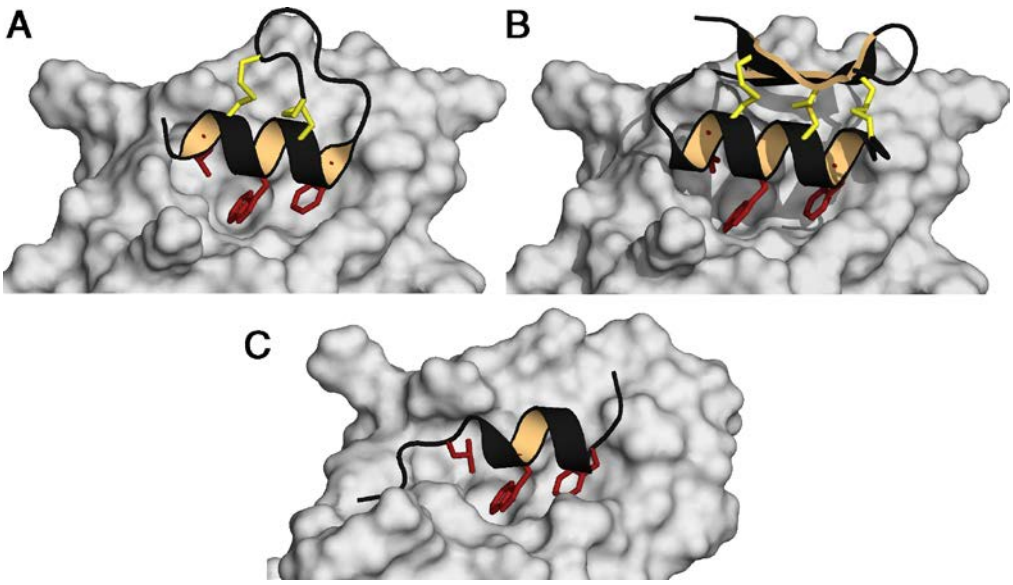

Figure 21: X-ray structural model of stingin 1 bound to HDM2 (A); stoppin (based on the structure of BmBKTtx1) overlaid onto stingin 1 bound to HDM2 (B); HDM2/p53 peptide; PDB entries: 1R1G, 3IUX, 1YCR (C)

Interestingly, the α-helices of apamin and wild type p53 peptide are positioned in respect to each other at a non-zero angle; apamin forms a regular α-helix and p53 start as a regular one and gets tighter towards C-terminus (Figure 21). BmBKTtx1 scorpion toxin is heavily stabilized by 3 disulfide bridges and is a markedly bigger peptide (18 vs. 27 amino acids).[71] Stoppin 1 construct has a K_d of 790 and 994 nM, correspondingly to HDM2 and HDMX, as compared to the wild type peptide in this assay 123 and 279 nM respectively.[71] It was suggested by the authors that his relatively low binding affinity is a consequence of different positions of the helices (stoppin 1 vs. p53);[71] this explanation is unsatisfactory, because it is not congruent with the apamin/HDM2 system.

Bcl-xl and Bcl-2 proteins were targeted by miniature proteins derived from avian pancreatic peptide. Remarkably, those designs managed to outcompete the wild type Bak peptide, which they aimed to mimic, by nearly 2 orders of magnitude achieving low nanomolar affinities (Table 1).[72] Moreover the group was able to reverse the selectivity of the designed peptide, which originally was bound Bcl-x_L one order of magnitude better than Bcl-2; the reversal of selectivity pushes the affinity into the high nanomolar range (Table 1).[72]

Peptide/Sequence	Affinity to Bcl-x_L (nM)	Affinity to Bcl-2 (nM)
Bak	180	6100
PPBH3-1	7	52
PPBH3-6	5400	540

Table 1: Affinities of Bak, PPH3-1 and PPH3-6 miniature proteins to Bcl-x_L and Bcl-2.

1.2.3 Protein-RNA interactions

Protein-RNA interactions are characterised by markedly different features compared to protein-protein interactions. Unlike protein-protein interfaces where protein surfaces evolved to contain large hydrophobic patches, RNA is composed of aromatic nucleobases attached *via* polar ribose sugars to a negatively charged phosphate backbone, limiting the opportunity to make hydrophobic interactions and making positively-charged amino acids essential.

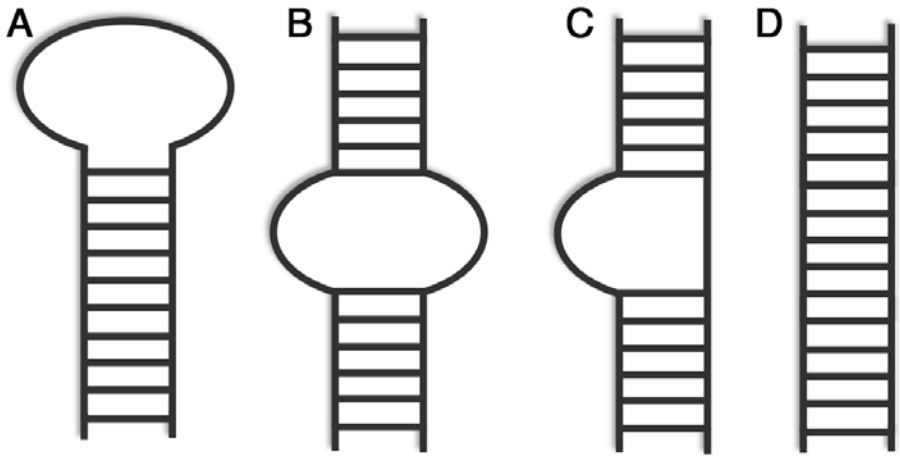

Figure 22: Four standard RNA secondary structure elements: (A) hairpin, (B) loop, (C) bulge, (D) duplex.

The conformational flexibility of single stranded RNA can allow access to a variety of secondary structures exposing different features that can differentiate RNA from less flexible double stranded DNA (Figure 22).[73]

Statistical analysis revealed that positively-charged arginine and lysine have the highest propensity to be found in the interface of protein-RNA complexes.[74-76] Other residues frequently reported include the aromatic amino acids tyrosine and tryptophan and the hydroxyl bearing amino acids serine and threonine. Residues that seem to have consistently low propensity are negatively charged (Glu, Asp) or hydrophobic (Ile, Leu, Val, Ala, Cys, Phe).[76] The high frequency of positive residues can be easily attributed to electrostatic interactions with the phosphate backbone. Tyrosine and tryptophan can intercalate/stack between aromatic nucleosides and form hydrogen bonds, as can serine and threonine. Hydrogen bonding seems to be very important for specificity, the hydroxyl groups of ribose are often found engaged in donating or accepting hydrogen bonds.[77] Hydrogen bonding is frequently mediated through water molecules, which seem to play more significant role than in protein-DNA or protein-protein interactions, with large water filled cavities frequently found as a part of a binding interface.[78]

The average protein-RNA interface area is reported to be similar to that of protein-protein interfaces[75,79] with most contacts being between amino acids and the phosphate backbone and ribose. Although hydrogen bonds are crucial, van der Waals dipolar interactions are even more numerous. The involvement of protein secondary structure elements is not frequent enough to assume high statistical significance, but it appears that helices are much less frequent than previously thought and β-hairpins predominate.[75] The

notable exception to this is in viral RNA, where α-helices have a higher propensity to be used.[75] Clearer trends may occur as more protein-RNA structures are solved.

1.2.3.1 Small molecule inhibitors targeting RNA

Small molecules that can target RNA and disrupt the protein-RNA interactions are relatively scarce and predominantly target interactions crucial for HIV-1 life cycle, oncogenic mRNA or the bacterial ribosome A-site. The HIV-1 interactions are the Rev Response Element – Regulator of Viron (RRE–Rev) and the Trans-Activation Response Element–Transactivator of Transcription (Tar–Tat) complexes. High throughput screening and rational design has yielded a few groups of compounds capable of binding Tar and RRE sequences with high affinity (Figure 23).[80-84] In turn, these have allowed NMR and crystallography studies to reveal structural details of their interactions providing clues as to how to improve their designs.[82,85-87]

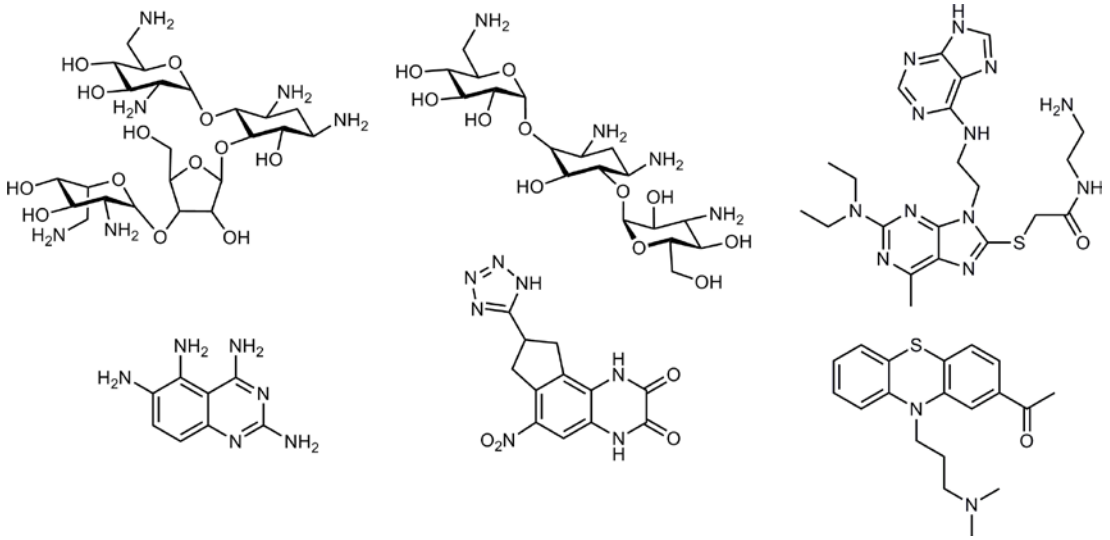

Figure 23: Various successful inhibitors of Tat-Tar and RRE-Rev interactions

Unfortunately, of the two crucial RNA-protein interactions RRE-Rev and Tat-Tar the latter one is more frequently featured in literature as a target of small molecule inhibitors research. One of the high throughput screenings revealed an *in vitro* potent binder[85]; the compound bound Tar RNA with K_d of 39 nM. The NMR structure revealed an interesting feature of the formed complex (Figure 24). Although the inhibitor binds approximately in the same place as the Tat peptide the RNA adopts a dramatically different conformation.

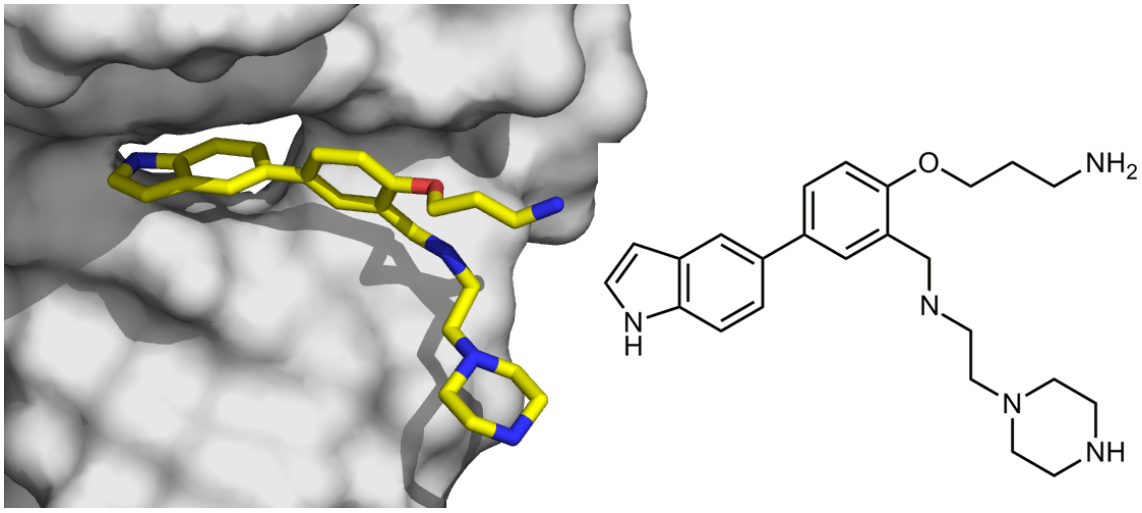

Figure 24: Tar-Tat interaction inhibitor bound to Tar (left); structure of the inhibitor (right). PDB entire: 1UTS.

Oncogenic mRNA (result of a transcription of fused genes *BCR-ABL* and *PAX3-FKHR*) has been targeted with aminoglycosides derivatives,[88] based on scaffolds like neomycin B or kanamycin. Most molecules used to bind to the A-site of the bacterial ribosome are also based on aminoglycosides (neamine dimers, neomycin E, geneticin).[89,90] They form a class of relatively old antibiotics which block bacterial translation (Figure 25).

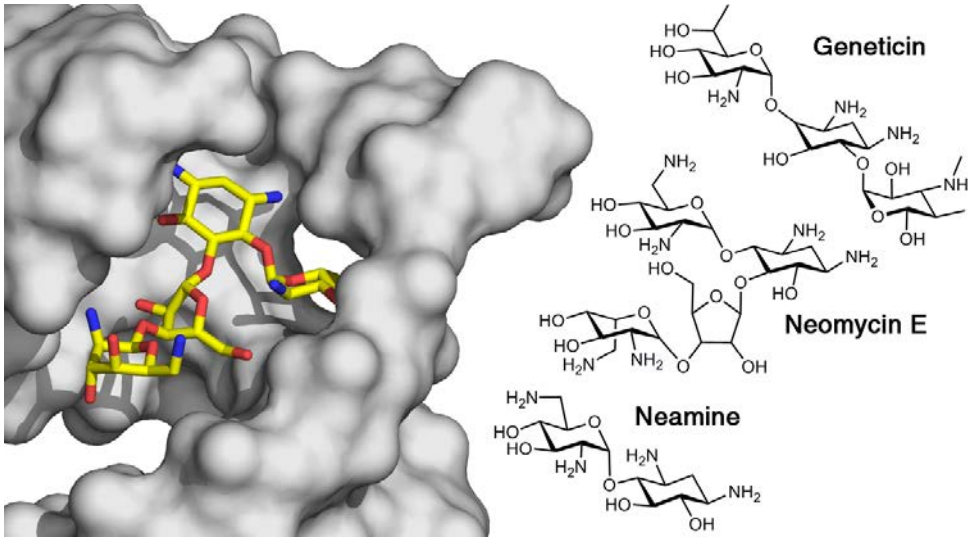

Figure 25: Structure of Neomycin E bound to RNA derived from ribosome A-site (left); structures of three common aminoglycoside antibiotics (right). PDB entries: 1J7T.

Newer targets include the RNA that causes myotonic dystrophy 2 or 1 *via* interaction between the r(CCUG) expansion in zinc finger 9 or rCUG expansion in the dystrophia myotonica protein kinase and muscleblind protein respectively, which result in an alternative, pathological splicing of the main muscle chloride channel and the insulin receptor. Both interactions have been successfully targeted by a modular approach where peptoid sidechains were functionalized with azido-aminoglycosides.[91,92] This study not only yielded a potent

binder, but created a basis for rational structure-activity relationship studies of further RNA targets.[91,92]

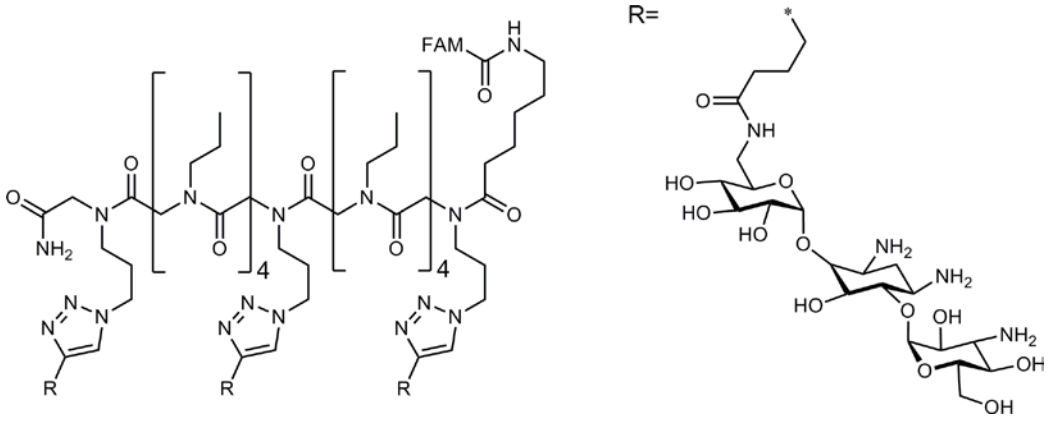

Figure 26: Structure of the peptoid binder of r(CCUG) expansion.

1.2.3.2 Peptides targeting RNA

Tar and RRE are the major focus of research towards RNA binding peptides. Peptides mimicking the original Rev sequence or comprised of new, independently evolved arginine rich sequences have been crosslinked with orthogonal amide bonds.[50,93] The work capitalizing on the evolved arginine rich sequences lead to a development of a peptide that has two fold better affinity and specificity to RRE (100 vs. 45 nM and specificity of 16 vs. 26 nM).[93]

Key Rev residues have also been transplanted to zinc finger scaffold.[94] This bulky designed peptide showed Zn^{2+} dependent helical content and binding affinity.[94] The affinities of a Rev^{wt} isolated helix and the designed zinc finger peptide were comparable at the assay conditions (330±40 nM vs. 330±20 nM).[94] Unfortunately, although no structure of this complex was solved by overlaying a structure of the Zinc Finger (Zif268) that was used as a starting point reveals how this relatively bulky miniprotein fits into the very deep binding grove of RRE and at the same time all key residues of Rev are retained (Figure 27).[94] That type of approach was employed and another modified Zinc Finger targeting RRE was developed and its structure was solved by NMR methods.[95,96] The binding affinity was improved to 55 nM and the actual sequence that forms the binding epitope bears very little resemblance to the original Rev; three residues are retained of which only one is considered to be the key residue.

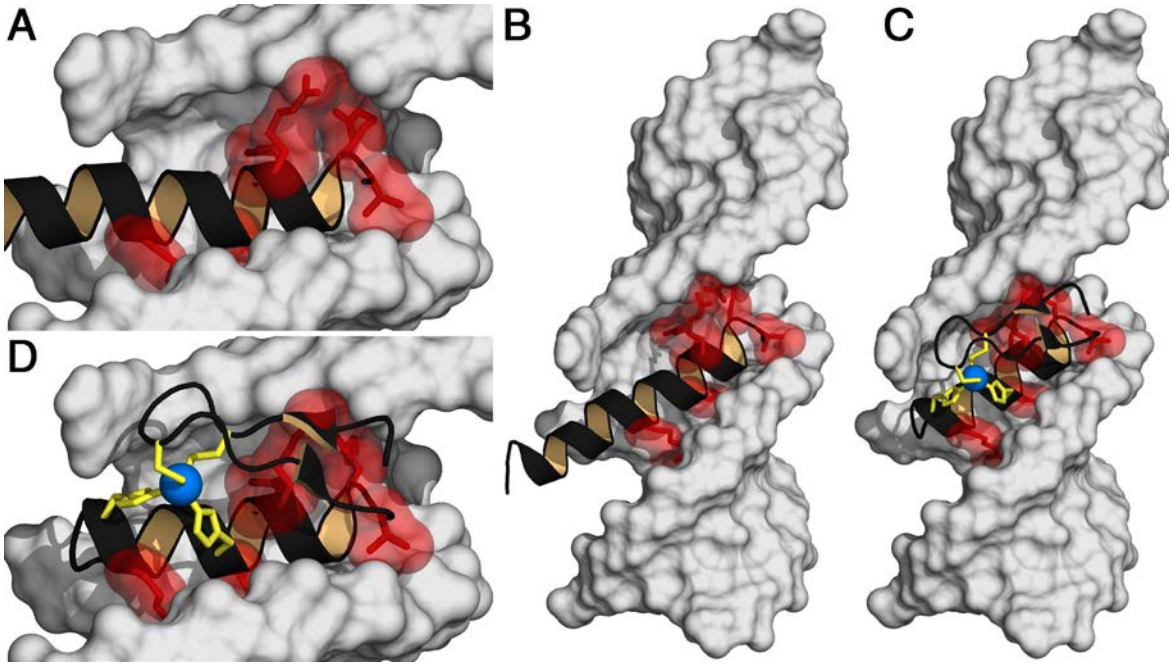

Figure 27: NMR structural model of the Rev peptide bound to RRE (A and B); Zinc Finger (Zif268) with Rev binding epitope transplanted overlaid onto Rev in complex with RRE structure (C and D). Key Rev residues are in red; two cysteines and histidines chelating a zinc ion are in yellow; zinc ion is in blue. PDB entries: 1ETF, 1A1J.

The sequence was not rationally designed but evolved through phage display.[95,96] Stunningly, by overlaying this Zinc Finger – Rev hybrid onto the Rev/RRE structure it becomes apparent that the beta-sheet part of the Zinc Finger clashes severely with RNA. This suggests a powerful adaptable potential of RRE structure which either undergoes dramatic rearrangement or binds to a very different amino acid patter and the sequence based overlay is misleading.

The importance of stereochemistry and directionality of the sequence of Rev peptides have been thoroughly investigated by Litovchick, A. *et al.*[97] D-Amino acid mirror image as well as reverse sequences of Rev were synthesized and their binding to RRE IIB RNA assessed.[97] Results show a remarkably small differences between these constructs, the highest being approximately four fold.[97] Constrained β-sheet peptides that mimicked a crucial β-sheet section of the Tat protein were synthesised and were found to bind Tar well[98-100] as did cysteine-constrained cyclic peptides that targeted the bovine immunodeficiency virus equivalent of Tar.[101] The latest peptide found by Robinson *et al.* had affinity of 1 nM (25 nM in the presence of tRNA excess) and its structure has been solved by NMR techniques (Figure 28).[100] The peptide incorporates D-Pro-L-Pro template to promote a β-hairpin formation (D-Pro at C- and L-Pro at N-terminus) and a disulfide bond in the middle of the sequence to further stabilize the interaction.

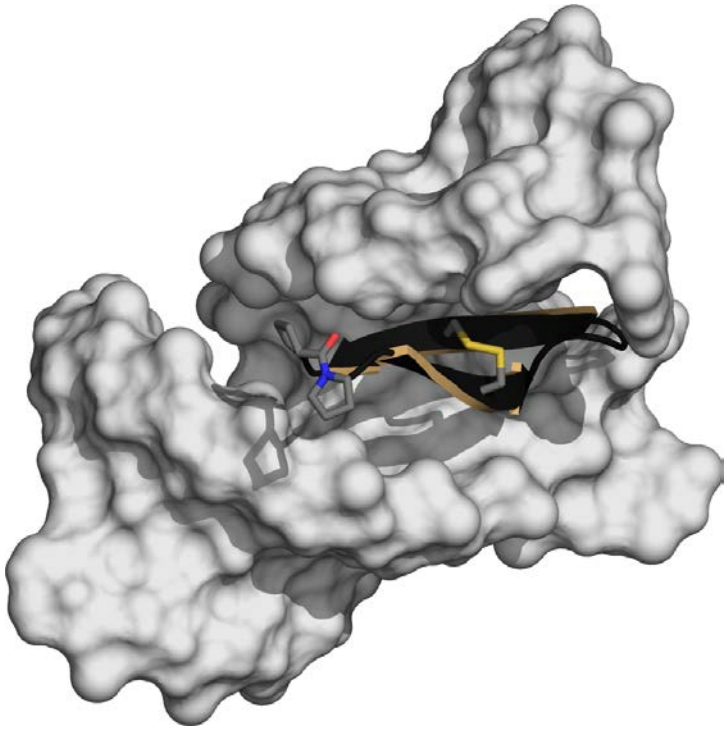

Figure 28: Structure of the cyclic Tar-derived evolved peptide bound to Tar RNA. D- and L-Pro are in gray and the disulfide bond is in yellow. PDB entries: 2KX5

Curiously, it was also possible to use a similar approach (β-sheets) to design inhibitors of Rev/RRE interaction by mimicking key residues position of Rev.[102] Small cyclic β-sheets capable of mimicking α-helical sidechain positions were designed and effectively targeted RRE with the affinity similar to the Rev^{wt}. A K_d of 100 nM was reported for the best designs and Rev peptide (in the presence of large excess of tRNA but in a low ionic strength buffer).[102]

A study of the acceptor stem of an *Escherichia coli* alanine tRNA identified a 10 residue peptide recognition element that was able to discriminate between G-U and G-C, U-G and inosine (I)-U base pairs.[103] In particular, peptides used to discriminate G-U versus I-U differ by only a single amino group, showing that this short peptide is capable of discriminating this minute change in the chemical composition of RNA (Figure 29).[103] Unfortunately no K_d or IC50 values are quoted and only offer a graph visualizing displacement of the wild type RNA (G-U containing) from a the RNA-peptide complex with the same RNA (G-U containing) or mutated (I-U containing).[103]

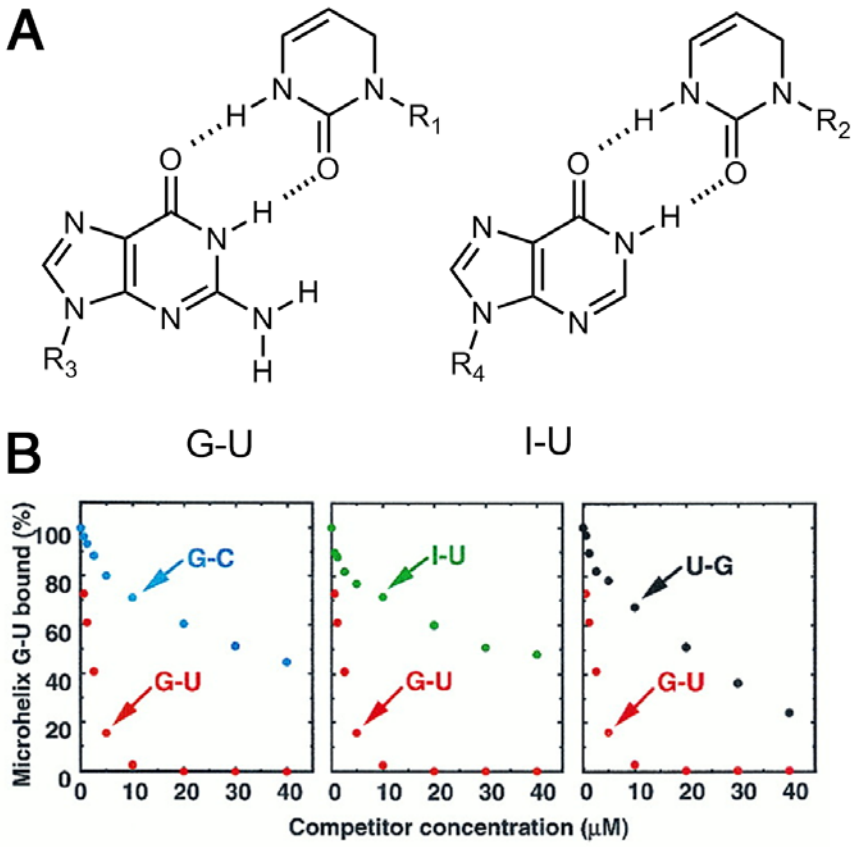

Figure 29: G-U and I-U base pairs (A); graphs showing displacement of wild type RNA (G-U containing) from the RNA-peptide complex with wild type (G-U containing) or mutated RNA (I-U or G-C or U-G containing); reproduced from[103] (B).

Another short peptide (three residues) was crosslinked with an intercalating group to form a so called helix-threading peptide and successfully bound to microRNA (miRNA) with the affinity as high as 52±17 nM.[104,105] Similarly, so called γ -AA (with sidechains protruding from nitrogen) peptides were able to bind well to RNA.[106]

1.3 HIV and AIDS

1.3.1 Overview

Human immunodeficiency virus type 1 (HIV-1) is a virus that belongs to the *Retroviridae* family.[107] It is a close relative of simian immunodeficiency virus (SIV), which is believed to be an HIV-1 ancestor that was transmitted from chimpanzee (*Pan troglodytes troglodytes*) to humans in a process called zoonosis.[108] The origins of HIV-1 can be traced to the late 19th or early 20th century and traditional hunting activities in Africa.[109,110] The current hypothesis is that hunters could be infected with SIV through the blood of killed chimpanzees and subsequently acquired mutations allowed the virus to cross the species barrier and find a

new stable host in humans.[109] It is established that HIV is the factor responsible for acquired immunodeficiency syndrome (AIDS); a disease that slowly depletes the immune systems defences leaving individuals prone to opportunistic infections. In 2009 AIDS was the cause of death for 1.8 million people. With 33.3 million people infected AIDS is a pandemic with no vaccine currently available.[111,112] AIDS is not only a major health problem in the developing countries it originates from (sub-Saharan Africa), developed ones are heavily affected by this disease.[112]

1.3.2 Biology of the HIV virus

HIV infects some lymphocytes including CD4+ T cells (mature helper T cells) with a particular preference for CD4+ T cells that recognise HIV.[113] The crucial role of CD4+ T cells in activating the immune response and the decline in number of these cells following the HIV infection is the direct cause of AIDS. The HIV virus is capable of recognizing CD4 protein, expressed abundantly in CD4+ T cells but also to lesser extent in macrophages and dendritic cells,[107,114] through interaction with HIV envelope glycol protein 120 (gp120).[107]

1.3.3 Structure

A typical HIV virion is 145±25 nm in diameter and approximately spherical.[115,116] It is surrounded by a lipid bilayer acquired from the host cell during budding (Figure 30) which is abundant in membrane embedded proteins including glycoprotein 41 (gp41, TM) bound to glycoprotein 120 (gp120, SU) which plays a crucial role in virus adsorption and entry.[117] Within the membrane, packed in protein 17 and protein 15, resides the viral capsid which is composed of protein 24 and encloses 2 copies of single stranded positive sense HIV RNA tightly bound to protein 7.[107,115] Three HIV-coded enzymes are also present inside the capsid: protein 10 or HIV protease, protein 66/51/reverse transcriptase and protein 32/integrase.[115] No crystal structure exists of the whole virus; however a low-resolution reconstruction of the 3D structure has been made based on an electron micrograph.[115,116]

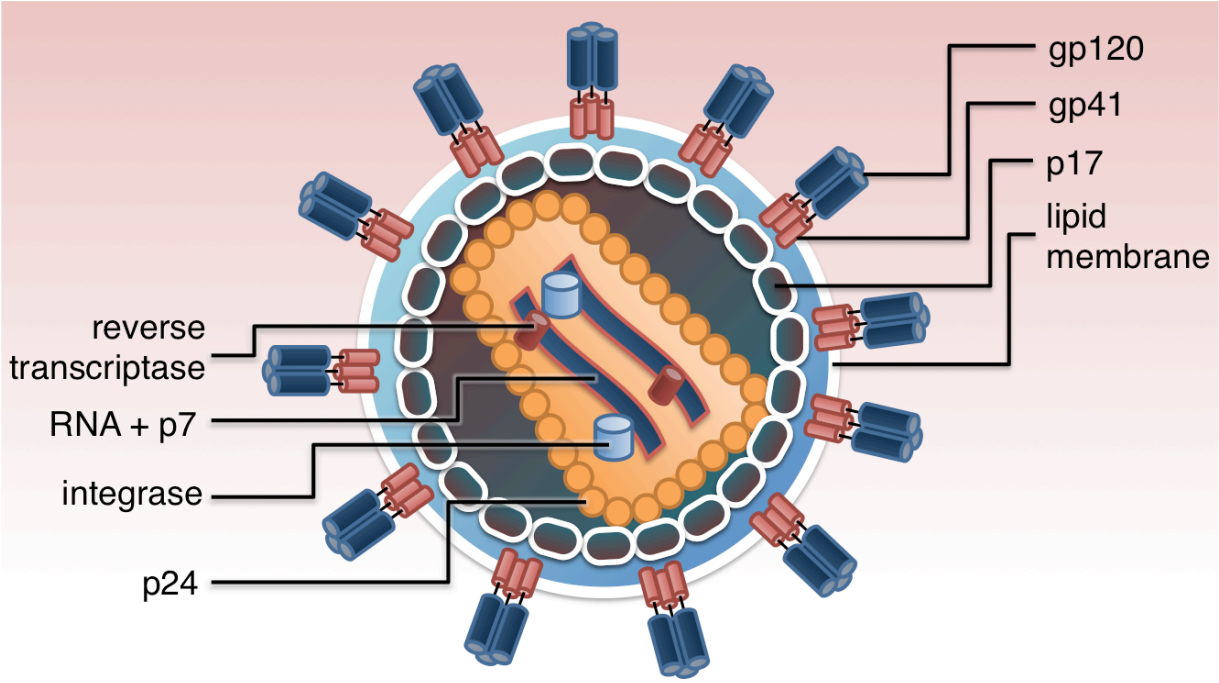

Figure 30: Schematic representation of the structure of the HIV-1 virion. Some proteins present inside were omitted for clarity.

1.3.4 Genome organization and coded proteins

The HIV genome is around 9 kbp long and possesses 3 open reading frames containing the genes: 1 – *gag, vif, nef*; 2 – *vpu*; 3 – *pol, vpr, and env* (Figure 31). Tat and rev are separated between the 1st/2nd and 2nd/3rd ORFs, respectively.[118] Viral pre-mRNA can be subjected to none, single or multiple splicing events giving rise to 9, 4 and 1.8 kbp splice variants, respectively.[118,119] Unspliced mRNA can either form the HIV genome or can be translated into Gag-Pol or Pol precursor multiproteins that are autocatalytically cleaved into the component proteins of the virion, while multiple splicing produces mRNA for the regulatory proteins Rev, Tat and Nef.[107] Interestingly the secondary structure of HIV RNA inversely corresponds to protein tertiary structure; unstructured regions of RNA correspond to highly structured regions on multiproteins and vice versa, this allows for pauses in translation and independent folding of domains/proteins of multiprotein HIV gene products.[118] Only multiply spliced mRNA does not contain the RRE sequence.[120] HIV-1 genome terminates in 5' and 3' long terminal repeats (LTR).

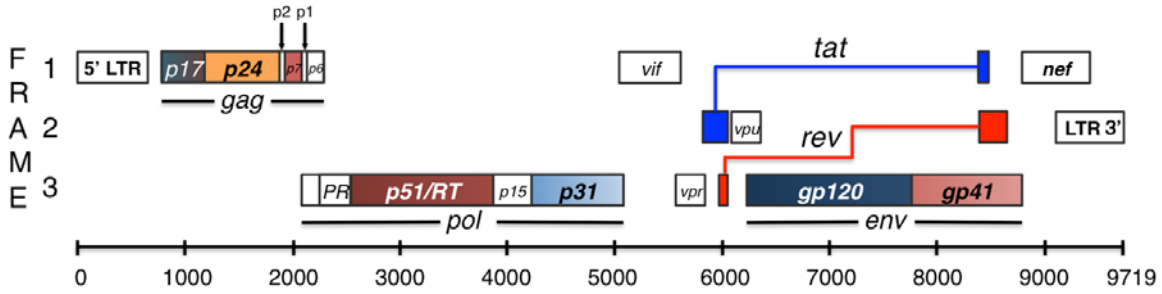

Figure 31: Map of the HIV-1 genome colour coded according coherently to Figure 30. *Tat* and *rev* genes are indicated in blue and red respectively.

1.3.5 Regulatory elements

Initial transcription and translation of viral genes is slow and only fully spliced HIV-1 mRNA is produced. The fully spliced variants include mRNA that codes for Tat protein, which is responsible for transcription initiation complex assembly. Tat binds to the Tar sequence present in the LTR.[121] Tat phosphorylates RNA polymerase II and recruits essential transcription factors, which increase viral gene expression by more then two orders of magnitude.[121] On the other hand, Rev (also translated from fully spliced mRNA) is responsible for the nuclear export of singly spliced and unspliced mRNA variants. [120,122] Rev is imported to the nucleus and binds the RRE sequence present in singly spliced and unspliced variants and promotes their export to the cytoplasm and subsequent translation.[120,122,123]

1.3.5.1 RRE/Rev

RRE is a 350 nucleotide long RNA sequence that has not yet been completely mapped (Figure 32).[118] It overlaps with the gp41 coding region and it contains a large number of RNA secondary structure elements.[118] The precise secondary and tertiary structure of the RRE is still a matter of debate and somewhat different RNA maps have been proposed in the literature.[118,124] However, it has been established that minimal RRE length required for function is approximately 250 bp and encompasses regions I to V.[125] Rev is a 116 amino acid long protein whose complete structure remains unknown.[126,127] Only the N-terminal region has been structurally characterized, which has been shown to form an eccentric helix-loop-helix motif.[127]

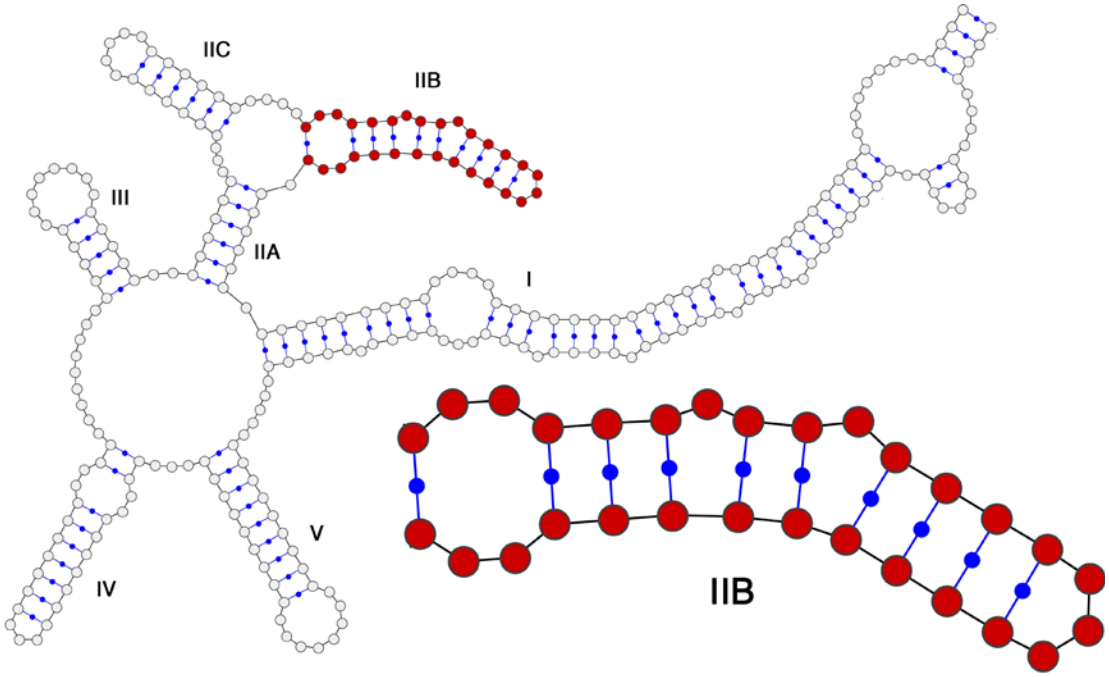

Figure 32: Topology of the core RRE element with binding sites annotated appropriately and RRE IIB indicated in red (right); enlargement of the RRE IIB element (bottom right)

This part of Rev contains the arginine-rich region known to bind with high affinity to RRE; a peptide derived from the arginine-rich sequence is capable of strongly interacting with RRE (Figure 33).[128-130] Rev has the highest affinity for the stem loop IIB part of the RRE sequence.[131] The N-terminal region of Rev is known to be responsible for oligomerization; in the current model, binding of the first molecule of Rev to RRE IIB nucleates oligomerization and binding to the remaining sites on the mRNA.[125,127] Nuclear localization sequences (NLS) which overlap with the arginine-rich region are concealed by this interaction and oligomerization organises nuclear export sequences into a multivalent array, which can interact sufficiently strongly with Chromosome Region Maintenance 1 protein (Crm1) protein to facilitate nuclear export.[127,132] This allows singly spliced and unspliced mRNAs to be exported from the nucleus before further splicing occurs, thus allowing the remaining HIV-1 genes to be expressed.[123]

The arginine-rich region of Rev spans residues 34 to 50 and its sequence is TRQARRNRRRRWRERQR, where the underlined residues are those which decrease affinity by more than 10-fold upon mutation to alanine (Figure 33).[133] The structure of this and similar peptides bound to RRE IIB have been solved by NMR and crystallography techniques.[133-135]

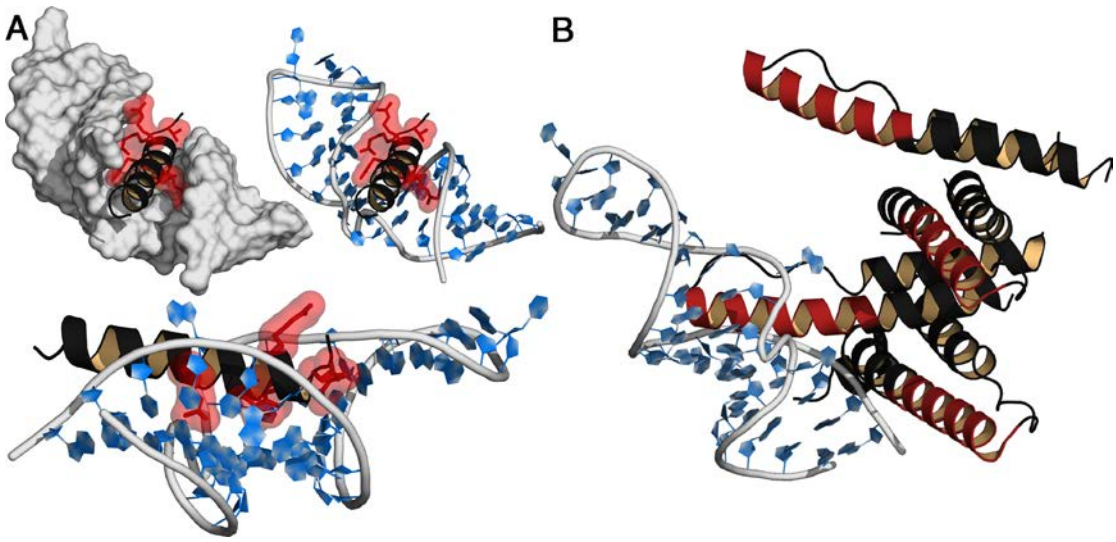

Figure 33: Structural models of Rev peptide bound to RRE IIB RNA (A) with key residues in red; N-terminal domains of Rev protein dimer of dimers aligned onto Rev bound to RRE IIB (B) with arginine-rich region in red. PDB entries: 1ETF, 3LPH.

Structural reports on the RRE/Rev interaction all describe a remarkable structural flexibility of the RNA structure, which in order to accommodate the Rev peptide, needs to widen by approximately 5 Å (Figure 34).[133] Moreover, in the complex the peptide sits in a very deep grove, with 270° of its surface surrounded by RNA (Figure 34).

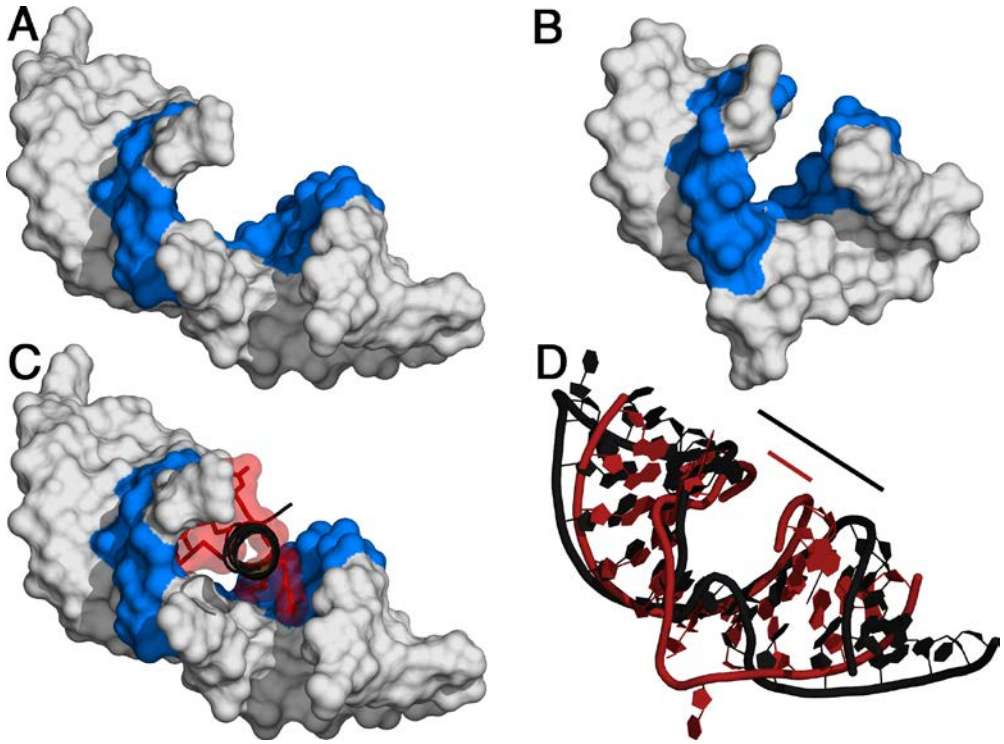

Figure 34: Structural models of the complex of Rev derived peptide (not shown) and RRE IIB (A); free RRE IIB (B); complex of Rev derived peptide and RRE IIB with residues that decrease binding affinity by more then 10 fold indicated in red (C); overlay of RREIIB free (red) and bound (black) to Rev derived peptide (not shown); black and red bars are proportional to the grove width. PDB entries: 1ETF, 1DUQ.

1.4 Photocontrol of Structure

1.4.1 Natural examples

During the course of life's history light not only has been used as a source of energy throughout all kingdoms of life, but light sensitive switching mechanisms have evolved independently many times. Even the simple unicellular organism *Chlamydomonas reinhardtii* possesses an "eyespot" that allows it to exhibit phototaxis and photophobia reactions.[136] The proteins, which carry out this function, are light-gated ion channels: channelrhodopsin-1 (ChR1) and channelrhodopsin-2 (ChR2).[137] They contain a light sensitive chromophore, retinal, which undergoes a reversible isomerisation from all-*trans* retinal to 13-*cis*-retinal. This conformational change in the chromphore induces a larger conformational change in the transmembrane region of the channel, opening it to both monovalent and divalent cations. Within micro seconds 13-*cis*-retinal thermally relaxes to its ground state, which eventually closes the channel.[137] A similar mechanism can be observed in the human retina. The retinal containing protein rhodopsin is a transmembrane protein, but unlike ChR it is not an ion channel, but a G-protein coupled receptor.[138] Interestingly the ground state of its retinal cofactor is 11-*cis*-retinal and upon absorption of light it changes configuration to all-*trans*. The resulting conformational change in the protein activates interacting G-protein on the cytosolic side of the membrane.[138] A different chromophore is used by phytochromes, a group of proteins found in plants, bacteria and fungi, which incorporate a billin-derived chromophore.[139] In most cases two thermodynamically stable forms of this chromophore exist that can be shifted between one another upon absorption of red or far-red photons (Pr and Pfr forms).[140,141] Therefore, unlike the retinal based light-sensing proteins, the change in chromophore conformation is not immediately reversed to the previous state.[140,141] Plants have evolved a whole ensemble of light sensing proteins. In addition to few different types of phytochromes, cryptochromes, phototropins and superchrome classes exist. Phototropins use flavin as their chromophore (contained within LOV domains),[142] the absorption of light and subsequent formation of a bond between flavin and a cysteine residue in its binding pocket abolishes an interaction essential for the stabilization of amphipathic, and partially water exposed helix.[143] Release of this helix can trigger kinase activity and transfer the signal further down a pathway (Figure 35).

Figure 35: Structural model of a LOV domain in the dark state (black); a model of the helix unfolding after blue light irradiation is schematically represented (red); flavin chromophore is visible in the protein hydrophobic core (red). PDB entries: 2VOU.

1.4.2 Photo-control of structure in chemistry

One of the earliest examples of light induced control of synthesized compounds structure involves a group of molecules called azobenzenes (Figure 36). These chemical compounds are capable of photo-isomerisation around the N=N double bond.[144] Usually the *trans* configuration is thermally stable and upon absorption of a photon of approximately 360 nm wavelength, the *trans* configuration switches to *cis*, creating a more compact molecule with end-to-end distances between *para* ring positions differing by approximately 3.5 Å compared to the *trans* form.[145] The *cis* form of the azobenzene thermally reverts to *trans* and this process can be accelerated by illumination with light matching the absorption maximum of the *cis* form (450 nm).

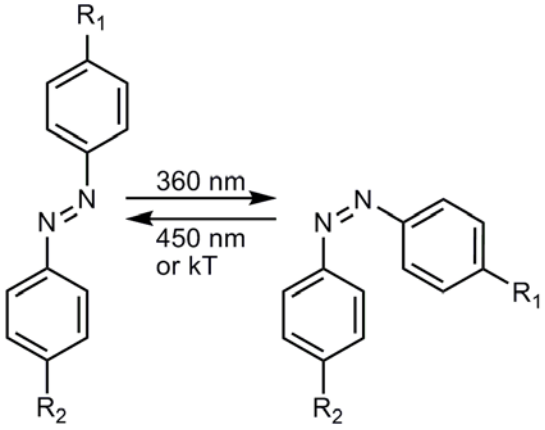

Figure 36: Azobenzene *trans-cis* photoisomerization; R_1 and R_2 indicate different possible substituents.

Different substituents around the aromatic ring systems can alter the overall properties of an azobenzene chromophore, affecting its reactivity, absorption maximum, thermal relaxation rate, and state *cis*/*trans* ratio at the photostationary state.[146] The *cis* form of azobenzene is usually thermally unstable, but when a bridged (constrained) version of azobenzene was synthesised, the *cis* state was found to be thermally stable and could be switched to *trans* by irradiation in a structural transition that is accompanied by a dramatic change in colour.[147]

Other popular photo-switchable molecules include derivatives of spiropyran (Figure 37). Upon irradiation of spiropyran it is isomerised to a merocyanine form. As a result the initially largely hydrophobic compound becomes zwitterionic. This change is thermally reversible and can be made more rapid by irradiation at the absorption maximum of merocyanine. Those properties have been used to control rheology by doping micelles with spiropyran-derived detergent, force light-induced patterning or control liquid-crystal phases.[148-150]

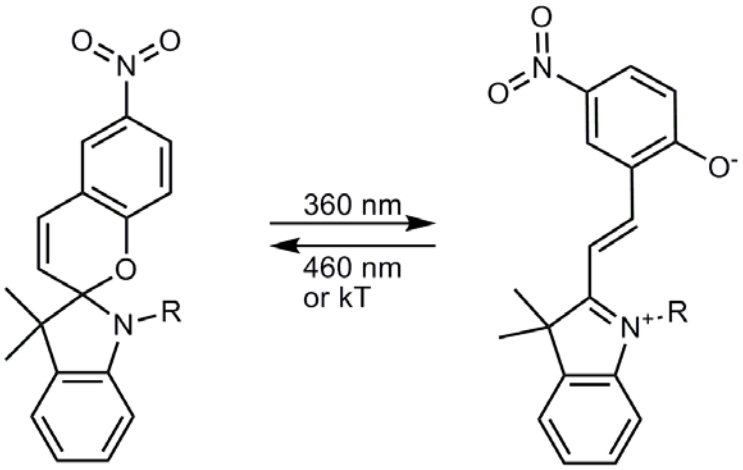

Figure 37: Photoisomerization of spiropyran to merocyanine; R indicates a possible substituent.

Researchers are keen to create light-driven motors, ideally producing unidirectional, light-driven motion at molecular level to mimic in naturally occurring motor complexes. Considerable successes have been achieved using both helical alkenes and rotaxanes. In the former approach a chiral, helical alkene is designed such that its two subunits are connected by a double bond and one type of rotation, *anti* or clockwise has lower energy barrier (Figure 37).[151] Light can then be used to isomerise the molecule around the double bond. A different wavelength of light is then used in order to rotate back to initial state.[151]

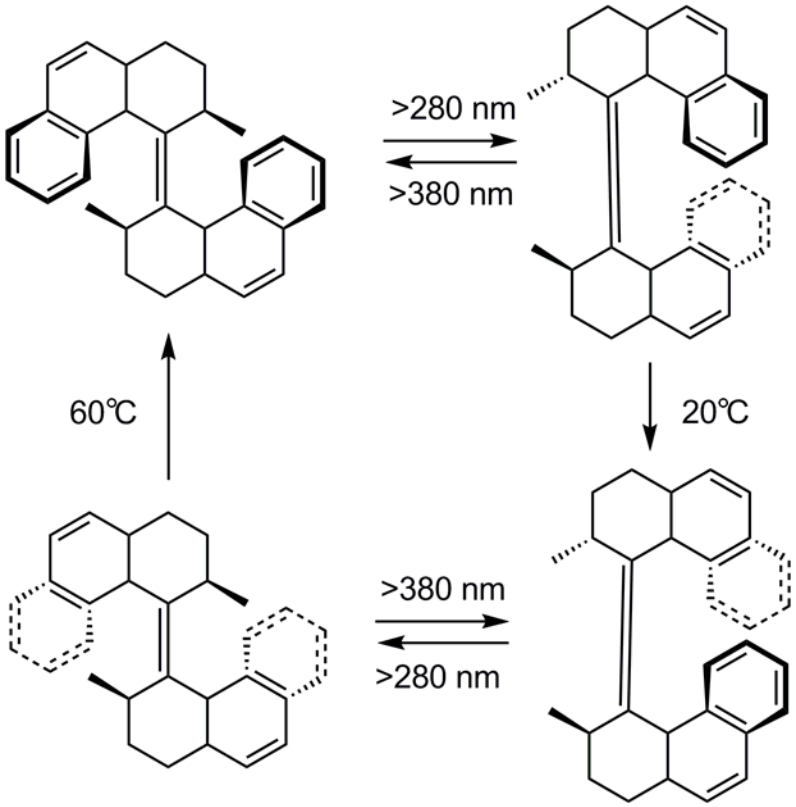

Figure 38: Helical alkene cycle. Optical filters cut-off and temperatures used for the isomerisation are noted appropriately.

In the latter approach macroassemblies called rotaxanes are used. Those systems are composed of long molecules with bulky chromophores at each end, trapped within another molecule in the shape of a closed ring (Figure 39).[152] The change in protonation states induced by light changes the binding preference of the ring molecule for chromophore over the other, this process is reversible by light and represents a prototypical molecular piston.[152]

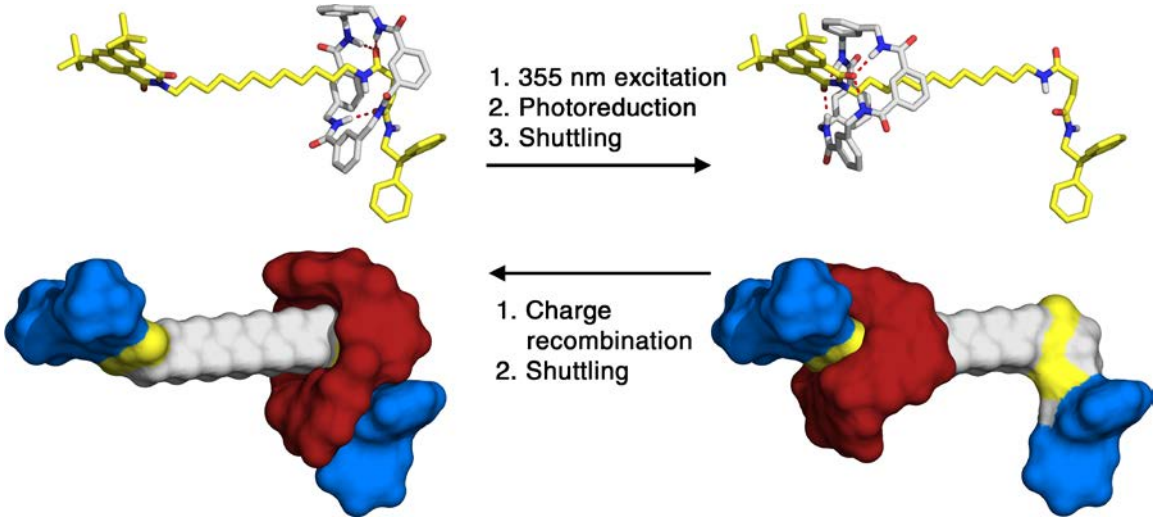

Figure 39: Representation of the molecular shuttle cycle; hydrogen bonds are indicated with a dashed red line (upper); surface representation with bulky "stopper" substituents in blue, macrocycle "shuttle" in red, hydrogen bonding interaction sites in yellow (lower).

1.4.3 Photocontrol of protein structure

1.4.3.1 Overview

Three typical approaches to protein-structure photocontrol are found in the published literature: photo-caging, conjugation of photo-switchable molecules and optogenetics. The former pair relies on the introduction of light-reactive molecule to otherwise natural system, while optogenetics capitalizes on naturally evolved proteins which, after appropriate genetic-engineering, serve as photoswitches in a variety of arrangements. The latter two have the significant advantage of possessing a degree of reversibility.

1.4.3.2 Photocaging

Photo-caged molecules are the oldest way of introducing light reactivity to polypeptides. The general strategy introduces a photo-labile group into a protein or peptide that either forces a peptide into certain conformation or masks amino acids that are responsible for interaction. Upon irradiation with light the chemical group is cleaved off. Usually peptides revert to their natural (or close to natural) function and therefore are activated, but there exist examples of the opposite switching effect. Recent research describes photocaging of an arginine-glycine-aspartate peptide widely recognised for its ability to promote cell adhesion (Figure 40).[153] In this case a 2-nitrobenzyl group was introduced at the peptide bond to create a peptoid incapable of binding its partner protein. Brief irradiation with 365 nm light removed the 2-nitrobenzyl group and the resulting peptide regained its ability to stimulate cell adhesion.[153]

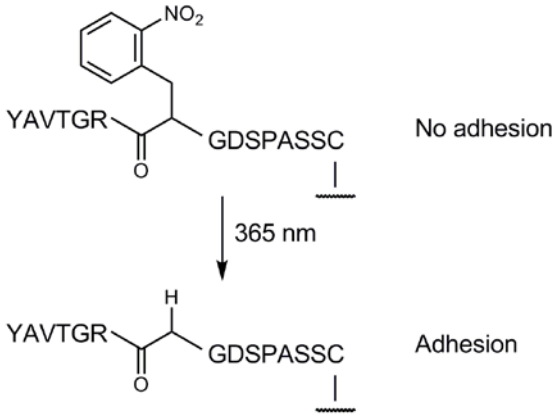

Figure 40: Photocleavage of the 2-nitrobenzyl blocked arginine-glycine-aspartate peptide.

Interestingly, despite UV light exposure taking place after the cells to be tested for adherence had been introduced to the experimental setup, this group reports no signs of deterioration induced by harmful radiation.[153]

Similarly, introducing the photolabile group directly at a peptide bond protected a peptide against proteolytic cleavage, this protection is lost and peptide degraded after UV irradiation in presence of a proteolytic enzyme.[154]

1.4.3.3 Attachment of a photo-switchable molecule

The photo-switchable molecules most widely employed in attempts to design photocontrolled proteins or peptides are spiropyrans and azobenzenes. The earliest attempts to use azobenzene derivatives were based on synthesizing poly-L-p-(phenylazo) phenylalanine.[155,156]

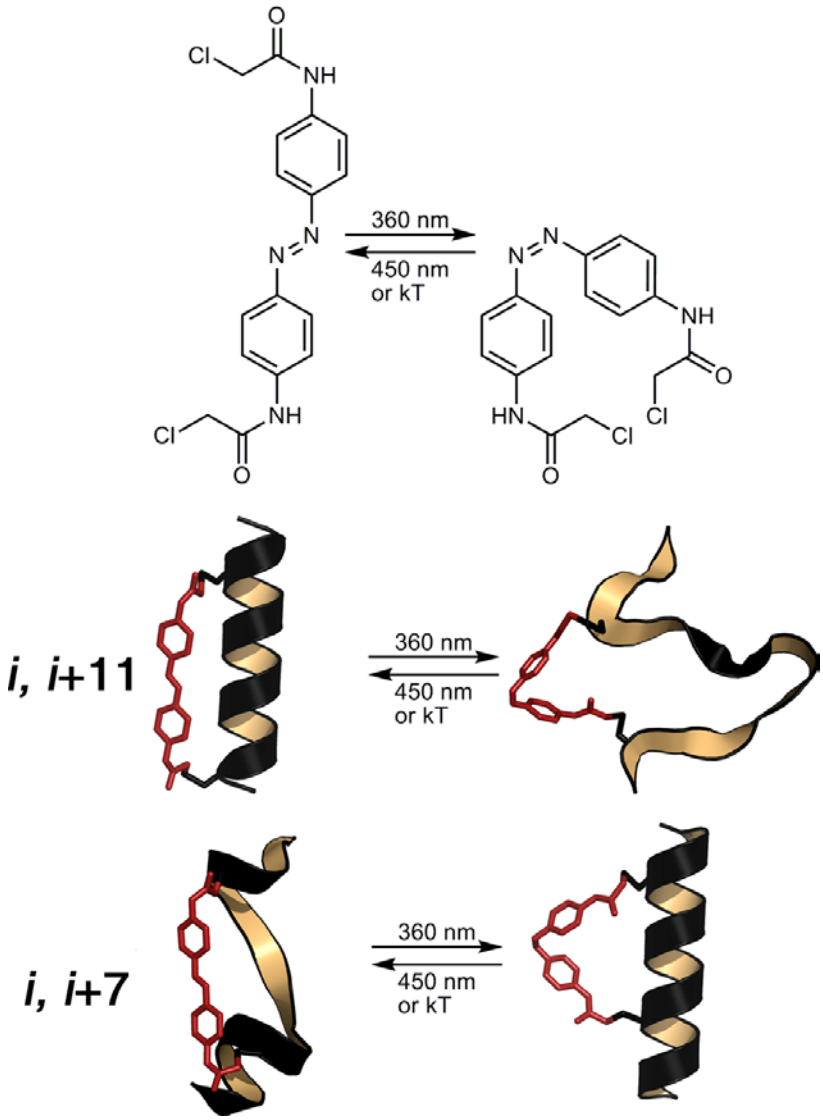

Figure 41: Photoisomerization of a cysteine reactive azobenzene crosslinker (top), peptide with cysteine reactive azobenzene crosslinker attached in *i*, *i*+11 (middle) and *i*, *i*+7 (bottom) configurations.

These polymeric, azobenzene-functionalised peptides were shown to undergo dramatic structural rearrangement upon irradiation.[155,156] The first reports of designs combining

azobenzenes with natural amino acids were based on cyclising peptides with azobenzene-derived amino acids.[157] Those constructs exhibited a transition from a disordered state to a β-turn when the azobenzene conformation switched from *trans* to *cis*.[157] A similar result is obtained when an azobenzene is coupled to two polypeptide chains that are prone to form a β-sheet.[158] The resulting functionalised peptide is a linear equivalent of the cyclic peptide and reacts analogously to irradiation. Cyclic azobenzene amino acid bearing peptides were recently applied to control muscle contraction.[159]

The next generation of photocontrolled peptides provided a range of sophisticated tools as they granted reversible control of the degree of helicity of a peptide sequence.[160,161] Across a number of studies a range of peptides have been shown to undergo conformational transitions upon irradiation. The difference in end-to-end distance of the *cis* compared to *trans* state of cysteine-reactive azobenzene derivatives allowed for a variety of designs where the degree of helicity can be up- or down-regulated depending on the arrangement of cysteines (Figure 41). In arrangements where end-to-end distance better matches the *trans* distance (e.g. *i*, *i*+11), stabilisation of an alpha helix in dark state is achieved. Conversely, spacing approximating distances in the *cis* configuration (e.g. *i*, *i*+7; *i*, *i*+4) yield a peptide a more helical peptide after exposure to UV light. This flexibility allows the design of a plethora of α-helical binding peptides using azobenzene to control their conformation and their affinity to relevant targets. Especially successful were BH3 region/domain derived peptides targeting Bcl-x_L and antiapoptotic members of Bcl-2 family.[162] Not only is their *in vitro* affinity before and after irradiation significantly different, but according to preliminary data, their activity *in vivo* can be controlled by light, so that apoptosis induction is achieved. (R.J. Mart, personal communication) An added benefit of structural reinforcement by azobenzene crosslinking is that the resulting peptides afford a degree of protection against proteolytic activity by their restricted conformations (unpublished data). The affinity of DNA-protein interactions both *in vitro* and *in vivo* can be controlled by azobenzene crosslinked helical peptides, as can collagen triple-helix formation.[163-166] Recent research expands the scope of this technique from peptides to full-size proteins; azobenzene photo-switching was shown to alter the structure of Fyn SH3 domain and subsequently to control enzymatic activity of a restriction enzyme.[167,168] Apart from such direct integration in protein structure, azobenzene crosslinkers have also been used to control the opening of ion channels by positioning an azobenzene derivative such that an agonist substituent can be pulled out or replaced in its binding site to modulate the channel activity.[169] This has been demonstrated for the nicotinic acetylcholine receptor,

the voltage-dependent K^+ channel Shaker and used to construct a light-gated glutamate receptor.[169-171] Further developments include the evolution of chromophore structures to modify spectral properties, alter thermal relaxation times and improve solubility.[161,172-174]

Similar sheet and helical control strategies were adopted using spiropyran derivatives.[175] Perhaps the most elegant example of this research involves modification of *E. coli* mechanosensitive channel (Mscl). This membrane protein opens up in response to a pressure change along the membrane, but it found that mutating members of a group of neutral residues inside the channel to charged ones also opens the protein channel.[176] Using the capability of spiropyran to shift from neutral to charged, an engineered, stable and reversible light responsive channel was created. Moreover, researchers were able to introduce membrane patches containing the photo-controlled channel to hollowed polymersomes and create artificial photo-controlled microcontainers.[176] Hemithioindigo derivatives have also used to control a channel by conjugation to a pore-forming bacterial peptide, gramicidin, which enabled the ionic current across the membrane to be modulated in response to light.[177]

The least "invasive" chromophore involves modification of the peptide bond itself. The carboxylic oxygen of the peptide bond is replaced with selenium or sulphur and these modified bonds can be isomerised from the usual *trans* configuration to *cis* by irradiation and the change in configuration of a single peptide bond can disrupt the secondary structure.[178,179] The biggest disadvantage of this strategy is the need to use very high-energy radiation to induce the isomerisation – 280 nm UV light is needed which make such engineered peptides unsuitable for *in vivo* applications and also some more sensitive *in vitro* were systems.[178,179]

1.4.3.4 Optogenetics

The ability of natural light-sensitive proteins to undergo sometimes a very dramatic shift in structure and their responsiveness to less energetic light makes them suitable for *in vivo* applications. Moreover, no chemical intervention is necessary as the appropriate genes can be transfected into cells. Channelrhodopsin is a particularly popular protein channel used extensively in neurological studies. Optogenetic techniques were applied to this protein to study the sleep/awakening transition.[180] Similarly, precise spatiotemporal control of neuronal activity allowed for identification of neurons crucial for the treatment Parkinsonism[181] and allowed an insight into learning mechanisms of mice.[182] Rhodopsins have been also reported as useful tools for control and analysis of other animal behaviour.[183] It is not only

neuroscience that benefits from optogenetic technology; a recent study employed engineered phytochromes to precisely control cellular trafficking.[184]

1.5 Programed Cell Death

1.5.1 Overview

Cell death may be due to an acute effect beyond the cells control or genetically encoded and directly regulated. This second option, programmed cell death, allows a multicellular organism to choose for some cells to die to benefit the whole organism. Various forms of programmed cell death are crucial for the homeostasis of a multicellular organism (although it has been observed in some unicellular species as well) allowing infected, oncogenic, damaged, detached or otherwise abnormal cells that could pose a threat to the whole organism or colony to be purged. The demise of some cells is a necessary developmental step or part of a tissue function.[185-188] The first documented examples of programmed cell death describe the process of cornification, where keratinocytes (live epithelial skin cells), undergo transformation into dead corneocytes that form a tight barrier around bodies of some vertebrates.[185] Once programmed cell death processes have been triggered, they can generally still be reversed until a certain threshold is crossed; but as yet no such points have been specified with a satisfactory level of accuracy and confidence. The current proposal of the Nomenclature Committee on Cell Death is to treat a cell as dead if its plasma membrane is ruptured or has lost its integrity or if the cell, including its nucleus, has been fragmented and/or engulfed by neighbouring cells. This cautious definition clearly describes cells long after the point of no return has been crossed, and even this practical definition fails to account for corneocytes (death epithelial cells) which do not meet either criteria. On the contrary, their lipid bilayer is heavily modified and reinforced to the point where it no longer resembles the structure it originated from.[189] Other more specific definitions not based on morphology have been brought forward, but in each case the proposed irreversible event has been refuted, these include: mitochondrial depolarization, massive activation of caspases, surface exposure of phosphatidylserine or loss of clonogenic survival (possibly to undergo mitosis).

1.5.2 Apoptosis

Although the programmed cell death itself was explicitly proposed in 1964, the term apoptosis was first coined six years later.[190,191] The process was not very well characterized

until the seminal work of Robert Horvitz describing the development of *Caenorhabditis elegans* and the role that apoptosis plays in it.[192] A series of gene knockouts revealed a complex system of proteins involved in controlling the fate of the 1031 cells of the hermaphroditic form of this nematode, of which 131 undergo apoptosis.[192] There is a remarkable resemblance between human apoptotic pathways and those in *C. elegans,* suggesting it is an old, evolutionarily conserved phenomenon among eukaryotes. Although it was originally thought that apoptosis-related genes were simply controlling tissue organism development, apoptosis is also a stress response phenomenon. The most prominent features of apoptosis include massive caspase activation, the involvement of mitochondria and mitochondrial proteins in signal transduction and a characteristic final morphology with the cell blebbing and forming of spherical apoptotic bodies. Unfortunately, there is little consensus in closely defining apoptosis, especially in the light of recent evidence of other processes with an apoptosis-like morphology, but without the use of canonical signal transduction pathways. NCCD recommends using the term apoptosis as it was originally conceived, to describe purely the morphological aspect ignoring that various pathways that may lead to the same final appearance.[189] Current opinion classifies apoptosis as belonging to the group of typical programmed cell deaths: programmed necrosis (necroptosis), autophagy (or rather autophagy related cell death), cornification and apoptosis rather than the atypical categories of mitotic catastrophe, anoikis, excitotoxicity, Wallerian degradation, paraptosis, pyroptosis, pyronecrosis and entosis.[189]

1.5.2.1 Morphology

By current definition all apoptotic cells feature the same characteristic morphology (Figure 42),[189] whose most prominent feature is massive cell membrane blebbing and formation of apoptotic bodies of various sizes.[191] The shape of those structures is either spherical or near-spherical and they contain apparently intact (not only morphologically but often also biochemically) cellular organelles or condensed fragments of chromatin derived from nuclear fragmentation (karyohexis); another hallmark of apoptosis.[191] In contrast to necrotic processes, the blebbing cell membrane remains intact although the overall cell volume is decreased (pyknosis) and it acquires a round shape. Pseudopods, if present, are retracted.[189,193] The presence of phosphatidylserine and other signalling molecules on the outer surface of the cell membrane and the release of nucleotides encourage the engulfing of apoptotic bodies by residing phagocytes. If the efficiency of this process is insufficient,

apoptotic bodies can acquire a necrotic morphology in a process termed secondary necrosis.[189,194]

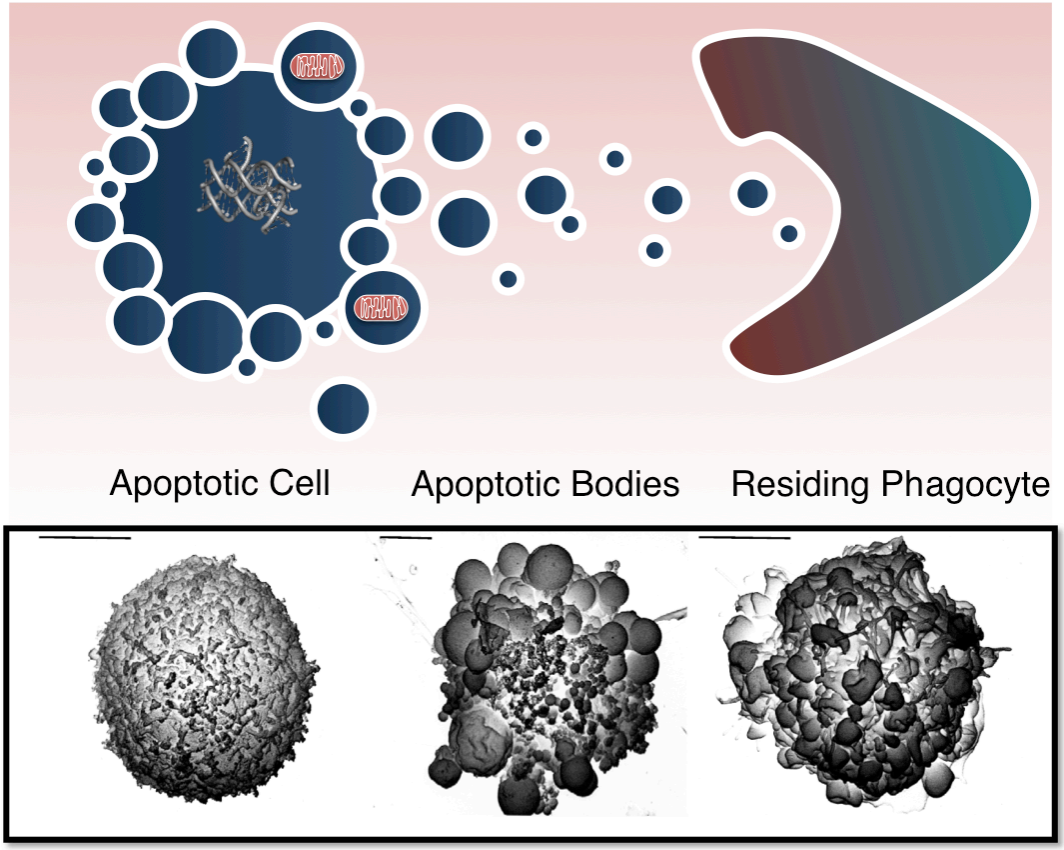

Figure 42: Top: Schematically depicted apoptotic morphology highlighting apoptotic blebbing and engulfment of apoptotic bodies by residing phagocytes. Apoptotic cells with (bottom from left to right) ROCK I protein disabled (no blebbing); Rho protein disabled; just apoptotic. Bottom: SEM picture adapted from reference.[195]

1.5.2.2 Extrinsic vs. intrinsic apoptosis

The apoptotic program operates by default in most cells and needs to be actively inhibited in order to maintain homeostasis.[196] Currently distinct intrinsic and extrinsic pathways are recognised, with different signal initiation pathways but converging at the effector level (massive caspase activation). The extrinsic apoptotic pathway is mediated *via* ligands and their cell membrane receptors, that belong to the TNF protein family and comprise the TNF-related apoptosis-inducing ligand receptors (TRAILR1/DR4 and TRAILR2/DR5), DR3 and DR6, CD95 (APO-1, Fas) and TNF receptor 1 (TNFR1) itself. Although their main function seems to be the reception and relaying of cell death signals, in some cases these proteins have been observed to stimulate necrosis or even cell survival.[193] A large group of ligands has been described so far, some of which are receptor specific and some bind to a few different targets. These include: Apo2L (TRAIL), TNF and FAS ligand (FasL, CD95L, and Apo1L). The TRAIL/TRAILR (both DR4 and DR5) ligand and receptor

pair exemplifies the signalling cascade of this pathway (Figure 43); DR TRAILR is a transmembrane protein containing a so-called death domain (DD) on the cytosolic site and a cysteine rich domain on the extracellular side. Receptors oligomerize prior to binding *via* interaction of Preligand Binding Assembly Domains (PLAD) of each subunit; such preassembly seems to be a general phenomenon for all death receptors. [197,198-200] TRAIL ligation then promotes assembly of even higher molecular weight oligomers, which in turn drives interaction with Fas Associated Death Domain (FADD) or TNF Receptor Associated Death Domain (TRADD) proteins through death domains (DD).[201,202] Association of one of those two factors promotes recruitment of initiatory caspases 8 and 10 through death effector domains (DED) and leads to their autoactivation. This response is moderated in some contexts by recruitment of the caspase mimicking, but catalytically inactive FLICE inhibitory protein (c-FLIP).[203,204] The cytosolic part of the multiprotein structure is referred to as the death inducing signalling complex (DISC) and the caspases it activates propagate the proteolytic signal to caspases 3, 6 and 7.

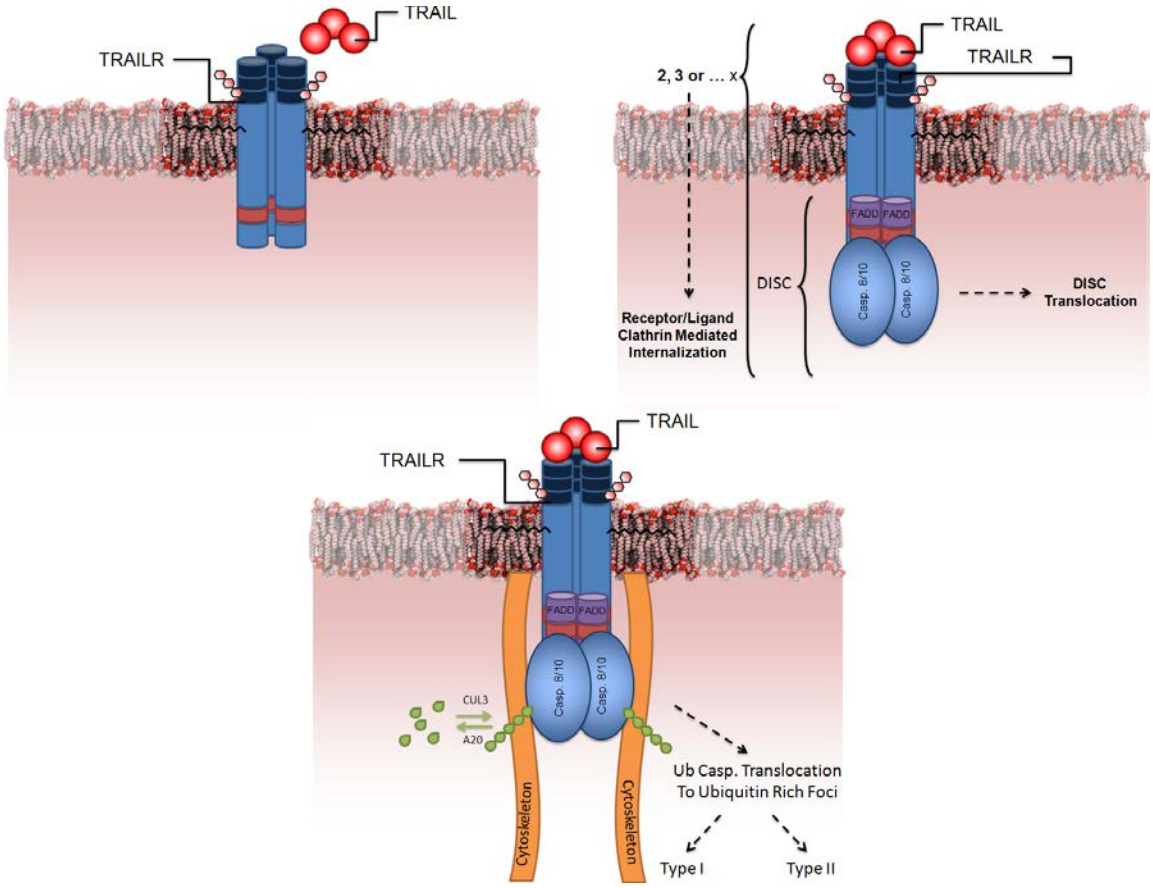

Figure 43: Model of TRAILR/TRAIL interaction and subsequent activation of the extrinsic apoptotic pathway.

The activation of effector caspases through DISC depends on the cell type. In type I cells there is sufficient caspase 8 or 10 activity present to propagate the proteolytic death signal further, but in type II cells lower levels necessitate stimulation of intrinsic pathway.[205] Current findings hint that in order to fully activate initiatory and effector caspases, it is necessary to polyubiquitinate bound caspases, which promotes translocation of the entire DISC to ubiquitin rich foci and increases their effective concentration.[206] Both type I and type II cells process the protein Bid to form truncated Bid (tBid), which activates the intrinsic apoptotic pathway.[205]

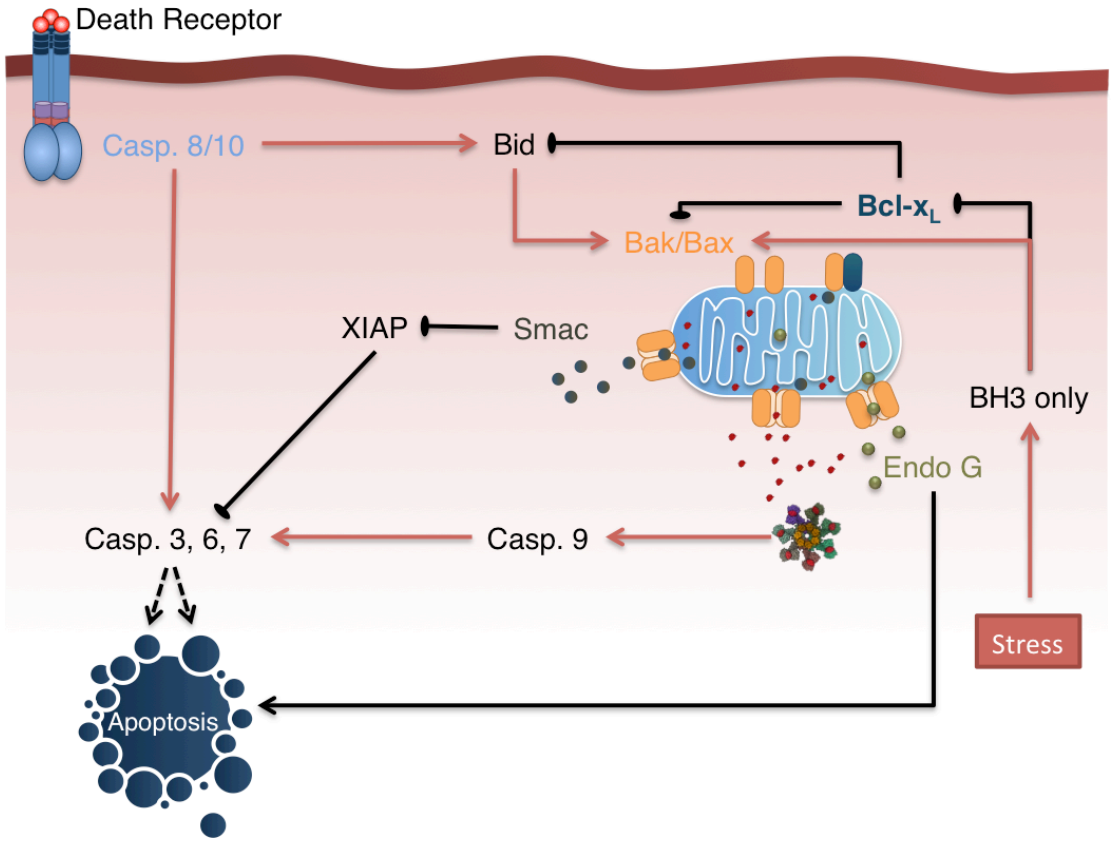

Figure 44: Intrinsic and extrinsic pathway crosstalk and convergence at final apoptosis execution itself (black arrow indicates inhibition and red arrows activation).

In contrast to the extrinsic apoptotic pathway, the intrinsic one occurs in response to intracellular stressors including UV light, ionizing radiation, infection, ER stress, heat shock and oxidative stress (Figure 44).[207] These stimuli are directly or indirectly sensed by so called BH3-only proteins that in turn activate proapoptotic multi-region Bcl-2 family members (Bak and Bax). Upon activation Bak/Bax homooligomerize and cause mitochondrial outer membrane permabilization, probably by forming pores whereby the soluble contents of mitochondria are released to cytosol.[208] Among them are cytochrome c, second mitochondria-

derived activator of caspases/direct IAP binding protein with low pI (SMAC/Diablo), apoptosis inducing factor (AIF) and endonuclease G (EndoG). SMAC/Diablo inhibits X-linked inhibitor of apoptosis (XIAP), which removes its inhibition of downstream caspases. Cytochrome c assembles into a complex with apoptotic protease-activating factor 1 (Apaf-1, Figure 45) known as the apoptosome which, after a large conformational change oligomerises into a 7-fold symmetrical multiprotein complex.

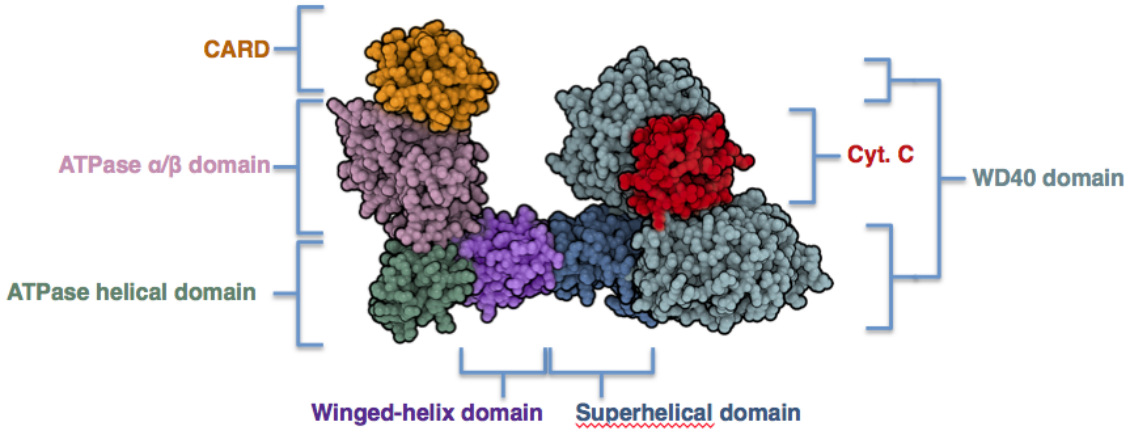

Figure 45: Monomer of an apoptosome; domains of Apaf-1 with cytochrome c bound are indicated. The CARD domain (yellow) is responsible for the homotypic interaction with CARD on caspase 9. The WD40 domain (grey) responsible for the binding of cytochrome c (red). The ATPase domains (pink and green) are responsible in this model for oligomer formation. Also illustrated are the winged-helix and superhelical domains (purple and blue). Apaf-1 structure adapted from reference 104.

Caspase 9 is the primary initiatory caspase of the intrinsic pathway and is activated by dimerization facilitated by the apoptosome (Figure 46). Two hypotheses to explain this have been brought forward so far; in the first the small amount of caspase 9 dimer already present in solution interacts with the CARD domains of the apoptosome *via* a homotypic interaction, which triggers formation of a stable dimer possibly *via* kinetic trapping of the activated dimer conformation. In the second model each subunit interacts with apoptosome separately and after such activation they assemble in solution. Once there is sufficient active caspase 9 in the cell the signal can propagate *via* proteolytic activation of effector caspases.

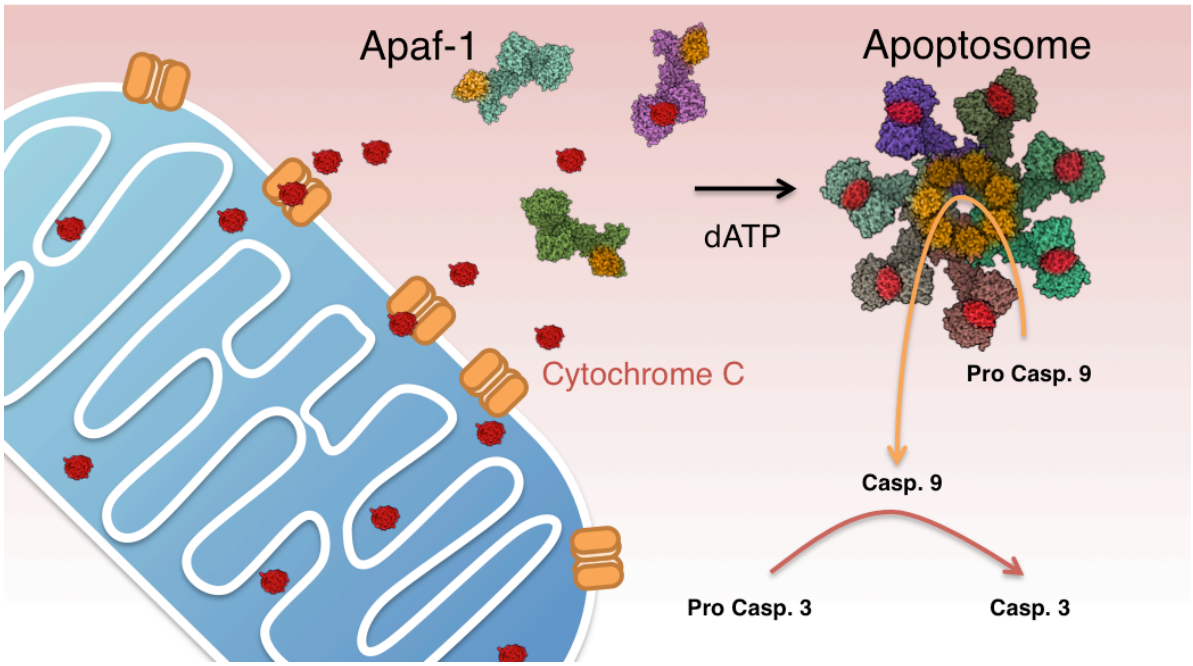

Figure 46: Apoptosome assembly triggered by cytochrome c release from the mitochondria caused by the other mitochondrial membrane permabilization; assembled apoptosome dimerizes caspase 8, which in turn triggers the proteolytic cascade of executor caspases. Structure of Apaf-1 and apoptosome adapted from.[209]

1.5.2.3 Caspase overview

Caspases are intracellular cysteine-dependent aspartate-specific proteases that carry out several distinct cellular processes.[210] They are a conserved and old eukaryotic family of proteins and share a common ancestor protein.[211] Although caspases are principly known for their crucial role in the propagation of apoptosis, they are also involved in proliferation and inflammation.[212] Caspases are broadly classified into two groups based on what their role is believed to be. Caspases 2, 3, 6, 7, 8, 9 and 10 are apoptotic and 1, 4, 5, 11, 12, 13 are most often related to inflammation.[212] Caspase 14 is thought to be solely engaged in keratinocyte maturation.[213] However, it is worth noting that some members of each group are implicated in multiple processes.[214] Classification within the group of caspases related to apoptosis distinguishes initiator (apical) caspases from effector (executioner, downstream) caspases. This grouping overlaps with mechanistic aspects of activation and structural features as the propeptide domain of initiator caspases with activity triggered by dimerization, rather then proteolytic cleavage as for downstream caspases whose zymogen is already dimeric (Figure 47). All caspases posses a catalytic dyad of a cysteine and a histidine which are used to cleave peptide bonds with specificity strongly biased towards substrates containing aspartic acid, with proteolysis commonly occurring after this residue.[210]

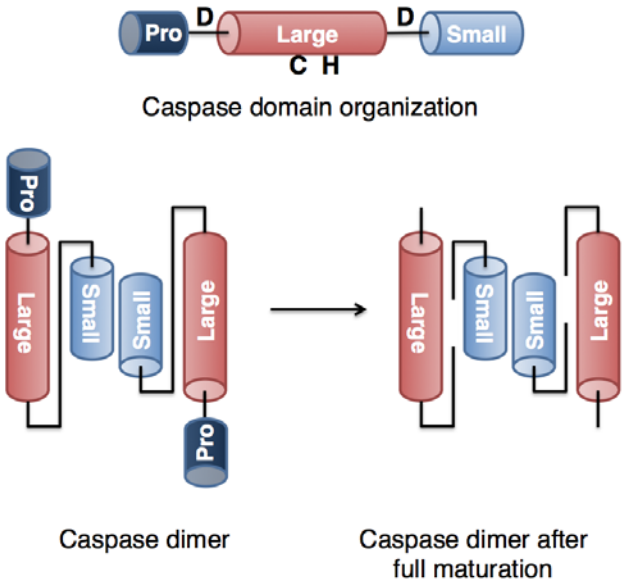

Figure 47: Caspases structural organization; three distinc domain indicated (propeptide, large domain and small domain) with residues that undergo cleavage (D) and the catalytic dyad (C and H).

Most caspases are able to cleave a wide range of proteins; nevertheless a certain tendency to target specific multiprotein complexes and proteins involved in specific pathways has been reported.[215] It is caspase activity within a cell which confers the characteristic apoptotic morphology;[216] apoptotic blebbing has been identified as being a consequence of the activation of Rho effector protein (ROCK I) by proteolytic cleavage.[195] ROCK I activation also causes increased condensed chromatin translocation. Cleavage of lamins by caspases contributes to the weakening of the nuclear envelope and fragmentation of DNA is achieved *via* caspase activated DNases that lead to the characteristic DNA ladder seen by electrophoresis techniques. Cellular shrinkage is attributed to the proteolysis of various cytoskeleton proteins.[217] The activity of caspases can be controlled by three different regulatory methods: the turnover of activated caspases is very high as they are rapidly directed to the proteosome,[214] "decoy" proteins like c-FLIP can compete with caspases for the adaptors (e.g. at DISC), and numerous inhibitors of apoptosis (IAP) directly or indirectly affect the activity of caspases.

1.5.2.4 Bcl-2 family

The main controllers of the intrinsic pathway of apoptosis are the B cell lymphoma 2 family of proteins (Bcl-2 family). This functionally and structurally diverse family of proteins is characterized by presence of one or more so called Bcl-2 homology regions (BH regions).[218] The family was named after the prototypical member whose gene was discovered

to be interrupted in ~60% cases of human follicular lymphomas; as the research was conducted on B cell lymphomas the gene was nicknamed *bcl-2*, which was later, according to the convention, transferred to its product the Bcl-2 protein.[219] The family can be divided into three distinct subgroups: multi-region antiapoptotic proteins possess BH1-4 regions; multi region proapoptotic proteins possess BH1-3 regions (according to more recent literature BH4 regions also present in sequences of multi region proapoptotic proteins), and proapoptotic BH3-only proteins (Table 2).[208] Although BH regions are persistently called "domains" in the literature, this is incorrect as domains are capable of folding independently and performing their function, which is not the case with most of BH regions.

Antiapoptotic	
Bcl-2	BH1 BH2 BH3 BH4 TM
Bcl-x_L	BH1 BH2 BH3 BH4 TM
Bcl-w	BH1 BH2 BH3 BH4 TM
Bcl-2A1	BH1 BH2 BH3 TM
Mcl1	BH1 BH2 BH3 BH4 TM
Bcl-b	BH1 BH2
Proapoptotic multi-region	
Bax	BH1 BH2 BH3 BH4 TM
Bak	BH1 BH2 BH3 BH4 TM
Bok	BH1 BH2 BH3 BH4 TM
Proapoptotic BH3-only	
Bid	BH3
Bik	BH3 TM
Bim	BH3 TM
Puma	BH3
Noxa	BH3
Bad	BH3
Hrk	BH3 TM
BNip1	BH3
BNip2	BH3
BNip3	BH3

Table 2: The overview of the Bcl-2 family members; gray colour indicates uncertainty in classification.

In addition to BH regions, the antiapoptotic group and multi-region proapoptotic group (Bak and Bax) share C-terminal transmembrane region/domain (TM) that allows them to locate to the surface of the outer mitochondrial membrane.[218] Bik is the sole BH3-only protein shown to possess a TM region.[220] Five antiapoptotic members of the Bcl-2 family are known, those are Bcl-2, Bcl-x_L, Bcl-2A1, Mcl1 and Bcl-b. As all five of them possess a TM region, they are usually found localized in the other mitochondrial membrane, but also on the endoplasmic reticulum; this subcellular localization is not surprising given the amount of continuity between these membranes.[218] Although initially Mcl-1 was believed to lack a BH4 region,

careful signature analysis allows a BH4 region/domain to be identified; likewise Bak and Bax.[221]

The two known proapoptotic multi-region proteins, Bak and Bax, are responsible for the permabilization of the outer mitochondrial membrane in a process that almost certainly involves a formation of protein pores (mitochondrial apoptotic channels, MAC), although neither the existence nor the nature of those pores has yet been established. Two models have been proposed involving either proteinaceous or lipidic pores, with current data mostly supporting up the latter hypothesis.[222-225] Another multi-region proapoptotic member of the Bcl-2 family Bcl-2 ovarian killer/Matador (Bok/Mtd) has yet to receive the attention given to Bak and Bax, but is suggested to act in a similar fashion. Its expression seems to be very tissue specific and it has been implicated in human placental development.[226-229]

Two distinct models and a compromise between them have been proposed to try to explain the interplay between Bcl-2 family members (Figure 48). The direct activation model subcategorizes BH3-only proteins into two groups; sensitizers (Bad, Bmf, Bik, Hrk, Noxa, Beclin-1) and activators (Bim, tBid, Puma). The role of activators is to trigger a structural change in Bax/Bak to ready them for oligomerization and pore formation. Sensitizers are proposed to liberate activators from unproductive complexes with antiapoptotic members of Bcl-2 family. In the displacement model, Bak and Bax are either already active, or the act of release from an inhibitory complex itself triggers activation. In the latter case the sole role of BH3-only proteins would be to displace Bax/Bak from complexes with antiapoptotic proteins. The third, ‘embedded together’, model suggests that both of those scenarios occur and both activation and release from antiapoptotic members of Bcl-2 play a role in the intrinsic pathway. Embedding the proteins together in the outer mitochondrial membrane alters the local concentrations and dissociation constants of the different complexes to introduce subtleties that might not be reflected by *in vitro* measurements. Regardless of which model best describes reality, and the process could differ according to tissue or environment, it is certain that BH3-only proteins activate apoptosis. To do so they need to sense an apoptotic signal, respond to it and relay it to other members of Bcl-2 family. Human BH3-only proteins have evolved a whole variety of solutions to this problem. For example, Bad is dephosporylated if the level of growth factors drops below a certain threshold.

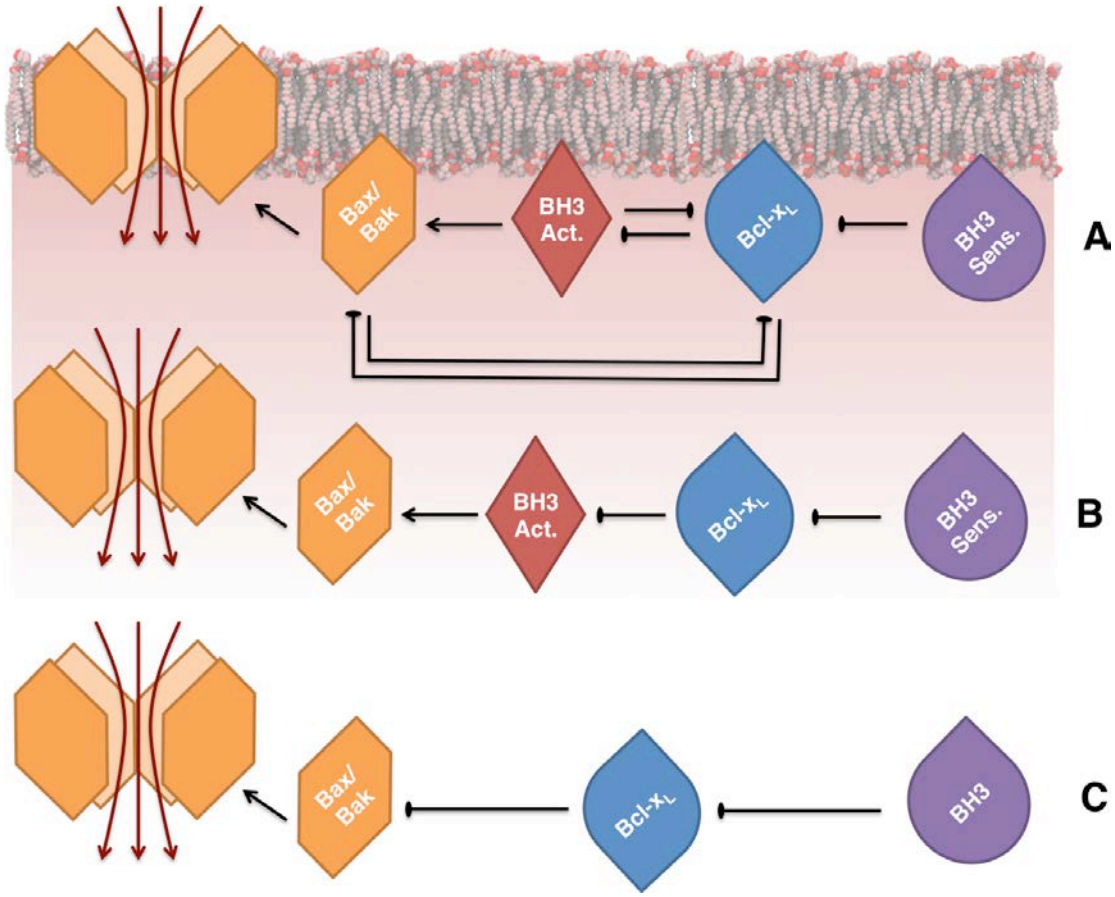

Figure 48: Three models of Bcl-2 proteins interplay. (A) Embedded together model; (B) direct activation model; (C) displacement model. "Bcl-x_L" stands for any antiapoptotic member of Bcl-2 family; "Act." stands for activator and "Sens." stands for sensitizer.

The lack of this strategically positioned phosphate group allows Bad to interact with Bcl-2 and Bcl-x_L, releasing proapoptotic proteins from complexes with these proteins.[230] Phosphorylated Bak not only has lower affinity to Bcl-x_L and Bcl-2, but also is sequestered from mitochondria by 14-3-3 proteins.[231,232] Bmf and Bim act as sensors of UV damage and cell detachment by being released from their resting complexes (myosin V motor chains for Bmf and microtubules for Bim) to displace proapoptotic proteins from complexes with Bcl-x_L, Bcl-w and Bcl-2.[233] Additionally, Bad and Bim synthesis is upregulated and their dephosphorylation accelerated when a cell is deprived of survival factors.[234] Puma, Noxa and Bik are not constitutively produced, but their translation is upregulated by p53 protein if certain apoptotic stimuli are present (DNA damage, oxidative stress).[220,235,236]

As noted previously, the intrinsic and extrinsic apoptotic pathways cross-talk at the level of Bid activation (Figure 22). Bid treated with caspase 8, caspase 3, calpain or granzyme B is truncated at the C-terminus to tBid which is postulated to insert into the outer mitochondrial membrane and induce Bax to permeabilize the membrane.[237] Although transferring apoptotic information from death receptors to Bax/Bak seems to dominate, there

are some cases where Bax activation through membrane embedded death receptors is not mediated by tBid. In addition, p53 itself is known to bind Bax and to induce activation of the intrinsic pathway. It seems that Bak responds to a variety of stimuli to surmount the conformational change activation barrier and induce pore formation; physical factors like temperature and acidity have been reported to be among them.[238] Apart from their roles as regulators of apoptosis, Bcl-2 family members have also been reported as engaged in various other cellular pathways; Bcl-xl and Bcl-2 are known to modulate mitochondrial fission and Bad is involved in the regulation of glycolysis.[230,239] In addition to the Bcl-2 family members already discussed, yet others have been identified (e.g. Bcl-g, Bcl-Rambo, Bcl-b) and probably many remain to be found.[238] Some of the additional members are known to be involved in apoptosis, but too little data is available for any rigorous classification.

1.5.2.5 Structural Biology of Bcl-2 family members

Unfortunately, structures of Bcl-2 family members are often elusive; of all the BH3-only proteins only that of Bid are known.[240] Experiments suggest that the other BH3-only proteins are natively unfolded and obtain more pronounced tertiary structure upon binding to other proteins, a quality that perhaps reflects the multitude of environments and adaptors they need to interact with in order to sense and relay apoptotic signals[241] (similar behaviour is observed for p53).[242] Analysis of the published structures allows a fairly rigorous classification of structural elements of the Bcl-2 family members, defining helices (numbered from α1 to α8), loops (some of which serve functions other than just connecting two helices), regions/domains (BH and TM) and binding grooves (Figure 49 and 50).

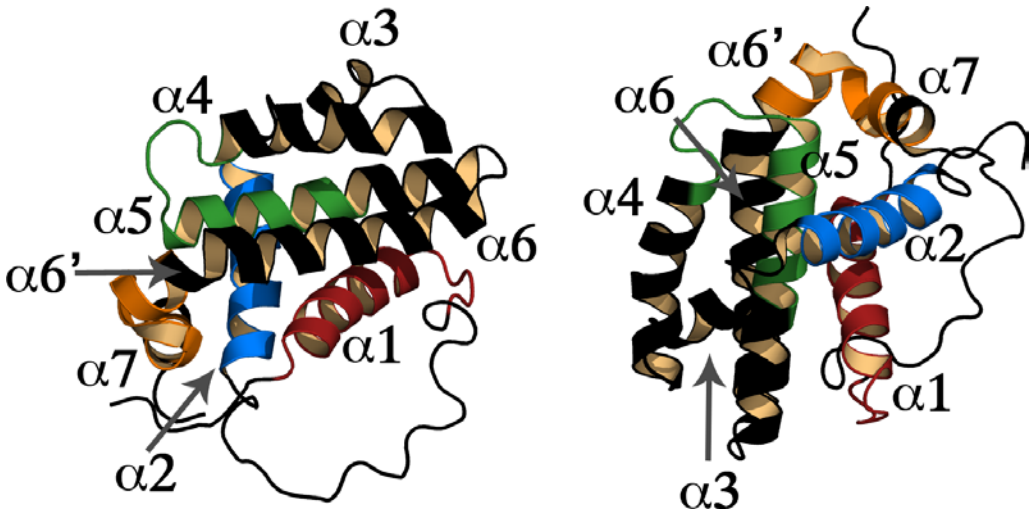

Figure 49: Bcl-2 family members helix annotation (based on Bcl-2). Please notice that last helix (α8) is not always present and is not represented on this structure. BH1 – green, BH2 – orange, BH3 – blue, BH4 – red. PDB entry: 1G5M.

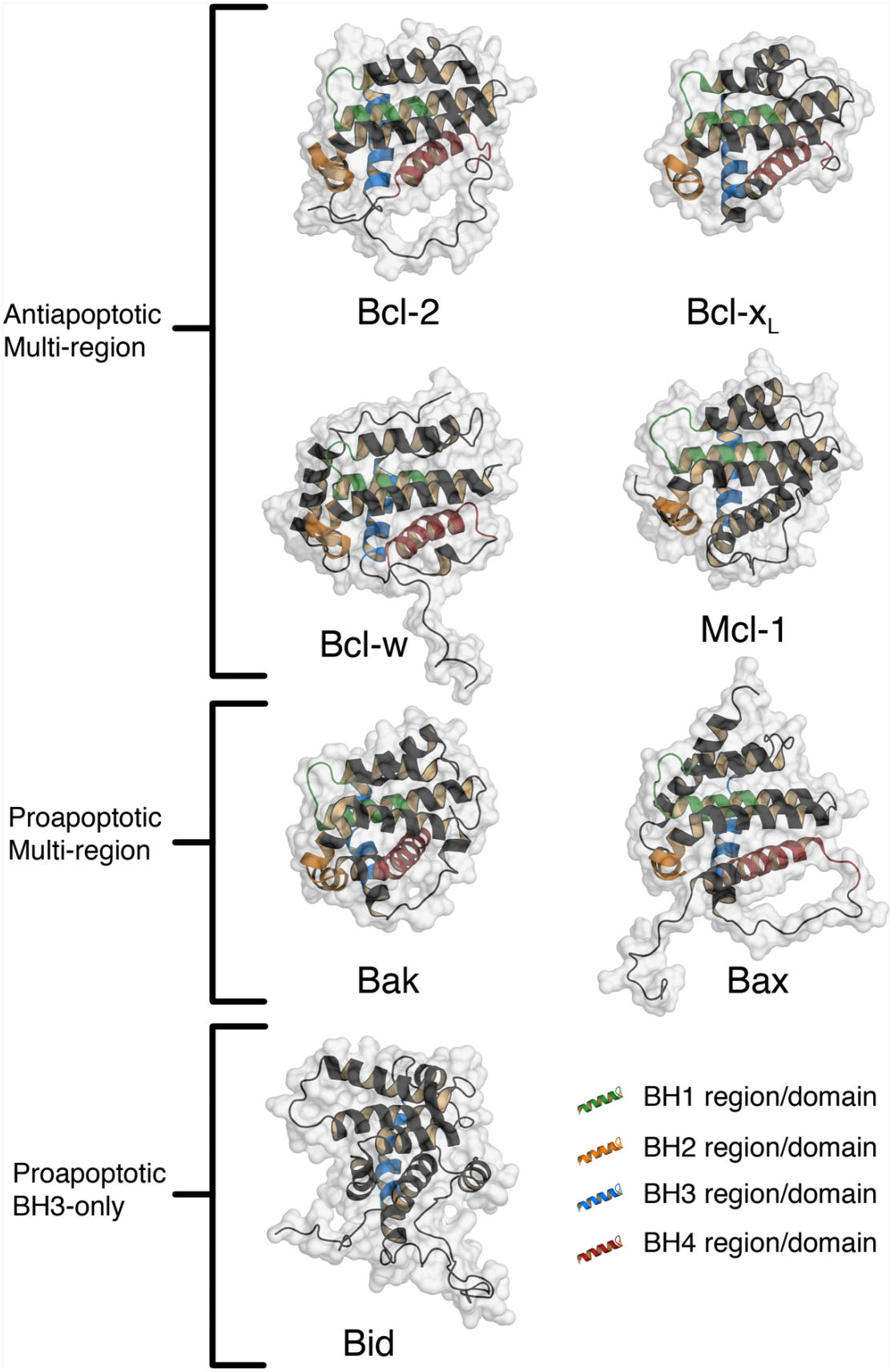

Figure 50: Structural comparison of Bcl-2 family members. PDB entries: 1WSX, 1R2D, 1O0L, 1G5M, 2JCN, 2K7W, 2BID.

Although the members of the family often show low sequence identity, the overall structure is conserved, for example Bcl-x_L and Bid share only 11.2% sequence identity but have an overall RMSD of 3.8 Å (Figure 50).[243] Known structures show all α–helical, globular proteins, connected predominantly by short loops with a fold organized around a central α5 helix that forms the hydrophobic core and is surrounded by amphipathic alpha helices. Although it is difficult to analyse the structure-sequence relationship between proteins that share so little sequence identity yet so much in architecture, a set of key residues that define the Bcl-2 family topology/fold has been identified. Data mining revealed 72 structurally equivalent residues 22, of which are at the same time of similar size make the same inter-residue contacts and share solvent accessibility with the Bcl-2 family helix bundle fold.[243] By extending this technique and considering the 72 positions only for viral Bcl-2 homologues sequences with known structures it was possible to successfully identify new viral homologues of antiapoptotic proteins.[243]

Unfortunately, the published literature lacks a coherent helix numbering scheme, α6' is sometimes referred to as α7, which accordingly changes numbering of the following helices (Figure 49). The notation used in this thesis is described in Figure 49. In some structures, α8 (Bax, Bcl-w) is also visible. This helix is often found in the canonical BH3 peptide binding grove formed by α3 and α4[244-246] unless displaced by a BH3 peptide. It is possible that the lack of a resolved α8 or its displacement from the BH3 binding grove in the published structures of Mcl-1, Bcl-x_L and Bcl-2 is an artifact of protein engineering, as all of those proteins were truncated at the C-terminus in order to improve their properties for NMR sample production. However, some structures of proteins in complexes with BH3 peptides show a "displaced" helix at full length C-terminal helix.[247-251] It is hypothesised that α8 is displaced upon binding of a BH3 domain and inserted into structure of its binding partner to compensate for its partial unfolding or that exposure of the loop between α7 and α8 facilitates membrane insertion.

Although most loops present in the structures of Bcl-2 family proteins are short and well ordered, there are two exceptions. The loop connecting α3 and α4 is flexible and allows for the significant reshaping of the binding grove (helices α3 and α4).[247,248] The disorder of this loop in the free protein is reflected in a lack of long range NOEs and the disappearance of amide N-H NMR peaks in a manner characteristic of fast exchange with water (unpublished data), but once a ligand is bound the loop seems to be very ordered. The other flexible loop

has a variable length and connects α1 and α2 (also BH4 and BH3). In all cases except Bcl-w it is found to be highly disordered and while its function has yet to be determined, it connects two BH region/domains and possibly works (at least in case of proapoptotic Bcl-2 family members) as a flexible tether allowing BH3 region/domain to reach its target. What is certain is that in the case of antiapoptotic proteins deletion of this loop, a common practice in structural studies, does not affect the binding of BH3 region/domain extracted peptides or *in vivo* activity.[247-249,252] Some studies suggest that phosphorylation within this loop abolishes Bcl-2/p53 interaction.[253]

A striking characteristic of Bcl-x_L is its ability to form domain swapped dimers (Figure 51). Three different arrangements are known, one of which has been suggested to be solely a crystallization artifact, due to a deliberately shortened loop connecting helix α1 with α2.[45,254,255] However, the binding of other proteins to the loop may effectively shorten that the average distance between those helices, leading to a natural situation similar to that seen in the crystal structure of the mutated protein. Proteins embedded in the outer mitochondrial membrane exist in an environment that is still very difficult to research or model, and could possibly engage aforementioned loop. In the second domain swap dimer an even more significant reorganization is observed to be triggered by temperature. The BH2 region/domain is swapped between two monomers and a long helix forms between BH1 and BH2.[256] Similar domain swap dimers have been reported for Bcl-w, but different helices are swapped.[257]

The fold of proapoptotic Bcl-2 family members requires a major reshaping to expose their BH3 domains (*vide* Bak/Bcl-x_L) as the amphipathic BH3 helix is packed centrally and some residues face into the hydrophobic core. This suggests that the protein itself must have a surmountable barrier to unfolding under certain conditions to adapt its shape for the binding purposes or to allow membrane insertion. This implies that at least 3 (likely more) very different conformations of Bak exist, and possibly even more if the previously discussed, often-removed α8 of Bcl-xL, Mcl-1 and Bcl-2 replaces partially unfolded Bak, Bax or Bid. There are no published structures of the antiapoptotic Bcl-2 family members in complex with full-size BH3-only proteins or Bak/Bax, despite a multitude of X-ray and NMR structures of complexes of BH3 domain derived peptides of various lengths.[258]

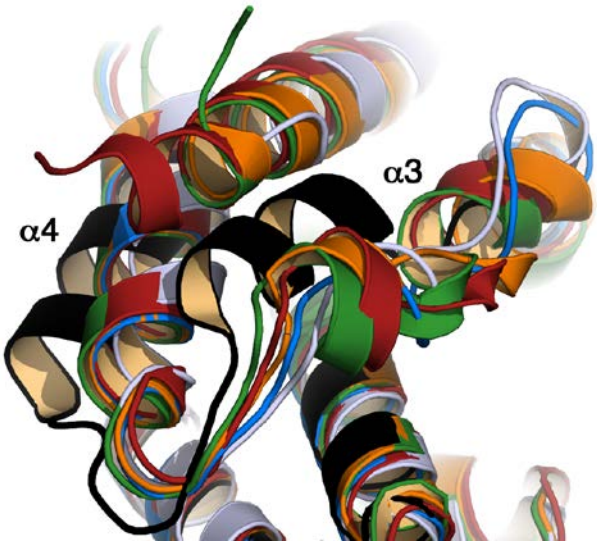

Figure 51: Overlaid structures of Bcl-x_L (black) and Bcl-x_L in complex with Bim (orange), Bak (light-purple), Bad (red), Bax (green) and Beclin-1 (blue) peptides derived from the respective BH3 region/domains. Helices α3 and α4 are indicated. PDB entries: 1R2D, 3PLX, 2P1L, 1BXL, 2BZW and 3PL7.

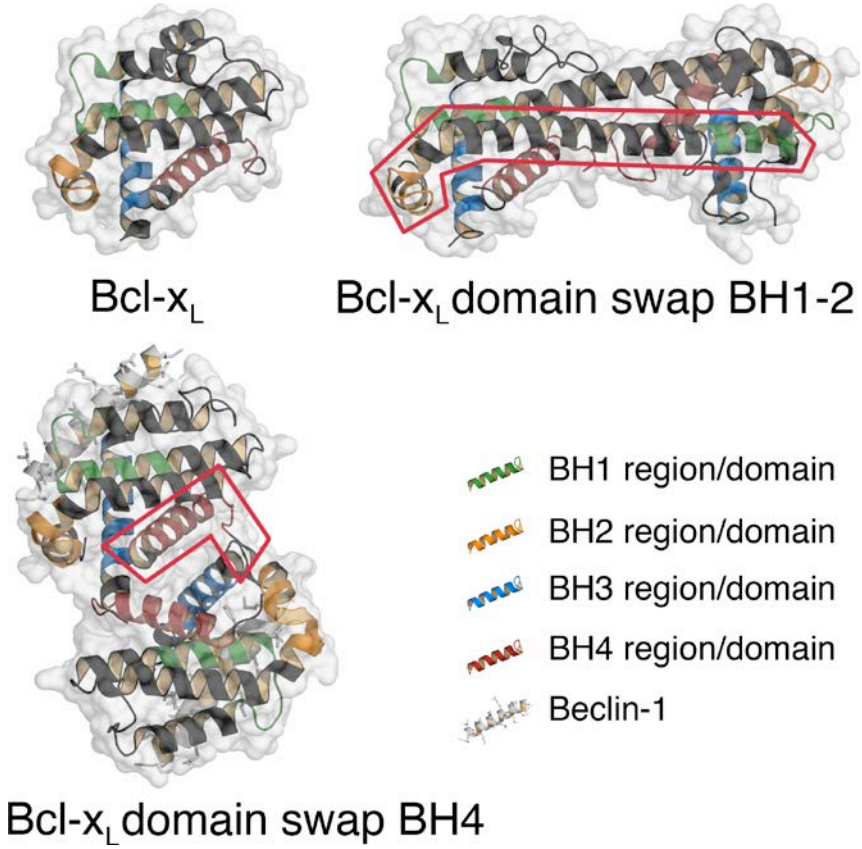

Figure 52: Comparison of different Bcl-x_L forms; Bcl-x_L monomer, Bcl-xL domain swap dimer (BH1-2) and Bcl-x_L domain swap dimer (BH4). PDB entries: 2P1L, 1R2D, 2B48.

As structures of new small molecule inhibitors of antiapoptotic members of the Bcl-2 family enter the populate Protein Data Bank they reveal a surprising degree of flexibility in

the canonical BH3 binding grove (between helices α3 and α4). Helix α4 moves towards the peptide or inhibitor with a concomitant shift of the helix α3 that widens the binding grove to adapt to the ligand and reshapes the connecting loop (Figure 52 and 53).

Upon closer examination of Bcl-x_L structures in complex with its four different ligands, we can see that not all of them perfectly overlay with each other as judged by the carbon alpha position (Figure 52 and 53). The most significant deviations are seen in the Bcl-x_L in complex with Beclin-1 and Bak (Figure 52). In the first case, it perhaps reflects this peptide's dissimilarity as compared to others. In the second case, quality of the NMR structure is not sufficient to determine if the change is genuine or is just an artifact of the structure solving algorithm. Pronounced differences can be found in side chain positions even when carbon alpha positions are very similar. This is a direct consequence of differing sidechains.

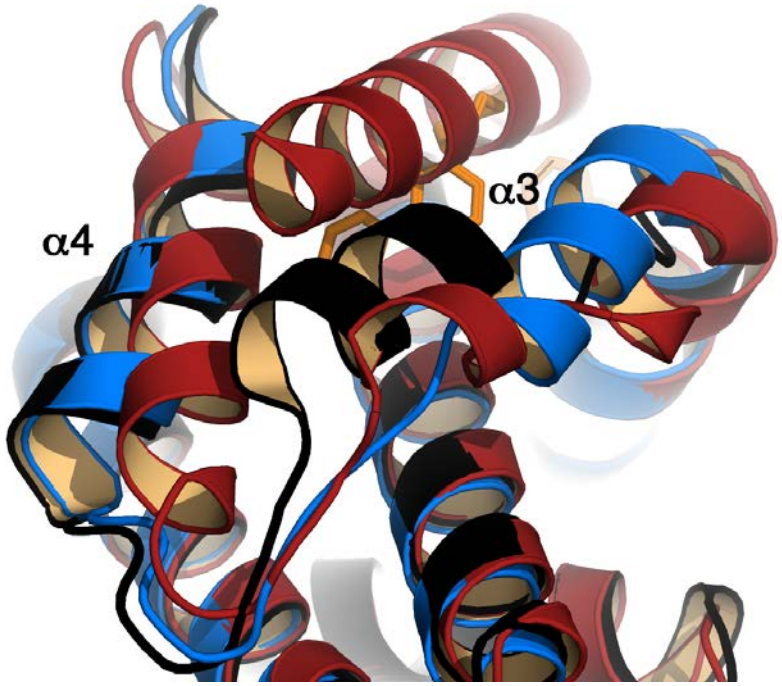

Figure 53: Overlaid structures of Bcl-x_L (black) and Bcl-x_L in complex with Bax (red) and a small molecule inhibitor (blue). The inhibitor is visible in the middle of the picture (orange). Helices α3 and α4 are indicated PDB entries: 1R2D, 3PLX and 3INQ.

Indeed, the adaptability of the BH3 binding is well illustrated by a structure of Bcl-x_L in complex with a small molecule ligand (Figure 53).[15] The small molecule used in this study bound well to Bcl-x_L, yet induced a conformational change that was not as pronounced as that observed upon BH3 peptide binding.[15] This structure lies somewhere between free Bcl-x_L structure and that of Bcl-x_L complexed to a BH3 peptide and sheds some light on how the transition in conformation upon binding might take place. It seems that first helix α3 widens

the grove when the C-terminal part of the peptide binds and then helix α4 follows to engage the N-terminal part (Figure 52 and 53).[15]

Both structural biology studies and alanine scanning have been used to reveal the residues that contribute most to the binding energy of the antiapoptotic protein Bcl-x_L to a Bak BH3 peptide (Figure 54). The most important residues on the Bak side are valine 74, arginine 76, leucine 78, isoleucine 81 and 85, aspartate 83. For Bcl-x_L, phenylalanine 97 and 105, tyrosine 101 and 195, valine 126 and 141, glutamate 129 and arginine 139 are most important.[247,258] The binding is driven predominantly by the hydrophobic effect, when leucine 78 of the Bak peptide was mutated to alanine binding was practically abolished.[247,258] However there are also two important salt bridges between arginine 76 and aspartate 83 on the peptide to glutamate 129 and arginine 139 of the protein.[247] The existence of those two salt bridges is somewhat controversial. The analysis of the only published structure of Bcl-x_L in complex with Bak (1BXL) reveals that the distance between those residues might not be sufficient to report an electrostatic interaction (*vide infra*). These patterns differ noticeably for other peptides binding to this grove on Bcl-x_L as well as between other antiapoptotic members of the Bcl-2 family.

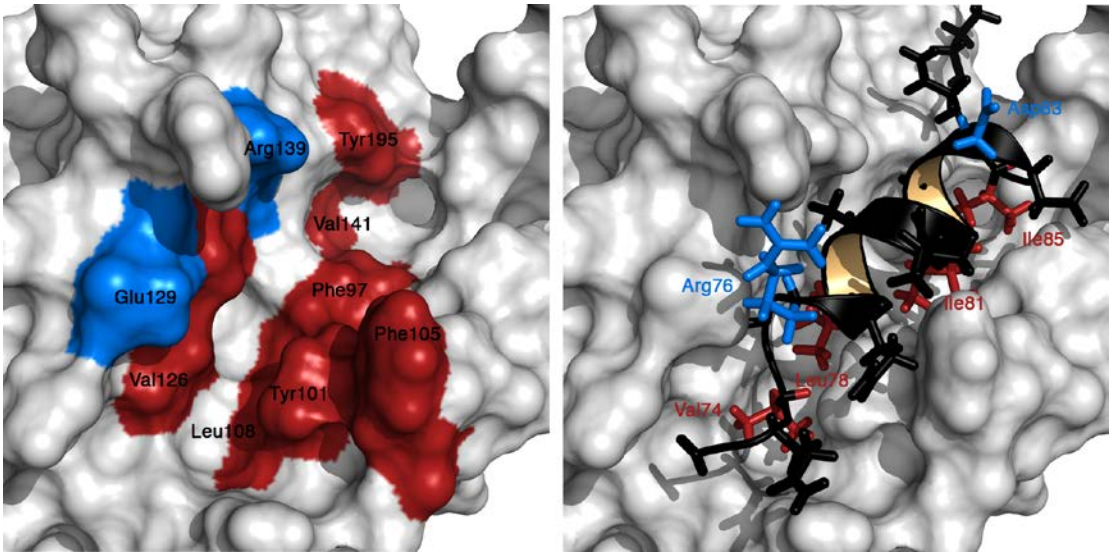

Figure 54: Surface representation of the Bcl-x_L in complex with Bak with residues contributing most to the binding energy on Bcl-x_L (left, Bak not visible); same surface representation with Bak helix and amino acid side chains of Bak contributing most to the binding energy (right).

One of the remarkable observations in studies detailing the interactions between Bcl-2 proteins and BH3 peptides is that one of the best binding Mcl-1 peptides is derived from the BH3 domain of Mcl-1 itself. This result is echoed by the finding that Bcl-w forms a self-inhibiting C-terminal α–helix, a feature that might be universal amongst Bcl-2 proteins.[244,245] This raises a very interesting question about binding equilibria of the whole population of

those proteins, especially when artificial peptides derived from BH3 regions are added to the system, further increasing the overall complexity.[59] This situation although not naturally occurring is of the highest interest due to possible medical applications of engineered BH3 peptides. It is difficult to predict the effect of BH3 peptides on the equilibrium of Bcl-2 family of proteins.

It seems that this group of proteins evolved around the same topology, the same fold and to contain similar regions (BH3), so that parts of their structure can be exchanged (although perhaps only transiently) between themselves either to activate/deactivate a protein or to allow for oligomer formation (MAC formation). One can further theorize that formation of oligomers may be one of the ways to deactivate selected Bcl-2 family members and clear them *via* relevant pathways. This model of virtually exchangeable protein units (domains/regions) shines a new light on how proteins in general may interact and it is a way more similar to some nucleic acid structures rather than canonical protein-protein interactions.

2 Results and Discussion

2.1 Photocontrol of Rev/RRE interaction

2.1.1 Aims and Objectives

The aim of this study was to synthesise a photoswitchable RNA binding peptide and evaluate its absolute and relative (irradiated vs. dark-adapted) affinities to a target RNA sequence. Since peptides derived from Rev can act as antagonists of the Rev-RRE interaction and serve as models for the development of methodology for the photocontrol of peptide-RNA interactions, the photoswitchable peptides were based on the sequence of the minimal Rev peptide retaining high affinity to RRE.[128-130]

2.1.2 Biosynthesized Rev Peptides

2.1.2.1 Design Cloning, Expression and Crosslinking of the Rev peptides

The initial peptide design was performed by Dr Sabine Kneissl. The Rev wild type peptide sequence was modified in order to accommodate the crosslinker (Table 3). Two appropriately spaced residues on the solvent exposed face of the Rev α-helix and opposite to the residues involved in RNA binding had to be mutated to cysteine. For the $Rev^{i, i+4}$ Ala 37 and Arg 41; for $Rev^{i, i+7}$ Arg 38 and Trp 45; for $Rev^{i, i+11}$ Arg 38 and Glu 49 were mutated to cysteine. Additionally, Arg 41 ($Rev^{i, i+7}$) and/or Arg 42 ($Rev^{i, i+7}$ and $Rev^{i, i+11}$) and/or Arg 42 and Trp 45 ($Rev^{i, i4}$ and $Rev^{i, i+11}$) were changed to alanine to prevent potentially unfavourable interactions with the crosslinker (Table 2). Peptides were produced in *E. coli* using *p*ET 31b expression system (Novagene) and all of them contain (depending on conditions) homoserine or homoserine lactone (Hs) at their C-terminus, which is an artefact amino acid left by CNBr cleavage of a peptide at a methionine site (*vide infra*) and their name/abbreviation will be presided by letter "b" designating their origin as biosynthesized peptides (Material and Methods).[259]

Peptide	Sequence
$bRev^{wt}$	^{33}DTRQARRNRRRRWRERQR^{50}–Hs
$bRev^{i, i+11}$	DTRQA**C**RNR**A**RR**A**RER**C**R–Hs
$bRev^{i, i+7}$	DTRQA**C**RN**AA**RR**C**RERQR–Hs
$bRev^{i, i+4}$	DTRQ**C**RRN**C**RRR**A**RERQR–Hs

Table 3: Sequence of modified bRev peptides and the wild type. Hs = homoserine or homoserine lactone

Oligonucleodites coding for bRev peptides were optimised for *E. coli* codon usage (Material and Methods)[260] and contained tandem repeats of each peptide separated by a methionine and overhangs to facilitate cloning *via* the *AlwN*1 cleavage site of *p*ET 31b plasmid. Initial clones were screened by *AlwN*1 restriction to detect the presence of its cognate restriction site which is abolished by successful ligation (Figure 33). The presence of the given insert was confirmed by DNA sequencing.

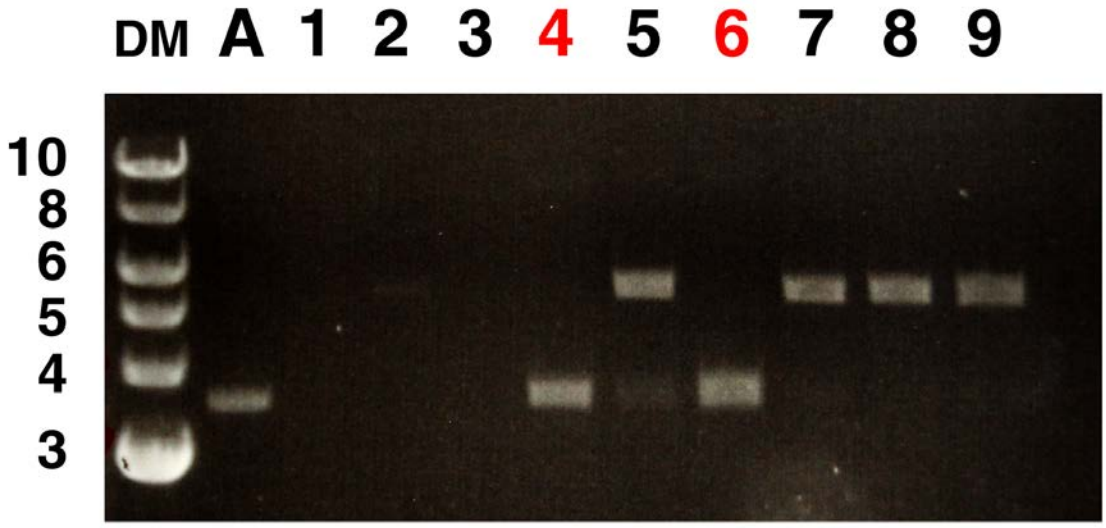

Figure 55: An exemplary bRev clones *AlwN*1 screening assay. DM - 1kbp DNA marker; A - positive control (plasmid with an insert); 1 - 9 clones screened; 4, 6 possible clones without an insert.

Interestingly despite of multiple attempts in varying conditions in all cases the final vector contained either frameshift mutations or complete deletion of the second peptide of the tandem.

Peptide	Sequence
bRev$^{i,\ i+11}$	MetDTRQACRNRARRARERCRMetLLEHHHHHHSTOP
bRev$^{i,\ i+7}$	MetDTRQACRNAARRCRERQRMetLLEHHHHHHSTOP
bRev$^{i,\ i+4}$	MetDTRQCRRNCRRRARERQRMetDTRQCRRNCCSSTTTT…
bRev^{wt}	MetDTRQARRNRRRRWRERQRMetAARAPPPPPLRSGCSTOP

Table 4: Exemplary sequencing results of Rev clones; the sequence is translated to the corresponding amino acid sequence and only the part after KSI is shown; Methionine and translation STOP is indicated in red.

Peptides were produced in *E. coli* BL21(DE3) expression system fused to ketosteroid isomerase (KSI), which promotes the formation of inclusion bodies, to enable separation of the product fusion peptide from soluble proteins before further purification in denaturing conditions (*vide infra*).[259,261] In the case of bRev$^{i,\ i+4}$ a mutation lead to expression of a "fusion protein" that considerably increased the size of produced protein and decreased the

final yield of the peptide (Table 4, Figure 56). Nevertheless all peptides apart from bRev^{wt} were biosynthesised and purified. Although different expression conditions, clones and cell lines were tried, no overexpression for bRev^{wt} gene was observed for any of the clones even though sequencing determined that they contained the correct insert sequence (Table 4, Figure 56). In order to provide positive control for further experiments Rev^{wt} synthesised by Fmoc protocols was provided by Dr Sabine Kneissl.

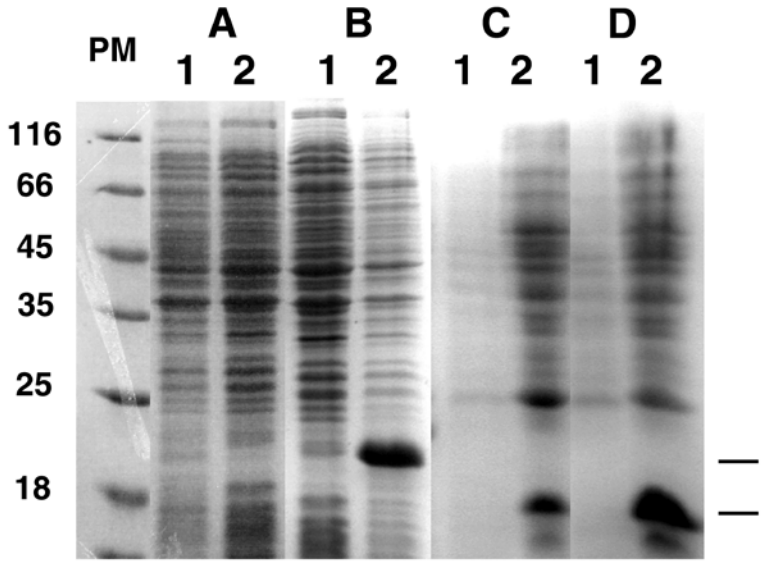

Figure 56: SDS-PAGE analysis of the Rev peptides expression. A - bRev^{wt}; bRev$^{i, i+4}$; bRev$^{i, i+7}$; bRev$^{i, i+11}$; 1 - cell culture before IPTG induction; 2 cell culture 6 h after IPTG induction.

Following expression, cells were harvested by a centrifugation and frozen frozen for storage (-80 °C). Pellets were then thawed and lysed by ultrasonication. After centrifugation the pellet containing the KSI fused peptides was washed (3x) (Material and Methods). Finally, the pellet was dissolved in a buffer containing 6M guanidinium hydrochloride, any remaining insoluble debris were removed by centrifugation and the supernatant was dialysed against water (Material and Methods). The pellets were treated with CNBr, liberating the desired peptides *via* cleavage at the methionine site, which as aforementioned produces a homoserine lactone at C-terminus. Peptides were purified by HPLC and their mass was confirmed by mass spectrometry (Figures 57 – 59, Table 5, Material and Methods). In all cases two peaks corresponding to each peptide were identified as the open and lactonized form of the C-terminal homoserine are observed under purification conditions. This suggests that a similar equilibrium might exist during binding measurements, effectively producing two populations of peptides. It is difficult to assess the significance of this change on the properties of the peptide and it is difficult to change the population distribution for the

duration of subsequent assays in order to assess possible different affinities of those two forms because the technique that was going to be used requires either acidic or basic condition both of which catalyze homoserine lactone to open form interconversion. Moreover determining the lactone/open ratio under given conditions is somewhat complicated as the HPLC technique changes them (most notably by the *p*H of the mobile phase). The presence of the lactone form considerably increases reactivity of the peptides especially towards amines.[261]

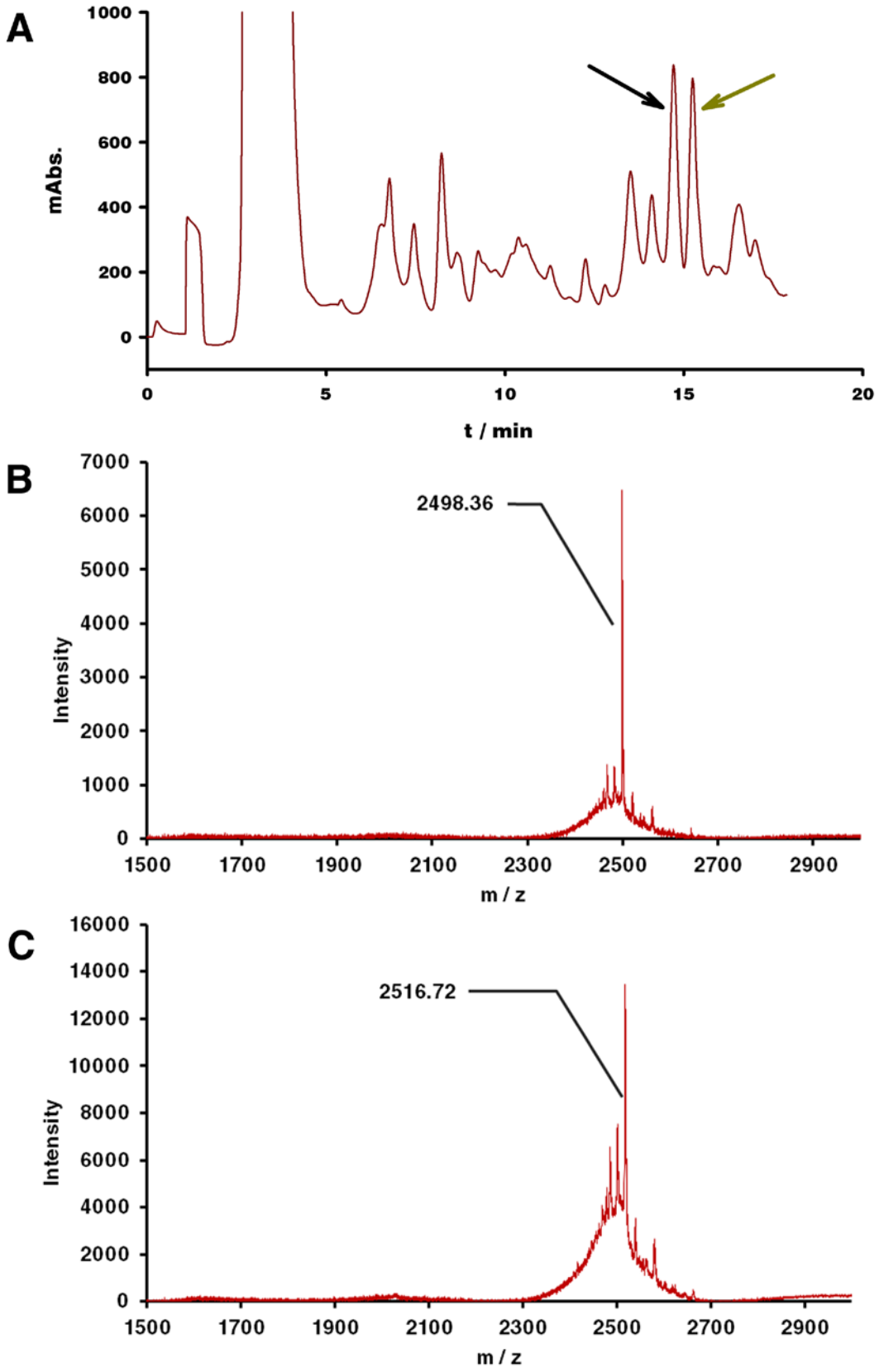

Figure 57: (A) HPLC chromatogram from the purification of the CNBr cleavage mixture of bRev$^{i, i+4}$ with the open homoserine form indicated with black arrow and closed with green. MALDI MS traces of bRev$^{i, i+4}$; (B) closed homoserine form and (C) open.

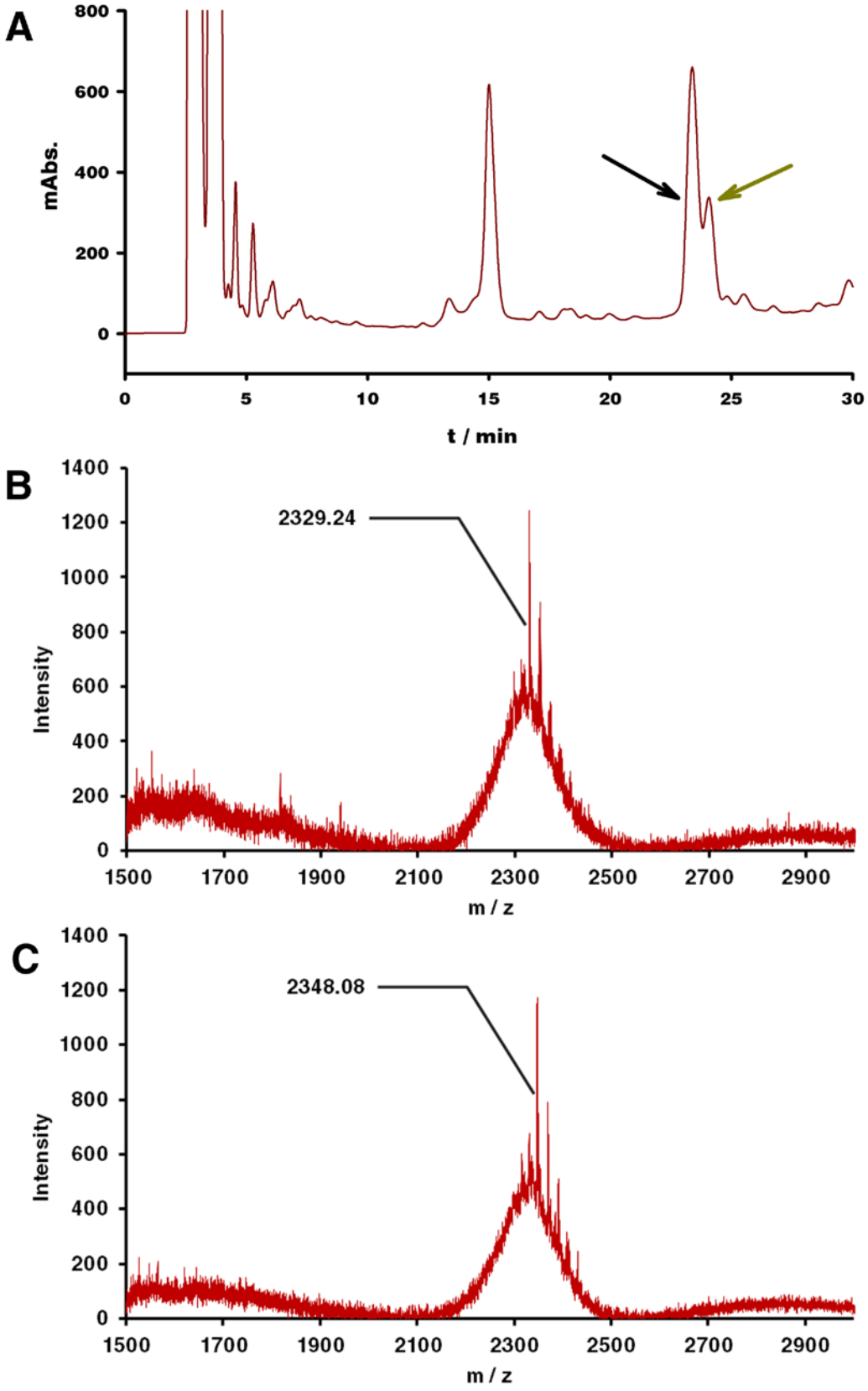

Figure 58: (A) HPLC chromatogram from the purification of the CNBr cleavage mixture of bRev$^{i, i+7}$ with the open homoserine form indicated with black arrow and closed with green. MALDI MS traces of bRev$^{i, i+7}$; (B) closed homoserine form and (C) open.

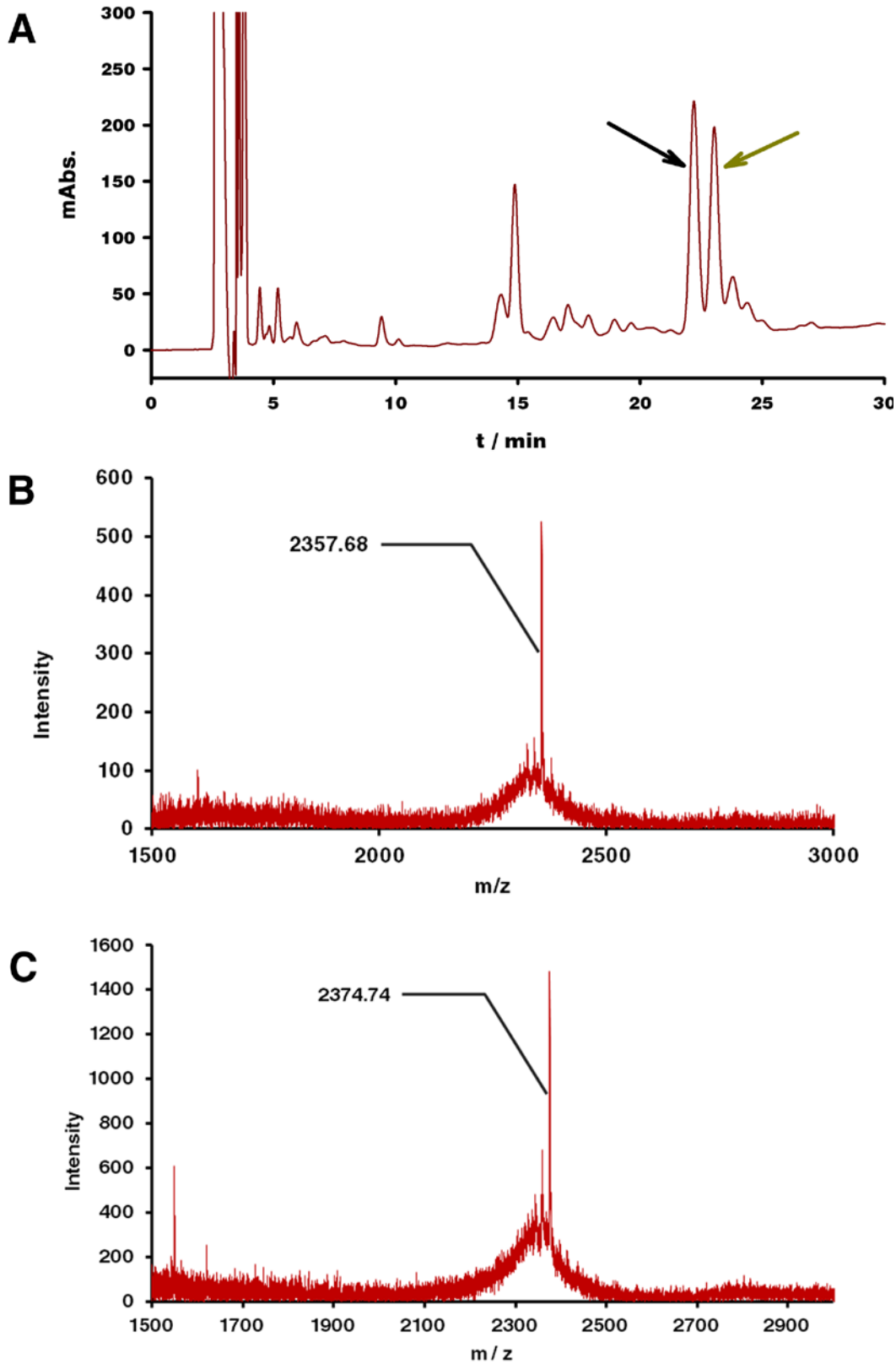

Figure 59: (A) HPLC chromatogram from the purification of the CNBr cleavage mixture of bRev$^{i, i+11}$ with the open homoserine form indicated with black arrow and closed with green. MALDI MS traces of bRev$^{i, i+11}$; (B) closed homoserine form and (C) open.

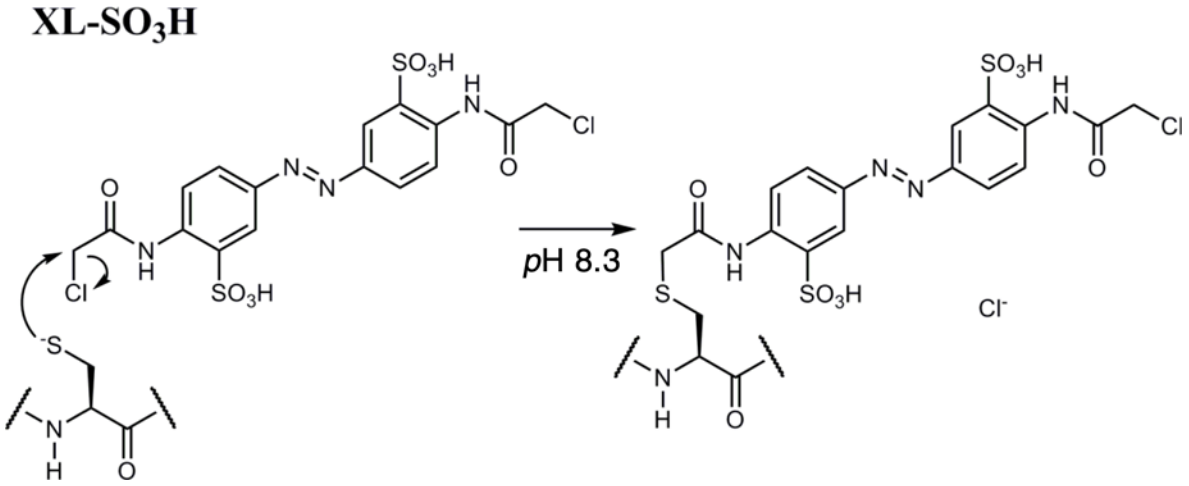

Figure 60: Scheme of the XL-SO_3H crosslinker used in this study and the crosslinking reaction mechanism.

Purified peptides were crosslinked similarly as it was described previously.[162] Briefly, peptide and the crosslinker were dissolved in 50 mM Tris:HCl, 2 mM TCEP, *p*H 8.3 to 1 mg/mL. The crosslinker solution was added in three portions followed by 2 h incubation at the room temperature and then overnight at 4 °C. In each case precipitation was observed and the precipitate had to be resuspended in 1:1 MeCN/H_2O before further purification (Material and Methods). From this moment on peptides crosslinked with the azobenzene derivative containing sulfonate substituents (Figure 60) will bear the name of the parent peptide suffixed with XL-SO_3H. Crosslinking $bRev^{i,\ i+4}$ proved difficult, altering the procedure by UV irradiating the reaction mixture or increasing temperature to improve the probability of correct cyclisation still yielded only minute amounts of crosslinked peptide (Material and Methods). All crosslinked products were subjected to HPLC purification and once again similar peaks corresponding to the open and closed form of homoserine were observed; the masses of crosslinked peptides were verified by mass spectrometry techniques (Figures 61 – 63, Table 4, Material and Methods).

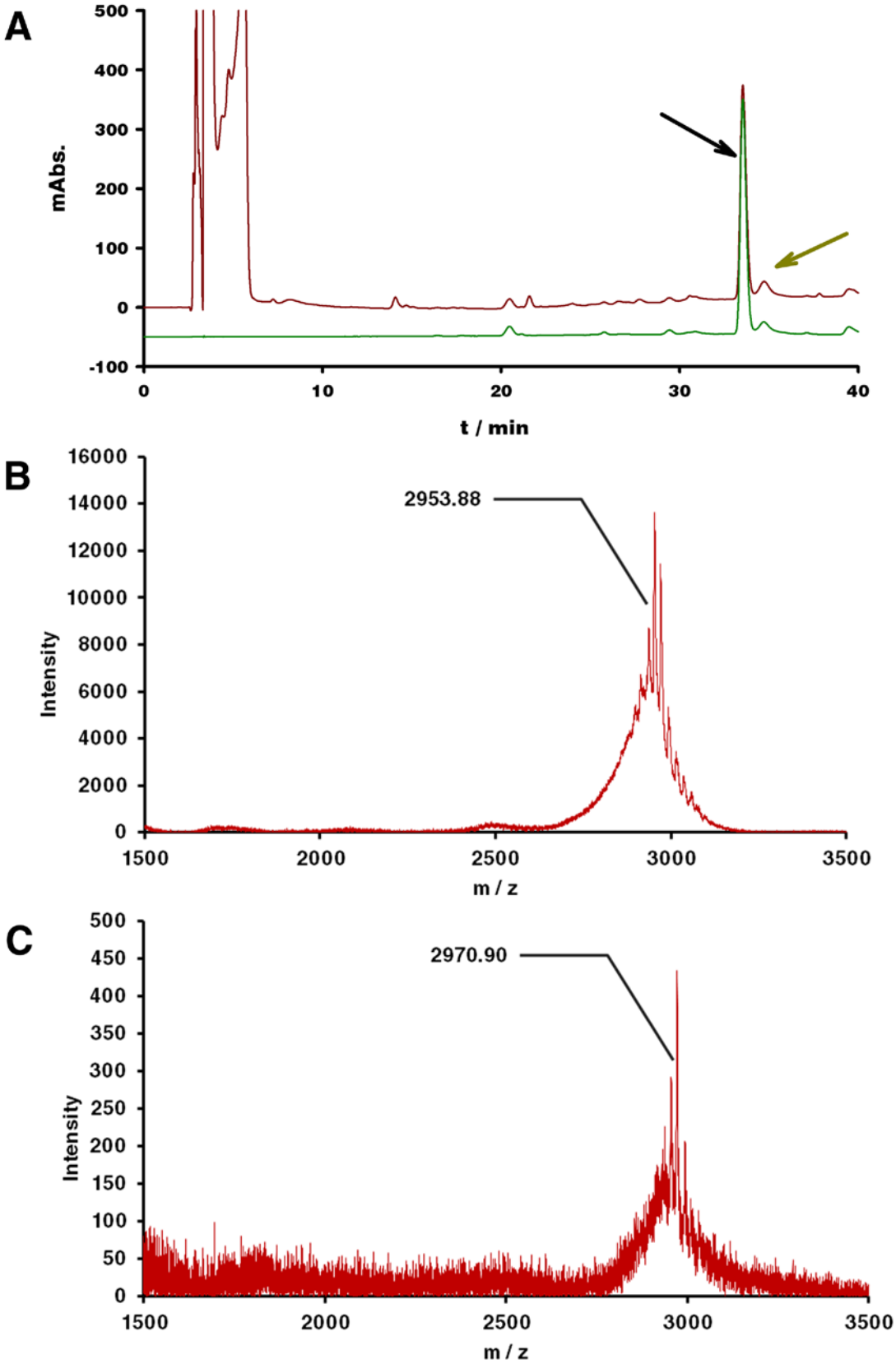

Figure 61: (A) HPLC chromatogram from the purification of the CNBr cleavage mixture of bRev$^{i, i+4}$-XL-SO$_3$H with the open homoserine form indicated with black arrow and closed with green (A). MALDI MS traces of bRev$^{i, i+4}$-XL-SO$_3$H; (B) closed homoserine form and (C) open.

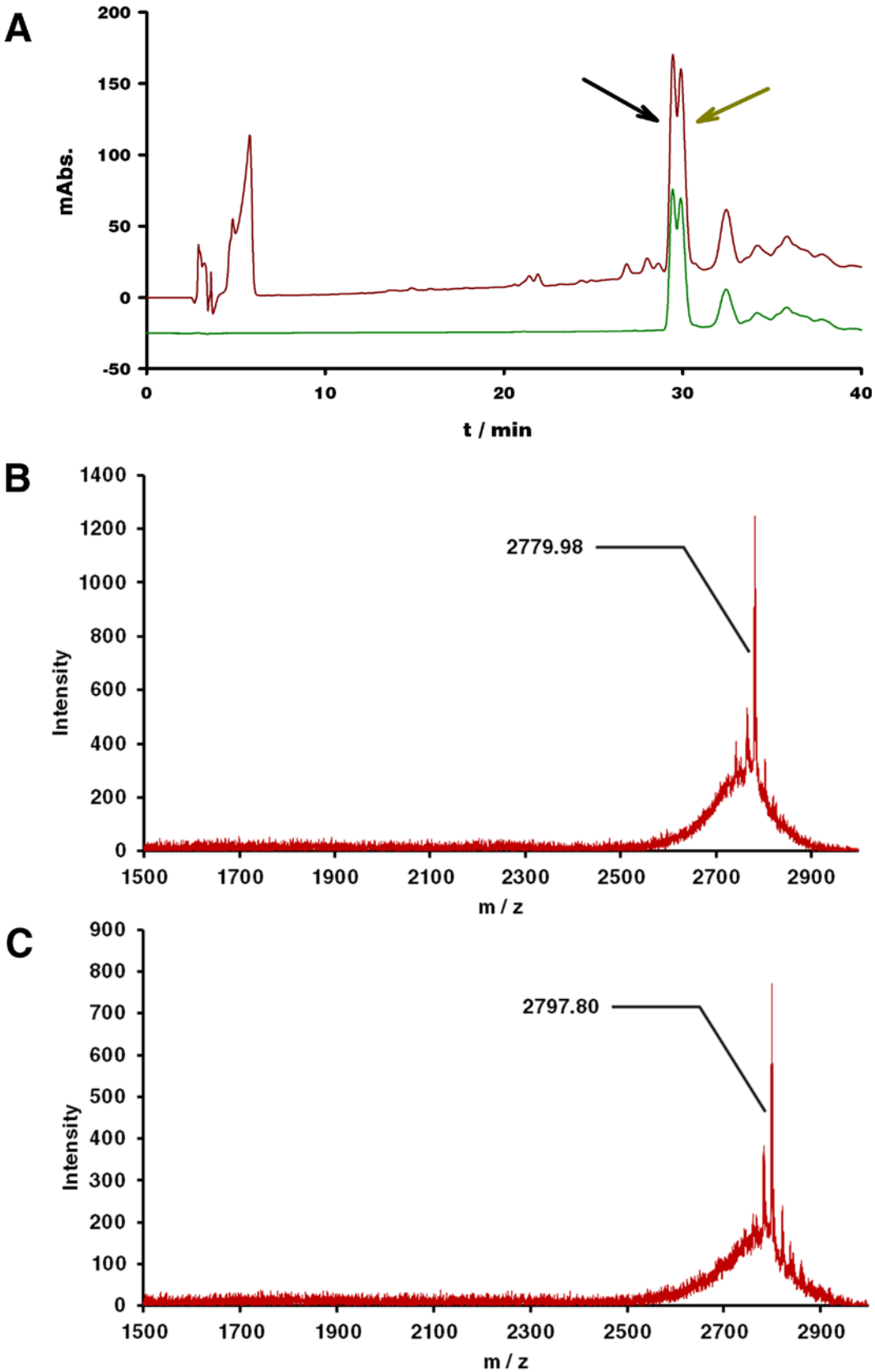

Figure 62: (A) HPLC chromatogram from the purification of the CNBr cleavage mixture of bRev$^{i, i+7}$-XL-SO$_3$H with the open homoserine form indicated with black arrow and closed with black (A). MALDI MS traces of bRev$^{i, i+7}$-XL-SO$_3$H; (B) closed homoserine form and (C) open.

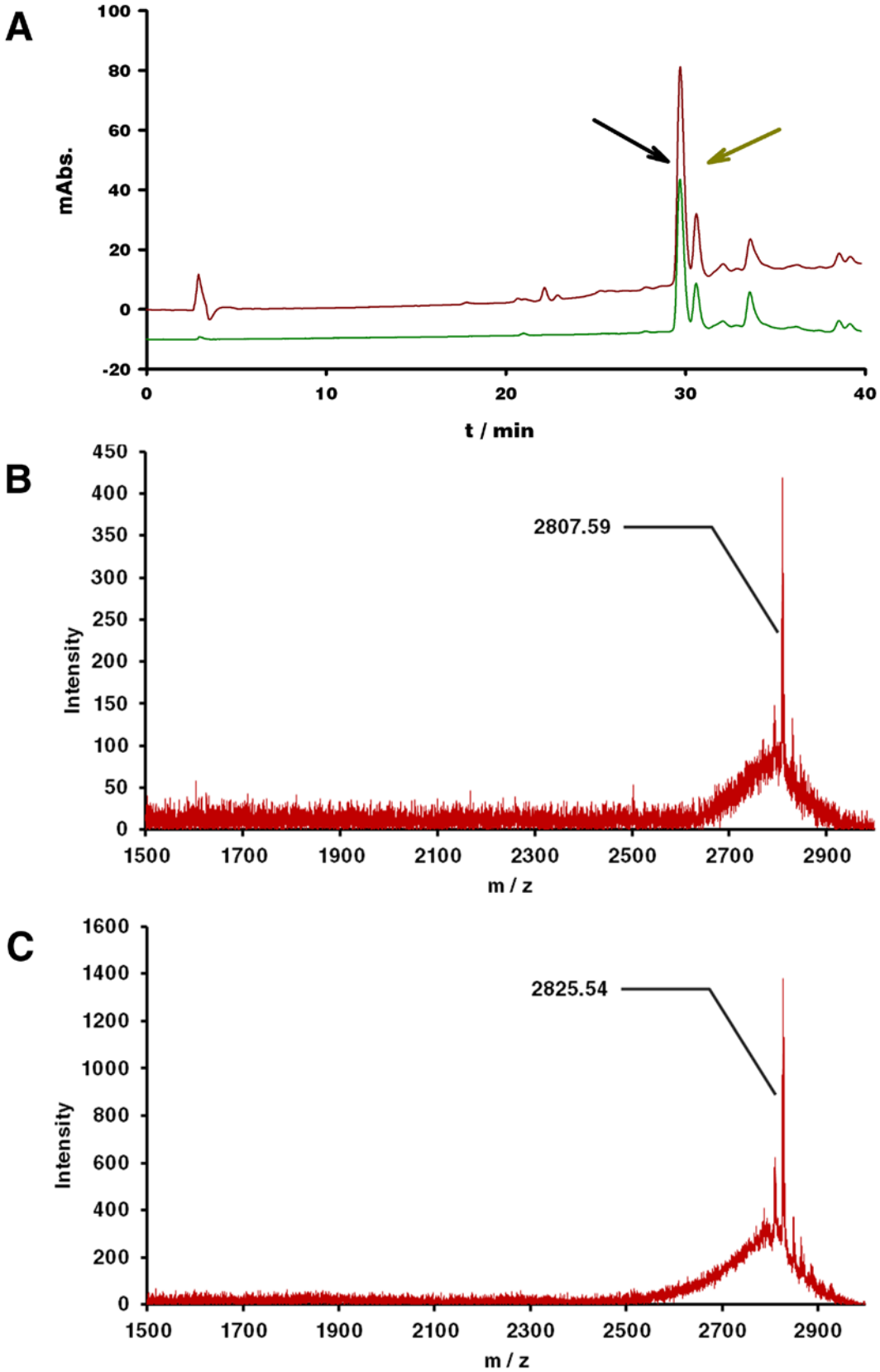

Figure 63: (A) HPLC chromatogram from the purification of the CNBr cleavage mixture of bRev$^{i, i+11}$-XL-SO$_3$H with the open homoserine form indicated with black arrow and closed with black (A). MALDI MS traces of bRev$^{i, i+11}$-XL-SO$_3$H; (B) closed homoserine form and (C) open.

Peptide	Calculated $[M+H]^+$	Observed $[M+H]^+$
$bRev^{i,i+4}$ (Hs-open)	2517.298[a]	2516.72
$bRev^{i,i+4}$ (Hs-lactone)	2499.287[a]	2498.36
$bRev^{i,i+7}$ (Hs-open)	2347.170[a]	2348.08
$bRev^{i,i+7}$ (Hs-lactone)	2329.159[a]	2329.24
$bRev^{i,i+11}$ (Hs-open)	2375.213[a]	2374.74
$bRev^{i,i+11}$ (Hs-lactone)	2357.202[a]	2357.68
$bRev^{i,i+4}$-XL-SO_3H (Hs-open)	2969.307	2970.90
$bRev^{i,i+4}$-XL- SO_3H (Hs-lactone)	2951.296	2953.88
$bRev^{i,i+7}$-XL-SO_3H (Hs-open)	2799.179	2797.80
$bRev^{i,i+7}$-XL- SO_3H (Hs-lactone)	2781.168	2779.98
$bRev^{i,i+11}$-XL-SO_3H (Hs-open)	2827.222	2825.54
$bRev^{i,i+11}$-XL- SO_3H (Hs-lactone)	2809.211	2807.59

Table 5: Observed and calculated masses of purified bRev peptides; [a] calculated assuming reduced cysteines

2.1.2.2 CD and UV

The kinetic parameters of peptides relaxing after UV irradiation were estimated by measuring relaxation rates (k) at various temperatures (T) and plotted. Relaxation rates were measured by following a decrease in 365 nm absorbance of the crosslinked peptides over time after irradiation with 365 nm UV light (Material and Methods) and fitting a first order kinetics equation (Equation 1) to that data. Measurements were performed using Chirascan CD spectrometer by either taking a full spectrum (180 to 600 nm) after each interval for temperatures 4 °C to 20 °C or by measuring at 208, 222 and 365 nm for temperatures 25 °C to 40 °C.

$$A_{365} = A_{365}^{ir.} e^{-kt} + A_{365}^{dark}$$

Equation 1: The first order kinetics equation fitted to a plot of 365 nm absorbance against time; where A_{365} is absorbance at any time; $A_{365}^{ir.}$ is absorbance after irradiation (t=0); A_{365}^{dark} is absorbance before irradiation; k is the rate constant and t is time.

To analyze the kinetic parameters the logarithmic (linear) version (Equation 3) of Arrhenius equation (Equation 2) was fitted to the plot of the natural logarithm of relaxation rates (lnk) against the reciprocal of temperature (1/T). This allows estimating activation energy (E_a) and preexponential factor A expressed as lnA (Figure 64).

$$k = Ae^{-\frac{E_a}{RT}}$$

Equation 2: Arrhenius equation; where k is a rate constant; A is a preexponential factor; E_a is an activation energy; R is the gas constant and T is temperature.

$$lnk = lnA - \frac{E_a}{RT}$$

Equation 3: Linear form of the Arrhenius equation.

Both peptides exhibit slower kinetics as compared to the free crosslinker which has been observed before. At 15 °C *i, i*+7 spacings makes the crosslinked peptide relax slower as compared to *i, i*+11 (Table 6). The slower relaxation is determined by the activation energy which is 5 kJ/mol lower for *i, i*+7 peptide, but at the same time lnA is higher and not enough so to compensate for the increase in the activation energy barrier. This is further discussed together with the results for Fmoc synthesized peptides (*vide infra*).

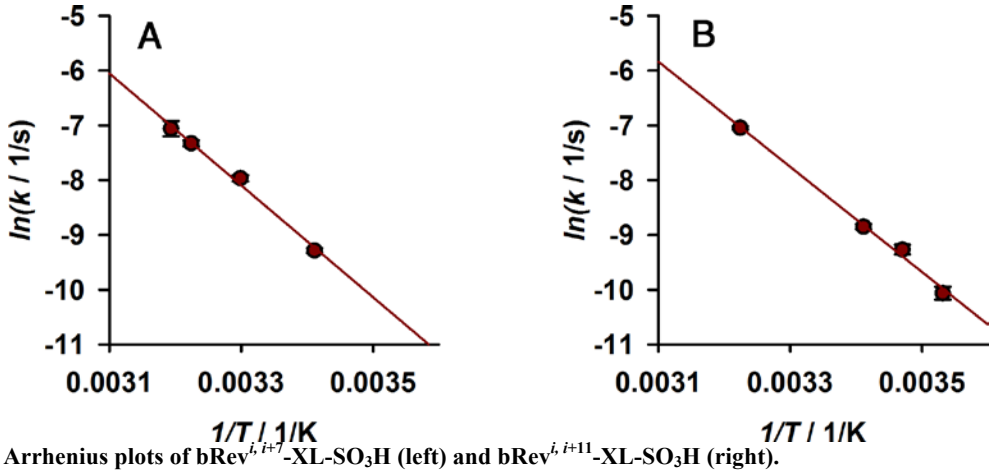

Figure 64: Arrhenius plots of bRev$^{i, i+7}$-XL-SO$_3$H (left) and bRev$^{i, i+11}$-XL-SO$_3$H (right).

Peptide	Half Life at 15 °C [min]	Activation Energy [kJ/mol]	lnA
bRev$^{i, i+7}$-XL-SO$_3$H	217.7 ± 16.5	83.0 ± 4.4	25.6 ± 1.78
bRev$^{i, i+11}$-XL-SO$_3$H	123.0 ± 11.3	78.0 ± 1.8	23.9 ± 1.84

Table 6: Kinetic parameters of UV$_{363}$ relaxation after UV$_{363}$ irradiation for all crosslinked bRev peptides.

Both bRev$^{i, i+7}$-XL-SO$_3$H and bRev$^{i, i+11}$-XL-SO$_3$H exhibit changes in their CD spectra upon irradiation, indicating structural changes (Figure 65). Although the effect is much stronger for bRev$^{i, i+11}$-XL-SO$_3$H, transition from more to less helical peptide can be observed for bRev$^{i, i+11}$-XL-SO$_3$H and *vice versa* for bRev$^{i, i+7}$-XL-SO$_3$H. Both characteristic alpha helical minima (208, 222 nm) and one maximum (190 nm) in their CD spectra change intensity and a peak wavelength shift is observed for 208 nm. Upon irradiation bRev$^{i, i+11}$-XL-SO$_3$H loses α-helical character; bRev$^{i, i+7}$ becomes more α-helical after irradiation. It is important to note that the conversion from *trans* to *cis* state of the crosslinker is not complete, and therefore the CD spectra of peptides after UV irradiation represent the contribution of both forms with the *trans-cis* conversion between 70 and 80% complete.

Interestingly, both peptides exhibit CD signals arising from the azobenzene: with maxima at 265 nm (small) and 360 nm and a minimum at 440 nm. bRev$^{i, i+11}$-XL-SO$_3$H also exhibits a small maximum at 250 nm. All signals apart from the one at 265 nm have been reported previously, however it was initially suggested that only the non-sulfonated

crosslinker exhibited them. Although the crosslinker itself is not chiral it contacts a chiral environment provided by the chiral peptide. The intensity of those signals is much stronger for bRev$^{i,\ i+11}$-XL-SO_3H then bRev$^{i,\ i+7}$-XL-SO_3H. Moreover, the change in intensity upon irradiation is reversed for bRev$^{i,\ i+11}$-XL-SO_3H compared to bRev$^{i,\ i+7}$-XL-SO_3H. Taken together this suggests a direct correlation between crosslinker CD signal intensity and the average amount of helicity. Indeed by examining the CD spectra of bRev$^{i,\ i+11}$-XL-SO_3H taken at different temperatures one can see a concomitant shift in helicity and intensity of the crosslinker CD signals (data not shown). This property might allow estimation of a helical content of crosslinked peptides in conditions that are not conducive for other techniques of structure determination.

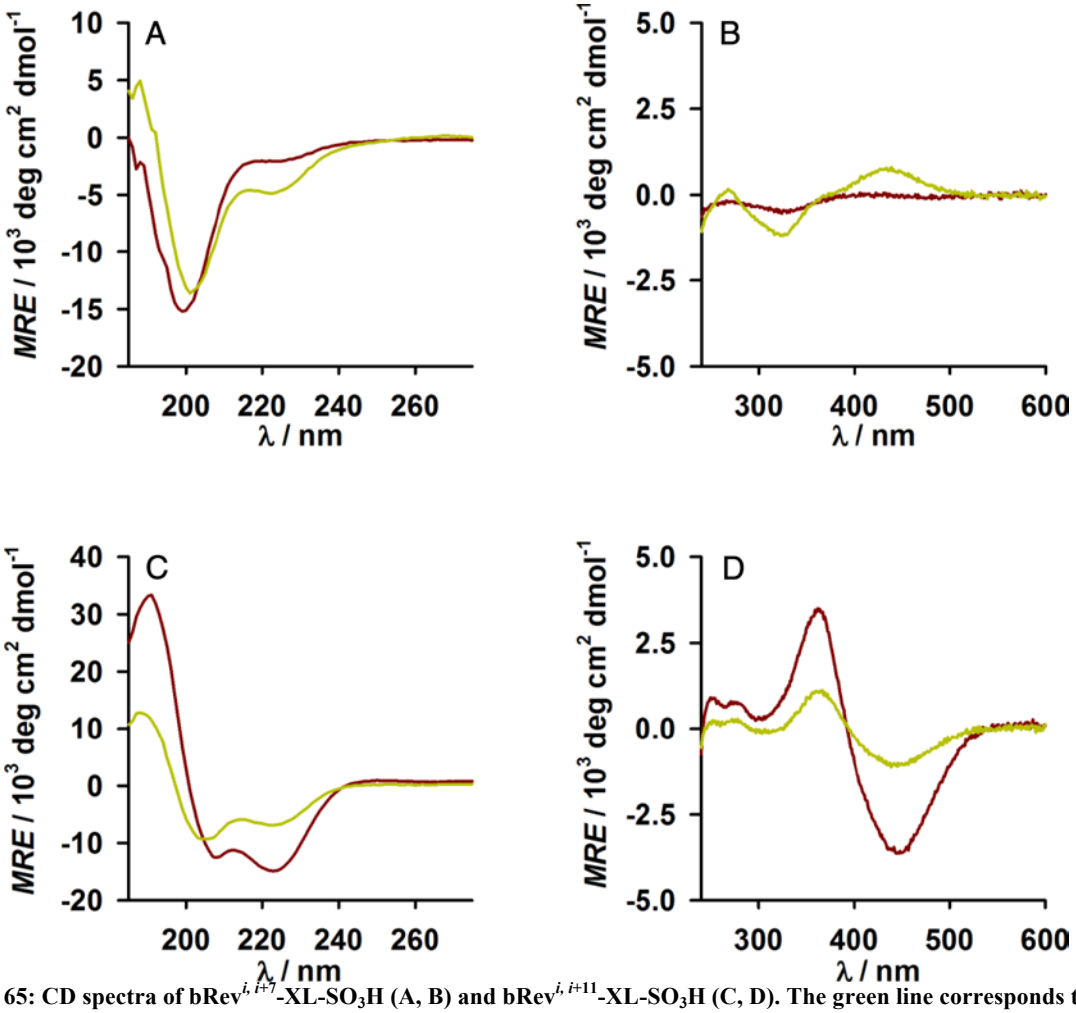

Figure 65: CD spectra of bRev$^{i,\ i+7}$-XL-SO_3H (A, B) and bRev$^{i,\ i+11}$-XL-SO_3H (C, D). The green line corresponds to the UV irradiated peptide and red to dark adapted.

2.1.2.3 Transcription and purification of RRE

RRE IIB sequence was synthesised using the Superscript kit (Ambion) according to manufacturer's specifications. Briefly, RNA polymerase and the four nucleotides are mixed together with the DNA template strand and the enzymatic synthesis is allowed to proceed at

37 °C. For the purpose of radioactive labelling of RRE RNA *in vitro* transcription mixture was supplemented with α-^{32}P-UTP. The following DNA sequence was used for the transcription:

5'AATTTAATACGACTCACTATA**GGGTCTGGGCGCAGCGCAAGCTGACGGTAC AGGCC**-3' (sense strand, coding strand, RRE sequence in bold font) and 5'-GGCCTGTACCGTCAGCTTGCGCTGCGCCCAGACCCTATAGTGAGAGTCGTATTAA ATT-3' (antisense strand, template). Before transcription the two complementary strands were annealed by heating up to 95 °C and cooling down to 15 °C over 16 h (Material and Methods). RRE IIB in vitro transcription gave a single product as determined by denaturing PAGE and visualized by exposing a phosphorimager plate followed by scanning with Typhoon Gel Imager (Figure 66). The product was refolded by heating up to 95 °C and cooling down to 15 °C over 16 h (Material and Methods).

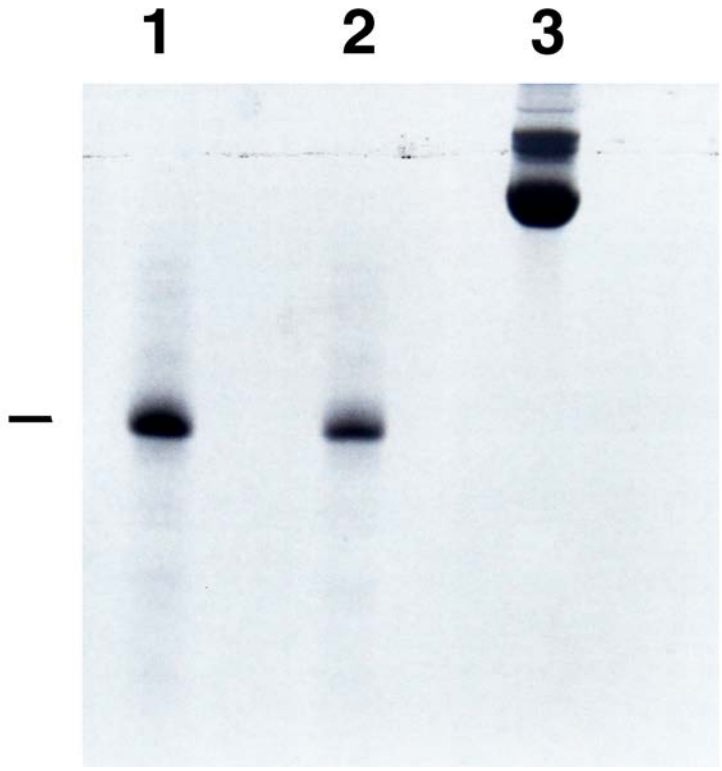

Figure 66: Denaturing-PAGE analysis of RRE *in vitro* transcription products. 1 and 2 - RRE *in vitro* transcription after 2 and 6 hours; 3 - pTRI-Xef control template transcription product. The black bar indicates the postion of RRE.

2.1.2.4 Electrophoretic mobility shift assay (EMSA) of RRE/Rev interaction

In order to estimate the binding affinities of the bRev peptides an electrophoresis mobility shift assay (EMSA) was established. In this experiment ^{32}P labelled RNA is titrated with a peptide. Any complexes formed upon binding of the ligand usually have sufficiently

different physical properties to migrate with different speed through a polyacrylamide gel as compared to free RNA. In this way a ligand concentration dependent binding curve can be plotted and a K_d estimated. The densitometric intensity of bands appearing at the RNA-peptide complex gel position or disappearing at the free RNA position can be used and plotted against the concentration of the peptide. Alternatively an upper limit for the K_d can be provided based on visual analysis the bands intensities.

A number of attempts were made to identify conditions suitable for visualizing the Rev-RRE interaction; however none showed any shift with either crosslinked or uncrosslinked peptides as a titrant. The only case where a shift was visible corresponded to the Rev^{wt} peptide and the K_d was estimated to be lower than 25 nM (Figure 67), which is in good agreement with previous reports.[129] The figure below summarizes all results (Figure 68). These results were in contrast to the initial findings of Dr Sabine Kneissl, and suggested that one of the following might be occurring: the conditions of the gel electrophoresis are not suitable for peptides crosslinked with the azobenzene derivative used in this study; presence of an additional reactive amino acid (homoserine lactone) interfered with the assay or the initial results were misleading and the system requires redesigning. Redesigning could be necessary due to suboptimal position of the crosslinker (mutation introducing cysteines), interference of the crosslinker with the Rev-RRE or compromised upon clash-relieving mutations.

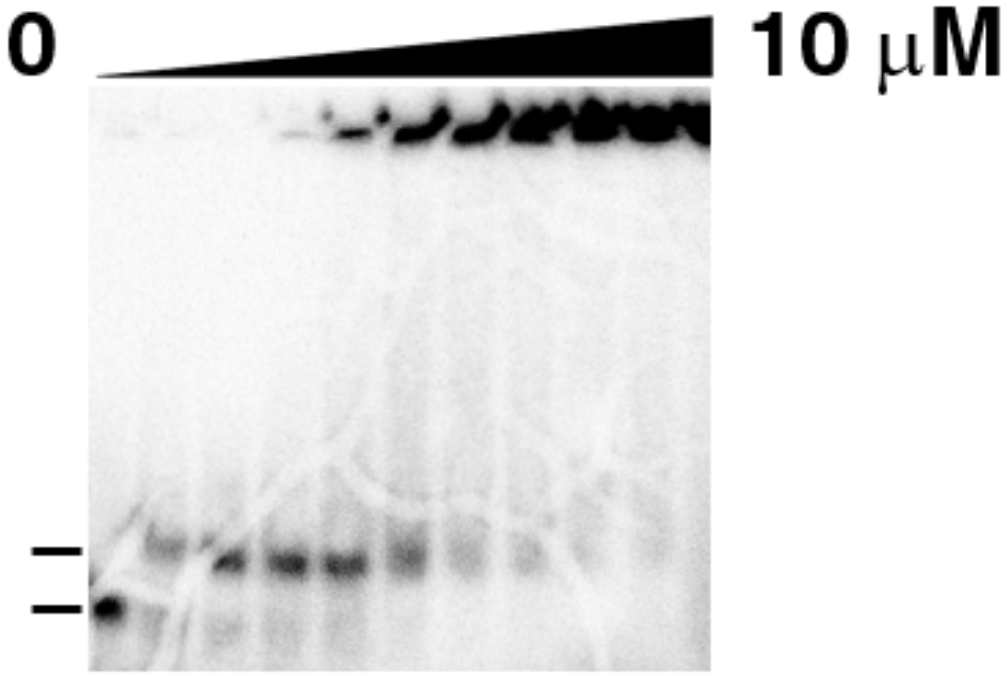

Figure 67: EMSA gel scan visualizing the interaction between RRE (constant concentration) and increasing concentrations (5, 10, 25, 50, 100, 200, 500, 1000, 2000, 5000 and 10000 nM) of Rev^{wt}. The black bars indicates the position of free RRE and bound to Rev^{wt}.

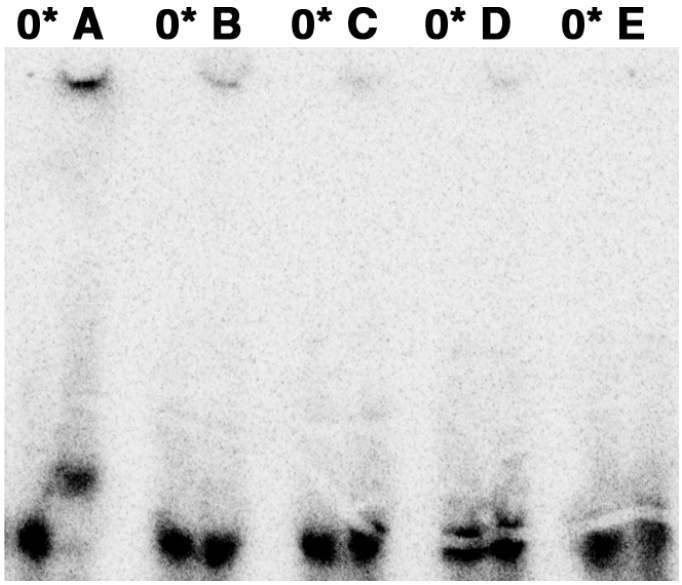

Figure 68: EMSA gel visualizing the interaction between (A) RRE and Rev^{wt} or both not crosslinked and crosslinked $bRev^{i, i+7}$ (B and D respectively) and $bRev^{i, i+11}$ (C and E respectively). Concentration of the RRE is held constant and Rev peptides are added at 250 nM concentration. Lanes labelled as 0* contain RRE without peptide.

2.1.3 Fmoc synthesized Rev peptides

2.1.3.1 Designing, synthesis and purification of Rev peptides

Due to low and variable yields, potential interference from the homoserine cleavage artefact and labelling limitations imposed by biosynthesis of peptides it was decided to optimize the Fmoc synthesis of Rev peptides. The same basic design as for the biosynthesized peptides was retained with two exceptions: peptides did not contain free C-terminal carboxylic acid but an amide at C-terminus (Table 3). Additionally $Rev^{i,\ i+4}$ was not synthesised as it was already known that its crosslinking was difficult. Since Rev peptides are rich in arginine, their chemical synthesis *via* Fmoc protocols might be troublesome; sidechain protected Fmoc-arginine is bulky and one of the most expensive amino acids. Initially synthesis of Rev peptides failed to provided purifiable mixtures, yielding only a distribution of arginine deletion mutants that could be indentified from the complex chromatograms by mass spectroscopy. Double coupling of arginines, a well calibrated and perfectly functional peptide synthesizer and longer cleavage durations in a large volume of trifluoroacetic acid with very fresh scavengers were crucial for success. Fmoc chemistry allowed for easy N-terminal carboxyfluorescein labelling and opened opportunities for the development of sensitive and "real-time" assays for the affinity measurements.

Peptides were synthesized on a CEM peptide synthesizer on a rink amide resin with HBTU as the activator. Coupling temperatures were adjusted from standard 75 °C to 50 °C for cysteines and arginines were additionally sidechain protected with acetic acid anhydride and double coupled (Material and Methods). After final deprotection the resin bound peptide was incubated in 20 ml of the cleavage cocktail (Material and Methods) for 6 h and filtered. After evaporation under nitrogen flow cold diethyl ether was added (49 ml) and this suspension was incubated at -20 °C overnight the precipitate was recovered by centrifugation, dissolved in water (+0.5% TFA) and purified *via* HPLC techniques (Material and Methods). When it was required resin-bound peptides before the final deprotection were labelled with fluoresceineamide (FAM). FAM labelling was performed by immersing the resin-bound peptide in 1.5 ml of DMF containing 50 mg of FAM, 20 mg of HOBT and 26 µl diisopropylcarbodiimide for at least 2 h prior to sidechain deprotection.

Eventually all six basic uncrosslinked peptides (Rev^{wt}, $Rev^{i,\ i+7}$, $Rev^{i,\ i+11}$, FAM-Rev^{wt}, FAM-$Rev^{i,\ i+7}$, FAM-$Rev^{i,\ i+11}$) were acquired in relatively high yields and their masses were verified by MS techniques (Figure 69–74, Table 7).

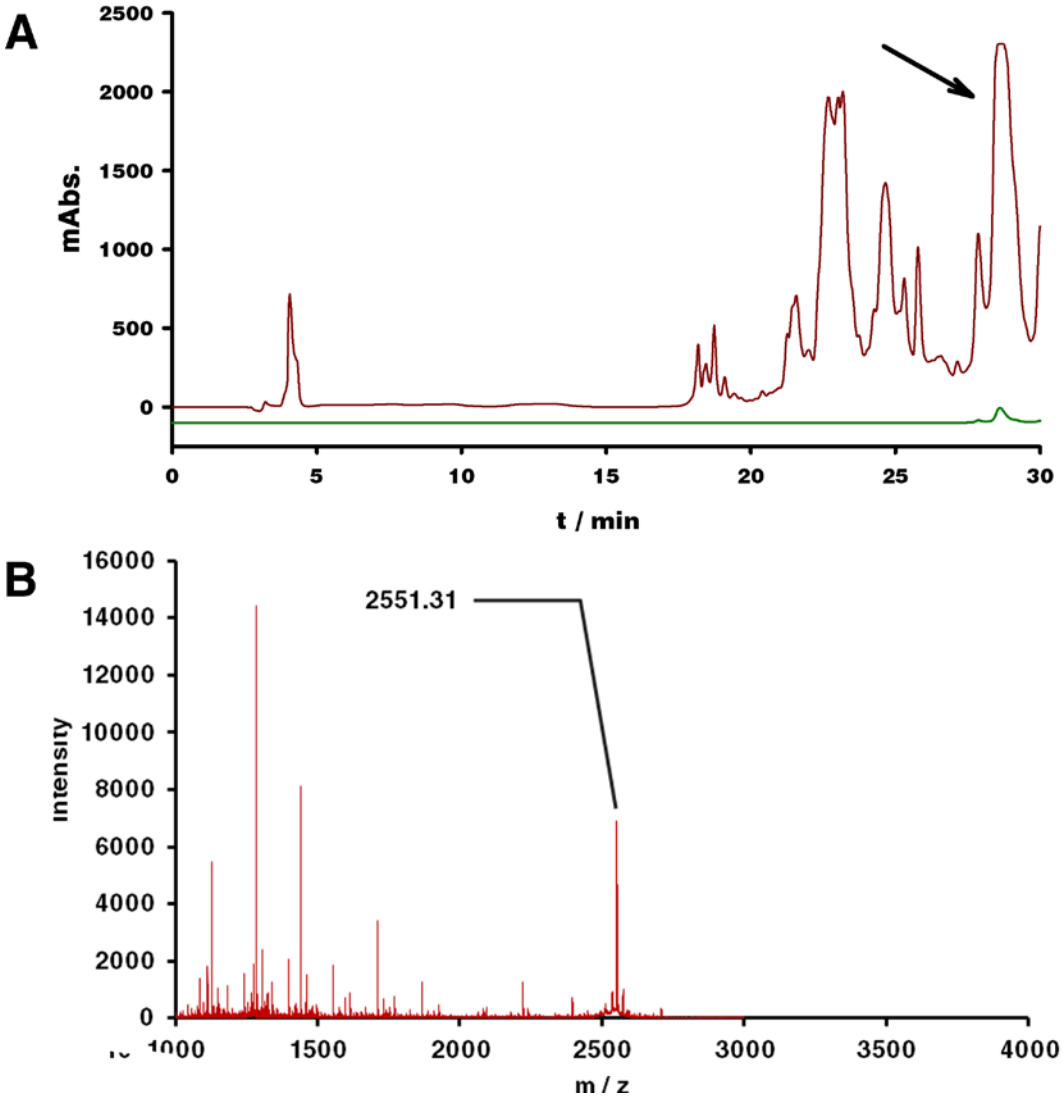

Figure 69: (A) HPLC chromatogram (red 210 nm, green 280 nm absorption) from the purification of the crude Rev^{wt} indicated with black arrow. (B) MALDI MS trace of Rev^{wt}.

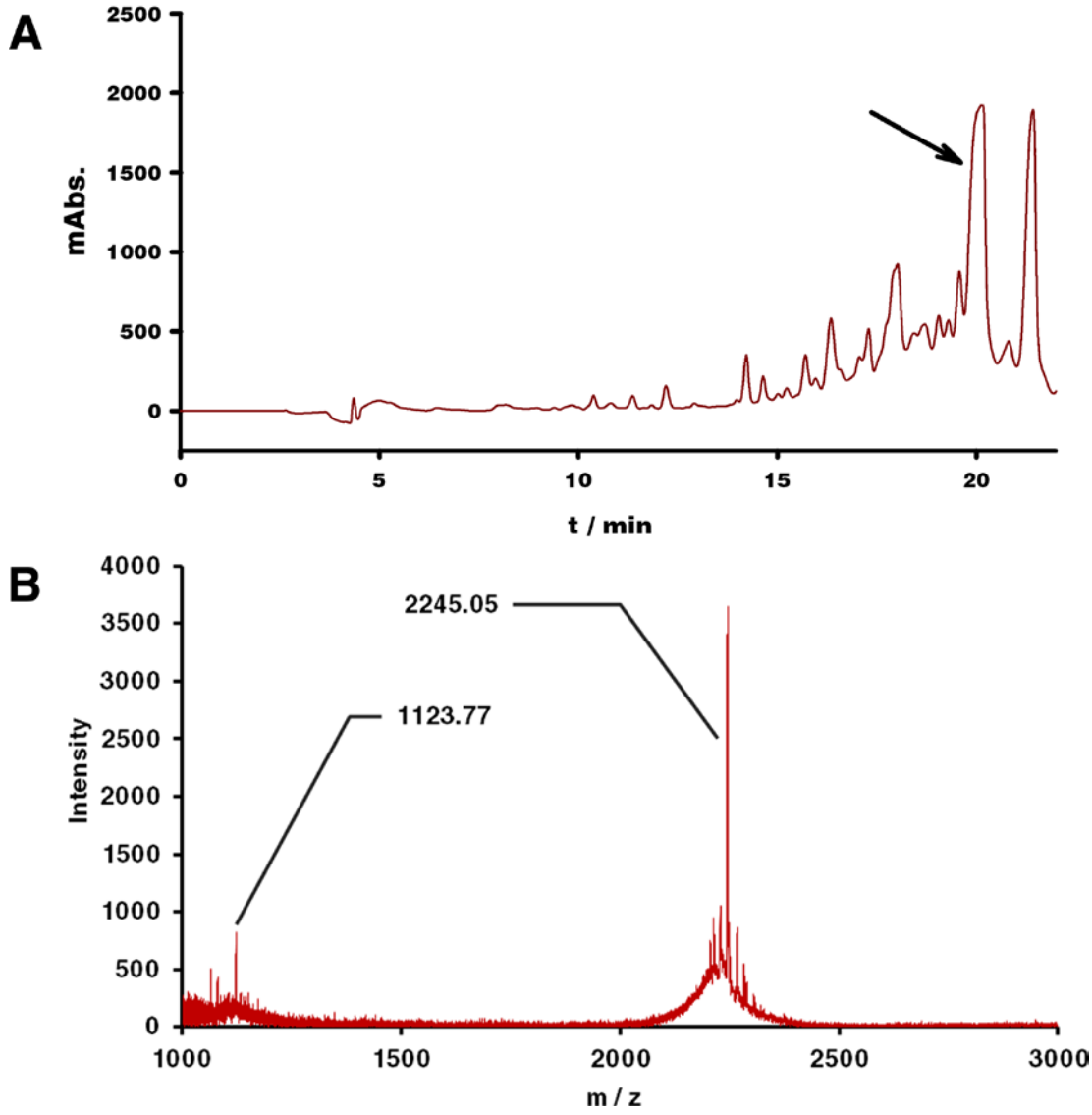

Figure 70: (A) HPLC chromatogram (red 210 nm absorption) from the purification of the crude Rev$^{i, i+7}$ indicated with black arrow. (B) MALDI MS trace of Rev$^{i, i+7}$.

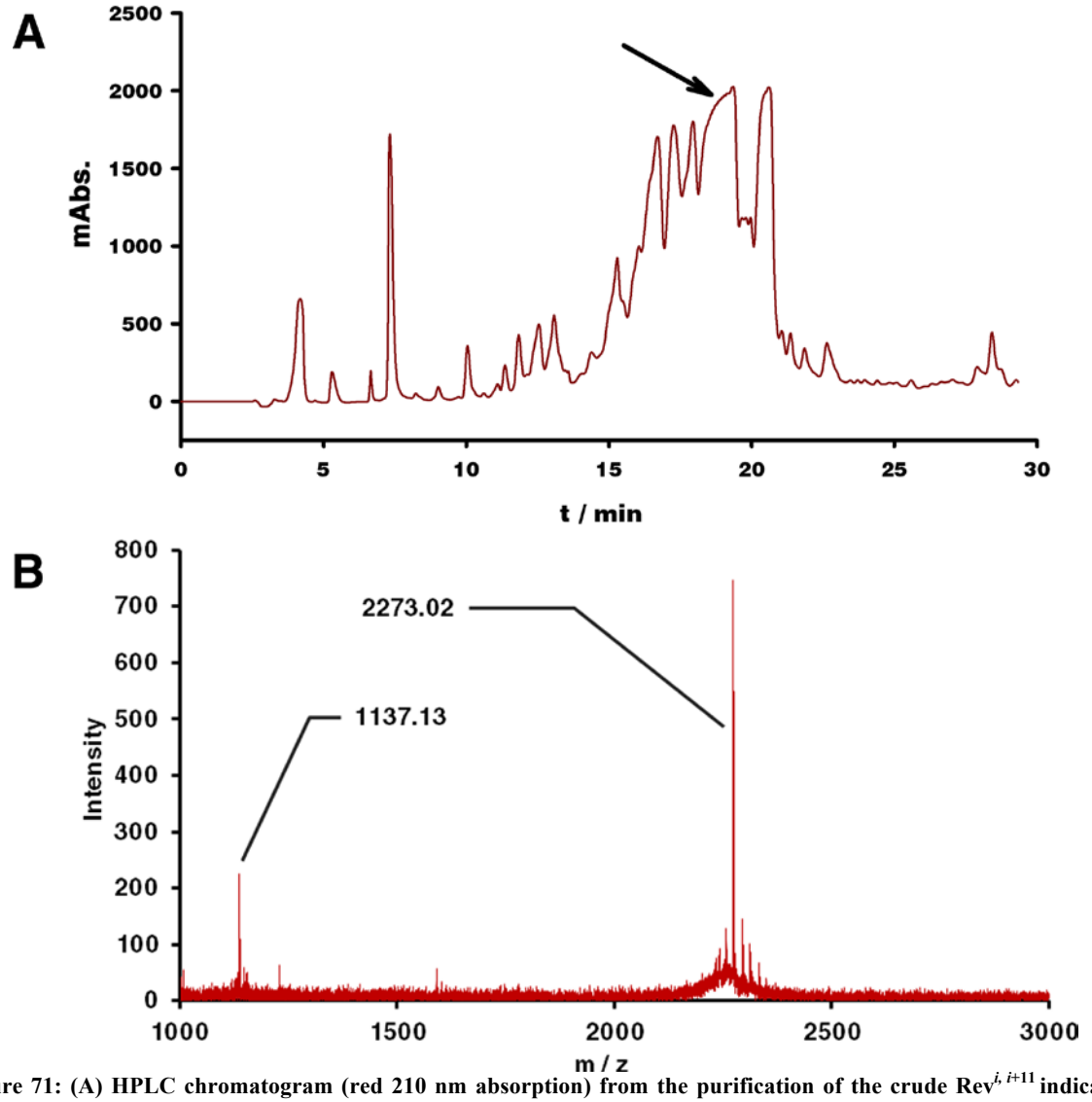

Figure 71: (A) HPLC chromatogram (red 210 nm absorption) from the purification of the crude Rev$^{i, i+11}$ indicated with black arrow. (B) MALDI MS trace of Rev$^{i, i+11}$.

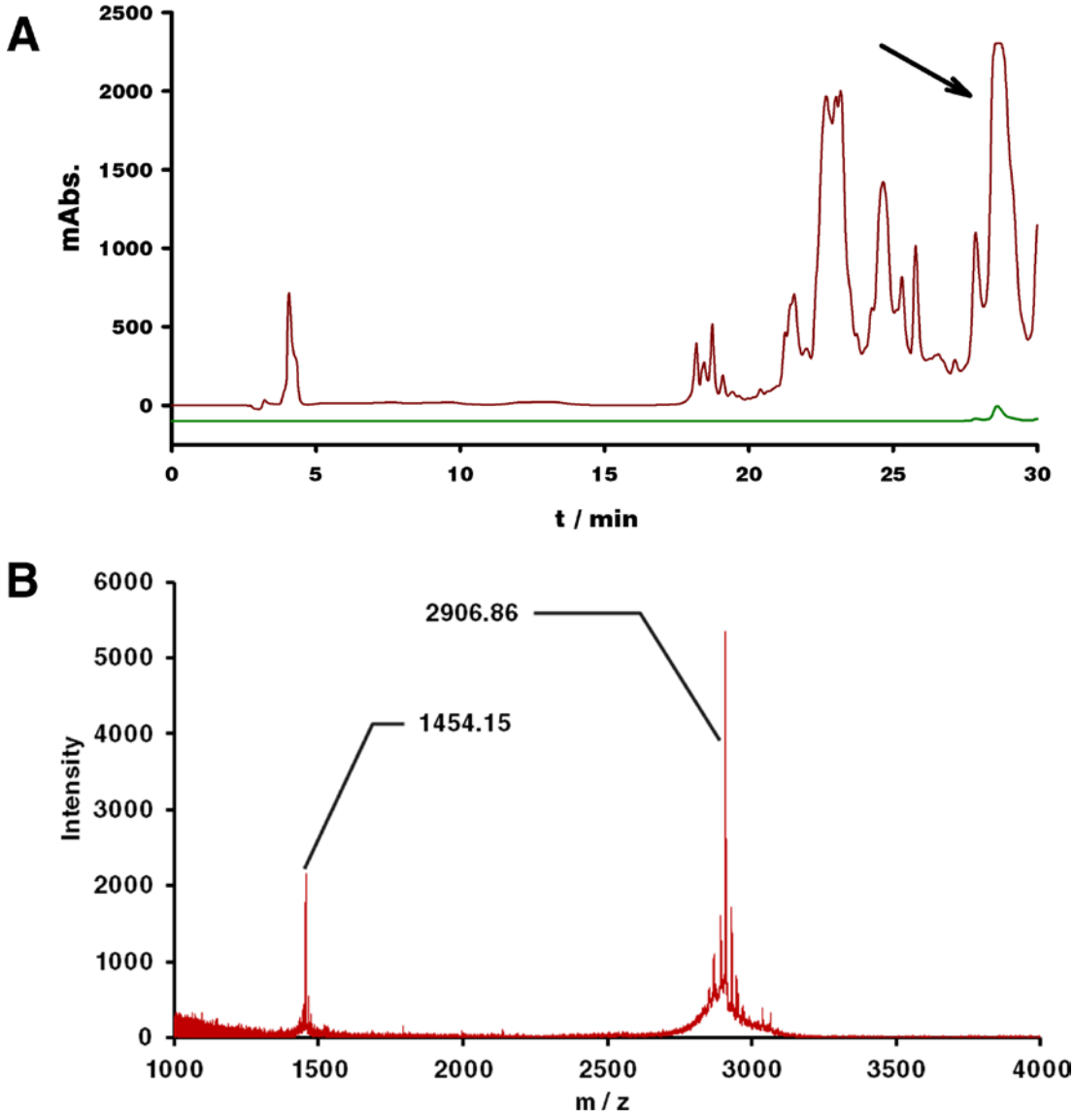

Figure 72: (A) HPLC chromatogram (red 210 nm, green 495 nm absorption) from the purification of the crude FAM-Rev^{wt} indicated with black arrow. (B) MALDI MS trace of FAM-Rev^{wt}.

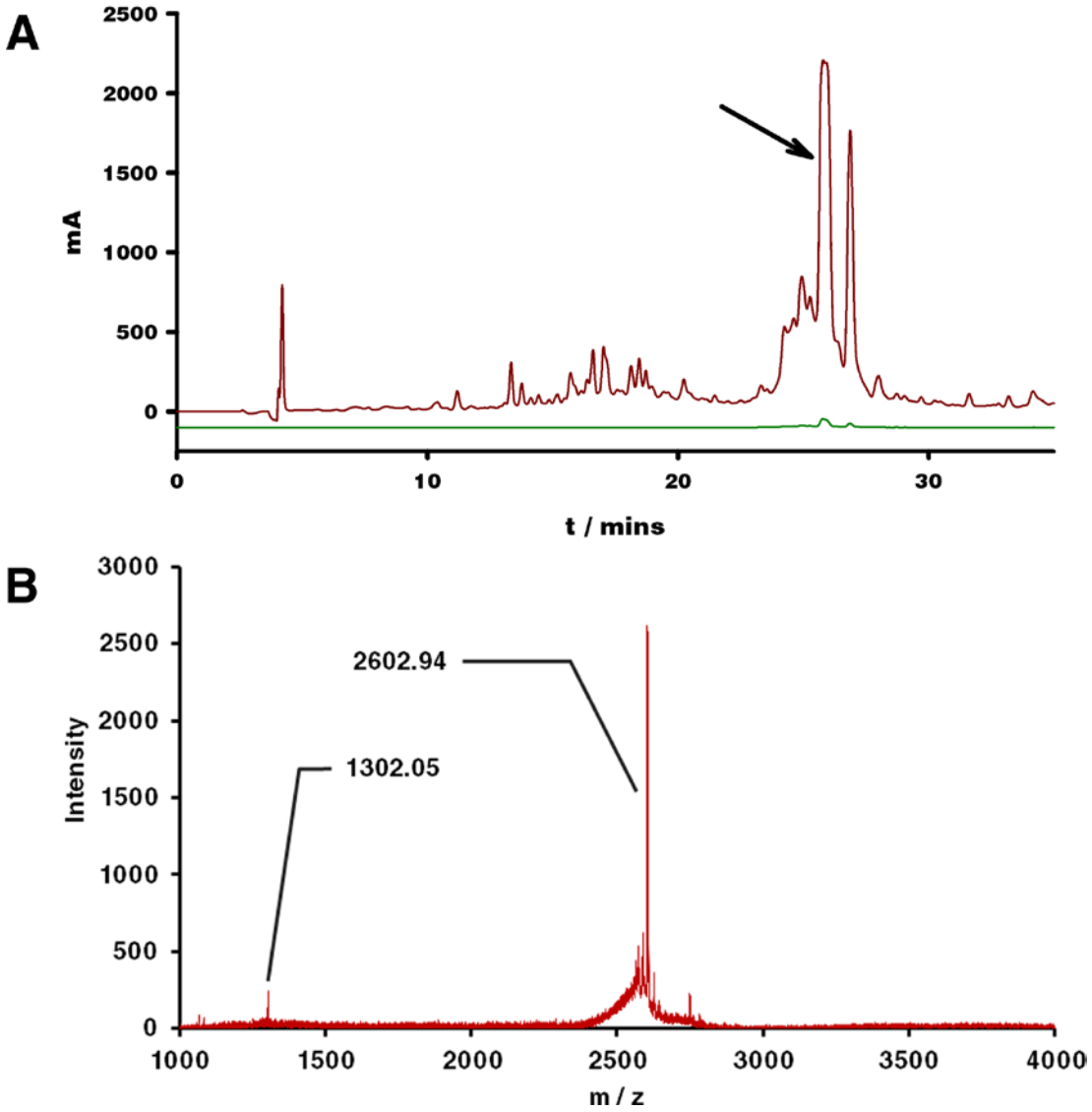

Figure 73: (A) HPLC chromatogram (red 210 nm, green 495 nm absorption) from the purification of the crude FAM-Rev$^{i, i+7}$ indicated with black arrow. (B) MALDI MS trace of FAM-Rev$^{i, i+7}$.

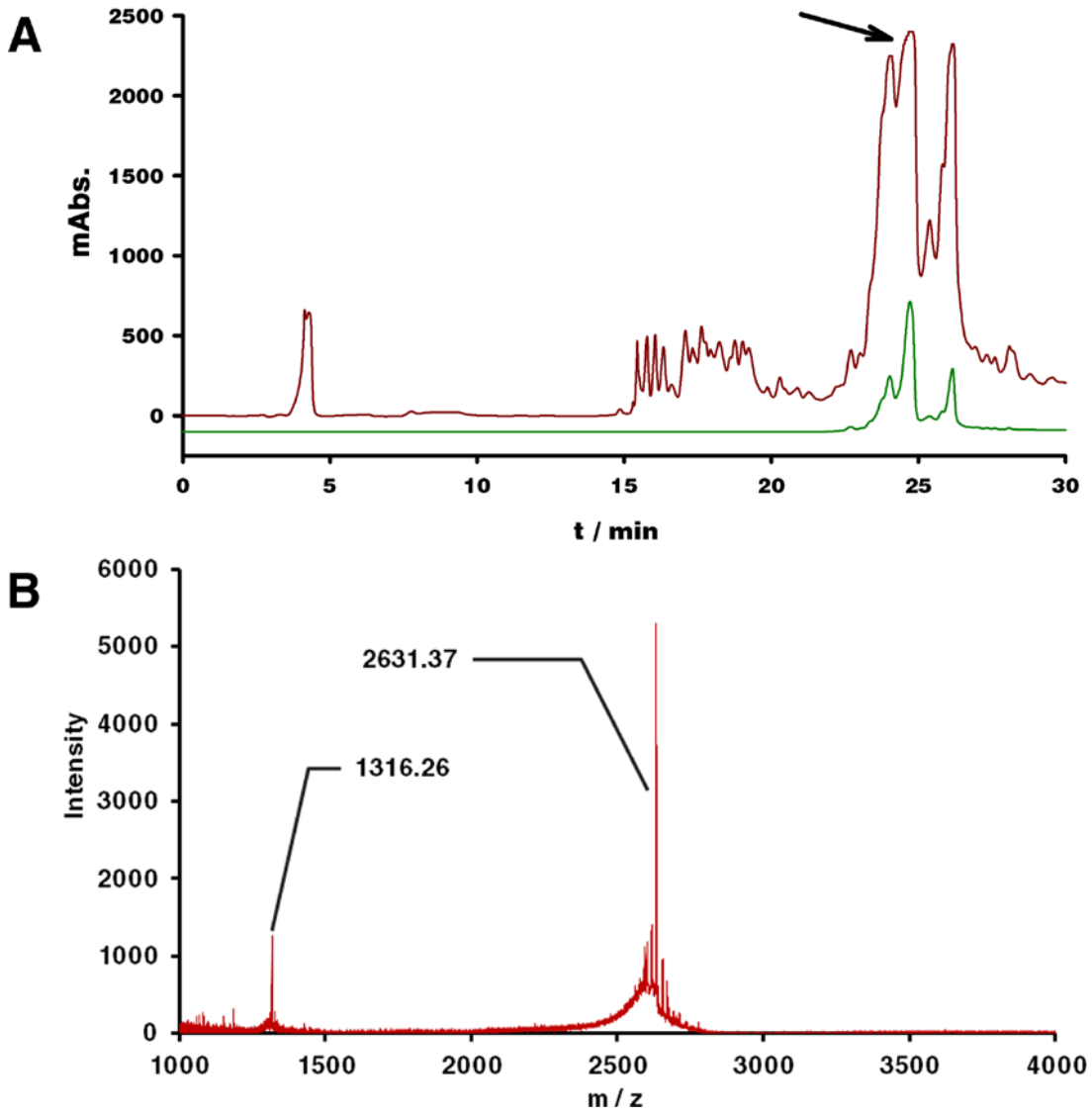

Figure 74: (A) HPLC chromatogram (red 210 nm, green 495 nm absorption) from the purification of the crude FAM-Rev$^{i, i+11}$ indicated with black arrow. (B) MALDI MS trace of FAM-Rev$^{i, i+11}$.

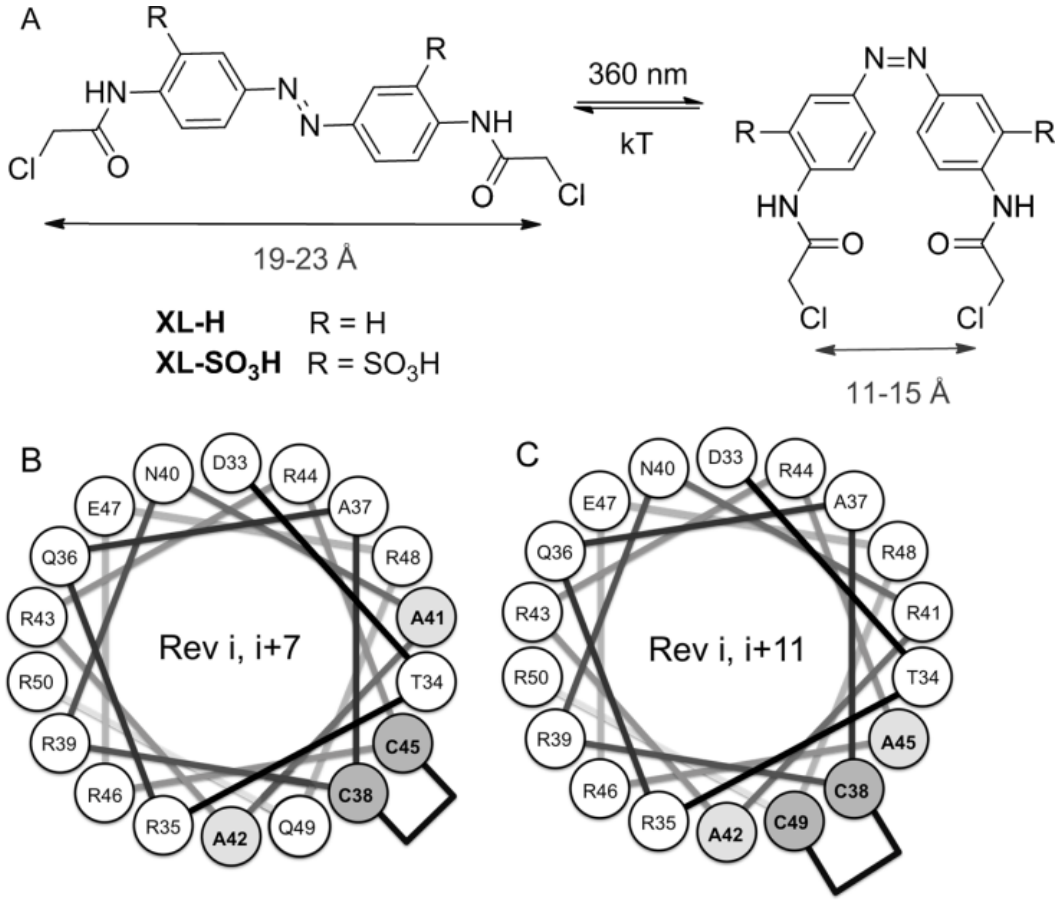

Figure 75: (A) Structure of azobenzene derived crosslinkers XL-H and XL-SO_3H and the interconversion between the *trans* and *cis* forms. Helical wheel representations of $Rev^{i, i+7}$ (B) and $Rev^{i,i+11}$ (C). Dark grey: cysteine residues introduced to allow cross-linking; light grey: residues replaced to avoid steric clash with the azobenzene cross-linker.

Crosslinking with the water soluble azobenzene crosslinker (XL-SO_3H) was achieved as it was described above and the same precipitation issues as for crosslinking biosynthesized Rev peptides was observed. After evaluation of binding affinities and molecular dynamics studies on the system were performed by Dr A. Ricci and Dr A. Brancale (*vide infra*)[262] it was decided to assess the effect of the less water soluble unsubstituted azobenzene crosslinker (XL-H) on peptide structure and affinity (Figure 75). A new, simpler and more efficient method of crosslinking was developed relaying on volatile THF rather than the more difficult to remove DMSO. Crosslinking with XL-H was accomplished by dissolving peptides at 1 mg/mL in water/THF mixture (1:1) buffered to *p*H 8.3 with Tris:HCl (50 mM) containing TCEP (2 mM). The same solvent was used to dissolve the XL-H crosslinker at concentration equivalent to a 10× molar excess over the peptide to be crosslinked. A single equivalent was added over the course of 2 hours, the remainder was added in further aliqouts over 3 hours of the solution volume. If precipitation occurred, additional water/TFA was added. After

overnight reaction, solutions were freeze-dried, resuspended in water containing 0.5% TFA, spun down and the supernatant was recovered. Crosslinked and labelled peptides were also synthesized bearing both crosslinker and fluorescein. All peptides were purified by HPLC techniques and their mass was verified by MALDI mass spectrometry (Figures 76–83, Table 7). The purity of peptides used in the subsequent FRET binding assay was confirmed by HPLC analysis (Figure 84–87).

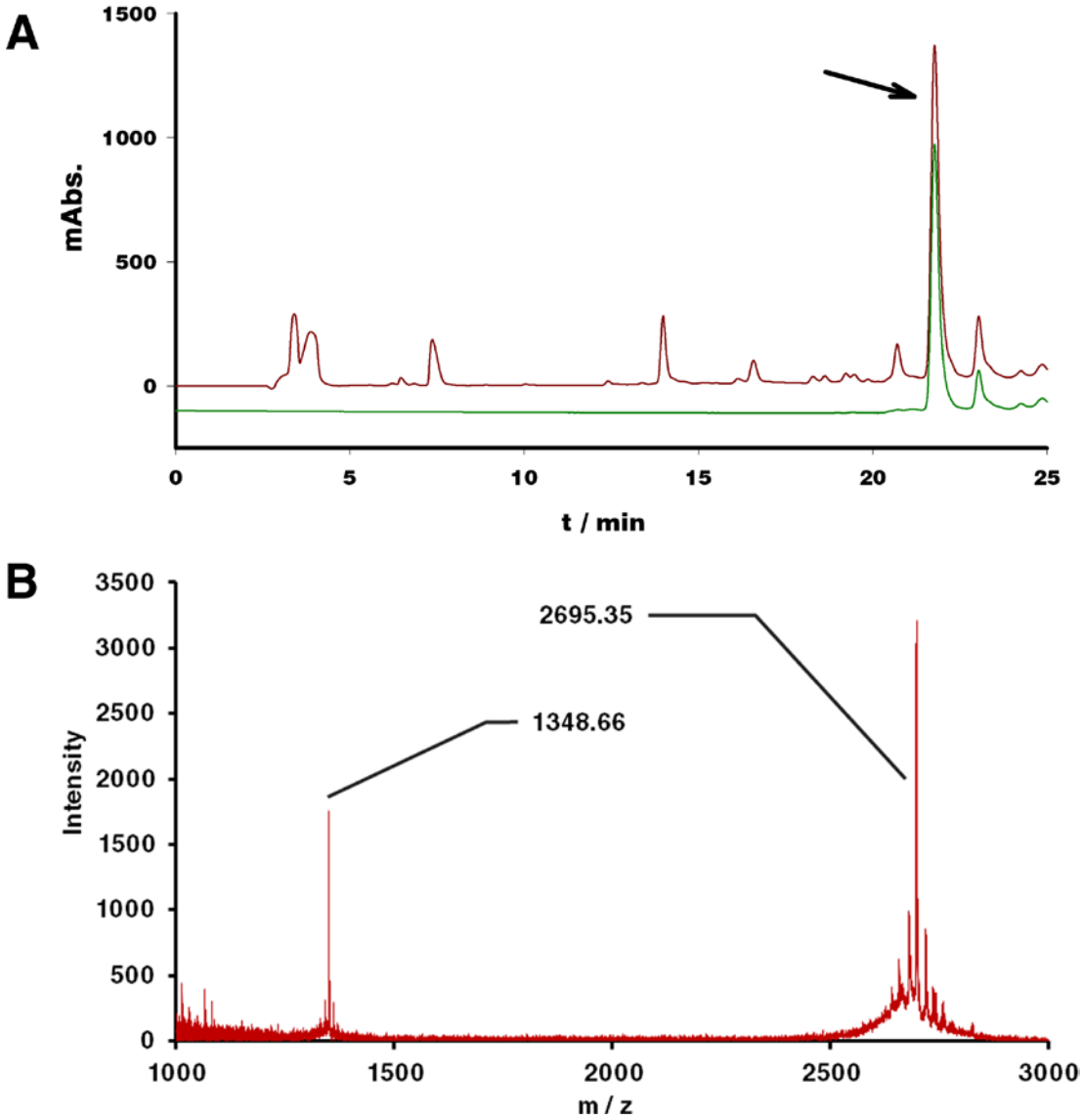

Figure 76: (A) HPLC chromatogram (red 210 nm, green 363 nm absorption) from the purification of crosslinking reaction of Rev$^{i, i+7}$-XL-SO$_3$H indicated with black arrow. (B) MALDI MS trace of Rev^{i+7}-XL-SO$_3$H.

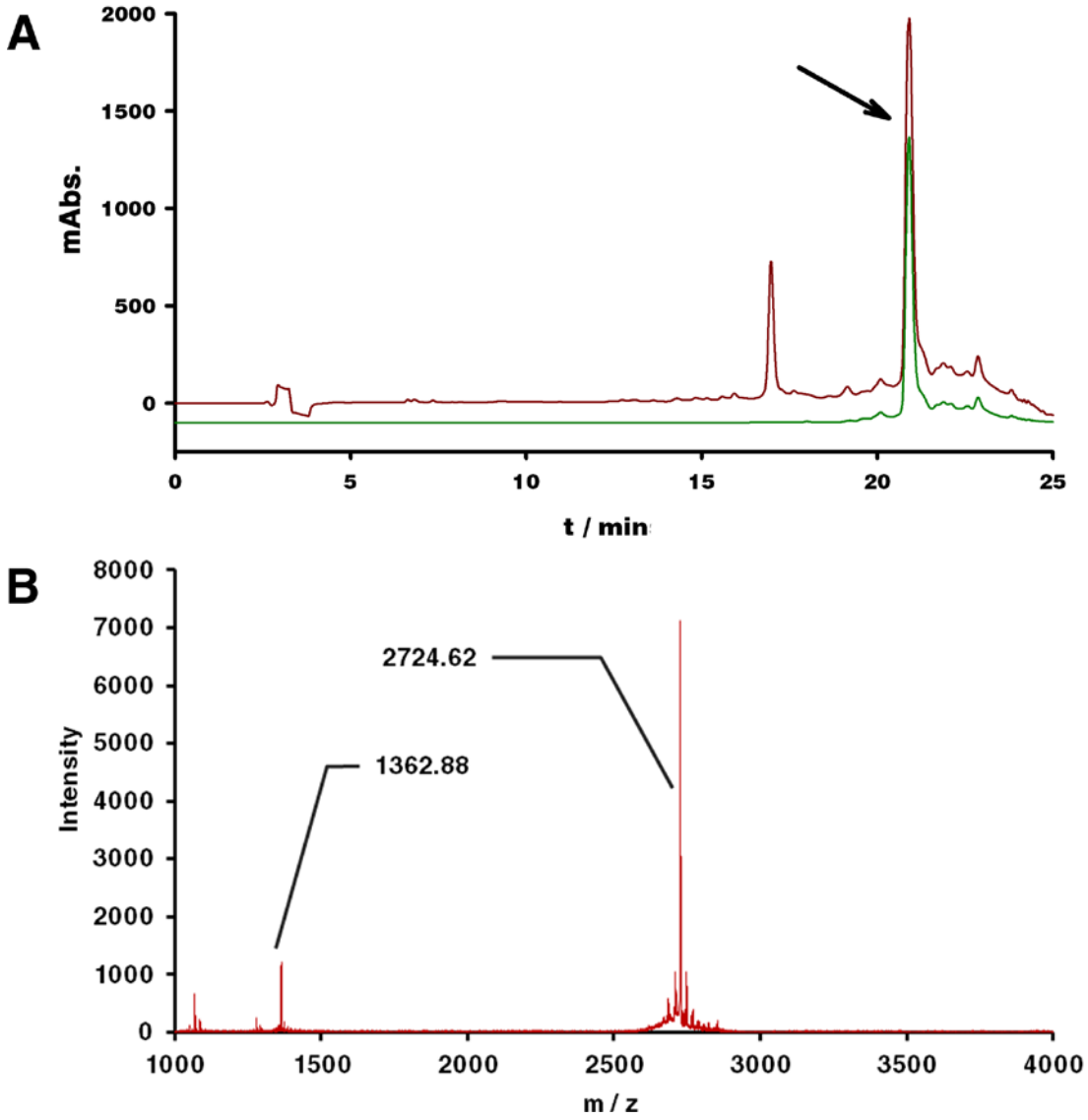

Figure 77: (A) HPLC chromatogram (red 210 nm, green 363 nm absorption) from the purification of crosslinking reaction of Rev$^{i, i+11}$-XL-SO_3H indicated with black arrow. (B) MALDI MS trace of Rev$^{i, i+11}$-XL-SO_3H.

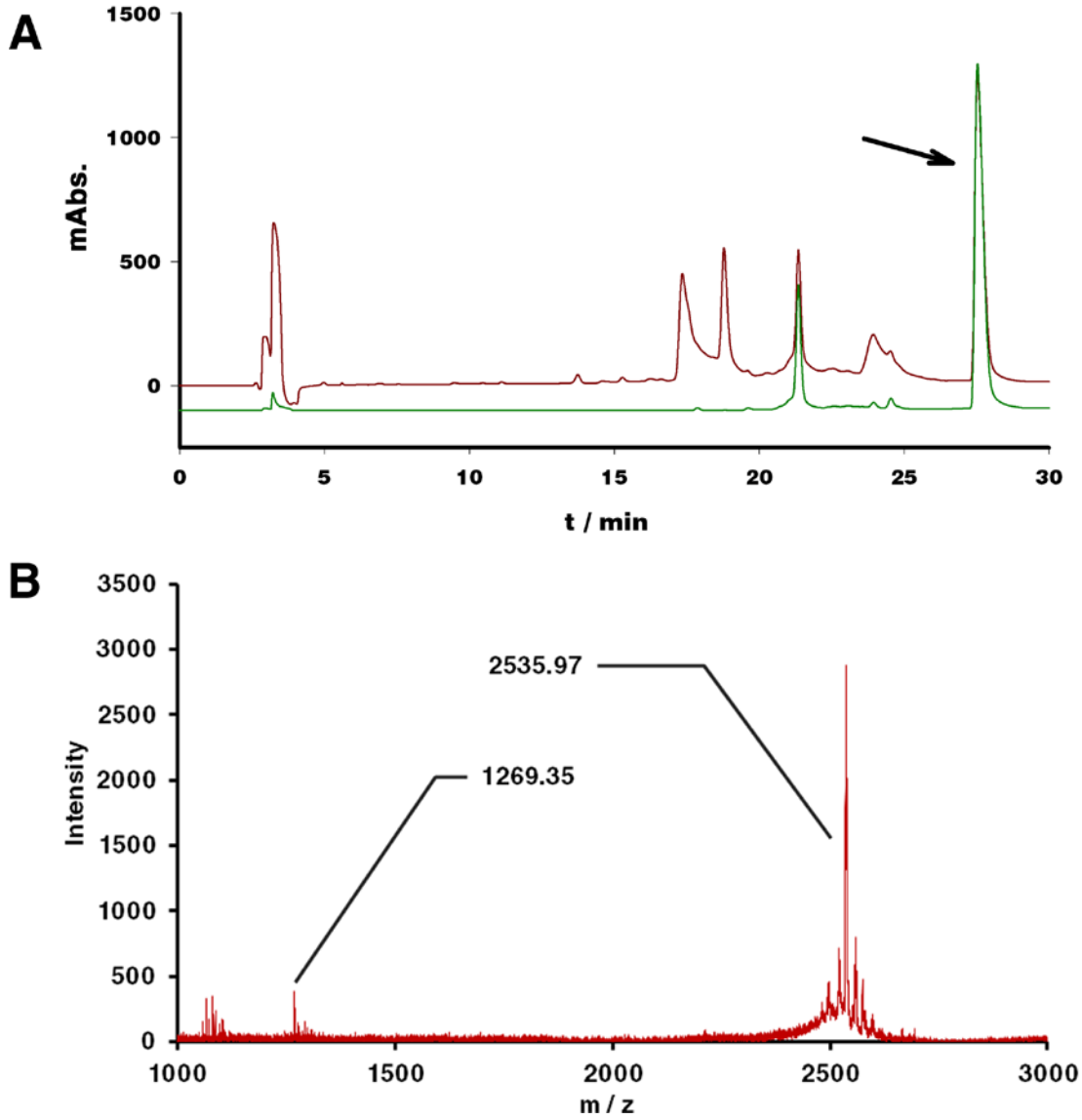

Figure 78: (A) HPLC chromatogram (red 210 nm, green 363 nm absorption) from the purification of crosslinking reaction of Rev$^{i, i+7}$-XL-H ^{indicated} with black arrow. (B) MALDI MS trace of Rev$^{i, i+7}$-XL-H.

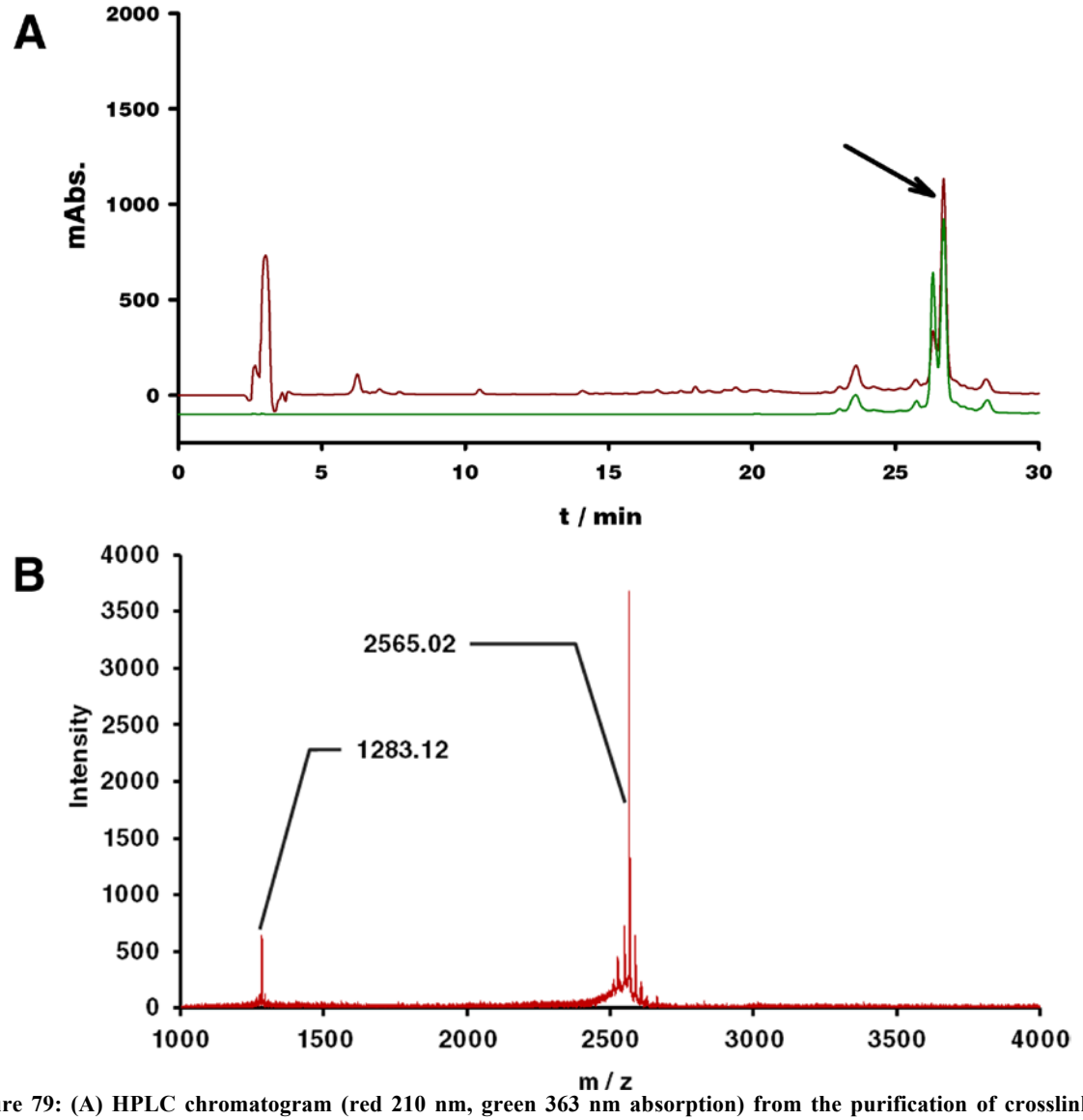

Figure 79: (A) HPLC chromatogram (red 210 nm, green 363 nm absorption) from the purification of crosslinking reaction of Rev$^{i, i+11}$-XL-H indicated with black arrow. (B) MALDI MS trace of Rev$^{i, i+11}$-XL-H.

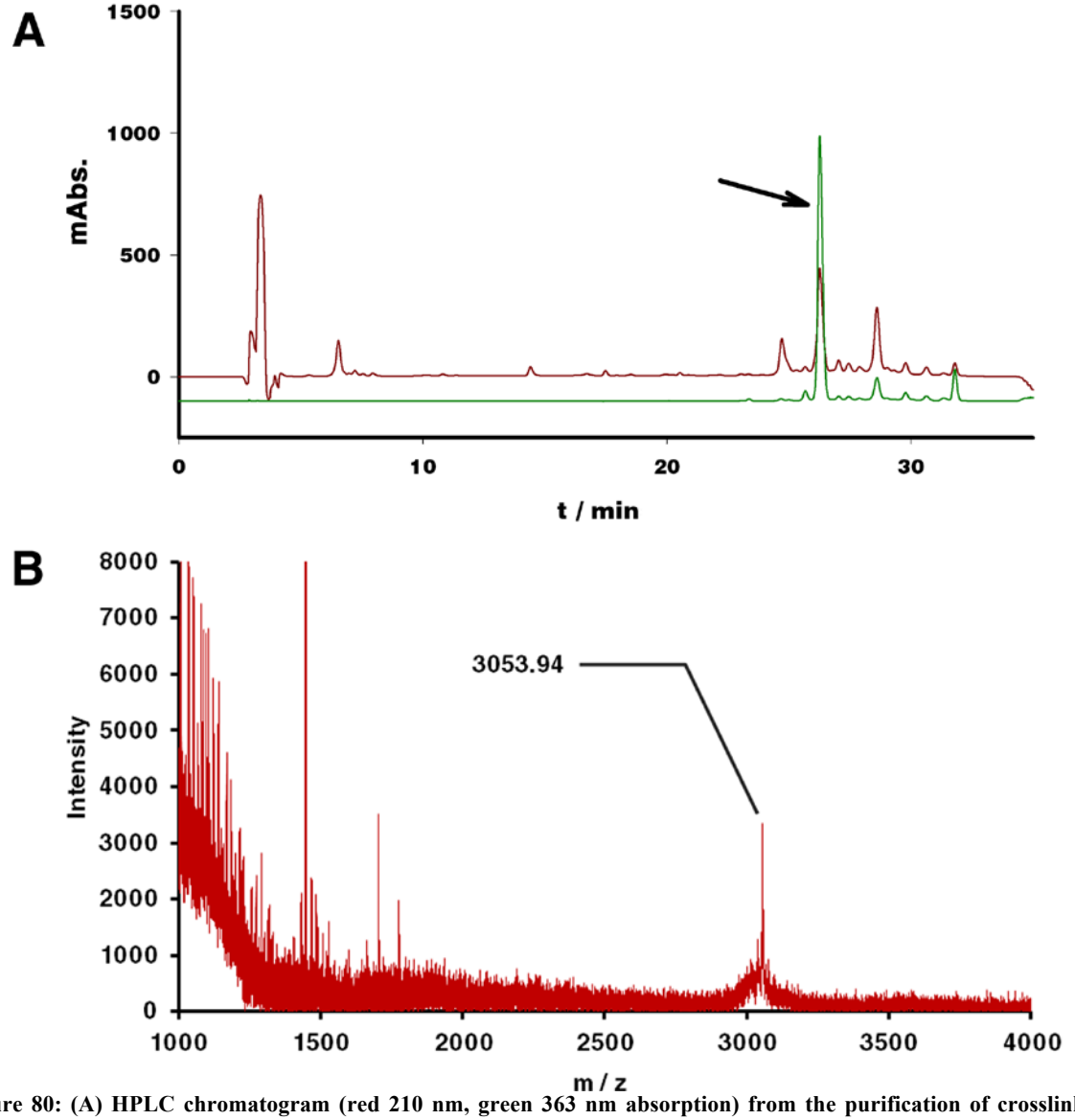

Figure 80: (A) HPLC chromatogram (red 210 nm, green 363 nm absorption) from the purification of crosslinking reaction of FAM-Rev$^{i, i+7}$-XL-SO_3H indicated with black arrow. (B) MALDI MS trace of FAM-Rev$^{i, i+7}$-XL-SO_3H.

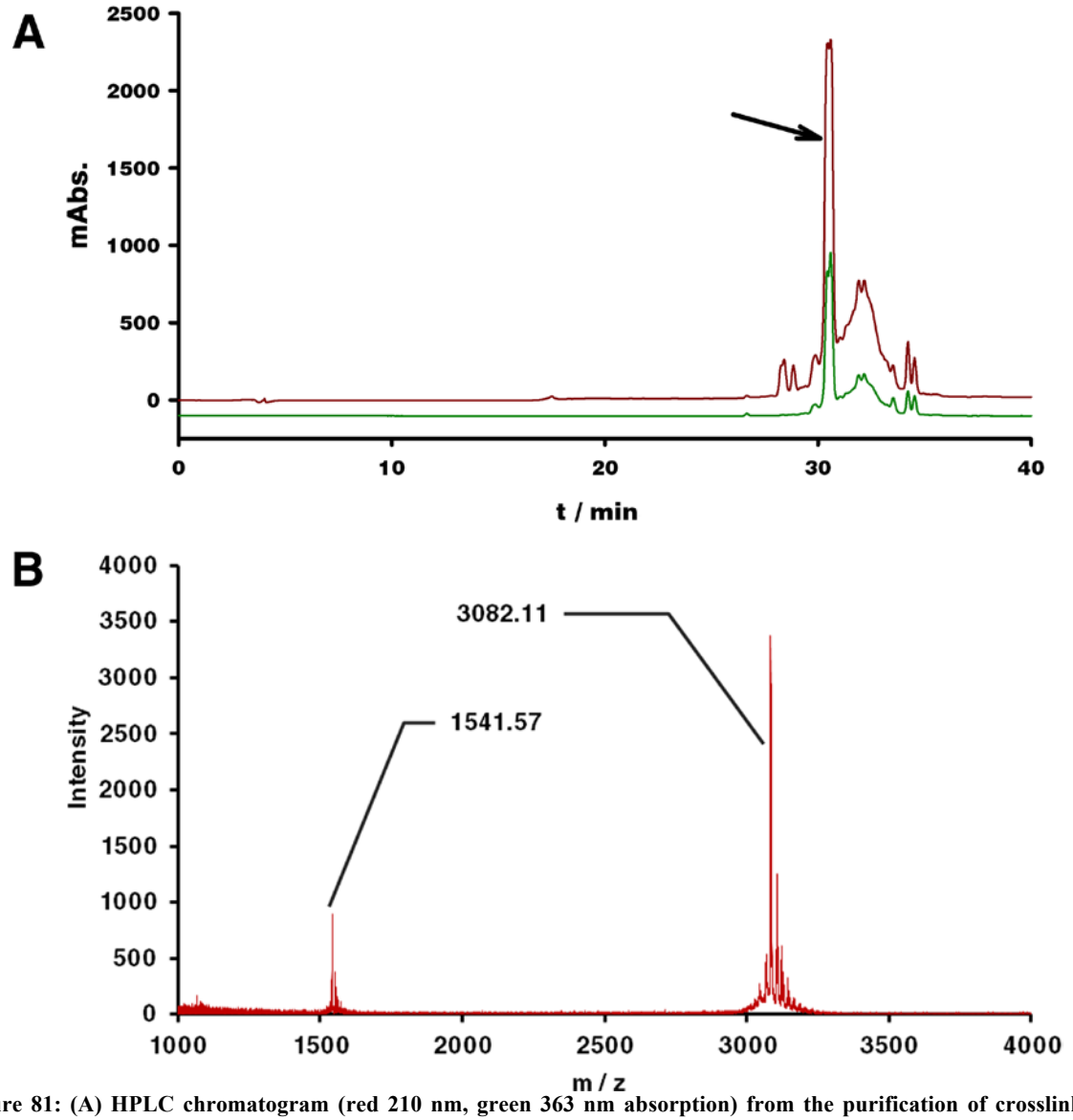

Figure 81: (A) HPLC chromatogram (red 210 nm, green 363 nm absorption) from the purification of crosslinking reaction of FAM-Rev$^{i, i+11}$-XL-SO_3H indicated with black arrow (B) MALDI MS trace of FAM-Rev$^{i, i+11}$XL-SO_3H.

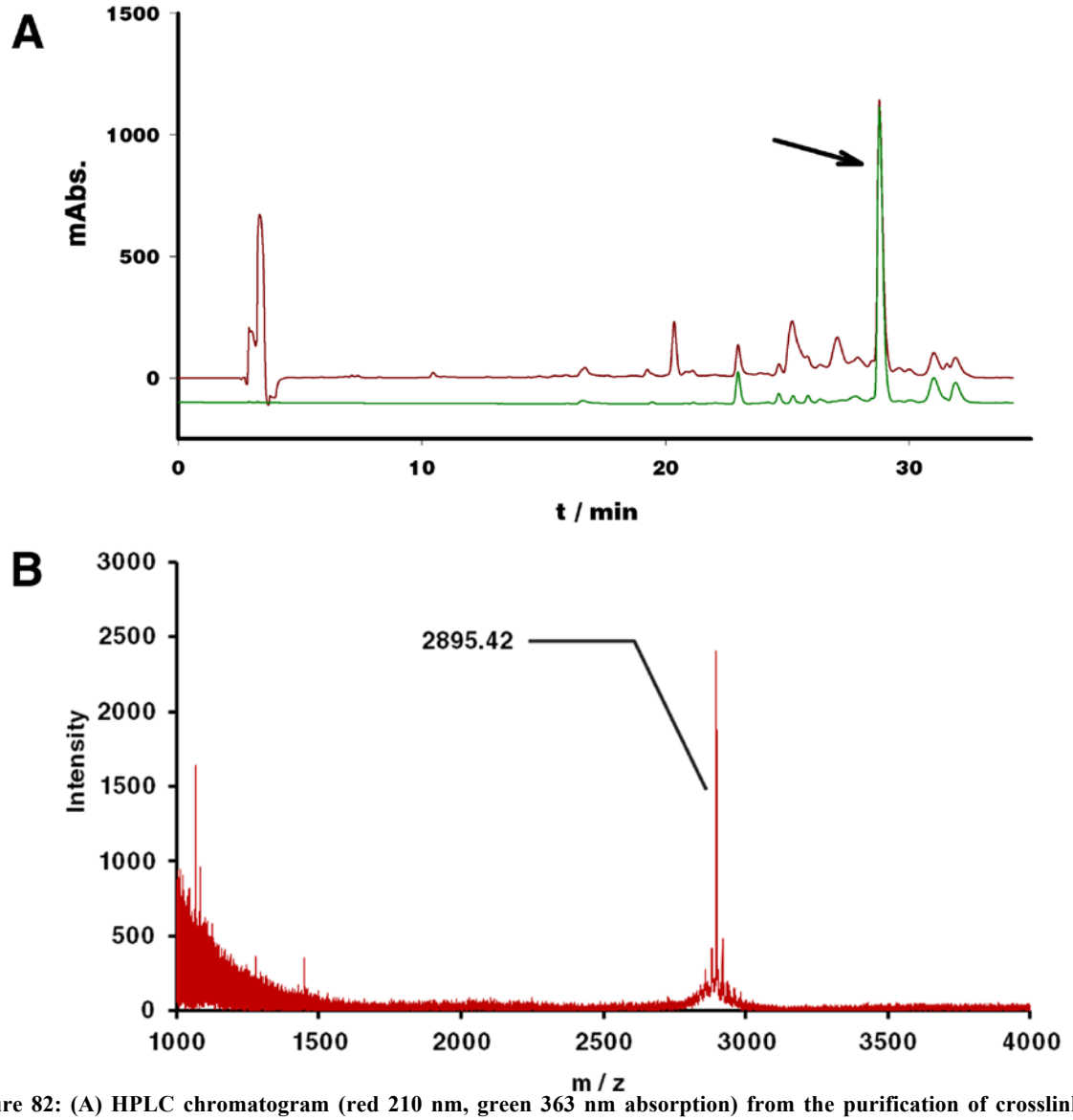

Figure 82: (A) HPLC chromatogram (red 210 nm, green 363 nm absorption) from the purification of crosslinking reaction of FAM-Rev$^{i, i+7}$-XL-H indicated with black arrow. (B) MALDI MS trace of FAM-Rev$^{i, i+7}$-XL-H.

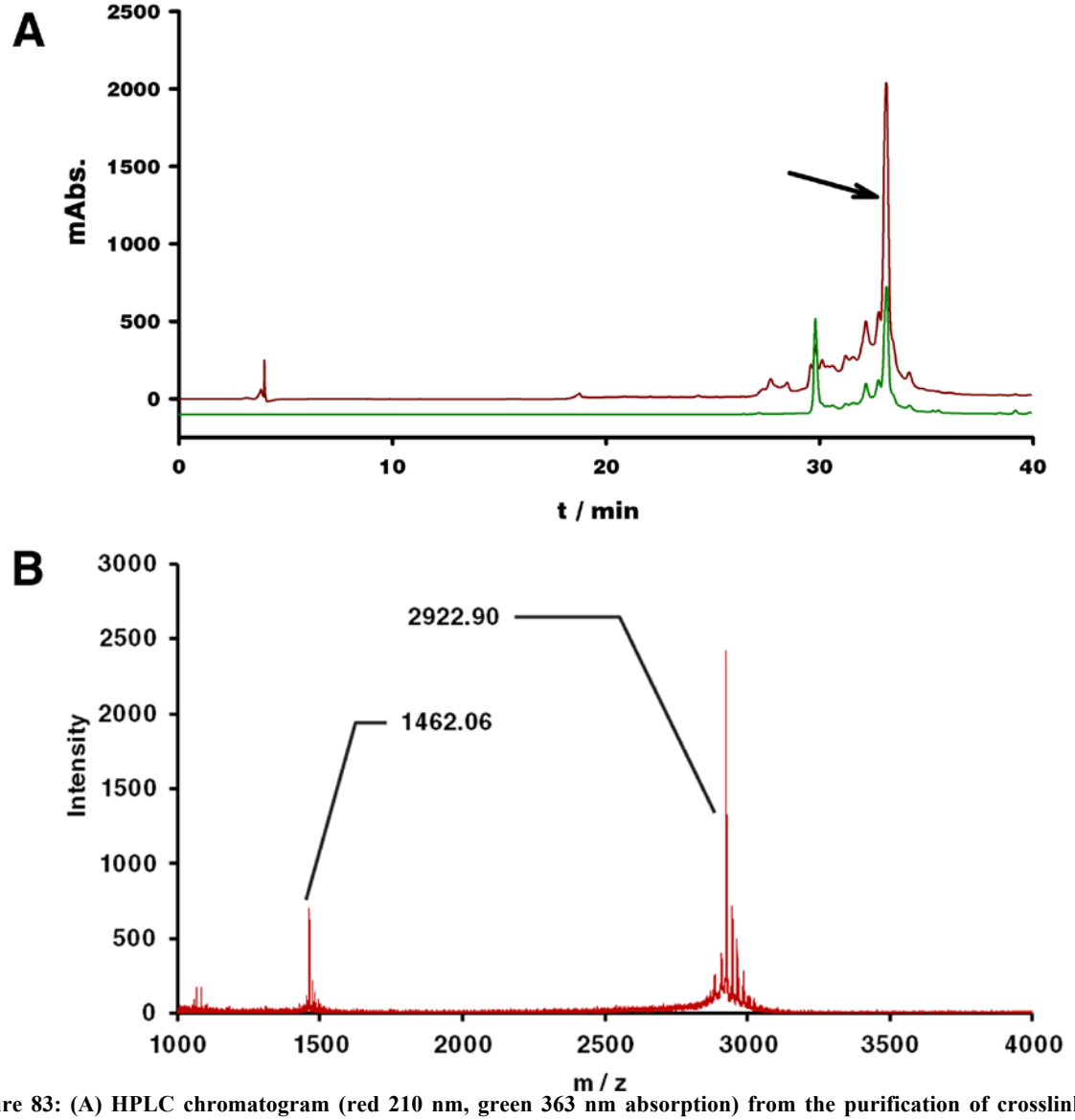

Figure 83: (A) HPLC chromatogram (red 210 nm, green 363 nm absorption) from the purification of crosslinking reaction of FAM-Rev$^{i, i+11}$-XL-H indicated with black arrow. (B) MALDI MS trace of FAM- Rev$^{i, i+11}$-XL-H.

Peptide	Calculated $[M+H]^+$	Observed $[M+H]^+$	Observed $[M+2H]^{2+}$
Rev^{wt}	2551.439	2551.31	N/A
FAM-Rev^{wt}	2909.487	2906.86	1454.15
$Rev^{i, i+7}$	2245.149[a]	2245.05	1123.77
$Rev^{i, i+11}$	2273.192[a]	2273.02	1137.13
FAM-$Rev^{i, i+7}$	2603.197[a]	2602.94	1302.05
FAM-$Rev^{i, i+11}$	2631.239[a]	2631.37	1316.26
$Rev^{i, i+7}$-XL-SO_3H	2697.159	2695.66	1348.66
$Rev^{i, i+7}$-XL-H	2537.245	2535.97	1269.35
$Rev^{i, i+11}$-XL-SO_3H	2725.201	2724.62	1362.88
$Rev^{i, i+11}$-XL-H	2565.288	2565.02	1283.12
FAM-$Rev^{i, i+7}$-XL-SO_3H	3055.207	3053.94	N/A
FAM-$Rev^{i, i+7}$-XL-H	2895.293	2895.42	N/A
FAM-$Rev^{i, i+11}$-XL-SO_3H	3083.249	3082.11	1541.57
FAM-$Rev^{i, i+11}$-XL-H	2923.336	2922.9	1462.06

Table 7: Observed and calculated masses of purified Rev peptides; [a] – calculated assuming reduced cysteines

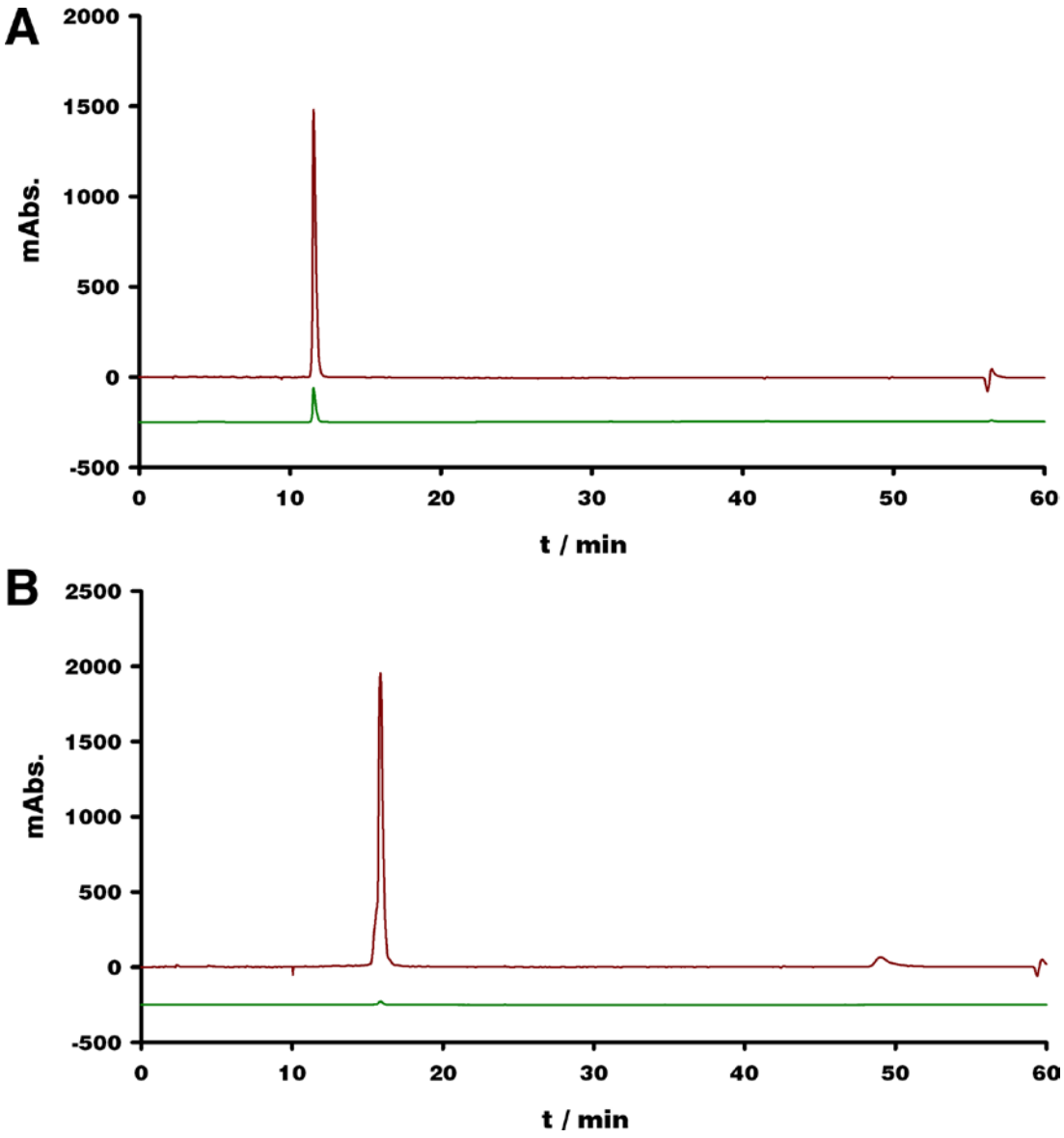

Figure 84: Analytical HPLC chromatograms of (A) Rev^{wt} and (B) FAM-Rev^{wt}; red 210 nm, green – 280 nm absorption for A; green- 495 nm absorption for B

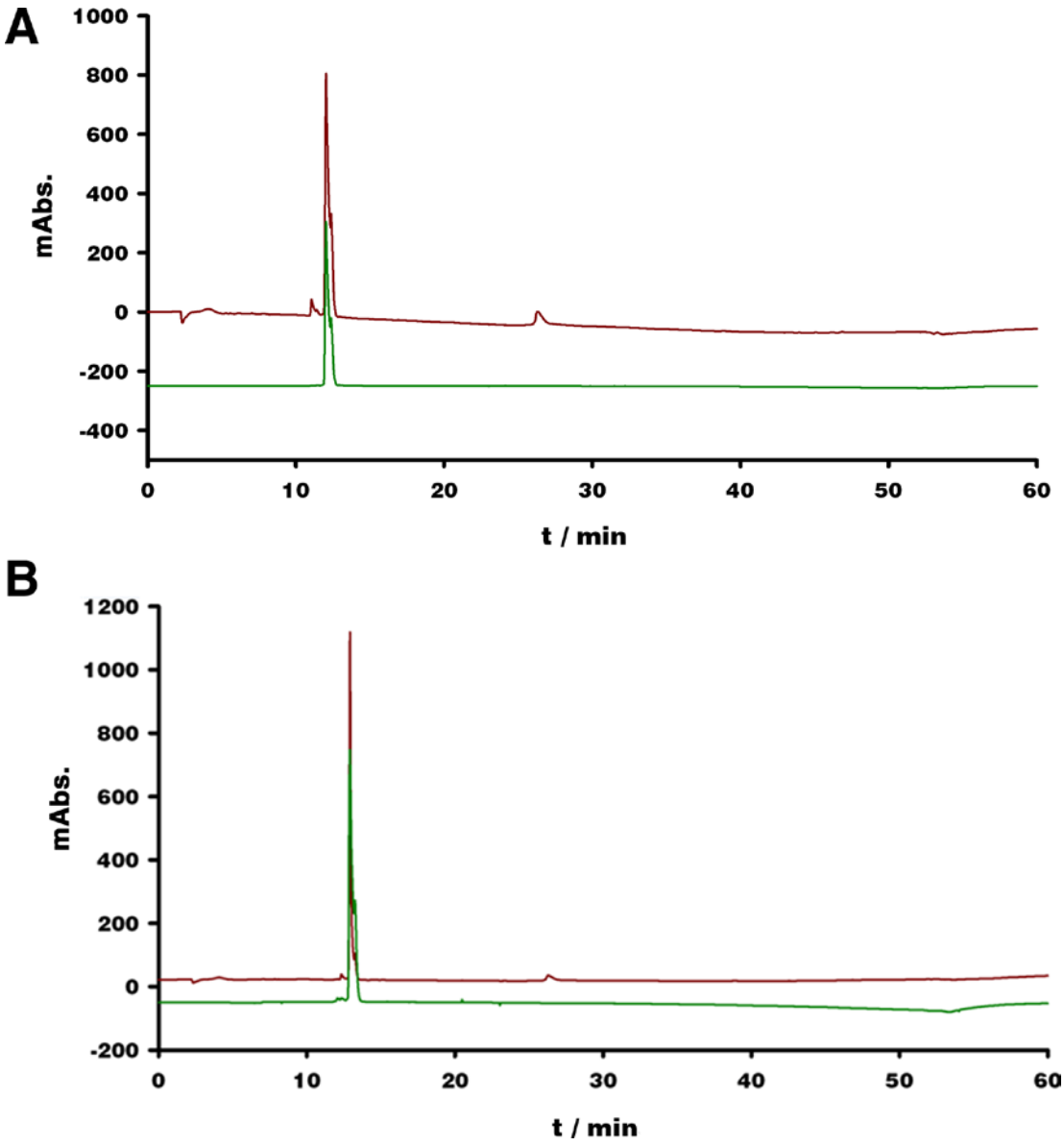

Figure 85: Analytical HPLC chromatograms of (A) Rev$^{i, i+7}$-XL-SO_3H and (B) Rev$^{i, i+7}$-XL-H; red 210 nm, green – 363 nm absorption.

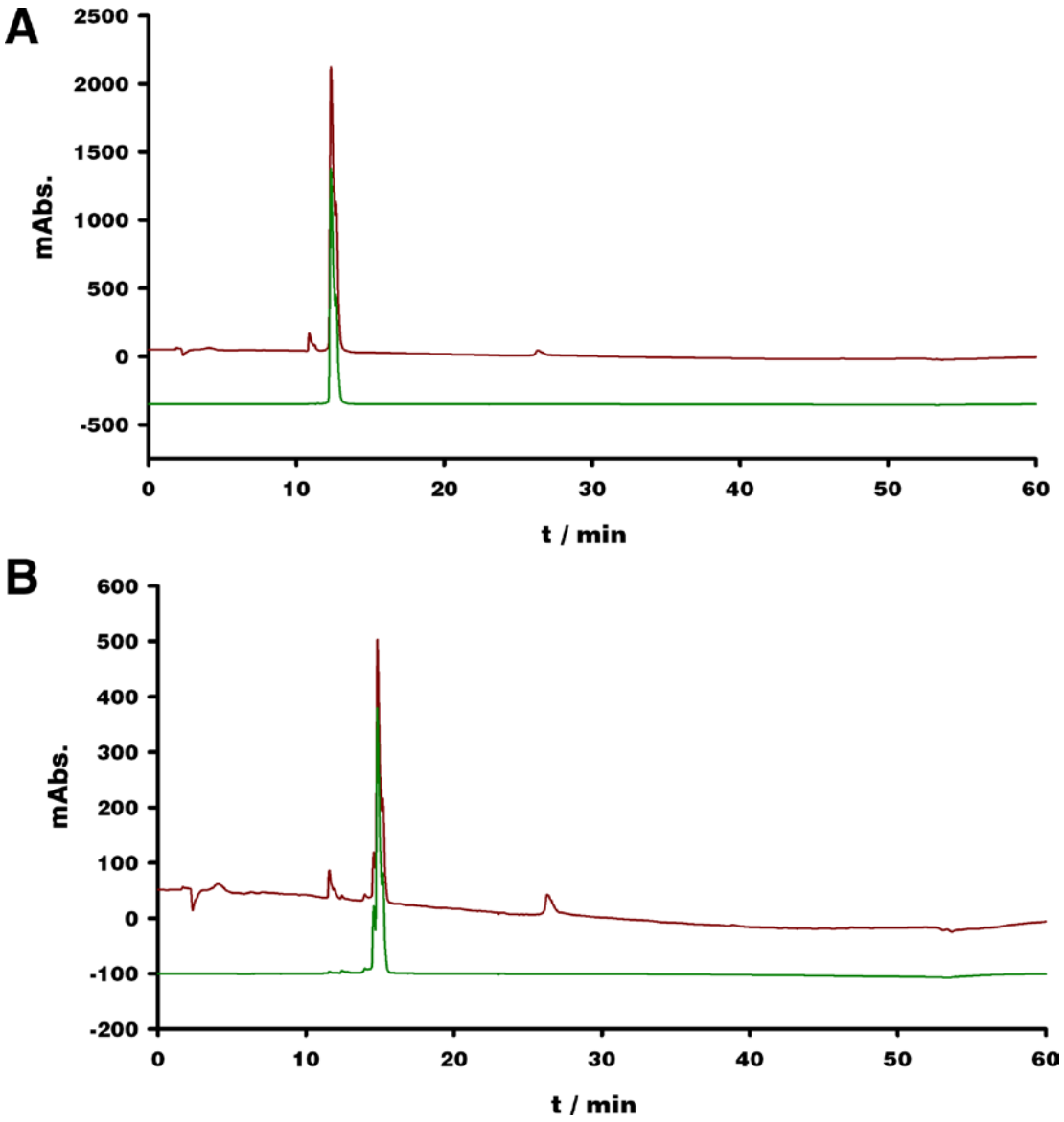

Figure 86: Analytical HPLC chromatogram of (A) Rev$^{i, i+11}$-XL-SO_3H and (B) Rev$^{i, i+11}$-XL-H; red 210 nm absorption; green – 363 nm absorption.

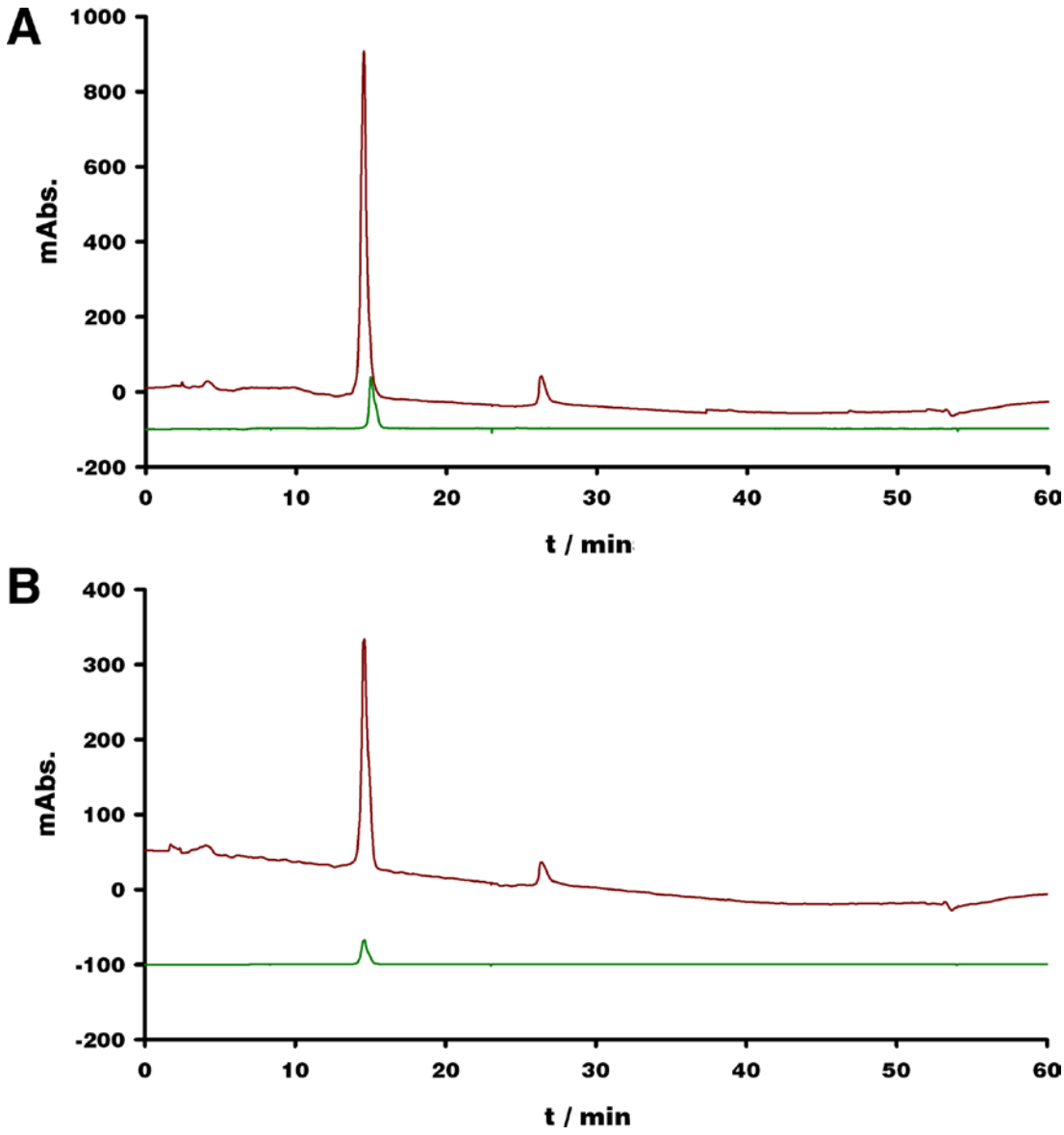

Figure 87: Analytical HPLC chromatograms of (A) FAM-Rev$^{i, i+7}$ and (B) FAM-Rev$^{i, i+11}$; red 210 nm, green 495 nm absorption.

2.1.3.2 CD and UV

In order to estimate activation energy and preexponential factors of the *cis-trans* relaxation, rate constants (k) for the *cis* to *trans* reversion were measured at different temperatures to create Arrhenius plots (Figure 88, 89). Differently to equivalent measurements for biosynthesized Rev peptides these were done four in parallel for each point and one was excluded. Measuremnts were perfomered on UV-VIS spectrometer by monitoring absorbtion at 365 nm as a function of time. The acquired data was fitted to Equation 1 to estimate first order relaxation constants (k). Interesting patterns arise in relation to both different spacings (*i, i*+7 vs. *i, i*+11) and different crosslinkers (XL-H vs. XL-SO_3H). At 15 °C both for *i, i*+7 and *i, i*+11 spacings XL-H peptides relax more quickly with half lives of 93 and 50 min respectively versus 299 and 135 min for XL-SO_3H (Table 8) nearly identical values were reported for biosynthesized bRev-XL-SO_3H peptides 217 and 123 min. It has been reported previously that peptides crosslinked with XL-H relax quicker than those with XL-SO_3H.[161] However, the activation energy values are similar between peptides with the same spacing irrespective of the crosslinker; for *i, i*+7: 88.4 and 93.4 (83.0 for bRev$^{i, i+7}$-XL-SO_3H); for *i, i*+11: 77.6 and 71.5 (78.0 for bRev$^{i, i+11}$-XL-SO_3H). Moreover the values of preexponential factors for both biosynthesized (for bRev$^{i, i+7}$ lnA is 25.6 and for bRev$^{i, i+11}$ lnA is 23.9) and Fmoc synthesized peptides (for bRev$^{i, i+7}$-XL-SO_3H lnA is 27.1 and for bRev$^{i, i+11}$-XL-SO_3H lnA is 23.0) are similar and keep the same described here tendencies (Table 6 and 8).

Fascinatingly, the shorter half lives at 15° C for XL-H peptides are achieved in an opposite way for each spacing. In the case of *i, i*+7 the activation energy is actually higher for XL-H, but it is compensated by higher preexponential factor values. For *i, i*+11 preexponential factors are lower for XL-H (slower kinetics) however it is compensated by lower activation energy.

$$k = \frac{k_B T}{h} e^{-\frac{\Delta H^{\ddagger}}{R}} e^{\frac{\Delta S^{\ddagger}}{RT}}$$

Equation 4: Eyring –Polanyi equation; where k is a rate constant; k_B is Boltzmann constant; h is Planck constant; $\Delta H^{\ddagger}$ is activation enthalpy; $\Delta S^{\ddagger}$ is activation entropy; E_a is an activation energy; R is the gas constant and T is temperature.

The relation between parameters of Arrhenius equation (equation 1) derived from gas collision theory and Eyring–Polanyi equation (equation 5) derived from transition state theory is following: preexponential factor values correlate to values of difference in entropy between the *cis*-peptide and the transition state ($\ln A \sim \Delta S^{\ddagger}$) according to equation 6 and activation energy to values of difference in enthalpy between the *cis*-peptide and the transition state ($E_a \sim \Delta H^{\ddagger}$) according to equation 7.

$$A = e^{-\Delta n} \frac{k_B T}{h} e^{\frac{\Delta S^{\ddagger}}{R}}$$

Equation 5: The relationship between activation entropy and preexponential factor

$$\Delta H^{\ddagger} = E_a + RT$$

Equation 6: The relationship between activation enthalpy and activation energy

Higher values of $\Delta S^{\ddagger}$ for *i, i*+7 could be explained by the increasing disorder of the peptide along the relaxation coordinate (*vide* CD results). The opposite is true for the *i, i*+11. Lower values of entropy reflect the increasing helicity (order) along the relaxation coordinate.

It is more complex to fully explain the differences in $\Delta H^{\ddagger}$ between the two cysteine spacings, because it not trivial to describe the exact geometry and rationalize the enthalpy difference between the transitions states of the two spacings. Nevertheless if we assume similar enthalpy of the transition state for *i*, *i*+7 and *i*, *i*+11 the following partial explanation can be provided: higher values of $\Delta H^{\ddagger}$ for *i*, *i*+7 reflect relatively low enthalpy of the peptide in the *cis*-state. The *cis* configuration "fits" *i*, *i*+7 spacing more, is less strained and therefore has lower enthalpy, which in effect increases $\Delta H^{\ddagger}$ (activation energy). Again, opposite is true for *i, i*+11 spacings were more strain is caused in the *cis* configuration as crosslinked peptides with this spacing are designed to be more strained in the *cis* configuration than *trans*. Overall peptides with *i, i*+7 spacings relax more slowly than *i, i*+11 spacings with the respective crosslinker.

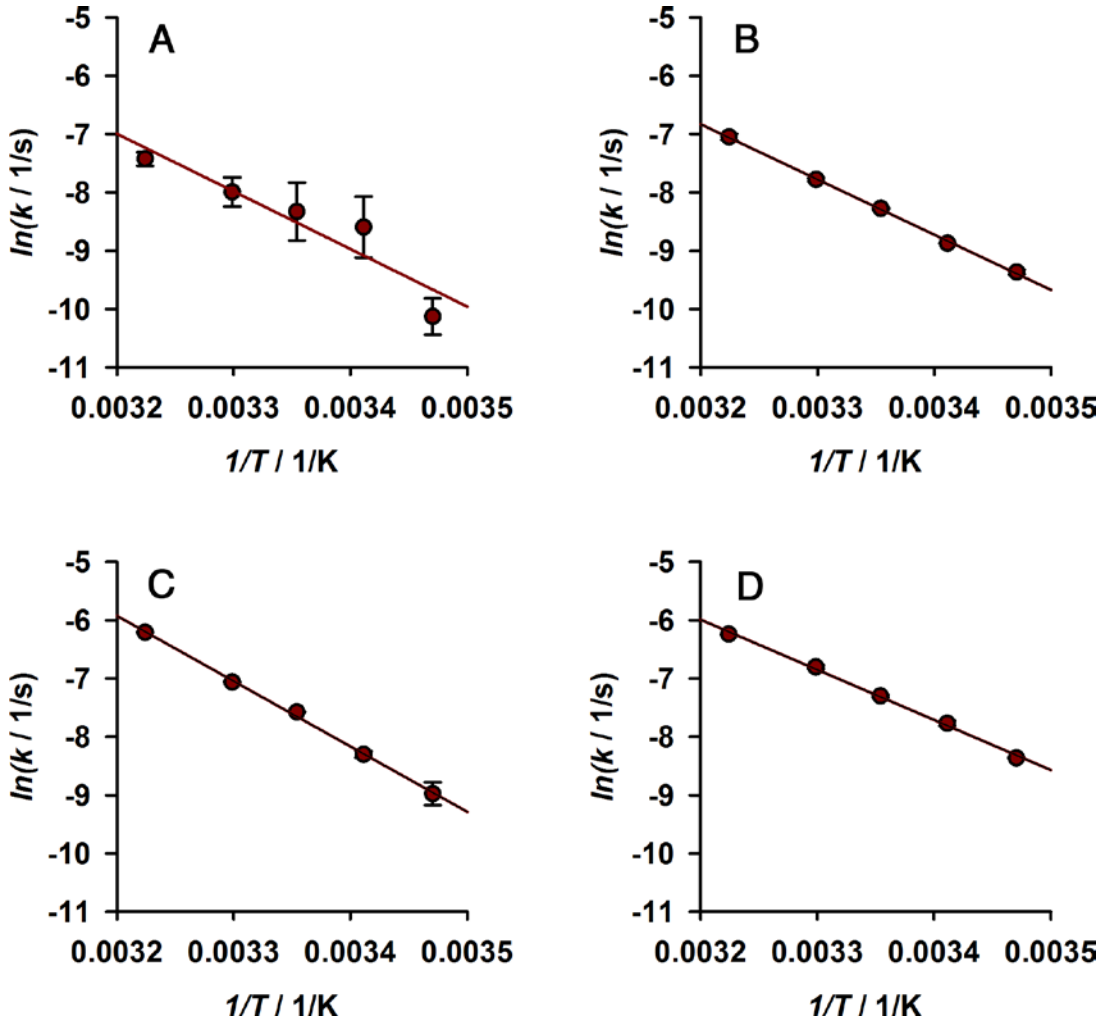

Figure 88: Arrhenius plots of UV_{363} relaxation after UV_{363} irradiation for (A) $Rev^{i, i+7}$-XL-SO_3H; (B) $Rev^{i, i+11}$-XL-SO_3H ; (C) $Rev^{i, i+7}$-XL-H; (D) $Rev^{i, i+11}$-XL-H.

Peptide	Half Life at 15 °C [min]	Activation Energy [kJ/mol]	lnA
$Rev^{i, i+7}$-XL-SO_3H	299 ±98.3	88.4±1.70	27.1±0.71
$Rev^{i, i+7}$-XL-H	93 ±16.9	93.4±4.90	30.0±1.94
$Rev^{i, i+11}$-XL-SO_3H	135 ±5.4	77.6±2.42	23.0±0.98
$Rev^{i, i+11}$-XL-H	50 ±0.7	71.5±0.46	21.5±0.19

Table 8: Kinetic parameters of UV_{363} relaxation after UV_{363} irradiation for all crosslinked Rev peptides.

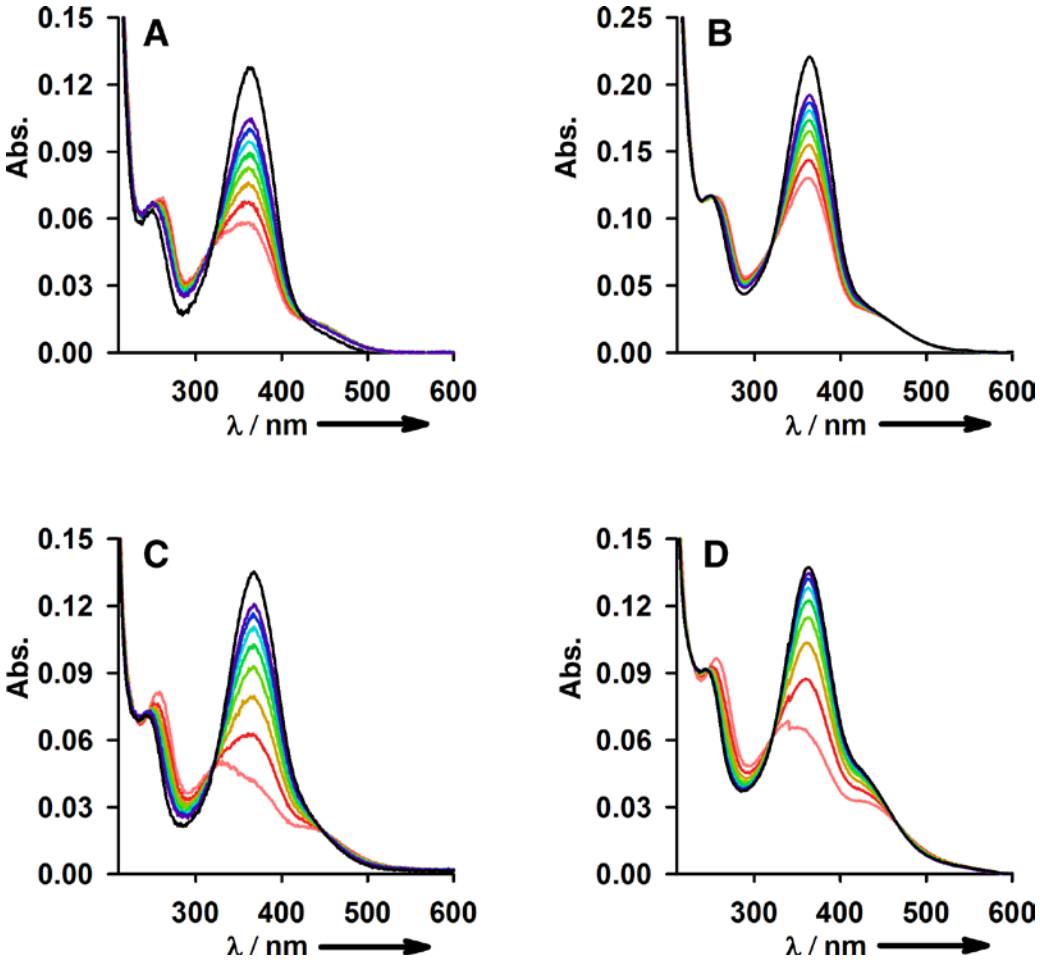

Figure 89: Absorption spectra (210 – 600 nm) of UV_{363} relaxation after UV_{363} irradiation taken at time 0 and then every 30 min for total 240 min for (A) Rev$^{i, i+7}$-XL-SO_3H; (B) Rev$^{i, i+11}$-XL-SO_3H; (C) Rev$^{i, i+7}$-XL-H; (D) Rev$^{i, i+11}$-XL-H. Dark adapted states are shown in black.

Both peptides with crosslinker XL-H and XL-SO_3H confer structural photocontrol, exhibiting CD spectra typical of an α-helix with minima at 208 and 222 nm in dark state for Rev$^{i,i+11}$XL and Rev$^{i,i+7}$XL after UV irradiation for (Figure 67 and 68). The expected reduction in helical character when Rev *i, i*+11 is irradiated and for Rev *i, i*+7 in the dark state is observed as a decrease in 222 nm signal intensity and a peak shift from 208 nm towards 202 nm; characteristic of short unstructured peptides. Both the absolute values and relative changes of helicity between XL-H and XL-SO_3H peptides are almost identical. This similarity suggests that both crosslinkers impose similar level of structure photocontrol. Compared to parent uncrosslinked peptides attachment of the crosslinker in *i, i*+7 causes further reduction of the α-helical character and irradiation restores the level of helicity present in the parent peptide (Figure 67). For the *i, i*+11 spacing attachment of the crosslinker increases helicity considerably and irradiation decreases it almost to the level of the parent uncrosslinked peptide.

Interestingly, for both crosslinkers and cysteine spacings, both relative and absolute helicity is somewhat higher for the Fmoc synthesized peptides compared to the biosynthesized ones (Figure 65 vs. 90). This suggests that either homoserine has a negative impact on the helical character of Rev peptides or that the C-terminal amide has a helix stabilising effect on Fmoc synthesized Rev peptides. As for the biosynthesized crosslinked peptides, CD signals corresponding to the crosslinker are visible and follow the same pattern of response to irradiation for Rev$^{i, i+11}$ XL with the sole exception that 265 nm signal is not visible for Rev$^{i, i+11}$ XL-H (Figure 65, 90, 91). However for Rev$^{i, i+7}$ XL-H the dark adapted state exhibits almost an exactly inverse CD signal for the crosslinker compared to irradiated state (Figure 91). This hints that although the peptide bond transitions CD regions are approximately identical the peptide might adopt very different conformations depending on the nature of the substituent present on the azobenzene croslinker as judged by the reversal of the 265, 360 and 440 nm CD signals present only when XL-H is used rather than XL-SO_3H. It is important to note that CD measures the average helical content, and may result from an equilibrium distribution of helical and non- or less-helical conformation, or peptide that is part strongly and constantly helical but, for instance, has disordered C and N-termini. For uncrosslinked peptides CD signals were only observed in the peptide bond region, the attached fluorescein is generally CD silent (Figure 90) except for FAM-Rev$^{i, i+11}$ where a small minimum around 510 nm and FAM-Rev^{wt} where an even smaller maximum around 500 nm are present (Figure 90); this feature of these two CD spectra might suggest that fluorescein might interact with these peptides.

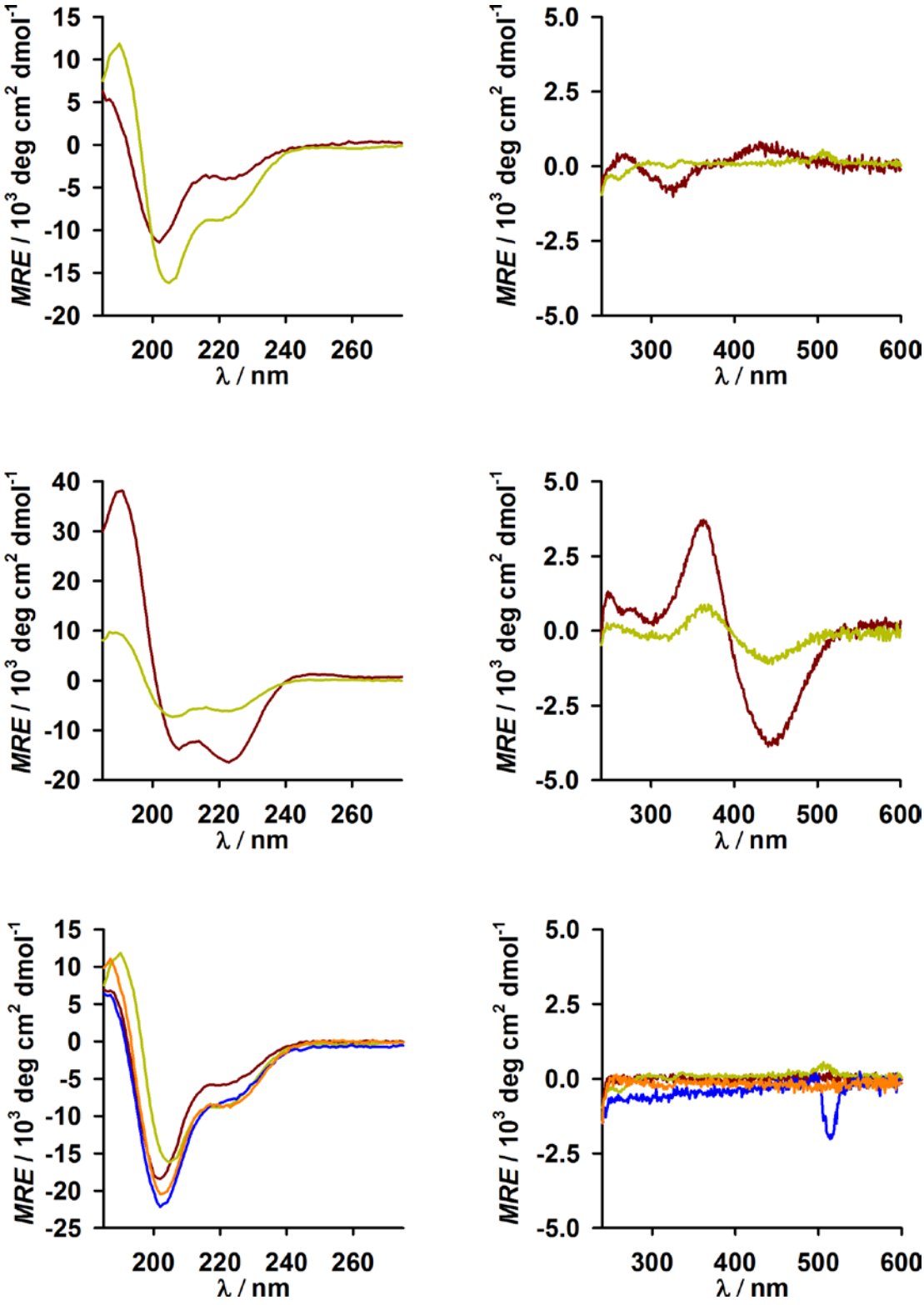

Figure 90: CD spectroscopy results, dark adapted (dark red) and after 3 min of UV_{363} irradiation (yellow-green) for $Rev^{i, i+7}$-XL-SO_3H (A and B); $Rev^{i, i+11}$-XL-SO_3H (C and D). Collection of CD plots for peptides not containing a crosslinker (E and F), Rev^{wt} (dark red); FAM-Rev^{wt} (yellow-green); FAM-$Rev^{i, i+7}$ (orange); FAM-$Rev^{i, i+11}$ (blue).

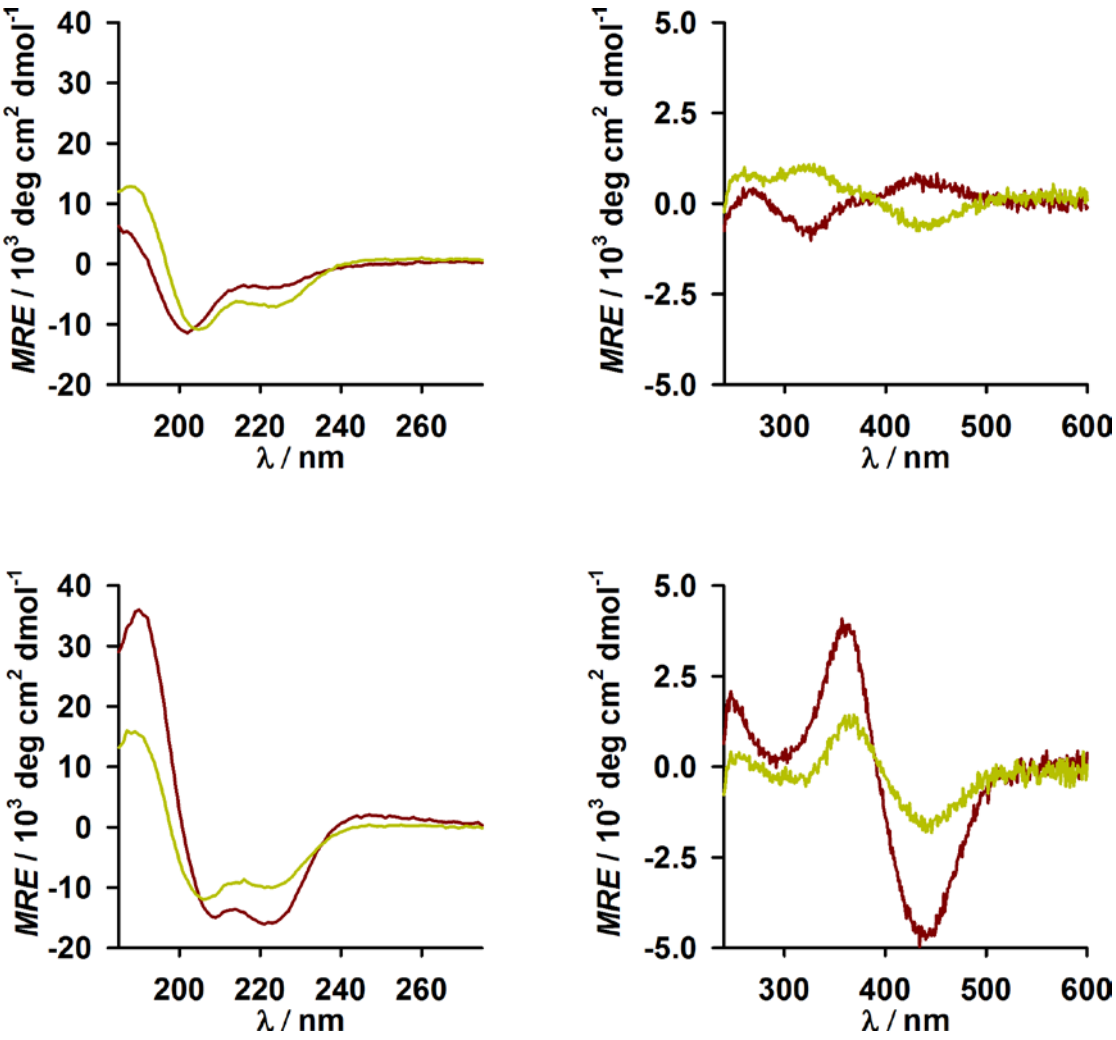

Figure 91: CD spectroscopy results, dark adapted (dark red) and after 3 min of UV_{363} irradiation (yellow-green) for $Rev^{i, i+11}$-XL-H (A and B); $Rev^{i, i+11}$-XL-H (C and D).

2.1.3.3 Rev/RRE FRET binding studies

A FRET binding assay was developed in order to measure the binding affinities between Rev peptides and RRE RNA. Peptides were labelled with fluorescein at their N-terminus and Cy3 labelled RRE was titrated in. The emission wavelength of the fluorescein matches the excitation wavelength of Cy3, so a decrease in fluorescein signal is observed as the increasing concentration of Cy3-RRE results in binding and intermolecular resonance energy transfer rather than fluorescence takes place. At first a direct binding assay was tested using doubly modified Rev peptides (crosslinker + FAM), unfortunately such peptides were poorly soluble and tended to form aggregates and precipitate. Interestingly, although they were soluble during the purification conditions, once freeze-dried they could not be redissolved in the same conditions. This suggests either an irreversible aggregation or very

slow redissolution process. This type of behaviour has been observed previously for some peptides (T. Fricke, personal communication) containing an octaarginine sequence. Also, a direct interaction with FAM and an anzobenzene crosslinker has been reported previously. Intra or intermolecular interactions between negatively charged crosslinker or fluorescein moieties and positively charged arginine sidechains or aromatic stacking between fluorescein and the crosslinker may cause such behaviour. Therefore an indirect, displacement assay was developed whereby FAM-Rev^{wt} bound to Cy3-RRE was titrated with crosslinked peptides without a fluorescent dye and the recovery of fluorescein signal as the wild type peptide was displaced from a complex with RRE RNA was observed.

Initial results indicated that the modifications introduced to allow the attachment of a crosslinker had a deleterious effect on the affinity of fluorescently-labelled $Rev^{i,i+7}$ and $Rev^{i,i+11}$ peptides to RRE RNA compared to the wild type sequence (Figure 92; Table 9). Additionally, mutation of Arg 38 ($Rev^{i,\ i+7}$ and $Rev^{i,\ i+11}$), Arg 41 ($Rev^{i,\ i+11}$) and Arg 42 ($Rev^{i,\ i+7}$ and $Rev^{i,\ i+11}$) to Ala also might have decreased the affinity. The CD spectra of all three peptides are similar suggesting that the helical content does not determine their affinity. Indeed, Rev^{wt} appeared to have the least helical character but was the best binder.

A saturable increase in fluorescence was observed when Rev^{wt} was released from the complex and a dissociation constant of 247 ± 19 nM could be estimated for the RRE complex of $Rev^{i,i+11}$-XL-SO_3H from the apparent IC_{50}, which was only a modest increase over the uncrosslinked peptide (Table 9 and 10). $Rev^{i,i+11}$-XL-SO_3H responds to irradiation and the affinity drops to 418 ± 28 nM. This change directly correlates to the decrese of an α-helical character as judged by CD spectroscopy. No RRE binding was observed for $Rev^{i,i+7}$-XL-SO_3H in either the dark or the irradiated state for concentrations of up to 1 μM (Figure 92). Simple inspection of superpositions of $Rev^{i,i+11}$-XL-SO_3H and Rev^{wt} with the published structure of the RRE complex did not suggest any obvious steric reasons for the low binding affinity. Also, literature precedent contains several bulkier custom peptides that have relatively good affinities to RRE; the Rev binding epitope transplanted to zinc finger protein had K_d of 330 ± 40 nM (vs. 330 ± 20 nM for the wild type under the assay conditions) and β sheet peptide 2 nM.[94,102] Additonaly, electrostatic repulsion between the crosslinked peptides and the negatively charged phosphate backbone seemed at first a likely cause for the low binding affinities of $Rev^{i,i+11}$-XL-SO_3H and $Rev^{i,i+7}$-XL-SO_3H but inspection of the overlaid structures suggested that distances between sulfonates and phosphate groups were too grate for a significant charge-charge repulsion. There could have been a more subtle effect at play

and to investigate this aspect, we performed a series of molecular dynamics (MD) studies of Rev$^{i,\ i+11}$ peptides crosslinked with both XL-H and XL-SO_3H bound to RRE structures. MD trajectories revealed that the sulfonate groups perturbed the α-helical structure of Rev$^{i,i+11}$-XL-SO_3H. Further details on this subject are provided at 2.1.3.4 Molecular dynamics simulation of Rev$^{i,\ i+11}$-XL-SO_3H and Rev$^{i,\ i+11}$-XL-H.

The non-sulfonated crosslinker (XL-H) produced peptides with substantially improved affinities; Rev$^{i,i+11}$-XL-H showed improved affinity for RRE with a K_d of 24 ± 9 nM (Table 10), approximately one order of magnitude lower than the dissociation constant measured for non-alkylated Rev$^{i,i+11}$ (K_d = 854 ± 40 nM) (Table 10). After irradiation with 363 nm light, the affinity decreased to 66 ±17 nM (Table 10). This threefold change in apparent affinity was in agreement with the observed change in the CD signal of Rev$^{i,i+11}$-XL-H (Figure 91) and with the low photo-conversion of only approximately 50% (Figure 89). A K_d of 1029 ± 65 nM could now be measured for Rev$^{i,i+7}$-XL-H (Table 2), while light-activated Rev$^{i,i+7}$-XL-H RRE complex showed a K_d of 607 ± 39 nM.

The absolute and relative affinities for peptides crosslinked with XL-H follow the trends expected from the degree of helicity present in their CD spectra, but whilst peptides crosslinked with XL-SO_3H demonstrate the expected relative affinities, they have poor absolute binding affinities compared to molecules incorporating XL-H. Such reductions in binding affinity when using XL-SO_3H would explain the results of EMSA studies performed on biosynthesized Rev peptides. Since the biosynthesized Rev peptides showed even lower helicity, the use of the XL-SO_3H and possibly the conditions used during electrophoresis disfavoured binding, reducing the affinity of the biosynthesized Rev peptides. Tellingly, the average helical content as judged by the CD is almost identical between peptides crosslinked with XL-SO_3H and XL-H, refuting previous assumptions that the major determinant of the success in this type of designs is high helicity of the Rev prior to binding.

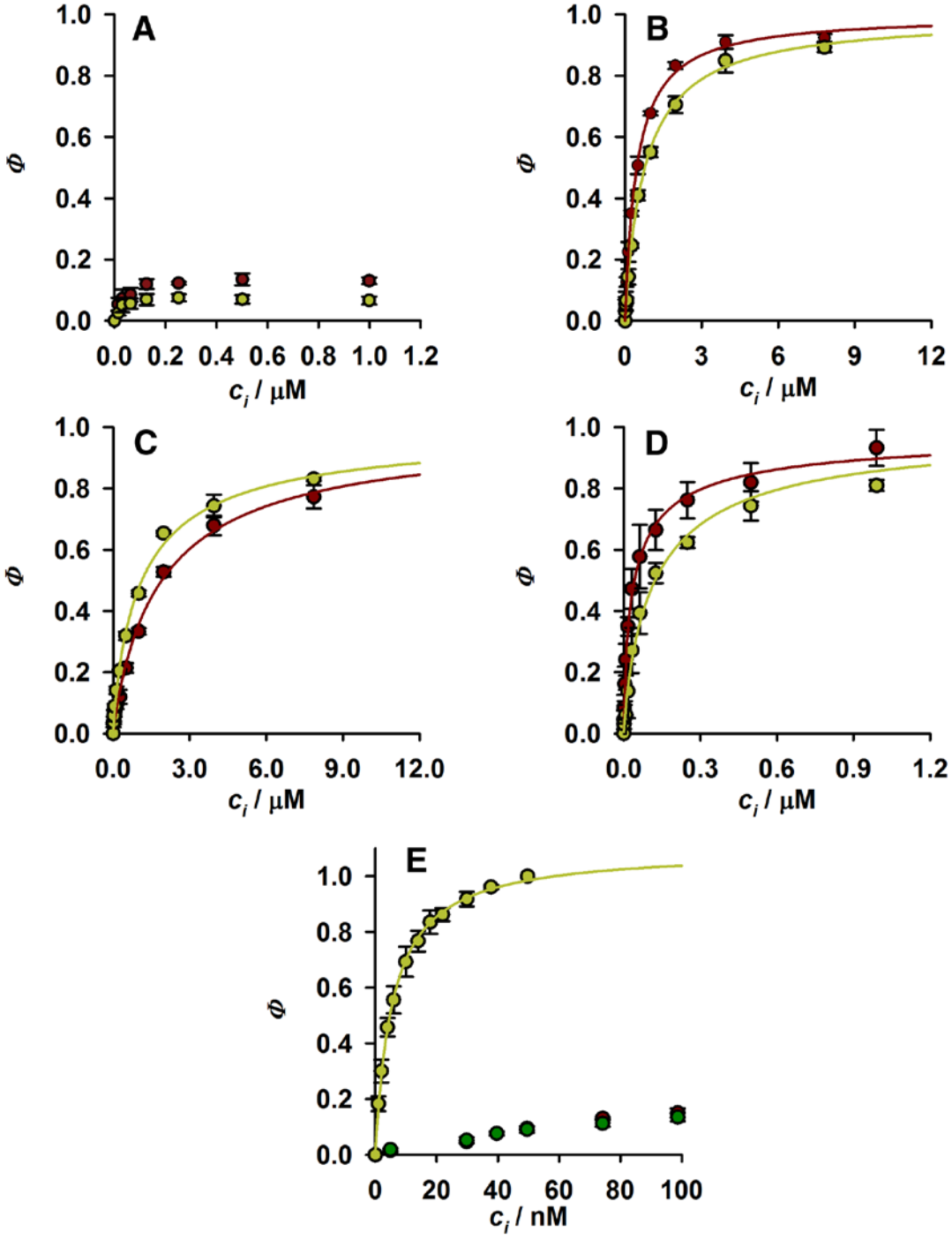

Figure 92: FAM Rev wt displacement plots, dark adapted (dark red) and after 3 min of UV_{363} irradiation (yellow-green) for (A) Rev$^{i,i+7}$XL-SO_3H; (B) Rev$^{i,i+11}$XL-SO_3H; (C) Rev$^{i,i+7}$XL-H; (D) Rev$^{i,i+11}$XL-H. (E) Direct binding curves for FAM-Rev^{wt} (yellow-green); FAM-Rev$^{i,i+7}$ (green); FAM-Rev$^{i,i+11}$ (dark-red).

Peptide	K_d
FAM-Rev^{wt}	5.8 ± 1.3 nM
FAM-Rev^{i+7}	561 ± 112 nM
FAM-Rev^{i+11}	424 ± 128 nM

Table 9: Binding affinities of the FAM labelled Rev peptides

Peptide	K_d
$Rev^{i, i+7}$-XL-SO_3H	NB
$Rev^{i, i+7}$-XL-SO_3H UV	NB
$Rev^{i, i+11}$-XL-SO_3H	247 ± 19 nM
$Rev^{i, i+11}$-XL-SO_3H UV	418 ± 23 nM
$Rev^{i, i+7}$-XL-H	1029 ± 65 nM
$Rev^{i, i+7}$-XL-H UV	607 ± 39 nM
$Rev^{i, i+11}$-XL-H	24 ± 9 nM
$Rev^{i, i+11}$-XL-H UV	66 ± 17 nM

Table 10: Binding affinities estimated form the FRET displacement assay for Rev peptides crosslinked with XL-SO_3H and XL-H, in dark and after UV_{363} irradiation (UV). NB signifies No Binding.

2.1.3.4 Molecular dynamics simulation of $Rev^{i, i+11}$-XL-SO_3H and $Rev^{i, i+11}$-XL-H

To fully explore the system Quantum mechanics (QM) calculations and molecular dynamics (MD) studies were performed in collaboration with Dr Andrea Brancale and Dr Antonio Ricci.[262] It was decided to consider all four possible crosslinker conformations obtainable from peptides incorporating XL-SO_3H (Figure 94). The azobenzene functionality has been extensively studied; particularly its *cis-trans* isomerisation and the dihedral angle formed between the two nitrogens of the azo group and the two *ipso* carbons of the aromatic moiety (Figure 93). In gas phase, it has been shown that these angles oscillate between two minima at ± 19° with a low-energy (0.3 kcal/mol) planar transition state.[263] Other studies[145] have described its conformation as planar, with a D_{2h} point group symmetry. Since our focus was on any change in geometry caused by the two sulfonate groups, we used HF/6-31G* and B3LYP/6-31+G** levels of theory to perform calculations rather than the more accurate MP2/6-31+G*. In particular, the choice of the HF/6-31G* basis set was made according to the requirements of the RESP and ESP charge derive (R.E.D.) software for deriving restrained electrostatic potential (RESP) charge values.[264] Minimum energy conformations were

obtained for both *anti* and *syn* isomers and their vibrational analyses (Figure 93, Table 10) were checked for consistency. Both levels of theory described the *anti*-disubstituted azobenzene as the preferred conformation over the *syn*. The planarity of the system was not perturbed by the sulfonate substituents, allowing to use the plain azobenzene parameters as reasonable approximation in the molecular dynamics (MD) experiments. R.E.D. software was used to derive RESP charges for the crosslinkers to model their properties in the MD set-up.

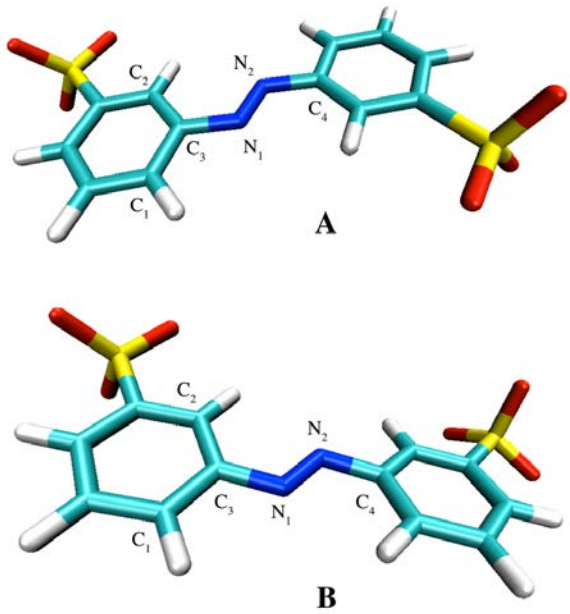

Figure 93: Geometries of the two isomers obtained from B3LYP/6-31+G* QM calculations. *Anti* conformation of XL-SO_3H (A); *syn* conformation of XL-SO_3H (B).

	HF/6-31G*		B3LYP/6-31+G**	
	A[b]	B[b]	A[b]	B[b]
ΔG (Kcal/mol)[a]	0.00	3.90	0.00	4.19
Bond (Å)				
C-N	1.42	1.42	1.45	1.42
N-N	1.22	1.22	1.23	1.22
Angle (Degrees)				
C_2-C_3-N_1	115	115.5	114	115
C_1-C_3-N_1	125	124.5	125	124
Dihedral (Degrees)				
C_1-C_3-N_1-N_2	180	174.5	179.5	174.5
C_3-N_1-N_2-C_4	180	179.5	179.9	179.5

Table 11: Relative Free Energy and selected geometrical parameters measures; [a] ΔG = Electronic and Nuclear Repulsion Energy + Zero Point Energy; [b] same nomenclature as Figure 93

The parameterization of the linker was achieved with the AmberTools package, using the General AMBER force field (GAFF).[265,266] Finally, the parameters obtained were added to the AMBER99 force-field in the GROMACS 4.5 package. To assure planarity of the azobenzene system, a further improper harmonic potential was added to the C-N-N-C and C-C-N-N dihedral. The value chosen for the force constant was calibrated to allow the C-C-N-N dihedral to oscillate between an angle of ±30° with a minimum energy value corresponding to the planar conformation (Table 11), in accordance with the experimental results.[267,268]

MD experiments were set up with the molecules to be studied were placed into a cubic TIP3P water box (with a 1.0 nm distance for the periodic boundary condition), neutralized with Cl^- or Na^+ ions and minimized according to the steepest descent algorithm. A 100 ps thermal equilibration (v-rescaling thermostat) followed by a 200 ps pressure equilibration (Berendsen rescaling algorithm) was initially performed keeping the positions of the heavy atoms fixed. Following short simulations, the positional restraint constants were gradually reduced to homogenize the systems. Finally, 5 ns production simulations were recorded for all the system investigated. The snapshots of averaged structures from the simulations of $Rev^{i, i+11}$-XL-H and the four conformers of $Rev^{i, i+11}$-XL-SO_3H are pictured in Figure 95.

These systems were also superimposed upon the non-substituted azobenzene in order to measure the deviation from the original linear structure (RMSD of the Cα, Figure 95) and the Rev peptide found in the Rev/RRE NMR structure (1ETF). MD simulations were run of the Rev/RRE complexes obtained to study the difference in interactions obtained between the various sulfonate substituted Rev peptide conformers and the RNA. Figures 96 and 97 allow comparison between the four Rev isomers/RRE complexes and the calculated non-substituted Rev/RRE structure.

MD trajectories revealed an interesting result: for $Rev^{i\ i+11}$-XL-SO_3H isomers the sulphate groups perturbed the α-helical structure of the isolated peptide, bending it out of a regular helix (Figure 94). This behaviour seemed to be caused not only by the steric hindrance of the sulfonate moieties, but also by the electrostatic interactions between these groups and the positively charged arginine sidechains of the Rev peptide. In some cases the intramolecular interaction between an arginine sidechain and the crosslinker sulfonates found in the isolated peptides is conserved, suggesting an impaired peptide-RNA binding interaction due to the reduced arginine/phosphate interactions. Moreover, the non-linear structure adopted by the substituted Rev peptide changes the morphology of the interaction RRE site, suggesting dynamic topological changes between the two interacting systems is at least partly

responsible for the different affinities. Taken together, these factors suggest a possible explanation for the different dissociation constants found as an ensemble of different contributions: *i)* intramolecular electrostatic attraction between arginine and sulfonate groups distorting the peptide before binding; *ii)* intermolecular electrostatic repulsion between sulfonate and phosphate groups on the RRE disfavouring binding and *iii)* distortion of the linearity of the α-helical secondary structure once bound (Figure 95 vs. 96).

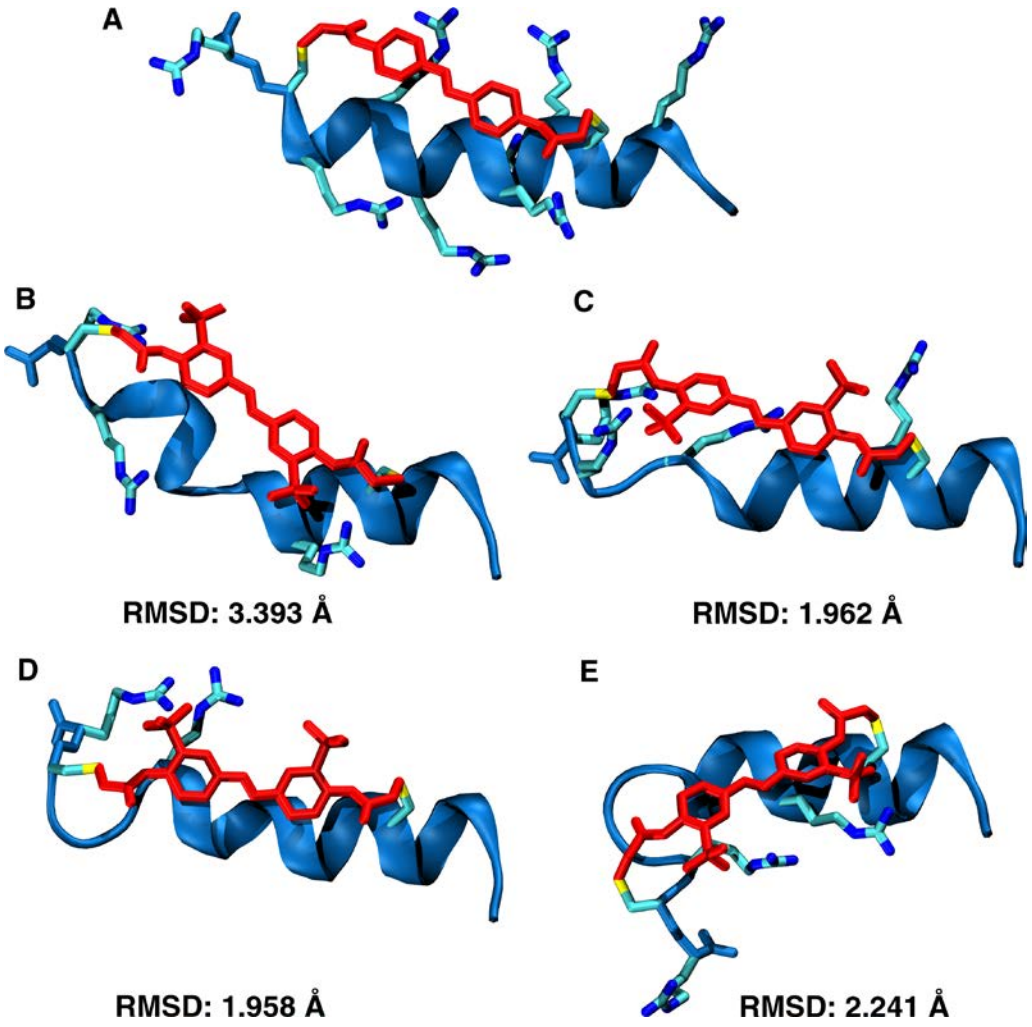

Figure 94: MD snapshots of (A) Rev^{i+11} XL-H and all the 4 possible Rev^{i+11} XL-SO$_3$H isomers (B – E) with the backbone RMDS of the respective MD simulations (all frames) from MD simulations (all frames) of Rev$^{i, i+11}$-XL-H is show below each isomer. Crosslinkers are in red; all arginine sidechains in Rev^{i+11} XL-H, and those interacting with sulfonates of XL-SO$_3$H in Rev$^{i, i+11}$-XL-SO$_3$H are in blue.

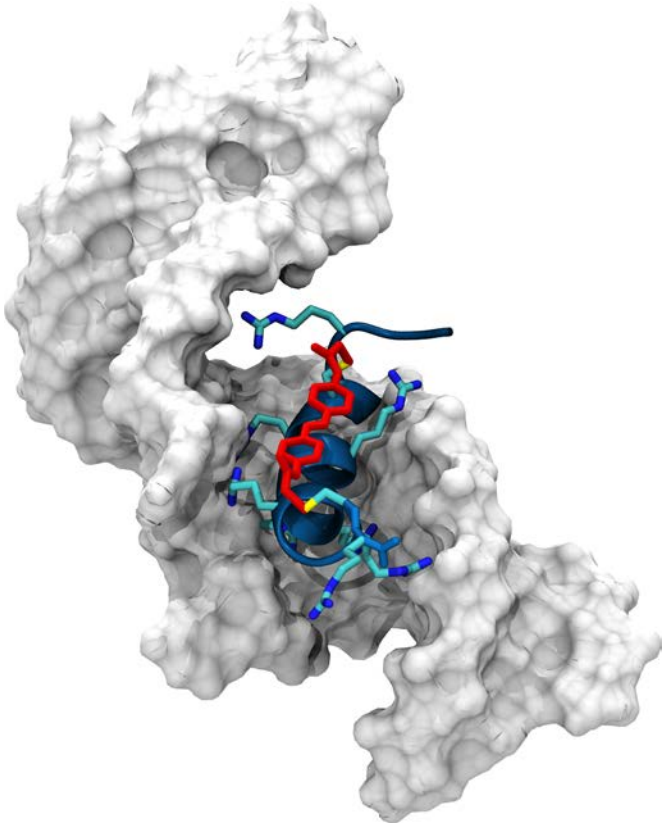

Figure 95: MD snapshot of the Rev$^{i, i+11}$ XL–H/RRE complex

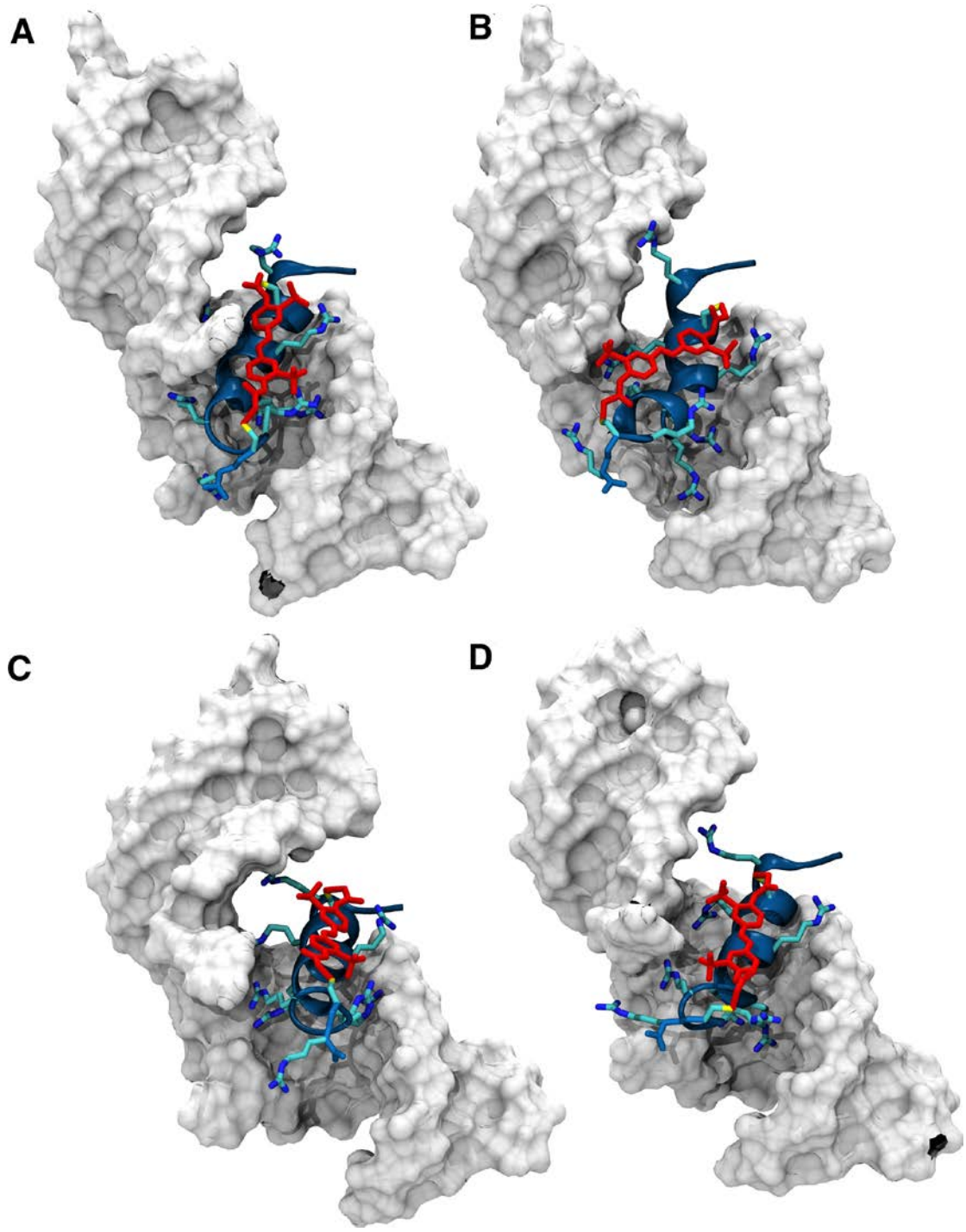

Figure 96: MD snapshots of the four different (A-D) Rev$^{i, i+11}$ XL-SO_3H/RRE complexes. Please note the intramolecular electrostatic interaction between the arginine and the sulfonate groups and the distortion of the binding site compared to Figure 71.

2.1.4 Conclusions

Peptides derived from the arginine rich region of Rev protein were both biosynthesized using the *E. coli* system and chemicly synthesized utilizing Fmoc chemistry. Optimized Fmoc synthesis allows production of higher amounts of Rev peptides and removes the possibility of C-terminal homoserine or homeserine lactone interference. Although the same biosynthetic approch was used earlier to produce Bak derived peptides in high quantities and double tandem oligonucleotides repeats was reported to be incorporated into the KSI plasmid (*p*ET-31b) with ease it is not the case for Rev derived peptides.[261] The KSI system is difficult to optimize for this particular sequence and is not an efficient source of these peptides but at the same time it proved to be good source of poly-arginine peptides (Thomas Fricke personal communication)[269]. In general, therefore, it seems that the *p*ET-31b expression system is not of general use and some careful consideration needs to be taken into account before utilising it for untried sequences.

It appears that although thorough biophysical characterization of the XL-SO_3H crosslinked biosynthesized peptides was performed and CD-derived structural features suggested a well designed system for further affinity studies, no binding to the RRE was detected. This problem was traced later to the crosslinker (XL-SO_3H) but this could not fully explain the lack of binding in the experimental set-up for the biosynthesized Rev derived peptides (EMSA).

The XL-SO_3H crosslinked Rev (Fmoc synthesized) had surprisingly low affinity to RRE. MD simulations performed on the Revs crosslinked with XL-H or XL-SO_3H both free and in complex with RRE provided a strong suggestion about the cause of this behaviour. Results of these simulations pointed unexpectedly not to charge-charge repulsion between phosphate backbone and sulphonates or steric hindrance, but to intramolecular interaction between the sulphonates and arginines. Strain imposed by those interactions translates into a bent conformation of the peptide, which together with the nonavailabality of intermolcularly interacting arginines to engage phosphate backbone is the most likely reason for the very low affinities. Those *in silico* findings were confirmed by crosslinking peptides with XL-H which, despite having virtually identical CD spectra, bind at least an order of magnitude more tightly. Previous reports suggest that a high helical content is the principal requirement for the construction of high affinity Rev peptides. However, our system illustrates that the obtained CD spectra, typical of a high degree of helicity, may not indicate an ideally preorganised conformation for tight binding and a relatively minor change in the substitution of the

crosslinker was observed to effect in a dramatic change in the affinity. An implication of this is the possibility that various sequences of peptides might interact with the crosslinker. Depending on the nature of the crosslinker and the given sequence those interactions might have various natures and are not necessarily limited to electrostatic interactions.

The biophotonic peptides derived from Rev targeting the HIV type 1 Rev Response Element present a first peptide/RNA system, whose stability can be tuned with light. The pitfalls of designing such systems have been clearly demonstrated and helpfully directed by computational studies. Although this research will serve as a base for future studies of photoswitchable peptides targeting RNA thorough computational evaluation of the next generations of biophotonic peptides designs (not only targeting RNA) is highly recommended.

The high binding affinities and relatively long-lived photoswitched conformations of the XL-H crosslinked peptides suggest that they can serve as useful probes and modulators of protein RNA interactions both *in vitro* and *in vivo* as part of a rapidly expanding toolbox of photo-controllable peptides.

A number of caveats need to be noted regarding the present study. Most importantly a rather limited Rev derived peptides sequence space has been explored which was a consequence of the aims of this study; it was intended to establish a prove-of-principle peptide photoswitch binding to RNA rather than rigorous optimization of the sequence. Moreover, not a full scope of available azobenzene crosslinkers was explored. Lastly, although the computational studies provide a probable model of photoswitchable Rev peptides in complex with RRE and generate interesting hypothesises, can not replace NMR of X-ray derived structural models.

A future study investigating structures of the photoswitchable peptides bound to RRE would be very interesting. NMR techniques derived structures could provide necessary details of the interactions. The issue of the distribution of four Rev^{i+11} XL-SO_3H isomers (free and in complex RRE bound) is an intriguing one which could be explored in further research. A reasonable approach to tackle this issue could be to use Carr-Purcell-Meiboom-Gill (CPMG) relaxation-dispersion methods of NMR spectroscopy. This would also allow to kinetically characterise the transition between all four isomers.

2.2 Solution structure of Bcl-x_L

2.2.1 Aims and Objectives

The aim of this study was to solve the NMR structure of Bcl-x_L protein in collaboration with Mathew Crump (University of Bristol). Whilst the structure of Bcl-x_L has previously been solved by both NMR and X-ray, it was important to provide a directly comparable model to the studies of the Bcl-x_L/$Bak_{I81F}^{i, i+11}$-XL-SO_3H that followed and has the potential to guide the more intricate problem of solving a solution structure of a protein-peptide complex. 2D and 3D NMR spectra were acquired with the aim of achieving high quality structure. The previously published NMR structure is of a low quality and it was important to provide an independent refinement of this vital structure.

2.2.2 Biosynthesis, purification and NMR sample preparation of Bcl-x_L

A loop and C-terminally truncated Bcl-x_L mutant (1-209 Δ45-84, Δ210-233) was used for the structural studies. These modifications have been shown not to disrupt the protein's activity *in vivo* or affinity to BH3 peptides *in vitro* and they were necessary for the preparation of high concentration samples for NMR studies.[248] Bcl-x_L was expressed in *E. coli* with a C-terminal His-tag that allowed straightforward affinity chromatography *via* NiNTA-sepharose to be applied. The protein was purified in one step to high homogeneity, dialysed into NMR sample buffer and concentrated to 1.2-1.5 mM (Figure 93, Materials and Methods).

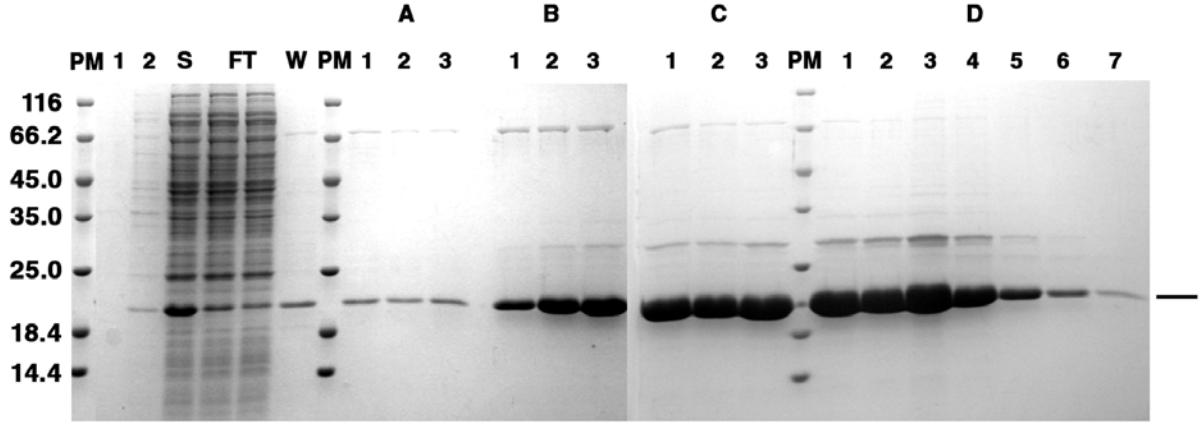

Figure 97: SDS-PAGE purification profile. PM – protein marker; 1 – cell culture sample prior to IPTG induction and 1 – five hours after; S – supernatant from cell debris separation; FT – flow through; A, B, C, D (1, 2, 3) fraction from NiNTA/Histag chromatography (A 100 mM, B 150 mM, C 200 mM, D 500 mM imidazole). The black bar indicates the postion of Bcl-x_L.

2.2.3 NMR spectra acquisition

Initial data were acquired on 600 MHz Varian Inova NMR spectrometer equipped with a room temperature probe. At the beginning of each set of NMR acquisitions an ^{1}H-^{15}N HSQC was taken as a reference point. It was followed by HNCA, HNCACB, HN(CO)CACB, HN(CO)CA, HNCO and HN(CA)CO experiments which allowed the backbone chemical shift assignments. Sidechain assignments were facilitated by C(CO)NH, HCC(CO)NH, HCCH-TOCSY, ^{13}C-HSQC spectra and in more difficult cases were aided by $^{13}C/^{15}N$ 3D separated NOESY. Inter-proton NOE signals were identified with the help of $^{13}C/^{15}N$ 3D separated NOESY. On two occasions additional ^{1}H-^{15}N HSQC spectra were recorded was taken to ensure that protein was unaffected by prolonged storage at 25 ºC.

2.2.4 NMR spectra assignment

Most of the Bcl-x_L ^{1}H-^{15}N HSQC signals arising from backbone or sidechain amides and tryptophan indole NH protons and were well separated suggesting a largely folded protein.[270] This is especially true for the tryptophan indole chains, predominantly localized in the hydrophobic core, which were spread across a wide chemical shift area; often a signature of a tightly packed, well formed protein interior. However some more central areas of the ^{1}H-^{15}N HSQC showed a significant signal overlap, suggesting a presence of flexible protein areas and posing difficulties in determining an unambiguous assignment.[270]

Ultimately, 96.4% of backbone amides signals were indentified. The only residues which did not have an identifiable related peak in ^{1}H-^{15}N HSQC spectrum were Met 1, Gln 3 and Arg 103, His 113, Gly 117, Ala 119 and Phe 123. The first pair correspond to the N-terminus, which often is flexible and the first methionine in *E. coli* can undergo a variety of modifications, further increasing heterogeneity.[271,272] It is also possible that the Met 1 is not present at all, as it often occurs in proteins biosynthesised in prokaryotic expression system. The second group are related to the BH3 domain binding site and most likely their poorly defined resonances represent its flexibility. A few residues could be assigned to more than one, or to a much broadened peak in the ^{1}H-^{15}N HSQC. Those are residues in the C-terminal part (Asn 197, Ala 199, Asn 207, Glu 208), which is known to be flexible and these residues probing a larger conformational space with what seems to be a few distinct minima.[248] Sidechain assignments reflected the difficulties in backbone assignments and had to be heavily aided by $^{13}C/^{15}N$ 3D separated NOESY in less well defined areas. Especially difficult were the inherently flexible parts of the protein between residues Asn 20 and Ser 43.

Although the signals associated with those residues were strong, unfortunately the overlap in proton-proton space was high. Fortunately, this loop is an artefact of the truncation of an even larger loop and both are expected to be unstructured.[248] Eventually 84.7% of carbon, 92.1% nitrogen and 77.5% of proton resonances were indentified. Interproton NOEs were manually extracted from $^{13}C/^{15}N$ 3D separated NOESY. Ambiguous peaks were deliberately left unassigned to be taken care of by Aria 2.3 algorithms.[273]

Atom type	% Assigned
Carbon	84.7
Nitrogen	92.1
Proton	77.5

Table 12: Assignment completeness for free Bcl-x_L; histidines from the C-terminal HisTag are excluded from the calculation

2.2.5 Bcl-x_L Structure Calculation

Interproton NOEs from $^{13}C/^{15}N$ 3D separated NOESY were used to calculate distance restraints by automatic Aria 2.3 protocols.[273,274] Those calculations were corrected for spin diffusion and log harmonic distance potential restraints were used in simulated annealing with CNS 1.21 as the structure calculation engine followed by structure refinement calculations using a thin shell of water.[275-277] Initial structures were used to correct misassigned NOEs by analysing those rejected or violated by Aria. In the final protocol 500 structures were calculated in the last iteration, 20 were water refined. These 20 structures were overlaid, coordinates averaged and the resulting structure underwent a brief energy minimazation to relieve clashes (CNS default protocols). Unless otherwise stated the comparisons or protein visualizations featuring a single structure are based on this average and energy minimazed structure.

2.2.6 Final Bcl-x_L structure ensemble

Calculations led to a structure ensemble with low RMSD values: 0.46 Å for backbone heavy atoms and 0.89 Å for all heavy atoms (Figure 98). Virtually all quality statistics surpass those of previously published Bcl-x_L NMR structures (Figure 99, Table 13 and 18).

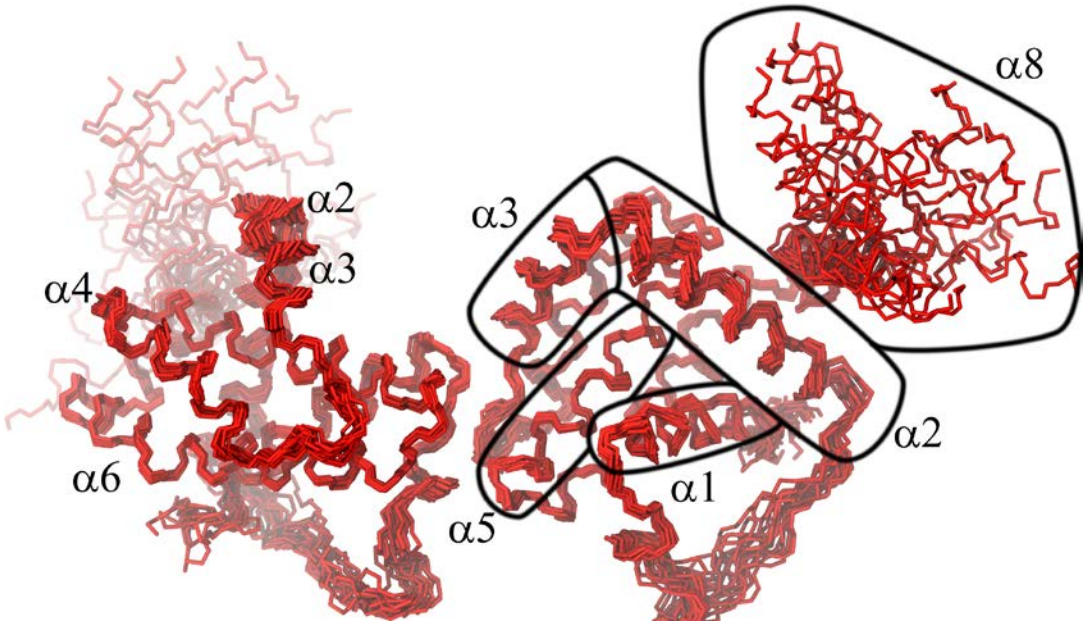

Figure 98: NMR ensemble of the Bcl-x_L. HisTag has been removed for clarity

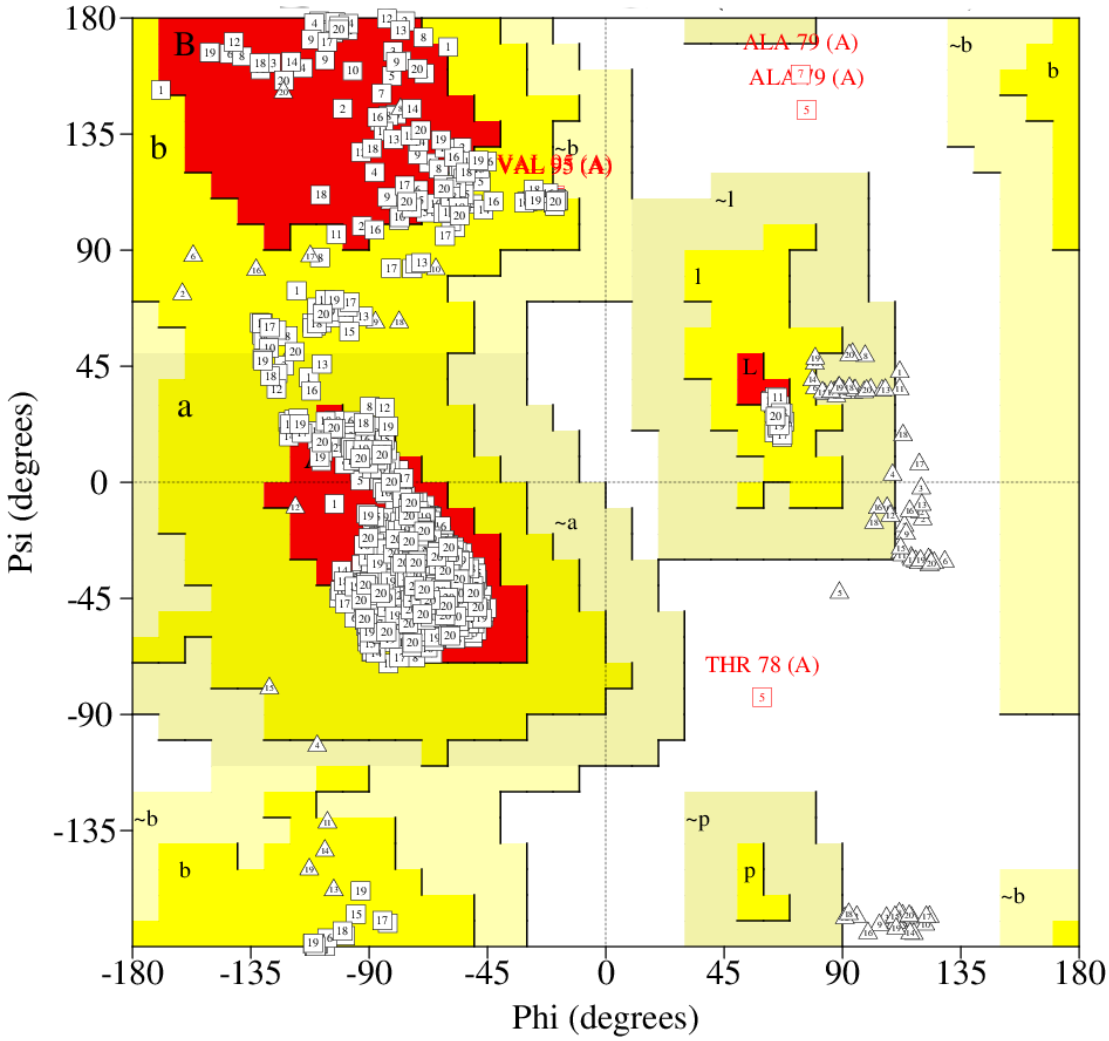

Figure 99: Ramachandran plot of the final water refined NMR ensemble of Bcl-x_L (residues 3 to 19 and 44 to 196 of Bcl-x_L and 72 to 87 of $Bak_{I81F}^{i,\,i+11}$-XL-SO_3H). Allowed region (red), additionally allowed regions (yellow) and generously allowed regions (pale yellow) are indicated. Glycines are indicated by triangles.

Statistic	Bcl-x_L
Number of restraints	
Total	3319
Unambiguous	2758
Ambiguous	561
Intra-residue	1084
Sequential	711
Short range	470
Long range	375
DANGLE φ /ψ	2x157
RMSD to restraints	
Bonds (Å)	0.003
Angles (°)	0.41
NOE (Å)	0.043
DANGLE (°)	0.36
Precision (RMSD to mean, Å)	
BB atoms (A.3-19,A.44-196,B.72-87)	0.46±0.08
AH atoms (A.3-19,A.44-196,B.72-87)	0.89±0.10
Ramachandran plot regions (%)	
Most favoured	92.9%
Additionally allowed	6.8%
Generously allowed	0.1%
Disallowed	0.1%
WHATCHECK Structure Z scores	
First-generation packing quality	1.853±0.559
Second-generation packing quality	4.658±1.236
Ramachandran plot appearance	-2.455±0.350
χ_1/ χ_2 rotamer normality	-0.826±0.473
Backbone conformation	0.194±0.225
WHATCHECK RMS Z-scores	
Bond lengths	1.000±0.001
Bond angles	0.285±0.003
Omega angle restrains	0.579±0.026
Side chain planarity	0.486±0.120
Improper dihedral distribution	0.453±0.082
Inside/Outside distribution	0.940±0.011

Table 13: Bcl-x_L NMR structural ensemble validation statistics. BB – backbone; AH – all heavy.

The RMSD between the published and the structure reported here is 3.67; it is a very large difference between structures that are supposed to be very similar.[248] On the other hand if we compare it to two crystal structures of Bcl-x_L the values are 1.80 Å for 1MAZ and 1.75 Å for 1R2D (Figure 76), i.e. the NMR structure described in this thesis more closely matches the high quality X-ray structures.[248,278] Nonetheless, there are some differences (the 2 crystal structures have RMSD values between each other of 0.65 Å). They cluster mainly around the binding sites, which as mentioned previously is flexible and responsive to changes in the local

environment, and mainly concern helix α3, which appears to be longer in our structure and slightly shifted towards α2. Taking into account the very good NMR quality statistics I hypothesise that the difference is due to a real structural diversity rather than being an artifact of NMR structure solution protocols. Such a change in a binding site known to reshape upon binding could be induced by the very high protein concentration essential to acquire NMR (molecular crowding) or the scarcity of ions in the NMR buffer, contrasted the rich and complex crystallization buffers. Additonaly, altered protein dynamics could contribute to the change in the binding site rendition.

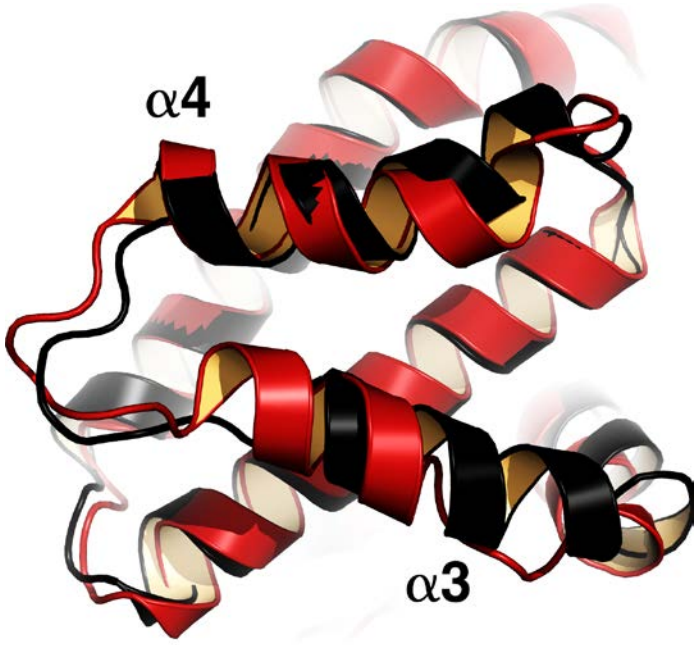

Figure 100: Comparison between shown here NMR structure of Bcl-x_L structure (black) and X-ray structure of the same protein (red). PDB entries: 1R2D.

In contrast to the published NMR structures of Bcl-x_L, the one reported herein appears with the C-terminal helix α8. Unfortunately, all crystal structures antiapoptotic members of the Bcl-2 family are based on deletion mutants deprived of the all residues beyond Gly 196 (helices α8 and α9) and therefore a direct comparison is not possible. Our NMR Bcl-x_L structure contains residues from Gly 196 to Arg 209 but not residues from Phe 210 to Lys 233; this part is predicted to form a transmembrane helix which is postulated to sometimes dock into the BH3 peptides binding site. The only structures available that are not truncated at equivalent positions are of Bcl-w and in this case helix α8 folds onto the canonical BH3 binding site (Figure 77).[244,245] For the NMR spectra of the structure reported here no, NOEs were found suggesting such a conformational state but the Histag and lack of mentioned C-

terminal sequence most likely precludes the folding of α8 on the rest of the protein as most likely it would be predominantly driven by interaction with the BH3 binding site.

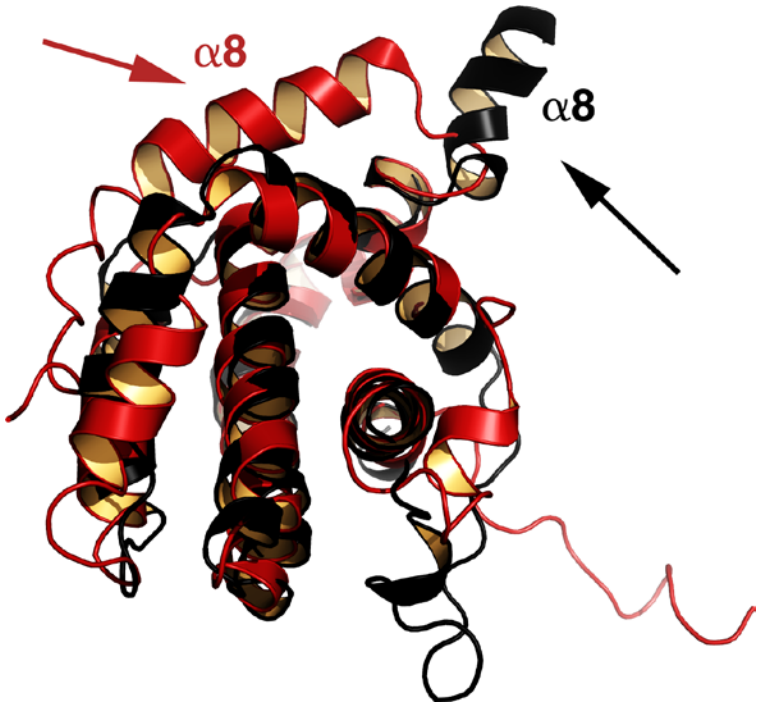

Figure 101: Comparison between the Bcl-x_L NMR structure and NMR structure of Bcl-w. Helix α8 is indicated in both proteins. Please notice the difference in length due to C-terminal truncation. PDB entries: 1O0L

2.2.7 Conclusions

The calculated Bcl-x_L NMR structure has very good quality statistics and served as a good base for subsequent structural studies of Bcl-x_L/$Bak_{I81F}^{i, i+11}$-XL-SO_3H. Using the same conditions allows for a direct structural comparison between the two structures and could also help in initial spectra assignment of the complexed structure. Moreover, the structure on its own is the best solution structure of Bcl-x_L determined so far, closely matching those acquired by X-ray techniques. The somewhat different binding site organization hints that it not only responds by reshaping when binding a BH3 sequence, but also can be influenced by the environment in monomeric state. This suggests that perhaps the environment could shape affinities for different BH3 peptides as reorganization cost to form differently shaped binding sites could be also dissimilar. A further study could assess what is or rather what are the structures of the binding site. This research should therefore concentrate on the investigation of dynamics of the binding site also in response to mutations as well as in response to binding events also away from the binding site.

2.3 NMR structural studies of Bcl-x_L in complex with $Bak_{I81F}^{i,i+11}$-XL-SO_3H

2.3.1 Aims and Objectives

The aim of this work was to gain a comprehensive structural description of Bcl-x_L/$Bak_{I81F}^{i,i+11}$-XL-SO_3H using NMR techniques, the first description of an azobenzene modified peptide bound to its target. Initial work involved calculation of a backbone amide chemical shift perturbation map by comparing ^1H-^1H HSQCs of free Bcl-x_L and in complex with $Bak_{I81F}^{i,i+11}$-XL-SO_3H. After acquiring a full set of NMR spectra and assigning them, molecular dynamics simulations using NMR-derived restraints were run in order to give the structure of Bcl-x_L/$Bak_{I81F}^{i,i+11}$-XL-SO_3H. The goal is to understand the similarities and differences between Bcl-x_L bound to wild type Bak peptide (Bak^{wt}) and $Bak_{I81F}^{i,i+11}$-XL-SO_3H.

2.3.2 Biosynthesis, purification and NMR sample preparation of Bcl-x_L/$Bak_{I81F}^{i,i+11}$-XL-SO_3H

The protein was biosynthesized and purified in the same way as for the free Bcl-x_L NMR studies it was then dialyzed into a suitable buffer for NMR acquisition and mixed with $Bak_{I81F}^{i,i+11}$-XL-SO_3H at a 1:1.1 ratio and concentrated. In later experiments to acquire isotope filtered spectra for $Bak_{I81F}^{i,i+11}$-XL-SO_3H resonance assignments the ratio used was either was 1:1.1 or 1:0.5 and the NMR buffer was prepared in oxygen free conditions and without β-mercaptoethanol. In order to prepare samples in D_2O, an equivalent buffer was prepared in D_2O (correcting for differences between *p*H and *p*D) and exchanged with the original protic one over a spin concentrator to a final concentration app. 1 mM.

2.3.3 Synthesis and purification of Bak derived peptides

$Bak_{I81F}^{i,i+11}$, FAM-$Bak^{i,i+11}$, FAM-$Bak_{I81F}^{i,i+11}$ and FAM-Bak_{wt} were synthesized according to standard Fmoc protocols and purified by HPLC techniques (Figure 102 and 103; Table 14). All were capped with an amide group at C-terminus and either modified with carboxyfluorescein or capped with an acetyl group at the N-terminus. FAM-$Bak_{I81F}^{i,i+11}$, FAM-$Bak^{i,i+11}$ and FAM-Bak^{wt} were provided by Dr Robert Mart. Crosslinking and purification of Bak peptides was achieved as it was described previously.[162] FAM-$Bak_{I81F}^{i,i+11}$-XL-SO_3H, FAM-$Bak^{i,i+11}$-XL-SO_3H were provided by Dr Robert Mart. Purity of the relevant peptides was confirmed by HPLC techniques (Figures 104-106).

Peptide	Sequence
Bak_{wt}	Ac–^{72}GQVGRQLAIIGDDINR^{87}–NH_2
$Bak^{i, i+11}$	Ac–GCVGRALAAIGDCINR–NH_2
$Bak_{I81F}{}^{i, i+11}$	Ac–GCVGRALAAFGDCINR–NH_2
FAM- $Bak^{i, i+11}$	FAM–GCVGRALAAIGDCINR–NH_2
FAM-$Bak_{I81F}{}^{i, i+11}$	FAM–GCVGRALAAFGDCINR–NH_2

Table 14: Sequences of Bak peptides used in this study.

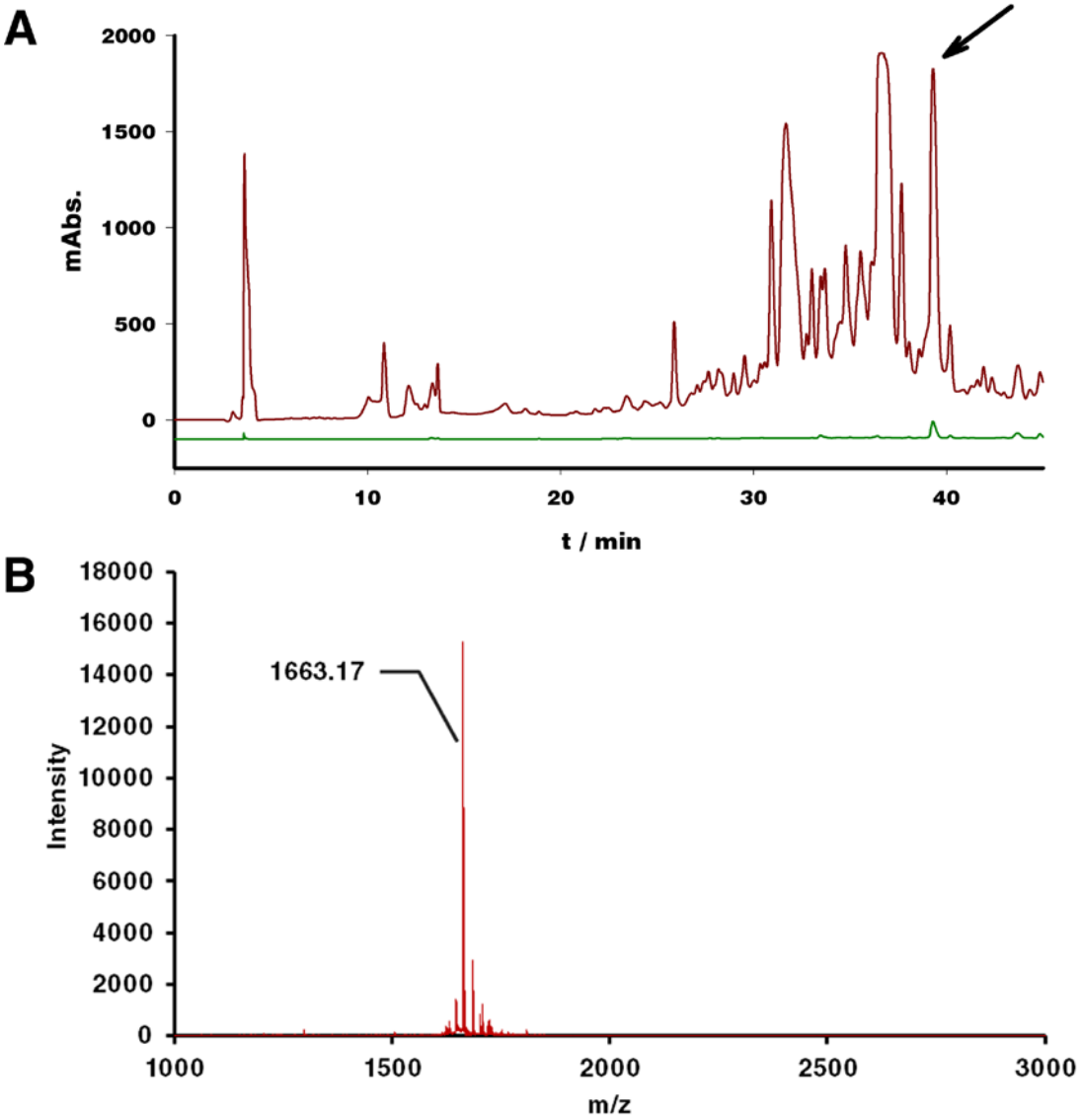

Figure 102: (A) HPLC chromatogram (red 210 nm, green 280 nm absorption) from the purification of crosslinking reaction of crude $Bak_{I81F}{}^{i, i+11}$ indicated with black arrow. (B) MALDI MS spectrum of $Bak_{I81F}{}^{i, i+11}$.

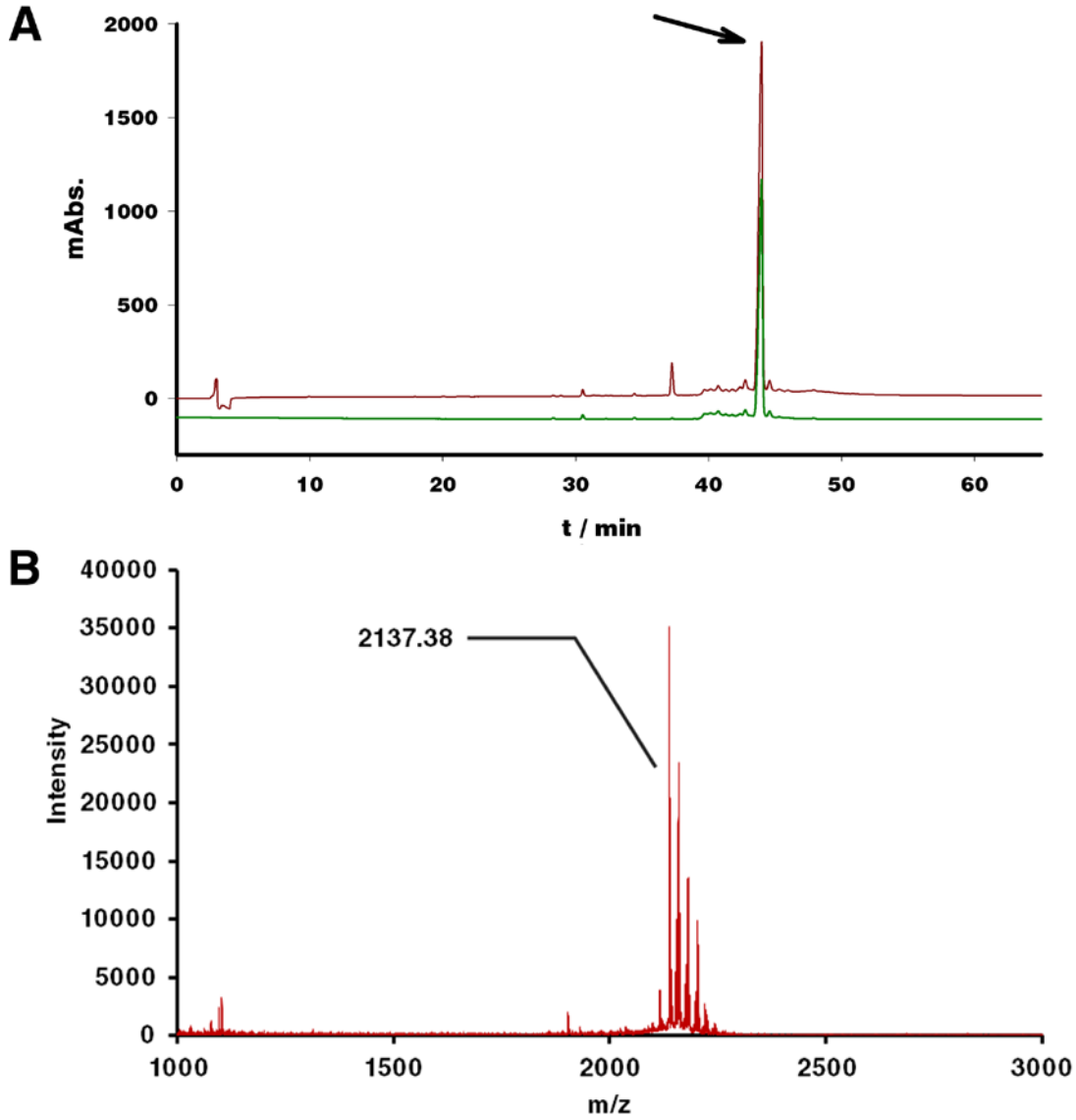

Figure 103: (A) HPLC chromatogram (red 210 nm, green 363 nm absorption) from the purification of crosslinking reaction of $Bak_{I81F}^{i, i+11}$-XL-SO_3H indicated with black arrow. (B) MALDI MS spectrum of $Bak_{I81F}^{i, i+11}$-XL-SO_3H.

Peptide	Observed		Calculated	
$Bak_{I81F}^{i, i+11}$	1663.805	$[M+H]^+$	1663.17	$[M+H]^+$
$Bak_{I81F}^{i, i+11}$-XL-SO_3H	2137.795	$[M+Na]^+$	2137.38*	$[M+Na]^+$

Table 15: Observed and calculated masses of purified Bak peptides.

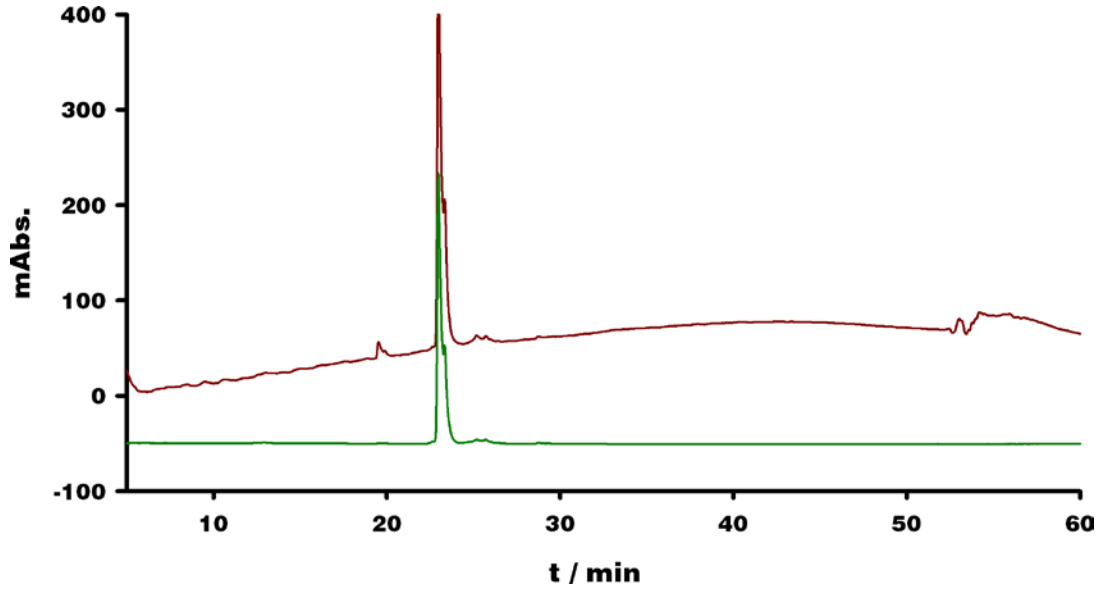

Figure 104: Analytical HPLC chromatograms of $Bak_{I81F}^{i, i+11}$-XL-SO_3H; red 210 nm, green 363 nm absorption.

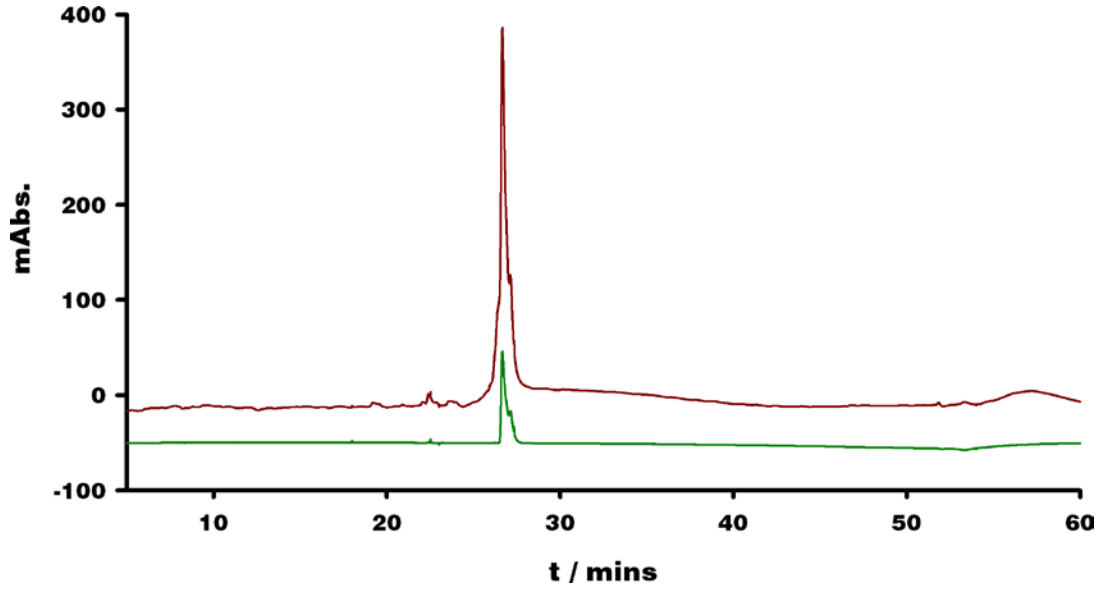

Figure 105: Analytical HPLC chromatograms of FAM $Bak_{I81F}^{i, i+11}$-XL-SO_3H; red 210 nm, green 363 nm absorption.

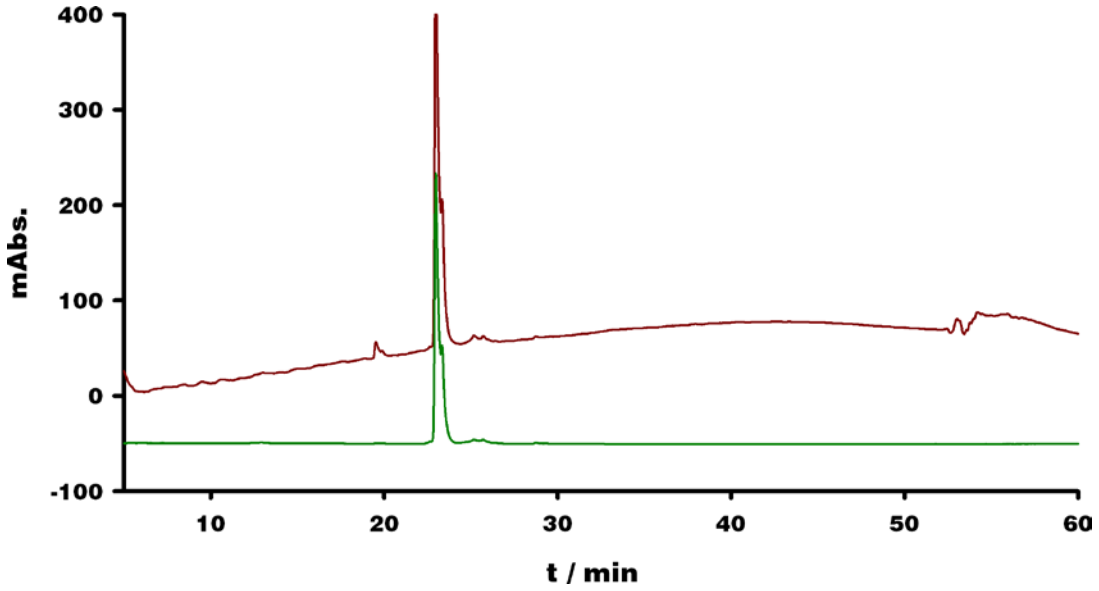

Figure 106: Analytical HPLC chromatograms of FAM Bak$^{i, i+11}$-XL-SO_3H; red 210 nm, green 363 nm absorption.

2.3.4 Preliminary chemical shift map of Bcl-x_L vs. Bcl-x_L/Bak$_{I81F}^{i, i+11}$-XL-SO_3H

Samples were prepared by Dr Sabine Kneissl in 10 mM sodium phosphate buffer *p*H 7.3, 5 mM 2-mercaptoethanol at a 0.3 mM concentration of Bcl-x_L (^{15}N labelled) and at 1:1.1 ratio with Bcl-x_L/Bak$_{I81F}^{i, i+11}$-XL-SO_3H were spiked with 5% D_2O (v/v). Spectra were acquired on a Varian INOVA 600 MHz NMR spectrometer (University of Bristol). The analysis of data, provided by Dr Sabine Kneissl, was done in cooperation with Dr Joel Loveridge and Dr Sabine Kneissl using Analysis 1.0.15 software.[162] Assignment of ^{1}H-^{15}N HSQC spectrum of Bcl-x_L in complex with Bak$_{I81F}^{i, i+11}$-XL-SO_3H was accomplished by comparing peak positions with known assignments for Bcl-x_L complexed with Bad, a 25 residue peptide derived from, a BH3 domain of the pro-apoptotic protein, which is known to bind in the same place as Bak^{wt}.[249]

The ^{1}H-^{15}N HSQC spectrum of Bcl-x_L in complex with Bak^{wt} exhibits a high degree of similarity with the ^{1}H-^{15}N HSQC spectrum of Bcl-x_L in complex with Bad or Bak^{wt}.[249,279] Assignments were confirmed using the ^{15}N-edited TOCSY-HSQC. The difference in peak positions between ^{1}H-^{15}N HSQC spectra of the Bcl-x_L and its complex with Bak$_{I81F}^{i, i+11}$-XL-SO_3H was defined as the length of a vector (Ω) connecting the centres of the corresponding peaks, calculated after weighting the differences between the ^{1}H and ^{15}N ppm dimensions (Equation 7, Figure 107).

$$\Omega = \langle \frac{1}{N} \{[\Delta(\delta H)]^2 + [0.102\Delta(\delta N)]^2\} \rangle^{\frac{1}{2}} 272$$

Equation 7: Equation use to calculate chemical shift perturbation

Residues were defined as having a significant shift if $\Omega/\Omega_{max} > 0.15$ (Ω_{max} =0.938 for Phe 131).

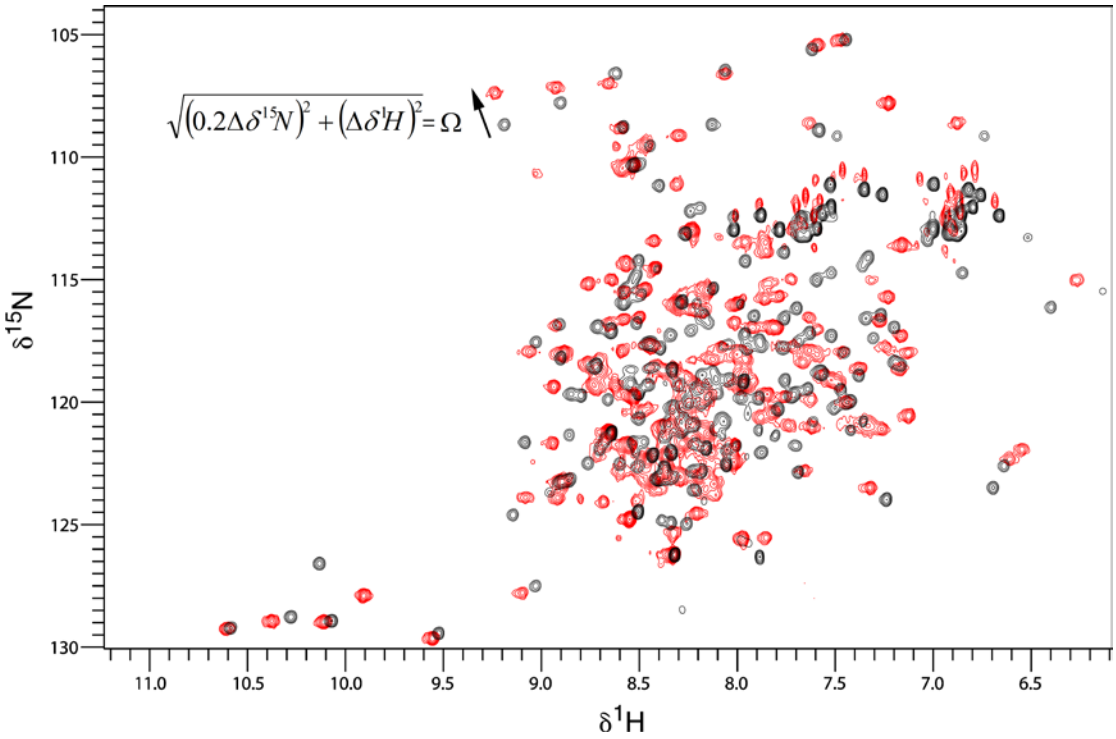

Figure 107: ^{1}H-^{15}N HSQC of free Bcl-x_L (gray) and a complex of Bcl-x_L with $Bak_{I81F}^{i, i+11}$-XL-SO_3H (red). The equation used to calculate the chemical shift perturbation Ω (upper right corner). Arrow indicates the change in Gln 94 peak position as an example.

It can be seen that the residues that show the highest values of Ω/Ω_{max} cluster around the binding sites (helices alpha α2, α3, α4) with the highest values of Ω/Ω_{max} (> 0.5) exhibited by residues Phe 97, Leu 99, Phe 105, Thr 109, Phe 131, Asp 133, Phe 146 (Figure 108 and 109). All of those residues are in the close proximity to the Bak^{wt} in the solution structure of its complex with Bcl-x_L, but none of them are recognized as directly important for binding.[247] It seems that the highest chemical shift perturbations are assosiated with residues that comprise the most mobile elements of the Bcl-x_L, areas that are known to significantly change position upon binding. It must be bourne in mind that values of chemical shift are directly related to the enviroment probed by the given nucleus, and in some cases nearby residues might be changed yet still be in a very similar type of enviroment giving misleadingly low values of Ω/Ω_{max}. Thus, it is possible that a partially buried residue becomes more solvent exposed during binding even as the binding site undergoes reshaping to accommodate a ligand yet in the final state/structure of the complex this residue packs agains

the peptide ligand and shows little change in the shift. This chemical shift perturabation mapping provides a strong evidence that $Bak_{I81F}^{i,\ i+11}$-XL-SO_3H adopts the same binding mode as the Bak^{wt}. This in turn suggests that the crosslinker, as intended, only influences the structure of the peptide, not the way it binds to the target.

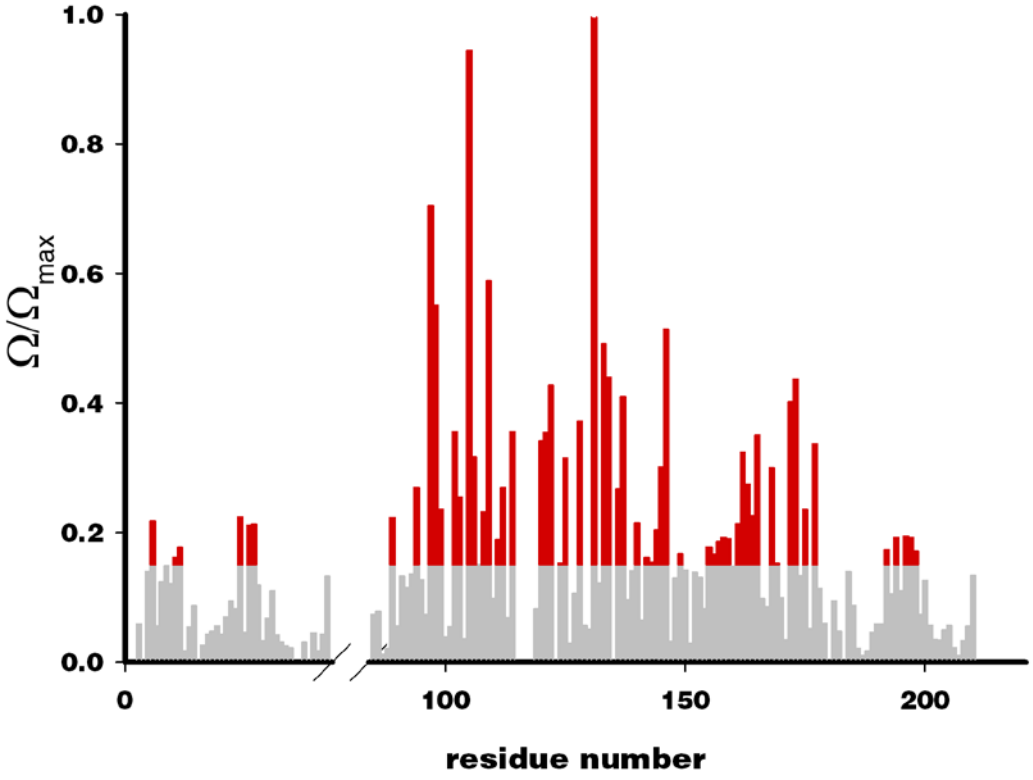

Figure 108: Backbone amide chemical shift perturbation map of free Bcl-x_L vs. Bcl-x_L in complex with $Bak_{I81F}^{i,\ i+11}$-XL-SO_3H; red colour indicates residues that have values of Ω/Ω_{max} higher then 0.15.

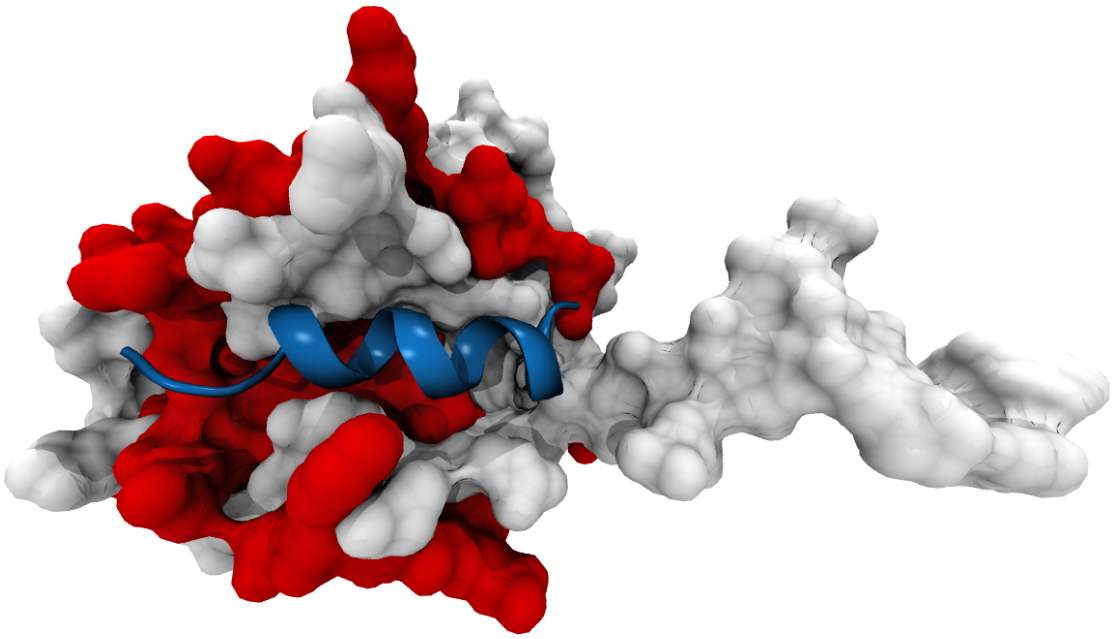

Figure 109: Chemical shift perturbation map from figure 108 transferred onto structure of Bcl-x_L in complex with *Bak^{wt}*.

2.3.5 Refined chemical shift map of Bcl-x_L vs. Bcl-x_L/Bak$_{I81F}^{i, i+11}$-XL-SO_3H

Using the NMR assignments of the Bcl-x_L (see Appendix) and Bcl-x_L/Bak$_{I81F}^{i, i+11}$-XL-SO_3H (see Appendix) acquired during the structure solving process a more refined chemical shift map has been calculated. It benefitted from using a more comprehensive set of NMR spectra that allowed to unequivocally determine the ^1H and ^{15}N resonances of the backbone amide for Bcl-x_L/Bak$_{I81F}^{i, i+11}$-XL-SO_3H and verify or correct some of the already published assignments for the free Bcl-x_L. The refined map confirmed with much higher greater accuracy that the crosslinked peptide adopts the same binding mode as the wild type (Figure 110 and 11). The significant shift differences clustered predominantly at the canonical BH3 binding grove located between helices α3 and α4. Some minor changes were also visible in the regions adjacent to structural elements that are known to reshape upon binding. Most importantly the residue with the highest Ω value is different, 1.185 for Leu 108 as opposed to 0.938 for Phe 131 in the previous map.[162] Moreover the overall characteristics of the chemical shift perturbation graph (see Figure 86) are different. Relatively similar values of Ω/Ω_{max} tend to cluster as opposed to the jagged plot in the previous case. The average value of Ω/Ω_{max} is higher: 0.172 vs. 0.159 even though the data was normalized to slightly higher value of Ω_{max} in the former case; suggesting that previously a number of assignments had been made based on imperfect assumption of proximity to either free Bcl-x_L peaks or in complex with Bad on ^1H-^{15}N HSQC spectra. The set of residues with the highest values of Ω/Ω_{max} (> 0.5) is also somewhat different and does not include Phe 91 despite it previously defining Ω_{max}.

The per residue RMSD values between free Bcl-x_L and Bcl-x_L/Bak^{wt} or Bcl-x_L/Bak$_{I81F}^{i, i+11}$-XL-SO_3H excluding the unstructured regions of the long loop region (Lys 19 to Glu 44) and flexible C-terminus (from Gly 196) onwards correlate very well to the chemical shift perturbation map (*vide infra*). This again suggests that this technique is better suited to recognize the mobile elements of a binding interaction than residues more directly engaged in this interaction.

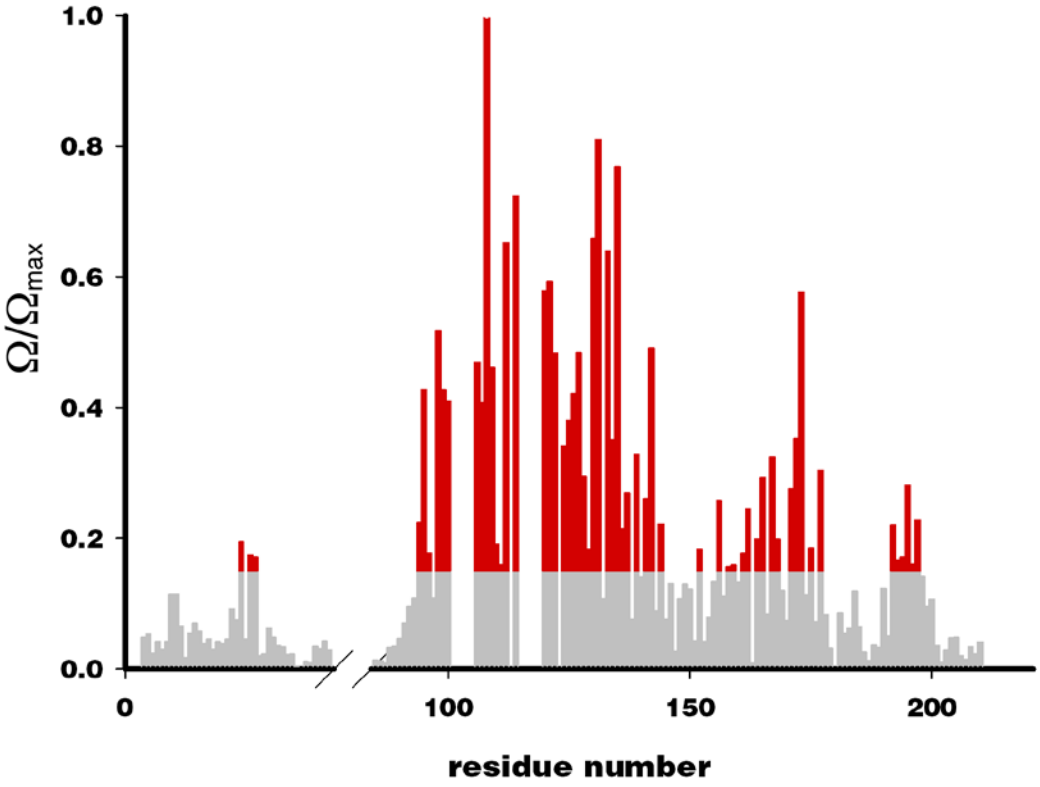

Figure 110: Refined backbone amide chemical shift perturbation map of free Bcl-x_L vs. Bcl-x_L in complex with $Bak_{I81F}^{\ i,\ i+11}$-XL-SO_3H; red colour indicates residues that have values of Ω/Ω_{max} higher then 0.15.

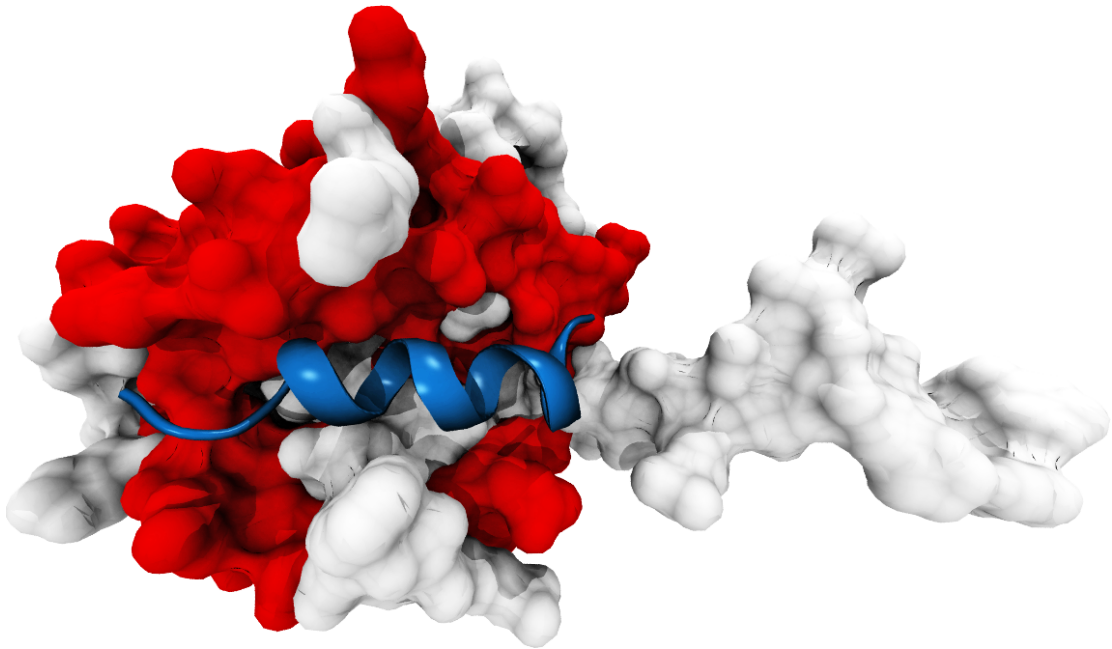

Figure 111: Chemical shift perturbation map from Figure 110 transferred onto structure of Bcl-x_L in complex with *$Bak_{I81F}^{\ i,\ i+11}$-XL-SO_3H*

2.3.6 NMR spectra acquisition

The first data set was acquired on 600 MHz Varian Inova NMR spectrometer equipped with a room temperature probe (University of Bristol). It included all spectra necessary for the backbone, sidechain and NOEs assignments (for the protein: ^{1}H-^{15}N HSQC, HNCA, HNCACB, HN(CO)CACB, HN(CO)CA, HNCO, CC(CO)NH, HCC(CO)NH, HCCH-TOCSY, ^{13}C HSQC, $^{13}C/^{15}N$ 3D separated NOESY; for the peptide: ^{13}C, ^{15}N F1, F2 filtered ^{1}H-^{1}H NOESY, ^{13}C, ^{15}N F1, F2 filtered ^{1}H-^{1}H TOCSY and F2 filtered ^{1}H-^{1}H NOESY.[280] However, it soon became apparent that the signal-to-noise ratio was insufficient for complete backbone and sidechain assignments, despite the methodology employed for data acquisition being the same as used for Bcl-x_L (apart from the additional filtered experiments). The reasons for this were traced to the different physical properties of the Bcl-x_L/$Bak_{I81F}^{i, i+11}$-XL-SO_3H complex, which is more prone to aggregation and therefore has a lower maximum attainable concentration; precipitation also shortens the useful life of the NMR sample. Most likely the conformational rearrangement induced by binding of the peptide makes Bcl-x_L more prone to domain swapping or similar higher order oligomer formation, a situation especially pronounced at the high concentration necessary for the NMR sample preparation. Such domain swapping has been observed before for some members of the Bcl-2 family and seems to be a hallmark of this group of proteins.[256,257] Therefore whilst these proteins are roughly spherical and middle-sized, solving their structures is far from routine due to these physical characteristics. Additionally, the presence of the crosslinker itself might influence the stability of the protein. Although any oligomers formed would tumble much more slowly in the solution, causing their NMR signal to be much weaker and broader and therefore either not visible or easily indistinguishable from the signal arising from the monomer of the complex it was decided to limit the time span for each data set acquired on every sample and instead prepare multiple fresh samples as often as necessary in order to avoid collecting data on possibly altered states of the protein complex.

In order to address the low signal-to-noise ratio problem, subsequent sets of NMR acquisition was performed on Varian VNMRS 600 MHz (^{1}H) spectrometer equipped with cryogenically cooled 5 mm z-gradient HCN probe. The same set of spectra as previously was acquired. Moreover a new set of $^{13}C/^{15}N$ separated 3D NOESY (dual acquisition) was obtained. Additionally, ^{13}C, ^{15}N F1, F2 filtered ^{1}H-^{1}H NOESY, ^{13}C, ^{15}N F2, F1 filtered ^{1}H-^{1}H TOCSY, F2 filtered ^{1}H-^{1}H NOESY and ^{15}N F1, F2 filtered ^{1}H-^{1}H TOCSY spectra were collected.[280] This new set of data allowed practically complete backbone and much improved

sidechain assignments. Initial interproton NOEs has been also indentified, but unfortunately both the quality and signal-to-noise ratio of the crucial ^{13}C 3D separated NOESY, from which most of the distance restrains are derived, was still insufficient for the structure calculation. The set of filtered experiments allowed only some assignments to be made of the $Bak_{I81F}^{i, i+11}$-XL-SO_3H. Only when Varian INOVA 900 MHz (^1H) spectrometer equipped with a cryogenically cooled 5 mm z-gradient HCN probe was employed for $^{13}C/^{15}N$ 3D separated NOESY, were the quality of spectra and signal intensity adequate for structure calculations were obtained. These spectra also allowed complete sidechain assignments.

Because the $Bak_{I81F}^{i, i+11}$-XL-SO_3H assignments remained a problem a new set of samples was prepared to solve this issue. Bcl-x_L/$Bak_{I81F}^{i, i+11}$-XL-SO_3H samples without 5 mM β-mercaptoethanol was prepared in oxygen free conditions including one sample with a higher Bcl-x_L to $Bak_{I81F}^{i, i+11}$-XL-SO_3H ratio to filter out signal arising from the free $Bak_{I81F}^{i, i+11}$-XL-SO_3H and another sample in D_2O without the reducing agent to facilitate identification of non-exchangeable protons. The same suite of filtered and half-filtered experiments was acquired and ^{13}C, ^{15}N filtered ^1H-^1H COSY were attempted, but no usable COSY spectra were collected from any of the aforementioned samples (Varian VNMRS 600 MHz (^1H) spectrometer equipped with cryogenically cooled 5 mm z-gradient HCN probe).

2.3.7 NMR spectra assignment

Backbone assignments posed a somewhat bigger challenge than in the case of free Bcl-x_L. The middle regions of the filtered ^1H-^{15}N HSQC exhibited a marginally higher level of overlap between peaks (e.g. residues Glu 32, Phe123, Val 141, Ile 182) similarly, as for the free Bcl-x_L. Residues Asn 157, Ala 159, Ala 160 and Ala 161 were represented by broad or multiple peaks. Some residues backbone resonance could not be assigned (Met 1, Ser 2, Gln 3, Tyr 101, Arg 102, Arg 103, Gly 117, and Glu 171) and these residues differed from those not assignable for Bcl-x_L. Again, the residues that could not be assigned belonged to the flexible parts of the Bcl-x_L protein. Unfortunately the HCCH-TOCSY, CC(CO)NH TOCSY and especially the HCC(CO)NH TOCSY spectra had a lower signal-to-noise ratio then their equivalents for the free protein and were on their own not sufficient for sidechain assignments. Assignment was therefore heavily aided by $^{13}C/^{15}N$ separated 3D NOESY (Figure 88 and 89) and in the case of some more difficult backbone assignments ^{15}N separated 3D NOESY provided useful information about $i, i \pm 1$ connectivity (Figure 112).

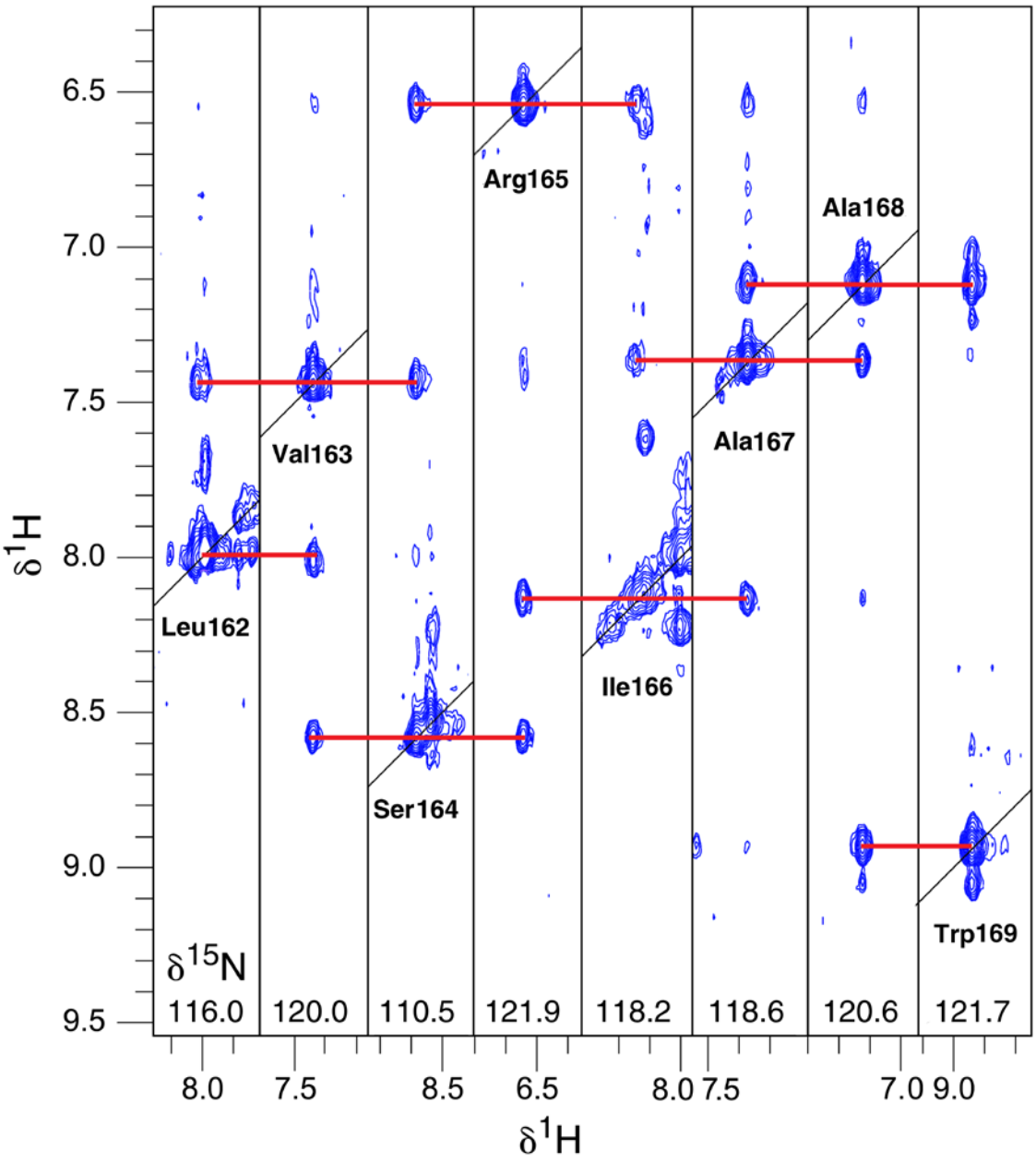

Figure 112: ^{15}N separated 3D NOESY of Bcl-x_L/$Bak_{I81F}^{i, i+11}$-XL-SO_3H complex. Strips of Leu162 to Trp169 backbone amides originating NOEs are shown and crosspeaks to adjacent residues are indicated with a red line.

As in the case of the free Bcl-x_L, interproton NOEs were manually assigned with less certain ones left unidentified for Aria automatic disambiguation protocols. Although the initial set of ^{13}C/^{15}N separated 3D NOESY had a very low signal-to-noise ratio, the set acquired on 900 MHz spectrometer provided a good quality, intense spectrum (Figure 113). Unfortunately due to limitations of the technique it is very difficult to obtain TOCSY type or filtered spectra on 900 MHz spectrometer.

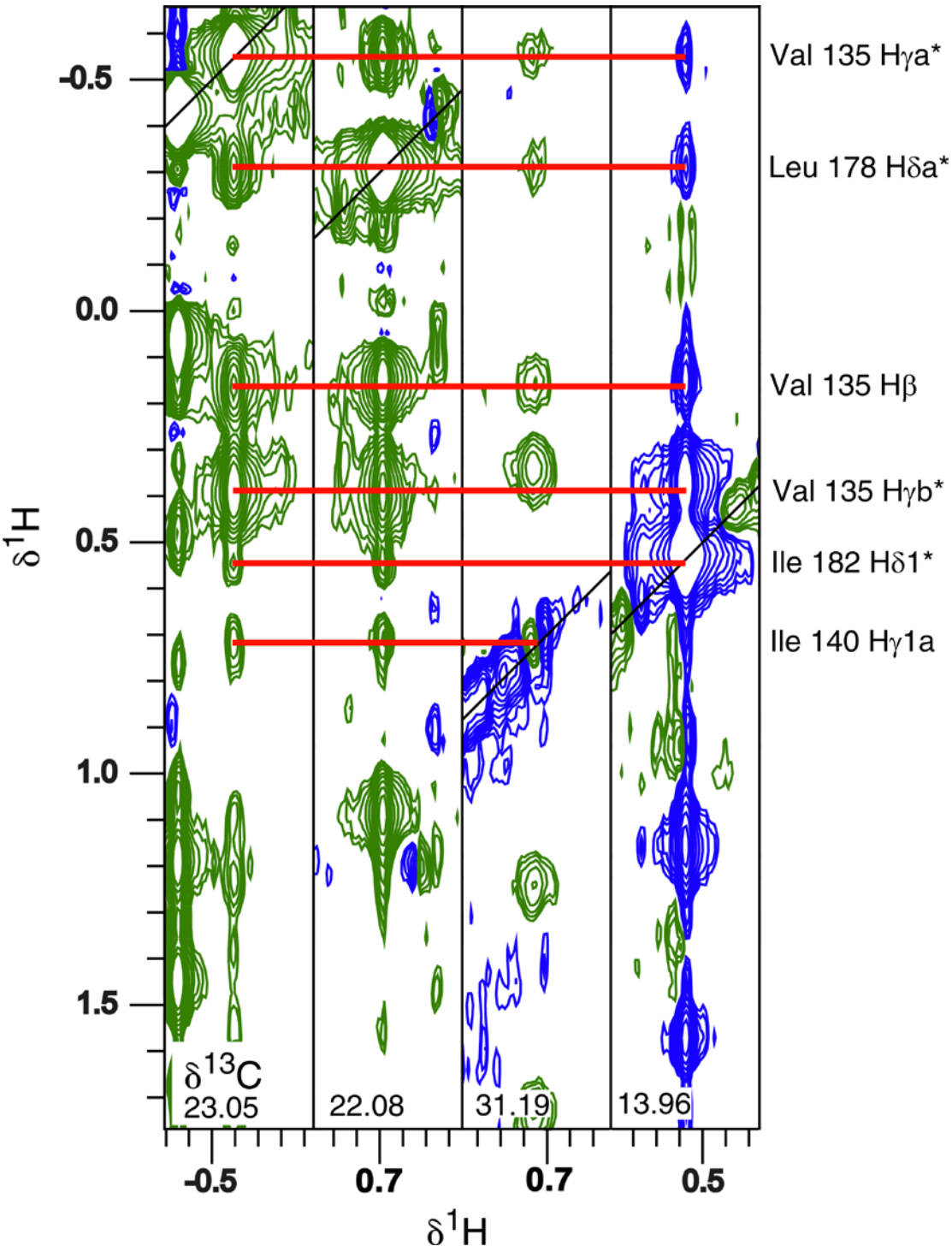

Figure 113: ^{15}C separated 3D NOESY of Bcl-x_L/ $Bak_{I81F}^{i, i+11}$-XL-SO_3H complex. Four strips are shown exemplifying the long range NOEs assignment. Corresponding crosspeaks are indicated with a red line with the given proton they match to on the right. Green peaks are positive and blue ones are negative.

Assignment of the non-isotopicaly labelled peptide resonances was achieved *via* the previously mentioned filtered experiments in which signals of protons attached to ^{13}C and/or ^{15}N are rejected, thus only those on isotopicly not labeled peptide are detected. In order to aid the assignment and also to solve the structure of the free $Bak_{I81F}^{i, i+11}$-XL-SO_3H a set of 2D

spectra of the peptide alone were acquired. Unfortunately at the concentrations required for NMR the peptide itself tends to aggregate and does not exhibit NMR spectra typical of an α-helix, despite the CD spectrum clearly indicating this structure. Interestingly in the earlier research it has been observed that at lower concentrations this peptide shows up as practically 100 % helical.[162] Although it was not possible to obtain any restraints that could be used to estimate the structure of the free peptide, these spectra aided in filtering out signal arising from the unbound peptide in the sample of the complex with Bcl-x_L. Only proton resonances of the non-isotopically labelled peptide could be accessed and 13 of a total of 16 residues, Cys 73 to Ile 85 (Table 16), yielded at least one associated resonance (87.5% assigned in total); the work posed a significant challenge due to high signal overlap, low intensity of some of the signals and influence of the crosslinker which shifts the usual patterns of amino acids into less obvious areas (Figure 114).

Bcl-x_L/$Bak_{I81F}^{i, i+11}$-XL-SO_3H protein-peptide NOEs were extracted primarily from ^{13}C separated 3D NOESY, but also from F2f ^{1}H-^{1}H NOESY (Figure 90) and to a lesser extent ^{15}N separated 3D NOESY. Most of those signals between residues 74 to 81 were clustered with no interproton peptide-protein NOEs identified for the first N-terminal residue (Gly 72) or the last 2 C-terminal residues (Asn 86, Arg 87). Residues on the water exposed/crosslinker shielded face showed no NOEs to Bcl-x_L.

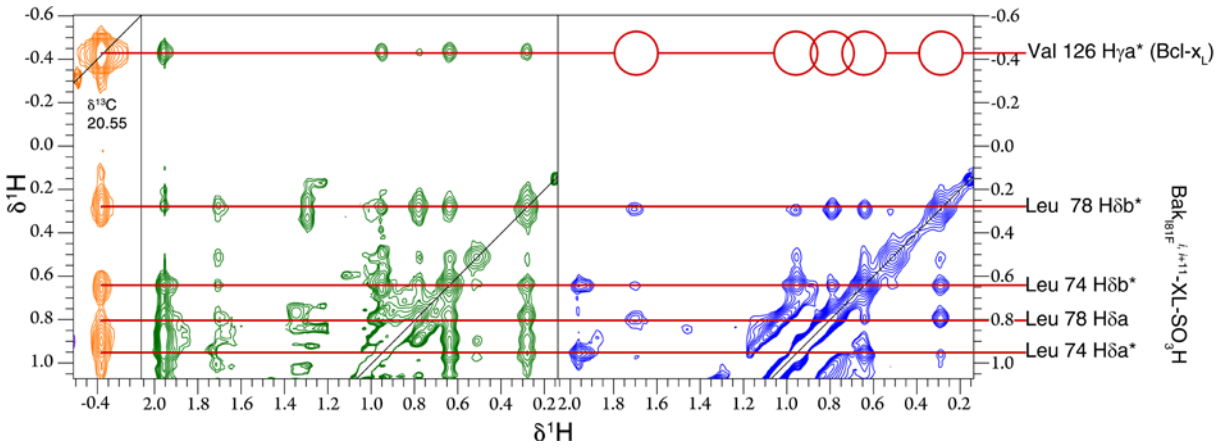

Figure 114: ^{15}C separated 3D NOESY (black), ^{13}C, ^{15}N F1, F2 filtered ^{1}H-^{1}H NOESY (blue) and F2f ^{1}H-^{1}H NOESY (green)of Bcl-x_L/$Bak_{I81F}^{i, i+11}$-XL-SO_3H complex. Strips exemplify protein-peptide NOEs by demonstrating Val 126 (Bcl-x_L) to Val 74 and Leu 78 ($Bak_{I81F}^{i, i+11}$-XL-SO_3H) NOEs. Corresponding crosspeaks are indicated with a red line with the given proton they match to on the right. Red circles indicate peptide-protein NOEs that are present in F2f ^{1}H-^{1}H NOESY and missing in ^{13}C, ^{15}N F1, F2 filtered ^{1}H-^{1}H NOESY.

Atom type	% assigned
Carbon	69.1
Nitrogen	69.3
Proton	85.6

Table 16: Assignment completeness for free Bcl-x_L in complex with $Bak_{I81F}^{i, i+11}$-XL-SO_3H

2.3.8 Bcl-x_L/ $Bak_{I81F}^{i, i+11}$-XL-SO_3H structure calculation

Interproton NOEs from the $^{13}C/^{15}N$ 3D separated NOESY spectrum were used to calculate intra-Bcl-x_L distance restraints. Protein-peptide distance restraints were extracted from the $^{13}C/^{15}N$ 3D separated NOESY and F2f ^{1}H-^{1}H NOESY spectra. Intramolecular restraints for the $Bak_{I81F}^{i, i+11}$-XL-SO_3H originated from the ^{13}C, ^{15}N F1, F2 filtered ^{1}H-^{1}H NOESY spectrum. All NOEs were selected manually, but were calculated by Aria 2.3 protocols correcting for spin diffusion with CNS 1.21 as the structure calculation engine and the structures refined in a thin shell of water.[273,275-277] In order to ensure that protein-peptide distance restraints are preserved by Aria during the simulated annealing iterations they were removed from the list of restraints that could be rejected by Aria to prevent ejection of the peptide from Bcl-x_L during the high-energy early stages of the simulation. Instead those restraints were manually analyzed after each calculation for violations and, if necessary, corrected. Initial structures were used to correct misassigned NOEs by analysing those restraints rejected by Aria or violated. The protocol for simulated annealing was modified to extend the cooling times as suggested in the literature in order to better deal with ambiguous assignments.[281] In the final protocol 500 structures were calculated in the last iteration, and the best 20 were water refined. Coordinates of those 20 structures were averaged and the resulting structure was subjected to a short energy minimization (CNS default settings) to ease any possible clashes and correct any aberrant geometries.

Initial crosslinker parameterization was achieved with the help Dr Marc van der Kamp (University of Bristol) and then refined.[282] There are two possible arrangements of the free crosslinker sulfonate groups: *syn* with both sulfonates on the same side of the crosslinker and *anti* with them opposite. However, when crosslinked to a peptide, the directionality of the constituent amino acids split each of the arrangements creating four possibilities. By varying parameters of the crosslinker in calculations at late stages of refinement it was found that only one *syn* and one *anti* arrangement appear in the final ensemble of structures. One of those ensembles, the *syn*, was violated to lower degree and had more favourable overall energies; therefore it was taken to better represent the arrangement of the crosslinker.

2.3.9 Final Bcl-x_L/ $Bak_{I81F}^{i, i+11}$-XL-SO_3H structure ensemble

The quality of the final structure surpasses most of the published NMR structures of Bcl-2 family proteins (Table 21). The RMSD between free Bcl-x_L (see previous chapter) and Bcl-x_L bound to $Bak_{I81F}^{i, i+11}$-XL-SO_3H backbone atoms is 2.27 Å. As in the case of similar structures with different BH3 peptides, the principal differences between free and complexed Bcl-x_L concern the parts of α2 and α3, α4 helices and the loop connecting them. In fact if those (residues 94-138) are excluded from the RMSD calculation the value drops to 0.96 Å with the structures virtually identical in respect to the non-binding site fragments.

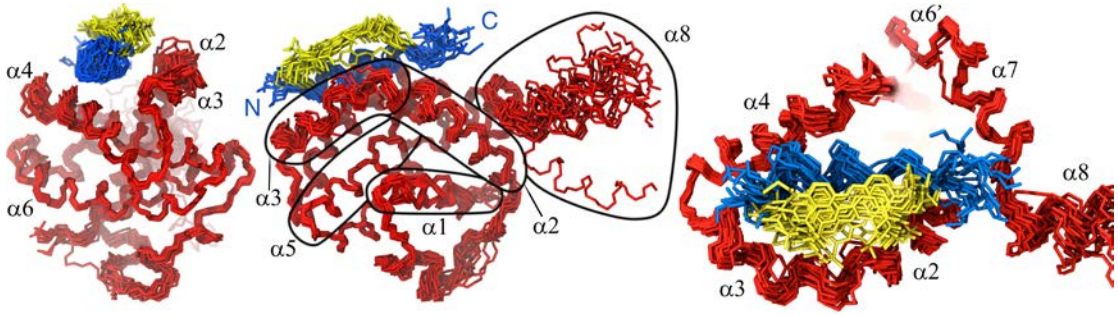

Figure 115: NMR ensemble of the Bcl-x_L in complex with $Bak_{I81F}^{i, i+11}$-XL-SO_3H. HisTag has been removed for clarity.

When compared to the published structures of wild type Bcl-x_L bound to Bak^{wt} the RMSD is 2.37 Å for backbone atoms, but care has to be taken in assessing similarities between those as the quality of the mentioned Bcl-x_L/Bak^{wt} structure is quite low.[247] The residues Phe 105 to His 113 which form part of helix α3 appear to be disordered in the published structure of Bcl-x_L/Bak^{wt} but appear in our structure as only slightly distorted helices, making this region more similar in secondary structure to some other structures of Bcl-x_L bound e.g. Bad (RMSD 1.59 Å and 1.19 Å when residues 94-138 are not considered) than the NMR structure with Bak^{wt} (Figure 118 and 94). Finally if we once again exclude the whole binding site (residues 93 – 134) from comparison, the RMSD value for backbone is 1.66 Å and 2.68 Å for all heavy atoms. This is a remarkable discrepancy, with the RMSD values between free Bcl-x_L and $Bak_{I81F}^{i, i+11}$-XL-SO_3H, crudely suggesting that my structure is more similar to the free Bcl-x_L than the Bcl-x_L/Bak^{wt} is. Other differences in our structure compared to the wild type structure include the transition between helix α2 and α3, which is smooth (there is no well defined breakpoint) in our structure as opposed to turning rapidly at residue 98 (Figure 118). Part of helix α2 (residues Leu 99 to Phe 105) which forms a bent helix connecting α2 and α3, however the same region in Bcl-x_L/Bak^{wt} NMR structure continues straight on as a part of α2. Differences in this region could either be a consequence

of the low quality of this region in the Bcl-x_L/Bak^{wt} structure (the disorderly character of this region contrasts well defined structure seen in virtually every other Bcl-x_L/BH3 peptide NMR or X-Ray structural model) or it could be attributed to the interaction between the one of the crosslinker sulfonate groups and Arg 100 or/and 103. However, such interactions are likely to be of transient nature as although mentioned arginines indeed occupy the right space for a possible interaction with the crosslinker in the final structures sulfonates are further then 5 Å from any possible positively charged residues with minimum for a salt bridge being 4 Å.[283].

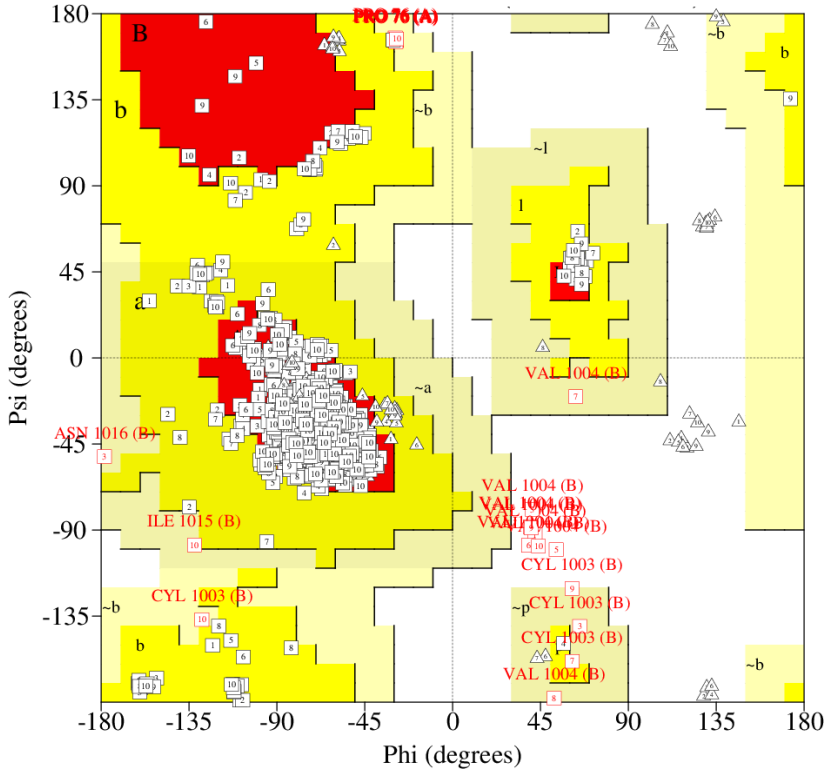

Figure 116: Ramachandran plot of the final water refined NMR ensemble of Bcl-x_L bound to Bak$_{I81F}^{i, i+11}$-XL-SO_3H (residues 3 to 19 and 44 to 196 of Bcl-x_L and 72 to 87 of Bak$_{I81F}^{i, i+11}$-XL-SO_3H). Allowed region (red), additionally allowed regions (yellow) and generously allowed regions (pale yellow) are indicated. Glycines are represented by triangles.

The last major difference between Bcl-x_L/Bak^{wt} and Bcl-x_L bound to Bak$_{I81F}^{i, i+11}$-XL-SO_3H is found in two closely interacting regions: helix α7 (residues 175 to 185) and the loop connecting helices α4 and α5. It appears that both the position of the helix α7 and the loop is significantly different between the structures. The change in the position of the loop possibly occurs *via* altered interactions with the peptide between the two structures with a concomitant shift in helix α7 as in both structures the helix and loop appear bound together *via* the interaction of a hydrophobic patch formed by a of Leu 178, Trp 181 on the loop with Val 135 of the helix wedged between them. Again, at least one residue in the Bcl-x_L/Bak^{wt} structure that forms the loop localizes far from the most probable Ramachandran space.

Table 17: Comparison of WhatIf Z-score quality analyses provided by the iCING server of common residues (1-19, 85-196) of the structural statistics of the NMR structures reported here and those of previously determined structures

	Free Bcl-x_L Structures		Bcl-x_L/Peptide Complex Structures		
Structure PDB ID	**Bcl-x_L** **2LPC**	Bcl-x_L 1LXL	**Bcl-x_L/$Bak_{I81F}^{i, i+11}$-XL** **2LP8**	Bcl-x_L/Bak^{wt} 1BXL	Bcl-x_L/Bad^{wt} 1G5J
First-generation packing quality	**1.811**	-1.590	**-0.139**	-1.592	-1.822
Second-generation packing quality	**4.467**	-1.740	**1.756**	-1.846	-2.303
Ramachandran plot appearance	**-2.405**	-5.378	**-4.826**	-6.212	-5.307
$\chi1/\chi2$ rotamer normality	**-0.887**	-7.763	**-4.960**	-6.970	-7.681
Backbone conformation	**0.104**	-1.044	**-1.124**	-0.446	-0.508
Ramachandran plot (%)					
Core	**92.7**	77.0	**88.0**	85.1	86.6
Allowed	**7.1**	18.9	**10.8**	11.2	10.6
Generous	**0.1**	1.6	**0.5**	3.0	2.1
Disallowed	**0.1**	2.5	**0.7**	0.7	0.7

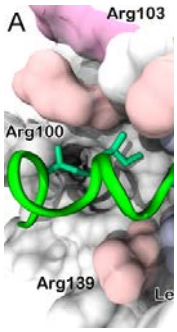

Figure 117: Comparison between a) Bak_{wt}/Bcl-x_L (1BXL), b) $Bak_{I81F}^{i, i+11}$-XL/Bcl-x_L (2LP8) c), and Bad/ Bcl-x_L (1G5J).

	Bcl-x_L/$Bak_{I81F}^{i, i+11}$-XL-SO_3H
Number of restraints	
Total	4641
Unambiguous	3939
Ambiguous	702
Intra-residue	1603
Sequential (|i-j|=1)	801
Short range (2≤|i-j|≤3)	1037
Medium range (4≤|i-j|<5)	729
Long range(5≤|i-j|)	471
DANGLE φ /ψ	154
RMSD to restraints	
Bonds (Å)	0.005
Angles (°)	0.65

Table 18: Bcl-x_L/$Bak_{I81F}^{i, i+11}$-XL-SO_3H NMR structural ensemble validation statistics. BB – backbone; AH – all heavy.

NOE (Å)	0.051
DANGLE (°)	0.97
Precision (RMSD to mean, Å)	
BB atoms (A.3-19,A.44-156,B.1001-1017)	0.57±0.09
AH atoms (A.3-19,A.44-156,B.1001-1017)	1.40±0.20
Ramachandran plot regions (%)	
Most favoured	88.2%
Additionally allowed	10.6%
Generously allowed	0.6%
Disallowed	0.5%
WHATCHECK Structure Z scores	
First-generation packing quality	-0.020± 0.633
Second-generation packing quality	1.365±1.355
Ramachandran plot appearance	-4.757±0.203
χ_1/ χ_2 rotamer normality	-5.054±0.285
Backbone conformation	-1.201±0.281
WHATCHECK RMS Z-scores	
Bond lengths	1.007±0.002
Bond angles	0.376±0.007
Omega angle restrains	0.928±0.045
Side chain planarity	0.682±0.232
Improper dihedral distribution	0.634±0.107
Inside/Outside distribution	1.039±0.011

Figure 118: (A) Comparison between Bcl-x_L in complex with $Bak_{I81F}^{i, i+11}$-XL-SO_3H (red, 2PL8) and Bak_{wt} (blue, 1BXL). (B) Comparison between Bcl-x_L in complex with $Bak_{I81F}^{i, i+11}$-XL (red, 2PL8) and Bad (blue, 1G5J). Only backbone and crosslinker atoms are shown for the peptides. Overlay of $Bak_{I81F}^{i, i+11}$-XL with (C) Bak_{wt} and (D) Bad based on alignment of the Bcl-x_L/$Bak_{I81F}^{i, i+11}$-XL-SO_3H (2PL8) model with Bcl-x_L/Bak_{wt} (1BXL) or Bcl-x_L/Bad (1G5J). Four peptide residues making important hydrophobic contacts are indicated in pink or blue. Fluorescence anisotropy binding assays

Curiously, the $Bak_{I81F}^{i, i+11}$-XL helix is shifted by approximately ½ helix pitch towards the N-terminus compared to Bak_{wt} (Figure 117, 118), similar to the position of the peptide in the Bad/Bcl-x_L complex.[249] Comparing the RMSD between $Bak_{I81F}^{i, i+11}$-XL and Bak^{wt} or Bad (1.88 and 0.731 respectively) suggests better agreement with the position of Bad rather than Bak^{wt}. The positions of the 'hot-spot' residues of $Bak_{I81F}^{i, i+11}$-XL that make hydrophobic contacts with Bcl-x_L are also different from those observed in the Bak^{wt} complex (Figure 117, 118, 119). Clearly, some of these distortions are imposed by the crosslinker that forces the hydrophobic side chains out of the positions adopted by the wild type peptide. However, instead of the side chains of the peptide adjusting their position to retain optimal hydrophobic interactions, the binding site of Bcl-x_L displays a large degree of plasticity and remodels itself to make more favourable interactions with the constrained peptide (Figure 118 and 119). Such adjustments may be expected for a protein with a large number of similar but distinct binding partners.[15,284-286]

The two salt bridges reported between Bxl-x_L and Bak^{wt} cannot be found in our structure, but the Bxl-x_L/Bak^{wt} structure places Glu 129 and Arg 76 and Arg 139 and Glu 83 further then 7 Å apart. They are closer to each other in our structure yet still not close enough to report a definite electrostatic interaction, indeed Arg 76 is revealed by alanine scanning not to contribute to the binding energy. Moreover, due to significant signal overlap it was not possible to determine chemical shift and NOE signal patterns for the last two C-terminal residues of the $Bak_{I81F}^{i, i+11}$-XL-SO_3H. In the published structure of Bcl-x_L bound to Bak^{wt} most assignments of the peptides C-terminal section are exactly the same as corresponding/similar residues in the N-terminal part suggesting, that the helical conformation, although most probable, was artificially forced.[247] It is difficult to assess which of those two models better represents reality.

Residue	Unambiguous peptide-protein NOEs
Gly 72	0
Cys 73	1
Val 74	20
Gly 75	8
Arg 76	1
Ala 77	2
Leu 78	13
Ala 79	19

Ala 80	0
Phe 81	7
Gly 82	3
Asp 83	0
Cys 84	1
Ile 85	3
Asn 86	0
Arg 87	0

Table 19: Number of peptide-protein NOE restraints by peptide residue

Residue	Backbone RMSD	Sidechain RMSD
Gly 72	2.27	-
Cys 73	1.21	1.26
Val 74	1.03	1.08
Gly 75	0.84	-
Arg 76	0.75	1.75
Ala 77	0.70	0.73
Leu 78	0.54	0.86
Ala 79	0.49	0.56
Ala 80	0.56	0.63
Phe 81	0.57	0.98
Gly 82	0.52	-
Asp 83	0.70	1.08
Cys 84	0.91	0.98
Ile 85	1.39	1.98
Asn 86	2.42	3.51
Arg 87	3.83	6.29
Azobenzene crosslinker	2.04	-

Table 20: Peptide backbone and sidechain RMSD values for all members of the ensemble

On average residues Val 74 and Cys 75 of the crosslinked peptide show dihedral angles slightly outside the alpha helical region of the Ramachandran plot (Figure 116). Those residues comprise the cysteine that is directly attached to the crosslinker and the adjacent valine. One possible explanation is that the presence of the crosslinker forces those residues into unusual conformations but because our structure is the first of this type reported no direct comparison exists to assess the normality of this deviation of in the Ramachandran plot. Nevertheless, the other side of crosslinker attachment (Cys 84) shows no Ramachandran distortion, however as it has been mentioned the C-terminus appear to be more disordered and shows much less contact to the protein therefore making this part of the $Bak_{I81F}^{i,\ i+11}$-XL-

SO_3H less strained overall and perhaps more able to relax into more favourable regions of Ramachandran plot.

Even though the crosslinker shows multiple NOEs to the peptide, no signals to Bcl-x_L were found and in the final ensemble of structures it was found not to interact directly with the protein (Figure 115, 117, 120). Although intuitively Arg 100 and 103 seem possible candidates for an interaction with the crosslinker sulfonate this was not confirmed by NOEs or the distance analysis on the ensemble structures. This finding is in contrast to those with a hydrocarbon staple which was shown to participate in hydrophobic interactions.[57,58,61] The final calculated ensemble of structures shows the crosslinker probing a fairly limited conformational space. It locates closer to the α3/α2 helix than α4 leaving a large open cavity between the crosslinker and helix α4. Conformational stability and positional bias towards α3/α2 is most likely caused by the interaction with Phe 83. Remarkably, and in contrast to our observations of azobenzene-modified peptides binding to RNA, attachment of the fairly bulky crosslinker seems not to interfere with the binding in steric or electronic fashion and only stabilizes the conformation of the peptide.

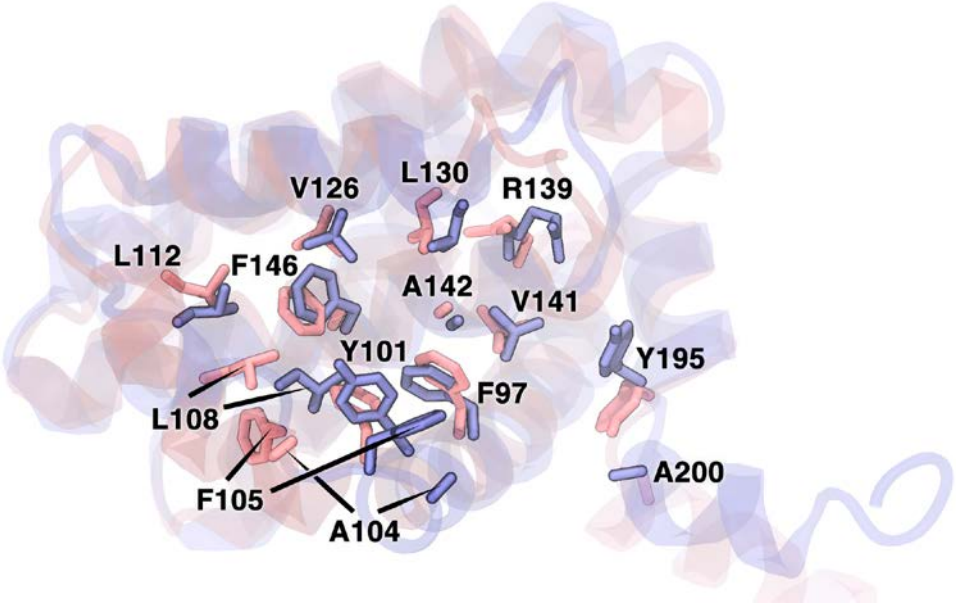

Figure 119: Comparison between BH3 binding site of Bcl-x_L in complex with $Bak_{I81F}^{i,i+11}$-XL-SO_3H (blue, 2PL8) and Bak^{wt} (pink, 1BXL). Hydrophobic residues forming the binding site are highlighted and labelled.

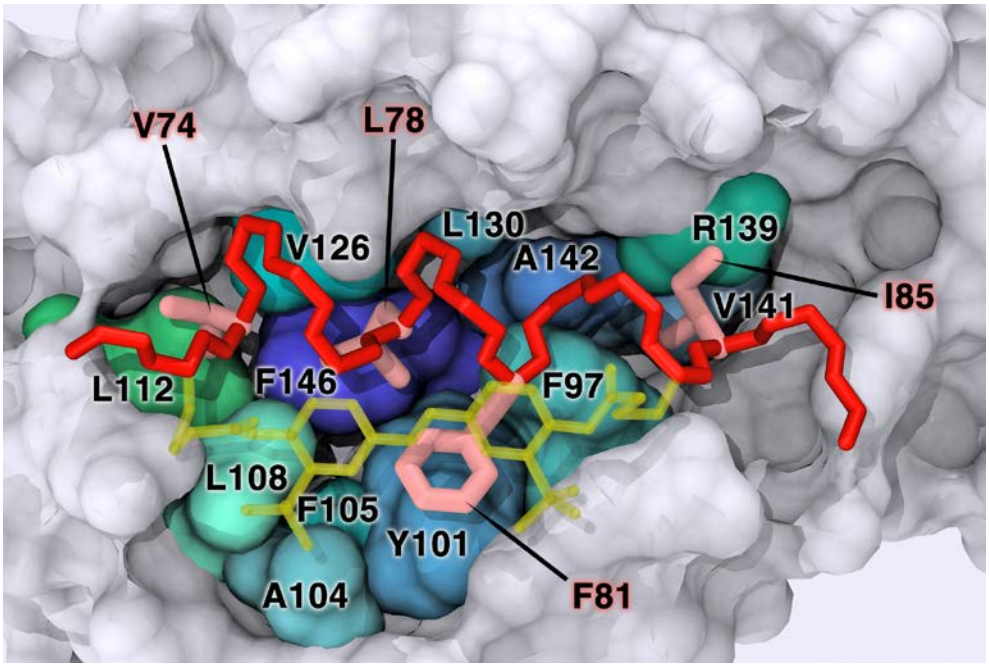

Figure 120: Structure of the Bcl-x_L/$Bak_{I81F}^{i, i+11}$-XL-SO_3H highlighting four peptide residues making important hydrophobic contacts (V74, L78, F81, and I85). Bcl-x_L hydrophobic residues within 4 Å from the peptide are labelled

Fascinatingly, Phe 83 which replaces leucine in the wild type Bak peptide (a modification reported to increase affinity to Bcl-x_L)[287] gave strong NOE signals to azobenzene crosslinker aromatic rings. In the calculated structure the phenylalanine was found outside the canonical hydrophobic pocket and flanked by hydrophobic patch formed by Tyr101, Phe105 and the crosslinker itself. Although it was proposed that the leucine to phenylalanine substitution increases affinity as it better fills the canonical pocket, it appears that in our system Phe 83 preferentially interacts with the crosslinker.

In order to test whether this isoleucine to phenylalanine substitution enhances affinity or decreases $Bak^{i, i+11}$-XL-SO_3H affinity to its target we performed fluorescence anisotropy binding studies. Fluorescence anisotropy binding measurements revealed that $Bak_{I81F}^{i, i+11}$-XL bound to Bcl-x_L with a K_D = 30.5 ± 5.7 nM, an affinity comparable to that measured for $Bak^{i, i+11}$-XL-SO_3H K_D = 15.2 ± 1.3 nM (Figure 121). By contrast, it is known that in the wild type peptide mutation from isoleucine to alanine (I81A) causes decreases from 340 nM to 17000 nM suggesting that the interaction is crucial.[247] Whilst Bak_{I81F} fits well into the binding pocket of the Bak^{wt}/Bcl-x_L structure, the presence of the azobenzene crosslinker and the remodelling of the binding site in the $Bak_{I81F}^{i,i+11}$-XL/Bcl-x_L complex clearly favour this different conformation of Phe81.

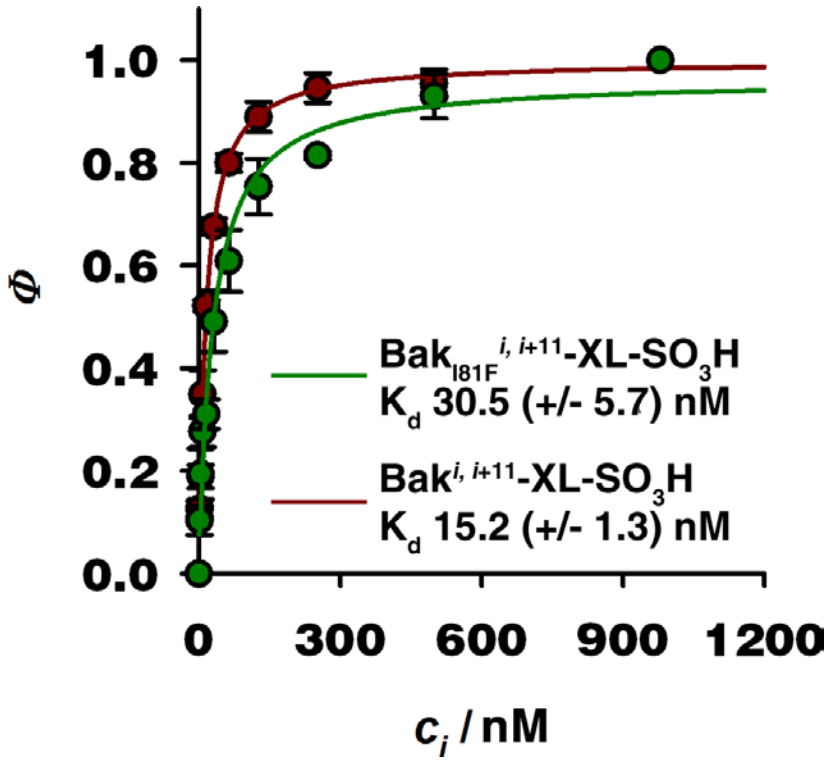

Figure 121: Fluorescence anisotropy binding curves for FAM-Bak$^{i, i+11}$-XL-SO$_3$H (red) and FAM-Bak$_{I81F}^{i, i+11}$-XL-SO$_3$H (green) titrated with Bcl-x$_L$.

2.3.10 Conclusions

Azobenzene-based biophotonically switchable peptides have now been shown to be able to bind a plethora of targets and effectively disrupt interactions between biomacromolecules. Their ability to undergo a photo-activated conformational rearrangement allows external, non-invasive control over their affinity to their respective targets and therefore the interaction in question. Keeping in mind the overwhelming complexity of the protein-protein interactions in a given cell and their crucial role in virtually every aspect of molecular biology, biophotonic switches can provide highly desirable tools to probe these interactions or form a new generation of photodynamic drugs. However, very few structural details are known about the binding of azobenzene peptides to target proteins. The orientation of the crosslinker, its dynamics, positive interactions or clashes with the protein or peptide structures, the effect on the structure of the peptide in its bound form all remained unexplored. The structure of Bak$_{I81F}^{i, i+11}$-XL-SO$_3$H in complex with Bcl-x$_L$ reported here is the first ever solved and begins to address those issues.

The chemical shift perturbation map previously produced from preliminary data supported the hypothesis that the biophotonic Bak binds in the same grove as wild type peptides, but this technique has limitations: The slow exchange regime of the interaction in

question posed a significant challenge and the hotspot residues of this interaction do not necessarily correlate to the highest values of Ω. Instead it was observed that the mobile elements tend to change the environment most and therefore their chemical shift is most perturbed. Completing the NMR assignment of the backbone and sidechains revealed that the most difficult areas to assign are within the binding site itself and the disappearance of some of the peaks there suggest a conformational exchange process taking place that broadens peaks till they are indistinguishable from noise.

The calculation of the structure underlined the issue of multiple orientations of the sulfonated azobenzene crosslinker. Although $Bak_{I81F}^{\ i,\ i+11}$-XL-SO_3H might adopt a single orientation of the crosslinker sulfonates in solution, when bound a different orientation might be more energetically stable and given enough time the system might relax into that conformation. Nevertheless it was difficult to rationalize the position that was found to give the lowest violation and *in silico* energy. One of the possibilities was suggested by the presence of arginines in the vicinity for hydrogen or electrostatic bonding, yet NOEs and the final calculated structure suggest these are improbable. In fact it was observed that the bulky crosslinker does not interfere directly with the binding site. At the same time the crosslinker does cause Phe 81 to somewhat flip out of the canonical binding site and interact with crosslinkers aromatic rings. Therefore although I81F mutation was on multiple occasions related to higher affinity this is not necessarily true for our peptide and might reflect a general tendency of azobenzene peptides. The observation that the crosslinker probes a rather limited conformational space may be a direct effect of the interaction with Phe. Moreover the final position of the crosslinker could be a direct consequence of the same interaction, suggesting that in case of possible steric clashes when designing new peptides it might be possible to bias the crosslinker over one side of the peptide by strategically positioning a Phe residue.

The NMR solution structure of the complex of $Bak_{I81F}^{\ i,\ i+11}$-XL-SO_3H and Bcl-x_L indicates that, while the photo-peptide bound to the canonical binding site, a remodeling of the binding site occurred that led to a shift in α-helical register and a perturbation of the complex structure of $Bak_{I81F}^{\ i,\ i+11}$-XL-SO_3H relative to that observed with the wild-type peptide. Although the peptide was expected to be helical it appears that the attachment of the crosslinker slightly distorts the peptide from an ideal α–helix at the point of attachment. The most interesting observation comes from the comparison between the structure of Bcl-x_L/$Bak_{I81F}^{\ i,\ i+11}$-XL-SO_3H with Bcl-x_L/Bak^{wt}. Not only are the binding sites quite different to each other, but the position of the crosslinked peptide is significantly shifted. The difference

between those two NMR structures is surprisingly greater then between Bcl-x_L/$Bak_{I81F}^{i, i+11}$-XL-SO_3H and free Bcl-x_L. This variability in structure between the two complexes cannot be attributed solely to the different solution conditions for NMR. Though the ability of the Bcl-x_L BH3 binding site to accommodate a whole spectrum of peptides and small molecules by means of conformational flexibility is well established, it is intriguing to imagine that the presence of Phe 81 and the crosslinker cooperatively cause the complex to form a different overall structure and this structure is more reminiscent of activator rather than sensitizer BH3 peptide complexes. Regrettably the Bcl-x_L/Bak^{wt} NMR structure is of rather poor quality and neither full NOE assignment nor even the ensemble of structures is available for closer inspection.[247] If, indeed, the differences arise from a miscalculated structure it would mean that one of the most cited structural papers about Bcl-2 family may be leading scientists astray. Nonetheless, the lack of a crystal structure of the same complex precludes any final statements about the reliability of the structure presented there, despite my structure surpassing the Bcl-x_L/Bak^{wt} in quality indicators it is, after all, of a somewhat different complex. The need to provide an "internal" control for NMR structure solution is apparent, as well as the need to critically evaluate influential structures.

3 Material and Methods

3.1 General remarks

Unless otherwise specified 15 MΩ H_2O was used. RNAse free H_2O (18 MΩ) from an ELGA LabWater purifier for all experiments involving RNA was used and care has been taken to avoid contamination of RNA samples. All chemicals for RNA work used for buffer preparation and purification were designated for RNA work only and were not used before for any other purpose.

3.2 Buffers and Media

3.2.1 FRET buffer[93]

30 mM HEPES
100 mM KCl
40 mM NaCl
10 mM NH_4OOCCH_3
10 mM Gdm:HCl
2 mM $MgCl_2$
0.5 mM EDTA
0.1% Triton X-100
*p*H 7.5
18 MΩ H_2O

3.2.2 CD and UV buffer (for Rev and Bak peptides)

5 mM K_2HPO_4/KH_2PO_4
*p*H 7.5
18 MΩ H_2O

3.2.3 Anisotropy buffer

100 mM Na_2HPO_4/NaH_2PO_4
10 mM NaCl
*p*H 7.5
18 MΩ H_2O

3.2.4 Annealing Buffer

100 mM KCl

2 mM $MgCl_2$

5 mM HEPES

*p*H 7.5

18 MΩ H_2O

3.2.5 Ampicillin Stock

10% Ampicilin

50% H_2O

50% CH_3CH_2OH

Stored at -20 °C

3.2.6 Kanamycin Stock

5% Kanamycin sulphate

Sterile filtered and stored at -20 °C

3.2.7 SDS-PAGE stacking (upper) gel buffer[288]

0.5 M Tris:HCl

*p*H 6.8

3.2.8 SDS-PAGE resolving (lower) gel buffer[288]

1.5 M Tris:HCl

*p*H 8.8

3.2.9 5x SDS-PAGE running (lower) gel buffer[288]

0.125 M Tris base (TRIZMA)

0.96 M Glycine

0.5 % SDS

H_2O to 1 L total volume

3.2.10 Acetic Acid saturated phenol

50% 2 M sodium acetate *p*H 4.2

50% phenol

Phenol phase recovered

3.2.11 Bcl-x_L purification buffers

Base:

100 mM Na_2HPO_4/NaH_2PO_4

500 mM NaCl

10 mM 2-Mercaptoethanol

*p*H 7.5

Lysis buffer is the Base buffer plus:

5 mM Imidazole

1 mM PMSF

Equilibration buffer is the Base buffer plus:

5 mM Imidazole

Wash buffer is the Base buffer plus:

50 mM Imidazole

Elution buffers (A, B, C and D) are the Base buffer plus:

100 mM Imidazole (A)

150 mM Imidazole (B)

200 mM Imidazole (C)

500 mM Imidazole (D)

3.2.12 Crosslinking buffer for XL-SO_3H

50 mM Tris:HCl

2 mM TCEP

*p*H 8.3

3.2.13 Crosslinking buffer for XL-H

25 mM Tris:HCl

2 mM TCEP

50% THF

*p*H 8.3

3.2.14 MALDI matrix

Sat. α-Cyano-hydroxy-cinnamic acid

50% MeCN

50% 18 MΩ H_2O

The suspension is mixed and centrifuged briefly to sediment the excess of matrix

3.2.15 10x M9 minimal salt

71.6 g Na_2HPO_4

30 g KH_2PO_4

5 g NaCl

H_2O to 1L total volume

Autoclaved

3.2.16 100x Trace elements solution

0.3 mM $ZnCl_2$

0.7 mM $FeCl_3 \cdot 6H_2O$

0.06 mM $CuCl_2 \cdot 2H_2O$

0.05 mM $MnCl_2 \cdot 4H_2O$

0.03 mM $Na_2B_4O_7 \cdot 10H_2O$

0.006 mM $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$

H_2O to 1L total volume

3.2.17 M9 minimal media

100 mL M9 salts

10 mL Sterile-filtered 5% NH_4Cl (or $^{15}NH_4Cl$)

20 mL Sterile-filtered 15-20% D-glucose (or uniformly ^{13}C labelled D-glucose)

10 mL Trace elements solution

0.1 mL Sterile-filtered 1 M $CaCl_2$

2 mL Sterile-filtered 1 M $MgSO_4$

H_2O to 1L total volume

3.2.18 5x Tris/Borate/EDTA (TBE) buffer

108 g Tris base (TRIZMA)

55 g Boric acid

40 mL 0.5 M EDTA (*p*H 8.0)

H_2O to bring it to 1 L total volume

3.2.19 50 x Tris/Acetic acid/EDTA (TAE) buffer

242g Tris base (TRIZMA)

57.1 mL Glacial CH_3COOH

100 mL 0.5 M EDTA (*p*H 8.0)

18 MΩ H_2O to 1 L total volume

3.2.20 10x DNA Loading Buffer

2.5 g Ficoll-400

1 ml Tris:HCl (*p*H 7.4)

2 mL 0.5 M EDTA (*p*H 8.0)

*p*H 7.4

15 MΩ H_2O to 10 ml total volume

3.2.21 4x Protein Loading Buffer [288]

2.4 ml 1 M Tris *p*H 6.8 (Upper gel buffer)

0.8 g SDS

4 mL 100% Glycerol

0.01% Bromophenol blue

1 mL 2-mercaptoethanol

H_2O to bring it to 10 ml

3.2.22 Lysogeny broth (LB)[289]

10 g Tryptone.

5 g Yeast extract.

10 g NaCl

H_2O to 1L total volume

Autoclaved and cooled to RT before proceeding.

3.2.23 LB Agar

1 g Tryptone

0.5 g Yeast extract

1 g NaCl

1.5 g Agar

H_2O to 100 ml total volume

Autoclaved, cooled to app. 50 °C, antibiotic added, mixed thoroughly and poured onto plastic Petri dishes.

3.3 Molecular Biology Methods

3.3.1 Cloning of Rev peptide DNA sequences

Sense and antisense oligonucleotides required for the biosynthesis of Rev peptides were acquired from Eurofin MWG Operon. Following phosphorylation by polynucleotide kinase (NEB) at 5' end they were annealed overnight (see Nucleotide Annealing section). Double stranded DNA with overhangs corresponding to those generated by restriction enzyme *AlwN*I (NEB) were ligated into *p*ET-31b plasmid pre-cut with the *AlwN*1 and dephosphorylated by Antarctic phosphatase (NEB). The resulting plasmid was propagated and extracted as described and positive clones were identified by performing *AlwN*1 cleavage and analysing the result *via* agarose electrophoresis.

bRev^{wt} sense:

5' GACACCCGTCAGGCGCGTCGTAACCGTCGTCGTCGTTGGCGTGAACGTCAGCG TATGGACACCCGTCAGGCGCGTCGTAACCGTCGTCGTCGTTGGCGTGAACGTCAG CGTATG 3'

bRev^{wt} antisense:

5' CATACGCTGACGTTCACGCCAACGACGACGACGGTTACGACGCGCCTGACGGG TGTCCATACGCTGACGTTCACGCCAACGACGACGACGGTTACGACGCGCCTGACG GGTGTC 3'

bRev$^{i,\ i+4}$sense:

5' GACACCCGTCAGTGCCGTCGTAACTGCCGTCGTCGTGCGCGTGAACGTCAGCG TATGGACACCCGTCAGTGCCGTCGTAACTGCCGTCGTCGTGCGCGGAACGTCAGC GTATG 3'

***bRev*$^{i,\ i+4}$** antisense:

5’CATACGCTGACGTTCACGCGCACGACGACGGCAGTTACGACGGCACTGACGGG
TGTCCATACGCTGACGTTCACGCGCACGACGACGGCAGTTACGACGGCACTGACG
GGTGTC 3’

***bRev*$^{i,\ i+7}$** sense:

5’ GACACCCGTCAGGCGTGCCGTAACGCGGCGCGTCGTTGCCGTGAACGTCAGCG
TATGGACACCCGTCAGGCGTGCCGTAACGCGGCGCGTCGTTGCCGTGAACGTCAG
CGTATG 3’

***bRev*$^{i,\ i+7}$** antisense:

5’ CATACGCTGACGTTCACGGCAACGACGCGCCGCGTTACGGCACGCCTGACGGG
TGTCCATACGCTGACGTTCACGGCAACGACGCGCCGCGTTACGGCACGCCTGACG
GGTGTC 3’

***bRev*$^{i,\ i+11}$** sense:

5’ GACACCCGTCAGGCGTGCCGTAACCGTGCGCGTCGTGCGCGTGAACGTTGCCG
TATGGACACCCGTCAGGCGTGCCGTAACCGTGCGCGTCGTGCGCGTGAACGTTGC
CGTATG 3’

***bRev*$^{i,\ i+11}$** antisense:

5’ CATACGGCAACGTTCACGCGCACGACGCGCACGGTTACGGCACGCCTGACGGG
TGTCCATACGGCAACGTTCACGCGCACGACGCGCACGGTTACGGCACGCCTGACG
GGTGTC 3’

3.3.2 In vitro transcription

A DNA RRE IIB template containing T7 promoter sequence was ordered from Eurofins. RRE IIB *in vitro* transcription was achieved according to the Megascript™ (Albion) protocol. Briefly, 20 µl of enzymatic reaction mixture was assembled containing four unlabeled NTPs, annealed RRE IIB template (below), 1x reaction buffer, T7 polymerase (enzyme mix). The mixture was supplemented with trace amount of ^{32}P-α-UTP. A control reaction using pTRI-Xef as a template was assembled separately.

Template for RRE IIB sense:
5'AATTTAATACGACTCACTATA'GGGTCTGGGCGCAGCGCAAGCTGACGGTACAG GCC 3'

Template for RRE IIB antisense:
5'GGCCTGTACCGTCAGCTTGCGCTGCGCCCAGACCCTATAGTGAGTCGTATTAAA TT3'

The reaction was incubated at 37 °C for 4 or 16 h. At the end the reaction DNAase I was added and the mixture was incubated for further 30 min, the mixture was then quenched with 5M sodium acetate and 100 mM EDTA (130 µl), and purified as described in the Purification Methods section.

3.4 Protein Biosynthesis

3.4.1 Purification Methods

3.4.1.1 Expressed peptides purification

Cells were resuspended in potassium phosphate buffer (50 mM, pH 7.0) with NaCl (100 mM) and sonicated on ice with 3 x 30 pulses (1 sec pulse at 40 % intensity, 1 sec pause) then centrifuged (20 min, 4°C, 12000 x g). The supernatant was removed and the pellet was resuspended in potassium phosphate buffer (50 mM, pH 7.0) with NaCl (100 mM) and 0.5 % Tween 20. It was stirred for 2 h and centrifuged (20 min, 4°C, 12000 x g). The supernatant was removed and the pellet was resuspended in potassium phosphate buffer (50 mM, pH 7.0) with NaCl (1 M). It was stirred for 2 h and centrifuged (20 min, 4°C, 12000 x g). The pellet was dissolved in guanidine hydrochloride (6 M) and stirred over night at 4 °C then centrifuged (30 min, 4°C, 25000 x g). The supernatant was dialysed at 4 °C against H_2O (3 L) using a 12-14 kDa cut off dialysis membrane. The suspension was centrifuged (20 min, 4 °C, 12000 x g) to remove precipitated protein, the pellet was resuspended in 15 ml 80 % formic acid and transferred to a 50 ml round bottom flask and CNBr (0.6 g) was added. The flask was wrapped in an aluminium foil and suspension was stirred under N_2 atmosphere for 18-22 h. The flask was attached to rotary evaporator and the solution was evaporated at 40 °C until a gel like consistency. The gel was resuspended in monobasic potassium phosphate (10 ml, 25

mM) containing one average crystal of TCEP adjusted to *p*H 7.5 with KOH. The suspension was stirred for 1 h, and the *p*H was adjusted again to 7.5. Finally it was stirred for a further 1 h and centrifuged (20 min, 4°C, 25000 x g). The supernatant was applied on C18 reverse phase column on HPLC and purified by a 10-100% methanol + 0.1% TFA gradient.

3.4.1.2 Synthesized peptide purification and purity

Peptides were purified by reversed phase HPLC using a Phemonenex Gemini 10 μm, 110 Å, 10×250 mm C18 column with gradients from 100% water (+ 0.1%) TFA to 100% acetonitrile (+ 0.1%) over 100 minutes at a flow rate of 5 mL. For analysing the purity of product peptides a Dionex Acclaim, 3 μm, 120 A, 4.6×150 mm C18 column was used with a gradient from 100% water (+ 0.1% TFA) to 100% acetonitrile (+ 0.1% TFA) over 50 minutes (1 mL/min). Peptide masses were verified as described below.

3.4.1.3 Bcl-x_L biosynthesis and purification

A *p*ET 19 plasmid containing a gene coding for loop truncated (1-209 Δ45-84, Δ210-233) Bcl-x_L was used to transform *E. coli* BL(21)DE3 strain. This bacterial culture was grown at 37 °C until OD_{600} reached 0.8, then expression was induced by addition of IPTG (0.5 mM). The culture was harvested after 6 h. Following one freeze-thaw cycle, the bacterial pellet was resuspended in sodium phosphate (100 mM), NaCl (500 mM), imidazole (5 mM), 2-mercaptoethanol (10 mM) *p*H 7.5 buffer. Cells were lysed by sonication and cell debris was removed by centrifugation. The supernatant containing Bcl-x_L was subjected to NiNTA affinity chromatography and the protein was eluted by applying a step gradient of imidazole (50, 100, 150, 200, 500 mM). Fractions containing Bcl-x_L were pooled and dialysed against buffer containing sodium phosphate (5 mM), 2-mercaptoethanol (5 mM). Protein was mixed with equimolar $Bak_{I81F}^{\ i,\ i+11}$-XL-SO_3H and concentrated to app. 1 mM. For binding studies expression was carried out in LB media and for the NMR sample preparation M9 media supplemented with $^{15}NH_4Cl$ and U^{13}C-D-glucose as sole sources of nitrogen and carbon respectively was used.

3.4.2 Purification of in vitro transcription products

Quenched *in vitro* transcription reaction mixtures were mixed with sodium acetate (150 μl, 0.3 M, *p*H 5.0) saturated phenol/chloroform mixture (1:1) and mixed vigorously for 1 min. The mix was centrifuged (1 min, 16 000 g) and the upper phase was recovered. It was mixed with pure chloroform (150 μl), mixed vigorously, centrifuged (1 min, 16 000 g) and

the upper phase was recovered. 2 volumes of ice-cold isopropanol were added to the upper phase and it was incubated for 2 h at -20 °C. The mixture was centrifuged (10 min, 0 °C, 16 000 g), the supernatant was discarded and the precipitate containing RRE IIB was resuspended in H_2O and stored at -80 °C till needed.

3.5 Analytical Methods

3.5.1 General acrylamide gels preparation protocol[288]

For SDS-PAGE an acrylamide:N, N'-methylenebisacrylamide ratio of 37.5:1 was used. For EMSA and urea gels this ratio was 29:1. 30% acrylamide mix, gel buffer stock (stacking, resolving, 1x TBE or 0.5x TBE) and 15 MΩ water were mixed and for SDS-PAGE the mix was supplemented with 10% SDS stock solution to 0.1% final concentration. 10% (w/v) ammonium persulfate stock solution was added to a final concentration of 0.1% and mixed thoroughly, followed by TEMED to a final concentration of 0.01% with further mixing. A gel was immediately cast. For SDS-PAGE resolving gels, the gel was first covered with isopropanol layer, which was removed after polymerization, washed, excess water was drained and the stacking gel was cast over the resolving gel.

3.5.2 SDS-PAGE[288]

Samples were prepared by mixing the solutions to be analyzed with protein loading buffer and incubating them for 10 min at 95 °C. Samples were then briefly centrifuged, allowed to cool to room temperature and loaded into gel wells. For all protein analysis 15% SDS-PAGE gels were used. Electrophoresis was carried out at 180 V for 1 h or until bromophenol blue migrated out of the gel.

3.5.3 EMSA

5% Acrylamide gels were prepared as described in the gel preparation section. Gels were pre-run in 0.5x TBE buffer for 1 h at 4 °C, 50 V. Buffer was circulated between the upper and lower chambers. Samples were prepared in one of the EMSA sample buffers (Buffers and Media Section) containing approximately 1 nM of RRE IIB in 25 µl and increasing amounts of one of the bRev peptides (0, 25, 50, 100, 200, 500, 1000, 2000, 5000 or 10000 nM). After 30 min incubation, the entire volume was loaded into the gel wells with a flat-end Hamilton syringe and the gel was run for 3 h at 4 °C, 225 V with buffer circulated between the upper and lower chambers. Gels were transferred to 3M filter paper, covered with

cling film and dried on a gel drier for 10 h at 65 °C. A phosphorimager was exposed to the gel for 10 h and subsequently scanned on Typhoon Gel Imager.

3.5.4 Urea gels

20% acrylamide gels with 8M urea and 1x TBE were prepared as described in previously. Samples of RRE IIb or pTRI-Xef were prepared in formamide gel loading buffer supplied with Megascript™ (Ambion) and were heated to 80 °C for 5 min, then immediately cooled to 4 °C before loading into pre-rinsed gel wells. Electrophoresis was run for 3 h at 300 V in 1x TBE. Excess buffer was removed from the gel with paper towels and the gel was tightly wrapped with cling film. A phosphorimager was exposed to the gel for 1 h and then scanned on Typhoon Gel Imager.

3.5.5 DNA agarose gels

Agarose gels were prepared by dissolving 1 - 1.5% agarose in 1x TEA buffer. DNA samples were mixed with DNA loading buffer and transferred into gel wells. Electrophoresis was run at 80 V for 45 to 60 min. Gel was placed in a solution of 1x TEA buffer and 5 mg/L ethidium bromide for staining and visualized under a transiluminator.

3.5.6 Fluorescence spectroscopy – Rev/RRE binding studies

Cy3 labelled RRE sequence (Cy3 5' GGUCUGGGCGCAGCGCAAGCUGACGGUACAGGCC 3') was purchased from Eurofins. All measurements were performed on LS 55 Luminescence Spectrometer (Perkin Elmer) equipped with F 25 temperature control unit (Julabo). The temperature was kept constant at 15 °C and solutions were allowed to equilibrate for 20 minutes before use. The cuvette was a standard 3 mL quartz cuvette and the initial volume of all assays was 2.5 mL. Fluorecsence assays were performed in FRET buffer at 15 °C. This buffer composition simulates the high ionic strength and the ionic composition found in nucleus and includes detergent to minimize possible adherence of the peptides to the walls of the cuvette.

At each measurement point three complete spectra from 510 to 535 nm were acquired and averaged (excitation/emission slits: 5/10 nm, scan speed: 50 nm/s, excitation wavelength: 495 nm). Concentrations and fluorescence intensities were corrected for the effects of dilution. All titrations were fitted independently in SigmaPlot and reported K_d values are an average of those fits. To estimate apparent binding affinities from IC50 values Cheng-Prussof equation was used.

$$K_d = \frac{IC50}{1 + \frac{[FAM\ Rev\ wt]}{K_d wt}}$$

Equation 8: Where K_d is an apparent dissociation constant; IC50 is the concentration of antagonist required to displace 50% of ligand; K_d wt is the apparent dissociation constant of FAM Rev^{wt} and [FAM Rev-wt] is the concentration of FAM Rev^{wt}.

The equation used to calculate IC50 values was a linear equivalent of the Hill slope equation used for displacement assays data after $\log_{10}$ transformation of concentration values and is given below.

$$NpF = \frac{maxNpF}{1 + (\frac{[Rev]}{IC50})^B}$$

Equation 9: Where [Rev] is the concentration of added Rev peptide; IC50 is the concentration of antagonist required to displace 50% of ligand; B is the equivalent of Hill slope; NpF is the value of fluorescence normalized to predicted fluorescence values at a saturation point and maxNpF is the maximum value of NpF, that by definition must equal 1.

The direct binding assay data was fitted to the single site binding equation.

$$NF = \frac{max\ NF}{1 + (\frac{[RRE]}{K_d})}$$

Equation 10: Where NF is the values of normalized fluorescence; [RRE] is the concentration of added Cy3 RRE; K_d is an apparent dissociation constant; maxNF is the predicted maximal values of normalized fluorescence

For data visualisation corresponding normalised points were averaged and standard deviation of the mean for three measurements at each data point was calculated, plots were then normalized to calculated saturation values (Sigma Plot). Due to high concentration of Cy3 RRE needed for experiments using FAM-labelled non-crosslinked peptides it was necessary to make corrections for a small degree of Cy3 emission overlapping with the FAM emission. None of those mathematical transformations effect data content and estimated IC50 or K_d.

3.5.7 UV spectroscopy

Measurements were performed on Jasco V-660 spectrometer equipped with DAC-743 R temperature control unit (JASCO Inc.) and AWC 100 heat exchanger (Julabo). The relaxation of the *cis* to *trans* form was observed by monitoring the recovery in absorption at 363 nm due to increasing quantities of the *trans* isomer. Absorbance readings were recorded every 120 seconds for total of 40 minutes (37, 30 °C), 80 minutes (25 °C) or 120 minutes (20, 15 °C) to provide suitable data to achieve good quality fits to equation below.

3.5.8 CD spectroscopy

All measurements were performed on Chirascan (Applied Photophysics) equipped with a TC 125 temperature controller (Quantum Northwest) and AWC 100 heat exchanger (Julabo). The temperature was held at 15 °C and solutions were allowed to equilibrate for 10 minutes before measurements except for those of the irradiated states of peptides, where they were recorded immediately after the exposure to UV light. Data is expressed as mean residue ellipticity (MRE) calculated using the following equation:

$$\mathbf{MRE} = \Theta/(\mathbf{n}\ \mathbf{c}\ \mathbf{l})$$

Equation 11: MRE Where Θ is measured ellipticity in mdeg; n is the number of amino acids minus 1; c is concentration in mol/L and l is the pathlength in mm

3.5.9 Mass Spectrometry

The masses of Rev peptides were checked by matrix assisted laser desorption ionisation using a MALDI micro MX, MALDI TOF MS (Waters) in reflectron mode. α-cyano-hydroxy-cinnaminic acid in 1:1 acetonitrile water (2 μL) was used as the matrix And mixed on a spot on a MALDI plate with analyte solution (2 μL).

3.5.10 NMR Spectroscopy

NMR spectra were acquired at 25 °C on 600 MHz (with room temperature or cryogenically cooled probe) and 800 Mhz (with cryogenically cooled probe) Varian Inova instruments. In all cases ^{13}C and ^{15}N labelled Bcl-x_L was used. For the purpose of backbone assignments a following spectra were collected ^{1}H-^{15}N-HSQC, HNCA, HN(CO)CA, CACBNH, CACB(CO)NH, HNCO, HN(CA)CO; sidechain assignment was facilitated by CCOHN, HCCOHN and HCCH-TOCSY, ^{13}C-HSQC and aided by 3D ^{1}H-^{15}N NOESY-HSQC and 3D ^{1}H-^{13}C NOESY-HSQC.

In order to assign the non-isotopically labelled $Bak_{I81F}^{\ i,\ i+11}$-XL-SO_3H we collected a set of ^{13}C, ^{15}N or just ^{15}N filtered ^{1}H-^{1}H NOESY, ^{1}H-^{1}H TOCSY. To assist the assignment of peptides aromatic residues a sample in D_2O was prepared and ^{13}C, ^{15}N filtered ^{1}H-^{1}H NOESY, ^{1}H-^{1}H TOCSY spectra were acquired. Interproton protein-protein distance restrains were derived from 3D ^{1}H-^{15}N NOESY-HSQC and ^{1}H-^{13}C NOESY-HSQC. Interproton peptide-peptide distance restrains were derived from ^{13}C, ^{15}N filtered ^{1}H-^{1}H NOESY. Interproton peptide-protein distance restrains were derived from F2F filtered ^{1}H-^{1}H NOESY, 3D ^{1}H,^{15}N NOESY-HSQC and 3D ^{1}H,^{13}C NOESY-HSQC spectra. Dihedral restraints were calculated using DANGLE and manually checked for consistency.[290] All spectra were processed in NMRPipe and analysed in ccpnNMR Analysis 2.1.5.[279,291]

3.6 Synthetic Chemistry Methods

3.6.1 Peptides Synthesis

All peptides were synthesized according to standard fluorenylmethylcarbamoyl (Fmoc) solid phase synthesis protocols using a CEM Liberty microwave-assisted peptide synthesizer. Protected amino acids with trityl (Trt), *tert*-butyl (^{t}Bu), butoxycarbonyl (Boc) or 2,2,4,6,7- pentamethyldihydrobenzofurane (Pbf) sidechain protecting groups as required, O-benzotriazole-N,N,N',N'-tetramethyl-uronium-hexafluoro-phosphate (HBTU), hydroxybenzotriazole (HOBt), N-methylpyrrolidinone (NMP) and dimethylformamide (DMF) were purchased from AGTC Bioproducts. Dichloromethane (DCM), trifluoroacetic acid (TFA) and diethyl ether were sourced from Fisher. Piperidine, acetic anhydride, tri-*iso*-propylsilane (TIS), N,N-diisopropylethylamine (DIEA), 4-(2-hydroxyethyl)-1-piperizineethanesulfonic acid (HEPES) and triscarboxyethylphosphine (TCEP) were purchased from Sigma Aldrich. Rink Amide resin (0.72 mMol/g) and 5(6)-carboxyfluorescein were purchased from NovaBiochem. The following solutions were used for peptide synthesis:

Activator	0.45M HBTU in DMF
Deprotection Mix	20% Piperidine, 0.1M HOBt in DMF
Activator Base	2M DIEA in NMP
Capping Solution	20% Acetic acid anhydride in DMF
Amino Acid Solutions	0.1 M Fmoc-protected amino acids in DMF;

Fmoc-Ala-OH
Fmoc-Arg(Pbf)-OH
Fmoc-Asp(^{t}Bu)-OH
Fmoc-Asn(Trt)-OH
Fmoc-Cys(Trt)-OH
Fmoc-Glu(^{t}Bu)-OH
Fmoc-Gln(Trt)-OH
Fmoc-Gly-OH
Fmoc-Ile-OH
Fmoc-Leu-OH
Fmoc-Phe-OH
Fmoc-Thr(^{t}Bu)-OH
Fmoc-Trp(Boc)-OH

Fmoc-Val-OH

Cleavage Cocktail 95% TFA, 2.5% TIS, 2.5% water

Resin (140 mg, 0.1 mMol scale) was transferred to the reaction chamber in a 1:1 mixture of DMF and DCM and allowed to swell for 15 minutes.

Following swelling and at each following step, deprotection was achieved by addition to washed resin of 7 mL of the deprotection mix described above, followed by microwave-assisted deprotection. Heating was provided for 180 seconds using 20 W to achieve a maximum temperature of 75 °C with 3 seconds of nitrogen bubbling every 7 seconds. The resin was then cooled and washed with DMF (3 × 5 mL).

Such is the difficult of constructing poly-arginine peptides that double couplings were used throughout. Coupling of amino acids was achieved by adding the required amino acid solution (2.5 mL), activator (1 mL) and of activator base (0.5 mL) followed by microwave-assisted coupling. The heating regime was varied according to the amino acid:

Arginine: No heating was provided for 1800 seconds, then heating was provided for 600 seconds using 20 W to achieve a maximum temperature of 75 °C. The resin was agitated throughout with 3 seconds of nitrogen bubbling every 7 seconds. The resin was then cooled and washed with of DMF (3 × 5 mL). Due to the hindering nature of the Pbf sidechain protecting group, arginine couplings were followed by a capping step. Capping solution (7 mL) was added allowed to react for 600 seconds without additional heating. The resin was then washed with of DMF (2 × 8 mL).

Cysteine: No heating was provided for 300 seconds, then heating was provided for 900 seconds using 20 W to achieve a maximum temperature of 50 °C. The resin was agitated throughout with 3 seconds of nitrogen bubbling every 7 seconds. The resin was then cooled and washed with of DMF (3 × 5 mL).

All other amino acids: Heating was provided for 600 seconds using 20 W to achieve a maximum temperature of 75 °C with 3 seconds of nitrogen bubbling every 7 seconds. The resin was then cooled and washed with of DMF (3 × 5 mL).

Final deprotection of the resin-bound peptide chain utilised a two-part deprotection. Deprotection mixture (7 mL) was added and heated using 20 W to achieve a maximum

temperature of 75 °C for 180 seconds. The solvent was then drained and the resin was washed with DMF (3 × 5 mL). A second addition of the deprotection mix (7 mL) was added and heated for 180 seconds using 20 W to achieve a maximum temperature of 75 °C with 3 seconds of nitrogen bubbling every 7 seconds. The resin was then cooled and again washed with DMF (3 × 5 mL).

When required, FAM labelling was achieved as previously reported.[162]

The resin was transferred to 50 mL Falcon tube, drained on a sintered glass filter and suspended in cleavage cocktail (20 mL). The cleavage reaction was carried out at room temperature for 6 h with shaking, then the resin was filtered and the filtrate was evaporated under nitrogen flow until the volume was reduced to approximately 1 mL. This suspension was diluted with 49 mL of cold diethyl ether to precipitate the product. After chilling overnight at -20 °C, the resulting precipitate was spun down, resuspended in water (+0.5 % TFA), filtered and subjected to HPLC purification.

3.6.2 Crosslinker Synthesis

Synthesis of XL-SO_3H was achieved as described previously.[161,292] XL-H was synthesised as detailed in the literature,[160] but the last step of chloride substitution with iodine was omitted.

3.6.3 Crosslinking

Crosslinking with XL-SO_3H was achieved by dissolving peptides at 1 mg/mL in *p*H 8.3 with Tris:HCl (50 mM) containing TCEP (2 mM). This buffer was also used to disolve the XL-SO_3H crosslinker at 10x molar excess over the peptide (app. 1 mg/ml). The crosslinker solution was transferred in three equal aliquots to the peptide solution. After each addition reaction was allowed to proceed for 2h at room temperature and finally the temperature was decreased to 4 °C. After an overnight incubation the reaction mixture was spun down and the supernatant was subjected to HPLC purification.

Crosslinking with XL-H was accomplished by dissolving peptides at 1 mg/mL in water/THF mixture (1:1) buffered to *p*H 8.3 with Tris:HCl (50 mM) containing TCEP (2 mM). The same solvent was used to dissolve the XL-H crosslinker at concentration equivalent to a 10× molar excess over the peptide to be crosslinked. A single equivalent was added over the course of 2 hours, the remainder was added in further aliqouts over 3 hours of

the solution volume. If precipitation occurred, additional water/TFA was added. After overnight reaction, solutions were freeze-dried, resuspended in water containing 0.5% TFA, spun down and the supernatant was purified by HPLC. The masses of the crosslinked peptides were verified as described below.

3.6.4 NMR statistics

Number of restraints and RMSD of restraints values were derived from final CNS 1.21 statistics.[275,293] The number of proton, nitrogen and carbon assignments was taken from Analysis 2.1.5.[279] All remaining statistics were calculated with the help of iCING server that internally or externally uses PROCHECK and WHATCHECK for quality statistics.[294,295] Z-scores were calculated by comparing appropriate statistic with the average derived from a database of high quality crystal structures.

3.7 Other Methods

3.7.1 UV light source

All irradiations were performed with a UV-P 280 light source (UV Light Technology Ltd.) coupled to 360 nm band pass filter (10 nm bandwidth). The lamp was allowed to warm up for 30 min before use. The achievement of photostationary state was monitored by Nanodrop 1000 spectrophotometer (Thermo Scientific) and always occurred in less than the standard 3 minute irradiation time. For binding assays peptides solutions were irradiated in thin-walled PCR tubes, which were showed to be transparent for the wavelength of light used.

3.7.2 Nucleotide Annealing

Equimolar amounts of each strand were prepared in the annealing buffer (100 mM KCl, 2 mM $MgCl_2$, 5 mM HEPES p*H* 7.5). Annealing was carried out in a PCR thermocycler with heated lid. Samples were initially warmed up to 95 °C for 10 min and then slow cooling program to 15 °C over 16 h was applied.

4 References

(1) Hao, B.; Gong, W.; Ferguson, T. K.; James, C. M.; Krzycki, J. A.; Chan, M. K. *Science* **2002**, *296*, 1462.
(2) Hatfield, D. L.; Gladyshev, V. N. *Mol. Cell. Biol.* **2002**, *22*, 3565.
(3) Zinoni, F.; Birkmann, A.; Stadtman, T. C.; Bock, A. *Proceedings of the National Academy of Sciences of the United States of America* **1986**, *83*, 4650.
(4) Whitford, D. *Proteins : structure and function*; J. Wiley & Sons: Hoboken, NJ, 2005.
(5) Grigoriev, A. *Nucleic Acids Res.* **2003**, *31*, 4157.
(6) Stumpf, M. P.; Thorne, T.; de Silva, E.; Stewart, R.; An, H. J.; Lappe, M.; Wiuf, C. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 6959.
(7) Lo Conte, L.; Chothia, C.; Janin, J. *J. Mol. Biol.* **1999**, *285*, 2177.
(8) Chene, P. *ChemMedChem* **2006**, *1*, 400.
(9) Keskin, O.; Ma, B.; Nussinov, R. *J. Mol. Biol.* **2005**, *345*, 1281.
(10) Shoemaker, B. A.; Panchenko, A. R. *PLoS Comp. Biol.* **2007**, *3*, e42.
(11) Jochim, A. L.; Arora, P. S. *ACS Chem. Biol.* **2010**, *5*, 919.
(12) Henchey, L. K.; Jochim, A. L.; Arora, P. S. *Curr. Opin. Chem. Biol.* **2008**, *12*, 692.
(13) Oltersdorf, T.; Elmore, S. W.; Shoemaker, A. R.; Armstrong, R. C.; Augeri, D. J.; Belli, B. A.; Bruncko, M.; Deckwerth, T. L.; Dinges, J.; Hajduk, P. J.; Joseph, M. K.; Kitada, S.; Korsmeyer, S. J.; Kunzer, A. R.; Letai, A.; Li, C.; Mitten, M. J.; Nettesheim, D. G.; Ng, S.; Nimmer, P. M.; O'Connor, J. M.; Oleksijew, A.; Petros, A. M.; Reed, J. C.; Shen, W.; Tahir, S. K.; Thompson, C. B.; Tomaselli, K. J.; Wang, B.; Wendt, M. D.; Zhang, H.; Fesik, S. W.; Rosenberg, S. H. *Nature* **2005**, *435*, 677.
(14) Hikita, H.; Takehara, T.; Shimizu, S.; Kodama, T.; Shigekawa, M.; Iwase, K.; Hosui, A.; Miyagi, T.; Tatsumi, T.; Ishida, H.; Li, W.; Kanto, T.; Hiramatsu, N.; Hayashi, N. *Hepatology* **2010**, *52*, 1310.
(15) Lee, E. F.; Czabotar, P. E.; Yang, H.; Sleebs, B. E.; Lessene, G.; Colman, P. M.; Smith, B. J.; Fairlie, W. D. *J. Biol. Chem.* **2009**, *284*, 30508.
(16) De Benedetti, A.; Graff, J. R. *Oncogene* **2004**, *23*, 3189.
(17) Marcotrigiano, J.; Gingras, A. C.; Sonenberg, N.; Burley, S. K. *Mol. Cell* **1999**, *3*, 707.
(18) Cencic, R.; Hall, D. R.; Robert, F.; Du, Y.; Min, J.; Li, L.; Qui, M.; Lewis, I.; Kurtkaya, S.; Dingledine, R.; Fu, H.; Kozakov, D.; Vajda, S.; Pelletier, J. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 1046.
(19) Chen, X.; Kopecky, D. J.; Mihalic, J.; Jeffries, S.; Min, X.; Heath, J.; Deignan, J.; Lai, S.; Fu, Z.; Guimaraes, C.; Shen, S.; Li, S.; Johnstone, S.; Thibault, S.; Xu, H.; Cardozo, M.; Shen, W.; Walker, N.; Kayser, F.; Wang, Z. *J. Med. Chem.* **2012**, *55*, 3837.
(20) Mosyak, L.; Zhang, Y.; Glasfeld, E.; Haney, S.; Stahl, M.; Seehra, J.; Somers, W. S. *EMBO J.* **2000**, *19*, 3179.
(21) Kozakov, D.; Hall, D. R.; Chuang, G. Y.; Cencic, R.; Brenke, R.; Grove, L. E.; Beglov, D.; Pelletier, J.; Whitty, A.; Vajda, S. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 13528.
(22) Kenny, C. H.; Ding, W.; Kelleher, K.; Benard, S.; Dushin, E. G.; Sutherland, A. G.; Mosyak, L.; Kriz, R.; Ellestad, G. *Anal. Biochem.* **2003**, *323*, 224.
(23) Rush, T. S., 3rd; Grant, J. A.; Mosyak, L.; Nicholls, A. *J. Med. Chem.* **2005**, *48*, 1489.
(24) Jennings, L. D.; Foreman, K. W.; Rush, T. S., 3rd; Tsao, D. H.; Mosyak, L.; Li, Y.; Sukhdeo, M. N.; Ding, W.; Dushin, E. G.; Kenny, C. H.; Moghazeh, S. L.; Petersen, P. J.; Ruzin, A. V.; Tuckman, M.; Sutherland, A. G. *Bioorganic & medicinal chemistry letters* **2004**, *14*, 1427.

(25) Lee, J. H.; Zhang, Q.; Jo, S.; Chai, S. C.; Oh, M.; Im, W.; Lu, H.; Lim, H. S. *J. Am. Chem. Soc.* **2011**, *133*, 676.
(26) Rew, Y.; Sun, D.; Gonzalez-Lopez De Turiso, F.; Bartberger, M. D.; Beck, H. P.; Canon, J.; Chen, A.; Chow, D.; Deignan, J.; Fox, B. M.; Gustin, D.; Huang, X.; Jiang, M.; Jiao, X.; Jin, L.; Kayser, F.; Kopecky, D. J.; Li, Y.; Lo, M. C.; Long, A. M.; Michelsen, K.; Oliner, J. D.; Osgood, T.; Ragains, M.; Saiki, A. Y.; Schneider, S.; Toteva, M.; Yakowec, P.; Yan, X.; Ye, Q.; Yu, D.; Zhao, X.; Zhou, J.; Medina, J. C.; Olson, S. H. *J. Med. Chem.* **2012**, *55*, 4936.
(27) Orner, B. P.; Ernst, J. T.; Hamilton, A. D. *J. Am. Chem. Soc.* **2001**, *123*, 5382.
(28) Davis, J. M.; Tsou, L. K.; Hamilton, A. D. *Chem. Soc. Rev.* **2007**, *36*, 326.
(29) Kutzki, O.; Park, H. S.; Ernst, J. T.; Orner, B. P.; Yin, H.; Hamilton, A. D. *J. Am. Chem. Soc.* **2002**, *124*, 11838.
(30) Ernst, J. T.; Kutzki, O.; Debnath, A. K.; Jiang, S.; Lu, H.; Hamilton, A. D. *Angew. Chem. Int. Ed.* **2002**, *41*, 278.
(31) Chen, L.; Yin, H.; Farooqi, B.; Sebti, S.; Hamilton, A. D.; Chen, J. *Mol. Cancer Ther.* **2005**, *4*, 1019.
(32) Yin, H.; Lee, G. I.; Park, H. S.; Payne, G. A.; Rodriguez, J. M.; Sebti, S. M.; Hamilton, A. D. *Angew. Chem. Int. Ed.* **2005**, *44*, 2704.
(33) Ernst, J. T.; Becerril, J.; Park, H. S.; Yin, H.; Hamilton, A. D. *Angew. Chem. Int. Ed.* **2003**, *42*, 535.
(34) Yap, J. L.; Cao, X.; Vanommeslaeghe, K.; Jung, K. Y.; Peddaboina, C.; Wilder, P. T.; Nan, A.; MacKerell, A. D., Jr.; Smythe, W. R.; Fletcher, S. *Org. Biomol. Chem* **2012**, *10*, 2928.
(35) Rodriguez, J. M.; Ross, N. T.; Katt, W. P.; Dhar, D.; Lee, G. I.; Hamilton, A. D. *ChemMedChem* **2009**, *4*, 649.
(36) Shaginian, A.; Whitby, L. R.; Hong, S.; Hwang, I.; Farooqi, B.; Searcey, M.; Chen, J.; Vogt, P. K.; Boger, D. L. *J. Am. Chem. Soc.* **2009**, *131*, 5564.
(37) Yin, H.; Lee, G. I.; Sedey, K. A.; Rodriguez, J. M.; Wang, H. G.; Sebti, S. M.; Hamilton, A. D. *J. Am. Chem. Soc.* **2005**, *127*, 5463.
(38) Seebach, D.; Gardiner, J. *Acc. Chem. Res.* **2008**, *41*, 1366.
(39) Frackenpohl, J.; Arvidsson, P. I.; Schreiber, J. V.; Seebach, D. *ChemBioChem* **2001**, *2*, 445.
(40) Martinek, T. A.; Fulop, F. *Chem. Soc. Rev.* **2012**, *41*, 687.
(41) Guo, L.; Almeida, A. M.; Zhang, W.; Reidenbach, A. G.; Choi, S. H.; Guzei, I. A.; Gellman, S. H. *J. Am. Chem. Soc.* **2010**, *132*, 7868.
(42) Price, J. L.; Horne, W. S.; Gellman, S. H. *J. Am. Chem. Soc.* **2010**, *132*, 12378.
(43) Horne, W. S.; Gellman, S. H. *Acc. Chem. Res.* **2008**, *41*, 1399.
(44) Sawada, T.; Gellman, S. H. *J. Am. Chem. Soc.* **2011**, *133*, 7336.
(45) Lee, E. F.; Sadowsky, J. D.; Smith, B. J.; Czabotar, P. E.; Peterson-Kaufman, K. J.; Colman, P. M.; Gellman, S. H.; Fairlie, W. D. *Angew. Chem. Int. Ed.* **2009**, *48*, 4318.
(46) Boersma, M. D.; Haase, H. S.; Peterson-Kaufman, K. J.; Lee, E. F.; Clarke, O. B.; Colman, P. M.; Smith, B. J.; Horne, W. S.; Fairlie, W. D.; Gellman, S. H. *J. Am. Chem. Soc.* **2012**, *134*, 315.
(47) Pellegrini, M.; Royo, M.; Chorev, M.; Mierke, D. F. *J. Pept. Res.* **1997**, *49*, 404.
(48) Leduc, A. M.; Trent, J. O.; Wittliff, J. L.; Bramlett, K. S.; Briggs, S. L.; Chirgadze, N. Y.; Wang, Y.; Burris, T. P.; Spatola, A. F. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 11273.
(49) Zaykov, A. N.; Ball, Z. T. *Chemical Communications (Cambridge)* **2011**, *47*, 10927.

(50) Harrison, R. S.; Shepherd, N. E.; Hoang, H. N.; Ruiz-Gomez, G.; Hill, T. A.; Driver, R. W.; Desai, V. S.; Young, P. R.; Abbenante, G.; Fairlie, D. P. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 11686.
(51) Harrison, R. S.; Ruiz-Gomez, G.; Hill, T. A.; Chow, S. Y.; Shepherd, N. E.; Lohman, R. J.; Abbenante, G.; Hoang, H. N.; Fairlie, D. P. *J. Med. Chem.* **2010**.
(52) Shepherd, N. E.; Hoang, H. N.; Abbenante, G.; Fairlie, D. P. *J. Am. Chem. Soc.* **2005**, *127*, 2974.
(53) Kritzer, J. A. *Nat. Chem. Biol.* **2010**, *6*, 566.
(54) Kim, Y. W.; Grossmann, T. N.; Verdine, G. L. *Nat. Protoc.* **2011**, *6*, 761.
(55) Walensky, L. D.; Kung, A. L.; Escher, I.; Malia, T. J.; Barbuto, S.; Wright, R. D.; Wagner, G.; Verdine, G. L.; Korsmeyer, S. J. *Science* **2004**, *305*, 1466.
(56) Walensky, L. D.; Pitter, K.; Morash, J.; Oh, K. J.; Barbuto, S.; Fisher, J.; Smith, E.; Verdine, G. L.; Korsmeyer, S. J. *Mol. Cell* **2006**, *24*, 199.
(57) Bhattacharya, S.; Zhang, H.; Debnath, A. K.; Cowburn, D. *J. Biol. Chem.* **2008**, *283*, 16274.
(58) Phillips, C.; Roberts, L. R.; Schade, M.; Bazin, R.; Bent, A.; Davies, N. L.; Moore, R.; Pannifer, A. D.; Pickford, A. R.; Prior, S. H.; Read, C. M.; Scott, A.; Brown, D. G.; Xu, B.; Irving, S. L. *J. Am. Chem. Soc.* **2011**, *133*, 9696.
(59) Stewart, M. L.; Fire, E.; Keating, A. E.; Walensky, L. D. *Nat. Chem. Biol.* **2010**, *6*, 595.
(60) Bernal, F.; Tyler, A. F.; Korsmeyer, S. J.; Walensky, L. D.; Verdine, G. L. *J. Am. Chem. Soc.* **2007**, *129*, 2456.
(61) Baek, S.; Kutchukian, P. S.; Verdine, G. L.; Huber, R.; Holak, T. A.; Lee, K. W.; Popowicz, G. M. *J. Am. Chem. Soc.* **2011**.
(62) Torres, O.; Yuksel, D.; Bernardina, M.; Kumar, K.; Bong, D. *ChemBioChem* **2008**, *9*, 1701.
(63) Madden, M. M.; Muppidi, A.; Li, Z.; Li, X.; Chen, J.; Lin, Q. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 1472.
(64) Patgiri, A.; Jochim, A. L.; Arora, P. S. *Acc. Chem. Res.* **2008**, *41*, 1289.
(65) Liu, J.; Wang, D.; Zheng, Q.; Lu, M.; Arora, P. S. *J. Am. Chem. Soc.* **2008**, *130*, 4334.
(66) Wang, D.; Liao, W.; Arora, P. S. *Angew. Chem. Int. Ed.* **2005**, *44*, 6525.
(67) Henchey, L. K.; Kushal, S.; Dubey, R.; Chapman, R. N.; Olenyuk, B. Z.; Arora, P. S. *J. Am. Chem. Soc.* **2010**, *132*, 941.
(68) Radhakrishnan, I.; Perez-Alvarado, G. C.; Parker, D.; Dyson, H. J.; Montminy, M. R.; Wright, P. E. *Cell* **1997**, *91*, 741.
(69) Rutledge, S. E.; Volkman, H. M.; Schepartz, A. *J. Am. Chem. Soc.* **2003**, *125*, 14336.
(70) Li, C.; Pazgier, M.; Liu, M.; Lu, W. Y.; Lu, W. *Angew. Chem. Int. Ed.* **2009**, *48*, 8712.
(71) Li, C.; Liu, M.; Monbo, J.; Zou, G.; Yuan, W.; Zella, D.; Lu, W. Y.; Lu, W. *J. Am. Chem. Soc.* **2008**, *130*, 13546.
(72) Gemperli, A. C.; Rutledge, S. E.; Maranda, A.; Schepartz, A. *J. Am. Chem. Soc.* **2005**, *127*, 1596.
(73) Chen, Y. C.; Lim, C. *Nucleic Acids Res.* **2008**, *36*, 7078.
(74) Jones, S.; Daley, D. T.; Luscombe, N. M.; Berman, H. M.; Thornton, J. M. *Nucleic Acids Res.* **2001**, *29*, 943.
(75) Ellis, J. J.; Broom, M.; Jones, S. *Proteins* **2007**, *66*, 903.
(76) Bahadur, R. P.; Zacharias, M.; Janin, J. *Nucleic Acids Res.* **2008**, *36*, 2705.
(77) Morozova, N.; Allers, J.; Myers, J.; Shamoo, Y. *Bioinformatics* **2006**, *22*, 2746.
(78) Sonavane, S.; Chakrabarti, P. *Nucleic Acids Res.* **2009**, *37*, 4613.
(79) Biswas, S.; Guharoy, M.; Chakrabarti, P. *Bioinformation* **2008**, *2*, 422.

(80) Mei, H. Y.; Cui, M.; Heldsinger, A.; Lemrow, S. M.; Loo, J. A.; Sannes-Lowery, K. A.; Sharmeen, L.; Czarnik, A. W. *Biochemistry* **1998**, *37*, 14204.
(81) Wang, S.; Huber, P. W.; Cui, M.; Czarnik, A. W.; Mei, H. Y. *Biochemistry* **1998**, *37*, 5549.
(82) Lind, K. E.; Du, Z.; Fujinaga, K.; Peterlin, B. M.; James, T. L. *Chemistry and Biology* **2002**, *9*, 185.
(83) Pang, R.; Zhang, C.; Yuan, D.; Yang, M. *Biorg. Med. Chem.* **2008**, *16*, 8178.
(84) Mayer, M.; James, T. L. *J. Am. Chem. Soc.* **2004**, *126*, 4453.
(85) Murchie, A. I.; Davis, B.; Isel, C.; Afshar, M.; Drysdale, M. J.; Bower, J.; Potter, A. J.; Starkey, I. D.; Swarbrick, T. M.; Mirza, S.; Prescott, C. D.; Vaglio, P.; Aboul-ela, F.; Karn, J. *J. Mol. Biol.* **2004**, *336*, 625.
(86) Du, Z.; Lind, K. E.; James, T. L. *Chem. Biol.* **2002**, *9*, 707.
(87) Faber, C.; Sticht, H.; Schweimer, K.; Rosch, P. *J. Biol. Chem.* **2000**, *275*, 20660.
(88) Sucheck, S. J.; Greenberg, W. A.; Tolbert, T. J.; Wong, C. H. *Angew. Chem. Int. Ed.* **2000**, *39*, 1080.
(89) Sucheck, S. J.; Wong, A. L.; Koeller, K. M.; Boehr, D. D.; Draker, K.-a.; Sears, P.; Wright, G. D.; Wong, C.-H. *J. Am. Chem. Soc.* **2000**, *122*, 5230.
(90) Poehlsgaard, J.; Douthwaite, S. *Nat. Rev. Microbiol.* **2005**, *3*, 870.
(91) Lee, M. M.; Pushechnikov, A.; Disney, M. D. *ACS Chem. Biol.* **2009**, *4*, 345.
(92) Disney, M. D.; Labuda, L. P.; Paul, D. J.; Poplawski, S. G.; Pushechnikov, A.; Tran, T.; Velagapudi, S. P.; Wu, M.; Childs-Disney, J. L. *J. Am. Chem. Soc.* **2008**, *130*, 11185.
(93) Mills, N. L.; Daugherty, M. D.; Frankel, A. D.; Guy, R. K. *J. Am. Chem. Soc.* **2006**, *128*, 3496.
(94) McColl, D. J.; Honchell, C. D.; Frankel, A. D. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 9521.
(95) Friesen, W. J.; Darby, M. K. *Nat. Struct. Biol.* **1998**, *5*, 543.
(96) Mishra, S. H.; Shelley, C. M.; Barrow, D. J., Jr.; Darby, M. K.; Germann, M. W. *Biopolymers* **2006**, *83*, 352.
(97) Litovchick, A.; Rando, R. R. *RNA* **2003**, *9*, 937.
(98) Leeper, T. C.; Athanassiou, Z.; Dias, R. L.; Robinson, J. A.; Varani, G. *Biochemistry* **2005**, *44*, 12362.
(99) Athanassiou, Z.; Patora, K.; Dias, R. L.; Moehle, K.; Robinson, J. A.; Varani, G. *Biochemistry* **2007**, *46*, 741.
(100) Davidson, A.; Patora-Komisarska, K.; Robinson, J. A.; Varani, G. *Nucleic Acids Res.* **2011**, *39*, 248.
(101) Burns, V. A.; Bobay, B. G.; Basso, A.; Cavanagh, J.; Melander, C. *Bioorganic & medicinal chemistry letters* **2008**, *18*, 565.
(102) Moehle, K.; Athanassiou, Z.; Patora, K.; Davidson, A.; Varani, G.; Robinson, J. A. *Angew. Chem. Int. Ed.* **2007**, *46*, 9101.
(103) Frugier, M.; Schimmel, P. *Proceedings of the National Academy of Sciences of the United States of America* **1997**, *94*, 11291.
(104) Krishnamurthy, M.; Gooch, B. D.; Beal, P. A. *Org. Biomol. Chem* **2006**, *4*, 639.
(105) Krishnamurthy, M.; Simon, K.; Orendt, A. M.; Beal, P. A. *Angew. Chem. Int. Ed.* **2007**, *46*, 7044.
(106) Niu, Y.; Jones, A. J.; Wu, H.; Varani, G.; Cai, J. *Org. Biomol. Chem* **2011**, *9*, 6604.
(107) Levy, J. A. *HIV and the pathogenesis of AIDS*; ASM Press: Washington, D.C., 1994.

(108) Gao, F.; Bailes, E.; Robertson, D. L.; Chen, Y.; Rodenburg, C. M.; Michael, S. F.; Cummins, L. B.; Arthur, L. O.; Peeters, M.; Shaw, G. M.; Sharp, P. M.; Hahn, B. H. *Nature* **1999**, *397*, 436.
(109) Wolfe, N. D.; Switzer, W. M.; Carr, J. K.; Bhullar, V. B.; Shanmugam, V.; Tamoufe, U.; Prosser, A. T.; Torimiro, J. N.; Wright, A.; Mpoudi-Ngole, E.; McCutchan, F. E.; Birx, D. L.; Folks, T. M.; Burke, D. S.; Heneine, W. *Lancet* **2004**, *363*, 932.
(110) Worobey, M.; Gemmel, M.; Teuwen, D. E.; Haselkorn, T.; Kunstman, K.; Bunce, M.; Muyembe, J. J.; Kabongo, J. M.; Kalengayi, R. M.; Van Marck, E.; Gilbert, M. T.; Wolinsky, S. M. *Nature* **2008**, *455*, 661.
(111) Barouch, D. H. *Nature* **2008**, *455*, 613.
(112) Joint United Nations Programme on HIV/AIDS. *AIDS scorecards : overview : UNAIDS report on the global AIDS epidemic 2010*; UNAIDS: [Geneva], 2010.
(113) Douek, D. C.; Brenchley, J. M.; Betts, M. R.; Ambrozak, D. R.; Hill, B. J.; Okamoto, Y.; Casazza, J. P.; Kuruppu, J.; Kunstman, K.; Wolinsky, S.; Grossman, Z.; Dybul, M.; Oxenius, A.; Price, D. A.; Connors, M.; Koup, R. A. *Nature* **2002**, *417*, 95.
(114) Cunningham, A. L.; Donaghy, H.; Harman, A. N.; Kim, M.; Turville, S. G. *Curr. Opin. Microbiol.* **2010**, *13*, 524.
(115) Briggs, J. A.; Wilk, T.; Welker, R.; Krausslich, H. G.; Fuller, S. D. *EMBO J.* **2003**, *22*, 1707.
(116) Briggs, J. A.; Grunewald, K.; Glass, B.; Forster, F.; Krausslich, H. G.; Fuller, S. D. *Structure* **2006**, *14*, 15.
(117) Burton, D. R. *Nature* **2006**, *441*, 817.
(118) Watts, J. M.; Dang, K. K.; Gorelick, R. J.; Leonard, C. W.; Bess, J. W., Jr.; Swanstrom, R.; Burch, C. L.; Weeks, K. M. *Nature* **2009**, *460*, 711.
(119) Purcell, D. F.; Martin, M. A. *J. Virol.* **1993**, *67*, 6365.
(120) Felber, B. K.; Hadzopoulou-Cladaras, M.; Cladaras, C.; Copeland, T.; Pavlakis, G. N. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 1495.
(121) Zhou, M.; Deng, L.; Kashanchi, F.; Brady, J. N.; Shatkin, A. J.; Kumar, A. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 12666.
(122) Malim, M. H.; Hauber, J.; Le, S. Y.; Maizel, J. V.; Cullen, B. R. *Nature* **1989**, *338*, 254.
(123) Cullen, B. R. *Nature* **2005**, *433*, 26.
(124) Pallesen, J.; Dong, M.; Besenbacher, F.; Kjems, J. *FEBS J.* **2009**, *276*, 4223.
(125) Daugherty, M. D.; Booth, D. S.; Jayaraman, B.; Cheng, Y.; Frankel, A. D. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 12481.
(126) Malim, M. H.; Hauber, J.; Fenrick, R.; Cullen, B. R. *Nature* **1988**, *335*, 181.
(127) Daugherty, M. D.; Liu, B.; Frankel, A. D. *Nat. Struct. Mol. Biol.* **2010**, *17*, 1337.
(128) Kjems, J.; Frankel, A. D.; Sharp, P. A. *Cell* **1991**, *67*, 169.
(129) Kjems, J.; Calnan, B. J.; Frankel, A. D.; Sharp, P. A. *EMBO J.* **1992**, *11*, 1119.
(130) Tan, R.; Chen, L.; Buettner, J. A.; Hudson, D.; Frankel, A. D. *Cell* **1993**, *73*, 1031.
(131) Tiley, L. S.; Malim, M. H.; Tewary, H. K.; Stockley, P. G.; Cullen, B. R. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 758.
(132) Daelemans, D.; Afonina, E.; Nilsson, J.; Werner, G.; Kjems, J.; De Clercq, E.; Pavlakis, G. N.; Vandamme, A. M. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 14440.
(133) Battiste, J. L.; Mao, H.; Rao, N. S.; Tan, R.; Muhandiram, D. R.; Kay, L. E.; Frankel, A. D.; Williamson, J. R. *Science* **1996**, *273*, 1547.
(134) Ippolito, J. A.; Steitz, T. A. *J. Mol. Biol.* **2000**, *295*, 711.
(135) Gosser, Y.; Hermann, T.; Majumdar, A.; Hu, W.; Frederick, R.; Jiang, F.; Xu, W.; Patel, D. J. *Nat. Struct. Biol.* **2001**, *8*, 146.

(136) Foster, K. W.; Saranak, J.; Patel, N.; Zarilli, G.; Okabe, M.; Kline, T.; Nakanishi, K. *Nature* **1984**, *311*, 756.
(137) Radu, I.; Bamann, C.; Nack, M.; Nagel, G.; Bamberg, E.; Heberle, J. *J. Am. Chem. Soc.* **2009**, *131*, 7313.
(138) Patel, A. B.; Crocker, E.; Eilers, M.; Hirshfeld, A.; Sheves, M.; Smith, S. O. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 10048.
(139) Lamparter, T. *FEBS Lett.* **2004**, *573*, 1.
(140) Nagatani, A. *Curr. Opin. Plant Biol.* **2010**, *13*, 565.
(141) Scheerer, P.; Michael, N.; Park, J. H.; Nagano, S.; Choe, H. W.; Inomata, K.; Borucki, B.; Krauss, N.; Lamparter, T. *Chemphyschem : a European journal of chemical physics and physical chemistry* **2010**, *11*, 1090.
(142) Pfeifer, A.; Mathes, T.; Lu, Y.; Hegemann, P.; Kottke, T. *Biochemistry* **2010**, *49*, 1024.
(143) Harper, S. M.; Neil, L. C.; Gardner, K. H. *Science* **2003**, *301*, 1541.
(144) Noble, A. *Annalen der Chemie und Pharmacie* **1856**, *98*, 253.
(145) Fliegl, H.; Kohn, A.; Hattig, C.; Ahlrichs, R. *J. Am. Chem. Soc.* **2003**, *125*, 9821.
(146) Beharry, A. A.; Woolley, G. A. *Chem. Soc. Rev.* **2011**, *40*, 4422.
(147) Siewertsen, R.; Neumann, H.; Buchheim-Stehn, B.; Herges, R.; Nather, C.; Renth, F.; Temps, F. *J. Am. Chem. Soc.* **2009**, *131*, 15594.
(148) Lee, H. Y.; Diehn, K. K.; Sun, K.; Chen, T.; Raghavan, S. R. *J. Am. Chem. Soc.* **2011**, *133*, 8461.
(149) Kojima, M.; Nakanishi, T.; Hirai, Y.; Yabu, H.; Shimomura, M. *Chemical Communications (Cambridge)* **2010**, *46*, 3970.
(150) Cabrera, I.; Shvartsman, F.; Veinberg, O.; Krongauz, V. *Science* **1984**, *226*, 341.
(151) Koumura, N.; Zijlstra, R. W.; van Delden, R. A.; Harada, N.; Feringa, B. L. *Nature* **1999**, *401*, 152.
(152) Brouwer, A. M.; Frochot, C.; Gatti, F. G.; Leigh, D. A.; Mottier, L.; Paolucci, F.; Roffia, S.; Wurpel, G. W. *Science* **2001**, *291*, 2124.
(153) Ohmuro-Matsuyama, Y.; Tatsu, Y. *Angew. Chem. Int. Ed.* **2008**, *47*, 7527.
(154) Nandy, S. K.; Agnes, R. S.; Lawrence, D. S. *Org. Lett.* **2007**, *9*, 2249.
(155) Goodman, M.; Kossoy, A. *J. Am. Chem. Soc.* **1966**, *88*, 5010.
(156) Goodman, M.; Falxa, M. L. *J. Am. Chem. Soc.* **1967**, *89*, 3863.
(157) Ulysse, L.; Cubillos, J.; Chmielewski, J. *J. Am. Chem. Soc.* **1995**, *117*, 8466.
(158) Dong, S. L.; Loweneck, M.; Schrader, T. E.; Schreier, W. J.; Zinth, W.; Moroder, L.; Renner, C. *Chemistry* **2006**, *12*, 1114.
(159) Hoppmann, C.; Schmieder, P.; Domaing, P.; Vogelreiter, G.; Eichhorst, J.; Wiesner, B.; Morano, I.; Ruck-Braun, K.; Beyermann, M. *Angew Chem Int Ed Engl* **2011**, *50*, 7699.
(160) Kumita, J. R.; Smart, O. S.; Woolley, G. A. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 3803.
(161) Zhang, Z.; Burns, D. C.; Kumita, J. R.; Smart, O. S.; Woolley, G. A. *Bioconj. Chem.* **2003**, *14*, 824.
(162) Kneissl, S.; Loveridge, E. J.; Williams, C.; Crump, M. P.; Allemann, R. K. *ChemBioChem* **2008**, *9*, 3046.
(163) Zhang, F.; Timm, K. A.; Arndt, K. M.; Woolley, G. A. *Angew Chem Int Ed Engl* **2010**, *49*, 3943.
(164) Guerrero, L.; Smart, O. S.; Weston, C. J.; Burns, D. C.; Woolley, G. A.; Allemann, R. K. *Angew. Chem. Int. Ed.* **2005**, *44*, 7778.

(165) Kusebauch, U.; Cadamuro, S. A.; Musiol, H. J.; Moroder, L.; Renner, C. *Chemistry* **2007**, *13*, 2966.
(166) Guerrero, L.; Smart, O. S.; Woolley, G. A.; Allemann, R. K. *J. Am. Chem. Soc.* **2005**, *127*, 15624.
(167) Zhang, F.; Zarrine-Afsar, A.; Al-Abdul-Wahid, M. S.; Prosser, R. S.; Davidson, A. R.; Woolley, G. A. *J. Am. Chem. Soc.* **2009**, *131*, 2283.
(168) Schierling, B.; Noel, A. J.; Wende, W.; Hien le, T.; Volkov, E.; Kubareva, E.; Oretskaya, T.; Kokkinidis, M.; Rompp, A.; Spengler, B.; Pingoud, A. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 1361.
(169) Volgraf, M.; Gorostiza, P.; Numano, R.; Kramer, R. H.; Isacoff, E. Y.; Trauner, D. *Nat. Chem. Biol.* **2006**, *2*, 47.
(170) Fortin, D. L.; Banghart, M. R.; Dunn, T. W.; Borges, K.; Wagenaar, D. A.; Gaudry, Q.; Karakossian, M. H.; Otis, T. S.; Kristan, W. B.; Trauner, D.; Kramer, R. H. *Nat. Methods* **2008**, *5*, 331.
(171) Fortin, D. L.; Dunn, T. W.; Fedorchak, A.; Allen, D.; Montpetit, R.; Banghart, M. R.; Trauner, D.; Adelman, J. P.; Kramer, R. H. *J. Neurophysiol.* **2011**, *106*, 488.
(172) Pozhidaeva, N.; Cormier, M. E.; Chaudhari, A.; Woolley, G. A. *Bioconj. Chem.* **2004**, *15*, 1297.
(173) Beharry, A. A.; Sadovski, O.; Woolley, G. A. *Org. Biomol. Chem* **2008**, *6*, 4323.
(174) Chi, L.; Sadovski, O.; Woolley, G. A. *Bioconj. Chem.* **2006**, *17*, 670.
(175) Fujimoto, K.; Amano, M.; Horibe, Y.; Inouye, M. *Org. Lett.* **2006**, *8*, 285.
(176) Kocer, A.; Walko, M.; Meijberg, W.; Feringa, B. L. *Science* **2005**, *309*, 755.
(177) Lougheed, T.; Borisenko, V.; Hennig, T.; Ruck-Braun, K.; Woolley, G. A. *Org. Biomol. Chem* **2004**, *2*, 2798.
(178) Huang, Y.; Cong, Z.; Yang, L.; Dong, S. *J. Pept. Sci.* **2008**, *14*, 1062.
(179) Huang, Y.; Jahreis, G.; Lucke, C.; Wildemann, D.; Fischer, G. *J. Am. Chem. Soc.* **2010**, *132*, 7578.
(180) Adamantidis, A. R.; Zhang, F.; Aravanis, A. M.; Deisseroth, K.; de Lecea, L. *Nature* **2007**, *450*, 420.
(181) Gradinaru, V.; Mogri, M.; Thompson, K. R.; Henderson, J. M.; Deisseroth, K. *Science* **2009**, *324*, 354.
(182) Huber, D.; Petreanu, L.; Ghitani, N.; Ranade, S.; Hromadka, T.; Mainen, Z.; Svoboda, K. *Nature* **2008**, *451*, 61.
(183) Airan, R. D.; Thompson, K. R.; Fenno, L. E.; Bernstein, H.; Deisseroth, K. *Nature* **2009**, *458*, 1025.
(184) Levskaya, A.; Weiner, O. D.; Lim, W. A.; Voigt, C. A. *Nature* **2009**, *461*, 997.
(185) Candi, E.; Schmidt, R.; Melino, G. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 328.
(186) Buttner, S.; Ruli, D.; Vogtle, F. N.; Galluzzi, L.; Moitzi, B.; Eisenberg, T.; Kepp, O.; Habernig, L.; Carmona-Gutierrez, D.; Rockenfeller, P.; Laun, P.; Breitenbach, M.; Khoury, C.; Frohlich, K. U.; Rechberger, G.; Meisinger, C.; Kroemer, G.; Madeo, F. *EMBO J.* **2011**, *30*, 2779.
(187) Portt, L.; Norman, G.; Clapp, C.; Greenwood, M.; Greenwood, M. T. *Biochim. Biophys. Acta* **2011**, *1813*, 238.
(188) Fulda, S.; Gorman, A. M.; Hori, O.; Samali, A. *Int. J. Cell. Biol.* **2010**, *2010*, 214074.
(189) Kroemer, G.; Galluzzi, L.; Vandenabeele, P.; Abrams, J.; Alnemri, E. S.; Baehrecke, E. H.; Blagosklonny, M. V.; El-Deiry, W. S.; Golstein, P.; Green, D. R.; Hengartner, M.; Knight, R. A.; Kumar, S.; Lipton, S. A.; Malorni, W.; Nunez, G.; Peter, M. E.; Tschopp, J.; Yuan, J.; Piacentini, M.; Zhivotovsky, B.; Melino, G. *Cell Death Differ.* **2009**, *16*, 3.

(190) Lockshin, R. A.; Williams, C. M. *J. Insect Physiol.* **1964**, *10*, 643.
(191) Kerr, J. F.; Wyllie, A. H.; Currie, A. R. *Br. J. Cancer* **1972**, *26*, 239.
(192) Ellis, H. M.; Horvitz, H. R. *Cell* **1986**, *44*, 817.
(193) Vandenabeele, P.; Galluzzi, L.; Vanden Berghe, T.; Kroemer, G. *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 700.
(194) Elliott, M. R.; Chekeni, F. B.; Trampont, P. C.; Lazarowski, E. R.; Kadl, A.; Walk, S. F.; Park, D.; Woodson, R. I.; Ostankovich, M.; Sharma, P.; Lysiak, J. J.; Harden, T. K.; Leitinger, N.; Ravichandran, K. S. *Nature* **2009**, *461*, 282.
(195) Coleman, M. L.; Sahai, E. A.; Yeo, M.; Bosch, M.; Dewar, A.; Olson, M. F. *Nat. Cell Biol.* **2001**, *3*, 339.
(196) Hedrick, S. M.; Ch'en, I. L.; Alves, B. N. *Immunol. Rev.* **2010**, *236*, 41.
(197) Clancy, L.; Mruk, K.; Archer, K.; Woelfel, M.; Mongkolsapaya, J.; Screaton, G.; Lenardo, M. J.; Chan, F. K. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 18099.
(198) Papoff, G.; Hausler, P.; Eramo, A.; Pagano, M. G.; Di Leve, G.; Signore, A.; Ruberti, G. *J. Biol. Chem.* **1999**, *274*, 38241.
(199) Chan, F. K.; Chun, H. J.; Zheng, L.; Siegel, R. M.; Bui, K. L.; Lenardo, M. J. *Science* **2000**, *288*, 2351.
(200) Chan, F. K. *Cytokine* **2007**, *37*, 101.
(201) Kischkel, F. C.; Lawrence, D. A.; Chuntharapai, A.; Schow, P.; Kim, K. J.; Ashkenazi, A. *Immunity* **2000**, *12*, 611.
(202) Wilson, N. S.; Dixit, V.; Ashkenazi, A. *Nat. Immunol.* **2009**, *10*, 348.
(203) Jin, T. G.; Kurakin, A.; Benhaga, N.; Abe, K.; Mohseni, M.; Sandra, F.; Song, K.; Kay, B. K.; Khosravi-Far, R. *J. Biol. Chem.* **2004**, *279*, 55594.
(204) Kataoka, T.; Tschopp, J. *Mol. Cell. Biol.* **2004**, *24*, 2627.
(205) Ozoren, N.; El-Deiry, W. S. *Neoplasia* **2002**, *4*, 551.
(206) Gonzalvez, F.; Ashkenazi, A. *Oncogene* **2010**, *29*, 4752.
(207) Rodriguez, D.; Rojas-Rivera, D.; Hetz, C. *Biochim. Biophys. Acta* **2011**, *1813*, 564.
(208) Shamas-Din, A.; Brahmbhatt, H.; Leber, B.; Andrews, D. W. *Biochim. Biophys. Acta* **2011**, *1813*, 508.
(209) Diemand, A. V.; Lupas, A. N. *J. Struct. Biol* **2006**, *156*, 230.
(210) Alnemri, E. S.; Livingston, D. J.; Nicholson, D. W.; Salvesen, G.; Thornberry, N. A.; Wong, W. W.; Yuan, J. *Cell* **1996**, *87*, 171.
(211) Wang, Y.; Gu, X. *Genetics* **2001**, *158*, 1311.
(212) Lamkanfi, M.; Festjens, N.; Declercq, W.; Vanden Berghe, T.; Vandenabeele, P. *Cell Death Differ.* **2007**, *14*, 44.
(213) Denecker, G.; Hoste, E.; Gilbert, B.; Hochepied, T.; Ovaere, P.; Lippens, S.; Van den Broecke, C.; Van Damme, P.; D'Herde, K.; Hachem, J. P.; Borgonie, G.; Presland, R. B.; Schoonjans, L.; Libert, C.; Vandekerckhove, J.; Gevaert, K.; Vandenabeele, P.; Declercq, W. *Nat. Cell Biol.* **2007**, *9*, 666.
(214) Pop, C.; Salvesen, G. S. *J. Biol. Chem.* **2009**, *284*, 21777.
(215) Mahrus, S.; Trinidad, J. C.; Barkan, D. T.; Sali, A.; Burlingame, A. L.; Wells, J. A. *Cell* **2008**, *134*, 866.
(216) Wyllie, A. H. *Mol. Neurobiol.* **2010**, *42*, 4.
(217) He, B.; Lu, N.; Zhou, Z. *Curr. Opin. Cell Biol.* **2009**, *21*, 900.
(218) Chipuk, J. E.; Moldoveanu, T.; Llambi, F.; Parsons, M. J.; Green, D. R. *Mol. Cell* **2010**, *37*, 299.
(219) Tsujimoto, Y.; Cossman, J.; Jaffe, E.; Croce, C. M. *Science* **1985**, *228*, 1440.
(220) Chinnadurai, G.; Vijayalingam, S.; Rashmi, R. *Oncogene* **2008**, *27 Suppl 1*, S20.

(221) Kvansakul, M.; Yang, H.; Fairlie, W. D.; Czabotar, P. E.; Fischer, S. F.; Perugini, M. A.; Huang, D. C.; Colman, P. M. *Cell Death Differ.* **2008**, *15*, 1564.
(222) Terrones, O.; Antonsson, B.; Yamaguchi, H.; Wang, H. G.; Liu, J.; Lee, R. M.; Herrmann, A.; Basanez, G. *J. Biol. Chem.* **2004**, *279*, 30081.
(223) Qian, S.; Wang, W.; Yang, L.; Huang, H. W. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 17379.
(224) Schafer, B.; Quispe, J.; Choudhary, V.; Chipuk, J. E.; Ajero, T. G.; Du, H.; Schneiter, R.; Kuwana, T. *Mol. Biol. Cell* **2009**, *20*, 2276.
(225) Bleicken, S.; Classen, M.; Padmavathi, P. V.; Ishikawa, T.; Zeth, K.; Steinhoff, H. J.; Bordignon, E. *J. Mol. Biol.* **2010**, *285*, 6636.
(226) Hsu, S. Y.; Kaipia, A.; McGee, E.; Lomeli, M.; Hsueh, A. J. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 12401.
(227) Rodriguez, J. M.; Glozak, M. A.; Ma, Y.; Cress, W. D. *J. Biol. Chem.* **2006**, *281*, 22729.
(228) Ray, J. E.; Garcia, J.; Jurisicova, A.; Caniggia, I. *Cell Death Differ.* **2010**, *17*, 846.
(229) Zeilstra, J.; Joosten, S. P.; Wensveen, F. M.; Dessing, M. C.; Schutze, D. M.; Eldering, E.; Spaargaren, M.; Pals, S. T. *Biochem. Biophys. Res. Commun.* **2011**, *406*, 1.
(230) Danial, N. N. *Oncogene* **2008**, *27 Suppl 1*, S53.
(231) Yang, J.; Li, J. H.; Wang, J.; Zhang, C. Y. *J. Theor. Biol.* **2010**, *266*, 231.
(232) Polzien, L.; Baljuls, A.; Rennefahrt, U. E.; Fischer, A.; Schmitz, W.; Zahedi, R. P.; Sickmann, A.; Metz, R.; Albert, S.; Benz, R.; Hekman, M.; Rapp, U. R. *J. Biol. Chem.* **2009**, *284*, 28004.
(233) Hausmann, M.; Leucht, K.; Ploner, C.; Kiessling, S.; Villunger, A.; Becker, H.; Hofmann, C.; Falk, W.; Krebs, M.; Kellermeier, S.; Fried, M.; Scholmerich, J.; Obermeier, F.; Rogler, G. *J. Biol. Chem.* **2011**, *286*, 26533.
(234) Gillings, A. S.; Balmanno, K.; Wiggins, C. M.; Johnson, M.; Cook, S. J. *FEBS J.* **2009**, *276*, 6050.
(235) Yu, J.; Zhang, L. *Oncogene* **2008**, *27 Suppl 1*, S71.
(236) Ploner, C.; Kofler, R.; Villunger, A. *Oncogene* **2008**, *27 Suppl 1*, S84.
(237) Billen, L. P.; Shamas-Din, A.; Andrews, D. W. *Oncogene* **2008**, *27 Suppl 1*, S93.
(238) Westphal, D.; Dewson, G.; Czabotar, P. E.; Kluck, R. M. *Biochim. Biophys. Acta* **2011**, *1813*, 521.
(239) Rolland, S. G.; Conradt, B. *Curr. Opin. Cell Biol.* **2010**, *22*, 852.
(240) Chou, J. J.; Li, H.; Salvesen, G. S.; Yuan, J.; Wagner, G. *Cell* **1999**, *96*, 615.
(241) Hinds, M. G.; Smits, C.; Fredericks-Short, R.; Risk, J. M.; Bailey, M.; Huang, D. C.; Day, C. L. *Cell Death Differ.* **2007**, *14*, 128.
(242) Wells, M.; Tidow, H.; Rutherford, T. J.; Markwick, P.; Jensen, M. R.; Mylonas, E.; Svergun, D. I.; Blackledge, M.; Fersht, A. R. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 5762.
(243) Lama, D.; Sankararamakrishnan, R. *Biochemistry* **2010**, *49*, 2574.
(244) Hinds, M. G.; Lackmann, M.; Skea, G. L.; Harrison, P. J.; Huang, D. C.; Day, C. L. *EMBO J.* **2003**, *22*, 1497.
(245) Denisov, A. Y.; Madiraju, M. S.; Chen, G.; Khadir, A.; Beauparlant, P.; Attardo, G.; Shore, G. C.; Gehring, K. *J. Biol. Chem.* **2003**, *278*, 21124.
(246) Suzuki, M.; Youle, R. J.; Tjandra, N. *Cell* **2000**, *103*, 645.
(247) Sattler, M.; Liang, H.; Nettesheim, D.; Meadows, R. P.; Harlan, J. E.; Eberstadt, M.; Yoon, H. S.; Shuker, S. B.; Chang, B. S.; Minn, A. J.; Thompson, C. B.; Fesik, S. W. *Science* **1997**, *275*, 983.

(248) Muchmore, S. W.; Sattler, M.; Liang, H.; Meadows, R. P.; Harlan, J. E.; Yoon, H. S.; Nettesheim, D.; Chang, B. S.; Thompson, C. B.; Wong, S. L.; Ng, S. L.; Fesik, S. W. *Nature* **1996**, *381*, 335.
(249) Petros, A. M.; Nettesheim, D. G.; Wang, Y.; Olejniczak, E. T.; Meadows, R. P.; Mack, J.; Swift, K.; Matayoshi, E. D.; Zhang, H.; Thompson, C. B.; Fesik, S. W. *Protein Sci.* **2000**, *9*, 2528.
(250) Petros, A. M.; Medek, A.; Nettesheim, D. G.; Kim, D. H.; Yoon, H. S.; Swift, K.; Matayoshi, E. D.; Oltersdorf, T.; Fesik, S. W. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 3012.
(251) Day, C. L.; Chen, L.; Richardson, S. J.; Harrison, P. J.; Huang, D. C.; Hinds, M. G. *J. Biol. Chem.* **2005**, *280*, 4738.
(252) Ku, B.; Liang, C.; Jung, J. U.; Oh, B. H. *Cell Res.* **2011**, *21*, 627.
(253) Deng, X.; Gao, F.; Flagg, T.; Anderson, J.; May, W. S. *Mol. Cell. Biol.* **2006**, *26*, 4421.
(254) Lee, E. F.; Smith, B. J.; Horne, W. S.; Mayer, K. N.; Evangelista, M.; Colman, P. M.; Gellman, S. H.; Fairlie, W. D. *ChemBioChem* **2011**, *12*, 2025.
(255) Oberstein, A.; Jeffrey, P. D.; Shi, Y. *J. Biol. Chem.* **2007**, *282*, 13123.
(256) O'Neill, J. W.; Manion, M. K.; Maguire, B.; Hockenbery, D. M. *J. Mol. Biol.* **2006**, *356*, 367.
(257) Lee, E. F.; Dewson, G.; Smith, B. J.; Evangelista, M.; Pettikiriarachchi, A.; Dogovski, C.; Perugini, M. A.; Colman, P. M.; Fairlie, W. D. *Structure* **2011**, *19*, 1467.
(258) Petros, A. M.; Olejniczak, E. T.; Fesik, S. W. *Biochim. Biophys. Acta* **2004**, *1644*, 83.
(259) Kuliopulos, A.; Walsh, C. T. *J. Am. Chem. Soc.* **1994**, *116*, 4599.
(260) Maloy, S.; Stewart, V.; Taylor, R. *Genetic Analysis of Pathogenic Bacteria*; Cold Spring Harbor Laboratory Press: NY, 1996.
(261) Fricke, T.; Mart, R. J.; L., W. C.; M., W.; J., E. R.; T., J. A.; Allemann, R. K. *Bioconj. Chem.* **2011**, *22*, 1763.
(262) Mart, R. J.; Wysoczanski, P.; Kneissl, S.; Ricci, A.; Brancale, A.; Allemann, R. K. *ChemBioChem* **2012**, *13*, 515.
(263) Kurita, N.; Ikegami, T.; Ishikawa, Y. *Chem. Phys. Lett.* **2002**, *360*, 349.
(264) Dupradeau, F. Y.; Pigache, A.; Zaffran, T.; Savineau, C.; Lelong, R.; Grivel, N.; Lelong, D.; Rosanski, W.; Cieplak, P. *PCCP* **2010**, *12*, 7821.
(265) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. *J. Mol. Graphics Model.* **2006**, *25*, 247.
(266) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157.
(267) Traetteberg, M.; Hilmo, I.; Hagen, K. *J. Mol. Struct.* **1977**, *39*, 231.
(268) Tsuji, T.; Takashima, H.; Takeuchi, H.; Egawa, T.; Konaka, S. *J. Phys. Chem. A* **2001**, *105*, 9347.
(269) Fricke, T., Cardiff University, 2012.
(270) Samuel, D.; K., K. S.; Srimathi, T.; Hsieh, H. C.; Yu, C. *J. Biol. Chem.* **2000**, *275*, 34968.
(271) Sherman, F.; Stewart, J. W.; Tsunasawa, S. *Bioessays* **1985**, *3*, 27.
(272) Gupta, N.; Tanner, S.; Jaitly, N.; Adkins, J. N.; Lipton, M.; Edwards, R.; Romine, M.; Osterman, A.; Bafna, V.; Smith, R. D.; Pevzner, P. A. *Genome Res.* **2007**, *17*, 1362.
(273) Linge, J. P.; S.I., O. D.; Nilges, M. *Methods Enzymol.* **2001**, *339*, 71.
(274) Nilges, M.; Bernard, A.; Bardiaux, B.; T., M.; Habeck, M.; Rieping, W. *Structure* **2008**, *16*, 1305.

(275) Brunger, A. T.; Adams, P. D.; Clore, G. M.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J.-S.; Kuszewski, J.; Nilges, N.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **1998**, *54*, 905.
(276) Linge, J. P.; Williams, M.; Spronk, C.; Bonvin, A.; Nilges, M. *Proteins* **2003**, *50*, 496.
(277) Linge, J. P.; Habeck, M.; Rieping, W.; Nilges, M. *J. Magn. Reson.* **2004**, *167*, 334.
(278) Manion, M. K.; O'Neill, J. W.; Giedt, C. D.; Kim, K. M.; Zhang, K. Y.; Hockenbery, D. M. *J. Biol. Chem.* **2004**, *279*, 2159.
(279) Fogh, R. H.; Boucher, W.; Vranken, W. F.; Pajon, A.; Stevens, T. J.; Bhat, T. N.; Westbrook, J.; Ionides, J. M. C.; Laue, E. D. *Bioinformatics* **2004**, *21*, 1678.
(280) Peterson, R. D.; Theimer, C. A.; Wu, H.; Feigon, J. *J. Biomol. NMR* **2004**, *28*, 59.
(281) Fossi, M.; Oschkinat, H.; Nilges, M.; Ball, L. J. *J. Magn. Reson.* **2005**, *175*, 92.
(282) Schuttelkopf A.W.; van Aalten, D. M. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2004**, *60*, 1355.
(283) Kumar, S.; Nussinov, R. *ChemBioChem* **2002**, *3*, 604.
(284) Yang, C. Y.; Wang, S. *ACS Med. Chem. Lett.* **2011**, *2*, 280.
(285) Xu, H.; Ye, H.; Osman, N. E.; Sadler, K.; Won, E. Y.; Chi, S. W.; Yoon, H. S. *Biochemistry* **2009**, *48*, 12159.
(286) Bharatham, N.; Chi, S. W.; Yoon, H. S. *PloS one* **2011**, *6*, 1.
(287) Chin, J. W.; Schepartz, A. *Angew. Chem. Int. Ed.* **2001**, *40*, 3806.
(288) Laemmli, U. K. *Nature* **1970**, *227*, 680.
(289) Bertani, G. *J. Bacteriol.* **1951**, *62*, 293.
(290) Cheung, M. S.; Maguire, M. L.; Stevens, T. J.; Broadhurst, R. W. *J. Magn. Reson. Imaging* **2010**, *202*, 223.
(291) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277.
(292) Burns, D. C.; Zhang, F.; Woolley, G. A. *Nat. Protoc.* **2007**, *2*, 251.
(293) Brunger, A. T. *Nat. Protoc.* **2007**, *2*, 2728.
(294) Laskowski, R. A.; Macarthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Crystallogr.* **1993**, *26*, 283.
(295) Hooft, R. W.; Vriend, G.; Sander, C.; Abola, E. E. *Nature* **1996**, *381*, 272.

5 APPENDIX

Number	Shift (ppm)	SD (ppm)	Assign. Name	Residue	Isotope
1	4.34	0.008	Ha	1Met	^{1}H
2	21.74	0.000	Ce	1Met	^{13}C
3	2.26	0.008	Hga	1Met	^{1}H
4	2.31	0.007	Hgb	1Met	^{1}H
5	1.91	0.002	Hb3	1Met	^{1}H
6	1.91	0.002	Hb2	1Met	^{1}H
7	36.36	0.008	Cg	1Met	^{13}C
8	30.58	0.000	Cb	1Met	^{13}C
9	56.61	0.034	Ca	1Met	^{13}C
10	175.76	0.000	C	1Met	^{13}C
11	1.19	0.006	He*	1Met	^{1}H
12	122.26	0.051	N	2Ser	^{15}N
13	7.95	0.003	H	2Ser	^{1}H
14	59.96	0.042	Ca	2Ser	^{13}C
15	64.96	0.084	Cb	2Ser	^{13}C
16	3.81	0.024	Hb2	2Ser	^{1}H
17	4.27	0.002	Ha	2Ser	^{1}H
18	3.83	0.027	Hb3	2Ser	^{1}H
19	7.67	0.003	He2b	3Gln	^{1}H
20	112.34	0.263	Ne2	3Gln	^{15}N
21	2.11	0.015	Hba	3Gln	^{1}H
22	2.31	0.002	Hg2	3Gln	^{1}H
23	2.11	0.015	Hbb	3Gln	^{1}H
24	2.31	0.002	Hg3	3Gln	^{1}H
25	33.67	0.050	Cg	3Gln	^{13}C
26	56.62	0.099	Ca	3Gln	^{13}C
27	29.22	0.023	Cb	3Gln	^{13}C
28	4.39	0.000	Ha	3Gln	^{1}H
29	179.08	0.023	C	3Gln	^{13}C
30	6.87	0.003	He2a	3Gln	^{1}H
31	179.57	0.022	Cd	3Gln	^{13}C
32	8.65	0.007	H	4Ser	^{1}H
33	117.10	0.086	N	4Ser	^{15}N
34	175.93	0.034	C	4Ser	^{13}C
35	58.99	0.039	Ca	4Ser	^{13}C
36	63.49	0.037	Cb	4Ser	^{13}C
37	4.47	0.015	Ha	4Ser	^{1}H
38	3.91	0.004	Hba	4Ser	^{1}H
39	4.06	0.008	Hbb	4Ser	^{1}H
40	8.60	0.010	H	5Asn	^{1}H
41	122.48	0.070	N	5Asn	^{15}N
42	8.02	0.005	Hd2b	5Asn	^{1}H

43	112.37	0.143	Nd2	5Asn	^{15}N
44	56.15	0.072	Ca	5Asn	^{13}C
45	38.36	0.055	Cb	5Asn	^{13}C
46	177.38	0.007	C	5Asn	^{13}C
47	4.74	0.012	Ha	5Asn	^{1}H
48	3.09	0.008	Hbb	5Asn	^{1}H
49	2.75	0.009	Hba	5Asn	^{1}H
50	7.50	0.006	Hd2a	5Asn	^{1}H
51	175.02	0.018	Cg	5Asn	^{13}C
52	120.02	0.074	N	6Arg	^{15}N
53	8.25	0.004	H	6Arg	^{1}H
54	59.86	0.088	Ca	6Arg	^{13}C
55	29.76	0.122	Cb	6Arg	^{13}C
56	177.33	0.023	C	6Arg	^{13}C
57	3.93	0.010	Ha	6Arg	^{1}H
58	3.19	0.011	Hd3	6Arg	^{1}H
59	3.19	0.011	Hd2	6Arg	^{1}H
60	43.26	0.023	Cd	6Arg	^{13}C
61	27.41	0.045	Cg	6Arg	^{13}C
62	1.64	0.003	Hgb	6Arg	^{1}H
63	1.64	0.004	Hga	6Arg	^{1}H
64	1.93	0.009	Hb3	6Arg	^{1}H
65	1.93	0.009	Hb2	6Arg	^{1}H
66	7.95	0.004	H	7Glu	^{1}H
67	117.80	0.072	N	7Glu	^{15}N
68	59.38	0.055	Ca	7Glu	^{13}C
69	179.39	0.028	C	7Glu	^{13}C
70	36.88	0.115	Cg	7Glu	^{13}C
71	29.62	0.076	Cb	7Glu	^{13}C
72	3.91	0.005	Ha	7Glu	^{1}H
73	2.10	0.010	Hbb	7Glu	^{1}H
74	2.00	0.008	Hba	7Glu	^{1}H
75	2.37	0.007	Hgb	7Glu	^{1}H
76	2.22	0.008	Hga	7Glu	^{1}H
77	119.79	0.057	N	8Leu	^{15}N
78	7.73	0.007	H	8Leu	^{1}H
79	58.12	0.056	Ca	8Leu	^{13}C
80	41.93	0.043	Cb	8Leu	^{13}C
81	178.37	0.018	C	8Leu	^{13}C
82	1.47	0.009	Hg	8Leu	^{1}H
83	23.52	0.112	Cdb	8Leu	^{13}C
84	26.24	0.072	Cg	8Leu	^{13}C
85	0.12	0.004	Hdb*	8Leu	^{1}H
86	24.29	0.106	Cda	8Leu	^{13}C
87	-0.02	0.005	Hda*	8Leu	^{1}H

88	1.79	0.012	Hbb	8Leu	^{1}H
89	1.29	0.011	Hba	8Leu	^{1}H
90	3.87	0.013	Ha	8Leu	^{1}H
91	7.76	0.006	H	9Val	^{1}H
92	117.82	0.040	N	9Val	^{15}N
93	67.77	0.099	Ca	9Val	^{13}C
94	32.43	0.041	Cb	9Val	^{13}C
95	177.08	0.046	C	9Val	^{13}C
96	3.68	0.009	Ha	9Val	^{1}H
97	1.22	0.015	Hgb*	9Val	^{1}H
98	1.07	0.008	Hga*	9Val	^{1}H
99	2.42	0.006	Hb	9Val	^{1}H
100	22.94	0.113	Cgb	9Val	^{13}C
101	24.14	0.119	Cga	9Val	^{13}C
102	31.40	0.068	Cb	10Val	^{13}C
103	8.33	0.008	H	10Val	^{1}H
104	117.32	0.071	N	10Val	^{15}N
105	67.03	0.073	Ca	10Val	^{13}C
106	178.59	0.042	C	10Val	^{13}C
107	2.93	0.011	Ha	10Val	^{1}H
108	0.86	0.010	Hgb*	10Val	^{1}H
109	0.13	0.006	Hga*	10Val	^{1}H
110	22.82	0.128	Cgb	10Val	^{13}C
111	20.18	0.049	Cga	10Val	^{13}C
112	1.76	0.005	Hb	10Val	^{1}H
113	43.35	0.093	Cb	11Asp	^{13}C
114	8.07	0.005	H	11Asp	^{1}H
115	120.86	0.096	N	11Asp	^{15}N
116	58.95	0.026	Ca	11Asp	^{13}C
117	177.50	0.012	C	11Asp	^{13}C
118	2.68	0.046	Hbb	11Asp	^{1}H
119	2.56	0.011	Hba	11Asp	^{1}H
120	4.08	0.009	Ha	11Asp	^{1}H
121	63.22	0.049	Ca	12Phe	^{13}C
122	120.29	0.054	N	12Phe	^{15}N
123	40.13	0.064	Cb	12Phe	^{13}C
124	8.33	0.005	H	12Phe	^{1}H
125	178.08	0.023	C	12Phe	^{13}C
126	3.13	0.009	Hba	12Phe	^{1}H
127	3.50	0.014	Hbb	12Phe	^{1}H
128	3.95	0.013	Ha	12Phe	^{1}H
129	7.09	0.014	Hd*	12Phe	^{1}H
130	131.49	0.078	Cd*	12Phe	^{13}C
131	131.37	0.083	Ce*	12Phe	^{13}C
132	7.47	0.005	He*	12Phe	^{1}H

133	116.84	0.064	N	13Leu	^{15}N
134	8.91	0.004	H	13Leu	^{1}H
135	58.29	0.059	Ca	13Leu	^{13}C
136	41.74	0.070	Cb	13Leu	^{13}C
137	179.45	0.018	C	13Leu	^{13}C
138	3.99	0.008	Ha	13Leu	^{1}H
139	0.95	0.011	Hd1*	13Leu	^{1}H
140	25.77	0.163	Cdb	13Leu	^{13}C
141	26.66	0.142	Cg	13Leu	^{13}C
142	0.95	0.010	Hd2*	13Leu	^{1}H
143	2.27	0.009	Hg	13Leu	^{1}H
144	1.82	0.010	Hbb	13Leu	^{1}H
145	21.85	0.049	Cda	13Leu	^{13}C
146	1.51	0.009	Hba	13Leu	^{1}H
147	8.50	0.005	H	14Ser	^{1}H
148	114.31	0.098	N	14Ser	^{15}N
149	62.59	0.000	Cb	14Ser	^{13}C
150	177.28	0.021	C	14Ser	^{13}C
151	62.64	0.065	Ca	14Ser	^{13}C
152	4.26	0.014	Ha	14Ser	^{1}H
153	3.96	0.011	Hbb	14Ser	^{1}H
154	3.83	0.003	Hba	14Ser	^{1}H
155	124.99	0.063	N	15Tyr	^{15}N
156	8.25	0.007	H	15Tyr	^{1}H
157	60.31	0.087	Ca	15Tyr	^{13}C
158	37.99	0.156	Cb	15Tyr	^{13}C
159	178.25	0.003	C	15Tyr	^{13}C
160	3.04	0.005	Hba	15Tyr	^{1}H
161	4.39	0.013	Ha	15Tyr	^{1}H
162	6.77	0.013	Hd*	15Tyr	^{1}H
163	132.18	0.073	Cd*	15Tyr	^{13}C
164	6.89	0.002	He*	15Tyr	^{1}H
165	117.41	0.000	Ce*	15Tyr	^{13}C
166	120.26	0.065	N	16Lys	^{15}N
167	8.33	0.005	H	16Lys	^{1}H
168	180.27	0.035	C	16Lys	^{13}C
169	57.06	0.042	Ca	16Lys	^{13}C
170	29.85	0.051	Cb	16Lys	^{13}C
171	3.83	0.017	Ha	16Lys	^{1}H
172	41.18	0.081	Cb	17Leu	^{13}C
173	27.93	0.158	Cg	17Leu	^{13}C
174	26.40	0.110	Cdb	17Leu	^{13}C
175	23.99	0.056	Cda	17Leu	^{13}C
176	2.14	0.014	Hbb	17Leu	^{1}H
177	4.08	0.009	Ha	17Leu	^{1}H

178	1.52	0.012	Hba	17Leu	^{1}H
179	0.93	0.010	Hdb*	17Leu	^{1}H
180	0.76	0.007	Hda*	17Leu	^{1}H
181	1.85	0.010	Hg	17Leu	^{1}H
182	118.56	0.046	N	17Leu	^{15}N
183	8.73	0.003	H	17Leu	^{1}H
184	58.69	0.101	Ca	17Leu	^{13}C
185	179.92	0.010	C	17Leu	^{13}C
186	8.39	0.019	H	18Ser	^{1}H
187	117.86	0.049	N	18Ser	^{15}N
188	61.78	0.089	Ca	18Ser	^{13}C
189	62.61	0.138	Cb	18Ser	^{13}C
190	178.61	0.002	C	18Ser	^{13}C
191	4.70	0.007	Ha	18Ser	^{1}H
192	4.21	0.012	Hbb	18Ser	^{1}H
193	4.10	0.005	Hba	18Ser	^{1}H
194	122.90	0.060	N	19Gln	^{15}N
195	7.69	0.005	H	19Gln	^{1}H
196	7.52	0.009	He2b	19Gln	^{1}H
197	6.85	0.008	He2a	19Gln	^{1}H
198	114.61	0.250	Ne2	19Gln	^{15}N
199	58.54	0.154	Ca	19Gln	^{13}C
200	28.80	0.195	Cb	19Gln	^{13}C
201	4.02	0.005	Ha	19Gln	^{1}H
202	34.46	0.106	Cg	19Gln	^{13}C
203	177.24	0.017	C	19Gln	^{13}C
204	2.11	0.010	Hgb	19Gln	^{1}H
205	2.00	0.027	Hga	19Gln	^{1}H
206	179.72	0.007	Cd	19Gln	^{13}C
207	42.27	0.136	Ce	20Lys	^{13}C
208	116.43	0.049	N	20Lys	^{15}N
209	7.27	0.003	H	20Lys	^{1}H
210	54.03	0.168	Ca	20Lys	^{13}C
211	32.66	0.150	Cb	20Lys	^{13}C
212	4.43	0.008	Ha	20Lys	^{1}H
213	1.68	0.009	Hdb	20Lys	^{1}H
214	3.03	0.015	He2	20Lys	^{1}H
215	2.12	0.008	Hbb	20Lys	^{1}H
216	28.27	0.070	Cd	20Lys	^{13}C
217	1.68	0.009	Hda	20Lys	^{1}H
218	3.03	0.015	He3	20Lys	^{1}H
219	25.11	0.071	Cg	20Lys	^{13}C
220	1.04	0.028	Hga	20Lys	^{1}H
221	1.78	0.025	Hba	20Lys	^{1}H
222	1.57	0.010	Hgb	20Lys	^{1}H

223	175.58	0.024	C	20Lys	^{13}C
224	7.62	0.005	H	21Gly	^{1}H
225	105.59	0.078	N	21Gly	^{15}N
226	45.32	0.077	Ca	21Gly	^{13}C
227	3.68	0.007	Haa	21Gly	^{1}H
228	4.00	0.006	Hab	21Gly	^{1}H
229	174.01	0.000	C	21Gly	^{13}C
230	2.79	0.008	Hba	22Tyr	^{1}H
231	5.05	0.006	Ha	22Tyr	^{1}H
232	3.22	0.008	Hbb	22Tyr	^{1}H
233	7.89	0.004	H	22Tyr	^{1}H
234	119.88	0.049	N	22Tyr	^{15}N
235	55.65	0.027	Ca	22Tyr	^{13}C
236	40.97	0.077	Cb	22Tyr	^{13}C
237	175.02	0.005	C	22Tyr	^{13}C
238	6.83	0.008	He*	22Tyr	^{1}H
239	131.92	0.053	Cd*	22Tyr	^{13}C
240	118.16	0.048	Ce*	22Tyr	^{13}C
241	7.07	0.006	Hd*	22Tyr	^{1}H
242	4.86	0.005	Ha	23Ser	^{1}H
243	3.83	0.003	Hb3	23Ser	^{1}H
244	3.83	0.004	Hb2	23Ser	^{1}H
245	115.41	0.046	N	23Ser	^{15}N
246	8.56	0.006	H	23Ser	^{1}H
247	64.78	0.047	Cb	23Ser	^{13}C
248	56.78	0.066	Ca	23Ser	^{13}C
249	174.48	0.032	C	23Ser	^{13}C
250	4.35	0.015	Ha	24Trp	^{1}H
251	3.47	0.010	Hbb	24Trp	^{1}H
252	3.20	0.008	Hba	24Trp	^{1}H
253	10.07	0.001	He1	24Trp	^{1}H
254	128.94	0.070	Ne1	24Trp	^{15}N
255	9.03	0.006	H	24Trp	^{1}H
256	127.53	0.057	N	24Trp	^{15}N
257	60.05	0.075	Ca	24Trp	^{13}C
258	30.82	0.071	Cb	24Trp	^{13}C
259	176.88	0.042	C	24Trp	^{13}C
260	7.29	0.015	He3	24Trp	^{1}H
261	7.03	0.011	Hd1	24Trp	^{1}H
262	6.72	0.008	Hz3	24Trp	^{1}H
263	6.81	0.009	Hh2	24Trp	^{1}H
264	119.19	0.067	Ce3	24Trp	^{13}C
265	131.43	0.000	Cd1	24Trp	^{13}C
266	7.29	0.007	Hz2	24Trp	^{1}H
267	121.18	0.102	Cz3	24Trp	^{13}C

268	122.83	0.026	Ch2	24Trp	^{13}C
269	114.37	0.040	Cz2	24Trp	^{13}C
270	60.03	0.044	Ca	25Ser	^{13}C
271	4.10	0.004	Ha	25Ser	^{1}H
272	3.87	0.014	Hb3	25Ser	^{1}H
273	3.87	0.013	Hb2	25Ser	^{1}H
274	112.24	0.081	N	25Ser	^{15}N
275	8.24	0.006	H	25Ser	^{1}H
276	62.93	0.132	Cb	25Ser	^{13}C
277	175.16	0.022	C	25Ser	^{13}C
278	34.24	0.070	Cg	26Gln	^{13}C
279	3.89	0.008	Ha	26Gln	^{1}H
280	1.84	0.007	Hga	26Gln	^{1}H
281	1.96	0.003	Hgb	26Gln	^{1}H
282	1.38	0.004	Hba	26Gln	^{1}H
283	1.50	0.008	Hbb	26Gln	^{1}H
284	7.47	0.005	H	26Gln	^{1}H
285	119.53	0.052	N	26Gln	^{15}N
286	7.35	0.005	He2b	26Gln	^{1}H
287	111.10	0.201	Ne2	26Gln	^{15}N
288	6.82	0.005	He2a	26Gln	^{1}H
289	56.96	0.106	Ca	26Gln	^{13}C
290	28.76	0.104	Cb	26Gln	^{13}C
291	176.07	0.030	C	26Gln	^{13}C
292	180.23	0.012	Cd	26Gln	^{13}C
293	3.04	0.008	Hbb	27Phe	^{1}H
294	4.59	0.006	Ha	27Phe	^{1}H
295	2.17	0.010	Hba	27Phe	^{1}H
296	116.67	0.047	N	27Phe	^{15}N
297	7.34	0.005	H	27Phe	^{1}H
298	39.91	0.056	Cb	27Phe	^{13}C
299	56.63	0.118	Ca	27Phe	^{13}C
300	129.28	0.012	Cz	27Phe	^{13}C
301	6.87	0.018	Hz	27Phe	^{1}H
302	130.98	0.053	Ce*	27Phe	^{13}C
303	7.06	0.006	He*	27Phe	^{1}H
304	175.11	0.014	C	27Phe	^{13}C
305	7.02	0.009	Hd*	27Phe	^{1}H
306	131.46	0.147	Cd*	27Phe	^{13}C
307	7.34	0.007	H	28Ser	^{1}H
308	114.12	0.093	N	28Ser	^{15}N
309	57.66	0.063	Ca	28Ser	^{13}C
310	63.93	0.079	Cb	28Ser	^{13}C
311	4.15	0.005	Ha	28Ser	^{1}H
312	3.18	0.014	Hba	28Ser	^{1}H

313	3.24	0.016	Hbb	28Ser	^{1}H
314	173.70	0.015	C	28Ser	^{13}C
315	8.37	0.005	H	29Asp	^{1}H
316	122.73	0.065	N	29Asp	^{15}N
317	54.19	0.101	Ca	29Asp	^{13}C
318	41.15	0.062	Cb	29Asp	^{13}C
319	4.60	0.010	Ha	29Asp	^{1}H
320	2.66	0.007	Hb3	29Asp	^{1}H
321	2.66	0.007	Hb2	29Asp	^{1}H
322	176.32	0.007	C	29Asp	^{13}C
323	7.97	0.002	H	30Val	^{1}H
324	119.20	0.089	N	30Val	^{15}N
325	62.42	0.040	Ca	30Val	^{13}C
326	33.02	0.079	Cb	30Val	^{13}C
327	0.90	0.003	Hg2*	30Val	^{1}H
328	4.10	0.002	Ha	30Val	^{1}H
329	0.90	0.003	Hg1*	30Val	^{1}H
330	2.07	0.021	Hb	30Val	^{1}H
331	21.09	0.347	Cga	30Val	^{13}C
332	20.50	0.090	Cga	30Val	^{13}C
333	176.23	0.001	C	30Val	^{13}C
334	8.50	0.010	H	31Glu	^{1}H
335	124.47	0.051	N	31Glu	^{15}N
336	56.71	0.098	Ca	31Glu	^{13}C
337	30.25	0.026	Cb	31Glu	^{13}C
338	36.37	0.000	Cg	31Glu	^{13}C
339	2.26	0.004	Hb3	31Glu	^{1}H
340	2.26	0.005	Hg2	31Glu	^{1}H
341	4.26	0.004	Ha	31Glu	^{1}H
342	176.53	0.007	C	31Glu	^{13}C
343	2.26	0.004	Hg3	31Glu	^{1}H
344	2.26	0.004	Hb2	31Glu	^{1}H
345	122.14	0.082	N	32Glu	^{15}N
346	8.43	0.003	H	32Glu	^{1}H
347	30.64	0.195	Cb	32Glu	^{13}C
348	56.73	0.064	Ca	32Glu	^{13}C
349	2.26	0.013	Hbb	32Glu	^{1}H
350	2.26	0.013	Hba	32Glu	^{1}H
351	36.33	0.031	Cg	32Glu	^{13}C
352	4.23	0.014	Ha	32Glu	^{1}H
353	1.94	0.008	Hg2	32Glu	^{1}H
354	1.94	0.008	Hg3	32Glu	^{1}H
355	176.17	0.017	C	32Glu	^{13}C
356	8.50	0.013	H	33Asn	^{1}H
357	119.56	0.103	N	33Asn	^{15}N

358	6.92	0.003	Hd2a	33Asn	^{1}H
359	112.70	0.218	Nd2	33Asn	^{15}N
360	7.60	0.015	Hd2b	33Asn	^{1}H
361	53.36	0.154	Ca	33Asn	^{13}C
362	38.99	0.109	Cb	33Asn	^{13}C
363	4.71	0.037	Ha	33Asn	^{1}H
364	2.83	0.013	Hbb	33Asn	^{1}H
365	2.73	0.011	Hba	33Asn	^{1}H
366	174.98	0.006	C	33Asn	^{13}C
367	177.04	0.005	Cg	33Asn	^{13}C
368	8.33	0.007	H	34Arg	^{1}H
369	122.04	0.048	N	34Arg	^{15}N
370	56.15	0.027	Ca	34Arg	^{13}C
371	30.89	0.081	Cb	34Arg	^{13}C
372	27.05	0.044	Cg	34Arg	^{13}C
373	4.41	0.009	Ha	34Arg	^{1}H
374	3.19	0.007	Hdb	34Arg	^{1}H
375	43.36	0.106	Cd	34Arg	^{13}C
376	1.62	0.010	Hg3	34Arg	^{1}H
377	1.77	0.003	Hba	34Arg	^{1}H
378	1.89	0.007	Hbb	34Arg	^{1}H
379	1.62	0.010	Hg2	34Arg	^{1}H
380	176.47	0.016	C	34Arg	^{13}C
381	3.19	0.003	Hda	34Arg	^{1}H
382	8.28	0.004	H	35Thr	^{1}H
383	115.85	0.051	N	35Thr	^{15}N
384	62.08	0.061	Ca	35Thr	^{13}C
385	69.83	0.062	Cb	35Thr	^{13}C
386	4.22	0.005	Hb	35Thr	^{1}H
387	21.60	0.123	Cg2	35Thr	^{13}C
388	1.20	0.005	Hg1	35Thr	^{1}H
389	174.43	0.008	C	35Thr	^{13}C
390	4.33	0.008	Ha	35Thr	^{1}H
391	8.40	0.004	H	36Glu	^{1}H
392	123.07	0.046	N	36Glu	^{15}N
393	56.40	0.164	Ca	36Glu	^{13}C
394	30.38	0.022	Cb	36Glu	^{13}C
395	36.22	0.000	Cg	36Glu	^{13}C
396	4.31	0.008	Ha	36Glu	^{1}H
397	2.25	0.002	Hba	36Glu	^{1}H
398	175.64	0.002	C	36Glu	^{13}C
399	8.31	0.007	H	37Ala	^{1}H
400	126.19	0.039	N	37Ala	^{15}N
401	50.53	0.098	Ca	37Ala	^{13}C
402	18.63	0.098	Cb	37Ala	^{13}C

403	1.35	0.002	Hb*	37Ala	^{1}H
404	4.61	0.002	Ha	37Ala	^{1}H
405	175.42	0.000	C	37Ala	^{13}C
406	3.66	0.006	Hda	38Pro	^{1}H
407	1.93	0.010	Hga	38Pro	^{1}H
408	63.22	0.110	Ca	38Pro	^{13}C
409	32.13	0.102	Cb	38Pro	^{13}C
410	27.58	0.051	Cg	38Pro	^{13}C
411	3.77	0.005	Hdb	38Pro	^{1}H
412	4.43	0.005	Ha	38Pro	^{1}H
413	50.61	0.042	Cd	38Pro	^{13}C
414	2.29	0.006	Hbb	38Pro	^{1}H
415	2.02	0.012	Hgb	38Pro	^{1}H
416	177.03	0.000	C	38Pro	^{13}C
417	1.94	0.011	Hba	38Pro	^{1}H
418	8.64	0.002	H	39Glu	^{1}H
419	121.30	0.093	N	39Glu	^{15}N
420	56.84	0.031	Ca	39Glu	^{13}C
421	30.44	0.228	Cb	39Glu	^{13}C
422	36.33	0.015	Cg	39Glu	^{13}C
423	4.30	0.009	Ha	39Glu	^{1}H
424	2.30	0.005	Hg2	39Glu	^{1}H
425	177.08	0.009	C	39Glu	^{13}C
426	1.99	0.023	Hba	39Glu	^{1}H
427	2.05	0.023	Hbb	39Glu	^{1}H
428	2.30	0.005	Hg3	39Glu	^{1}H
429	110.32	0.034	N	40Gly	^{15}N
430	8.52	0.003	H	40Gly	^{1}H
431	45.49	0.010	Ca	40Gly	^{13}C
432	4.08	0.007	Hab	40Gly	^{1}H
433	4.08	0.007	Haa	40Gly	^{1}H
434	174.60	0.037	C	40Gly	^{13}C
435	8.27	0.004	H	41Thr	^{1}H
436	113.18	0.040	N	41Thr	^{15}N
437	62.27	0.111	Ca	41Thr	^{13}C
438	69.92	0.032	Cb	41Thr	^{13}C
439	4.39	0.024	Ha	41Thr	^{1}H
440	1.20	0.005	Hg1	41Thr	^{1}H
441	21.79	0.000	Cg2	41Thr	^{13}C
442	175.50	0.040	C	41Thr	^{13}C
443	4.31	0.005	Hb	41Thr	^{1}H
444	123.14	0.060	N	42Glu	^{15}N
445	8.85	0.010	H	42Glu	^{1}H
446	57.49	0.055	Ca	42Glu	^{13}C
447	29.92	0.190	Cb	42Glu	^{13}C

448	2.31	0.006	Hgb	42Glu	^{1}H
449	36.31	0.031	Cg	42Glu	^{13}C
450	4.37	0.009	Ha	42Glu	^{1}H
451	1.96	0.006	Hba	42Glu	^{1}H
452	177.16	0.053	C	42Glu	^{13}C
453	2.31	0.005	Hga	42Glu	^{1}H
454	2.07	0.013	Hbb	42Glu	^{1}H
455	115.68	0.117	N	43Ser	^{15}N
456	8.51	0.016	H	43Ser	^{1}H
457	58.92	0.044	Ca	43Ser	^{13}C
458	63.71	0.044	Cb	43Ser	^{13}C
459	4.43	0.007	Ha	43Ser	^{1}H
460	3.98	0.012	Hbb	43Ser	^{1}H
461	3.89	0.011	Hba	43Ser	^{1}H
462	175.67	0.022	C	43Ser	^{13}C
463	29.41	0.040	Cb	44Glu	^{13}C
464	123.10	0.044	N	44Glu	^{15}N
465	8.35	0.023	H	44Glu	^{1}H
466	36.05	0.000	Cg	44Glu	^{13}C
467	59.07	0.051	Ca	44Glu	^{13}C
468	2.32	0.001	Hgb	44Glu	^{1}H
469	2.32	0.003	Hga	44Glu	^{1}H
470	3.99	0.004	Ha	44Glu	^{1}H
471	2.06	0.010	Hba	44Glu	^{1}H
472	177.92	0.014	C	44Glu	^{13}C
473	2.06	0.010	Hbb	44Glu	^{1}H
474	120.96	0.067	N	45Ala	^{15}N
475	8.23	0.005	H	45Ala	^{1}H
476	55.16	0.081	Ca	45Ala	^{13}C
477	18.58	0.148	Cb	45Ala	^{13}C
478	4.13	0.005	Ha	45Ala	^{1}H
479	1.47	0.009	Hb*	45Ala	^{1}H
480	180.64	0.004	C	45Ala	^{13}C
481	118.87	0.057	N	46Val	^{15}N
482	7.57	0.009	H	46Val	^{1}H
483	66.19	0.030	Ca	46Val	^{13}C
484	31.81	0.068	Cb	46Val	^{13}C
485	1.04	0.006	Hgb*	46Val	^{1}H
486	0.55	0.004	Hga*	46Val	^{1}H
487	3.50	0.004	Ha	46Val	^{1}H
488	1.96	0.006	Hb	46Val	^{1}H
489	22.45	0.074	Cga	46Val	^{13}C
490	22.91	0.201	Cgb	46Val	^{13}C
491	176.72	0.015	C	46Val	^{13}C
492	32.51	0.084	Cb	47Lys	^{13}C

493	117.92	0.047	N	47Lys	^{15}N
494	7.44	0.008	H	47Lys	^{1}H
495	59.80	0.063	Ca	47Lys	^{13}C
496	1.60	0.022	Hba	47Lys	^{1}H
497	1.87	0.008	Hbb	47Lys	^{1}H
498	2.72	0.008	Heb	47Lys	^{1}H
499	1.13	0.010	Hga	47Lys	^{1}H
500	1.31	0.000	Hgb	47Lys	^{1}H
501	2.72	0.008	Hea	47Lys	^{1}H
502	30.20	0.000	Cd	47Lys	^{13}C
503	3.72	0.008	Ha	47Lys	^{1}H
504	178.89	0.014	C	47Lys	^{13}C
505	116.49	0.075	N	48Gln	^{15}N
506	8.17	0.004	H	48Gln	^{1}H
507	7.88	0.002	He2b	48Gln	^{1}H
508	6.65	0.032	He2a	48Gln	^{1}H
509	112.09	0.276	Ne2	48Gln	^{15}N
510	59.03	0.085	Ca	48Gln	^{13}C
511	28.10	0.068	Cb	48Gln	^{13}C
512	2.37	0.004	Hgb	48Gln	^{1}H
513	33.21	0.100	Cg	48Gln	^{13}C
514	2.37	0.004	Hga	48Gln	^{1}H
515	1.98	0.008	Hba	48Gln	^{1}H
516	2.08	0.007	Hbb	48Gln	^{1}H
517	3.71	0.011	Ha	48Gln	^{1}H
518	177.63	0.039	C	48Gln	^{13}C
519	179.78	0.069	Cd	48Gln	^{13}C
520	120.36	0.078	N	49Ala	^{15}N
521	7.79	0.007	H	49Ala	^{1}H
522	54.93	0.055	Ca	49Ala	^{13}C
523	18.71	0.061	Cb	49Ala	^{13}C
524	1.46	0.008	Hb*	49Ala	^{1}H
525	4.23	0.004	Ha	49Ala	^{1}H
526	179.81	0.045	C	49Ala	^{13}C
527	23.49	0.115	Cdb	50Leu	^{13}C
528	8.49	0.006	H	50Leu	^{1}H
529	120.59	0.052	N	50Leu	^{15}N
530	58.58	0.052	Ca	50Leu	^{13}C
531	42.28	0.075	Cb	50Leu	^{13}C
532	23.33	0.068	Cda	50Leu	^{13}C
533	4.15	0.007	Ha	50Leu	^{1}H
534	0.12	0.006	Hdb*	50Leu	^{1}H
535	-0.37	0.004	Hda*	50Leu	^{1}H
536	1.23	0.009	Hg	50Leu	^{1}H
537	26.97	0.092	Cg	50Leu	^{13}C

538	1.50	0.020	Hbb	50Leu	^{1}H
539	1.43	0.018	Hba	50Leu	^{1}H
540	179.23	0.009	C	50Leu	^{13}C
541	30.89	0.091	Cb	51Arg	^{13}C
542	118.93	0.060	N	51Arg	^{15}N
543	8.17	0.008	H	51Arg	^{1}H
544	60.95	0.034	Ca	51Arg	^{13}C
545	178.62	0.027	C	51Arg	^{13}C
546	3.77	0.009	Ha	51Arg	^{1}H
547	116.53	0.059	N	52Glu	^{15}N
548	7.91	0.006	H	52Glu	^{1}H
549	59.38	0.038	Ca	52Glu	^{13}C
550	36.76	0.047	Cg	52Glu	^{13}C
551	4.16	0.006	Ha	52Glu	^{1}H
552	29.64	0.036	Cb	52Glu	^{13}C
553	2.38	0.005	Hga	52Glu	^{1}H
554	2.17	0.022	Hb3	52Glu	^{1}H
555	2.59	0.007	Hgb	52Glu	^{1}H
556	2.17	0.020	Hb2	52Glu	^{1}H
557	179.94	0.001	C	52Glu	^{13}C
558	123.23	0.084	N	53Ala	^{15}N
559	8.90	0.004	H	53Ala	^{1}H
560	55.11	0.025	Ca	53Ala	^{13}C
561	19.48	0.097	Cb	53Ala	^{13}C
562	1.69	0.009	Hb*	53Ala	^{1}H
563	4.47	0.004	Ha	53Ala	^{1}H
564	181.68	0.067	C	53Ala	^{13}C
565	9.19	0.003	H	54Gly	^{1}H
566	108.71	0.059	N	54Gly	^{15}N
567	47.11	0.093	Ca	54Gly	^{13}C
568	4.05	0.009	Hab	54Gly	^{1}H
569	175.33	0.000	C	54Gly	^{13}C
570	4.04	0.020	Haa	54Gly	^{1}H
571	39.84	0.047	Cb	55Asp	^{13}C
572	122.89	0.072	N	55Asp	^{15}N
573	8.22	0.007	H	55Asp	^{1}H
574	57.75	0.016	Ca	55Asp	^{13}C
575	179.02	0.036	C	55Asp	^{13}C
576	2.75	0.005	Hba	55Asp	^{1}H
577	3.04	0.014	Hbb	55Asp	^{1}H
578	120.75	0.088	N	56Glu	^{15}N
579	8.28	0.003	H	56Glu	^{1}H
580	59.60	0.026	Ca	56Glu	^{13}C
581	36.37	0.040	Cg	56Glu	^{13}C
582	178.47	0.007	C	56Glu	^{13}C

583	29.78	0.093	Cb	56Glu	^{13}C
584	2.30	0.007	Hb3	56Glu	^{1}H
585	2.30	0.007	Hb2	56Glu	^{1}H
586	2.44	0.006	Hga	56Glu	^{1}H
587	4.18	0.011	Ha	56Glu	^{1}H
588	2.44	0.006	Hgb	56Glu	^{1}H
589	8.67	0.006	H	57Phe	^{1}H
590	121.60	0.058	N	57Phe	^{15}N
591	61.91	0.119	Ca	57Phe	^{13}C
592	177.01	0.004	C	57Phe	^{13}C
593	40.03	0.091	Cb	57Phe	^{13}C
594	3.45	0.019	Hb3	57Phe	^{1}H
595	3.45	0.019	Hb2	57Phe	^{1}H
596	7.41	0.019	Hd*	57Phe	^{1}H
597	131.76	0.079	Cd*	57Phe	^{13}C
598	131.91	0.005	Ce*	57Phe	^{13}C
599	7.25	0.009	He*	57Phe	^{1}H
600	4.34	0.005	Ha	57Phe	^{1}H
601	8.14	0.008	H	58Glu	^{1}H
602	116.77	0.045	N	58Glu	^{15}N
603	29.78	0.124	Cb	58Glu	^{13}C
604	58.79	0.068	Ca	58Glu	^{13}C
605	36.69	0.080	Cg	58Glu	^{13}C
606	177.66	0.021	C	58Glu	^{13}C
607	3.67	0.007	Ha	58Glu	^{1}H
608	2.14	0.003	Hbb	58Glu	^{1}H
609	2.62	0.005	Hgb	58Glu	^{1}H
610	2.01	0.007	Hga	58Glu	^{1}H
611	2.02	0.008	Hba	58Glu	^{1}H
612	7.51	0.008	H	59Leu	^{1}H
613	117.34	0.054	N	59Leu	^{15}N
614	56.97	0.049	Ca	59Leu	^{13}C
615	42.98	0.051	Cb	59Leu	^{13}C
616	178.95	0.015	C	59Leu	^{13}C
617	24.55	0.063	Cdb	59Leu	^{13}C
618	4.20	0.009	Ha	59Leu	^{1}H
619	1.79	0.004	Hbb	59Leu	^{1}H
620	1.68	0.013	Hg	59Leu	^{1}H
621	1.70	0.011	Hba	59Leu	^{1}H
622	27.38	0.264	Cg	59Leu	^{13}C
623	24.55	0.063	Cda	59Leu	^{13}C
624	0.91	0.008	Hda*	59Leu	^{1}H
625	0.91	0.008	Hdb*	59Leu	^{1}H
626	57.90	0.087	Ca	60Arg	^{13}C
627	31.15	0.059	Cb	60Arg	^{13}C

628	117.78	0.096	N	60Arg	^{15}N
629	8.07	0.011	H	60Arg	^{1}H
630	177.34	0.043	C	60Arg	^{13}C
631	27.41	0.125	Cg	60Arg	^{13}C
632	43.38	0.085	Cd	60Arg	^{13}C
633	1.15	0.005	Hga	60Arg	^{1}H
634	1.48	0.008	Hgb	60Arg	^{1}H
635	3.02	0.002	Hd3	60Arg	^{1}H
636	1.25	0.013	Hba	60Arg	^{1}H
637	3.02	0.003	Hd2	60Arg	^{1}H
638	3.96	0.003	Ha	60Arg	^{1}H
639	1.52	0.011	Hbb	60Arg	^{1}H
640	2.24	0.007	Hba	61Tyr	^{1}H
641	2.79	0.006	Hbb	61Tyr	^{1}H
642	8.05	0.007	H	61Tyr	^{1}H
643	119.06	0.055	N	61Tyr	^{15}N
644	56.29	0.027	Ca	61Tyr	^{13}C
645	37.02	0.069	Cb	61Tyr	^{13}C
646	4.74	0.011	Ha	61Tyr	^{1}H
647	175.21	0.020	C	61Tyr	^{13}C
648	133.46	0.032	Cd*	61Tyr	^{13}C
649	6.89	0.007	Hd*	61Tyr	^{1}H
650	6.77	0.006	He*	61Tyr	^{1}H
651	117.44	0.058	Ce*	61Tyr	^{13}C
652	7.42	0.006	H	62Arg	^{1}H
653	119.96	0.059	N	62Arg	^{15}N
654	30.14	0.059	Cb	62Arg	^{13}C
655	59.02	0.184	Ca	62Arg	^{13}C
656	4.06	0.006	Ha	62Arg	^{1}H
657	178.18	0.000	C	62Arg	^{13}C
658	1.72	0.005	Hg2	62Arg	^{1}H
659	3.28	0.005	Hd2	62Arg	^{1}H
660	1.72	0.005	Hg3	62Arg	^{1}H
661	3.28	0.005	Hd3	62Arg	^{1}H
662	27.10	0.082	Cg	62Arg	^{13}C
663	1.97	0.005	Hbb	62Arg	^{1}H
664	43.34	0.051	Cd	62Arg	^{13}C
665	1.82	0.015	Hba	62Arg	^{1}H
666	177.39	0.011	C	63Arg	^{13}C
667	43.10	0.115	Cd	63Arg	^{13}C
668	4.22	0.003	Ha	63Arg	^{1}H
669	57.73	0.057	Ca	63Arg	^{13}C
670	1.64	0.003	Hg3	63Arg	^{1}H
671	1.64	0.003	Hg2	63Arg	^{1}H
672	3.21	0.009	Hd3	63Arg	^{1}H

673	1.82	0.015	Hba	63Arg	^{1}H
674	3.21	0.009	Hd2	63Arg	^{1}H
675	27.28	0.116	Cg	63Arg	^{13}C
676	29.18	0.157	Cb	63Arg	^{13}C
677	1.97	0.006	Hbb	63Arg	^{1}H
678	122.00	0.061	N	64Ala	^{15}N
679	7.87	0.008	H	64Ala	^{1}H
680	54.58	0.121	Ca	64Ala	^{13}C
681	18.64	0.112	Cb	64Ala	^{13}C
682	1.30	0.005	Hb*	64Ala	^{1}H
683	4.10	0.004	Ha	64Ala	^{1}H
684	179.97	0.014	C	64Ala	^{13}C
685	8.58	0.007	H	65Phe	^{1}H
686	115.86	0.077	N	65Phe	^{15}N
687	58.22	0.090	Ca	65Phe	^{13}C
688	38.88	0.057	Cb	65Phe	^{13}C
689	177.38	0.017	C	65Phe	^{13}C
690	3.38	0.008	Hbb	65Phe	^{1}H
691	3.15	0.008	Hba	65Phe	^{1}H
692	4.80	0.014	Ha	65Phe	^{1}H
693	7.33	0.012	Hd*	65Phe	^{1}H
694	131.01	0.059	Cd*	65Phe	^{13}C
695	116.94	0.193	N	66Ser	^{15}N
696	8.23	0.025	H	66Ser	^{1}H
697	61.76	0.094	Ca	66Ser	^{13}C
698	63.01	0.048	Cb	66Ser	^{13}C
699	176.47	0.038	C	66Ser	^{13}C
700	4.22	0.008	Ha	66Ser	^{1}H
701	3.98	0.004	Hb2	66Ser	^{1}H
702	3.98	0.004	Hb3	66Ser	^{1}H
703	8.48	0.006	H	67Asp	^{1}H
704	121.62	0.117	N	67Asp	^{15}N
705	56.83	0.019	Ca	67Asp	^{13}C
706	40.18	0.076	Cb	67Asp	^{13}C
707	178.40	0.042	C	67Asp	^{13}C
708	2.71	0.002	Hb3	67Asp	^{1}H
709	2.71	0.003	Hb2	67Asp	^{1}H
710	4.49	0.009	Ha	67Asp	^{1}H
711	121.20	0.071	N	68Leu	^{15}N
712	7.89	0.016	H	68Leu	^{1}H
713	57.97	0.088	Ca	68Leu	^{13}C
714	42.56	0.041	Cb	68Leu	^{13}C
715	179.05	0.017	C	68Leu	^{13}C
716	1.73	0.005	Hbb	68Leu	^{1}H
717	1.58	0.007	Hg	68Leu	^{1}H

718	4.10	0.005	Ha	68Leu	^{1}H
719	1.46	0.006	Hba	68Leu	^{1}H
720	24.48	0.295	Cdb	68Leu	^{13}C
721	24.56	0.100	Cdb	68Leu	^{13}C
722	0.60	0.011	Hdb*	68Leu	^{1}H
723	27.04	0.050	Cg	68Leu	^{13}C
724	0.60	0.014	Hda*	68Leu	^{1}H
725	4.09	0.028	Ha	69Thr	^{1}H
726	68.44	0.186	Cb	69Thr	^{13}C
727	1.15	0.006	Hg2*	69Thr	^{1}H
728	65.84	0.114	Ca	69Thr	^{13}C
729	4.31	0.008	Hb	69Thr	^{1}H
730	110.28	0.084	N	69Thr	^{15}N
731	8.49	0.012	H	69Thr	^{1}H
732	177.46	0.041	C	69Thr	^{13}C
733	22.69	0.026	Cg2	69Thr	^{13}C
734	7.75	0.006	H	70Ser	^{1}H
735	116.58	0.099	N	70Ser	^{15}N
736	60.68	0.056	Ca	70Ser	^{13}C
737	63.30	0.129	Cb	70Ser	^{13}C
738	4.04	0.002	Hb2	70Ser	^{1}H
739	4.04	0.002	Hb3	70Ser	^{1}H
740	4.33	0.013	Ha	70Ser	^{1}H
741	175.61	0.028	C	70Ser	^{13}C
742	34.13	0.075	Cg	71Gln	^{13}C
743	119.13	0.081	N	71Gln	^{15}N
744	7.75	0.005	H	71Gln	^{1}H
745	7.25	0.003	He2b	71Gln	^{1}H
746	111.43	0.274	Ne2	71Gln	^{15}N
747	6.77	0.030	He2a	71Gln	^{1}H
748	57.41	0.075	Ca	71Gln	^{13}C
749	29.53	0.121	Cb	71Gln	^{13}C
750	2.50	0.008	Hgb	71Gln	^{1}H
751	176.34	0.017	C	71Gln	^{13}C
752	2.12	0.022	Hbb	71Gln	^{1}H
753	4.11	0.004	Ha	71Gln	^{1}H
754	2.38	0.006	Hga	71Gln	^{1}H
755	2.12	0.022	Hba	71Gln	^{1}H
756	179.83	0.038	Cd	71Gln	^{13}C
757	55.44	0.097	Ca	72Leu	^{13}C
758	118.80	0.067	N	72Leu	^{15}N
759	7.58	0.004	H	72Leu	^{1}H
760	43.89	0.068	Cb	72Leu	^{13}C
761	1.54	0.008	Hba	72Leu	^{1}H
762	1.96	0.012	Hbb	72Leu	^{1}H

763	0.95	0.006	Hda*	72Leu	^{1}H
764	4.31	0.006	Ha	72Leu	^{1}H
765	174.29	0.000	C	72Leu	^{13}C
766	1.08	0.013	Hdb*	72Leu	^{1}H
767	1.64	0.006	Hg	72Leu	^{1}H
768	27.15	0.075	Cdb	72Leu	^{13}C
769	23.86	0.048	Cda	72Leu	^{13}C
770	27.12	0.116	Cg	72Leu	^{13}C
771	4.64	0.006	Ha	73His	^{1}H
772	3.02	0.005	Hba	73His	^{1}H
773	3.25	0.006	Hbb	73His	^{1}H
774	28.75	0.049	Cb	73His	^{13}C
775	54.96	0.022	Ca	73His	^{13}C
776	174.45	0.000	C	73His	^{13}C
777	122.47	0.040	N	74Ile	^{15}N
778	8.77	0.010	H	74Ile	^{1}H
779	61.96	0.057	Ca	74Ile	^{13}C
780	14.40	0.122	Cd1	74Ile	^{13}C
781	0.90	0.014	Hg2*	74Ile	^{1}H
782	0.94	0.008	Hd1*	74Ile	^{1}H
783	17.50	0.162	Cg2	74Ile	^{13}C
784	1.39	0.007	Hg1a	74Ile	^{1}H
785	28.32	0.082	Cg1	74Ile	^{13}C
786	1.59	0.004	Hg1b	74Ile	^{1}H
787	1.79	0.005	Hb	74Ile	^{1}H
788	4.21	0.006	Ha	74Ile	^{1}H
789	38.95	0.120	Cb	74Ile	^{13}C
790	7.95	0.005	H	75Thr	^{1}H
791	120.54	0.072	N	75Thr	^{15}N
792	1.52	0.002	Hg2*	75Thr	^{1}H
793	23.02	0.024	Cg2	75Thr	^{13}C
794	4.34	0.004	Hb	75Thr	^{1}H
795	4.32	0.014	Ha	75Thr	^{1}H
796	4.33	0.012	Hab	77Gly	^{1}H
797	3.68	0.008	Haa	77Gly	^{1}H
798	44.79	0.105	Ca	77Gly	^{13}C
799	130.04	0.033	N	78Thr	^{15}N
800	8.80	0.012	H	78Thr	^{1}H
801	4.15	0.007	Hb	78Thr	^{1}H
802	1.52	0.007	Hg2*	78Thr	^{1}H
803	4.15	0.006	Ha	78Thr	^{1}H
804	51.12	0.080	Ca	79Ala	^{13}C
805	177.64	0.000	C	79Ala	^{13}C
806	21.52	0.043	Cb	79Ala	^{13}C
807	1.65	0.003	Hb*	79Ala	^{1}H

808	4.83	0.000	Ha	79Ala	^{1}H
809	121.37	0.069	N	80Tyr	^{15}N
810	8.86	0.011	H	80Tyr	^{1}H
811	60.85	0.013	Ca	80Tyr	^{13}C
812	4.65	0.000	Ha	80Tyr	^{1}H
813	38.16	0.072	Cb	80Tyr	^{13}C
814	176.56	0.017	C	80Tyr	^{13}C
815	7.03	0.009	He*	80Tyr	^{1}H
816	2.76	0.022	Hbb	80Tyr	^{1}H
817	2.51	0.008	Hba	80Tyr	^{1}H
818	6.56	0.008	Hd*	80Tyr	^{1}H
819	118.18	0.029	Ce*	80Tyr	^{13}C
820	132.47	0.011	Cd*	80Tyr	^{13}C
821	115.12	0.088	N	81Gln	^{15}N
822	8.53	0.007	H	81Gln	^{1}H
823	7.78	0.002	He2b	81Gln	^{1}H
824	7.00	0.003	He2a	81Gln	^{1}H
825	112.66	0.251	Ne2	81Gln	^{15}N
826	59.91	0.076	Ca	81Gln	^{13}C
827	28.02	0.066	Cb	81Gln	^{13}C
828	178.33	0.040	C	81Gln	^{13}C
829	3.80	0.016	Ha	81Gln	^{1}H
830	2.04	0.011	Hba	81Gln	^{1}H
831	2.13	0.005	Hbb	81Gln	^{1}H
832	2.52	0.005	Hg2	81Gln	^{1}H
833	34.76	0.142	Cg	81Gln	^{13}C
834	2.52	0.005	Hg3	81Gln	^{1}H
835	180.79	0.011	Cd	81Gln	^{13}C
836	7.36	0.012	H	82Ser	^{1}H
837	114.42	0.075	N	82Ser	^{15}N
838	3.92	0.008	Hba	82Ser	^{1}H
839	3.92	0.010	Hbb	82Ser	^{1}H
840	4.25	0.002	Ha	82Ser	^{1}H
841	62.09	0.000	Ca	82Ser	^{13}C
842	63.34	0.000	Cb	82Ser	^{13}C
843	3.94	0.008	Ha	83Phe	^{1}H
844	2.67	0.010	Hbb	83Phe	^{1}H
845	61.34	0.036	Ca	83Phe	^{13}C
846	175.14	0.001	C	83Phe	^{13}C
847	39.43	0.067	Cb	83Phe	^{13}C
848	2.47	0.029	Hba	83Phe	^{1}H
849	132.16	0.031	Cd*	83Phe	^{13}C
850	7.01	0.007	Hd*	83Phe	^{1}H
851	119.92	0.067	N	84Glu	^{15}N
852	8.66	0.003	H	84Glu	^{1}H

853	59.15	0.048	Ca	84Glu	^{13}C
854	30.46	0.082	Cb	84Glu	^{13}C
855	177.01	0.002	C	84Glu	^{13}C
856	1.91	0.008	Hbb	84Glu	^{1}H
857	36.20	0.092	Cg	84Glu	^{13}C
858	3.33	0.009	Ha	84Glu	^{1}H
859	2.23	0.008	Hgb	84Glu	^{1}H
860	2.14	0.006	Hga	84Glu	^{1}H
861	1.31	0.008	Hba	84Glu	^{1}H
862	59.10	0.057	Ca	85Gln	^{13}C
863	115.03	0.058	N	85Gln	^{15}N
864	7.59	0.005	H	85Gln	^{1}H
865	6.79	0.019	He2a	85Gln	^{1}H
866	7.52	0.001	He2b	85Gln	^{1}H
867	111.76	0.309	Ne2	85Gln	^{15}N
868	28.59	0.066	Cb	85Gln	^{13}C
869	33.89	0.082	Cg	85Gln	^{13}C
870	3.89	0.009	Ha	85Gln	^{1}H
871	2.10	0.019	Hba	85Gln	^{1}H
872	2.48	0.004	Hgb	85Gln	^{1}H
873	2.40	0.006	Hga	85Gln	^{1}H
874	2.13	0.007	Hbb	85Gln	^{1}H
875	178.57	0.010	C	85Gln	^{13}C
876	180.22	0.007	Cd	85Gln	^{13}C
877	31.80	0.090	Cb	86Val	^{13}C
878	7.16	0.007	H	86Val	^{1}H
879	118.48	0.044	N	86Val	^{15}N
880	66.10	0.068	Ca	86Val	^{13}C
881	23.38	0.086	Cgb	86Val	^{13}C
882	1.00	0.014	Hgb*	86Val	^{1}H
883	22.12	0.108	Cga	86Val	^{13}C
884	1.88	0.005	Hb	86Val	^{1}H
885	0.65	0.015	Hga*	86Val	^{1}H
886	3.64	0.003	Ha	86Val	^{1}H
887	177.90	0.000	C	86Val	^{13}C
888	119.83	0.105	N	87Val	^{15}N
889	7.98	0.003	H	87Val	^{1}H
890	66.39	0.082	Ca	87Val	^{13}C
891	31.04	0.118	Cb	87Val	^{13}C
892	22.68	0.048	Cga	87Val	^{13}C
893	0.15	0.008	Hgb*	87Val	^{1}H
894	-0.02	0.004	Hga*	87Val	^{1}H
895	23.09	0.240	Cgb	87Val	^{13}C
896	1.12	0.007	Hb	87Val	^{1}H
897	3.03	0.006	Ha	87Val	^{1}H

898	177.85	0.015	C	87Val	^{13}C
899	7.49	0.004	Hd2b	88Asn	^{1}H
900	6.73	0.008	Hd2a	88Asn	^{1}H
901	109.11	0.065	Nd2	88Asn	^{15}N
902	117.63	0.099	N	88Asn	^{15}N
903	55.45	0.104	Ca	88Asn	^{13}C
904	8.44	0.004	H	88Asn	^{1}H
905	37.33	0.059	Cb	88Asn	^{13}C
906	4.28	0.006	Ha	88Asn	^{1}H
907	2.66	0.004	Hba	88Asn	^{1}H
908	2.98	0.010	Hbb	88Asn	^{1}H
909	178.99	0.067	C	88Asn	^{13}C
910	175.98	0.000	Cg	88Asn	^{13}C
911	118.94	0.104	N	89Glu	^{15}N
912	7.38	0.006	H	89Glu	^{1}H
913	59.02	0.119	Ca	89Glu	^{13}C
914	29.12	0.152	Cb	89Glu	^{13}C
915	4.00	0.005	Ha	89Glu	^{1}H
916	35.80	0.070	Cg	89Glu	^{13}C
917	179.30	0.024	C	89Glu	^{13}C
918	2.31	0.010	Hgb	89Glu	^{1}H
919	2.29	0.009	Hga	89Glu	^{1}H
920	2.08	0.012	Hbb	89Glu	^{1}H
921	2.06	0.013	Hba	89Glu	^{1}H
922	120.80	0.068	N	90Leu	^{15}N
923	7.61	0.007	H	90Leu	^{1}H
924	57.49	0.040	Ca	90Leu	^{13}C
925	1.10	0.009	Hba	90Leu	^{1}H
926	41.50	0.108	Cb	90Leu	^{13}C
927	3.82	0.011	Ha	90Leu	^{1}H
928	1.18	0.026	Hbb	90Leu	^{1}H
929	0.98	0.008	Hg	90Leu	^{1}H
930	24.31	0.049	Cdb	90Leu	^{13}C
931	27.00	0.048	Cg	90Leu	^{13}C
932	0.48	0.011	Hdb*	90Leu	^{1}H
933	0.43	0.009	Hda*	90Leu	^{1}H
934	23.93	0.126	Cda	90Leu	^{13}C
935	178.25	0.032	C	90Leu	^{13}C
936	7.02	0.005	H	91Phe	^{1}H
937	113.37	0.059	N	91Phe	^{15}N
938	39.47	0.055	Cb	91Phe	^{13}C
939	57.50	0.048	Ca	91Phe	^{13}C
940	6.70	0.005	He*	91Phe	^{1}H
941	129.94	0.000	Ce*	91Phe	^{13}C
942	4.53	0.012	Ha	91Phe	^{1}H

943	3.28	0.009	Hbb	91Phe	^{1}H
944	2.41	0.011	Hba	91Phe	^{1}H
945	176.47	0.026	C	91Phe	^{13}C
946	7.15	0.008	Hd*	91Phe	^{1}H
947	132.18	0.072	Cd*	91Phe	^{13}C
948	123.99	0.046	N	92Arg	^{15}N
949	7.24	0.003	H	92Arg	^{1}H
950	59.74	0.092	Ca	92Arg	^{13}C
951	30.17	0.062	Cb	92Arg	^{13}C
952	3.96	0.007	Ha	92Arg	^{1}H
953	1.62	0.013	Hg3	92Arg	^{1}H
954	1.62	0.013	Hg2	92Arg	^{1}H
955	26.37	0.058	Cg	92Arg	^{13}C
956	3.20	0.004	Hd3	92Arg	^{1}H
957	43.21	0.016	Cd	92Arg	^{13}C
958	3.20	0.004	Hd2	92Arg	^{1}H
959	1.97	0.013	Hbb	92Arg	^{1}H
960	1.74	0.007	Hba	92Arg	^{1}H
961	177.15	0.025	C	92Arg	^{13}C
962	4.82	0.004	Ha	93Asp	^{1}H
963	8.71	0.003	H	93Asp	^{1}H
964	116.87	0.049	N	93Asp	^{15}N
965	54.07	0.015	Ca	93Asp	^{13}C
966	40.75	0.034	Cb	93Asp	^{13}C
967	2.85	0.005	Hbb	93Asp	^{1}H
968	2.48	0.010	Hba	93Asp	^{1}H
969	175.59	0.022	C	93Asp	^{13}C
970	7.58	0.005	H	94Gly	^{1}H
971	108.98	0.088	N	94Gly	^{15}N
972	43.82	0.045	Ca	94Gly	^{13}C
973	3.68	0.003	Haa	94Gly	^{1}H
974	4.48	0.029	Hab	94Gly	^{1}H
975	172.64	0.026	C	94Gly	^{13}C
976	3.54	0.008	Ha	95Val	^{1}H
977	0.36	0.012	Hgb*	95Val	^{1}H
978	-0.66	0.004	Hga*	95Val	^{1}H
979	21.10	0.106	Cgb	95Val	^{13}C
980	22.97	0.114	Cga	95Val	^{13}C
981	0.34	0.019	Hb	95Val	^{1}H
982	62.88	0.074	Ca	95Val	^{13}C
983	8.29	0.003	H	95Val	^{1}H
984	32.15	0.103	Cb	95Val	^{13}C
985	120.78	0.063	N	95Val	^{15}N
986	174.45	0.013	C	95Val	^{13}C
987	116.20	0.070	N	96Asn	^{15}N

988	6.40	0.005	H	96Asn	^{1}H
989	113.20	0.049	Nd2	96Asn	^{15}N
990	6.51	0.004	Hd2a	96Asn	^{1}H
991	51.78	0.080	Ca	96Asn	^{13}C
992	39.85	0.051	Cb	96Asn	^{13}C
993	3.29	0.009	Hbb	96Asn	^{1}H
994	4.07	0.006	Ha	96Asn	^{1}H
995	3.03	0.014	Hba	96Asn	^{1}H
996	175.24	0.014	C	96Asn	^{13}C
997	7.65	0.005	Hd2b	96Asn	^{1}H
998	10.28	0.003	He1	97Trp	^{1}H
999	128.74	0.046	Ne1	97Trp	^{15}N
1000	8.53	0.009	H	97Trp	^{1}H
1001	118.78	0.106	N	97Trp	^{15}N
1002	124.91	0.040	Ch2	97Trp	^{13}C
1003	6.24	0.011	Hh2	97Trp	^{1}H
1004	7.17	0.007	Hz2	97Trp	^{1}H
1005	61.08	0.073	Ca	97Trp	^{13}C
1006	30.68	0.220	Cb	97Trp	^{13}C
1007	4.59	0.012	Ha	97Trp	^{1}H
1008	3.66	0.014	Hbb	97Trp	^{1}H
1009	3.53	0.014	Hba	97Trp	^{1}H
1010	178.21	0.037	C	97Trp	^{13}C
1011	7.69	0.005	Hd1	97Trp	^{1}H
1012	128.47	0.061	Cd1	97Trp	^{13}C
1013	115.06	0.046	Cz2	97Trp	^{13}C
1014	6.81	0.005	He3	97Trp	^{1}H
1015	117.29	0.042	Ce3	97Trp	^{13}C
1016	4.01	0.034	Ha2	98Gly	^{1}H
1017	4.01	0.032	Ha3	98Gly	^{1}H
1018	106.57	0.061	N	98Gly	^{15}N
1019	8.63	0.009	H	98Gly	^{1}H
1020	47.51	0.095	Ca	98Gly	^{13}C
1021	177.48	0.060	C	98Gly	^{13}C
1022	31.36	0.086	Cb	99Arg	^{13}C
1023	8.21	0.008	H	99Arg	^{1}H
1024	121.63	0.124	N	99Arg	^{15}N
1025	58.49	0.038	Ca	99Arg	^{13}C
1026	1.46	0.006	Hba	99Arg	^{1}H
1027	2.96	0.010	Hd2	99Arg	^{1}H
1028	43.83	0.087	Cd	99Arg	^{13}C
1029	1.54	0.016	Hbb	99Arg	^{1}H
1030	2.96	0.010	Hd3	99Arg	^{1}H
1031	1.79	0.011	Hgb	99Arg	^{1}H
1032	1.52	0.024	Hga	99Arg	^{1}H

1033	28.67	0.044	Cg	99Arg	^{13}C
1034	177.69	0.015	C	99Arg	^{13}C
1035	4.00	0.008	Ha	99Arg	^{1}H
1036	7.81	0.006	H	100Ile	^{1}H
1037	121.38	0.069	N	100Ile	^{15}N
1038	67.15	0.105	Ca	100Ile	^{13}C
1039	37.10	0.096	Cb	100Ile	^{13}C
1040	2.46	0.009	Hb	100Ile	^{1}H
1041	3.99	0.012	Ha	100Ile	^{1}H
1042	1.30	0.005	Hg2*	100Ile	^{1}H
1043	0.31	0.004	Hd1*	100Ile	^{1}H
1044	30.78	0.065	Cg1	100Ile	^{13}C
1045	1.74	0.006	Hg1b	100Ile	^{1}H
1046	0.79	0.006	Hg1a	100Ile	^{1}H
1047	14.25	0.095	Cd1	100Ile	^{13}C
1048	17.79	0.072	Cg2	100Ile	^{13}C
1049	178.14	0.062	C	100Ile	^{13}C
1050	121.44	0.074	N	101Val	^{15}N
1051	8.38	0.009	H	101Val	^{1}H
1052	68.16	0.041	Ca	101Val	^{13}C
1053	0.83	0.015	Hga*	101Val	^{1}H
1054	1.12	0.007	Hgb*	101Val	^{1}H
1055	31.42	0.041	Cb	101Val	^{13}C
1056	3.46	0.007	Ha	101Val	^{1}H
1057	21.41	0.054	Cga	101Val	^{13}C
1058	2.32	0.007	Hb	101Val	^{1}H
1059	23.77	0.038	Cgb	101Val	^{13}C
1060	177.75	0.037	C	101Val	^{13}C
1061	7.69	0.006	H	102Ala	^{1}H
1062	121.86	0.100	N	102Ala	^{15}N
1063	55.26	0.078	Ca	102Ala	^{13}C
1064	18.63	0.104	Cb	102Ala	^{13}C
1065	3.17	0.006	Ha	102Ala	^{1}H
1066	1.22	0.006	Hb*	102Ala	^{1}H
1067	178.22	0.007	C	102Ala	^{13}C
1068	8.10	0.006	H	103Phe	^{1}H
1069	119.56	0.068	N	103Phe	^{15}N
1070	60.81	0.220	Ca	103Phe	^{13}C
1071	38.59	0.092	Cb	103Phe	^{13}C
1072	3.20	0.009	Hba	103Phe	^{1}H
1073	176.26	0.071	C	103Phe	^{13}C
1074	3.99	0.013	Ha	103Phe	^{1}H
1075	3.84	0.011	Hbb	103Phe	^{1}H
1076	7.02	0.007	Hd*	103Phe	^{1}H
1077	132.26	0.085	Cd*	103Phe	^{13}C

1078	8.09	0.006	H	104Phe	^{1}H
1079	120.69	0.072	N	104Phe	^{15}N
1080	62.99	0.055	Ca	104Phe	^{13}C
1081	36.63	0.082	Cb	104Phe	^{13}C
1082	1.23	0.011	Hba	104Phe	^{1}H
1083	3.35	0.005	Ha	104Phe	^{1}H
1084	2.58	0.010	Hbb	104Phe	^{1}H
1085	177.60	0.000	C	104Phe	^{13}C
1086	7.16	0.016	Hd*	104Phe	^{1}H
1087	131.61	0.094	Cd*	104Phe	^{13}C
1088	114.78	0.072	N	105Ser	^{15}N
1089	8.51	0.003	H	105Ser	^{1}H
1090	63.77	0.040	Ca	105Ser	^{13}C
1091	4.44	0.018	Ha	105Ser	^{1}H
1092	63.41	0.002	Cb	105Ser	^{13}C
1093	177.60	0.014	C	105Ser	^{13}C
1094	4.04	0.018	Hb3	105Ser	^{1}H
1095	4.04	0.018	Hb2	105Ser	^{1}H
1096	38.64	0.090	Cb	106Phe	^{13}C
1097	3.03	0.008	Hbb	106Phe	^{1}H
1098	2.80	0.006	Hba	106Phe	^{1}H
1099	9.15	0.004	H	106Phe	^{1}H
1100	124.63	0.052	N	106Phe	^{15}N
1101	60.19	0.120	Ca	106Phe	^{13}C
1102	176.64	0.034	C	106Phe	^{13}C
1103	4.07	0.005	Ha	106Phe	^{1}H
1104	130.61	0.032	Ce*	106Phe	^{13}C
1105	6.79	0.012	Hd*	106Phe	^{1}H
1106	130.62	0.032	Cd*	106Phe	^{13}C
1107	7.00	0.023	He*	106Phe	^{1}H
1108	106.50	0.032	N	107Gly	^{15}N
1109	8.06	0.004	H	107Gly	^{1}H
1110	48.00	0.067	Ca	107Gly	^{13}C
1111	3.33	0.011	Haa	107Gly	^{1}H
1112	4.27	0.006	Hab	107Gly	^{1}H
1113	176.65	0.000	C	107Gly	^{13}C
1114	107.79	0.065	N	108Gly	^{15}N
1115	8.90	0.003	H	108Gly	^{1}H
1116	3.95	0.006	Hab	108Gly	^{1}H
1117	47.93	0.071	Ca	108Gly	^{13}C
1118	3.73	0.011	Haa	108Gly	^{1}H
1119	174.59	0.000	C	108Gly	^{13}C
1120	8.34	0.010	H	109Ala	^{1}H
1121	124.96	0.062	N	109Ala	^{15}N
1122	55.10	0.040	Ca	109Ala	^{13}C

1123	18.42	0.073	Cb	109Ala	^{13}C
1124	4.22	0.005	Ha	109Ala	^{1}H
1125	1.65	0.003	Hb*	109Ala	^{1}H
1126	181.11	0.020	C	109Ala	^{13}C
1127	8.30	0.012	H	110Leu	^{1}H
1128	119.34	0.094	N	110Leu	^{15}N
1129	42.56	0.058	Cb	110Leu	^{13}C
1130	57.40	0.083	Ca	110Leu	^{13}C
1131	25.62	0.056	Cda	110Leu	^{13}C
1132	0.61	0.006	Hda*	110Leu	^{1}H
1133	23.77	0.067	Cdb	110Leu	^{13}C
1134	0.81	0.012	Hdb*	110Leu	^{1}H
1135	4.02	0.010	Ha	110Leu	^{1}H
1136	1.34	0.018	Hba	110Leu	^{1}H
1137	1.73	0.026	Hbb	110Leu	^{1}H
1138	26.04	0.152	Cg	110Leu	^{13}C
1139	1.44	0.010	Hg	110Leu	^{1}H
1140	180.10	0.014	C	110Leu	^{13}C
1141	26.72	0.104	Cb	111Cys	^{13}C
1142	8.21	0.003	H	111Cys	^{1}H
1143	64.66	0.087	Ca	111Cys	^{13}C
1144	119.47	0.040	N	111Cys	^{15}N
1145	3.21	0.004	Hbb	111Cys	^{1}H
1146	2.70	0.009	Hba	111Cys	^{1}H
1147	3.82	0.009	Ha	111Cys	^{1}H
1148	175.95	0.028	C	111Cys	^{13}C
1149	8.11	0.007	H	112Val	^{1}H
1150	119.09	0.055	N	112Val	^{15}N
1151	67.32	0.067	Ca	112Val	^{13}C
1152	31.94	0.067	Cb	112Val	^{13}C
1153	1.04	0.005	Hga*	112Val	^{1}H
1154	23.09	0.050	Cga	112Val	^{13}C
1155	3.38	0.005	Ha	112Val	^{1}H
1156	22.31	0.063	Cgb	112Val	^{13}C
1157	2.17	0.011	Hb	112Val	^{1}H
1158	1.05	0.004	Hgb*	112Val	^{1}H
1159	177.36	0.040	C	112Val	^{13}C
1160	119.45	0.058	N	113Glu	^{15}N
1161	8.19	0.004	H	113Glu	^{1}H
1162	59.34	0.053	Ca	113Glu	^{13}C
1163	2.04	0.002	Hb2	113Glu	^{1}H
1164	2.04	0.002	Hb3	113Glu	^{1}H
1165	29.33	0.095	Cb	113Glu	^{13}C
1166	36.30	0.096	Cg	113Glu	^{13}C
1167	4.09	0.013	Ha	113Glu	^{1}H

1168	178.26	0.006	C	113Glu	^{13}C
1169	2.20	0.030	Hga	113Glu	^{1}H
1170	2.36	0.014	Hgb	113Glu	^{1}H
1171	7.76	0.007	H	114Ser	^{1}H
1172	113.85	0.029	N	114Ser	^{15}N
1173	63.25	0.076	Ca	114Ser	^{13}C
1174	3.89	0.026	Ha	114Ser	^{1}H
1175	62.23	0.101	Cb	114Ser	^{13}C
1176	3.68	0.012	Hb3	114Ser	^{1}H
1177	3.68	0.012	Hb2	114Ser	^{1}H
1178	175.49	0.022	C	114Ser	^{13}C
1179	7.42	0.006	H	115Val	^{1}H
1180	121.09	0.048	N	115Val	^{15}N
1181	66.75	0.072	Ca	115Val	^{13}C
1182	31.41	0.107	Cb	115Val	^{13}C
1183	0.35	0.005	Hgb*	115Val	^{1}H
1184	21.36	0.083	Cga	115Val	^{13}C
1185	22.84	0.096	Cgb	115Val	^{13}C
1186	3.20	0.006	Ha	115Val	^{1}H
1187	1.66	0.008	Hb	115Val	^{1}H
1188	0.13	0.003	Hga*	115Val	^{1}H
1189	179.47	0.032	C	115Val	^{13}C
1190	8.37	0.003	H	116Asp	^{1}H
1191	57.47	0.036	Ca	116Asp	^{13}C
1192	123.02	0.065	N	116Asp	^{15}N
1193	41.00	0.074	Cb	116Asp	^{13}C
1194	4.40	0.008	Ha	116Asp	^{1}H
1195	2.92	0.004	Hbb	116Asp	^{1}H
1196	2.70	0.008	Hba	116Asp	^{1}H
1197	178.03	0.019	C	116Asp	^{13}C
1198	25.77	0.058	Cg	117Lys	^{13}C
1199	29.03	0.127	Cd	117Lys	^{13}C
1200	2.99	0.009	He3	117Lys	^{1}H
1201	2.99	0.009	He2	117Lys	^{1}H
1202	4.38	0.010	Ha	117Lys	^{1}H
1203	1.65	0.008	Hgb	117Lys	^{1}H
1204	3.00	0.017	Heb	117Lys	^{1}H
1205	1.68	0.022	Hb3	117Lys	^{1}H
1206	3.00	0.017	Hea	117Lys	^{1}H
1207	1.46	0.007	Hga	117Lys	^{1}H
1208	1.68	0.021	Hb2	117Lys	^{1}H
1209	7.70	0.005	H	117Lys	^{1}H
1210	116.21	0.057	N	117Lys	^{15}N
1211	55.68	0.072	Ca	117Lys	^{13}C
1212	32.32	0.151	Cb	117Lys	^{13}C

1213	42.43	0.070	Ce	117Lys	^{13}C
1214	175.05	0.011	C	117Lys	^{13}C
1215	2.12	0.004	Hd2	117Lys	^{1}H
1216	2.12	0.004	Hd3	117Lys	^{1}H
1217	114.27	0.071	N	118Glu	^{15}N
1218	7.96	0.005	H	118Glu	^{1}H
1219	57.77	0.095	Ca	118Glu	^{13}C
1220	26.54	0.008	Cb	118Glu	^{13}C
1221	36.92	0.012	Cg	118Glu	^{13}C
1222	3.97	0.009	Ha	118Glu	^{1}H
1223	176.06	0.016	C	118Glu	^{13}C
1224	2.24	0.003	Hga	118Glu	^{1}H
1225	2.41	0.002	Hba	118Glu	^{1}H
1226	2.32	0.008	Hgb	118Glu	^{1}H
1227	8.54	0.003	H	119Met	^{1}H
1228	119.26	0.066	N	119Met	^{15}N
1229	53.68	0.077	Ca	119Met	^{13}C
1230	31.87	0.344	Cb	119Met	^{13}C
1231	31.79	0.184	Cg	119Met	^{13}C
1232	4.82	0.004	Ha	119Met	^{1}H
1233	2.55	0.011	Hgb	119Met	^{1}H
1234	1.90	0.012	Hba	119Met	^{1}H
1235	1.97	0.011	Hbb	119Met	^{1}H
1236	2.42	0.021	Hga	119Met	^{1}H
1237	176.82	0.013	C	119Met	^{13}C
1238	2.12	0.005	He*	119Met	^{1}H
1239	17.69	0.032	Ce	119Met	^{13}C
1240	111.02	0.229	Ne2	120Gln	^{15}N
1241	7.00	0.012	He2a	120Gln	^{1}H
1242	7.53	0.008	He2b	120Gln	^{1}H
1243	8.80	0.002	H	120Gln	^{1}H
1244	119.77	0.072	N	120Gln	^{15}N
1245	60.17	0.088	Ca	120Gln	^{13}C
1246	27.67	0.091	Cb	120Gln	^{13}C
1247	34.18	0.064	Cg	120Gln	^{13}C
1248	2.50	0.006	Hgb	120Gln	^{1}H
1249	177.47	0.022	C	120Gln	^{13}C
1250	1.98	0.008	Hb3	120Gln	^{1}H
1251	1.98	0.009	Hb2	120Gln	^{1}H
1252	2.31	0.004	Hga	120Gln	^{1}H
1253	3.49	0.005	Ha	120Gln	^{1}H
1254	179.85	0.050	Cd	120Gln	^{13}C
1255	7.96	0.006	H	121Val	^{1}H
1256	117.29	0.062	N	121Val	^{15}N
1257	63.61	0.037	Ca	121Val	^{13}C

1258	31.49	0.126	Cb	121Val	^{13}C
1259	2.21	0.004	Hb	121Val	^{1}H
1260	4.06	0.008	Ha	121Val	^{1}H
1261	0.93	0.023	Hgb*	121Val	^{1}H
1262	0.89	0.008	Hga*	121Val	^{1}H
1263	20.07	0.063	Cga	121Val	^{13}C
1264	20.05	0.049	Cgb	121Val	^{13}C
1265	175.68	0.017	C	121Val	^{13}C
1266	117.44	0.048	N	122Leu	^{15}N
1267	55.53	0.097	Ca	122Leu	^{13}C
1268	7.89	0.005	H	122Leu	^{1}H
1269	41.37	0.069	Cb	122Leu	^{13}C
1270	1.25	0.011	Hba	122Leu	^{1}H
1271	3.87	0.012	Ha	122Leu	^{1}H
1272	0.68	0.005	Hda*	122Leu	^{1}H
1273	1.88	0.009	Hbb	122Leu	^{1}H
1274	0.78	0.005	Hdb*	122Leu	^{1}H
1275	1.57	0.006	Hg	122Leu	^{1}H
1276	26.40	0.079	Cdb	122Leu	^{13}C
1277	22.81	0.078	Cda	122Leu	^{13}C
1278	27.07	0.050	Cg	122Leu	^{13}C
1279	177.96	0.075	C	122Leu	^{13}C
1280	67.82	0.089	Ca	123Val	^{13}C
1281	30.89	0.169	Cb	123Val	^{13}C
1282	7.44	0.006	H	123Val	^{1}H
1283	120.03	0.071	N	123Val	^{15}N
1284	21.78	0.095	Cga	123Val	^{13}C
1285	23.29	0.072	Cgb	123Val	^{13}C
1286	3.38	0.011	Ha	123Val	^{1}H
1287	0.41	0.011	Hga*	123Val	^{1}H
1288	1.77	0.011	Hb	123Val	^{1}H
1289	0.43	0.012	Hgb*	123Val	^{1}H
1290	177.04	0.010	C	123Val	^{13}C
1291	8.39	0.010	H	124Ser	^{1}H
1292	111.17	0.065	N	124Ser	^{15}N
1293	61.25	0.065	Ca	124Ser	^{13}C
1294	62.40	0.129	Cb	124Ser	^{13}C
1295	177.07	0.073	C	124Ser	^{13}C
1296	3.84	0.012	Hb2	124Ser	^{1}H
1297	3.98	0.005	Ha	124Ser	^{1}H
1298	3.84	0.011	Hb3	124Ser	^{1}H
1299	29.85	0.140	Cb	125Arg	^{13}C
1300	123.53	0.049	N	125Arg	^{15}N
1301	6.70	0.007	H	125Arg	^{1}H
1302	58.74	0.035	Ca	125Arg	^{13}C

1303	175.75	0.011	C	125Arg	^{13}C
1304	44.08	0.105	Cd	125Arg	^{13}C
1305	2.78	0.014	Hda	125Arg	^{1}H
1306	1.09	0.020	Hba	125Arg	^{1}H
1307	3.20	0.005	Hdb	125Arg	^{1}H
1308	1.49	0.014	Hbb	125Arg	^{1}H
1309	3.71	0.006	Ha	125Arg	^{1}H
1310	25.77	0.109	Cg	125Arg	^{13}C
1311	0.78	0.024	Hga	125Arg	^{1}H
1312	1.26	0.012	Hgb	125Arg	^{1}H
1313	118.80	0.088	N	126Ile	^{15}N
1314	8.12	0.005	H	126Ile	^{1}H
1315	37.67	0.050	Cb	126Ile	^{13}C
1316	64.96	0.036	Ca	126Ile	^{13}C
1317	177.97	0.038	C	126Ile	^{13}C
1318	0.69	0.005	Hg2*	126Ile	^{1}H
1319	2.86	0.005	Ha	126Ile	^{1}H
1320	0.54	0.005	Hg1a	126Ile	^{1}H
1321	29.90	0.081	Cg1	126Ile	^{13}C
1322	1.22	0.008	Hg1b	126Ile	^{1}H
1323	1.62	0.014	Hb	126Ile	^{1}H
1324	18.33	0.046	Cg2	126Ile	^{13}C
1325	0.53	0.005	Hd1*	126Ile	^{1}H
1326	14.73	0.057	Cd1	126Ile	^{13}C
1327	7.71	0.007	H	127Ala	^{1}H
1328	55.29	0.150	Ca	127Ala	^{13}C
1329	119.67	0.082	N	127Ala	^{15}N
1330	18.52	0.144	Cb	127Ala	^{13}C
1331	3.90	0.010	Ha	127Ala	^{1}H
1332	179.40	0.033	C	127Ala	^{13}C
1333	1.50	0.006	Hb*	127Ala	^{1}H
1334	7.36	0.003	H	128Ala	^{1}H
1335	120.85	0.066	N	128Ala	^{15}N
1336	17.82	0.141	Cb	128Ala	^{13}C
1337	55.26	0.102	Ca	128Ala	^{13}C
1338	1.57	0.010	Hb*	128Ala	^{1}H
1339	4.17	0.004	Ha	128Ala	^{1}H
1340	181.04	0.026	C	128Ala	^{13}C
1341	3.00	0.014	Hbb	129Trp	^{1}H
1342	2.68	0.010	Hba	129Trp	^{1}H
1343	126.61	0.065	Ne1	129Trp	^{15}N
1344	10.14	0.003	He1	129Trp	^{1}H
1345	9.09	0.006	H	129Trp	^{1}H
1346	121.65	0.070	N	129Trp	^{15}N
1347	57.79	0.029	Ca	129Trp	^{13}C

1348	28.92	0.084	Cb	129Trp	^{13}C
1349	179.56	0.033	C	129Trp	^{13}C
1350	124.34	0.069	Cd1	129Trp	^{13}C
1351	124.00	0.030	Ch2	129Trp	^{13}C
1352	7.17	0.022	Hd1	129Trp	^{1}H
1353	7.42	0.011	Hz2	129Trp	^{1}H
1354	6.85	0.010	Hh2	129Trp	^{1}H
1355	113.65	0.031	Cz2	129Trp	^{13}C
1356	6.60	0.021	Hz3	129Trp	^{1}H
1357	120.01	0.036	Cz3	129Trp	^{13}C
1358	34.93	0.085	Cb	130Met	^{13}C
1359	9.03	0.009	H	130Met	^{1}H
1360	117.54	0.085	N	130Met	^{15}N
1361	60.18	0.055	Ca	130Met	^{13}C
1362	178.17	0.025	C	130Met	^{13}C
1363	33.01	0.057	Cg	130Met	^{13}C
1364	4.35	0.009	Ha	130Met	^{1}H
1365	2.05	0.004	Hba	130Met	^{1}H
1366	2.38	0.014	Hbb	130Met	^{1}H
1367	2.16	0.010	Hga	130Met	^{1}H
1368	2.88	0.010	Hgb	130Met	^{1}H
1369	1.46	0.007	He*	130Met	^{1}H
1370	16.60	0.033	Ce	130Met	^{13}C
1371	8.27	0.010	H	131Ala	^{1}H
1372	120.97	0.085	N	131Ala	^{15}N
1373	17.63	0.120	Cb	131Ala	^{13}C
1374	4.10	0.009	Ha	131Ala	^{1}H
1375	1.60	0.006	Hb*	131Ala	^{1}H
1376	56.18	0.050	Ca	131Ala	^{13}C
1377	176.25	0.045	C	131Ala	^{13}C
1378	67.31	0.118	Ca	132Thr	^{13}C
1379	69.18	0.323	Cb	132Thr	^{13}C
1380	117.21	0.056	N	132Thr	^{15}N
1381	175.03	0.024	C	132Thr	^{13}C
1382	8.65	0.007	H	132Thr	^{1}H
1383	21.19	0.173	Cg2	132Thr	^{13}C
1384	4.29	0.007	Hb	132Thr	^{1}H
1385	3.96	0.014	Ha	132Thr	^{1}H
1386	1.02	0.015	Hg2*	132Thr	^{1}H
1387	8.96	0.008	H	133Tyr	^{1}H
1388	123.70	0.107	N	133Tyr	^{15}N
1389	63.58	0.084	Ca	133Tyr	^{13}C
1390	179.55	0.032	C	133Tyr	^{13}C
1391	39.31	0.075	Cb	133Tyr	^{13}C
1392	4.02	0.018	Ha	133Tyr	^{1}H

1393	3.49	0.017	Hba	133Tyr	^{1}H
1394	3.75	0.009	Hbb	133Tyr	^{1}H
1395	7.08	0.024	Hd*	133Tyr	^{1}H
1396	131.36	0.055	Cd*	133Tyr	^{13}C
1397	6.79	0.028	He*	133Tyr	^{1}H
1398	116.59	0.000	Ce*	133Tyr	^{13}C
1399	8.85	0.006	H	134Leu	^{1}H
1400	119.67	0.094	N	134Leu	^{15}N
1401	59.59	0.109	Ca	134Leu	^{13}C
1402	42.90	0.080	Cb	134Leu	^{13}C
1403	1.15	0.008	Hda*	134Leu	^{1}H
1404	1.87	0.006	Hba	134Leu	^{1}H
1405	2.14	0.010	Hbb	134Leu	^{1}H
1406	178.32	0.030	C	134Leu	^{13}C
1407	26.16	0.078	Cda	134Leu	^{13}C
1408	1.16	0.018	Hdb*	134Leu	^{1}H
1409	26.14	0.080	Cdb	134Leu	^{13}C
1410	3.83	0.007	Ha	134Leu	^{1}H
1411	1.87	0.006	Hg	134Leu	^{1}H
1412	28.41	0.085	Cg	134Leu	^{13}C
1413	8.45	0.007	H	135Asn	^{1}H
1414	117.66	0.086	N	135Asn	^{15}N
1415	112.73	0.228	Nd2	135Asn	^{15}N
1416	7.69	0.005	Hd2b	135Asn	^{1}H
1417	56.88	0.104	Ca	135Asn	^{13}C
1418	39.22	0.084	Cb	135Asn	^{13}C
1419	4.35	0.010	Ha	135Asn	^{1}H
1420	3.01	0.007	Hbb	135Asn	^{1}H
1421	2.91	0.005	Hba	135Asn	^{1}H
1422	176.41	0.024	C	135Asn	^{13}C
1423	176.33	0.005	Cg	135Asn	^{13}C
1424	7.01	0.010	Hd2a	135Asn	^{1}H
1425	115.94	0.041	N	136Asp	^{15}N
1426	8.58	0.003	H	136Asp	^{1}H
1427	40.86	0.103	Cb	136Asp	^{13}C
1428	177.80	0.006	C	136Asp	^{13}C
1429	56.50	0.146	Ca	136Asp	^{13}C
1430	4.34	0.011	Ha	136Asp	^{1}H
1431	2.02	0.008	Hba	136Asp	^{1}H
1432	2.42	0.011	Hbb	136Asp	^{1}H
1433	8.20	0.012	H	137His	^{1}H
1434	112.07	0.058	N	137His	^{15}N
1435	56.53	0.019	Ca	137His	^{13}C
1436	30.89	0.086	Cb	137His	^{13}C
1437	175.78	0.039	C	137His	^{13}C

1438	2.27	0.012	Hba	137His	^{1}H
1439	2.46	0.008	Hbb	137His	^{1}H
1440	4.55	0.009	Ha	137His	^{1}H
1441	43.47	0.121	Cb	138Leu	^{13}C
1442	117.36	0.080	N	138Leu	^{15}N
1443	7.31	0.019	H	138Leu	^{1}H
1444	55.70	0.060	Ca	138Leu	^{13}C
1445	176.37	0.012	C	138Leu	^{13}C
1446	2.13	0.010	Hbb	138Leu	^{1}H
1447	4.55	0.009	Ha	138Leu	^{1}H
1448	-0.23	0.004	Hda*	138Leu	^{1}H
1449	24.86	0.101	Cdb	138Leu	^{13}C
1450	22.02	0.081	Cda	138Leu	^{13}C
1451	0.29	0.005	Hdb*	138Leu	^{1}H
1452	1.15	0.010	Hba	138Leu	^{1}H
1453	26.41	0.017	Cg	138Leu	^{13}C
1454	1.19	0.039	Hg	138Leu	^{1}H
1455	121.92	0.067	N	139Glu	^{15}N
1456	8.71	0.026	H	139Glu	^{1}H
1457	60.53	0.178	Ca	139Glu	^{13}C
1458	27.77	0.065	Cb	139Glu	^{13}C
1459	173.82	0.000	C	139Glu	^{13}C
1460	2.25	0.007	Hba	139Glu	^{1}H
1461	2.39	0.011	Hg2	139Glu	^{1}H
1462	4.35	0.007	Ha	139Glu	^{1}H
1463	35.86	0.034	Cg	139Glu	^{13}C
1464	2.25	0.007	Hbb	139Glu	^{1}H
1465	2.39	0.011	Hg3	139Glu	^{1}H
1466	179.16	0.000	C	140Pro	^{13}C
1467	66.58	0.159	Ca	140Pro	^{13}C
1468	49.92	0.083	Cd	140Pro	^{13}C
1469	30.75	0.074	Cb	140Pro	^{13}C
1470	1.98	0.006	Hgb	140Pro	^{1}H
1471	3.61	0.010	Hdb	140Pro	^{1}H
1472	4.23	0.006	Ha	140Pro	^{1}H
1473	1.87	0.010	Hga	140Pro	^{1}H
1474	2.28	0.010	Hbb	140Pro	^{1}H
1475	3.18	0.006	Hda	140Pro	^{1}H
1476	1.78	0.006	Hba	140Pro	^{1}H
1477	28.70	0.080	Cg	140Pro	^{13}C
1478	9.52	0.003	He1	141Trp	^{1}H
1479	129.42	0.036	Ne1	141Trp	^{15}N
1480	118.37	0.042	N	141Trp	^{15}N
1481	7.20	0.004	H	141Trp	^{1}H
1482	62.20	0.077	Ca	141Trp	^{13}C

1483	5.41	0.004	Hh2	141Trp	^{1}H
1484	120.66	0.000	Ch2	141Trp	^{13}C
1485	3.15	0.010	Hba	141Trp	^{1}H
1486	179.41	0.004	C	141Trp	^{13}C
1487	3.86	0.006	Ha	141Trp	^{1}H
1488	28.93	0.113	Cb	141Trp	^{13}C
1489	3.17	0.009	Hbb	141Trp	^{1}H
1490	7.08	0.010	Hd1	141Trp	^{1}H
1491	126.30	0.046	Cd1	141Trp	^{13}C
1492	6.97	0.015	Hz2	141Trp	^{1}H
1493	114.04	0.028	Cz2	141Trp	^{13}C
1494	122.55	0.055	N	142Ile	^{15}N
1495	8.51	0.006	H	142Ile	^{1}H
1496	178.92	0.035	C	142Ile	^{13}C
1497	66.24	0.046	Ca	142Ile	^{13}C
1498	0.48	0.013	Hg1a	142Ile	^{1}H
1499	1.61	0.008	Hg1b	142Ile	^{1}H
1500	0.46	0.005	Hg2*	142Ile	^{1}H
1501	29.00	0.087	Cg1	142Ile	^{13}C
1502	0.63	0.004	Hd1*	142Ile	^{1}H
1503	1.91	0.013	Hb	142Ile	^{1}H
1504	16.26	0.091	Cg2	142Ile	^{13}C
1505	38.31	0.073	Cb	142Ile	^{13}C
1506	13.94	0.057	Cd1	142Ile	^{13}C
1507	2.89	0.007	Ha	142Ile	^{1}H
1508	33.58	0.101	Cg	143Gln	^{13}C
1509	3.92	0.008	Ha	143Gln	^{1}H
1510	116.78	0.073	N	143Gln	^{15}N
1511	8.51	0.005	H	143Gln	^{1}H
1512	8.01	0.010	He2b	143Gln	^{1}H
1513	112.76	0.221	Ne2	143Gln	^{15}N
1514	6.85	0.006	He2a	143Gln	^{1}H
1515	58.57	0.097	Ca	143Gln	^{13}C
1516	28.07	0.117	Cb	143Gln	^{13}C
1517	2.53	0.007	Hg3	143Gln	^{1}H
1518	2.53	0.007	Hg2	143Gln	^{1}H
1519	179.45	0.003	C	143Gln	^{13}C
1520	2.13	0.008	Hbb	143Gln	^{1}H
1521	2.04	0.010	Hba	143Gln	^{1}H
1522	180.13	0.035	Cd	143Gln	^{13}C
1523	7.65	0.005	H	144Glu	^{1}H
1524	119.47	0.030	N	144Glu	^{15}N
1525	57.90	0.076	Ca	144Glu	^{13}C
1526	29.92	0.044	Cb	144Glu	^{13}C
1527	36.37	0.042	Cg	144Glu	^{13}C

1528	4.08	0.003	Ha	144Glu	^{1}H
1529	2.24	0.009	Hgb	144Glu	^{1}H
1530	2.14	0.008	Hga	144Glu	^{1}H
1531	176.78	0.005	C	144Glu	^{13}C
1532	1.87	0.010	Hbb	144Glu	^{1}H
1533	1.83	0.015	Hba	144Glu	^{1}H
1534	116.90	0.052	N	145Asn	^{15}N
1535	7.20	0.005	H	145Asn	^{1}H
1536	6.14	0.031	Hd2b	145Asn	^{1}H
1537	5.90	0.005	Hd2a	145Asn	^{1}H
1538	115.38	0.053	Nd2	145Asn	^{15}N
1539	53.65	0.077	Ca	145Asn	^{13}C
1540	38.20	0.088	Cb	145Asn	^{13}C
1541	0.98	0.026	Hba	145Asn	^{1}H
1542	4.45	0.007	Ha	145Asn	^{1}H
1543	2.22	0.017	Hbb	145Asn	^{1}H
1544	173.89	0.008	C	145Asn	^{13}C
1545	105.17	0.085	N	146Gly	^{15}N
1546	7.44	0.010	H	146Gly	^{1}H
1547	45.72	0.038	Ca	146Gly	^{13}C
1548	4.51	0.013	Hab	146Gly	^{1}H
1549	3.72	0.005	Haa	146Gly	^{1}H
1550	175.98	0.000	C	146Gly	^{13}C
1551	108.77	0.051	N	147Gly	^{15}N
1552	8.58	0.003	H	147Gly	^{1}H
1553	44.53	0.100	Ca	147Gly	^{13}C
1554	4.52	0.008	Hab	147Gly	^{1}H
1555	3.67	0.011	Haa	147Gly	^{1}H
1556	173.37	0.000	C	147Gly	^{13}C
1557	10.58	0.004	He1	148Trp	^{1}H
1558	129.16	0.056	Ne1	148Trp	^{15}N
1559	8.71	0.004	H	148Trp	^{1}H
1560	60.56	0.082	Ca	148Trp	^{13}C
1561	118.43	0.025	N	148Trp	^{15}N
1562	3.53	0.008	Hbb	148Trp	^{1}H
1563	4.43	0.010	Ha	148Trp	^{1}H
1564	3.13	0.008	Hba	148Trp	^{1}H
1565	29.87	0.108	Cb	148Trp	^{13}C
1566	178.43	0.002	C	148Trp	^{13}C
1567	126.81	0.084	Cd1	148Trp	^{13}C
1568	7.47	0.015	Hd1	148Trp	^{1}H
1569	7.27	0.006	Hz2	148Trp	^{1}H
1570	6.80	0.010	Hh2	148Trp	^{1}H
1571	113.82	0.036	Cz2	148Trp	^{13}C
1572	123.08	0.009	Ch2	148Trp	^{13}C

1573	124.24	0.000	Ce3	148Trp	^{13}C
1574	7.23	0.045	He3	148Trp	^{1}H
1575	131.76	0.000	Cz3	148Trp	^{13}C
1576	7.16	0.015	Hz3	148Trp	^{1}H
1577	8.90	0.004	H	149Asp	^{1}H
1578	118.21	0.056	N	149Asp	^{15}N
1579	57.85	0.063	Ca	149Asp	^{13}C
1580	40.38	0.097	Cb	149Asp	^{13}C
1581	4.45	0.011	Ha	149Asp	^{1}H
1582	2.71	0.014	Hba	149Asp	^{1}H
1583	2.71	0.014	Hbb	149Asp	^{1}H
1584	179.63	0.011	C	149Asp	^{13}C
1585	117.18	0.056	N	150Thr	^{15}N
1586	7.63	0.007	H	150Thr	^{1}H
1587	66.38	0.162	Ca	150Thr	^{13}C
1588	68.34	0.162	Cb	150Thr	^{13}C
1589	1.45	0.005	Hg2*	150Thr	^{1}H
1590	22.63	0.053	Cg2	150Thr	^{13}C
1591	4.01	0.004	Ha	150Thr	^{1}H
1592	4.45	0.013	Hb	150Thr	^{1}H
1593	175.22	0.013	C	150Thr	^{13}C
1594	6.64	0.004	H	151Phe	^{1}H
1595	122.60	0.070	N	151Phe	^{15}N
1596	61.85	0.046	Ca	151Phe	^{13}C
1597	37.95	0.070	Cb	151Phe	^{13}C
1598	6.58	0.003	Hz	151Phe	^{1}H
1599	128.29	0.082	Cz	151Phe	^{13}C
1600	176.70	0.013	C	151Phe	^{13}C
1601	4.42	0.012	Ha	151Phe	^{1}H
1602	3.20	0.012	Hbb	151Phe	^{1}H
1603	3.11	0.009	Hba	151Phe	^{1}H
1604	6.81	0.010	Hd*	151Phe	^{1}H
1605	131.22	0.050	Cd*	151Phe	^{13}C
1606	130.30	0.201	Ce*	151Phe	^{13}C
1607	6.68	0.026	He*	151Phe	^{1}H
1608	117.68	0.148	N	152Val	^{15}N
1609	66.78	0.036	Ca	152Val	^{13}C
1610	7.84	0.032	H	152Val	^{1}H
1611	3.06	0.006	Ha	152Val	^{1}H
1612	2.19	0.013	Hb	152Val	^{1}H
1613	0.89	0.009	Hga*	152Val	^{1}H
1614	1.14	0.007	Hgb*	152Val	^{1}H
1615	21.72	0.180	Cga	152Val	^{13}C
1616	31.89	0.116	Cb	152Val	^{13}C
1617	24.66	0.052	Cgb	152Val	^{13}C

1618	178.31	0.004	C	152Val	^{13}C
1619	29.57	0.042	Cb	153Glu	^{13}C
1620	36.03	0.000	Cg	153Glu	^{13}C
1621	3.93	0.009	Ha	153Glu	^{1}H
1622	120.16	0.090	N	153Glu	^{15}N
1623	7.48	0.018	H	153Glu	^{1}H
1624	59.29	0.055	Ca	153Glu	^{13}C
1625	178.16	0.033	C	153Glu	^{13}C
1626	2.18	0.028	Hga	153Glu	^{1}H
1627	2.35	0.012	Hgb	153Glu	^{1}H
1628	4.05	0.005	Ha	154Leu	^{1}H
1629	23.24	0.075	Cda	154Leu	^{13}C
1630	0.87	0.009	Hda*	154Leu	^{1}H
1631	1.10	0.014	Hba	154Leu	^{1}H
1632	1.55	0.026	Hbb	154Leu	^{1}H
1633	7.70	0.019	H	154Leu	^{1}H
1634	117.52	0.099	N	154Leu	^{15}N
1635	56.84	0.066	Ca	154Leu	^{13}C
1636	43.74	0.067	Cb	154Leu	^{13}C
1637	0.88	0.010	Hdb*	154Leu	^{1}H
1638	25.70	0.103	Cdb	154Leu	^{13}C
1639	178.56	0.101	C	154Leu	^{13}C
1640	2.71	0.008	Hbb	155Tyr	^{1}H
1641	1.77	0.029	Hba	155Tyr	^{1}H
1642	8.21	0.004	H	155Tyr	^{1}H
1643	116.47	0.304	N	155Tyr	^{15}N
1644	59.64	0.153	Ca	155Tyr	^{13}C
1645	38.89	0.058	Cb	155Tyr	^{13}C
1646	4.23	0.006	Ha	155Tyr	^{1}H
1647	176.83	0.031	C	155Tyr	^{13}C
1648	6.81	0.004	He*	155Tyr	^{1}H
1649	117.36	0.044	Ce*	155Tyr	^{13}C
1650	6.85	0.005	Hd*	155Tyr	^{1}H
1651	133.25	0.036	Cd*	155Tyr	^{13}C
1652	3.76	0.011	Haa	156Gly	^{1}H
1653	4.09	0.022	Hab	156Gly	^{1}H
1654	8.12	0.015	H	156Gly	^{1}H
1655	108.68	0.059	N	156Gly	^{15}N
1656	45.63	0.076	Ca	156Gly	^{13}C
1657	173.89	0.043	C	156Gly	^{13}C
1658	38.73	0.080	Cb	157Asn	^{13}C
1659	4.64	0.013	Ha	157Asn	^{1}H
1660	118.54	0.082	N	157Asn	^{15}N
1661	8.43	0.015	H	157Asn	^{1}H
1662	53.48	0.030	Ca	157Asn	^{13}C

1663	2.85	0.018	Hba	157Asn	^{1}H
1664	2.85	0.015	Hbb	157Asn	^{1}H
1665	175.44	0.005	C	157Asn	^{13}C
1666	38.84	0.039	Cb	158Asn	^{13}C
1667	8.44	0.024	H	158Asn	^{1}H
1668	119.25	0.075	N	158Asn	^{15}N
1669	7.65	0.009	Hd2b	158Asn	^{1}H
1670	112.95	0.204	Nd2	158Asn	^{15}N
1671	6.89	0.047	Hd2a	158Asn	^{1}H
1672	53.84	0.018	Ca	158Asn	^{13}C
1673	4.63	0.008	Ha	158Asn	^{1}H
1674	2.79	0.000	Hba	158Asn	^{1}H
1675	2.79	0.000	Hbb	158Asn	^{1}H
1676	175.57	0.016	C	158Asn	^{13}C
1677	123.56	0.042	N	159Ala	^{15}N
1678	8.22	0.008	H	159Ala	^{1}H
1679	18.79	0.064	Cb	159Ala	^{13}C
1680	53.73	0.046	Ca	159Ala	^{13}C
1681	1.38	0.007	Hb*	159Ala	^{1}H
1682	4.19	0.008	Ha	159Ala	^{1}H
1683	178.45	0.036	C	159Ala	^{13}C
1684	8.16	0.008	H	160Ala	^{1}H
1685	121.92	0.100	N	160Ala	^{15}N
1686	53.13	0.031	Ca	160Ala	^{13}C
1687	18.62	0.037	Cb	160Ala	^{13}C
1688	1.41	0.005	Hb*	160Ala	^{1}H
1689	4.24	0.012	Ha	160Ala	^{1}H
1690	178.49	0.085	C	160Ala	^{13}C
1691	122.55	0.055	N	161Ala	^{15}N
1692	8.05	0.002	H	161Ala	^{1}H
1693	53.81	0.048	Ca	161Ala	^{13}C
1694	1.41	0.004	Hb*	161Ala	^{1}H
1695	18.84	0.069	Cb	161Ala	^{13}C
1696	4.16	0.008	Ha	161Ala	^{1}H
1697	178.99	0.052	C	161Ala	^{13}C
1698	36.15	0.000	Cg	162Glu	^{13}C
1699	2.29	0.001	Hb2	162Glu	^{1}H
1700	2.29	0.001	Hg3	162Glu	^{1}H
1701	2.29	0.001	Hg2	162Glu	^{1}H
1702	4.17	0.012	Ha	162Glu	^{1}H
1703	2.29	0.001	Hb3	162Glu	^{1}H
1704	8.33	0.003	H	162Glu	^{1}H
1705	118.71	0.087	N	162Glu	^{15}N
1706	57.79	0.059	Ca	162Glu	^{13}C
1707	29.77	0.031	Cb	162Glu	^{13}C

1708	177.63	0.023	C	162Glu	^{13}C
1709	4.38	0.010	Ha	163Ser	^{1}H
1710	3.93	0.002	Hbb	163Ser	^{1}H
1711	3.93	0.001	Hba	163Ser	^{1}H
1712	8.12	0.003	H	163Ser	^{1}H
1713	115.34	0.046	N	163Ser	^{15}N
1714	59.42	0.053	Ca	163Ser	^{13}C
1715	63.46	0.072	Cb	163Ser	^{13}C
1716	175.24	0.028	C	163Ser	^{13}C
1717	43.42	0.000	Cd	164Arg	^{13}C
1718	27.48	0.000	Cg	164Arg	^{13}C
1719	121.90	0.060	N	164Arg	^{15}N
1720	8.05	0.002	H	164Arg	^{1}H
1721	30.51	0.068	Cb	164Arg	^{13}C
1722	56.77	0.038	Ca	164Arg	^{13}C
1723	4.33	0.000	Ha	164Arg	^{1}H
1724	3.17	0.000	Hd3	164Arg	^{1}H
1725	3.17	0.000	Hd2	164Arg	^{1}H
1726	176.81	0.018	C	164Arg	^{13}C
1727	1.66	0.000	Hg3	164Arg	^{1}H
1728	1.86	0.000	Hb3	164Arg	^{1}H
1729	1.86	0.000	Hb2	164Arg	^{1}H
1730	1.66	0.000	Hg2	164Arg	^{1}H
1731	121.16	0.044	N	165Lys	^{15}N
1732	8.11	0.002	H	165Lys	^{1}H
1733	57.27	0.030	Ca	165Lys	^{13}C
1734	32.85	0.046	Cb	165Lys	^{13}C
1735	29.36	0.070	Cd	165Lys	^{13}C
1736	25.05	0.055	Cg	165Lys	^{13}C
1737	4.25	0.000	Ha	165Lys	^{1}H
1738	2.98	0.008	He3	165Lys	^{1}H
1739	1.83	0.000	Hb3	165Lys	^{1}H
1740	2.98	0.008	He2	165Lys	^{1}H
1741	1.66	0.027	Hd3	165Lys	^{1}H
1742	1.66	0.027	Hd2	165Lys	^{1}H
1743	1.83	0.000	Hb2	165Lys	^{1}H
1744	1.45	0.000	Hg2	165Lys	^{1}H
1745	1.45	0.000	Hg3	165Lys	^{1}H
1746	177.58	0.051	C	165Lys	^{13}C
1747	8.44	0.003	H	166Gly	^{1}H
1748	109.47	0.040	N	166Gly	^{15}N
1749	45.69	0.043	Ca	166Gly	^{13}C
1750	3.98	0.001	Ha2	166Gly	^{1}H
1751	3.98	0.001	Ha3	166Gly	^{1}H
1752	174.70	0.000	C	166Gly	^{13}C

1753	119.70	0.121	N	167Gln	^{15}N
1754	8.14	0.010	H	167Gln	^{1}H
1755	6.86	0.006	He2a	167Gln	^{1}H
1756	111.92	0.201	Ne2	167Gln	^{15}N
1757	7.57	0.003	He2b	167Gln	^{1}H
1758	56.16	0.108	Ca	167Gln	^{13}C
1759	29.40	0.027	Cb	167Gln	^{13}C
1760	33.79	0.065	Cg	167Gln	^{13}C
1761	4.34	0.001	Ha	167Gln	^{1}H
1762	2.36	0.003	Hg2	167Gln	^{1}H
1763	2.06	0.002	Hba	167Gln	^{1}H
1764	2.36	0.003	Hg3	167Gln	^{1}H
1765	2.14	0.001	Hbb	167Gln	^{1}H
1766	176.09	0.054	C	167Gln	^{13}C
1767	180.52	0.001	Cd	167Gln	^{13}C
1768	8.53	0.005	H	168Glu	^{1}H
1769	121.76	0.098	N	168Glu	^{15}N
1770	57.34	0.036	Ca	168Glu	^{13}C
1771	1.93	0.000	Hba	168Glu	^{1}H
1772	2.29	0.000	Hg2	168Glu	^{1}H
1773	2.05	0.000	Hbb	168Glu	^{1}H
1774	2.29	0.000	Hg3	168Glu	^{1}H
1775	29.99	0.072	Cb	168Glu	^{13}C
1776	36.32	0.000	Cg	168Glu	^{13}C
1777	4.22	0.000	Ha	168Glu	^{1}H
1778	176.75	0.022	C	168Glu	^{13}C
1779	3.17	0.009	Hd3	169Arg	^{1}H
1780	3.17	0.009	Hd2	169Arg	^{1}H
1781	1.78	0.005	Hb2	169Arg	^{1}H
1782	1.60	0.000	Hga	169Arg	^{1}H
1783	1.64	0.001	Hgb	169Arg	^{1}H
1784	4.29	0.008	Ha	169Arg	^{1}H
1785	1.78	0.005	Hb3	169Arg	^{1}H
1786	121.11	0.154	N	169Arg	^{15}N
1787	30.52	0.087	Cb	169Arg	^{13}C
1788	56.34	0.037	Ca	169Arg	^{13}C
1789	8.28	0.003	H	169Arg	^{1}H
1790	43.43	0.015	Cd	169Arg	^{13}C
1791	27.29	0.066	Cg	169Arg	^{13}C
1792	177.59	0.047	C	169Arg	^{13}C
1793	42.42	0.119	Cb	170Leu	^{13}C
1794	55.29	0.057	Ca	170Leu	^{13}C
1795	122.85	0.043	N	170Leu	^{15}N
1796	8.18	0.003	H	170Leu	^{1}H
1797	4.40	0.003	Ha	170Leu	^{1}H

1798	0.92	0.003	Hdb*	170Leu	^{1}H
1799	0.86	0.002	Hda*	170Leu	^{1}H
1800	1.63	0.014	Hba	170Leu	^{1}H
1801	1.63	0.015	Hg	170Leu	^{1}H
1802	1.63	0.015	Hbb	170Leu	^{1}H
1803	26.98	0.018	Cg	170Leu	^{13}C
1804	23.32	0.079	Cda	170Leu	^{13}C
1805	25.12	0.039	Cdb	170Leu	^{13}C
1806	177.36	0.005	C	170Leu	^{13}C
1807	8.35	0.002	H	171Glu	^{1}H
1808	120.82	0.061	N	171Glu	^{15}N
1809	56.72	0.024	Ca	171Glu	^{13}C
1810	30.29	0.000	Cb	171Glu	^{13}C
1811	176.26	0.000	C	171Glu	^{13}C

Table 21: Bcl-x_L NMR assignments. Residues are numbered ignoring the loop deletion between residues 44-85; SD is standard deviation.

Number	Shift (ppm)	SD (ppm)	Assign Name	Residue	Isotope
1	176.53	0.000	C	3Gln	^{13}C
2	56.75	0.059	Ca	3Gln	^{13}C
3	29.32	0.051	Cb	3Gln	^{13}C
4	33.92	0.047	Cg	3Gln	^{13}C
5	4.37	0.004	Ha	3Gln	^{1}H
6	2.11	0.005	Hbb	3Gln	^{1}H
7	2.11	0.005	Hba	3Gln	^{1}H
8	2.29	0.009	Hga	3Gln	^{1}H
9	2.31	0.017	Hgb	3Gln	^{1}H
10	112.53	0.020	Ne2	3Gln	^{15}N
11	6.86	0.005	He2a	3Gln	^{1}H
12	7.69	0.006	He2b	3Gln	^{1}H
13	116.82	0.064	N	4Ser	^{15}N
14	8.64	0.018	H	4Ser	^{1}H
15	59.15	0.039	Ca	4Ser	^{13}C
16	63.46	0.048	Cb	4Ser	^{13}C
17	4.04	0.007	Hbb	4Ser	^{1}H
18	3.91	0.005	Hba	4Ser	^{1}H
19	4.45	0.005	Ha	4Ser	^{1}H
20	176.10	0.000	C	4Ser	^{13}C
21	122.82	0.057	N	5Asn	^{15}N
22	8.59	0.015	H	5Asn	^{1}H
23	56.22	0.017	Ca	5Asn	^{13}C
24	176.37	0.000	C	5Asn	^{13}C
25	38.47	0.015	Cb	5Asn	^{13}C
26	2.69	0.006	Hba	5Asn	^{1}H
27	3.12	0.010	Hbb	5Asn	^{1}H

28	4.04	0.004	Ha	5Asn	^{1}H
29	112.68	0.036	Nd2	5Asn	^{15}N
30	7.55	0.012	Hd2a	5Asn	^{1}H
31	8.04	0.009	Hd2b	5Asn	^{1}H
32	8.23	0.013	H	6Arg	^{1}H
33	120.09	0.092	N	6Arg	^{15}N
34	59.91	0.051	Ca	6Arg	^{13}C
35	177.29	0.000	C	6Arg	^{13}C
36	3.18	0.008	Hdb	6Arg	^{1}H
37	3.18	0.009	Hda	6Arg	^{1}H
38	3.92	0.010	Ha	6Arg	^{1}H
39	2.35	0.010	Hbb	6Arg	^{1}H
40	2.35	0.012	Hba	6Arg	^{1}H
41	1.63	0.014	Hga	6Arg	^{1}H
42	1.63	0.014	Hgb	6Arg	^{1}H
43	43.25	0.021	Cd	6Arg	^{13}C
44	36.98	0.057	Cg	7Glu	^{13}C
45	7.98	0.012	H	7Glu	^{1}H
46	117.91	0.082	N	7Glu	^{15}N
47	59.47	0.049	Ca	7Glu	^{13}C
48	29.69	0.029	Cb	7Glu	^{13}C
49	179.30	0.000	C	7Glu	^{13}C
50	2.20	0.005	Hga	7Glu	^{1}H
51	2.37	0.005	Hgb	7Glu	^{1}H
52	3.88	0.010	Ha	7Glu	^{1}H
53	2.00	0.001	Hba	7Glu	^{1}H
54	2.10	0.008	Hbb	7Glu	^{1}H
55	119.77	0.074	N	8Leu	^{15}N
56	7.75	0.014	H	8Leu	^{1}H
57	58.25	0.062	Ca	8Leu	^{13}C
58	42.10	0.061	Cb	8Leu	^{13}C
59	0.15	0.007	Hdb*	8Leu	^{1}H
60	24.41	0.028	Cda	8Leu	^{13}C
61	0.06	0.007	Hda*	8Leu	^{1}H
62	23.75	0.022	Cdb	8Leu	^{13}C
63	3.86	0.007	Ha	8Leu	^{1}H
64	1.48	0.005	Hg	8Leu	^{1}H
65	1.33	0.008	Hba	8Leu	^{1}H
66	1.80	0.006	Hbb	8Leu	^{1}H
67	26.22	0.076	Cg	8Leu	^{13}C
68	178.38	0.000	C	8Leu	^{13}C
69	117.74	0.055	N	9Val	^{15}N
70	7.72	0.009	H	9Val	^{1}H
71	67.82	0.067	Ca	9Val	^{13}C
72	22.84	0.030	Cgb	9Val	^{13}C

73	1.20	0.006	Hgb*	9Val	^{1}H
74	2.36	0.004	Hb	9Val	^{1}H
75	1.07	0.006	Hga*	9Val	^{1}H
76	3.67	0.004	Ha	9Val	^{1}H
77	24.22	0.038	Cga	9Val	^{13}C
78	32.50	0.046	Cb	9Val	^{13}C
79	179.84	0.000	C	9Val	^{13}C
80	8.43	0.008	H	10Val	^{1}H
81	67.11	0.146	Ca	10Val	^{13}C
82	117.76	0.038	N	10Val	^{15}N
83	2.81	0.008	Ha	10Val	^{1}H
84	31.47	0.077	Cb	10Val	^{13}C
85	1.70	0.006	Hb	10Val	^{1}H
86	22.88	0.035	Cgb	10Val	^{13}C
87	0.78	0.004	Hgb*	10Val	^{1}H
88	19.83	0.093	Cga	10Val	^{13}C
89	-0.01	0.005	Hga*	10Val	^{1}H
90	178.67	0.000	C	10Val	^{13}C
91	121.14	0.048	N	11Asp	^{15}N
92	8.19	0.008	H	11Asp	^{1}H
93	59.00	0.049	Ca	11Asp	^{13}C
94	43.39	0.031	Cb	11Asp	^{13}C
95	2.56	0.007	Hba	11Asp	^{1}H
96	2.69	0.011	Hbb	11Asp	^{1}H
97	4.06	0.006	Ha	11Asp	^{1}H
98	177.49	0.000	C	11Asp	^{13}C
99	120.22	0.061	N	12Phe	^{15}N
100	8.25	0.009	H	12Phe	^{1}H
101	40.08	0.055	Cb	12Phe	^{13}C
102	63.43	0.069	Ca	12Phe	^{13}C
103	178.01	0.000	C	12Phe	^{13}C
104	3.96	0.011	Ha	12Phe	^{1}H
105	3.12	0.009	Hba	12Phe	^{1}H
106	3.53	0.008	Hbb	12Phe	^{1}H
107	116.85	0.067	N	13Leu	^{15}N
108	8.92	0.007	H	13Leu	^{1}H
109	58.42	0.064	Ca	13Leu	^{13}C
110	179.51	0.000	C	13Leu	^{13}C
111	3.98	0.006	Ha	13Leu	^{1}H
112	1.53	0.019	Hba	13Leu	^{1}H
113	1.78	0.010	Hbb	13Leu	^{1}H
114	41.73	0.058	Cb	13Leu	^{13}C
115	2.25	0.006	Hg	13Leu	^{1}H
116	25.62	0.039	Cd1	13Leu	^{13}C
117	0.95	0.011	Hdb*	13Leu	^{1}H

118	26.95	0.024	Cg	13Leu	^{13}C
119	25.62	0.034	Cd2	13Leu	^{13}C
120	0.95	0.011	Hda*	13Leu	^{1}H
121	114.27	0.057	N	14Ser	^{15}N
122	8.56	0.015	H	14Ser	^{1}H
123	62.93	0.065	Ca	14Ser	^{13}C
124	177.17	0.000	C	14Ser	^{13}C
125	3.85	0.002	Hba	14Ser	^{1}H
126	4.27	0.006	Ha	14Ser	^{1}H
127	62.68	0.038	Cb	14Ser	^{13}C
128	3.98	0.005	Hbb	14Ser	^{1}H
129	125.29	0.067	N	15Tyr	^{15}N
130	8.31	0.014	H	15Tyr	^{1}H
131	60.21	0.069	Ca	15Tyr	^{13}C
132	178.27	0.000	C	15Tyr	^{13}C
133	38.15	0.113	Cb	15Tyr	^{13}C
134	3.07	0.003	Hba	15Tyr	^{1}H
135	3.07	0.003	Hbb	15Tyr	^{1}H
136	4.39	0.005	Ha	15Tyr	^{1}H
137	120.17	0.050	N	16Lys	^{15}N
138	8.26	0.013	H	16Lys	^{1}H
139	57.44	0.066	Ca	16Lys	^{13}C
140	180.42	0.000	C	16Lys	^{13}C
141	41.62	0.029	Ce	16Lys	^{13}C
142	2.95	0.004	Heb	16Lys	^{1}H
143	2.61	0.005	Hea	16Lys	^{1}H
144	3.80	0.007	Ha	16Lys	^{1}H
145	24.26	0.024	Cg	16Lys	^{13}C
146	1.26	0.011	Hgb	16Lys	^{1}H
147	1.26	0.011	Hga	16Lys	^{1}H
148	26.68	0.031	Cd	16Lys	^{13}C
149	1.26	0.005	Hda	16Lys	^{1}H
150	1.51	0.010	Hdb	16Lys	^{1}H
151	30.04	0.030	Cb	16Lys	^{13}C
152	1.62	0.007	Hba	16Lys	^{1}H
153	1.79	0.004	Hbb	16Lys	^{1}H
154	8.69	0.010	H	17Leu	^{1}H
155	118.80	0.051	N	17Leu	^{15}N
156	58.79	0.068	Ca	17Leu	^{13}C
157	41.42	0.062	Cb	17Leu	^{13}C
158	0.98	0.008	Hdb*	17Leu	^{1}H
159	0.79	0.007	Hda*	17Leu	^{1}H
160	24.27	0.027	Cda	17Leu	^{13}C
161	26.50	0.032	Cdb	17Leu	^{13}C
162	1.86	0.006	Hg	17Leu	^{1}H

163	27.98	0.080	Cg	17Leu	^{13}C
164	1.51	0.008	Hba	17Leu	^{1}H
165	2.15	0.006	Hbb	17Leu	^{1}H
166	4.06	0.008	Ha	17Leu	^{1}H
167	177.11	0.000	C	17Leu	^{13}C
168	117.76	0.068	N	18Ser	^{15}N
169	8.44	0.006	H	18Ser	^{1}H
170	61.70	0.049	Ca	18Ser	^{13}C
171	62.86	0.051	Cb	18Ser	^{13}C
172	178.50	0.000	C	18Ser	^{13}C
173	4.10	0.004	Hba	18Ser	^{1}H
174	4.22	0.004	Hbb	18Ser	^{1}H
175	4.71	0.006	Ha	18Ser	^{1}H
176	122.82	0.056	N	19Gln	^{15}N
177	7.65	0.008	H	19Gln	^{1}H
178	58.50	0.063	Ca	19Gln	^{13}C
179	28.99	0.029	Cb	19Gln	^{13}C
180	177.17	0.000	C	19Gln	^{13}C
181	4.02	0.008	Ha	19Gln	^{1}H
182	1.92	0.004	Hba	19Gln	^{1}H
183	2.13	0.011	Hbb	19Gln	^{1}H
184	34.49	0.017	Cg	19Gln	^{13}C
185	2.01	0.007	Hga	19Gln	^{1}H
186	2.10	0.008	Hgb	19Gln	^{1}H
187	114.59	0.134	Ne2	19Gln	^{15}N
188	7.45	0.013	He2b	19Gln	^{1}H
189	6.76	0.006	He2a	19Gln	^{1}H
190	179.56	0.025	Cd	19Gln	^{13}C
191	116.69	0.029	N	20Lys	^{15}N
192	7.27	0.006	H	20Lys	^{1}H
193	54.15	0.162	Ca	20Lys	^{13}C
194	32.92	0.035	Cb	20Lys	^{13}C
195	25.21	0.037	Cg	20Lys	^{13}C
196	28.31	0.114	Cd	20Lys	^{13}C
197	1.68	0.005	Hda	20Lys	^{1}H
198	1.68	0.006	Hdb	20Lys	^{1}H
199	1.57	0.016	Hgb	20Lys	^{1}H
200	1.52	0.010	Hga	20Lys	^{1}H
201	4.43	0.005	Ha	20Lys	^{1}H
202	42.38	0.084	Ce	20Lys	^{13}C
203	3.01	0.004	Hea	20Lys	^{1}H
204	3.01	0.004	Heb	20Lys	^{1}H
205	2.07	0.006	Hbb	20Lys	^{1}H
206	1.70	0.008	Hba	20Lys	^{1}H
207	175.54	0.000	C	20Lys	^{13}C

208	105.46	0.032	N	21Gly	^{15}N
209	7.59	0.006	H	21Gly	^{1}H
210	45.46	0.076	Ca	21Gly	^{13}C
211	174.02	0.000	C	21Gly	^{13}C
212	3.97	0.005	Hab	21Gly	^{1}H
213	3.67	0.002	Haa	21Gly	^{1}H
214	119.67	0.050	N	22Tyr	^{15}N
215	7.84	0.004	H	22Tyr	^{1}H
216	55.44	0.043	Ca	22Tyr	^{13}C
217	41.22	0.065	Cb	22Tyr	^{13}C
218	5.12	0.007	Ha	22Tyr	^{1}H
219	3.28	0.008	Hbb	22Tyr	^{1}H
220	175.04	0.000	C	22Tyr	^{13}C
221	2.81	0.009	Hba	22Tyr	^{1}H
222	115.03	0.028	N	23Ser	^{15}N
223	8.63	0.007	H	23Ser	^{1}H
224	65.09	0.070	Cb	23Ser	^{13}C
225	56.71	0.057	Ca	23Ser	^{13}C
226	3.85	0.008	Hbb	23Ser	^{1}H
227	3.83	0.005	Hba	23Ser	^{1}H
228	4.94	0.004	Ha	23Ser	^{1}H
229	174.56	0.000	C	23Ser	^{13}C
230	9.09	0.008	H	24Trp	^{1}H
231	127.81	0.076	N	24Trp	^{15}N
232	128.95	0.063	Ne1	24Trp	^{15}N
233	10.10	0.008	He1	24Trp	^{1}H
234	60.59	0.049	Ca	24Trp	^{13}C
235	30.89	0.041	Cb	24Trp	^{13}C
236	3.21	0.007	Hba	24Trp	^{1}H
237	3.51	0.007	Hbb	24Trp	^{1}H
238	177.00	0.000	C	24Trp	^{13}C
239	4.30	0.011	Ha	24Trp	^{1}H
240	111.11	0.052	N	25Ser	^{15}N
241	8.30	0.006	H	25Ser	^{1}H
242	62.97	0.042	Cb	25Ser	^{13}C
243	60.44	0.072	Ca	25Ser	^{13}C
244	4.08	0.006	Ha	25Ser	^{1}H
245	3.90	0.009	Hba	25Ser	^{1}H
246	3.90	0.009	Hbb	25Ser	^{1}H
247	175.30	0.000	C	25Ser	^{13}C
248	6.84	0.008	He2a	26Gln	^{1}H
249	7.48	0.004	H	26Gln	^{1}H
250	119.25	0.058	N	26Gln	^{15}N
251	111.07	0.218	Ne2	26Gln	^{15}N
252	7.35	0.006	He2b	26Gln	^{1}H

253	57.26	0.068	Ca	26Gln	^{13}C
254	28.74	0.106	Cb	26Gln	^{13}C
255	34.43	0.084	Cg	26Gln	^{13}C
256	3.86	0.009	Ha	26Gln	^{1}H
257	1.40	0.005	Hbb	26Gln	^{1}H
258	1.23	0.007	Hba	26Gln	^{1}H
259	1.84	0.005	Hga	26Gln	^{1}H
260	1.94	0.007	Hgb	26Gln	^{1}H
261	176.32	0.000	C	26Gln	^{13}C
262	180.23	0.016	Cd	26Gln	^{13}C
263	115.74	0.050	N	27Phe	^{15}N
264	7.22	0.014	H	27Phe	^{1}H
265	56.47	0.035	Ca	27Phe	^{13}C
266	40.31	0.057	Cb	27Phe	^{13}C
267	4.60	0.006	Ha	27Phe	^{1}H
268	3.08	0.009	Hbb	27Phe	^{1}H
269	2.08	0.011	Hba	27Phe	^{1}H
270	174.85	0.000	C	27Phe	^{13}C
271	113.53	0.078	N	28Ser	^{15}N
272	7.16	0.011	H	28Ser	^{1}H
273	63.95	0.069	Cb	28Ser	^{13}C
274	57.45	0.073	Ca	28Ser	^{13}C
275	4.09	0.007	Ha	28Ser	^{1}H
276	3.06	0.008	Hbb	28Ser	^{1}H
277	2.99	0.006	Hba	28Ser	^{1}H
278	173.63	0.000	C	28Ser	^{13}C
279	122.86	0.049	N	29Asp	^{15}N
280	8.36	0.004	H	29Asp	^{1}H
281	54.11	0.070	Ca	29Asp	^{13}C
282	41.25	0.076	Cb	29Asp	^{13}C
283	2.66	0.006	Hbb	29Asp	^{1}H
284	2.66	0.006	Hba	29Asp	^{1}H
285	4.59	0.007	Ha	29Asp	^{1}H
286	176.30	0.000	C	29Asp	^{13}C
287	119.30	0.055	N	30Val	^{15}N
288	7.97	0.006	H	30Val	^{1}H
289	62.42	0.091	Ca	30Val	^{13}C
290	33.19	0.076	Cb	30Val	^{13}C
291	2.07	0.005	Hb	30Val	^{1}H
292	0.90	0.008	Hgb*	30Val	^{1}H
293	0.90	0.009	Hga*	30Val	^{1}H
294	4.09	0.005	Ha	30Val	^{1}H
295	21.38	0.045	Cga	30Val	^{13}C
296	20.58	0.045	Cga	30Val	^{13}C
297	176.19	0.000	C	30Val	^{13}C

298	124.78	0.039	N	31Glu	^{15}N
299	8.55	0.007	H	31Glu	^{1}H
300	56.58	0.042	Ca	31Glu	^{13}C
301	30.34	0.065	Cb	31Glu	^{13}C
302	36.49	0.018	Cg	31Glu	^{13}C
303	4.29	0.006	Ha	31Glu	^{1}H
304	2.27	0.006	Hgb	31Glu	^{1}H
305	2.26	0.006	Hga	31Glu	^{1}H
306	2.04	0.001	Hbb	31Glu	^{1}H
307	1.93	0.013	Hba	31Glu	^{1}H
308	176.50	0.000	C	31Glu	^{13}C
309	8.46	0.005	H	32Glu	^{1}H
310	122.42	0.028	N	32Glu	^{15}N
311	56.67	0.055	Ca	32Glu	^{13}C
312	30.50	0.073	Cb	32Glu	^{13}C
313	36.34	0.071	Cg	32Glu	^{13}C
314	4.23	0.006	Ha	32Glu	^{1}H
315	2.26	0.009	Hgb	32Glu	^{1}H
316	2.26	0.009	Hga	32Glu	^{1}H
317	1.92	0.005	Hba	32Glu	^{1}H
318	2.04	0.010	Hbb	32Glu	^{1}H
319	176.16	0.000	C	32Glu	^{13}C
320	119.78	0.045	N	33Asn	^{15}N
321	8.53	0.005	H	33Asn	^{1}H
322	112.81	0.205	Nd2	33Asn	^{15}N
323	7.60	0.006	Hd2b	33Asn	^{1}H
324	6.93	0.004	Hd2a	33Asn	^{1}H
325	39.00	0.080	Cb	33Asn	^{13}C
326	53.40	0.105	Ca	33Asn	^{13}C
327	2.83	0.010	Hbb	33Asn	^{1}H
328	2.73	0.011	Hba	33Asn	^{1}H
329	4.69	0.017	Ha	33Asn	^{1}H
330	175.00	0.000	C	33Asn	^{13}C
331	177.07	0.033	Cg	33Asn	^{13}C
332	122.22	0.066	N	34Arg	^{15}N
333	8.37	0.004	H	34Arg	^{1}H
334	27.33	0.066	Cg	34Arg	^{13}C
335	30.96	0.038	Cb	34Arg	^{13}C
336	56.15	0.030	Ca	34Arg	^{13}C
337	43.41	0.060	Cd	34Arg	^{13}C
338	4.42	0.005	Ha	34Arg	^{1}H
339	1.61	0.007	Hga	34Arg	^{1}H
340	1.61	0.007	Hgb	34Arg	^{1}H
341	1.77	0.010	Hba	34Arg	^{1}H
342	1.89	0.006	Hbb	34Arg	^{1}H

343	3.19	0.007	Hda	34Arg	^{1}H
344	3.19	0.007	Hdb	34Arg	^{1}H
345	176.46	0.000	C	34Arg	^{13}C
346	8.30	0.004	H	35Thr	^{1}H
347	115.99	0.041	N	35Thr	^{15}N
348	69.85	0.116	Cb	35Thr	^{13}C
349	62.11	0.073	Ca	35Thr	^{13}C
350	21.89	0.165	Cg2	35Thr	^{13}C
351	1.20	0.004	Hg2*	35Thr	^{1}H
352	4.33	0.005	Ha	35Thr	^{1}H
353	4.22	0.004	Hb	35Thr	^{1}H
354	174.41	0.000	C	35Thr	^{13}C
355	123.18	0.073	N	36Glu	^{15}N
356	8.42	0.004	H	36Glu	^{1}H
357	30.43	0.034	Cb	36Glu	^{13}C
358	56.35	0.026	Ca	36Glu	^{13}C
359	36.28	0.054	Cg	36Glu	^{13}C
360	4.31	0.008	Ha	36Glu	^{1}H
361	2.06	0.002	Hbb	36Glu	^{1}H
362	1.92	0.002	Hba	36Glu	^{1}H
363	2.26	0.003	Hga	36Glu	^{1}H
364	2.26	0.003	Hgb	36Glu	^{1}H
365	175.64	0.000	C	36Glu	^{13}C
366	126.24	0.050	N	37Ala	^{15}N
367	8.32	0.004	H	37Ala	^{1}H
368	50.60	0.088	Ca	37Ala	^{13}C
369	18.66	0.062	Cb	37Ala	^{13}C
370	1.35	0.003	Hb*	37Ala	^{1}H
371	4.61	0.004	Ha	37Ala	^{1}H
372	63.23	0.137	Ca	38Pro	^{13}C
373	4.43	0.008	Ha	38Pro	^{1}H
374	3.77	0.005	Hdb	38Pro	^{1}H
375	50.68	0.091	Cd	38Pro	^{13}C
376	3.66	0.003	Hda	38Pro	^{1}H
377	32.22	0.055	Cb	38Pro	^{13}C
378	2.29	0.008	Hbb	38Pro	^{1}H
379	1.94	0.004	Hba	38Pro	^{1}H
380	27.59	0.086	Cg	38Pro	^{13}C
381	177.06	0.000	C	38Pro	^{13}C
382	2.03	0.011	Hgb	38Pro	^{1}H
383	2.02	0.007	Hga	38Pro	^{1}H
384	121.25	0.020	N	39Glu	^{15}N
385	8.65	0.007	H	39Glu	^{1}H
386	30.52	0.041	Cb	39Glu	^{13}C
387	56.99	0.085	Ca	39Glu	^{13}C

388	36.48	0.046	Cg	39Glu	^{13}C
389	4.30	0.004	Ha	39Glu	^{1}H
390	2.07	0.002	Hbb	39Glu	^{1}H
391	1.99	0.002	Hba	39Glu	^{1}H
392	2.31	0.004	Hga	39Glu	^{1}H
393	2.31	0.004	Hgb	39Glu	^{1}H
394	177.12	0.000	C	39Glu	^{13}C
395	110.35	0.059	N	40Gly	^{15}N
396	8.53	0.007	H	40Gly	^{1}H
397	45.64	0.094	Ca	40Gly	^{13}C
398	4.09	0.007	Haa	40Gly	^{1}H
399	4.09	0.007	Hab	40Gly	^{1}H
400	174.65	0.000	C	40Gly	^{13}C
401	113.03	0.052	N	41Thr	^{15}N
402	8.23	0.004	H	41Thr	^{1}H
403	21.88	0.079	Cg2	41Thr	^{13}C
404	70.01	0.127	Cb	41Thr	^{13}C
405	62.43	0.104	Ca	41Thr	^{13}C
406	1.20	0.007	Hg2*	41Thr	^{1}H
407	4.39	0.006	Ha	41Thr	^{1}H
408	4.32	0.008	Hb	41Thr	^{1}H
409	175.57	0.000	C	41Thr	^{13}C
410	123.33	0.133	N	42Glu	^{15}N
411	8.89	0.007	H	42Glu	^{1}H
412	29.84	0.113	Cb	42Glu	^{13}C
413	57.76	0.076	Ca	42Glu	^{13}C
414	36.48	0.056	Cg	42Glu	^{13}C
415	4.34	0.004	Ha	42Glu	^{1}H
416	2.32	0.008	Hga	42Glu	^{1}H
417	2.32	0.008	Hgb	42Glu	^{1}H
418	2.10	0.007	Hbb	42Glu	^{1}H
419	2.00	0.002	Hba	42Glu	^{1}H
420	177.28	0.000	C	42Glu	^{13}C
421	115.44	0.053	N	43Ser	^{15}N
422	8.47	0.007	H	43Ser	^{1}H
423	59.06	0.059	Ca	43Ser	^{13}C
424	63.78	0.034	Cb	43Ser	^{13}C
425	4.42	0.003	Ha	43Ser	^{1}H
426	3.92	0.011	Hbb	43Ser	^{1}H
427	3.88	0.009	Hba	43Ser	^{1}H
428	175.80	0.000	C	43Ser	^{13}C
429	123.11	0.073	N	44Glu	^{15}N
430	8.28	0.009	H	44Glu	^{1}H
431	59.11	0.049	Ca	44Glu	^{13}C
432	3.99	0.008	Ha	44Glu	^{1}H

433	2.32	0.006	Hga	44Glu	^{1}H
434	2.32	0.006	Hgb	44Glu	^{1}H
435	2.05	0.006	Hbb	44Glu	^{1}H
436	2.05	0.004	Hba	44Glu	^{1}H
437	36.20	0.040	Cg	44Glu	^{13}C
438	29.56	0.012	Cb	44Glu	^{13}C
439	177.93	0.000	C	44Glu	^{13}C
440	120.85	0.079	N	45Ala	^{15}N
441	8.23	0.005	H	45Ala	^{1}H
442	18.72	0.047	Cb	45Ala	^{13}C
443	1.49	0.009	Hb*	45Ala	^{1}H
444	4.15	0.005	Ha	45Ala	^{1}H
445	55.24	0.045	Ca	45Ala	^{13}C
446	180.64	0.000	C	45Ala	^{13}C
447	118.91	0.035	N	46Val	^{15}N
448	7.55	0.007	H	46Val	^{1}H
449	66.23	0.111	Ca	46Val	^{13}C
450	3.50	0.006	Ha	46Val	^{1}H
451	22.92	0.031	Cgb	46Val	^{13}C
452	1.02	0.006	Hgb*	46Val	^{1}H
453	0.49	0.005	Hga*	46Val	^{1}H
454	22.39	0.080	Cga	46Val	^{13}C
455	31.87	0.084	Cb	46Val	^{13}C
456	1.93	0.007	Hb	46Val	^{1}H
457	176.77	0.000	C	46Val	^{13}C
458	7.46	0.007	H	47Lys	^{1}H
459	117.94	0.040	N	47Lys	^{15}N
460	59.83	0.044	Ca	47Lys	^{13}C
461	178.95	0.000	C	47Lys	^{13}C
462	3.74	0.009	Ha	47Lys	^{1}H
463	42.30	0.028	Ce	47Lys	^{13}C
464	2.54	0.005	Hea	47Lys	^{1}H
465	2.71	0.007	Heb	47Lys	^{1}H
466	32.57	0.031	Cb	47Lys	^{13}C
467	1.87	0.011	Hbb	47Lys	^{1}H
468	1.62	0.007	Hba	47Lys	^{1}H
469	30.45	0.028	Cd	47Lys	^{13}C
470	26.79	0.051	Cg	47Lys	^{13}C
471	1.56	0.005	Hdb	47Lys	^{1}H
472	1.42	0.005	Hda	47Lys	^{1}H
473	1.31	0.009	Hgb	47Lys	^{1}H
474	1.15	0.014	Hga	47Lys	^{1}H
475	7.88	0.007	He2b	48Gln	^{1}H
476	6.68	0.007	He2a	48Gln	^{1}H
477	112.30	0.181	Ne2	48Gln	^{15}N

478	116.29	0.069	N	48Gln	^{15}N
479	8.18	0.009	H	48Gln	^{1}H
480	59.08	0.078	Ca	48Gln	^{13}C
481	3.72	0.005	Ha	48Gln	^{1}H
482	33.32	0.087	Cg	48Gln	^{13}C
483	2.39	0.005	Hga	48Gln	^{1}H
484	2.39	0.004	Hgb	48Gln	^{1}H
485	2.09	0.007	Hbb	48Gln	^{1}H
486	2.00	0.012	Hba	48Gln	^{1}H
487	177.59	0.000	C	48Gln	^{13}C
488	179.75	0.016	Cd	48Gln	^{13}C
489	28.16	0.097	Cb	48Gln	^{13}C
490	120.45	0.040	N	49Ala	^{15}N
491	7.81	0.007	H	49Ala	^{1}H
492	55.09	0.055	Ca	49Ala	^{13}C
493	18.77	0.067	Cb	49Ala	^{13}C
494	1.45	0.008	Hb*	49Ala	^{1}H
495	179.63	0.000	C	49Ala	^{13}C
496	4.27	0.010	Ha	49Ala	^{1}H
497	120.43	0.065	N	50Leu	^{15}N
498	8.48	0.010	H	50Leu	^{1}H
499	58.65	0.078	Ca	50Leu	^{13}C
500	42.39	0.069	Cb	50Leu	^{13}C
501	-0.43	0.005	Hda*	50Leu	^{1}H
502	23.17	0.082	Cda	50Leu	^{13}C
503	23.57	0.069	Cdb	50Leu	^{13}C
504	0.07	0.005	Hdb*	50Leu	^{1}H
505	26.83	0.093	Cg	50Leu	^{13}C
506	1.19	0.003	Hg	50Leu	^{1}H
507	1.45	0.007	Hbb	50Leu	^{1}H
508	1.45	0.006	Hba	50Leu	^{1}H
509	4.15	0.008	Ha	50Leu	^{1}H
510	179.31	0.000	C	50Leu	^{13}C
511	118.90	0.066	N	51Arg	^{15}N
512	8.25	0.008	H	51Arg	^{1}H
513	61.07	0.066	Ca	51Arg	^{13}C
514	178.53	0.000	C	51Arg	^{13}C
515	3.82	0.007	Ha	51Arg	^{1}H
516	44.36	0.025	Cd	51Arg	^{13}C
517	2.21	0.007	Hdb	51Arg	^{1}H
518	27.31	0.037	Cg	51Arg	^{13}C
519	0.65	0.005	Hga	51Arg	^{1}H
520	1.98	0.007	Hgb	51Arg	^{1}H
521	1.98	0.004	Hda	51Arg	^{1}H
522	30.93	0.032	Cb	51Arg	^{13}C

523	1.10	0.007	Hba	51Arg	^{1}H
524	1.70	0.010	Hbb	51Arg	^{1}H
525	116.70	0.068	N	52Glu	^{15}N
526	8.00	0.007	H	52Glu	^{1}H
527	36.97	0.044	Cg	52Glu	^{13}C
528	59.60	0.060	Ca	52Glu	^{13}C
529	180.04	0.000	C	52Glu	^{13}C
530	2.60	0.004	Hgb	52Glu	^{1}H
531	2.39	0.004	Hga	52Glu	^{1}H
532	4.17	0.006	Ha	52Glu	^{1}H
533	2.17	0.011	Hba	52Glu	^{1}H
534	2.29	0.001	Hbb	52Glu	^{1}H
535	29.71	0.019	Cb	52Glu	^{13}C
536	8.90	0.010	H	53Ala	^{1}H
537	123.84	0.071	N	53Ala	^{15}N
538	19.17	0.044	Cb	53Ala	^{13}C
539	55.33	0.053	Ca	53Ala	^{13}C
540	4.50	0.012	Ha	53Ala	^{1}H
541	1.68	0.005	Hb*	53Ala	^{1}H
542	181.21	0.000	C	53Ala	^{13}C
543	9.23	0.009	H	54Gly	^{1}H
544	107.44	0.083	N	54Gly	^{15}N
545	47.13	0.093	Ca	54Gly	^{13}C
546	176.29	0.000	C	54Gly	^{13}C
547	4.05	0.008	Haa	54Gly	^{1}H
548	4.07	0.015	Hab	54Gly	^{1}H
549	124.09	0.038	N	55Asp	^{15}N
550	8.67	0.009	H	55Asp	^{1}H
551	57.97	0.065	Ca	55Asp	^{13}C
552	39.80	0.043	Cb	55Asp	^{13}C
553	4.55	0.004	Ha	55Asp	^{1}H
554	3.10	0.004	Hbb	55Asp	^{1}H
555	2.74	0.006	Hba	55Asp	^{1}H
556	179.13	0.000	C	55Asp	^{13}C
557	121.67	0.049	N	56Glu	^{15}N
558	8.28	0.009	H	56Glu	^{1}H
559	59.69	0.047	Ca	56Glu	^{13}C
560	36.34	0.059	Cg	56Glu	^{13}C
561	178.66	0.000	C	56Glu	^{13}C
562	29.57	0.009	Cb	56Glu	^{13}C
563	2.32	0.006	Hbb	56Glu	^{1}H
564	2.32	0.006	Hba	56Glu	^{1}H
565	2.45	0.003	Hgb	56Glu	^{1}H
566	2.32	0.007	Hga	56Glu	^{1}H
567	4.17	0.004	Ha	56Glu	^{1}H

568	8.70	0.012	H	57Phe	^{1}H
569	122.19	0.055	N	57Phe	^{15}N
570	62.54	0.072	Ca	57Phe	^{13}C
571	176.94	0.000	C	57Phe	^{13}C
572	40.42	0.020	Cb	57Phe	^{13}C
573	3.28	0.012	Hba	57Phe	^{1}H
574	3.70	0.007	Hbb	57Phe	^{1}H
575	4.06	0.002	Ha	57Phe	^{1}H
576	8.69	0.007	H	58Glu	^{1}H
577	118.26	0.034	N	58Glu	^{15}N
578	59.74	0.077	Ca	58Glu	^{13}C
579	177.69	0.000	C	58Glu	^{13}C
580	30.01	0.015	Cb	58Glu	^{13}C
581	2.28	0.005	Hbb	58Glu	^{1}H
582	2.23	0.003	Hba	58Glu	^{1}H
583	37.30	0.042	Cg	58Glu	^{13}C
584	2.56	0.003	Hgb	58Glu	^{1}H
585	2.29	0.008	Hga	58Glu	^{1}H
586	3.79	0.003	Ha	58Glu	^{1}H
587	118.68	0.033	N	59Leu	^{15}N
588	7.94	0.007	H	59Leu	^{1}H
589	58.03	0.081	Ca	59Leu	^{13}C
590	24.95	0.204	Cdb	59Leu	^{13}C
591	0.90	0.007	Hda*	59Leu	^{1}H
592	0.91	0.006	Hdb*	59Leu	^{1}H
593	24.49	0.157	Cda	59Leu	^{13}C
594	4.03	0.004	Ha	59Leu	^{1}H
595	1.77	0.006	Hba	59Leu	^{1}H
596	1.77	0.006	Hbb	59Leu	^{1}H
597	1.63	0.004	Hg	59Leu	^{1}H
598	42.27	0.072	Cb	59Leu	^{13}C
599	27.22	0.026	Cg	59Leu	^{13}C
600	179.38	0.000	C	59Leu	^{13}C
601	117.23	0.116	N	60Arg	^{15}N
602	7.59	0.008	H	60Arg	^{1}H
603	58.72	0.068	Ca	60Arg	^{13}C
604	30.43	0.000	Cb	60Arg	^{13}C
605	3.77	0.009	Ha	60Arg	^{1}H
606	44.00	0.023	Cd	60Arg	^{13}C
607	2.89	0.004	Hdb	60Arg	^{1}H
608	2.83	0.001	Hda	60Arg	^{1}H
609	1.25	0.002	Hga	60Arg	^{1}H
610	1.25	0.002	Hgb	60Arg	^{1}H
611	121.70	0.066	N	64Ala	^{15}N
612	8.08	0.005	H	64Ala	^{1}H

613	1.50	0.004	Hb*	64Ala	^{1}H
614	19.54	0.050	Cb	64Ala	^{13}C
615	53.76	0.067	Ca	64Ala	^{13}C
616	4.38	0.011	Ha	64Ala	^{1}H
617	176.28	0.000	C	65Phe	^{13}C
618	115.36	0.078	N	65Phe	^{15}N
619	8.19	0.008	H	65Phe	^{1}H
620	39.03	0.020	Cb	65Phe	^{13}C
621	3.34	0.004	Hbb	65Phe	^{1}H
622	3.23	0.000	Hba	65Phe	^{1}H
623	115.19	0.044	N	66Ser	^{15}N
624	7.83	0.011	H	66Ser	^{1}H
625	60.57	0.064	Ca	66Ser	^{13}C
626	63.36	0.087	Cb	66Ser	^{13}C
627	3.97	0.004	Hba	66Ser	^{1}H
628	4.33	0.003	Ha	66Ser	^{1}H
629	175.46	0.000	C	66Ser	^{13}C
630	3.97	0.000	Hbb	66Ser	^{1}H
631	124.04	0.064	N	67Asp	^{15}N
632	8.48	0.012	H	67Asp	^{1}H
633	55.30	0.076	Ca	67Asp	^{13}C
634	40.58	0.058	Cb	67Asp	^{13}C
635	2.95	0.003	Hbb	67Asp	^{1}H
636	2.67	0.007	Hba	67Asp	^{1}H
637	4.66	0.013	Ha	67Asp	^{1}H
638	177.57	0.000	C	67Asp	^{13}C
639	122.45	0.067	N	68Leu	^{15}N
640	9.03	0.010	H	68Leu	^{1}H
641	58.41	0.072	Ca	68Leu	^{13}C
642	23.86	0.036	Cdb	68Leu	^{13}C
643	0.91	0.008	Hda*	68Leu	^{1}H
644	0.93	0.008	Hdb*	68Leu	^{1}H
645	1.41	0.004	Hba	68Leu	^{1}H
646	4.03	0.007	Ha	68Leu	^{1}H
647	1.93	0.009	Hbb	68Leu	^{1}H
648	26.14	0.033	Cda	68Leu	^{13}C
649	43.05	0.040	Cb	68Leu	^{13}C
650	179.27	0.000	C	68Leu	^{13}C
651	1.95	0.023	Hg	68Leu	^{1}H
652	26.96	0.020	Cg	68Leu	^{13}C
653	9.01	0.016	H	69Thr	^{1}H
654	110.57	0.067	N	69Thr	^{15}N
655	66.65	0.077	Ca	69Thr	^{13}C
656	68.12	0.087	Cb	69Thr	^{13}C
657	4.16	0.005	Hb	69Thr	^{1}H

658	3.82	0.011	Ha	69Thr	^{1}H
659	1.12	0.011	Hg2*	69Thr	^{1}H
660	22.20	0.034	Cg2	69Thr	^{13}C
661	178.07	0.000	C	69Thr	^{13}C
662	117.64	0.068	N	70Ser	^{15}N
663	7.77	0.011	H	70Ser	^{1}H
664	61.14	0.134	Ca	70Ser	^{13}C
665	3.94	0.007	Hba	70Ser	^{1}H
666	4.35	0.007	Ha	70Ser	^{1}H
667	3.94	0.007	Hbb	70Ser	^{1}H
668	62.78	0.061	Cb	70Ser	^{13}C
669	178.47	0.000	C	70Ser	^{13}C
670	111.39	0.029	Ne2	71Gln	^{15}N
671	7.39	0.008	He2b	71Gln	^{1}H
672	119.35	0.050	N	71Gln	^{15}N
673	7.94	0.008	H	71Gln	^{1}H
674	58.32	0.058	Ca	71Gln	^{13}C
675	177.01	0.000	C	71Gln	^{13}C
676	4.03	0.006	Ha	71Gln	^{1}H
677	34.69	0.020	Cg	71Gln	^{13}C
678	2.41	0.005	Hga	71Gln	^{1}H
679	2.52	0.006	Hgb	71Gln	^{1}H
680	6.78	0.004	He2a	71Gln	^{1}H
681	2.11	0.013	Hba	71Gln	^{1}H
682	2.11	0.013	Hbb	71Gln	^{1}H
683	115.01	0.060	N	72Leu	^{15}N
684	7.71	0.010	H	72Leu	^{1}H
685	54.49	0.036	Ca	72Leu	^{13}C
686	42.49	0.051	Cb	72Leu	^{13}C
687	23.10	0.014	Cda	72Leu	^{13}C
688	0.87	0.004	Hda*	72Leu	^{1}H
689	26.61	0.030	Cdb	72Leu	^{13}C
690	1.05	0.010	Hdb*	72Leu	^{1}H
691	4.33	0.006	Ha	72Leu	^{1}H
692	1.83	0.004	Hg	72Leu	^{1}H
693	27.12	0.074	Cg	72Leu	^{13}C
694	1.84	0.008	Hbb	72Leu	^{1}H
695	1.59	0.010	Hba	72Leu	^{1}H
696	118.08	0.088	N	73His	^{15}N
697	7.59	0.007	H	73His	^{1}H
698	56.19	0.055	Ca	73His	^{13}C
699	27.20	0.148	Cb	73His	^{13}C
700	174.83	0.000	C	73His	^{13}C
701	3.23	0.010	Hba	73His	^{1}H
702	3.23	0.010	Hbb	73His	^{1}H

703	4.34	0.006	Ha	73His	^{1}H
704	118.49	0.098	N	74Ile	^{15}N
705	8.40	0.008	H	74Ile	^{1}H
706	62.20	0.059	Ca	74Ile	^{13}C
707	37.71	0.033	Cb	74Ile	^{13}C
708	12.00	0.024	Cd1	74Ile	^{13}C
709	0.82	0.009	Hd1*	74Ile	^{1}H
710	1.81	0.009	Hb	74Ile	^{1}H
711	0.80	0.008	Hg2*	74Ile	^{1}H
712	17.81	0.024	Cg2	74Ile	^{13}C
713	3.88	0.011	Ha	74Ile	^{1}H
714	175.45	0.000	C	74Ile	^{13}C
715	27.43	0.029	Cg1	74Ile	^{13}C
716	1.42	0.009	Hg1b	74Ile	^{1}H
717	1.41	0.006	Hg1a	74Ile	^{1}H
718	115.05	0.035	N	75Thr	^{15}N
719	7.30	0.010	H	75Thr	^{1}H
720	59.16	0.033	Ca	75Thr	^{13}C
721	4.98	0.010	Ha	75Thr	^{1}H
722	22.31	0.012	Cg2	75Thr	^{13}C
723	1.25	0.004	Hg2*	75Thr	^{1}H
724	71.03	0.038	Cb	75Thr	^{13}C
725	4.54	0.005	Hb	75Thr	^{1}H
726	50.55	0.022	Cd	76Pro	^{13}C
727	4.12	0.007	Hdb	76Pro	^{1}H
728	3.93	0.008	Hda	76Pro	^{1}H
729	28.98	0.036	Cg	76Pro	^{13}C
730	2.42	0.007	Hgb	76Pro	^{1}H
731	1.70	0.007	Hga	76Pro	^{1}H
732	67.04	0.021	Ca	76Pro	^{13}C
733	4.38	0.005	Ha	76Pro	^{1}H
734	31.56	0.059	Cb	76Pro	^{13}C
735	2.29	0.014	Hbb	76Pro	^{1}H
736	2.10	0.008	Hba	76Pro	^{1}H
737	46.13	0.095	Ca	77Gly	^{13}C
738	175.29	0.000	C	77Gly	^{13}C
739	4.23	0.008	Hab	77Gly	^{1}H
740	3.97	0.006	Haa	77Gly	^{1}H
741	108.62	0.048	N	78Thr	^{15}N
742	7.63	0.007	H	78Thr	^{1}H
743	61.61	0.074	Ca	78Thr	^{13}C
744	1.18	0.005	Hg2*	78Thr	^{1}H
745	21.50	0.036	Cg2	78Thr	^{13}C
746	4.87	0.018	Ha	78Thr	^{1}H
747	4.76	0.033	Hb	78Thr	^{1}H

748	69.83	0.000	Cb	78Thr	^{13}C
749	175.81	0.000	C	78Thr	^{13}C
750	125.56	0.096	N	79Ala	^{15}N
751	7.85	0.006	H	79Ala	^{1}H
752	56.44	0.101	Ca	79Ala	^{13}C
753	20.44	0.032	Cb	79Ala	^{13}C
754	1.80	0.004	Hb*	79Ala	^{1}H
755	4.12	0.008	Ha	79Ala	^{1}H
756	177.75	0.000	C	79Ala	^{13}C
757	8.87	0.009	H	80Tyr	^{1}H
758	117.96	0.045	N	80Tyr	^{15}N
759	61.65	0.055	Ca	80Tyr	^{13}C
760	37.16	0.010	Cb	80Tyr	^{13}C
761	176.60	0.000	C	80Tyr	^{13}C
762	2.83	0.005	Ha	80Tyr	^{1}H
763	2.45	0.015	Hbb	80Tyr	^{1}H
764	2.42	0.016	Hba	80Tyr	^{1}H
765	132.12	0.048	Cd*	80Tyr	^{13}C
766	5.76	0.005	Hd*	80Tyr	^{1}H
767	117.82	0.070	Ce*	80Tyr	^{13}C
768	6.49	0.003	He*	80Tyr	^{1}H
769	117.80	0.051	N	81Gln	^{15}N
770	8.08	0.007	H	81Gln	^{1}H
771	6.90	0.004	He2a	81Gln	^{1}H
772	111.81	0.225	Ne2	81Gln	^{15}N
773	7.65	0.004	He2b	81Gln	^{1}H
774	59.29	0.068	Ca	81Gln	^{13}C
775	28.16	0.113	Cb	81Gln	^{13}C
776	34.75	0.084	Cg	81Gln	^{13}C
777	2.49	0.014	Hgb	81Gln	^{1}H
778	2.44	0.005	Hga	81Gln	^{1}H
779	3.83	0.005	Ha	81Gln	^{1}H
780	2.19	0.006	Hbb	81Gln	^{1}H
781	2.09	0.003	Hba	81Gln	^{1}H
782	179.28	0.000	C	81Gln	^{13}C
783	180.63	0.004	Cd	81Gln	^{13}C
784	115.76	0.076	N	82Ser	^{15}N
785	7.85	0.009	H	82Ser	^{1}H
786	62.99	0.064	Ca	82Ser	^{13}C
787	174.49	0.000	C	82Ser	^{13}C
788	4.28	0.005	Ha	82Ser	^{1}H
789	63.27	0.036	Cb	82Ser	^{13}C
790	3.84	0.008	Hba	82Ser	^{1}H
791	4.03	0.006	Hbb	82Ser	^{1}H
792	39.96	0.022	Cb	83Phe	^{13}C

793	122.57	0.034	N	83Phe	^{15}N
794	8.46	0.008	H	83Phe	^{1}H
795	62.82	0.053	Ca	83Phe	^{13}C
796	2.93	0.005	Hba	83Phe	^{1}H
797	3.14	0.011	Hbb	83Phe	^{1}H
798	3.93	0.007	Ha	83Phe	^{1}H
799	175.54	0.000	C	83Phe	^{13}C
800	117.92	0.054	N	84Glu	^{15}N
801	8.58	0.006	H	84Glu	^{1}H
802	59.32	0.051	Ca	84Glu	^{13}C
803	36.82	0.061	Cg	84Glu	^{13}C
804	1.95	0.007	Hga	84Glu	^{1}H
805	2.01	0.006	Hgb	84Glu	^{1}H
806	1.72	0.007	Hbb	84Glu	^{1}H
807	1.53	0.005	Hba	84Glu	^{1}H
808	3.47	0.005	Ha	84Glu	^{1}H
809	30.04	0.075	Cb	84Glu	^{13}C
810	177.76	0.000	C	84Glu	^{13}C
811	7.46	0.005	He2b	85Gln	^{1}H
812	110.84	0.225	Ne2	85Gln	^{15}N
813	6.78	0.006	He2a	85Gln	^{1}H
814	116.90	0.071	N	85Gln	^{15}N
815	7.82	0.006	H	85Gln	^{1}H
816	59.11	0.059	Ca	85Gln	^{13}C
817	28.64	0.052	Cb	85Gln	^{13}C
818	34.20	0.051	Cg	85Gln	^{13}C
819	3.86	0.007	Ha	85Gln	^{1}H
820	2.56	0.005	Hgb	85Gln	^{1}H
821	2.30	0.003	Hga	85Gln	^{1}H
822	2.17	0.006	Hbb	85Gln	^{1}H
823	2.01	0.012	Hba	85Gln	^{1}H
824	179.12	0.000	C	85Gln	^{13}C
825	180.07	0.001	Cd	85Gln	^{13}C
826	118.35	0.050	N	86Val	^{15}N
827	7.66	0.008	H	86Val	^{1}H
828	66.67	0.081	Ca	86Val	^{13}C
829	-0.42	0.005	Hga*	86Val	^{1}H
830	20.60	0.063	Cga	86Val	^{13}C
831	0.90	0.007	Hgb*	86Val	^{1}H
832	3.22	0.007	Ha	86Val	^{1}H
833	1.56	0.011	Hb	86Val	^{1}H
834	23.07	0.053	Cgb	86Val	^{13}C
835	31.72	0.021	Cb	86Val	^{13}C
836	177.65	0.000	C	86Val	^{13}C
837	116.93	0.076	N	87Val	^{15}N

838	7.92	0.008	H	87Val	^{1}H
839	66.67	0.078	Ca	87Val	^{13}C
840	21.30	0.023	Cga	87Val	^{13}C
841	0.11	0.008	Hga*	87Val	^{1}H
842	0.21	0.006	Hgb*	87Val	^{1}H
843	1.36	0.004	Hb	87Val	^{1}H
844	3.15	0.006	Ha	87Val	^{1}H
845	21.83	0.024	Cgb	87Val	^{13}C
846	31.14	0.031	Cb	87Val	^{13}C
847	176.86	0.000	C	87Val	^{13}C
848	116.21	0.039	N	88Asn	^{15}N
849	8.23	0.005	H	88Asn	^{1}H
850	7.50	0.008	Hd2b	88Asn	^{1}H
851	111.25	0.227	Nd2	88Asn	^{15}N
852	6.93	0.006	Hd2a	88Asn	^{1}H
853	37.56	0.057	Cb	88Asn	^{13}C
854	55.19	0.051	Ca	88Asn	^{13}C
855	4.32	0.004	Ha	88Asn	^{1}H
856	2.84	0.005	Hbb	88Asn	^{1}H
857	2.57	0.004	Hba	88Asn	^{1}H
858	178.46	0.000	C	88Asn	^{13}C
859	176.52	0.007	Cg	88Asn	^{13}C
860	118.59	0.102	N	89Glu	^{15}N
861	7.16	0.008	H	89Glu	^{1}H
862	58.46	0.093	Ca	89Glu	^{13}C
863	36.10	0.065	Cg	89Glu	^{13}C
864	4.03	0.008	Ha	89Glu	^{1}H
865	2.41	0.004	Hgb	89Glu	^{1}H
866	2.05	0.016	Hga	89Glu	^{1}H
867	177.34	0.000	C	89Glu	^{13}C
868	29.97	0.035	Cb	89Glu	^{13}C
869	1.99	0.011	Hbb	89Glu	^{1}H
870	1.74	0.012	Hba	89Glu	^{1}H
871	123.83	0.073	N	90Leu	^{15}N
872	8.10	0.011	H	90Leu	^{1}H
873	57.82	0.057	Ca	90Leu	^{13}C
874	0.25	0.010	Hda*	90Leu	^{1}H
875	25.71	0.015	Cda	90Leu	^{13}C
876	3.66	0.004	Ha	90Leu	^{1}H
877	0.34	0.006	Hdb*	90Leu	^{1}H
878	42.35	0.071	Cb	90Leu	^{13}C
879	1.51	0.011	Hbb	90Leu	^{1}H
880	0.97	0.006	Hba	90Leu	^{1}H
881	-0.02	0.007	Hg	90Leu	^{1}H
882	22.54	0.021	Cdb	90Leu	^{13}C

883	26.80	0.083	Cg	90Leu	^{13}C
884	175.77	0.000	C	90Leu	^{13}C
885	108.65	0.022	N	91Phe	^{15}N
886	6.87	0.005	H	91Phe	^{1}H
887	38.62	0.024	Cb	91Phe	^{13}C
888	56.74	0.046	Ca	91Phe	^{13}C
889	177.36	0.000	C	91Phe	^{13}C
890	5.15	0.004	Ha	91Phe	^{1}H
891	3.25	0.009	Hbb	91Phe	^{1}H
892	2.49	0.008	Hba	91Phe	^{1}H
893	123.52	0.032	N	92Arg	^{15}N
894	7.32	0.008	H	92Arg	^{1}H
895	29.89	0.015	Cb	92Arg	^{13}C
896	60.21	0.077	Ca	92Arg	^{13}C
897	3.89	0.004	Ha	92Arg	^{1}H
898	26.03	0.030	Cg	92Arg	^{13}C
899	1.53	0.006	Hga	92Arg	^{1}H
900	1.53	0.006	Hgb	92Arg	^{1}H
901	43.37	0.031	Cd	92Arg	^{13}C
902	3.17	0.008	Hda	92Arg	^{1}H
903	3.18	0.012	Hdb	92Arg	^{1}H
904	177.64	0.000	C	92Arg	^{13}C
905	2.16	0.005	Hbb	92Arg	^{1}H
906	1.70	0.007	Hba	92Arg	^{1}H
907	113.43	0.057	N	93Asp	^{15}N
908	8.42	0.006	H	93Asp	^{1}H
909	40.29	0.069	Cb	93Asp	^{13}C
910	52.75	0.029	Ca	93Asp	^{13}C
911	2.80	0.004	Hbb	93Asp	^{1}H
912	2.59	0.005	Hba	93Asp	^{1}H
913	4.78	0.005	Ha	93Asp	^{1}H
914	175.71	0.000	C	93Asp	^{13}C
915	107.77	0.048	N	94Gly	^{15}N
916	7.22	0.005	H	94Gly	^{1}H
917	44.90	0.084	Ca	94Gly	^{13}C
918	3.69	0.005	Haa	94Gly	^{1}H
919	4.35	0.005	Hab	94Gly	^{1}H
920	170.50	0.000	C	94Gly	^{13}C
921	116.92	0.078	N	95Val	^{15}N
922	7.79	0.005	H	95Val	^{1}H
923	61.71	0.061	Ca	95Val	^{13}C
924	3.63	0.007	Ha	95Val	^{1}H
925	20.85	0.075	Cgb	95Val	^{13}C
926	0.40	0.005	Hgb*	95Val	^{1}H
927	0.16	0.012	Hb	95Val	^{1}H

928	-0.55	0.006	Hga*	95Val	^{1}H
929	23.05	0.032	Cga	95Val	^{13}C
930	34.27	0.077	Cb	95Val	^{13}C
931	174.18	0.000	C	95Val	^{13}C
932	115.06	0.052	N	96Asn	^{15}N
933	6.26	0.006	H	96Asn	^{1}H
934	52.34	0.110	Ca	96Asn	^{13}C
935	174.67	0.000	C	96Asn	^{13}C
936	39.77	0.041	Cb	96Asn	^{13}C
937	3.01	0.003	Hba	96Asn	^{1}H
938	3.26	0.006	Hbb	96Asn	^{1}H
939	3.97	0.007	Ha	96Asn	^{1}H
940	7.45	0.011	Hd2b	96Asn	^{1}H
941	7.31	0.014	Hd2a	96Asn	^{1}H
942	114.85	0.017	Nd2	96Asn	^{15}N
943	118.32	0.077	N	97Trp	^{15}N
944	8.21	0.021	H	97Trp	^{1}H
945	10.37	0.009	He1	97Trp	^{1}H
946	128.94	0.047	Ne1	97Trp	^{15}N
947	60.65	0.038	Ca	97Trp	^{13}C
948	31.88	0.088	Cb	97Trp	^{13}C
949	178.52	0.000	C	97Trp	^{13}C
950	3.82	0.019	Ha	97Trp	^{1}H
951	3.83	0.010	Hba	97Trp	^{1}H
952	3.83	0.015	Hbb	97Trp	^{1}H
953	8.66	0.013	H	98Gly	^{1}H
954	107.03	0.050	N	98Gly	^{15}N
955	47.74	0.076	Ca	98Gly	^{13}C
956	176.75	0.000	C	98Gly	^{13}C
957	3.87	0.012	Haa	98Gly	^{1}H
958	4.47	0.009	Hab	98Gly	^{1}H
959	120.21	0.044	N	99Arg	^{15}N
960	8.46	0.009	H	99Arg	^{1}H
961	60.01	0.081	Ca	99Arg	^{13}C
962	177.86	0.000	C	99Arg	^{13}C
963	3.77	0.008	Ha	99Arg	^{1}H
964	43.76	0.041	Cd	99Arg	^{13}C
965	3.16	0.020	Hdb	99Arg	^{1}H
966	2.54	0.015	Hda	99Arg	^{1}H
967	1.39	0.003	Hba	99Arg	^{1}H
968	31.62	0.000	Cb	99Arg	^{13}C
969	1.51	0.000	Hbb	99Arg	^{1}H
970	120.63	0.040	N	100Ile	^{15}N
971	7.88	0.008	H	100Ile	^{1}H
972	67.56	0.071	Ca	100Ile	^{13}C

973	14.33	0.031	Cd1	100Ile	^{13}C
974	0.35	0.004	Hd1*	100Ile	^{1}H
975	0.73	0.008	Hg1a	100Ile	^{1}H
976	1.74	0.007	Hg1b	100Ile	^{1}H
977	1.25	0.006	Hg2*	100Ile	^{1}H
978	17.60	0.056	Cg2	100Ile	^{13}C
979	2.42	0.006	Hb	100Ile	^{1}H
980	3.96	0.011	Ha	100Ile	^{1}H
981	31.23	0.096	Cg1	100Ile	^{13}C
982	37.61	0.048	Cb	100Ile	^{13}C
983	177.97	0.000	C	100Ile	^{13}C
984	122.80	0.028	N	101Val	^{15}N
985	8.47	0.008	H	101Val	^{1}H
986	68.27	0.060	Ca	101Val	^{13}C
987	3.43	0.006	Ha	101Val	^{1}H
988	21.46	0.079	Cga	101Val	^{13}C
989	0.77	0.005	Hga*	101Val	^{1}H
990	1.05	0.005	Hgb*	101Val	^{1}H
991	23.43	0.056	Cgb	101Val	^{13}C
992	2.46	0.009	Hb	101Val	^{1}H
993	178.38	0.000	C	101Val	^{13}C
994	31.35	0.020	Cb	101Val	^{13}C
995	123.57	0.024	N	102Ala	^{15}N
996	8.16	0.003	H	102Ala	^{1}H
997	55.18	0.056	Ca	102Ala	^{13}C
998	1.27	0.008	Hb*	102Ala	^{1}H
999	2.86	0.007	Ha	102Ala	^{1}H
1000	19.38	0.066	Cb	102Ala	^{13}C
1001	177.55	0.000	C	102Ala	^{13}C
1002	119.59	0.050	N	103Phe	^{15}N
1003	8.21	0.018	H	103Phe	^{1}H
1004	61.02	0.080	Ca	103Phe	^{13}C
1005	176.40	0.000	C	103Phe	^{13}C
1006	39.03	0.031	Cb	103Phe	^{13}C
1007	3.40	0.006	Hba	103Phe	^{1}H
1008	3.91	0.006	Hbb	103Phe	^{1}H
1009	3.93	0.012	Ha	103Phe	^{1}H
1010	8.27	0.007	H	104Phe	^{1}H
1011	119.76	0.028	N	104Phe	^{15}N
1012	63.02	0.082	Ca	104Phe	^{13}C
1013	37.28	0.042	Cb	104Phe	^{13}C
1014	177.64	0.000	C	104Phe	^{13}C
1015	3.39	0.007	Ha	104Phe	^{1}H
1016	1.41	0.008	Hba	104Phe	^{1}H
1017	2.53	0.007	Hbb	104Phe	^{1}H

1018	114.44	0.054	N	105Ser	^{15}N
1019	8.56	0.009	H	105Ser	^{1}H
1020	63.51	0.094	Ca	105Ser	^{13}C
1021	177.36	0.000	C	105Ser	^{13}C
1022	4.45	0.017	Ha	105Ser	^{1}H
1023	3.78	0.009	Hba	105Ser	^{1}H
1024	63.54	0.032	Cb	105Ser	^{13}C
1025	3.89	0.008	Hbb	105Ser	^{1}H
1026	9.08	0.010	H	106Phe	^{1}H
1027	123.95	0.059	N	106Phe	^{15}N
1028	60.77	0.073	Ca	106Phe	^{13}C
1029	39.25	0.068	Cb	106Phe	^{13}C
1030	175.91	0.000	C	106Phe	^{13}C
1031	4.24	0.002	Ha	106Phe	^{1}H
1032	3.20	0.010	Hbb	106Phe	^{1}H
1033	3.05	0.007	Hba	106Phe	^{1}H
1034	106.60	0.032	N	107Gly	^{15}N
1035	8.07	0.006	H	107Gly	^{1}H
1036	48.12	0.094	Ca	107Gly	^{13}C
1037	3.19	0.013	Haa	107Gly	^{1}H
1038	4.20	0.006	Hab	107Gly	^{1}H
1039	176.34	0.000	C	107Gly	^{13}C
1040	107.14	0.030	N	108Gly	^{15}N
1041	8.92	0.007	H	108Gly	^{1}H
1042	47.98	0.090	Ca	108Gly	^{13}C
1043	174.28	0.000	C	108Gly	^{13}C
1044	3.86	0.013	Hab	108Gly	^{1}H
1045	3.73	0.004	Haa	108Gly	^{1}H
1046	124.54	0.058	N	109Ala	^{15}N
1047	8.20	0.006	H	109Ala	^{1}H
1048	55.01	0.087	Ca	109Ala	^{13}C
1049	18.53	0.063	Cb	109Ala	^{13}C
1050	1.42	0.004	Hb*	109Ala	^{1}H
1051	4.18	0.003	Ha	109Ala	^{1}H
1052	180.62	0.000	C	109Ala	^{13}C
1053	118.74	0.065	N	110Leu	^{15}N
1054	8.37	0.008	H	110Leu	^{1}H
1055	42.14	0.148	Cb	110Leu	^{13}C
1056	57.50	0.061	Ca	110Leu	^{13}C
1057	24.56	0.025	Cda	110Leu	^{13}C
1058	0.71	0.012	Hda*	110Leu	^{1}H
1059	0.76	0.015	Hdb*	110Leu	^{1}H
1060	3.83	0.005	Ha	110Leu	^{1}H
1061	1.38	0.007	Hg	110Leu	^{1}H
1062	1.31	0.006	Hba	110Leu	^{1}H

1063	1.45	0.010	Hbb	110Leu	^{1}H
1064	26.48	0.026	Cg	110Leu	^{13}C
1065	179.52	0.000	C	110Leu	^{13}C
1066	24.56	0.018	Cda	110Leu	^{13}C
1067	64.64	0.052	Ca	111Cys	^{13}C
1068	8.23	0.008	H	111Cys	^{1}H
1069	119.29	0.031	N	111Cys	^{15}N
1070	26.77	0.017	Cb	111Cys	^{13}C
1071	175.62	0.000	C	111Cys	^{13}C
1072	3.77	0.009	Ha	111Cys	^{1}H
1073	2.79	0.006	Hba	111Cys	^{1}H
1074	3.13	0.005	Hbb	111Cys	^{1}H
1075	118.07	0.042	N	112Val	^{15}N
1076	8.00	0.006	H	112Val	^{1}H
1077	67.34	0.063	Ca	112Val	^{13}C
1078	3.40	0.008	Ha	112Val	^{1}H
1079	32.02	0.064	Cb	112Val	^{13}C
1080	2.13	0.006	Hb	112Val	^{1}H
1081	22.44	0.023	Cgb	112Val	^{13}C
1082	1.05	0.008	Hgb*	112Val	^{1}H
1083	22.99	0.029	Cga	112Val	^{13}C
1084	1.03	0.007	Hga*	112Val	^{1}H
1085	177.55	0.000	C	112Val	^{13}C
1086	119.38	0.088	N	113Glu	^{15}N
1087	8.22	0.006	H	113Glu	^{1}H
1088	59.52	0.061	Ca	113Glu	^{13}C
1089	178.34	0.000	C	113Glu	^{13}C
1090	4.13	0.007	Ha	113Glu	^{1}H
1091	36.40	0.032	Cg	113Glu	^{13}C
1092	2.24	0.008	Hga	113Glu	^{1}H
1093	2.40	0.013	Hgb	113Glu	^{1}H
1094	29.70	0.034	Cb	113Glu	^{13}C
1095	1.94	0.007	Hba	113Glu	^{1}H
1096	2.11	0.010	Hbb	113Glu	^{1}H
1097	113.85	0.034	N	114Ser	^{15}N
1098	7.84	0.008	H	114Ser	^{1}H
1099	63.69	0.016	Ca	114Ser	^{13}C
1100	62.04	0.039	Cb	114Ser	^{13}C
1101	175.56	0.000	C	114Ser	^{13}C
1102	3.84	0.009	Ha	114Ser	^{1}H
1103	3.67	0.010	Hba	114Ser	^{1}H
1104	3.83	0.013	Hbb	114Ser	^{1}H
1105	121.11	0.082	N	115Val	^{15}N
1106	7.26	0.009	H	115Val	^{1}H
1107	66.82	0.087	Ca	115Val	^{13}C

1108	31.52	0.065	Cb	115Val	^{13}C
1109	1.61	0.007	Hb	115Val	^{1}H
1110	21.32	0.091	Cga	115Val	^{13}C
1111	0.01	0.004	Hga*	115Val	^{1}H
1112	0.22	0.004	Hgb*	115Val	^{1}H
1113	22.97	0.036	Cgb	115Val	^{13}C
1114	175.58	0.000	C	115Val	^{13}C
1115	3.17	0.005	Ha	115Val	^{1}H
1116	122.08	0.057	N	116Asp	^{15}N
1117	8.13	0.010	H	116Asp	^{1}H
1118	57.52	0.055	Ca	116Asp	^{13}C
1119	41.44	0.052	Cb	116Asp	^{13}C
1120	4.39	0.003	Ha	116Asp	^{1}H
1121	2.93	0.006	Hbb	116Asp	^{1}H
1122	2.73	0.003	Hba	116Asp	^{1}H
1123	177.94	0.000	C	116Asp	^{13}C
1124	115.68	0.057	N	117Lys	^{15}N
1125	7.78	0.006	H	117Lys	^{1}H
1126	55.12	0.036	Ca	117Lys	^{13}C
1127	31.74	0.059	Cb	117Lys	^{13}C
1128	25.63	0.023	Cg	117Lys	^{13}C
1129	1.60	0.009	Hgb	117Lys	^{1}H
1130	1.48	0.011	Hga	117Lys	^{1}H
1131	4.46	0.009	Ha	117Lys	^{1}H
1132	1.68	0.007	Hdb	117Lys	^{1}H
1133	28.88	0.013	Cd	117Lys	^{13}C
1134	1.61	0.007	Hda	117Lys	^{1}H
1135	42.83	0.062	Ce	117Lys	^{13}C
1136	3.01	0.008	Hea	117Lys	^{1}H
1137	3.01	0.008	Heb	117Lys	^{1}H
1138	175.15	0.000	C	117Lys	^{13}C
1139	2.21	0.007	Hbb	117Lys	^{1}H
1140	1.68	0.012	Hba	117Lys	^{1}H
1141	113.47	0.039	N	118Glu	^{15}N
1142	7.85	0.005	H	118Glu	^{1}H
1143	58.05	0.080	Ca	118Glu	^{13}C
1144	26.58	0.039	Cb	118Glu	^{13}C
1145	2.36	0.011	Hbb	118Glu	^{1}H
1146	2.27	0.015	Hba	118Glu	^{1}H
1147	2.24	0.006	Hgb	118Glu	^{1}H
1148	2.24	0.006	Hga	118Glu	^{1}H
1149	4.02	0.005	Ha	118Glu	^{1}H
1150	37.15	0.076	Cg	118Glu	^{13}C
1151	176.10	0.000	C	118Glu	^{13}C
1152	119.28	0.047	N	119Met	^{15}N

1153	8.73	0.006	H	119Met	^{1}H
1154	33.22	0.052	Cb	119Met	^{13}C
1155	53.90	0.080	Ca	119Met	^{13}C
1156	32.26	0.023	Cg	119Met	^{13}C
1157	4.84	0.008	Ha	119Met	^{1}H
1158	176.79	0.000	C	119Met	^{13}C
1159	1.97	0.008	Hbb	119Met	^{1}H
1160	1.89	0.004	Hba	119Met	^{1}H
1161	2.56	0.005	Hgb	119Met	^{1}H
1162	2.41	0.011	Hga	119Met	^{1}H
1163	17.63	0.023	Ce	119Met	^{13}C
1164	2.11	0.007	He*	119Met	^{1}H
1165	111.27	0.235	Ne2	120Gln	^{15}N
1166	7.59	0.011	He2b	120Gln	^{1}H
1167	7.06	0.008	He2a	120Gln	^{1}H
1168	119.39	0.056	N	120Gln	^{15}N
1169	8.92	0.008	H	120Gln	^{1}H
1170	27.74	0.060	Cb	120Gln	^{13}C
1171	60.12	0.083	Ca	120Gln	^{13}C
1172	34.07	0.058	Cg	120Gln	^{13}C
1173	3.46	0.010	Ha	120Gln	^{1}H
1174	2.01	0.004	Hba	120Gln	^{1}H
1175	2.01	0.005	Hbb	120Gln	^{1}H
1176	2.37	0.003	Hga	120Gln	^{1}H
1177	2.54	0.004	Hgb	120Gln	^{1}H
1178	177.22	0.000	C	120Gln	^{13}C
1179	180.02	0.011	Cd	120Gln	^{13}C
1180	118.29	0.065	N	121Val	^{15}N
1181	7.96	0.008	H	121Val	^{1}H
1182	63.57	0.078	Ca	121Val	^{13}C
1183	19.92	0.071	Cga	121Val	^{13}C
1184	0.91	0.014	Hga*	121Val	^{1}H
1185	2.21	0.003	Hb	121Val	^{1}H
1186	0.94	0.012	Hgb*	121Val	^{1}H
1187	20.32	0.171	Cga	121Val	^{13}C
1188	4.06	0.004	Ha	121Val	^{1}H
1189	31.51	0.076	Cb	121Val	^{13}C
1190	177.44	0.000	C	121Val	^{13}C
1191	8.00	0.012	H	122Leu	^{1}H
1192	116.05	0.072	N	122Leu	^{15}N
1193	55.92	0.044	Ca	122Leu	^{13}C
1194	23.79	0.031	Cdb	122Leu	^{13}C
1195	0.98	0.009	Hdb*	122Leu	^{1}H
1196	26.22	0.017	Cda	122Leu	^{13}C
1197	0.86	0.011	Hda*	122Leu	^{1}H

1198	177.91	0.000	C	122Leu	^{13}C
1199	3.89	0.010	Ha	122Leu	^{1}H
1200	1.96	0.006	Hbb	122Leu	^{1}H
1201	41.91	0.018	Cb	122Leu	^{13}C
1202	1.25	0.007	Hba	122Leu	^{1}H
1203	1.69	0.008	Hg	122Leu	^{1}H
1204	119.99	0.056	N	123Val	^{15}N
1205	7.43	0.009	H	123Val	^{1}H
1206	68.20	0.072	Ca	123Val	^{13}C
1207	0.44	0.010	Hgb*	123Val	^{1}H
1208	0.42	0.008	Hga*	123Val	^{1}H
1209	23.20	0.417	Cga	123Val	^{13}C
1210	21.93	0.026	Cgb	123Val	^{13}C
1211	1.78	0.007	Hb	123Val	^{1}H
1212	3.34	0.005	Ha	123Val	^{1}H
1213	176.88	0.000	C	123Val	^{13}C
1214	31.13	0.013	Cb	123Val	^{13}C
1215	110.53	0.037	N	124Ser	^{15}N
1216	8.58	0.005	H	124Ser	^{1}H
1217	61.20	0.063	Ca	124Ser	^{13}C
1218	62.47	0.032	Cb	124Ser	^{13}C
1219	177.36	0.000	C	124Ser	^{13}C
1220	3.76	0.008	Hba	124Ser	^{1}H
1221	3.82	0.012	Hbb	124Ser	^{1}H
1222	3.92	0.013	Ha	124Ser	^{1}H
1223	121.90	0.058	N	125Arg	^{15}N
1224	6.54	0.008	H	125Arg	^{1}H
1225	58.54	0.043	Ca	125Arg	^{13}C
1226	175.05	0.000	C	125Arg	^{13}C
1227	3.60	0.005	Ha	125Arg	^{1}H
1228	29.28	0.067	Cb	125Arg	^{13}C
1229	1.03	0.007	Hba	125Arg	^{1}H
1230	1.13	0.006	Hbb	125Arg	^{1}H
1231	44.04	0.022	Cd	125Arg	^{13}C
1232	3.02	0.005	Hda	125Arg	^{1}H
1233	3.24	0.004	Hdb	125Arg	^{1}H
1234	27.10	0.016	Cg	125Arg	^{13}C
1235	1.54	0.008	Hga	125Arg	^{1}H
1236	1.73	0.004	Hgb	125Arg	^{1}H
1237	118.19	0.060	N	126Ile	^{15}N
1238	8.14	0.008	H	126Ile	^{1}H
1239	65.01	0.098	Ca	126Ile	^{13}C
1240	15.03	0.025	Cd1	126Ile	^{13}C
1241	0.56	0.006	Hd1*	126Ile	^{1}H
1242	0.67	0.009	Hg2*	126Ile	^{1}H

1243	18.16	0.046	Cg2	126Ile	^{13}C
1244	2.75	0.006	Ha	126Ile	^{1}H
1245	1.53	0.010	Hb	126Ile	^{1}H
1246	38.04	0.063	Cb	126Ile	^{13}C
1247	0.62	0.032	Hg1a	126Ile	^{1}H
1248	30.66	0.102	Cg1	126Ile	^{13}C
1249	1.33	0.010	Hg1b	126Ile	^{1}H
1250	177.94	0.000	C	126Ile	^{13}C
1251	118.58	0.078	N	127Ala	^{15}N
1252	7.37	0.007	H	127Ala	^{1}H
1253	18.76	0.067	Cb	127Ala	^{13}C
1254	55.14	0.042	Ca	127Ala	^{13}C
1255	1.49	0.009	Hb*	127Ala	^{1}H
1256	3.85	0.008	Ha	127Ala	^{1}H
1257	179.17	0.000	C	127Ala	^{13}C
1258	120.59	0.065	N	128Ala	^{15}N
1259	7.12	0.007	H	128Ala	^{1}H
1260	55.22	0.061	Ca	128Ala	^{13}C
1261	17.85	0.071	Cb	128Ala	^{13}C
1262	1.51	0.005	Hb*	128Ala	^{1}H
1263	4.18	0.004	Ha	128Ala	^{1}H
1264	180.82	0.000	C	128Ala	^{13}C
1265	8.93	0.007	H	129Trp	^{1}H
1266	121.71	0.047	N	129Trp	^{15}N
1267	127.95	0.045	Ne1	129Trp	^{15}N
1268	9.91	0.006	He1	129Trp	^{1}H
1269	58.00	0.007	Ca	129Trp	^{13}C
1270	28.59	0.012	Cb	129Trp	^{13}C
1271	181.45	0.000	C	129Trp	^{13}C
1272	4.74	0.006	Ha	129Trp	^{1}H
1273	2.68	0.010	Hba	129Trp	^{1}H
1274	2.85	0.008	Hbb	129Trp	^{1}H
1275	117.93	0.041	N	130Met	^{15}N
1276	9.05	0.006	H	130Met	^{1}H
1277	60.29	0.071	Ca	130Met	^{13}C
1278	177.76	0.000	C	130Met	^{13}C
1279	4.38	0.007	Ha	130Met	^{1}H
1280	2.34	0.006	Hbb	130Met	^{1}H
1281	2.02	0.017	Hba	130Met	^{1}H
1282	35.40	0.039	Cb	130Met	^{13}C
1283	33.07	0.030	Cg	130Met	^{13}C
1284	2.09	0.003	Hga	130Met	^{1}H
1285	2.89	0.006	Hgb	130Met	^{1}H
1286	16.55	0.016	Ce	130Met	^{13}C
1287	1.41	0.006	He*	130Met	^{1}H

1288	121.76	0.052	N	131Ala	^{15}N
1289	8.56	0.005	H	131Ala	^{1}H
1290	17.70	0.057	Cb	131Ala	^{13}C
1291	56.36	0.070	Ca	131Ala	^{13}C
1292	1.59	0.005	Hb*	131Ala	^{1}H
1293	4.04	0.006	Ha	131Ala	^{1}H
1294	179.02	0.000	C	131Ala	^{13}C
1295	115.13	0.057	N	132Thr	^{15}N
1296	8.76	0.009	H	132Thr	^{1}H
1297	67.29	0.097	Ca	132Thr	^{13}C
1298	69.25	0.149	Cb	132Thr	^{13}C
1299	4.41	0.007	Hb	132Thr	^{1}H
1300	1.43	0.004	Hg2*	132Thr	^{1}H
1301	4.08	0.005	Ha	132Thr	^{1}H
1302	22.26	0.035	Cg2	132Thr	^{13}C
1303	176.16	0.000	C	132Thr	^{13}C
1304	122.94	0.029	N	133Tyr	^{15}N
1305	8.29	0.008	H	133Tyr	^{1}H
1306	63.81	0.065	Ca	133Tyr	^{13}C
1307	39.22	0.032	Cb	133Tyr	^{13}C
1308	179.17	0.000	C	133Tyr	^{13}C
1309	4.18	0.006	Ha	133Tyr	^{1}H
1310	3.57	0.006	Hbb	133Tyr	^{1}H
1311	3.27	0.008	Hba	133Tyr	^{1}H
1312	8.72	0.009	H	134Leu	^{1}H
1313	119.45	0.053	N	134Leu	^{15}N
1314	59.29	0.076	Ca	134Leu	^{13}C
1315	42.79	0.088	Cb	134Leu	^{13}C
1316	1.17	0.008	Hdb*	134Leu	^{1}H
1317	26.50	0.022	Cdb	134Leu	^{13}C
1318	26.63	0.042	Cda	134Leu	^{13}C
1319	1.15	0.008	Hda*	134Leu	^{1}H
1320	1.83	0.009	Hg	134Leu	^{1}H
1321	1.89	0.017	Hba	134Leu	^{1}H
1322	2.01	0.013	Hbb	134Leu	^{1}H
1323	3.86	0.009	Ha	134Leu	^{1}H
1324	28.28	0.013	Cg	134Leu	^{13}C
1325	177.70	0.000	C	134Leu	^{13}C
1326	112.35	0.025	Nd2	135Asn	^{15}N
1327	7.03	0.005	Hd2a	135Asn	^{1}H
1328	116.57	0.044	N	135Asn	^{15}N
1329	8.50	0.005	H	135Asn	^{1}H
1330	56.67	0.049	Ca	135Asn	^{13}C
1331	39.25	0.032	Cb	135Asn	^{13}C
1332	3.02	0.004	Hbb	135Asn	^{1}H

1333	2.89	0.008	Hba	135Asn	^{1}H
1334	4.35	0.011	Ha	135Asn	^{1}H
1335	176.53	0.000	C	135Asn	^{13}C
1336	7.70	0.010	Hd2b	135Asn	^{1}H
1337	115.54	0.049	N	136Asp	^{15}N
1338	8.57	0.006	H	136Asp	^{1}H
1339	40.84	0.082	Cb	136Asp	^{13}C
1340	56.38	0.063	Ca	136Asp	^{13}C
1341	4.36	0.008	Ha	136Asp	^{1}H
1342	2.39	0.006	Hbb	136Asp	^{1}H
1343	1.95	0.008	Hba	136Asp	^{1}H
1344	177.39	0.000	C	136Asp	^{13}C
1345	113.43	0.066	N	137His	^{15}N
1346	7.96	0.007	H	137His	^{1}H
1347	56.63	0.051	Ca	137His	^{13}C
1348	30.91	0.043	Cb	137His	^{13}C
1349	175.88	0.000	C	137His	^{13}C
1350	4.58	0.005	Ha	137His	^{1}H
1351	1.85	0.004	Hba	137His	^{1}H
1352	2.39	0.006	Hbb	137His	^{1}H
1353	117.77	0.034	N	138Leu	^{15}N
1354	7.23	0.006	H	138Leu	^{1}H
1355	55.94	0.040	Ca	138Leu	^{13}C
1356	22.08	0.060	Cda	138Leu	^{13}C
1357	-0.31	0.004	Hda*	138Leu	^{1}H
1358	0.15	0.007	Hdb*	138Leu	^{1}H
1359	4.45	0.009	Ha	138Leu	^{1}H
1360	24.75	0.049	Cdb	138Leu	^{13}C
1361	26.18	0.038	Cg	138Leu	^{13}C
1362	1.09	0.006	Hg	138Leu	^{1}H
1363	43.16	0.026	Cb	138Leu	^{13}C
1364	2.04	0.007	Hbb	138Leu	^{1}H
1365	1.08	0.007	Hba	138Leu	^{1}H
1366	177.10	0.000	C	138Leu	^{13}C
1367	121.78	0.055	N	139Glu	^{15}N
1368	8.69	0.008	H	139Glu	^{1}H
1369	60.97	0.036	Ca	139Glu	^{13}C
1370	4.28	0.013	Ha	139Glu	^{1}H
1371	36.20	0.030	Cg	139Glu	^{13}C
1372	2.39	0.010	Hgb	139Glu	^{1}H
1373	2.34	0.004	Hga	139Glu	^{1}H
1374	27.73	0.024	Cb	139Glu	^{13}C
1375	2.28	0.003	Hbb	139Glu	^{1}H
1376	2.21	0.003	Hba	139Glu	^{1}H
1377	66.51	0.068	Ca	140Pro	^{13}C

1378	4.21	0.010	Ha	140Pro	^{1}H
1379	179.20	0.000	C	140Pro	^{13}C
1380	49.87	0.025	Cd	140Pro	^{13}C
1381	30.83	0.014	Cb	140Pro	^{13}C
1382	3.52	0.005	Hdb	140Pro	^{1}H
1383	3.14	0.004	Hda	140Pro	^{1}H
1384	2.27	0.004	Hbb	140Pro	^{1}H
1385	1.76	0.007	Hba	140Pro	^{1}H
1386	28.73	0.045	Cg	140Pro	^{13}C
1387	1.94	0.004	Hgb	140Pro	^{1}H
1388	1.85	0.005	Hga	140Pro	^{1}H
1389	117.99	0.062	N	141Trp	^{15}N
1390	7.13	0.008	H	141Trp	^{1}H
1391	129.64	0.064	Ne1	141Trp	^{15}N
1392	9.56	0.005	He1	141Trp	^{1}H
1393	62.30	0.069	Ca	141Trp	^{13}C
1394	3.84	0.005	Ha	141Trp	^{1}H
1395	179.33	0.000	C	141Trp	^{13}C
1396	3.14	0.007	Hbb	141Trp	^{1}H
1397	3.14	0.007	Hba	141Trp	^{1}H
1398	28.90	0.027	Cb	141Trp	^{13}C
1399	122.78	0.036	N	142Ile	^{15}N
1400	8.45	0.008	H	142Ile	^{1}H
1401	66.27	0.062	Ca	142Ile	^{13}C
1402	16.09	0.023	Cg2	142Ile	^{13}C
1403	0.38	0.007	Hg2*	142Ile	^{1}H
1404	38.41	0.046	Cb	142Ile	^{13}C
1405	1.87	0.006	Hb	142Ile	^{1}H
1406	2.87	0.005	Ha	142Ile	^{1}H
1407	13.96	0.026	Cd1	142Ile	^{13}C
1408	0.53	0.009	Hd1*	142Ile	^{1}H
1409	179.07	0.000	C	142Ile	^{13}C
1410	29.05	0.024	Cg1	142Ile	^{13}C
1411	0.50	0.018	Hg1a	142Ile	^{1}H
1412	1.57	0.005	Hg1b	142Ile	^{1}H
1413	112.75	0.209	Ne2	143Gln	^{15}N
1414	8.01	0.009	He2b	143Gln	^{1}H
1415	6.85	0.004	He2a	143Gln	^{1}H
1416	116.65	0.034	N	143Gln	^{15}N
1417	8.58	0.007	H	143Gln	^{1}H
1418	58.57	0.059	Ca	143Gln	^{13}C
1419	28.17	0.035	Cb	143Gln	^{13}C
1420	33.68	0.100	Cg	143Gln	^{13}C
1421	3.92	0.003	Ha	143Gln	^{1}H
1422	2.53	0.004	Hga	143Gln	^{1}H

1423	2.53	0.004	Hgb	143Gln	^{1}H
1424	2.12	0.006	Hbb	143Gln	^{1}H
1425	2.03	0.005	Hba	143Gln	^{1}H
1426	179.27	0.000	C	143Gln	^{13}C
1427	180.10	0.004	Cd	143Gln	^{13}C
1428	7.52	0.007	H	144Glu	^{1}H
1429	119.09	0.039	N	144Glu	^{15}N
1430	57.75	0.067	Ca	144Glu	^{13}C
1431	36.45	0.058	Cg	144Glu	^{13}C
1432	30.03	0.016	Cb	144Glu	^{13}C
1433	4.12	0.006	Ha	144Glu	^{1}H
1434	2.25	0.003	Hgb	144Glu	^{1}H
1435	2.15	0.014	Hga	144Glu	^{1}H
1436	1.85	0.006	Hbb	144Glu	^{1}H
1437	1.85	0.006	Hba	144Glu	^{1}H
1438	176.69	0.000	C	144Glu	^{13}C
1439	117.32	0.035	N	145Asn	^{15}N
1440	7.16	0.007	H	145Asn	^{1}H
1441	53.67	0.050	Ca	145Asn	^{13}C
1442	38.37	0.054	Cb	145Asn	^{13}C
1443	4.48	0.006	Ha	145Asn	^{1}H
1444	173.81	0.000	C	145Asn	^{13}C
1445	0.88	0.008	Hba	145Asn	^{1}H
1446	2.21	0.008	Hbb	145Asn	^{1}H
1447	115.77	0.004	Nd2	145Asn	^{15}N
1448	5.96	0.004	Hd2a	145Asn	^{1}H
1449	6.14	0.006	Hd2b	145Asn	^{1}H
1450	105.26	0.055	N	146Gly	^{15}N
1451	7.48	0.008	H	146Gly	^{1}H
1452	45.78	0.085	Ca	146Gly	^{13}C
1453	4.59	0.004	Hab	146Gly	^{1}H
1454	3.69	0.007	Haa	146Gly	^{1}H
1455	176.10	0.000	C	146Gly	^{13}C
1456	108.82	0.026	N	147Gly	^{15}N
1457	8.59	0.005	H	147Gly	^{1}H
1458	44.74	0.084	Ca	147Gly	^{13}C
1459	4.51	0.011	Hab	147Gly	^{1}H
1460	3.66	0.008	Haa	147Gly	^{1}H
1461	173.41	0.000	C	147Gly	^{13}C
1462	129.19	0.034	Ne1	148Trp	^{15}N
1463	10.60	0.008	He1	148Trp	^{1}H
1464	118.45	0.025	N	148Trp	^{15}N
1465	8.76	0.008	H	148Trp	^{1}H
1466	60.58	0.050	Ca	148Trp	^{13}C
1467	29.87	0.043	Cb	148Trp	^{13}C

1468	3.12	0.005	Hba	148Trp	^{1}H
1469	3.54	0.007	Hbb	148Trp	^{1}H
1470	4.48	0.011	Ha	148Trp	^{1}H
1471	178.50	0.000	C	148Trp	^{13}C
1472	113.99	0.000	Cz2	148Trp	^{13}C
1473	118.02	0.057	N	149Asp	^{15}N
1474	8.88	0.006	H	149Asp	^{1}H
1475	57.92	0.100	Ca	149Asp	^{13}C
1476	40.56	0.111	Cb	149Asp	^{13}C
1477	4.45	0.005	Ha	149Asp	^{1}H
1478	2.73	0.007	Hbb	149Asp	^{1}H
1479	2.73	0.007	Hba	149Asp	^{1}H
1480	179.70	0.000	C	149Asp	^{13}C
1481	117.89	0.084	N	150Thr	^{15}N
1482	7.63	0.007	H	150Thr	^{1}H
1483	66.72	0.094	Ca	150Thr	^{13}C
1484	68.40	0.084	Cb	150Thr	^{13}C
1485	22.54	0.028	Cg2	150Thr	^{13}C
1486	1.45	0.005	Hg2*	150Thr	^{1}H
1487	4.01	0.005	Ha	150Thr	^{1}H
1488	4.46	0.007	Hb	150Thr	^{1}H
1489	175.00	0.000	C	150Thr	^{13}C
1490	122.34	0.031	N	151Phe	^{15}N
1491	6.60	0.006	H	151Phe	^{1}H
1492	62.53	0.070	Ca	151Phe	^{13}C
1493	38.31	0.055	Cb	151Phe	^{13}C
1494	177.00	0.000	C	151Phe	^{13}C
1495	3.12	0.005	Ha	151Phe	^{1}H
1496	1.55	0.004	Hba	151Phe	^{1}H
1497	3.30	0.008	Hbb	151Phe	^{1}H
1498	118.29	0.043	N	152Val	^{15}N
1499	8.11	0.006	H	152Val	^{1}H
1500	67.03	0.103	Ca	152Val	^{13}C
1501	3.10	0.021	Ha	152Val	^{1}H
1502	21.99	0.095	Cga	152Val	^{13}C
1503	1.17	0.005	Hgb*	152Val	^{1}H
1504	0.91	0.006	Hga*	152Val	^{1}H
1505	178.49	0.000	C	152Val	^{13}C
1506	2.23	0.005	Hb	152Val	^{1}H
1507	31.87	0.023	Cb	152Val	^{13}C
1508	24.83	0.026	Cgb	152Val	^{13}C
1509	121.03	0.040	N	153Glu	^{15}N
1510	7.62	0.006	H	153Glu	^{1}H
1511	59.58	0.076	Ca	153Glu	^{13}C
1512	36.28	0.061	Cg	153Glu	^{13}C

1513	29.66	0.039	Cb	153Glu	^{13}C
1514	3.92	0.005	Ha	153Glu	^{1}H
1515	2.21	0.011	Hga	153Glu	^{1}H
1516	2.37	0.006	Hgb	153Glu	^{1}H
1517	178.37	0.000	C	153Glu	^{13}C
1518	2.18	0.013	Hba	153Glu	^{1}H
1519	2.20	0.008	Hbb	153Glu	^{1}H
1520	116.60	0.057	N	154Leu	^{15}N
1521	7.64	0.010	H	154Leu	^{1}H
1522	56.71	0.053	Ca	154Leu	^{13}C
1523	44.81	0.057	Cb	154Leu	^{13}C
1524	22.17	0.041	Cdb	154Leu	^{13}C
1525	0.79	0.004	Hda*	154Leu	^{1}H
1526	0.80	0.006	Hdb*	154Leu	^{1}H
1527	4.05	0.005	Ha	154Leu	^{1}H
1528	26.22	0.028	Cdb	154Leu	^{13}C
1529	178.63	0.000	C	154Leu	^{13}C
1530	1.64	0.009	Hbb	154Leu	^{1}H
1531	0.82	0.014	Hba	154Leu	^{1}H
1532	26.55	0.022	Cg	154Leu	^{13}C
1533	1.78	0.006	Hg	154Leu	^{1}H
1534	114.59	0.053	N	155Tyr	^{15}N
1535	8.41	0.008	H	155Tyr	^{1}H
1536	59.64	0.053	Ca	155Tyr	^{13}C
1537	176.32	0.000	C	155Tyr	^{13}C
1538	39.26	0.034	Cb	155Tyr	^{13}C
1539	1.56	0.005	Hba	155Tyr	^{1}H
1540	2.74	0.006	Hbb	155Tyr	^{1}H
1541	4.19	0.007	Ha	155Tyr	^{1}H
1542	109.18	0.069	N	156Gly	^{15}N
1543	8.29	0.009	H	156Gly	^{1}H
1544	45.57	0.089	Ca	156Gly	^{13}C
1545	4.18	0.008	Hab	156Gly	^{1}H
1546	3.79	0.011	Haa	156Gly	^{1}H
1547	173.65	0.000	C	156Gly	^{13}C
1548	113.20	0.248	Nd2	157Asn	^{15}N
1549	7.72	0.007	Hd2b	157Asn	^{1}H
1550	6.89	0.014	Hd2a	157Asn	^{1}H
1551	8.70	0.013	H	157Asn	^{1}H
1552	118.72	0.087	N	157Asn	^{15}N
1553	53.56	0.034	Ca	157Asn	^{13}C
1554	38.87	0.121	Cb	157Asn	^{13}C
1555	2.86	0.007	Hbb	157Asn	^{1}H
1556	2.78	0.005	Hba	157Asn	^{1}H
1557	4.64	0.005	Ha	157Asn	^{1}H

1558	175.37	0.000	C	157Asn	^{13}C
1559	177.35	0.032	Cg	157Asn	^{13}C
1560	7.67	0.006	Hd2b	158Asn	^{1}H
1561	113.21	0.015	Nd2	158Asn	^{15}N
1562	8.62	0.004	H	158Asn	^{1}H
1563	119.47	0.079	N	158Asn	^{15}N
1564	53.75	0.035	Ca	158Asn	^{13}C
1565	39.09	0.057	Cb	158Asn	^{13}C
1566	2.72	0.003	Hba	158Asn	^{1}H
1567	2.89	0.005	Hbb	158Asn	^{1}H
1568	4.65	0.010	Ha	158Asn	^{1}H
1569	178.08	0.000	C	158Asn	^{13}C
1570	6.92	0.004	Hd2a	158Asn	^{1}H
1571	123.43	0.119	N	159Ala	^{15}N
1572	8.12	0.018	H	159Ala	^{1}H
1573	19.15	0.073	Cb	159Ala	^{13}C
1574	53.81	0.051	Ca	159Ala	^{13}C
1575	1.36	0.004	Hb*	159Ala	^{1}H
1576	4.17	0.005	Ha	159Ala	^{1}H
1577	175.01	0.000	C	159Ala	^{13}C
1578	121.76	0.030	N	160Ala	^{15}N
1579	8.27	0.011	H	160Ala	^{1}H
1580	53.29	0.115	Ca	160Ala	^{13}C
1581	1.43	0.007	Hb*	160Ala	^{1}H
1582	4.24	0.004	Ha	160Ala	^{1}H
1583	18.92	0.074	Cb	160Ala	^{13}C
1584	177.69	0.000	C	160Ala	^{13}C
1585	122.47	0.026	N	161Ala	^{15}N
1586	8.02	0.004	H	161Ala	^{1}H
1587	53.66	0.027	Ca	161Ala	^{13}C
1588	4.18	0.004	Ha	161Ala	^{1}H
1589	1.42	0.008	Hb*	161Ala	^{1}H
1590	19.22	0.071	Cb	161Ala	^{13}C
1591	178.47	0.000	C	161Ala	^{13}C
1592	118.80	0.083	N	162Glu	^{15}N
1593	8.32	0.007	H	162Glu	^{1}H
1594	57.70	0.067	Ca	162Glu	^{13}C
1595	29.96	0.031	Cb	162Glu	^{13}C
1596	4.18	0.004	Ha	162Glu	^{1}H
1597	2.02	0.010	Hba	162Glu	^{1}H
1598	2.07	0.014	Hbb	162Glu	^{1}H
1599	2.31	0.008	Hga	162Glu	^{1}H
1600	2.31	0.008	Hgb	162Glu	^{1}H
1601	36.32	0.011	Cg	162Glu	^{13}C
1602	177.01	0.000	C	162Glu	^{13}C

1603	115.53	0.080	N	163Ser	^{15}N
1604	8.15	0.009	H	163Ser	^{1}H
1605	63.57	0.097	Cb	163Ser	^{13}C
1606	59.41	0.058	Ca	163Ser	^{13}C
1607	3.94	0.010	Hbb	163Ser	^{1}H
1608	4.38	0.006	Ha	163Ser	^{1}H
1609	3.94	0.010	Hba	163Ser	^{1}H
1610	174.86	0.000	C	163Ser	^{13}C
1611	122.19	0.020	N	164Arg	^{15}N
1612	8.05	0.008	H	164Arg	^{1}H
1613	56.75	0.022	Ca	164Arg	^{13}C
1614	30.64	0.060	Cb	164Arg	^{13}C
1615	4.32	0.010	Ha	164Arg	^{1}H
1616	3.18	0.005	Hda	164Arg	^{1}H
1617	3.18	0.005	Hdb	164Arg	^{1}H
1618	43.55	0.000	Cd	164Arg	^{13}C
1619	1.80	0.003	Hba	164Arg	^{1}H
1620	1.92	0.002	Hbb	164Arg	^{1}H
1621	1.63	0.003	Hga	164Arg	^{1}H
1622	1.69	0.003	Hgb	164Arg	^{1}H
1623	121.38	0.037	N	165Lys	^{15}N
1624	8.15	0.006	H	165Lys	^{1}H
1625	57.33	0.052	Ca	165Lys	^{13}C
1626	33.00	0.036	Cb	165Lys	^{13}C
1627	4.25	0.005	Ha	165Lys	^{1}H
1628	1.81	0.009	Hba	165Lys	^{1}H
1629	1.86	0.017	Hbb	165Lys	^{1}H
1630	25.07	0.053	Cg	165Lys	^{13}C
1631	42.32	0.058	Ce	165Lys	^{13}C
1632	2.97	0.004	Hea	165Lys	^{1}H
1633	2.97	0.004	Heb	165Lys	^{1}H
1634	1.68	0.007	Hda	165Lys	^{1}H
1635	1.68	0.007	Hdb	165Lys	^{1}H
1636	1.51	0.010	Hgb	165Lys	^{1}H
1637	1.44	0.007	Hga	165Lys	^{1}H
1638	29.32	0.003	Cd	165Lys	^{13}C
1639	109.57	0.117	N	166Gly	^{15}N
1640	8.47	0.004	H	166Gly	^{1}H
1641	45.77	0.055	Ca	166Gly	^{13}C
1642	176.21	0.000	C	166Gly	^{13}C
1643	3.98	0.011	Haa	166Gly	^{1}H
1644	3.98	0.011	Hab	166Gly	^{1}H
1645	119.82	0.030	N	167Gln	^{15}N
1646	8.16	0.005	H	167Gln	^{1}H
1647	112.10	0.224	Ne2	167Gln	^{15}N

1648	7.57	0.002	He2b	167Gln	^{1}H
1649	6.87	0.005	He2a	167Gln	^{1}H
1650	29.53	0.072	Cb	167Gln	^{13}C
1651	56.24	0.078	Ca	167Gln	^{13}C
1652	2.15	0.007	Hbb	167Gln	^{1}H
1653	2.01	0.005	Hba	167Gln	^{1}H
1654	4.32	0.004	Ha	167Gln	^{1}H
1655	33.98	0.101	Cg	167Gln	^{13}C
1656	2.36	0.011	Hgb	167Gln	^{1}H
1657	2.36	0.011	Hga	167Gln	^{1}H
1658	179.56	0.012	Cd	167Gln	^{13}C
1659	121.81	0.021	N	168Glu	^{15}N
1660	8.55	0.004	H	168Glu	^{1}H
1661	57.38	0.044	Ca	168Glu	^{13}C
1662	30.09	0.007	Cb	168Glu	^{13}C
1663	4.22	0.007	Ha	168Glu	^{1}H
1664	2.28	0.000	Hga	168Glu	^{1}H
1665	2.28	0.000	Hgb	168Glu	^{1}H
1666	2.01	0.003	Hba	168Glu	^{1}H
1667	2.01	0.003	Hbb	168Glu	^{1}H
1668	121.36	0.014	N	169Arg	^{15}N
1669	8.30	0.006	H	169Arg	^{1}H
1670	56.26	0.021	Ca	169Arg	^{13}C
1671	30.95	0.165	Cb	169Arg	^{13}C
1672	4.42	0.006	Ha	169Arg	^{1}H
1673	27.39	0.000	Cg	169Arg	^{13}C
1674	1.61	0.007	Hga	169Arg	^{1}H
1675	1.61	0.010	Hgb	169Arg	^{1}H
1676	43.54	0.052	Cd	169Arg	^{13}C
1677	3.19	0.003	Hda	169Arg	^{1}H
1678	3.19	0.003	Hdb	169Arg	^{1}H
1679	1.88	0.014	Hbb	169Arg	^{1}H
1680	1.76	0.004	Hba	169Arg	^{1}H
1681	123.13	0.027	N	170Leu	^{15}N
1682	8.19	0.008	H	170Leu	^{1}H
1683	55.39	0.065	Ca	170Leu	^{13}C
1684	42.45	0.075	Cb	170Leu	^{13}C
1685	23.59	0.039	Cda	170Leu	^{13}C
1686	25.18	0.027	Cdb	170Leu	^{13}C
1687	4.30	0.003	Ha	170Leu	^{1}H
1688	0.83	0.007	Hda*	170Leu	^{1}H
1689	0.89	0.006	Hdb*	170Leu	^{1}H
1690	1.60	0.004	Hbb	170Leu	^{1}H
1691	1.51	0.004	Hba	170Leu	^{1}H
1692	1.59	0.004	Hg	170Leu	^{1}H

1693	27.22	0.010	Cg	170Leu	^{13}C
1694	4.49	0.004	Ha	1003Cyl	^{1}H
1695	3.17	0.005	Hb1	1003Cyl	^{1}H
1696	3.08	0.014	Hb2	1003Cyl	^{1}H
1697	10.46	0.003	H3	1003Cyl	^{1}H
1698	3.73	0.004	H12	1003Cyl	^{1}H
1699	3.42	0.010	H11	1003Cyl	^{1}H
1700	8.30	0.003	H5	1003Cyl	^{1}H
1701	7.87	0.002	H7	1003Cyl	^{1}H
1702	8.43	0.010	H8	1003Cyl	^{1}H
1703	8.68	0.006	H18	1003Cyl	^{1}H
1704	8.44	0.007	H14	1003Cyl	^{1}H
1705	3.82	0.001	H21	1003Cyl	^{1}H
1706	3.74	0.000	H22	1003Cyl	^{1}H
1707	7.91	0.005	H13	1003Cyl	^{1}H
1708	8.30	0.003	H15	1003Cyl	^{1}H
1709	0.96	0.006	Hgb*	1004Val	^{1}H
1710	0.65	0.006	Hga*	1004Val	^{1}H
1711	3.44	0.008	Ha	1004Val	^{1}H
1712	1.96	0.009	Hb	1004Val	^{1}H
1713	7.55	0.007	H	1004Val	^{1}H
1714	3.71	0.005	Haa	1005Gly	^{1}H
1715	3.90	0.011	Hab	1005Gly	^{1}H
1716	8.41	0.007	H	1005Gly	^{1}H
1717	3.84	0.007	Ha	1006Arg	^{1}H
1718	3.26	0.007	Hdb	1006Arg	^{1}H
1719	1.73	0.009	Hgb	1006Arg	^{1}H
1720	3.16	0.009	Hda	1006Arg	^{1}H
1721	1.65	0.011	Hga	1006Arg	^{1}H
1722	1.73	0.010	Hba	1006Arg	^{1}H
1723	1.82	0.010	Hbb	1006Arg	^{1}H
1724	7.70	0.005	H	1006Arg	^{1}H
1725	2.39	0.012	Ha	1007Ala	^{1}H
1726	7.84	0.005	H	1007Ala	^{1}H
1727	0.53	0.005	Hb*	1007Ala	^{1}H
1728	0.29	0.006	Hda*	1008Leu	^{1}H
1729	0.79	0.006	Hdb*	1008Leu	^{1}H
1730	1.70	0.006	Hg	1008Leu	^{1}H
1731	3.46	0.009	Ha	1008Leu	^{1}H
1732	3.52	0.014	Ha	1009Ala	^{1}H
1733	1.31	0.008	Hb*	1009Ala	^{1}H
1734	7.61	0.011	H	1009Ala	^{1}H
1735	3.93	0.006	Ha	1010Ala	^{1}H
1736	7.10	0.006	H	1010Ala	^{1}H
1737	1.06	0.013	Hb*	1010Ala	^{1}H

1738	7.03	0.008	Hd*	1011Phe	^{1}H
1739	6.91	0.007	He*	1011Phe	^{1}H
1740	6.02	0.008	Hz	1011Phe	^{1}H
1741	3.67	0.014	Hab	1012Gly	^{1}H
1742	3.11	0.016	Haa	1012Gly	^{1}H
1743	8.88	0.007	H	1012Gly	^{1}H
1744	4.41	0.005	Ha	1013Asp	^{1}H
1745	2.83	0.010	Hbb	1013Asp	^{1}H
1746	2.66	0.008	Hba	1013Asp	^{1}H
1747	8.57	0.008	H	1013Asp	^{1}H
1748	2.95	0.008	Hba	1014Cys	^{1}H
1749	4.38	0.007	Ha	1014Cys	^{1}H
1750	3.70	0.011	Hbb	1014Cys	^{1}H
1751	7.85	0.004	H	1014Cys	^{1}H
1752	3.69	0.000	Ha	1015Ile	^{1}H
1753	1.35	0.008	Hg2*	1015Ile	^{1}H
1754	2.57	0.002	Hb	1015Ile	^{1}H
1755	1.85	0.001	Hg1a	1015Ile	^{1}H
1756	1.99	0.004	Hg1b	1015Ile	^{1}H
1757	0.77	0.005	Hd1*	1015Ile	^{1}H

Table 22: Bcl-x_L/$Bak_{I81F}^{i, i+11}$-XL-SO_3H NMR assignments. Residues in Bcl-x_L are numbered ignoring the loop deletion between residues 44-85; resiudes in $Bak_{I81F}^{i, i+11}$-XL-SO_3H are numbered from 1001 (being the acetyl group) to 1017; SD is standard deviation.

Residue	Ω	Ω/Ω_{max}
4Ser	0.06	0.05
5Asn	0.06	0.05
6Arg	0.03	0.02
7Glu	0.05	0.04
8Leu	0.03	0.03
9Val	0.05	0.04
10Val	0.14	0.11
11Asp	0.14	0.11
12Phe	0.08	0.07
13Leu	0.02	0.02
14Ser	0.06	0.05
15Tyr	0.08	0.07
16Lys	0.07	0.06
17Leu	0.05	0.04
18Ser	0.05	0.05
19Gln	0.04	0.03
20Lys	0.05	0.04
21Gly	0.05	0.04
22Tyr	0.05	0.05
23Ser	0.11	0.09

24Trp	0.09	0.07
25Ser	0.23	0.19
26Gln	0.05	0.05
27Phe	0.21	0.17
28Ser	0.20	0.17
29Asp	0.02	0.02
30Val	0.03	0.02
31Glu	0.07	0.06
32Glu	0.06	0.05
33Asn	0.04	0.04
34Arg	0.04	0.03
35Thr	0.03	0.02
36Glu	0.03	0.02
37Ala	0.01	0.00
39Glu	0.01	0.01
40Gly	0.01	0.01
41Thr	0.04	0.04
42Glu	0.04	0.03
43Ser	0.05	0.04
44Glu	0.03	0.03
45Ala	0.02	0.01
46Val	0.01	0.01
47Lys	0.01	0.01
48Gln	0.04	0.03
49Ala	0.04	0.03
50Leu	0.05	0.05
51Arg	0.08	0.07
52Glu	0.11	0.10
53Ala	0.13	0.11
54Gly	0.26	0.22
55Asp	0.51	0.43
56Glu	0.21	0.18
57Phe	0.13	0.11
58Glu	0.61	0.52
59Leu	0.51	0.43
60Arg	0.49	0.41
66Ser	0.56	0.47
67Asp	0.48	0.41
68Leu	1.18	1.00
69Thr	0.55	0.46
70Ser	0.23	0.19
71Gln	0.19	0.16
72Leu	0.77	0.65
74Ile	0.86	0.72
80Tyr	0.69	0.58

81Gln	0.70	0.59
82Ser	0.57	0.48
84Glu	0.40	0.34
85Gln	0.45	0.38
86Val	0.50	0.42
87Val	0.57	0.48
88Asn	0.35	0.29
89Glu	0.22	0.18
90Leu	0.78	0.66
91Phe	0.96	0.81
92Arg	0.13	0.11
93Asp	0.76	0.64
94Gly	0.42	0.35
95Val	0.91	0.77
96Asn	0.25	0.21
97Trp	0.32	0.27
98Gly	0.09	0.08
99Arg	0.39	0.33
100Ile	0.17	0.14
101Val	0.31	0.26
102Ala	0.58	0.49
103Phe	0.11	0.09
104Phe	0.26	0.22
105Ser	0.09	0.08
106Phe	0.15	0.13
107Gly	0.03	0.03
108Gly	0.13	0.11
109Ala	0.15	0.13
110Leu	0.14	0.12
111Cys	0.05	0.04
112Val	0.22	0.18
113Glu	0.05	0.04
114Ser	0.09	0.08
115Val	0.16	0.13
116Asp	0.30	0.26
117Lys	0.13	0.11
118Glu	0.19	0.16
119Met	0.19	0.16
120Gln	0.16	0.13
121Val	0.21	0.18
122Leu	0.29	0.24
123Val	0.01	0.01
124Ser	0.24	0.20
125Arg	0.35	0.29
126Ile	0.10	0.08

127Ala	0.38	0.32
128Ala	0.23	0.20
129Trp	0.14	0.12
130Met	0.09	0.07
131Ala	0.33	0.28
132Thr	0.42	0.35
133Tyr	0.68	0.58
134Leu	0.13	0.11
135Asn	0.22	0.18
136Asp	0.08	0.07
137His	0.36	0.30
138Leu	0.10	0.08
139Glu	0.04	0.03
141Trp	0.10	0.09
142Ile	0.06	0.05
143Gln	0.07	0.06
144Glu	0.14	0.12
145Asn	0.08	0.06
146Gly	0.03	0.03
147Gly	0.02	0.01
148Trp	0.04	0.04
149Asp	0.04	0.03
150Thr	0.15	0.12
151Phe	0.06	0.05
152Val	0.26	0.22
153Glu	0.20	0.17
154Leu	0.20	0.17
155Tyr	0.33	0.28
156Gly	0.19	0.16
157Asn	0.27	0.23
158Asn	0.17	0.14
159Ala	0.11	0.10
160Ala	0.13	0.11
161Ala	0.04	0.04
162Glu	0.01	0.01
163Ser	0.03	0.03
164Arg	0.06	0.05
165Lys	0.06	0.05
166Gly	0.02	0.02
167Gln	0.02	0.01
168Glu	0.04	0.03
169Arg	0.03	0.02
170Leu	0.05	0.04

Table 23: Isotope-weighted chemical shift changes (Ω) in the ^{1}H, ^{15}N HSQC spectrum between Bcl-x_L and Bcl-x_L/$Bak_{I81F}^{i, i+11}$-XL-SO_3H. Residues in Bcl-x_L are numbered ignoring the loop deletion between residues 44.