University Of Southampton
FACULTY OF MEDICINE, HEALTH AND LIFE SCIENCES
School of Biological Sciences

Probing the Interaction of Chemokines With Their Receptors, Using NMR Spectroscopy.

By

Joseph Edward Long

Thesis submitted for the degree of Doctor of Philosophy

October 2004

UNIVERSITY OF SOUTHAMPTON

Abstract

FACULTY OF MEDICINE, HEALTH AND LIFE SCIENCES
SCHOOL OF BIOLOGICAL SCIENCES

Doctor of Philosophy

PROBING THE INTERACTION OF CHEMOKINES WITH THEIR RECEPTORS, USING NMR SPECTROSCOPY.

by Joseph Edward Long

Chemokines are a part of the inflammatory response and are involved predominantly in the migration of leukocytes. They have been implicated in diseases such as asthma, dermatitis, rheumatoid arthritis, AIDS and multiple sclerosis. Their receptors are seven transmembrane helix G-coupled proteins which are expressed on the surface of leukocytes. To date over 50 chemokines and 20 receptors have been identified. The disparity between the number of chemokines and receptors is balanced by promiscuity in the ligand-receptor interactions, with most receptors recognising more than one chemokine and several chemokines binding to more than one receptor.

The family of chemokines known as the eotaxin family has three members, CCL11, CCL24 and CCL26 which all activate the receptor CCR3 only. It is hoped that the characterisation of this family's binding will give a greater insight into the complex chemokine/receptor relationship.

Recombinant CCL24 was expressed in *E.coli* and labelled with the NMR active isotope ^{15}N before being isolated in a highly purified form.

Peptides to mimic the extracellular loops of the receptor were designed and synthesised. The peptides for extracellular loops 1 and 3 were dissolved in both DMSO and the membrane mimic DPC and their structures solved by NMR spectroscopy. Loop 2 was also investigated, but failed to produce assignable NMR spectra due to reasons unknown.

^{15}N-labelled CCL24 was titrated into loops 1 and 3 suspended in DPC micelles and the resulting systems were studied by NMR spectroscopy. The results gained enabled the building of possible models of the interaction site.

To Research

Sonnets by Shakespeare

Why is my verse so barren of new pride,
So far from variation or quick change?
Why with the time do I not glance aside
To new-found methods and to compounds strange?
Why write I still all one, ever the same,
And keep invention in a noted weed,
That every word doth almost tell my name,
Showing their birth and where they did proceed?
O, know, sweet love, I always write of you,
And you and love are still my argument;
So all my best is dressing old words new,
Spending again what is already spent:
For as the sun is daily new and old,
So is my love still telling what is told.

No longer mourn for me when I am dead
Then you shall hear the surly sullen bell
Give warning to the world that I am fled
From this vile world, with vilest worms to dwell:
Nay, if you read this line, remember not
The hand that writ it; for I love you so
That I in your sweet thoughts would be forgot
If thinking on me then should make you woe.
O, if, I say, you look upon this verse
When I perhaps compounded am with clay,
Do not so much as my poor name rehearse.
But let your love even with my life decay,
Lest the wise world should look into your moan
And mock you with me after I am gone.

Acknowledgements

I would like to thank Matt Crump and Tom Ceska for the opportunity to carry out this research and for their advice along the way.

I am deeply grateful to Stuart Findlow who has enabled this work to progress and to Jörn Werner who helped make it correct.

To Big Ed Hyland who helped me sort out my jumbled thoughts, into a language almost resembling English.

To Viv Perkins and Alistair Henry for their tuition in cloning and protein expression.

To Pottage and all those who have helped from the Jordan, Lee, Gore, Wood, and Cooper groups and to those who came for a drink especially Graham, Paul, Monty and Rob.

To Neil Wells for his support through my time at Southampton and for his discussions on how to complete a PhD.

To my friends, even those who have had the indecency to get married and to the lads especially JG for allowing me to complete before him.

Lastly to the members of the hospital teams at Birmingham (Prem Mahendra) and Southampton (Andrew Duncombe) who enable the dance to go on.

Declaration

The research described in this thesis was carried out under the supervision of Dr M.Crump and Dr I.S.Findlow at the University of Southampton between November 2000 and October 2004

Abbreviations

ACP	Acyl carrier protein
ACPS	Holo-acyl carrier protein synthase
act	Benzoisochromanequinone actinorhodin
AU	Absorbance unit
t-BOC	*tert*-butoxycarbonyl
bp	Base pairs
C	Celcius
CC	Chemokine with the first two cysteines adjacent
CD	Circular dichroism
cDNA	Complementary DNA
CIP	Calf intestinal alkaline phosphatase
cm	Centimeter
COSY	Correlated spectroscopy
CSD	Chemical shift deviation
C-terminus	Carboxyl-terminus
CXC	Chemokine with one amino acid between the first two cysteines
δ	Chemical shift
δ_{Bound}	Chemical shift of Bound
$\delta_{Unbound}$	Chemical shift of Unbound
$\Delta\delta_{max}$	Maximum chemical shift deviation
$\Delta\delta_{obs}$	Observed chemical shift deviation
DQF-COSY	Double-quantum filtered correlation spectroscopy
DPC	n-Dodecylphosphocholine
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DSS	2,2-Dimethyl-2-silapentane-5-sulfonate sodium salt
dNTPS	Deoxynucleioside triphosphates
E.coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid.
FAS	Fatty acid synthase
FID	Free induction decay
FPLC	Fast protein liquid chromatography

g	Gram
g	Acceleration of gravity
GSH	Glutathione (reduced form)
GSSH	Glutathione (oxidised form)
GdmCl	Guanidine hydrochloride
HIV-1	Human immunodeficiency virus type 1
HPLC	High pressure liquid chromatography
HPP-CFC	High proliferative potential colony-forming cell
HSQC	Heteronuclear single-quantum correlation
Hz	Hertz
IgE	Immunoglobulin E
IPTG	Isopropyl-β-D-thiogalactosidase
ITC	Isothermal titration calorimetry
J	Spin-spin coupling constant
kbp	Kilo base pair
K_d	Dissociation constant
kDa	Kilo daltons
k_{off}	Rate constant off
k_{on}	Rate constant on
L	Litre
LB	Luria-Bertani Broth
MCAT	Malonyl CoA: ACP transacylase
mg	Milligrams
mL	Milliliters
MHC	Major histocompatibility complex
min	Minutes
mM	Millimolar
N-terminus	Amino-terminus
ng	Nanogram
nM	Nanomolar
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
NOESY	Nuclear Overhauser effect spectroscopy

OD	Optical density
P_{Bound}	Population in the bound state
$P_{Unbound}$	Population in the unbound state
PCR	Polymerase Chain Reaction
pM	Picomolar
PMN	Polymorphonuclear granulocyte neutrophils
PMSF	Phenyl methyl sulfonyl fluoride
ppm	Parts per million
RMSD	Squared root of mean square deviations
ROESY	Rotating-frame Overhauser effect spectroscopy
s	Second
Sc	*Streptomyces coelicolor*
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
τ_m	Mixing time
τ_c	Rotational correlation time
TAE	Tris-Acetate-EDTA
TEMED	Tetramethylethylenediamine
Taq	*Thermophilus aquaticus*
TFE	Trifluoroethanol
TMS	Tetramethylsilane
TOCSY	Total correlation spectroscopy
Tris	Tris-(hydromethyl)aminomethane
µL	Microliter
µM	Micromolar
UV	Ultraviolet
vMIPII	Viral monocyte inflammatory protein-II
ω_o	Larmour frequency
w	Watts
w/v	Weight to volume

1 Introduction

1.1 The Human Immune System

The human immune system is a remarkably complex defence system that protects against invading pathogenic microorganisms and cancer [1]. The immune response consists of two fundamentally different mechanisms, termed the innate response and the adaptive response. The innate immune response does not alter on repeated exposure to a stimulus, whereas the adaptive response improves with each exposure to the same stimulus. The immune responses are produced primarily by leukocytes. Leukocytes all arise from the pluripotent stem cells, but via two main lines of differentiation (Figure 1.1); the myeloid lineage (which produces phagocytes and accessory cells) and the lymphoid lineage (which produces lymphocytes). These are discussed in more detail in the following sections.

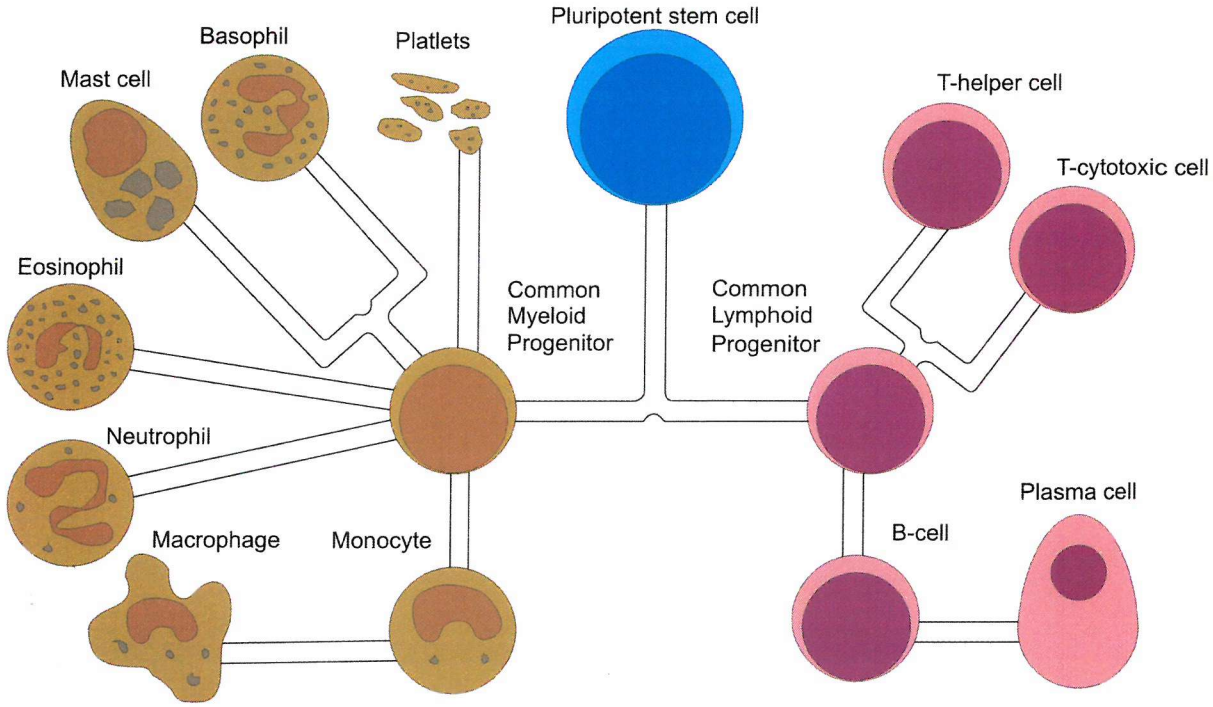


Figure 1.1 The lineage differentiation of the leukocytes. The pluripotent stem cell is shown in blue and differentiates via the common myeloid and lymphoid progenitors into the cells of the immune system. The myeloid cells shown in brown are the accessory and phagocytic cells. The lymphoid cells shown in red are the lymphocytes.

1.2 Phagocytes

Phagocytes form part of the innate immune response as they use non-specific recognition systems allowing them to bind to a variety of microbial targets. Phagocytes bind to,

internalize and then kill micro-organisms. There are three types of phagocyte: the mononuclear phagocytes, the polymorphonuclear granulocyte neutrophils and the eosinophils.

The mononuclear phagocytes are released from the bone marrow into the blood, where they circulate as monocytes. Upon migration into the surrounding tissue monocytes develop into macrophages, which are responsible for phagocytosis and also antigen presentation to the lymphocytes. The role of monocytes is unclear, but it is believed that they act as a circulating reserve of macrophages.

Often referred to as neutrophils or PMN, the polymorphonuclear granulocyte neutrophils migrate into the tissue spaces in response to certain stimuli. Neutrophils are short-lived cells which engulf foreign material, destroy it and then die. Bone marrow stem cells increase the neutrophil release rate in response to many types of infection and it is these cells which are normally the first to arrive at a site of inflammation.

The specific function of the eosinophils remains unclear, with many hypotheses proposed over the years regarding their role in the immune response [2-8]. The modern view is that, like neutrophils, they are motile phagocytic cells that can migrate from the blood into the tissue spaces [9]. Once there they defend against large potentially harmful organisms, such as nematodes and schistosomes. As eosinophils affect cells and tissues in many ways depending upon the environment in which they find themselves and the range of stimuli which act upon them, it is deemed appropriate to assess their role in each situation individually [9, 10].

1.3 Accessory Cells

The accessory cells are the basophils, mast cells and platelets. Basophils and mast cells are both involved in the mediation of inflammation through the use of preformed chemicals, such as chondroitin sulfates, proteases and histamine which are stored in cytoplasmic granules. Upon antigen crosslinking of the antibody immunoglobulin E (IgE), degranulation is induced and the preformed mediators are released extracellularly. The major difference between basophils and mast cells is that basophils are short-lived cells that mature in the bone marrow and circulate in the blood until they are recruited to sites of

inflammation or immune responses. In contrast, mast cells complete their differentiation in the vascular tissues, where they are long-lived and can retain their ability to proliferate under certain conditions [1].

In addition to their role in blood clotting, platelets also release inflammatory mediators either when activated during thrombogenesis or by means of antigen-antibody complexes.

1.4 Lymphocytes

Lymphocytes are the cells that are mostly responsible for the adaptive immune response and they are split into two main sub-groups: the B-lymphocytes and the T-lymphocytes. The B-lymphocytes (B-cells) mature within the bone marrow and upon release each expresses a unique antigen-binding receptor (antibody) on its membrane. When a naïve B-cell first binds the antigen that matches its membrane-bound antibody, it causes the cell to divide rapidly, with its progeny becoming either memory B-cells or effector B-cells (also known as plasma cells). Memory B-cells have a longer life span than naïve B-cells and continue to express the same membrane-bound antibody. Plasma cells produce the antibody in a secreted form and even though they live for just a few days they secrete enormous amounts of these antibodies, with rates of up to 2000 antibodies per second having been recorded [1].

T-lymphocytes (T-cells) also originate in the bone marrow, but migrate to the thymus gland to mature, rather than maturing in the bone marrow. Once mature, the T-cell expresses a unique antigen binding molecule, called the T-cell receptor, on its membrane. This receptor can only recognise antigens that are bound to cell-membrane proteins called major histocompatibility complex (MHC) molecules. When a naïve T-cell encounters the antigen it proliferates and differentiates into effector T-cells (T-helper and T-cytotoxic cells) and memory T-cells, which perform a similar task to that of the memory B-cells.

T-helper cells, like most of the cells in the immune system, secrete various growth factors known collectively as cytokines. These T-helper cell secreted cytokines play an important role in the activation of B-cells, macrophages and various other cells that participate in immune responses. Differences in the type of cytokine produced by the T-helper cell will result in different types of immune response.

Under the influence of T-helper cell derived cytokines, T-cytotoxic cells proliferate and differentiate into cytotoxic T-lymphocytes that perform a vital function in monitoring the cells of the body and eliminating any problem cells, such as virus-infected cells and tumour cells. The cytotoxic T-lymphocytes are also responsible for killing cells from a foreign tissue graft.

1.5 Chemokines

1.5.1 Introduction

Both the innate and adaptive responses are regulated by a group of soluble proteins called cytokines. The chemoattractant cytokines or chemokines were first discovered in 1977 [11] and comprise an entire super-family of cytokine proteins that are involved in the molecular regulation of cell trafficking. The chemokines are released at the point of infection to recruit leukocytes to the area. Their activities are mediated through the binding and activation of seven transmembrane helix G-coupled receptors that are expressed at the cell surface of leukocyte sub-populations. It is the study of the molecular interactions of these chemokines that forms the basis of this work.

Chemokines are perhaps the most complex of the G-protein coupled receptor ligands because of their large number, overlapping receptor specificity and extensive phylogenetic divergence [12]. As the number of family members identified increased, various short-lived collective terms were used to describe them including, "the platelet factor-4 family" [13], "the small inducible cytokine family" [14] and "the intercrines" [15]. Finally, the term "chemokine", a neologism short for "chemotactic cytokine", was accepted as the standard [16].

Chemokines can be classified into four major groups according to the number and spacing of conserved cysteines (Figure 1.2). α chemokines have the first two cysteines separated by one residue (CXC) and β chemokines have the first two cysteines adjacent (CC). γ chemokines have only two cysteines in the entire sequence (C), and δ chemokines have the first two cysteines separated by three residues (CX_3C). α, β and δ chemokines mostly have four cysteines, although there are a few examples that have more.

α Chemokines	CXC:	...CXC.....C.....C
β .Chemokines	CC:	...CC.....C.....C
γ Chemokines	C:	...C.....C
δ Chemokines	CX_3C:	...CXXXC.....C.....C

Figure 1.2 Structural classification of the chemokine family based upon the spacing and number of cysteines present.

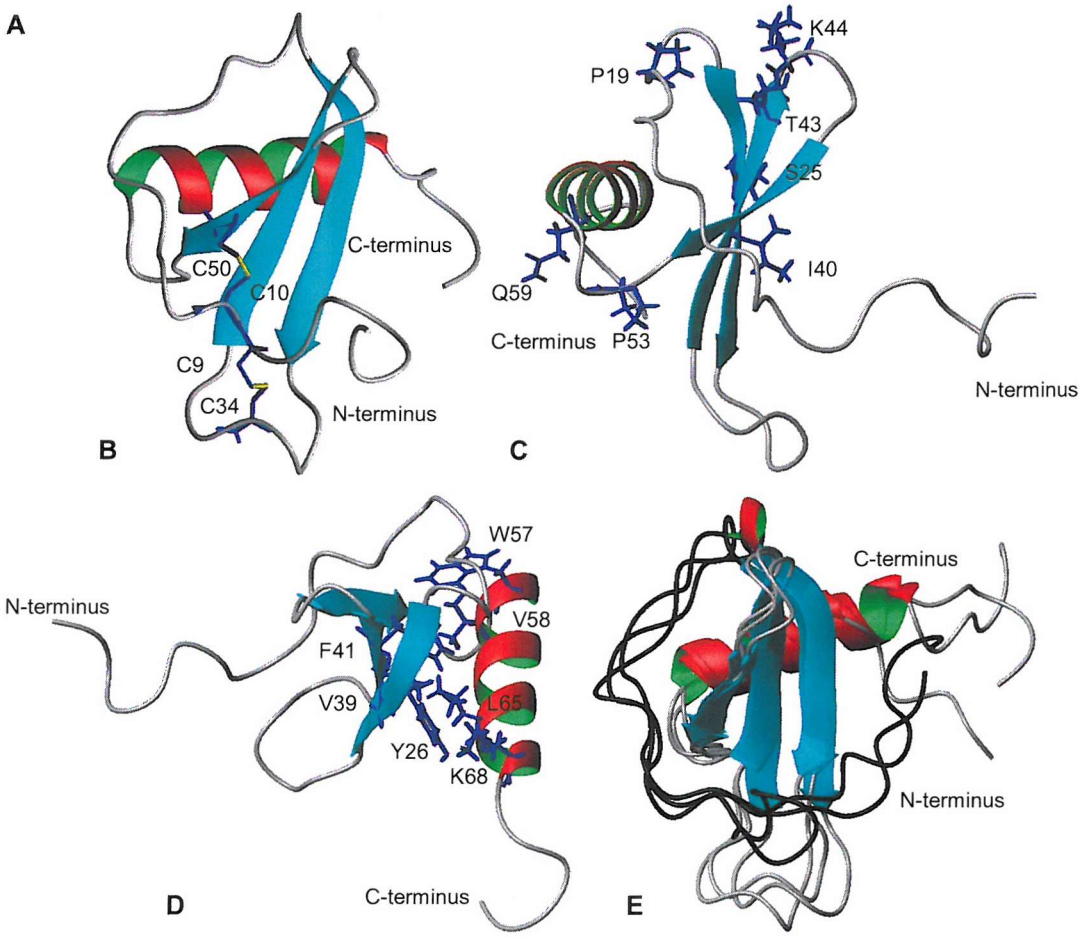
Due to the co-discovery of chemokines and their receptors by multiple groups, there has been a proliferation of chemokine aliases. The now accepted nomenclature is to denote the family; CXC, CC, C or CX_3C, followed by R for the receptors or L for the ligands. The family members are then sequentially numbered. All currently known chemokines and their receptors are listed in the appendix with their accepted nomenclature and their former names [12, 17-19].

1.5.2 CC Chemokine Structure

There is a subfamily of chemokines called the eotaxin family, which was defined by the member's ability to only activate CCR3. It is the eotaxin family that is the focus of this research and it consists of three members: CCL11 (eotaxin-1), CCL24 (eotaxin-2) and CCL26 (eotaxin-3). The alignment of the primary amino acid sequences (Figure 1.3A) and the sequence similarity between the family members suggests that they are closely related with respective similarities of 55% (CCL11 to CCL24), 54% (CCL11 to CCL26) and 52% (CCL24 to CCL26).

The tertiary structures of all the CC chemokines solved to date are very similar. The chemokine tertiary structure consists of a flexible N-terminus, a three stranded anti-parallel β sheet and a C-terminal α helix. Two disulfide bonds connect the flexible N-terminus to the body of the protein (Figure 1.3B).

```
CCL11 GSHM--GPASVPTTCCFNLANRKIPLQRLESYRRITSGKCPQKAVIFKTKLAKDICADPKKKWVQDSMKYLDQKSPTPKP-   74
CCL24 --GS----VVIPSPCCMFFVSKRIPENRVVSYQLSSRSTCLKAGVIFTTKKGQQFCGDPKQEWVQRYMKNLDAKQKKASPR   73
CCL26 -----TRGSDISKTCCFQVSHKPLPWTWVRSYEFTSN-SCSQRAVIFTTKRGKKVCTHPRKKWVQKYISLL--KTPKQL--   71
CC                 CC                  C              C
CON                CC       P     SY       C    VIF TK     C  P   WVQ     L  K
ruler -4...0.......10.......20.......30.......40.......50.......60.......70......
```



F

	CCL11	CCL24	CCL26
CCL11	0.24	0.91	0.97
CCL24	0.91	0.32	1.43
CCL26	0.97	1.43	0.31

Figure 1.3 The conserved residues and the structures of the eotaxin family [20-22] (PDB ID: 1EOT, 2EOT, 1EIG, 1EIH, 1G2S, 1G2T). A. The sequence alignment of the solved structures of the eotaxin family (CCL11, CCL24, CCL26), highlighting the conserved cysteines (CC) and the conserved residues of the eotaxin family (CON). The negative residues are those residues left after the removal of a purification tag used on the recombinant forms. B. The conserved disulphide bonds for CCL11 shown in yellow (C9, C10, C34 and C50). C. The conserved surface residues shown in blue for CCL11 (P19, S25, I40, T43, K44, P53 and Q59). D. The conserved interior residues shown in blue for CCL11 (Y26, V39, F41, W57, V58, L65 and K68). E. The published mean average structures of the eotaxin family were superimposed using the conserved secondary structure elements of CCL11 (L23-R28, V39-T43, I49-A51 and K56-L65), CCL24 (V21-L26, V37-T41, S47-G49 and E54-L63) and CCL26 (V24-F29, V39-T43, V49-T51 and K56-L65). The N-terminus is coloured black for clarity. F. The RMSDs in angstroms for the published NMR ensembles superimposed using the same residues as in Figure E.

The eotaxin family has fourteen conserved residues in addition to the four cysteines and these are shown as being surface or buried residues in Figure 1.3C and D respectively. The buried residues are clustered around the contact between the α helix and the β sheet. The surface residues seem more randomly distributed although no conserved residues apart from the cysteines occur in the N-terminus region, which is the proposed binding and activation site. Analysis of the published tertiary structures for the family shows that they are closely related in structure, with CCL11 more closely related to CCL24 and CCL26 than either is to the other (Figure 1.3E and F).

1.5.3 CC Chemokine Receptors

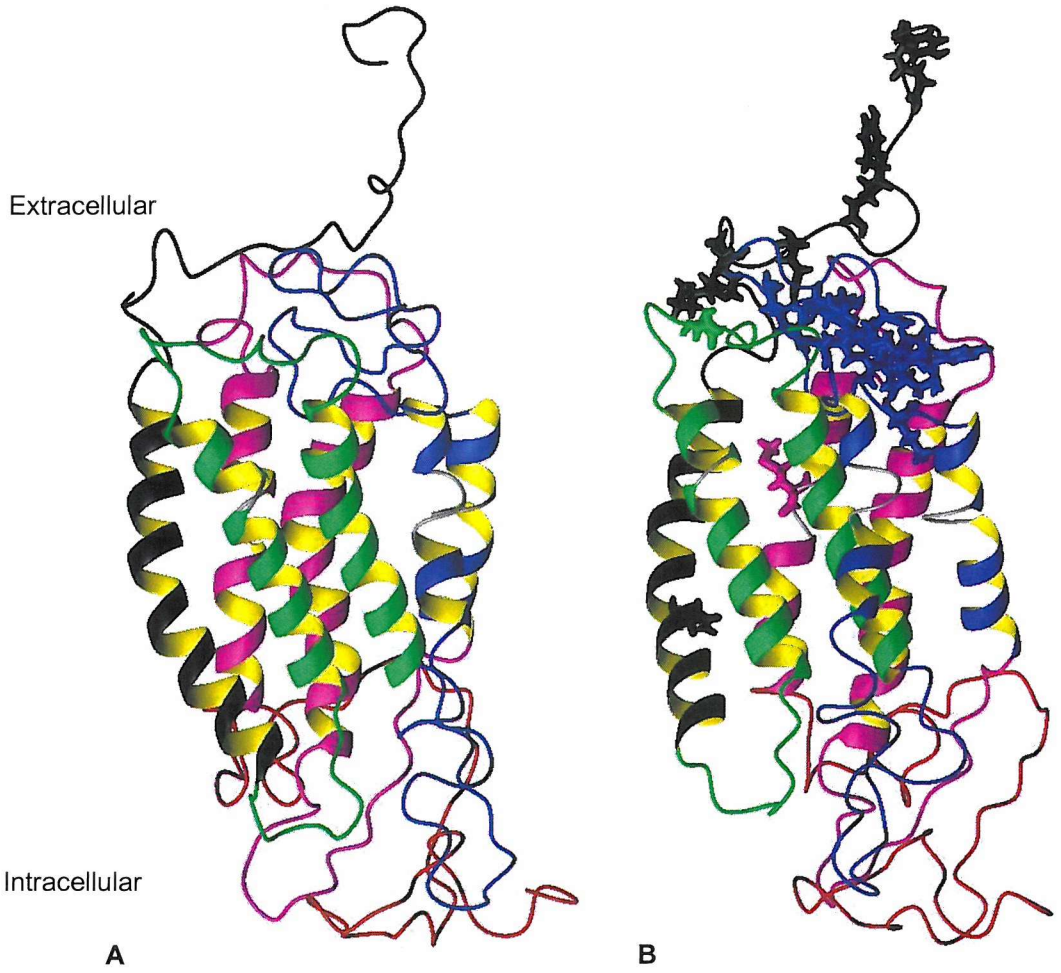
Each CC chemokine receptor has a distinct chemokine and leukocyte specificity, details of which are shown in the appendix. There is a 25-80% amino acid identity between the chemokine receptors which indicates a common ancestor. It is important to note that many other G-protein coupled receptors also have approximately 25% amino acid identity with chemokine receptors, suggesting that the structural boundary is not sharp [12].

The three-dimensional structure of chemokine receptors remains unknown, but models have been constructed for the transmembrane domains based on analogy with rhodopsin (Figure 1.4A) [23-26]. Models of the extracellular and intracellular domains are speculative and should be treated with care. Rhodopsin has one disulphide bond between extracellular loops one and two, and this is a highly conserved feature of G-protein coupled receptors. Chemokine receptors have two extra conserved cysteines in their extracellular domains, one in the N-terminus and the other in the third extracellular loop. It is partly because of these two extra cysteines that the models should be treated with care as a common assumption is to form a disulfide bond between them, but there is no experimental evidence in the literature for this; rather it is an assumption based on the oxidative environment they are in. It has been shown for CCR6 that the cysteines exist as free thiols [27]. It has also been shown that the normal chemotactic function of CXCL8 was mediated by one or more thiol residues [28, 29].

The mechanism by which the G-coupled receptors signal is believed to involve a structural rearrangement. In the high resolution structure of rhodopsin solved in its ground-state the DRY sequence which binds the G-proteins is hidden in the membrane, at the intracellular

end of helix three. The ligand for rhodopsin is 11-*cis*-retinylidene which is completely buried within the protein and undergoes a structural change upon activation with light. This change in structure of the ligand is believed to cause a change in the structure of rhodopsin, which then exposes the DRY sequence, allowing the G-proteins to bind.

It has been reported that CCR2, CCR5 and CXCR4 form homodimers [30-32] and in the case of CCR2, a dimer has been implicated as the functional form of the receptor. Dimerisation may be required for signalling.



Black: N-terminus and helix 1.

Blue: Intracellular loop 2, extracellular loop 2, helices 4 and 5.

Green: Intracellular loop 1, extracellular loop 1, helices 2 and 3.

Purple: Intracellular loop 3, extracellular loop 3, helices 6 and 7.

Red: C-terminus.

Figure 1.4 A. A model of the CCR5 receptor based on rhodopsin. The intracellular and extracellular loops are purely speculative. B. The model viewed after a 90° turn with the side chains of the residues implicated in binding chemokines highlighted [33-37]. The majority of these residues are located either on the N-terminus or extracellular loop 2.

1.5.4 Binding of Chemokines to their Receptors

The binding between chemokines and their receptors has been investigated predominantly through chimeras and mutagenesis. The data has suggested that the interaction involves at least two domains, one for docking and the other for triggering. A current model proposes that the N-terminus of the receptor binds to the chemokine, which then allows the N-terminus of the chemokine to bind and activate the receptor (Figure 1.5) [38-41].

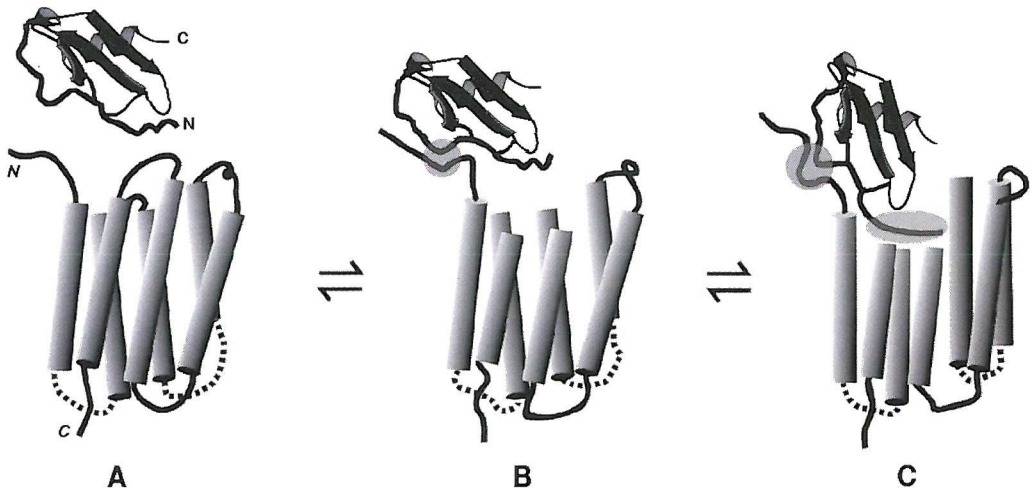


Figure 1.5 Possible model for the binding and activation of the chemokine receptors [40]. A. The separate chemokine and receptor. B. The N-terminus of the receptor first binding to the chemokine. C. The N-terminus of the chemokine activates the receptor. Extracellular loops 1 and 3 of the receptor have been omitted from figures B and C for clarity.

Whilst this model explains the results in certain systems, it is not general enough to explain all published results. For example, Monteclaro showed using chimeras that the N-terminus of CCR1 was not important for binding CCL3 whereas the N-terminus of CCR2 was the only strong binding site for CCL2 [39, 41]. An attempt to generalise any specific interaction has to be treated with caution as can be seen by a natural mutation (A29S) in CCR5 which was tested against several chemokines. Whilst the mutant showed no measurable binding to CCL4 surprisingly there was still signalling, whereas with CCL3 the mutant no longer bound nor signalled, but the interaction with CCL8 was unaffected [42].

When the results from numerous studies are considered it would appear that all the extracellular domains of the receptor are important for the binding of the chemokine, but which domain has the greater effect depends upon the specific system. The data available

on CCR3 is limited and is discussed briefly in section 1.5.5 and in more detail in chapter 4. The receptor that has been studied the most is CCR5 due to its role as the co-receptor of human immunodeficiency virus type 1 (HIV-1) [43, 44] and Figure 1.4B shows those residues of CCR5 that have been implicated in the binding of chemokines [33-37]. The majority of these residues are located on either the N-terminus or extracellular loop 2 with only three of the residues located elsewhere. Of these three residues one is in extracellular loop 1 and the other two are buried in the membrane.

In terms of signalling, results normally implicate extracellular loop 2 as the domain involved, although residues in extracellular loops 1 and 3 and helices 2 and 3 have been shown to be important [39, 45-48]. With so much evidence pointing towards loop 2, it is possible that the other residues play a structural role rather than mediating signalling induced from an interaction with the chemokine. One very interesting mutant in helix two is T82K, which despite being hidden in the membrane shows constitutive activation. It is possible that this mutant is causing a structural rearrangement which is exposing the DRY sequence.

The CC chemokines have been studied mainly by truncation, elongation and mutagenesis. Truncation has shown that the chemokine N-terminus is involved in the activation of the receptor but is not the major binding site. For example, the removal of the first eight residues from CCL26 results in a only a three-fold reduction in its binding to CCR3, but no signalling is observed [49]. The N-terminus is also involved in the selectivity of the chemokines, with a natural truncation of the first two residues of CCL4 causing it to start signalling via non-native partners CCR1 and CCR2 whilst retaining its ability to signal via its native CCR5 receptor [50].

Elongation of the N-terminus has varying effects with the addition of a single methionine to CCL5 producing an antagonist for CCR5, whereas the addition of aminooxypentane to CCL3 actually increases the binding to CCR5 [51, 52]. The addition of a fluorescent tag to the N-terminus did not affect the binding or activation of the receptor, but the fluorescence was quenched by iodide (which cannot enter the membrane) suggesting that the N-terminus of the chemokine does not enter the membrane, but interacts with the extracellular loops only [53].

A mutagenesis study of CCL5 showed that the residues important in binding CCR1, CCR3 and CCR5 were different [54] and as such care must be taken in the extrapolation of results from specific systems. It is interesting to see the general binding sites as defined by mutagenesis of several different systems, shown in Figure 1.6. Almost all the residues involved in binding are on the flexible regions of the protein, with only one residue located on regions of secondary structure.

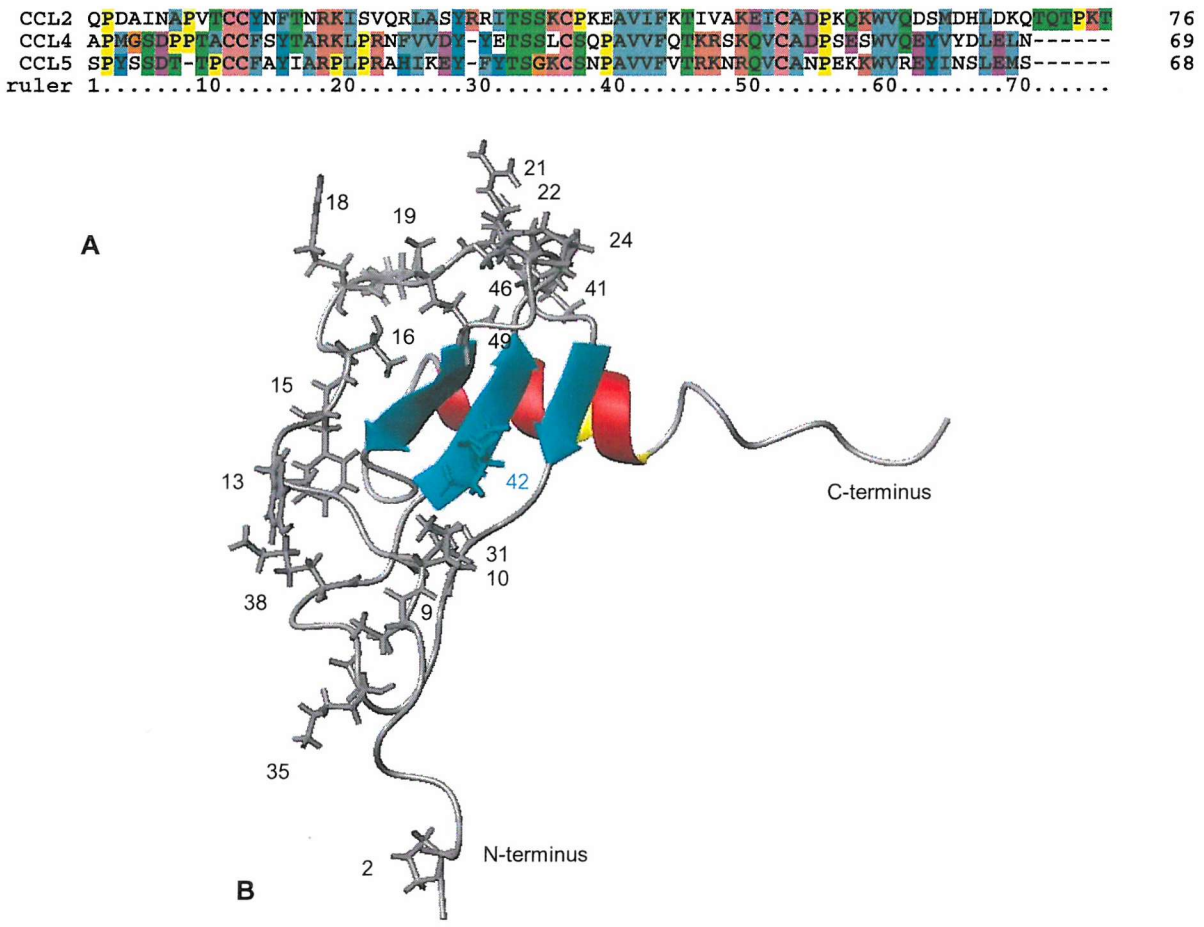


Figure 1.6 A. Sequence alignment of CCL2, CCL4, and CCL5. B. CCL2 with the residues implicated in receptor binding for one or more of CCL2, CCL4 and CCL5 from several studies. These residues were standardised using the sequence alignment. The residues are 2, 9, 10, 13, 15, 16, 18, 19, 21, 22, 24, 31, 35, 38, 42, 46, 47 and 49 [54-60]. The residue 42, coloured in blue, is the only residue implicated in binding that is on any of the secondary structural elements.

1.5.5 CCR3 and the Eotaxin Family

As discussed in section 1.5.2 the eotaxin family (CCL11, CCL24 and CCL26) is the focus of this research and this family is specific to CCR3. The data available on CCR3 is limited

to a few studies. Chimeras of CCR3 with CCR1 implied that the N-terminus and the third extracellular loop of CCR3 were involved in the binding of CCL11, but that neither functions as the activation site [61]. This result was supported by Datta who made a soluble mimic of CCR3 comprising its N-terminus and third extracellular loop [62]. Removing either the N-terminus or the third extracellular loop from the mimic resulted in a lowering of its binding to CCL11. The mimic bound CCL11 with a K_d of 3μM compared with the native K_d of 0.52nM [63]. This suggests that there may be another binding site not included in the mimic and from studies with mutants of CCL11 it was suggested that R16 and K17 on CCL11 may be involved in binding to this unknown site [62]. Whether the missing binding site is formed from another part of the receptor or post-translational modifications is uncertain. There are two post-translational modifications that are believed to occur at the N-terminus of CCR5: sulfation of the tyrosine residues and O-glycosylation on S6 [64-66]. As these post-translational modifications are not included in the mimic it is unclear whether they are involved in the binding. Interestingly, one of the fundamental design points was the formation of a disulphide bond between the N-terminus and the third extracellular loop. As discussed in section 1.5.3 this disulphide bond might not actually be present *in vivo*.

1.6 Nuclear Magnetic Resonance Spectroscopy of Proteins

Nuclear magnetic resonance spectroscopy, more commonly known as NMR spectroscopy, was first observed in 1946 by two groups working independently. Purcell *et al* detected weak radio-frequency signals that were generated by the nuclei of atoms in paraffin wax [67], whilst Bloch *et al* observed radio signals from the atomic nuclei in water [68].

The use of NMR spectroscopy for the analysis of protein structures is a relatively new application, with the first globular protein structure, proteinase inhibitor II, solved in 1984 [69] as compared with the first X-ray structure, myoglobin, solved in 1960 [70]. It was only in 1988 that the first "high resolution" NMR structure of a protein, alpha-amylase inhibitor tendamistat, was solved [71]. NMR spectroscopy has since been shown to be a very powerful tool for studying the structure and ligand binding of peptides and proteins in solution, with thousands of structures having been solved by increasingly complex applications of NMR spectroscopy.

1.7 Nuclear Magnetic Resonance Theory

The full explanation of NMR theory is not necessary to understand its application in this work. There are many text books on the subject, from simple introductions [72, 73] to more rigorous quantum mechanical derivations [74, 75]. This section attempts to explain a few select points, introducing certain experiments in terms of the information they contain and the parameters that vary between each sample.

For a NMR signal to be observed from an isotope the nuclear spin quantum number (spin) must not be equal to zero and for sharp NMR signals spin should be equal to a half. Spin is a fundamental property of the nucleus and is derived from the number of protons and neutrons in an isotope.

Table 1.1 shows the nuclear spins and natural abundance of the nuclei most generally found in proteins. Oxygen and sulphur both have a NMR active isotope, but neither isotope has a spin equal to a half; this combined with their low sensitivity means oxygen and sulphur are not commonly used in the study of proteins. Hydrogen, carbon and nitrogen all have an isotope with spin of a half, but it is only for hydrogen that this NMR active isotope is the most abundant. Despite the low natural abundance of ^{13}C it is possible to observe a NMR signal. In contrast the low natural abundance of ^{15}N combined with its low sensitivity means that it cannot be used. Where the observation of ^{13}C and ^{15}N is required in the study of proteins, the concentration of these NMR active isotopes is enriched to allow useful spectra to be obtained.

Isotope	Spin	Natural Abundance (%)	Sensitivity Relative to Hydrogen
^{1}H	½	99.99	100
^{2}H	1	0.01	15
^{12}C	0	98.93	n/a
^{13}C	½	1.07	25
^{14}N	1	99.63	7
^{15}N	½	0.37	10
^{16}O	0	99.75	n/a
^{17}O	$^{5}/_{2}$	0.04	14
^{32}S	0	95	n/a
^{33}S	$^{3}/_{2}$	0.75	7

Table 1.1 The nuclear spins of the common nuclei in proteins detailing in (a) the isotope (b) the spin (c) the natural abundance and (d) the sensitivity relative to ^{1}H [76, 77].

To observe a NMR signal the protein sample is placed in a magnet and a radio-frequency pulse applied. The intensity of the signal generated from a pulse is cyclic with respect to the length of time the pulse is applied for. The length of time that a pulse needs to be applied in order get the maximum signal is called the 90 degree pulse and must be calibrated for each sample. The calibration is normally done by finding the pulse length that equates to the 180 degree pulse and dividing the time by two; the significance of the 180 degree pulse is that it gives zero signal, which is easier to observe experimentally than a maximum. Pulse sequences are made up of combinations of pulses and delays of varying lengths that enable different aspects of the NMR signal to be selected.

The standard way to measure the NMR signal is by recording the free induction decay (FID). The FID is the weak oscillating voltage that the nuclear spins induce in a surrounding coil when perturbed from their equilibrium position. As the sample returns to its equilibrium position the signal decays through the effects of relaxation T1 (re-establishment of the equilibrium) and T2 (loss of phase coherence). The FID is therefore a time dependent oscillating voltage. The FID can be transformed to the frequency domain by the mathematical procedure of Fourier transformation. The NMR signal that arises contains different information about the environment of each of the nuclei, depending on the pulse sequence used. The pulse sequences used in this study are shown in the appendix and will not be discussed in detail. The experiments will be discussed in terms of the information they contain and the parameters that require consideration in running them. First, some general concepts that apply to all of them require discussion.

The most basic NMR experiment is the one-dimensional spectrum which normally contains information about the chemical shift (δ), the peak area, the peak intensity and the spin-spin coupling constants (J) of the observed nuclei. The chemical shift is the frequency of the peak and is proportional to the field strength of the spectrometer's magnet. It is measured relative to a standard such as 2,2-Dimethyl-2-silapentane-5-sulfonate sodium salt (DSS) or tetramethylsilane (TMS). Due to the fact that the frequency of the peak is field dependent, it is normally stated in parts per million (ppm) where:

$$\text{Chemical shift of a peak (ppm)} = \frac{(\text{Frequency of a peak} - \text{Frequency of the standard}) * 10^6}{\text{Frequency of the standard}}$$

The chemical shift is related to the environment in which the nucleus is found and because of this the chemical shift of the nucleus changes depending on the chemical sequence and the structure of the molecule. Figure 1.7 shows the general chemical shifts of the hydrogens as found in aqueous solutions of proteins.

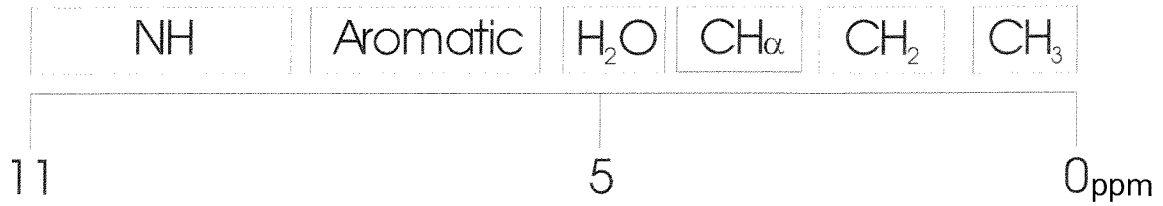


Figure 1.7 The approximate chemical shift ranges of the hydrogen resonances observed in proteins in ppm. The box widths define the chemical shift range for the particular hydrogen type shown.

The area under the peak is proportional to the number of nuclei that contribute to the peak, but the peak intensity can vary greatly depending upon the relaxation rate of the nuclei; the faster the nuclei relax the broader the peak.

Spin-spin coupling is caused by a through bond interaction between a nucleus and its nearest neighbours and results in the peaks being split into multiplets of n+1 where n is the number of adjacent splitting nuclei.

Water constitutes a large proportion of an aqueous protein sample and the hydrogens of water give a signal in the same region (~5ppm) as those from the protein. The water signal is much more intense than those signals from the protein, to such an extent that without correction the water peak would force the protein peaks into the noise level. The methods used to reduce the water peak are known as water suppression and two different types are found in the pulse sequences used in this work, either presaturation or water-gate, but both of them involve selecting only the signals from water and destroying them. As the name implies presaturation is applied before the start of the pulse sequence and involves using a low powered continuous pulse at the water frequency. The water-gate is found at the end of the pulse sequence and is a shaped pulse which has to be calibrated for each sample, to give the minimum signal from the water.

Due to the size of proteins and the small chemical shift dispersion of the hydrogen nuclei, the one-dimensional ^1H spectrum normally contains many overlapping peaks making

analysis difficult or impossible. A two-dimensional spectrum separates the peaks into a second frequency dimension, which decreases the amount of overlap and makes the spectrum more interpretable. All two-dimensional sequences have the same basic format and can be divided into four well-defined units. These units are the preparation, evolution, mixing and detection periods and are shown for the simplest two-dimensional pulse sequence in Figure 1.8. The preparation period consists of a delay that ensures that the sample is at equilibrium, followed by a pulse or a series of pulses that then perturb the system. The detection period is the time period during which the recording of the FID takes place. The evolution period is an incremented stepped-variable time delay which gives the second dimension and the mixing period comprises a pulse or series of pulses that perturb the system to make the second dimension different from the first. From system to system it is normal to vary the number of increments in the evolution period, depending on the resolution required in that dimension and the time available for the experiment. The length of time of the mixing period also varies and is normally dependent on the information required and the size of the protein being studied.

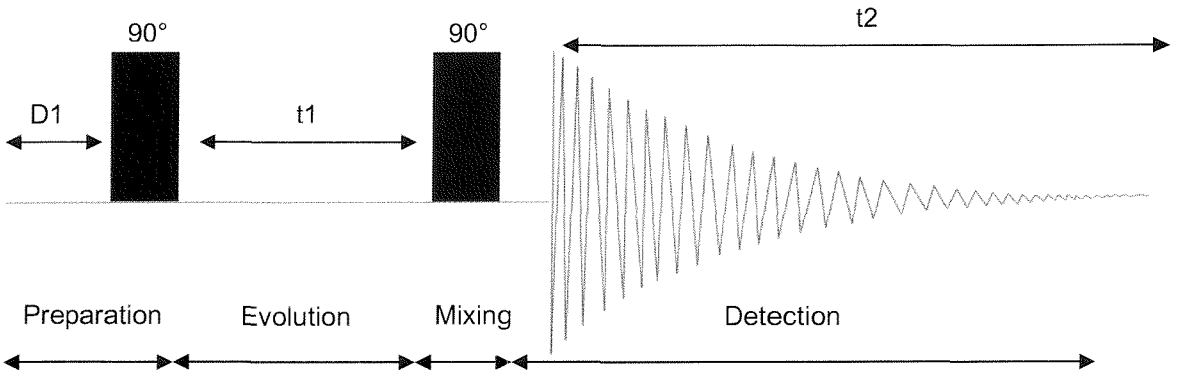


Figure 1.8 The simplest pulse sequence for a two-dimensional NMR experiment, showing the preparation period consisting of a delay of duration D1 and a 90° pulse. The evolution period consists of a time t1 and in this case the mixing period is simply a 90° pulse. The detection period consists of a time t2.

A third dimension can be added and this is commonly done if the sample is isotope-enriched with ^{15}N or ^{13}C. A basic three-dimensional pulse sequence is similar to that of a two-dimensional pulse sequence with the addition of another set of evolution and mixing periods. A three-dimensional spectrum causes further separation of the peaks which normally enables unambiguous assignment.

1.7.1 Correlated Spectroscopy (COSY)

Correlated spectroscopy (COSY) correlates hydrogen nuclei with mutual spin-spin coupling using a mixing period of a single pulse (Figure 1.8). The spectrum reveals only connectivities between hydrogens connected by three covalent bonds or less. Figure 1.9B shows a representation of the COSY spectrum for tryptophan (Figure 1.9A). The spectrum is symmetrical about the diagonal and for one side of the diagonal the peaks have been labelled. The resonance labelling is such that H2:H1 represents the connectivity between hydrogens, with the first resonance (H2) on the F2 axis and the second resonance (H1) on the F1 axis. The linking of the HA to the HN chemical shift gives rise to the "fingerprint" region which is some times referred to in the literature.

COSY enables the identification of the ten chemical shifts corresponding to the ten hydrogens on tryptophan. These chemical shifts can then be placed into three spin systems dependent on their coupling. The spin systems are:

(i) HE1 and HD1

(ii) HN, HA and HB#

(iii) HE3, HZ3, HH2 and HZ2

Once the chemical shifts are placed in their spin system, some of them can be assigned to hydrogens based on their expected values from Figure 1.7. This works well for spin systems (i) and (ii) however spin system (iii) has no distinctive chemical shift and requires another experiment for unambiguous assignment (section 1.7.3).

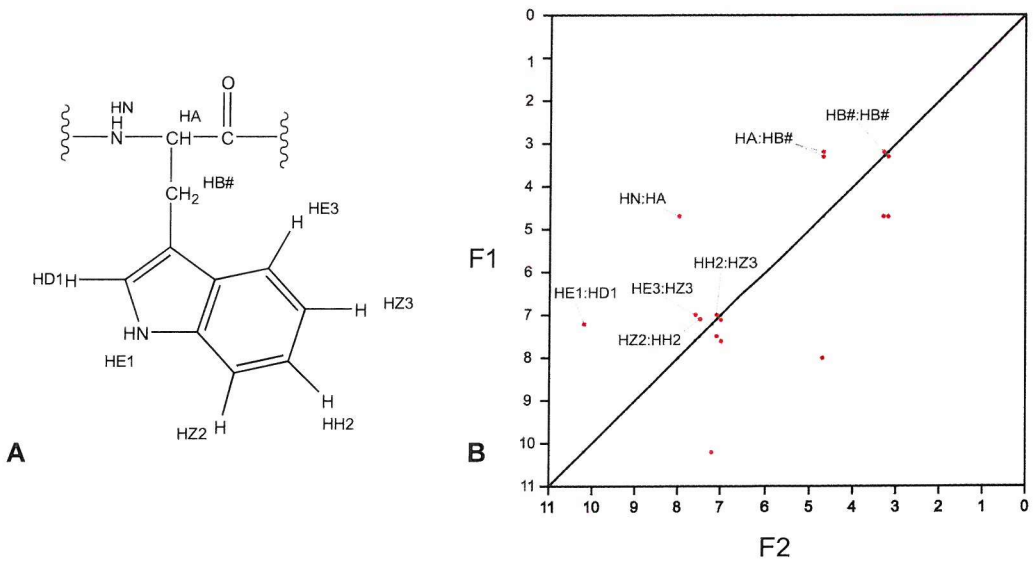


Figure 1.9 A. The structure and hydrogen nomenclature of tryptophan. B. Representation of an idealised COSY spectrum of tryptophan showing all the cross-peaks that could be observed.

1.7.2 Total Correlation Spectroscopy (TOCSY)

Total correlation spectroscopy (TOCSY) correlates coupled homonuclear spins across an extended network of *J*-coupled spins. This means the spectrum gives the same assignment information as the COSY but correlates the linking of all hydrogens in the same spin-system, even if they do not share mutual couplings. The advantages of this form of spectroscopy are that it makes the spin systems easier to identify, removes ambiguity around the diagonal and visually it closely resembles the NOESY (section 1.7.3), aiding interpretation. The disadvantage is that there are more peaks in the spectrum, which can start to overlap. The mixing period, which is a series of pulses, allows the correlation of all the hydrogens in a coupled system. If the mixing time (τ_m) is very short then the correlation is just between adjacent nuclei; as the mixing time increases the correlation spreads further down a coupled system (Figure 1.10). Figure 1.11 shows a representation of the TOCSY spectrum for tryptophan (Figure 1.9A).

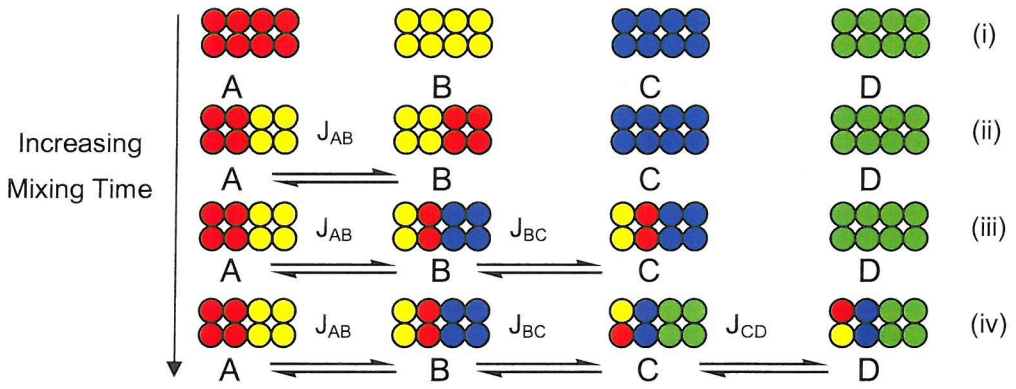


Figure 1.10 The effect of increasing the mixing time has upon the number of correlations observed for a spin coupled system (A-B-C-D). (i). Representation of equilibrium with each of the atoms depicted with eight colour balls. (ii). At a short mixing time there is an interaction between A and B that results in the transfer of some of their coloured balls. (iii). At a longer mixing time some of the balls that were originally transferred from A to B will start to be transferred from B to C hence A and C are now showing coupling to each other. (iv). At an even longer mixing time the balls that were transferred from A to B to C will now be transferred onto D linking the whole system together.

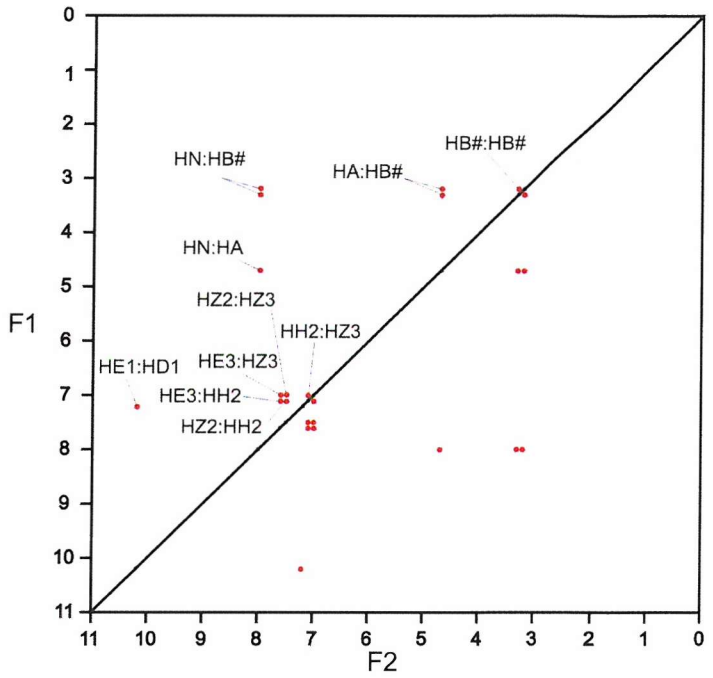


Figure 1.11 Representation of an idealised TOCSY spectrum of tryptophan showing all the cross-peaks that could be observed.

1.7.3 Nuclear Overhauser Effect Spectroscopy (NOESY)

Nuclear Overhauser effect spectroscopy (NOESY) separates the peaks based on the nuclear Overhauser effect (NOE) which is seen between nuclei that are close in space, whether through bond or not. The size and strength of the NOE depends on two criteria: the distance between the two hydrogens and the rate at which a molecule rotates in solution, which is defined by its rotational correlation time (τ_c). The rotational correlation time is the average time required for the molecule to rotate through an angle of one radian about any axis.

It is possible to calculate the theoretical variation of the homonuclear two-spin system NOE with the molecular tumbling rate, which is defined as $\omega_0\tau_c$ (where ω_0 is the spectrometer observation frequency). For molecules that are tumbling fast the NOE will be positive to a maximum value of 50% whereas molecules that are tumbling slowly have a NOE that is negative up to a maximum of 100% (Figure 1.12A). There is very little NOE build up when $\lg(\omega_0\tau_c)$ is around one and this would normally correlate to molecules with a molecular mass around 1.5kDa.

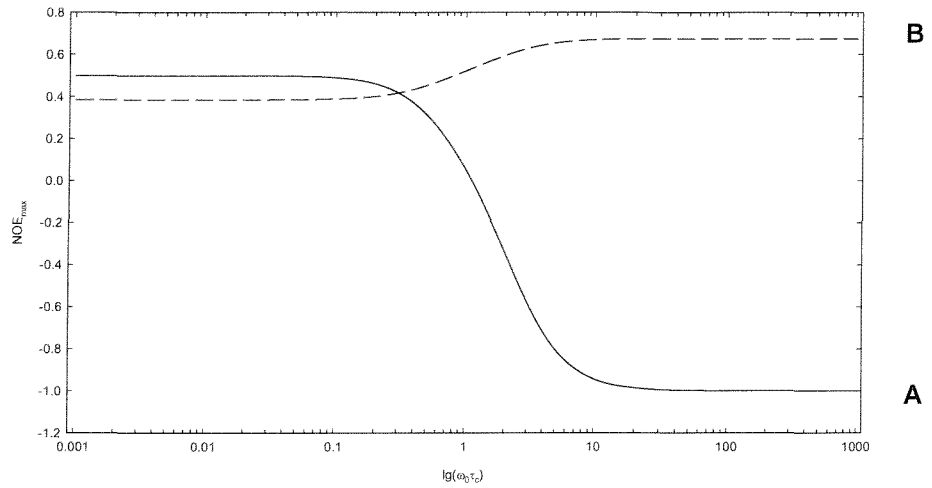


Figure 1.12 The variation of the theoretical maximum of the NOE for a homonuclear two-spin system as a function of the molecular tumbling rate ($\omega_0\tau_c$). A. The solid line shows the variation when the NOE is measured via a transient method as in the NOESY experiment. B. The dashed line shows the variation when the NOE is measured using a rotating frame as in the ROESY experiment [78, 79].

The NOE is dependent on the distance between the nuclei; the closer hydrogens are to each other the stronger the NOE and hence the NOE can be interpreted in terms of the spatial proximity between hydrogens. Whilst it is theoretically possible to calculate the exact distance based on the intensity of the NOE, other experimental factors make such a determination impractical. For structural calculations the NOEs are therefore grouped together on the basis of their strength and then given a distance range within which each pair of hydrogens must exist.

The mixing time used for the NOESY is related not just to how big the protein is but also how far apart the hydrogens to be observed are. A short mixing time will result in only the nearest hydrogens generating NOEs. Increasing the length of the mixing time will enable NOEs to hydrogens further out to build up, until such time as the signal starts to decay. Figure 1.13A shows a representation of a possible NOESY spectrum for tryptophan (Figure 1.9A).

The spectrum enables the linking of the spin systems due to the NOEs from HE1 to HZ2 and HE3 to HB#. The identification of HZ2 and HE3 enables the unique assignment of HH2 and HZ3 from the COSY.

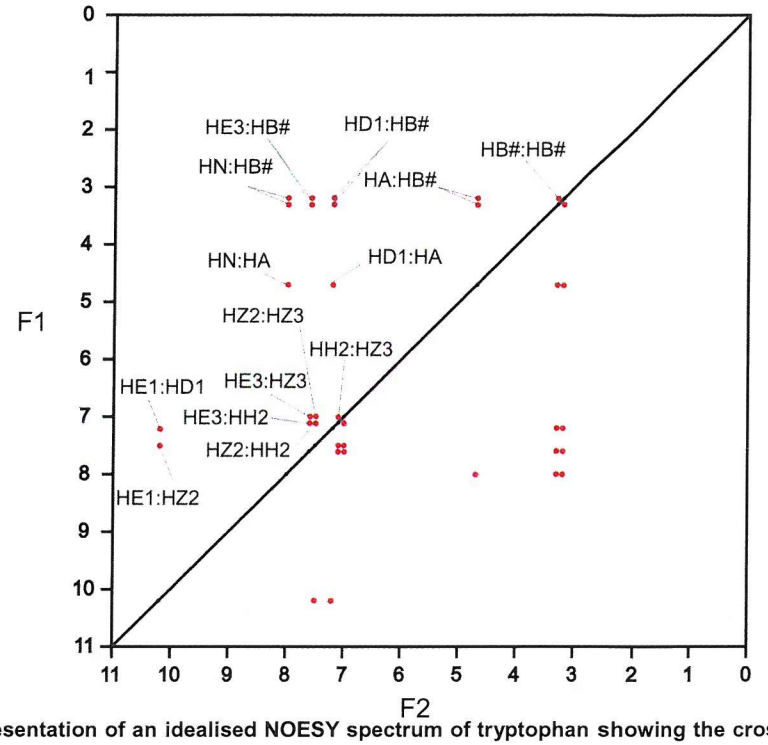


Figure 1.13 Representation of an idealised NOESY spectrum of tryptophan showing the cross-peaks that could be observed.

1.7.4 Rotating-Frame Overhauser Effect Spectroscopy (ROESY)

Rotating-frame Overhauser effect spectroscopy (ROESY) also establishes NOEs and hence spatial proximity between hydrogens. The ROESY differs from the NOESY in the manner in which it achieves the NOE and is suitable for any molecule (Figure 1.12B). However it is only normally used for proteins when $\omega_0\tau_c$ is around one and the NOESY is of no use, because the maximum NOE possible for ROESY is only 60%. The mixing period consists of a series of pulses and can be varied in length in the same way as the NOESY.

1.7.5 Heteronuclear Single-Quantum Correlation (HSQC)

Heteronuclear single-quantum correlation (HSQC) shows the correlation between coupled heteronuclear spins across a single bond and hence can be used to identify directly connected nuclei. The most commonly used version for protein work correlates nitrogen with hydrogen (^{15}N-HSQC) although the correlation of carbon with hydrogen (^{13}C-HSQC) is also used.

A ^{15}N-HSQC correlates each nitrogen atom with the directly attached hydrogen, so in proteins this means the spectrum only contains the signals from the amide backbone (except proline) and the side chains NHs of asparagine, glutamine, tryptophan, arginine

and histidine. This along with the larger distribution of NMR frequencies in the nitrogen dimension when compared to that of hydrogen, means that the peaks in a ^{15}N-HSQC spectrum have a much greater dispersion than those in the homonuclear spectra.

One of the main uses of ^{15}N-HSQC is for the study of ligand binding where the comparison of the spectra for the protein before and after addition of the ligand can show the site of interaction as discussed in section 1.10.

1.7.6 ^{15}N-Filtered Experiments

These experiments selectively remove the signals of hydrogens that are attached to ^{15}N such that the 2D-hydrogen spectrum produced only contains signals in the amide region that come from ^{14}N. These experiments are used in titrations where one component has been selectively labelled with ^{15}N and the other has not, so for example the observation of the peptide in ^{15}N-protein ^{14}N-peptide mixtures. A similar experiment is possible for the side chains using ^{13}C and ^{12}C.

1.7.7 Three-Dimensional Experiments

These experiments spread the information from a two-dimensional experiment into a third dimension. An example of this is a NOESY-TOCSY, which is a NOESY spectrum split into the third dimension by a TOCSY and although the resolution is different it should contain the same information as the two-dimensional TOCSY and NOESY spectra or the three dimensional TOCSY-NOESY, which is a TOCSY split into the third dimension by a NOESY. The other three-dimensional experiments commonly used to aid in the unambiguous assignment of ^{15}N-proteins are the HSQC-TOCSY, which is a HSQC split by a TOCSY and the HSQC-NOESY which is a HSQC split by a NOESY.

1.8 Sequential Assignment using Two-Dimensional NMR Spectra

The assignment of proteins from two dimensional spectra is a time consuming task which has been superseded through the introduction of isotope-enrichment, new NMR experiments and automatic assignment by computer programs [80]. As the isotopic enrichment of synthetic peptides is expensive, the traditional approach of homonuclear sequential assignment from two-dimensional NMR spectra was employed in this project. This technique was pioneered by Wuthrich and co-workers and is described in detail

elsewhere [72, 81]. In brief, it normally requires the collection of at least two data sets: one data set supplies through bond information (such as the TOCSY) and the other data set supplies through space information (such as the NOESY). The first stage in the assignment is the identification of an amino acid side chain, which is done using the TOCSY. Which side chain is identified depends on the amino acid composition of the protein being studied and the personal preference of the assigner; however it is normally easiest to start with a unique side chain or one that has distinctive chemical shifts and as such tryptophan is a common favourite. Once a side chain has been found and chemical shifts of the alpha-hydrogen and the amide-hydrogen assigned on the TOCSY then the through space information on the NOESY is used. The NOESY should contain the same information as the TOCSY plus information about the surrounding amino acids, as shown in Figure 1.14. The NOESY should have at least a NOE from the amide-hydrogen to the preceding amino acids side chain if not also to the proceeding amide-hydrogen.

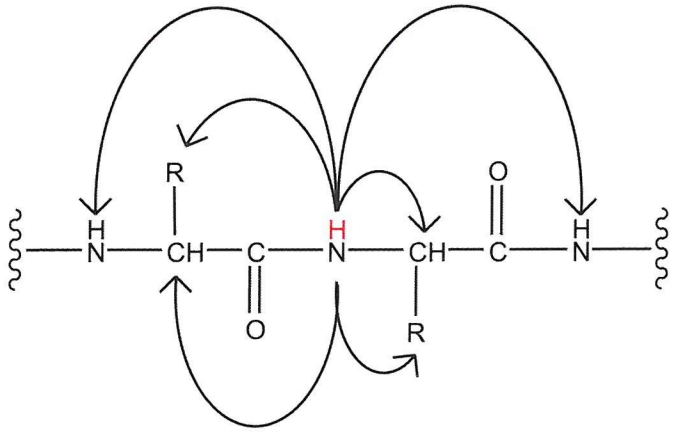


Figure 1.14 The NOEs that might be observed from an amide hydrogen (red) to the surrounding residues.

Once a nearest neighbour has been established then the process can be repeated, until the protein is fully assigned or the assignment becomes ambiguous, in which case another identifiable side chain is chosen and the process repeated. This methodology was used to assign two mutants of the same protein as described in appendix VI.

1.9 Secondary Structure from Chemical Shift Data

The deviation of chemical shifts from standard tables of either values for randomly coiled tetrapeptides [81] or averaged values from NMR databases [82-84] can give information about secondary structure. The chemical shift index was developed by Wishart *et al* and

works by comparing the alpha-hydrogen chemical shift for each residue to a set of random coil standards and assigning a number based on the outcome [85]. If the chemical shift is within 0.1ppm of the random coil then the residue is designated 0; if the deviation is greater than this then the residue is either designated +1 if the deviation is positive or -1 if the deviation is negative. A local density of non-zero indices which exceeds 70% is then required to define regions of structure and can be measured over several residues. A minimum of three consecutive "+1s" is required for a beta sheet and a minimum of four (not necessarily consecutive) "-1s" is required for an alpha helix. The chemical shift index is an empirical method but has been shown to be more than 90% effective when predicting secondary structure in proteins.

The Andersen plot [86] is a more reliable method for extracting information on possible secondary structure in peptides, using the chemical shift data rather than the chemical shift index. Analysis of the data for partially folded peptides, protein fragments and other peptides showed that there was a correlation between the change in the amide hydrogen chemical shift with temperature (amide temperature coefficient) and the amide hydrogen chemical shift deviation (CSD) for structured peptides. The analysis is facilitated by plotting the amide hydrogen CSD for the lowest temperature against the amide temperature coefficient. For peptides the occurrence of large chemical shift deviations and abnormal gradients are diagnostic of partial structuring at lower temperatures, which becomes increasingly random on warming. These plots can be used as justification of structure when they have a correlation coefficient greater than 0.7 and amide hydrogen CSD values of greater that 0.3.

1.10 Study of Protein-Peptide Interactions by NMR Spectroscopy

The study of protein-peptide interactions is possible because of the sensitivity of the NMR signal of a nucleus to its immediate environment [87]. When the peptide binds to the protein it will perturb the environment around the residues involved and hence change the nuclear magnetic resonances in the nearest neighbours. Experimentally one of the most common methods is to label the protein with ^{15}N and record ^{15}N-HSQCs as the unlabelled peptide is titrated in. Depending on the disassociation constant or exchange rate there are three modes of exchange observed in NMR spectroscopy, termed fast, intermediate and slow. These are typified by three modes of peak change as a titration progresses.

The exchange between two sites is not unique to proteins and is discussed in various text books [74, 88]. In this case there are two states, bound and unbound, which have the fractional populations p_{bound} and $p_{unbound}$ (with the sum being equal to 1) and two first order rate constants k_{on} and k_{off} such that:

$$\text{Bound} \underset{k_{on}}{\overset{k_{off}}{\rightleftharpoons}} \text{Unbound}$$

The simulated spectra in Figure 1.15 show the differences that can be expected for slow, intermediate and fast exchange.

In NMR spectra slow exchange (pM to low μM K_d) is characterised by two sets of resonances which relate to the bound and unbound forms. This means that during the course of a titration the peaks relating to the unbound form will disappear and the bound ones appear (if the chemical shift (i.e. frequency) difference is large compared to K_{off}). With fast exchange (μM to mM K_d) the bound and unbound forms give a single resonance at a weighted average frequency, so during the course of a titration the peaks will move from the unbound form to the bound form. The rate at which the peaks move in the fast exchange limit is related to the strength of the binding; the stronger the binding then a 1:1 complex is more easily reached. For any single mode binding interaction all the residues involved should reach the bound form at the same rate. With intermediate exchange the resonances broaden and eventually disappear from the spectra, so during the course of a titration the intensities of the peaks will vary with some peaks disappearing and sometimes appearing in the spectrum [89].

The effect the interaction has upon the ^{14}N-peptide can be seen by using ^{15}N-filtered experiments and comparing the amide region to that observed in standards without protein present.

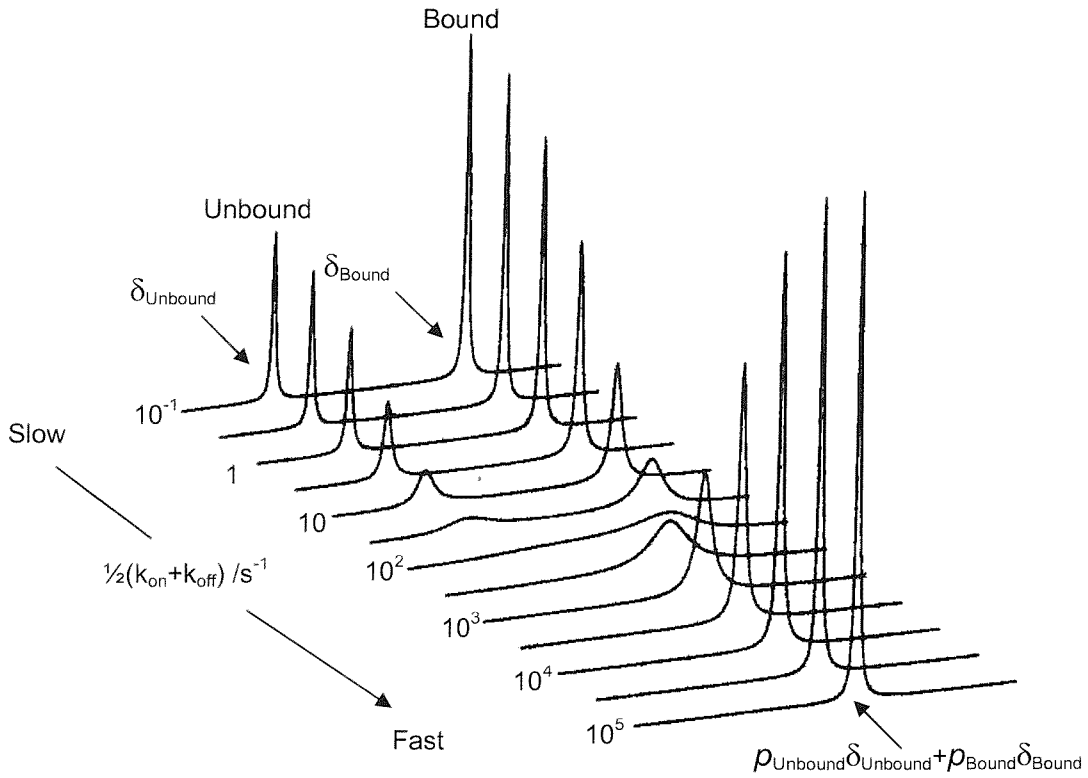


Figure 1.15 Simulated spectra of the effect of slow, intermediate and fast exchange on the resonances from the bound and unbound forms of a protein. For slow exchange two peaks are observed which relate to the bound and unbound forms. In fast exchange a single peak is observed at a weighted chemical shift between the bound and unbound forms. In intermediate exchange the peaks broaden and can even disappear from the spectrum [88].

1.11 Study of Membrane Proteins

Study of the receptors in isolation is not a trivial task due to the difficulties in expressing, purifying and handling transmembrane proteins. Even if the protein can be obtained in abundance, there are problems in structural characterisation by any technique [90]. The major limitation for those techniques that use the solid-phase, such as crystallography and microscopy, is the production of highly ordered samples. NMR spectroscopy can be used to obtain structural information in the solid, liquid-crystalline and solution phases and as such is a powerful tool. There are however limitations when using this technique such as the size of the protein, the relatively small chemical shift dispersion of residues in hydrophobic helices and the slow tumbling rate if the protein is studied in a membrane mimic, such as micelles [91].

One method of study is to synthesise peptides corresponding to various receptor domains and examine these using NMR spectroscopy [92]. It is common in the literature for the receptor domains to be defined, but closer examination often reveals that there is no

definitive proof to support these arguments and in a lot of cases little or no reasoning is given. Thus in order for the domains to be studied, they must first be identified. The first tool for the identification is the use of transmembrane prediction programs, which whilst varying slightly in their methodologies generally use hydrophobicity plots to predict the transmembrane regions. The algorithms tend to highlight the general positions of the transmembrane helices, but the outputs of several programs often differ in their predictions of the precise location. This lack of consensus on exact locations is probably in part a reflection of the diffuse nature of the membrane, which does not have a defined boundary but rather a series of probabilities. The information from the programs can be further analysed by hand using empirical methods based upon known structural information. In an analysis of high resolution structures, the amino acid proline occurred most frequently as the second residue of a helix, with almost 90% of prolines in helices occurring in the first turn [93, 94]. As the transmembrane regions of interest are believed to be helical, this information can be used to predict where the first turns occur. It has been noted that certain amino acids are more likely to be found at the membrane edge and are believed to act as anchors for the protein. The two most commonly found residues are tryptophan and tyrosine. Arginine and lysine are also commonly found at the interface. This is believed to be due to their side chains, which consist of a hydrophobic chain and positively charged head group, so they can "snorkel" within the membrane and hence give the protein some degree of flexibility in its position within the membrane [95-97].

Once the amino acid sequence of a domain has been predicted there are several ways to design the peptide. For the extracellular regions, a linear peptide corresponding to the sequence can be synthesised. Whilst linear peptides can model extracellular regions that only contact the membrane at one end, they are not a good representation of the loops where both ends contact the membrane. In order to model loops the linear peptides can be cyclized by the addition of cysteines at each end which can then be oxidised to form the disulfide bond. However, this method leads to the two ends lying close together spatially and also restricts the dynamics of the loop, which is not necessarily a true representation of the native system. Yeagle has published several papers that have modelled the extracellular regions of both bovine rhodopsin and bacteriorhodopsin using peptides [98, 99]. Examination of the peptides sequences show that they correspond not just to the extracellular region, but also to a small part of the helix on either side. In the case of rhodopsin the peptides were water soluble, although they aggregated over time and their

structure in DMSO was shown to match the more recent high resolution X-ray structure [26].

Micelles have been shown to be applicable for use in the study of membrane proteins, via NMR spectroscopy [100-102]. In recent years there have been several attempts to study the extracellular loops of the G-coupled receptors using peptides suspended in the membrane mimic n-dodecylphosphocholine (DPC) [103-110]. The results of several of these studies have enabled the study of the isolated extracellular loops with small peptides and organic molecules to be performed [105, 107, 109, 110]. To date no NMR analysis using extracellular loop mimics in DPC to study the interaction between the extracellular loop and a protein ligand has been reported.

1.12 Project Overview

CCL11 and CCL2 have a 65% amino acid sequence identity. However, CCL11 has a very high specificity for the CCR3 receptor, whereas CCL2 is unable to activate the CCR3 receptor and instead activates the CCR2 and CCR4 receptors. In contrast, CCL24 and CCL26 only have a 34-38% sequence identity to CCL11, but also have a high specificity to the CCR3 receptor when compared to other known receptors. The difference in sequence identity between CCL11, CCL24 and CCL26 makes these chemokines good comparators when studying the binding and activation mode of the chemokines.

In light of the constraints in studying membrane proteins, this project has involved the synthesis of peptides corresponding to the smaller regions of the extracellular loops of CCR3 and examination of the structure of these synthesised loops, via NMR spectroscopy in both DMSO and DPC. The peptide loops in DPC were titrated against a ^{15}N-labelled chemokine in an attempt to distinguish the sites of interaction. It is hoped that the results will shed light upon the interactions of the chemokines.

2 Design and Characterisation of Peptides Corresponding to the Extracellular Parts of CCR3

2.1 Introduction

The published results for the binding mode of CCL11 with CCR3 agree that the N-terminus of CCR3 is important in the binding of CCL11. It is not, however, the activation site. There must be at least one other site on CCR3 involved in the binding and activation and this is presumed to be extracellular [61, 111, 112]. Chimeric experiments implicated loop 3 as being involved in the binding of CCL11, but not the activation [61], whereas a study using NMR spectroscopy did not reveal any interaction between CCL11 and a peptide that corresponded to the published sequence of loop 3 [111]. Although a mimic containing both the N-terminus and loop 3 showed a decrease in binding if either section was removed [62]. The NMR study also could not detect any interaction between CCL11 and a peptide corresponding to the published sequence of loop 1 [111].

In contrast, the titration of CCL11 with a peptide whose sequence corresponded to loop 2 resulted in a precipitate which was suggested by the authors as being caused by an interaction [111]. The NMR study used two different forms of the peptides; one type was linear, the other was cyclised by a disulfide bond formed from cysteines that were added to either end of the peptide. Although both forms of the peptides gave the same results, the sequences only corresponded to the proposed extracellular parts of CCR3 and the peptide structures were not published. There was also no information that suggested these were mimicking the structure of the extracellular loops *in vivo*.

Whilst the number of transmembrane proteins being solved to atomic resolution is increasing there is still a lack of published structural information for the G-coupled receptors, with the only known structures being rhodopsin [26, 113] and bacteriorhodopsin [114]. In both cases Yeagle has shown that peptides corresponding to the extracellular loops could form independent turns equivalent to those shown in the crystal structure [98, 115] and in the case of rhodopsin this was achieved before the publication of the crystal structure. The peptide sequences used were slightly longer than those of the published extracellular loops. Consequently, it was proposed that if the peptides used to model CCR3 in the NMR study were extended to incorporate the amino acids either side, beyond

the published extracellular loop sequences of CCR3, then the peptides might adopt a more accurate fold in solution and generate a native interaction with the chemokines.

2.2 Design of First Generation Extracellular CCR3 Loop 1

The sequence shown in Figure 2.1 was designed by Dr M.Crump to mimic loop 1 and was synthesised by Southampton Polypeptides using solid phase t-BOC chemistry. The N-terminus was acetylated and the C-terminus was produced as the amide as these mimic the peptide bond.

^{91}IHYVRGHNW^{100}VFGH^{105}GM

Figure 2.1 The amino acid sequence of the first generation peptide synthesised to mimic loop 1. The numbering refers to the positions of the amino acids on CCR3. The peptide was synthesised by t-BOC chemistry with the N-terminus acetylated and the C-terminal produced as an amide.

2.2.1 NMR Study of the First Generation Extracellular CCR3 Loop 1

The experiments were performed on a 6mM peptide sample (20mM sodium acetate, 1mM sodium azide, 10% D_2O pH5.09). At 30°C the 2D experiments COSY, TOCSY and NOESY were run. Both the COSY and TOCSY spectra were run using the standard experiments and whilst most of the resonances could be assigned to spin-systems using the TOCSY, very few could be assigned to residues. The COSY was barely usable due to its low resolution, poor signal to noise and poor water suppression, which was partially a fault of the pulse sequence itself. Disappointingly the NOESY resulted in very few NOEs despite experiments being run using mixing times of 0.15, 0.3, 0.4, and 0.5s. This lack of NOE build up could have been due to the size of the peptide (1.8kDa) which might have placed the molecule in the region where there is very little NOE build-up (section .3) so a ROESY with a mixing time of 75ms was acquired. This also resulted in very few NOE peaks. With no through space information the spectra were unassignable.

Due to these problems the temperature was lowered to 8°C, in the hope that this would reduce the molecular tumbling rate sufficiently to allow NOE build up. The NOESY was run with a mixing time of 0.4s and showed a significant improvement in the build up of

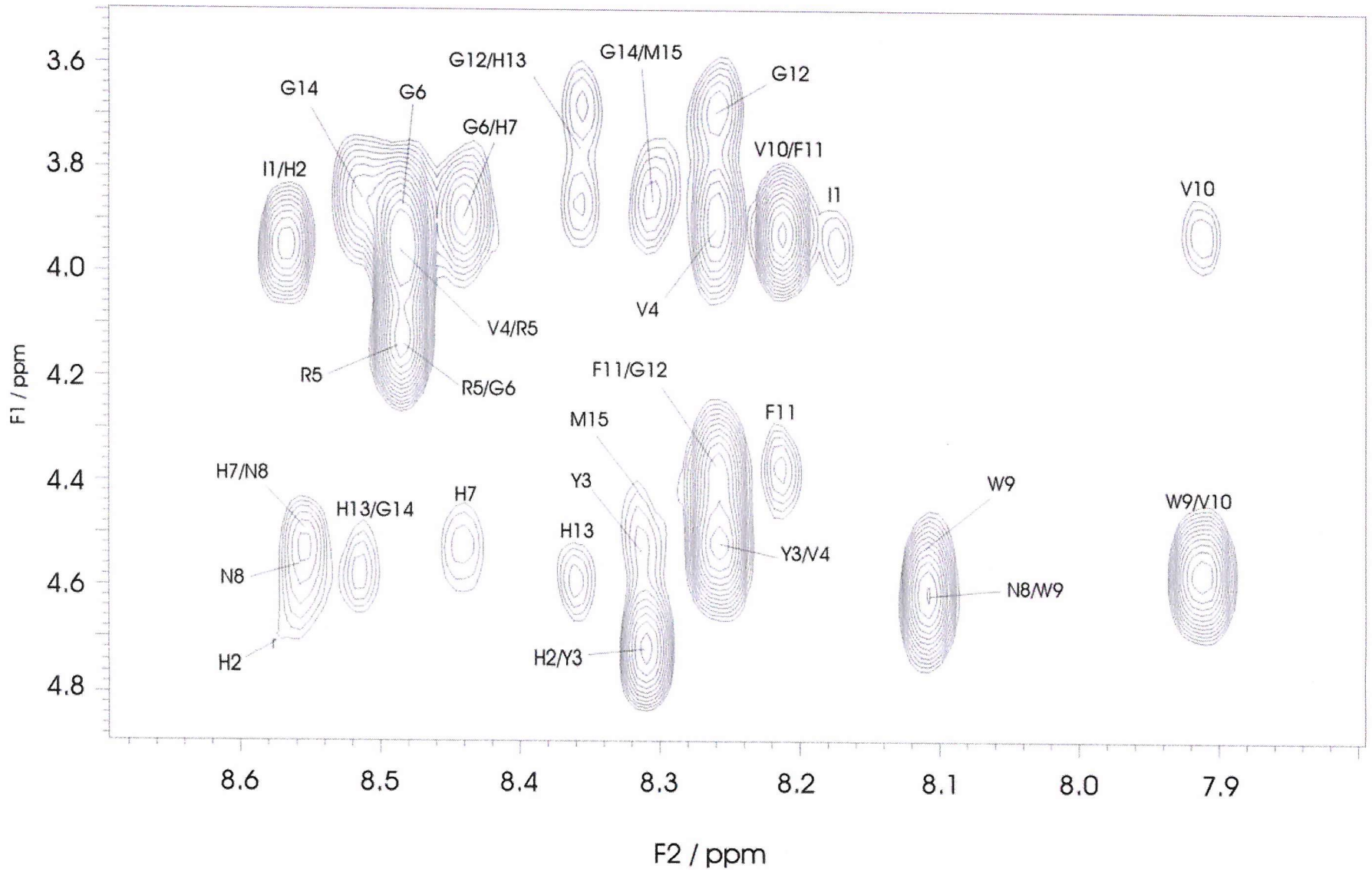


Figure 2.2 The first generation loop 1 NOESY spectrum (mixing time of 0.4s) at 8°C showing the region between 7.8 to 8.7ppm and 3.5 to 4.9ppm. The spectrum had 4096 points in F2 and zero filled from 128 to 2048 points in F1. The Varian window functions were LB (-5), GF (0.04), LB1(-2) and GF1(0.006).

NOEs, so a TOCSY was acquired using the standard experiment. The resulting spectra were assigned and this enabled the assignment of the TOCSY at 30°C.

The NOESY at 8°C showed 272 NOEs of which 66 were sequential and 6 were tentatively assigned as non-sequential. The remainder of the NOEs were intraresidue. The six non-sequential NOEs were Y3HE#:R5HA, Y3HD#:V10HG#, H7HD:V4HG#, W9HB#:R5HB#, W9HE1:H2HB# and W9HE1:H2HB#. If these assignments were correct then the presumed loop structure would start approximately at Y3 and end at V10. However, the spectra were not recorded at a very high resolution or good signal to noise and the non-sequential peaks were very weak and close to the noise level. The presence of an impurity in the sample also meant that the assignment of non-sequential NOEs had to be treated with care. The most convincing peak was W9HB#:R5HB#, but this cross-peak was only observed in the higher resolution F2 dimension and not the F1. Due to the uncertainty in the assignment of the NOEs, it was not possible to draw any conclusion about the structure without further analysis.

2.2.2 Analysis using Deviation from Random Coil

As discussed in section 1.9 the secondary structure of a protein can be predicted using the chemical shift index. The results of applying the chemical shift index to this peptide (Figure 2.3) suggested at 8°C that there was a helix from R5 to M15, which was not present at 30°C.

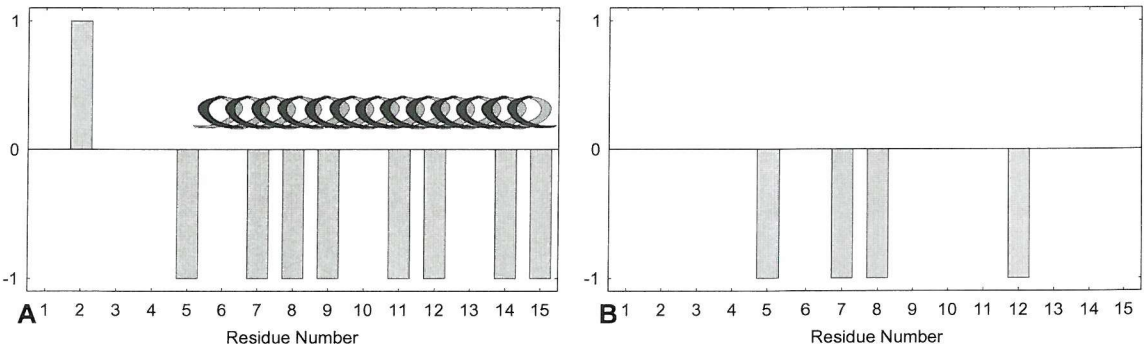


Figure 2.3 The chemical shift index analysis for loop 1. A. Analysis at 8°C which predicted a helix from R5 to M15. B. Analysis at 30°C which predicted no secondary structure.

The Andersen plot is a more reliable method for gaining structural information from the chemical shifts of peptides than the chemical shift index. For this peptide the Andersen

plot was constructed using the data collected at 8°C and 30°C (Figure 2.4). The plot did not give any justification for structure as there was not an amide with a significant CSD (0.3ppm) and the correlation coefficient was only 0.1246 rather than the 0.7 required.

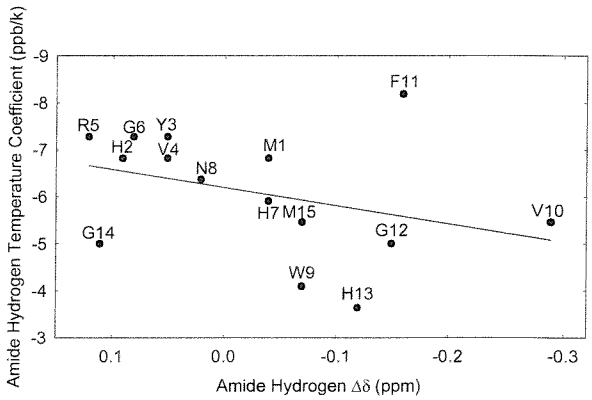


Figure 2.4 The Anderson plot for loop 1. This analysis plots the amide proton CSD against the amide proton temperature coefficient. The line is the least-squares linear regression. For peptides these plots can be used as a diagnostic tool for partial structuring when they have a correlation coefficient greater than 0.7 and amide proton CSD greater that 0.3. The plot showed that there was not an amide proton CSD greater than 0.3 and the correlation coefficient was only 0.1246.

2.2.3 Conclusion

The occurrence of six very weak peaks in the NOESY that could be assigned as non-sequential NOEs suggested the peptide might sample a folded conformation, although the presence of impurities and the poor quality of the spectra did not make a convincing argument. Analysis of the chemical shift data using the chemical shift index showed a possible alpha helix at 8°C, which was not present at 30°C. The Andersen plot which is a more reliable method of predicting structure in peptides than the chemical shift index, failed to give any corroborating evidence. This suggests that if the NOEs are the product of a real interaction, then the interaction must be very weak, and that there is not a single folded species in solution.

2.3 The N-terminus of CCR3

A peptide corresponding to the first thirty-five amino acids at the N-terminus of CCR3 had already been characterised [111]. The same peptide was studied here as a training exercise in assignment and with the view of using it in future work. The sequence (Figure 2.5) was synthesised by Southampton Polypeptides using solid phase t-BOC chemistry. The N-

terminus was unprotected as acetylation resulted in purification problems. The C-terminus was produced as the amide as this mimics the peptide bond.

1 10 20 30 35
MTTSLDTVETFGTTSYYDDVGLLSEKADTRALMAQ

Figure 2.5 The amino acid sequence of the peptide synthesised to mimic the N-terminus of CCR3 with a mutation of C24S. The numbering refers to the positions of the amino acids on CCR3. The peptide was synthesised by t-BOC chemistry. The N-terminus was unprotected and the C-terminus produced as an amide.

2.3.1 NMR Study of the N-terminus of CCR3

The experiments were performed on a 3mM peptide sample (20mM sodium acetate, 1mM sodium azide, 10% D_2O pH5.09). The 2D experiments COSY, TOCSY and NOESY were recorded using the standard pulse sequences at both 8°C and 30°C. At both temperatures the spectra were similar and assignable. At 8°C the NOESY spectra showed 274 intraresidue, 144 sequential and 45 short range NOEs. There were no long range NOEs.

As with loop 1 the chemical shift data was analysed by both the chemical shift index (Figure 2.6) and the Andersen plot (Figure 2.7). The chemical shift index suggested there were two helices at 8°C (T2 to T14 and V20 to E25) which disappeared at 30°C. The Andersen plot does not support this though with there being only one amide proton with a significant CSD (0.3ppm) and the correlation coefficient was only 0.06 rather than the 0.7 required.

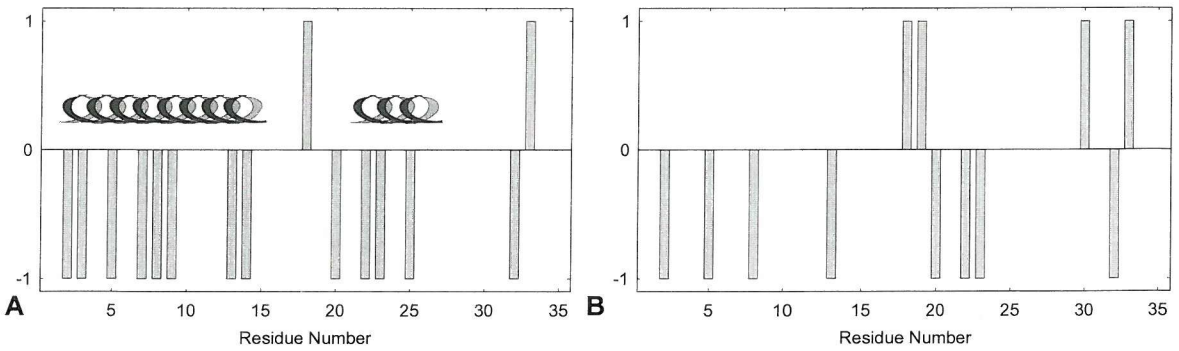


Figure 2.6 The chemical shift index analysis for the N-terminus. A. Analysis at 8°C predicting a helix from T2 to T14 and V20 to E25. B. Analysis at 30°C which does not predict any secondary structure.

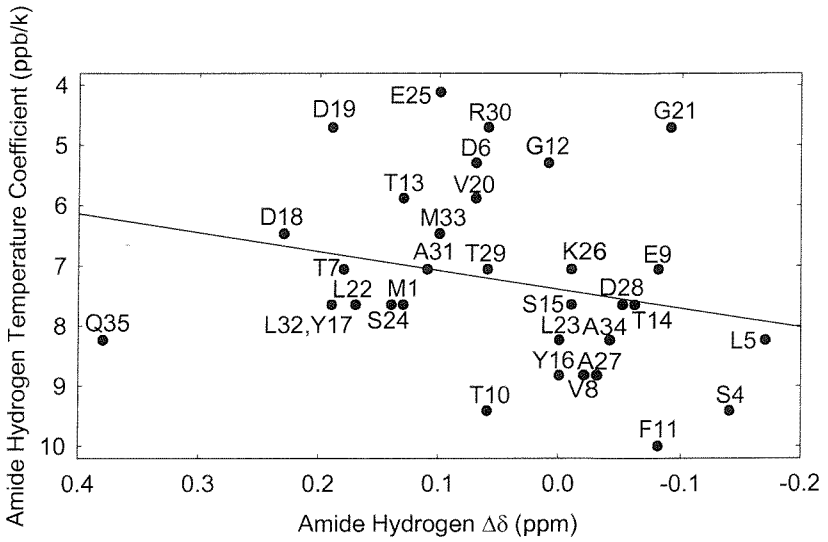


Figure 2.7 The Anderson plot for the N-terminus. This analysis plots the amide proton CSD against the amide proton temperature coefficient. The line is the least-squares linear regression. For peptides these plots can be used as a diagnostic tool for partial structuring when they have a correlation coefficient greater than 0.7 and amide proton CSD greater that 0.3. The plot showed that there was only one amide proton CSD greater than 0.3 and the correlation coefficient was only 0.06.

2.3.2 Conclusion

Whilst there were some short range NOEs suggesting a secondary structure, which was supported by the chemical shift index analysis, the Andersen plot failed to give any evidence of structure. As such the analysis of the N-terminus seemed to suggest that there was no tertiary structure, although there might be a small amount of secondary structure. This largely agrees with the previously published results where only two non-sequential NOEs were observed, one between V8:F11 and the other between V20:L22 [111]. These two NOEs would relate the two helices predicted at 8°C by the chemical shift index.

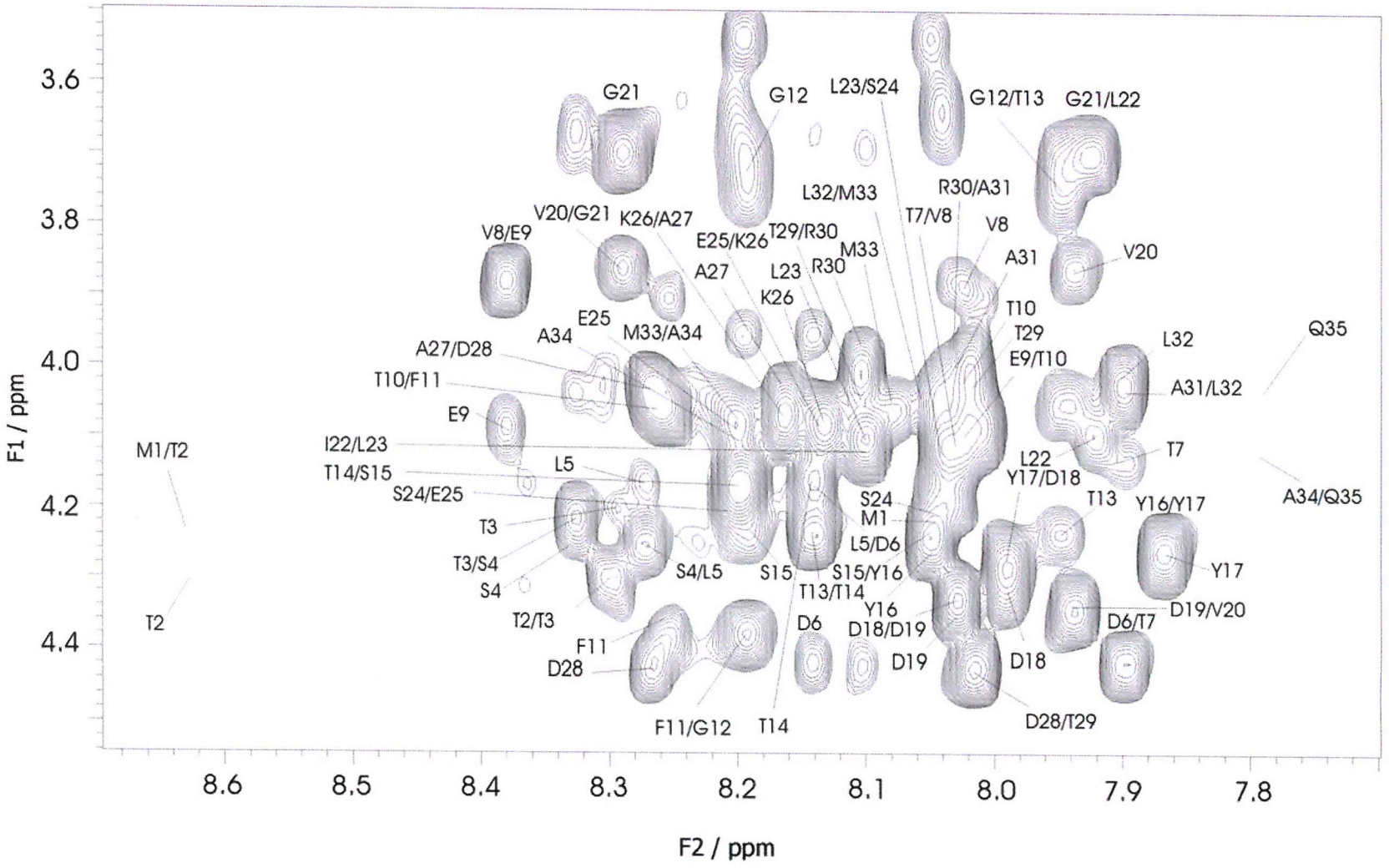


Figure 2.8 The N-terminus NOESY spectrum (mixing time of 0.4s) at 8°C showing the region between 7.7 to 8.7ppm and 3.5 to 4.55ppm. The spectrum had 32000 points in F2 and zero filled from 256 to 2048 points in F1. The Varian window functions were LB (0), GF (0.06), LB1(-10) and GF1(0.016).

2.4 Loop Redesign

With the first generation design for loop 1 failing to produce a convincing loop structure the problem was reassessed. The published sequences of CCR3 were derived from DNA database screens without any consideration given to post-translational modification. There was no signalling sequence given and it is unclear whether these proteins should have one, as examination of the published family members shows only CCR7 being accredited with such a sequence. From a sequence alignment (Figure 2.9) this would correspond to residues 1 to 7 on CCR3 which might explain why a peptide corresponding to a truncated receptor N-terminus (residues 8-35) gave a very similar K_d for binding CCL11 as the full length N-terminus [111]. It was also noted that the extracellular regions were stated, but there was no published methodology so it was unclear how these regions had been chosen [116-118].

There are several transmembrane prediction programs freely available on the worldwide web [119] and the sequence of CCR3 was analysed using TMpred [120], SOSUI [121-124], HMMTOP [125] and TMAP [126, 127]. When the outputs from these programs were compared with each other (Figure 2.10) it could clearly be seen that whilst they agreed on the general location of the extracellular regions they differed in the precise position and length. The only major difference in the outputs was that TMAP predicted that the C-terminus would be extracellular rather than intracellular.

The alignment of all CC chemokine specific receptors (Figure 2.9) also highlighted the difference in the proposed extracellular regions, with there being an agreement on the general location, although not the precise position and length.

The proposed method of study relies upon designing peptides whose sequences correspond to the extracellular regions of CCR3 and a short stretch of transmembrane helix at either end. In order to accomplish this it is obvious that the precise position and length of the extracellular loops must be identified. As already discussed in section 1.11 this is possible by examining the sequence around the predicted location of the extracellular loops and identifying the membrane anchors. The amino acids known to act as membrane anchors are tryptophan, tyrosine, arginine and lysine [95-97]. The presence of proline is also important because 90% of prolines in helices occur in the first turn [93, 94].

To take advantage of the study by Yeagle on bovine rhodopsin [99] its sequence was aligned with CCR3 (Figure 2.11). This alignment indicated the areas that corresponded to those peptides which had formed independent turns in solution. It should be noted that whilst these proteins are both G-coupled and should therefore have similar intracellular regions they react to different stimuli, with bovine rhodopsin's stimulus being the change in conformation of retinal upon activation with light, whilst a CC chemokine receptor's stimulus is the binding of a CC chemokine. So despite both proteins belonging to the same family, their transmembrane and extracellular loops could be very different. It was hoped that examination of the published structures for the peptides by Yeagle [99] would indicate the number of turns of the transmembrane helix used and allow for a judgement on how many residues should be added to each end of the loop.

Utilising all this information the following loops were designed.

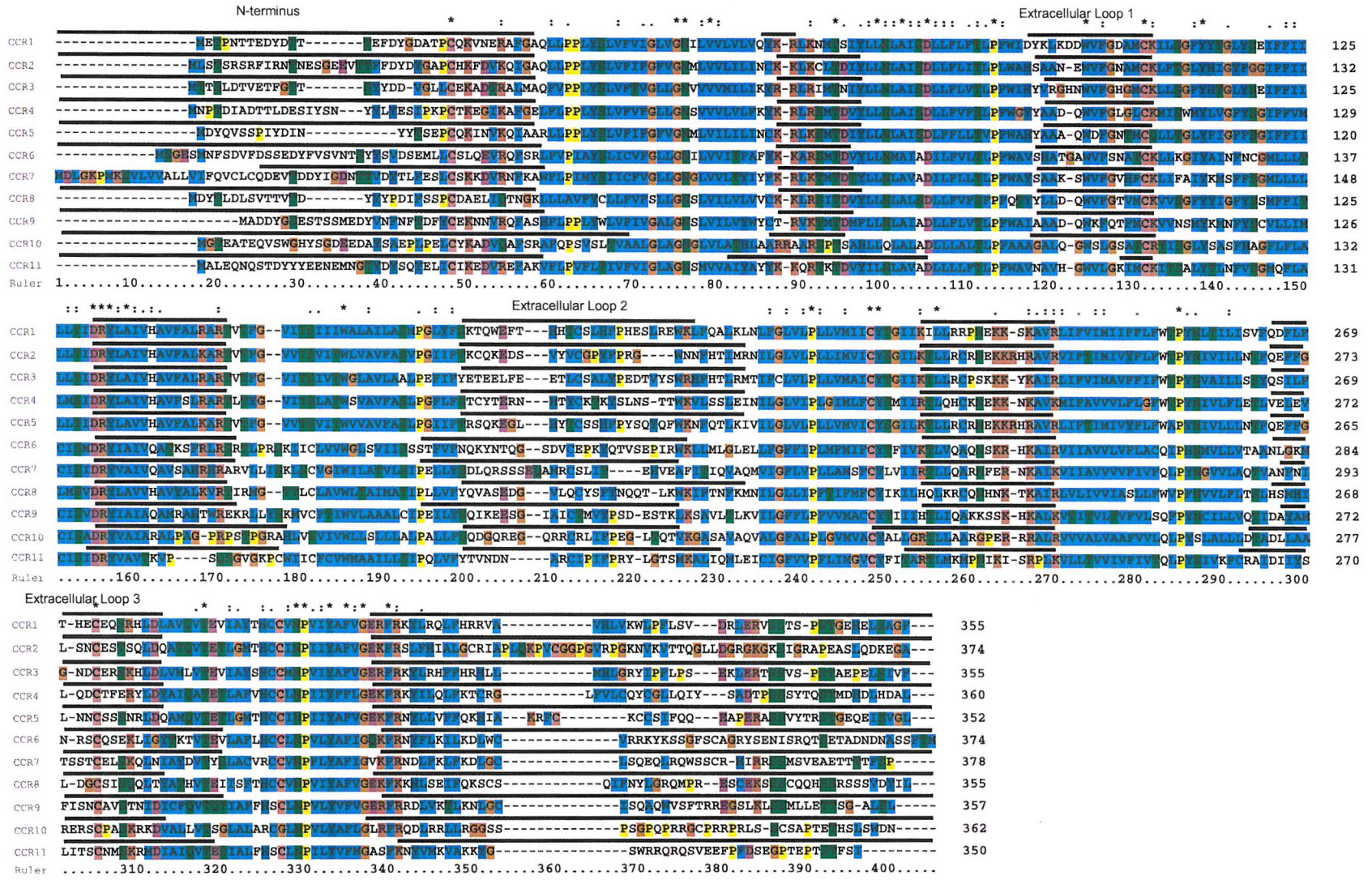
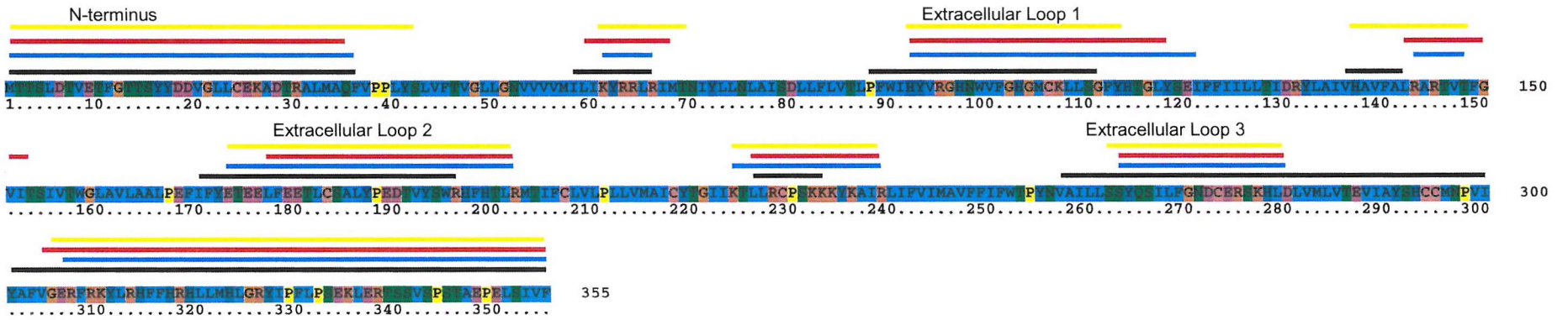


Figure 2.9 Alignment of the CC chemokine receptors, showing their published intracellular and extracellular regions which are marked by the black lines [116, 128-137].



Lines above sequence:

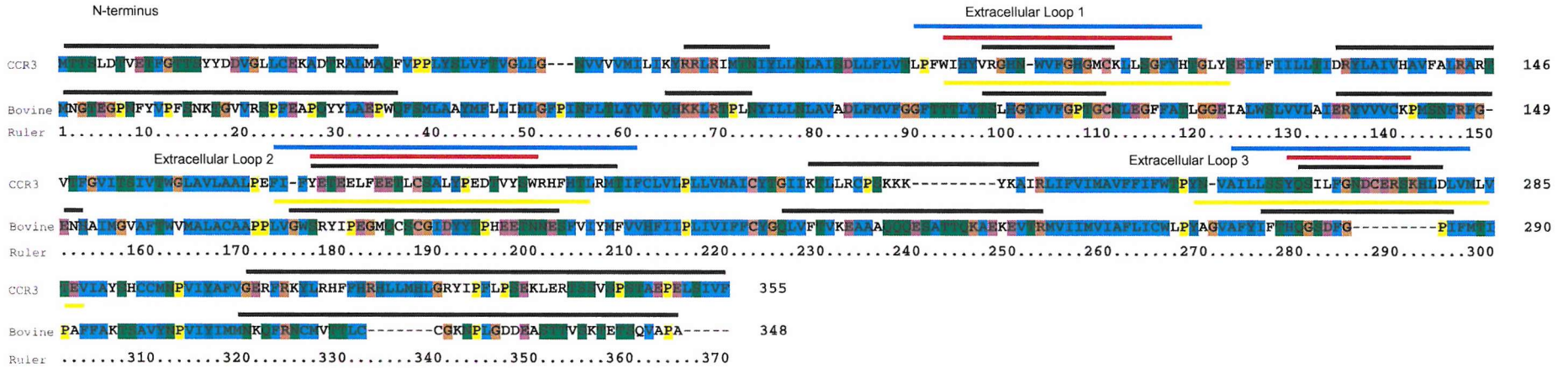
Black: TMAP.

Blue: HMMTOP.

Red: SOSUI.

Yellow: TMpred.

Figure 2.10 The proposed intracellular and extracellular regions of CCR3 from four transmembrane prediction programs [120-127].



Lines above sequence:

Black: Published extracellular and intracellular loops (for bovine rhodopsin these are from the crystal structure).

Blue: Sequence synthesised into peptides.

Red: Proposed extracellular loops.

Yellow: Sequence used in Yeagle's peptides.

Figure 2.11 Alignment of CCR3 against Bovine Rhodopsin showing Yeagle's peptides for comparison to those synthesised in this work [26, 115, 118].

2.4.1 Extracellular Loop 1

The alignment of the amino acid sequences of the CC chemokine receptors (Figure 2.12) showed that P88 was conserved throughout the series and W90 was partially conserved. From this information it was proposed that P88 formed the C-terminus of transmembrane helix 2 and W90 might act as the membrane anchor at the N-terminus of loop 1. The C-terminus of loop 1 might have been the partially conserved K107, but the prediction programs all indicated that the C-terminus was further along the sequence (Figure 2.13). Y113 was fully conserved so this was chosen as the membrane anchor at the C-terminus of loop 1.

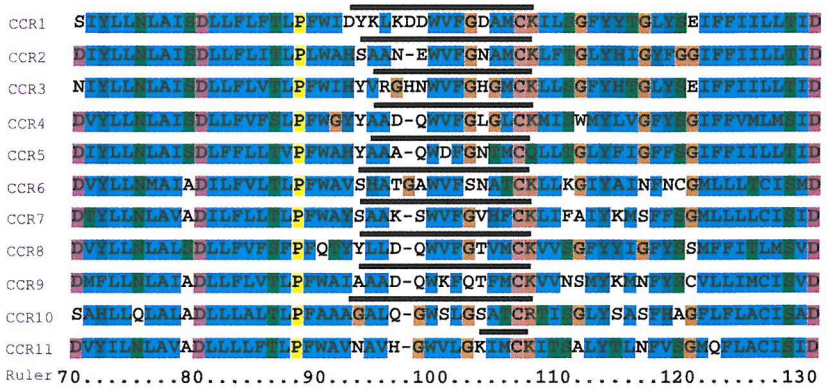
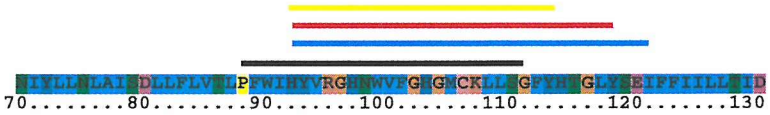


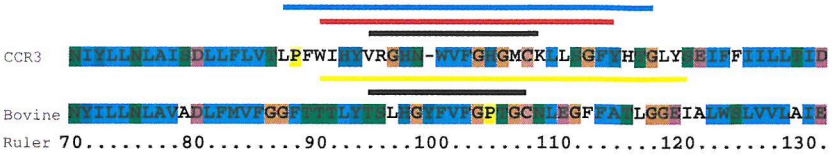
Figure 2.12 Alignment of CCR receptors showing the regions that corresponded to N70-130 on CCR3. The black line above each sequence shows the published extracellular loop 1 [116, 128-137].



Lines above sequence:
Black: TMAP.
Blue: HMMTOP.
Red: SOSUI.
Yellow: TMpred.

Figure 2.13 Extracellular loop 1 of CCR3 as predicted using data from four transmembrane prediction programs [120-127].

This analysis suggested that loop 1 would extend from W90 to Y113. The addition of three amino acids at each end to mimic one turn of the transmembrane helix would give a sequence of 30 amino acids (Figure 2.16A). This would be of a similar length to the peptide used by Yeagle (Figure 2.14) [115] but more balanced around the loop; as can be seen from Figure 2.15 Yeagle's peptide was lop-sided on the C-terminal end.



Lines above sequence:

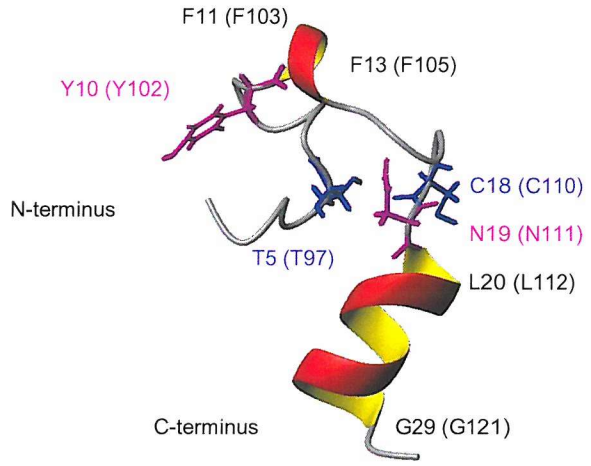
Black: Published intracellular loops 1 (for bovine rhodopsin this is from the crystal structure).

Blue: Sequence synthesised into a peptide.

Red: Proposed extracellular loop.

Yellow: Sequence used in Yeagle's peptide.

Figure 2.14 Alignment of CCR3 against bovine rhodopsin showing the region between N70 and D130 on CCR3 [26, 115, 118].



Black: Other residues mentioned in the text.

Blue: Last residues of the loop as taken from the crystal structure.

Purple: The end of the helices as defined by Yeagle.

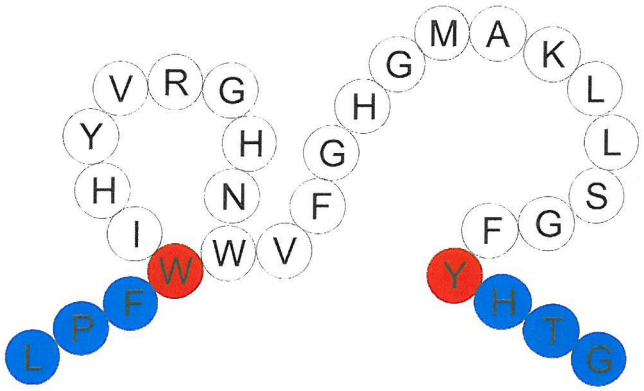
Figure 2.15 Structure of Yeagle's peptide corresponding to extracellular loop 1 of bovine rhodopsin (T93-I123) [115] (PDB ID: 1EDS). Yeagle defines this as a helix turn helix with the first helix ending at Y10 and then a turn to the second helix starting at N19. The modelling program defines the N-terminus as being random coiled for the first three residues and then a series of bends and turns to the first helix from F11 to F13. In contrast, the C-terminus is random coiled for the last two residues before forming a helix running from G29 to L20. Between the two helices there is a set of bends and turns.

The sequence (Figure 2.16A) was synthesised by Southampton Polypeptides using solid phase t-BOC chemistry. The N-terminus was unprotected and the C-terminus was produced as the amide as this mimics the peptide bond. The NMR spectra showed the peptide to be the correct sequence and to be pure.

Loop 1 would contain a pair of tryptophans separated by eight amino acids which can cause a peptide to form a β-hairpin via a tryptophan zipper [138]. This would mean that the loop may in fact form two loops (Figure 2.16B).

A

B



Blue: Residues believed to be from the transmembrane helix.

Red: Residues believed to be the membrane anchors.

Figure 2.16 A. The amino acid sequence of the peptide synthesised to mimic extracellular loop 1. The numbering refers to the positions of the amino acids on CCR3. B. A graphical representation of the proposed structure of the loop.

2.4.2 Extracellular Loop 2

The alignment of the amino acid sequences of the CC chemokine receptors (Figure 2.17) demonstrated that P167 was almost fully conserved, but there appeared to be no conserved membrane anchor. Careful examination of the alignment showed that there was a possible membrane anchor for each receptor within a few amino acids of each other. So despite not aligning perfectly it was proposed that Y172 acts as the N-terminal anchor of loop 2 for CCR3.

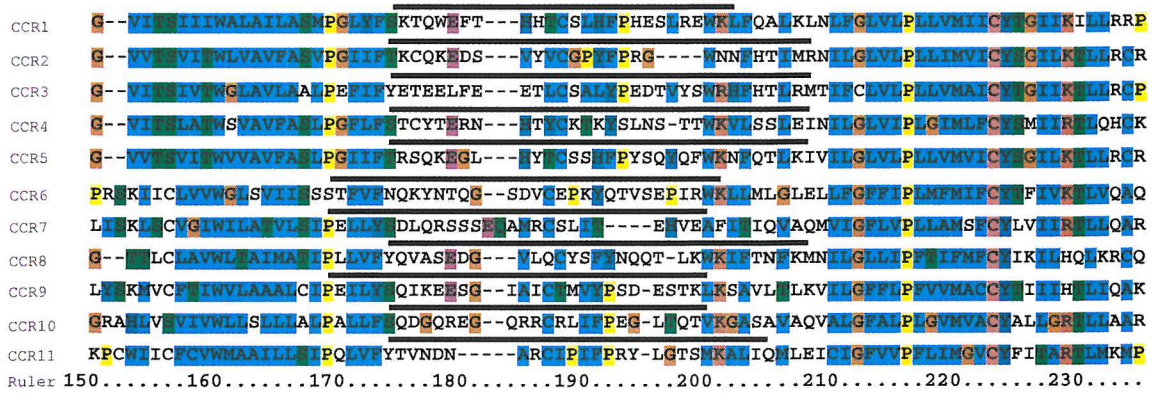
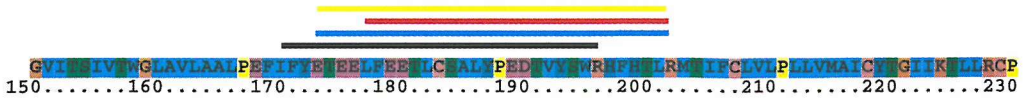


Figure 2.17 Alignment of CCR receptors showing the regions that corresponded to G150-P230 on CCR3. The black line above each sequence shows the published extracellular loop 2 [116, 128-137].



Lines above sequence:

Black: TMAP.

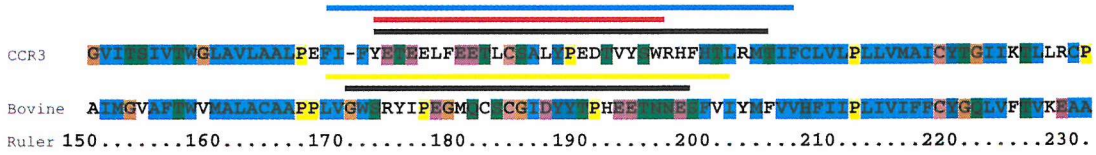
Blue: HMMTOP.

Red: SOSUI.

Yellow: TMpred.

Figure 2.18 Extracellular loop 2 of CCR3 as predicted using data from transmembrane prediction programs [120-127].

Two amino acids were highlighted as possible C-terminus membrane anchors for loop 2 of CCR3 (Figure 2.18). Three of the prediction programs chose R202, and one R196. R196 aligned with a partially conserved lysine and W195 is also partially conserved, so R196 was a logical choice (Figure 2.17). However, ending the loop at R196 would lead to a very long transmembrane helix and the receptor alignment showed that the amino acids proceeding R196 were particularly unconserved. The proposed hypothesis was that loop 2 acted as a signalling loop section 1.5.4. Upon binding of the chemokine the C-terminus membrane anchor could possible change between R196 and R202.



Lines above sequence:

Black: Published intracellular loops 2 (for bovine rhodopsin this is from the crystal structure).

Blue: Sequence synthesised into a peptide.

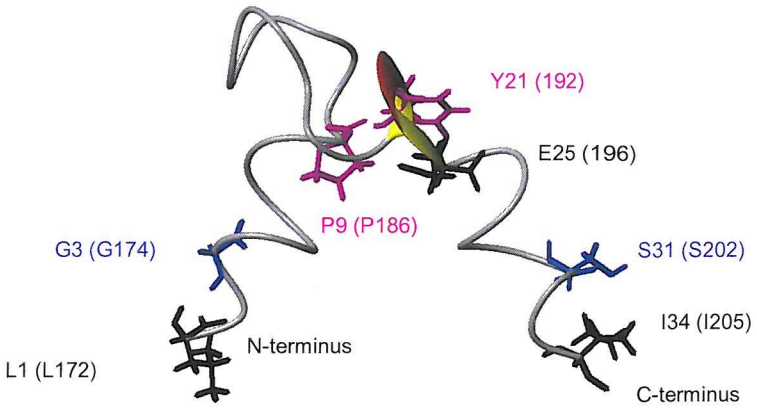
Red: Proposed extracellular loop.

Yellow: Sequence used in Yeagle's peptide.

Figure 2.19 Alignment of CCR3 against bovine rhodopsin showing the region between G150 and P230 on CCR3 [26, 115, 118].

This analysis suggested that loop 2 would extend from Y172 to R202. The addition of three amino acids at the N-terminus and four amino acids to the C-terminus to mimic one turn of the transmembrane helix would give a sequence of 38 amino acids (Figure 2.21A).

Analysis with Yeagle's structure (Figure 2.20) was not particularly insightful. It was possible to imagine that residues L1 to P9 formed a helix and the same for residues E25 to I34 however this was not picked up by the modelling program.



Black: Other residues mentioned in the text.

Blue: Last residues of the loop as taken from the crystal structure.

Purple: The end of the helices as defined by Yeagle.

Figure 2.20 Structure of a Yeagle's peptide corresponding to extracellular loop 2 of bovine rhodopsin (L172-I205) [115] (PDB ID: 1EDV). Yeagle defines this as a helix turn helix with the first helix ending at P9 and then a turn to the second helix starting at Y21. The modelling program defines the N-terminus as being random coiled for the first two residues and then a series of bends and turns until a helix between Y21 and H24. There is then another series of bends and turns until the last two residues which are random coiled.

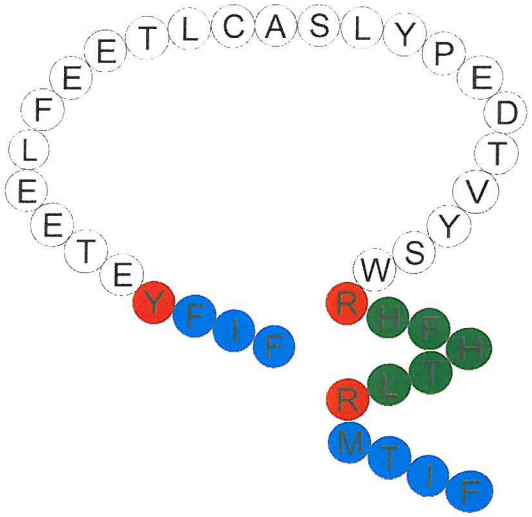
The sequence (Figure 2.21A) was synthesised by Southampton Polypeptides using solid phase t-BOC chemistry. The C-terminus was produced as the amide as this mimics the peptide bond and the N-terminus was unprotected.

A

169 180 190 200 206

FIFYETEELFEETLASALYPEDTVYSWRHFHTLRMTIF

B



Blue: Residues believed to be from the transmembrane helix.

Green: Residues believed to translocated upon binding.

Red: Residues believed to be the membrane anchors.

Figure 2.21 A. The amino acid sequence of the peptide synthesised to mimic extracellular loop 2. The numbering refers to the positions of the amino acids on CCR3. B. A graphical representation of the proposed structure of the loop.

2.4.3 Extracellular Loop 3

The alignment of the amino acid sequences of the CC chemokine receptors (Figure 2.22) showed P254 to be fully conserved, followed by a partially conserved tyrosine. This was unlikely to be the N-terminus membrane anchor as the residues following were all hydrophobic by nature. The sequence alignment does not reveal any other conserved membrane anchors although Y264 was a possibility for CCR3. However the sequence alignment did show a strange set of residues aligning with S262, suggesting this residue might be of importance. So whilst it was believed that Y264 was the membrane anchor, the design started from S262.

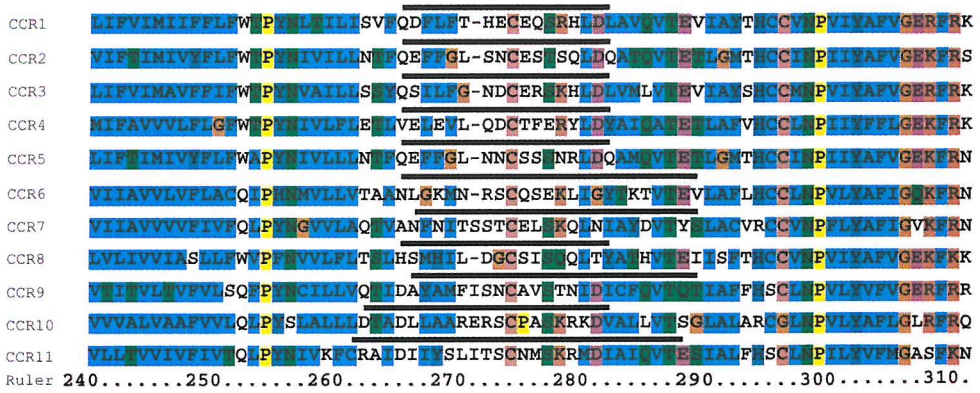
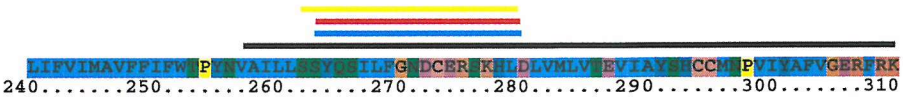


Figure 2.22 Alignment of CCR receptors showing the regions that corresponded to L240-K310 on CCR3. The black line above each sequence shows the published extracellular loop 3 [116, 128-137].

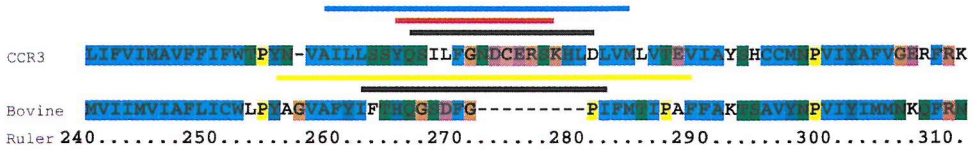
K277 was partially conserved and would be a logical choice to be the C-terminal membrane anchor on first inspection. However, all the transmembrane prediction programs tested terminated the loop at D280 (apart from TMAP) (Figure 2.23). As aspartic acid is hydrophilic it would be strange if it was to be found just inside the membrane. Therefore D280 was chosen as the C-terminus for loop 3.



Lines above sequence:
Black: TMAP.
Blue: HMMTOP.
Red: SOSUI.
Yellow: TMpred.

Figure 2.23 Extracellular loop 3 of CCR3 as predicted using data from transmembrane prediction programs [120-127].

While this analysis suggested that the membrane anchors for loop 3 would make the loop extend from Y264 to K277, the unconserved nature of sequence alignments suggested that the interesting aspect of this loop might actually stretch from S262 to D280. As such these where chosen as the outer points before the addition of four amino acids at the N-terminus and three amino acids to the C-terminus to mimic the turns of the transmembrane helix. This gave a sequence of 26 amino acids (Figure 2.26) which was much shorter than the corresponding peptide by Yeagle.



Lines above sequence:

Black: Published intracellular loops 3 (for bovine rhodopsin this is from the crystal structure).

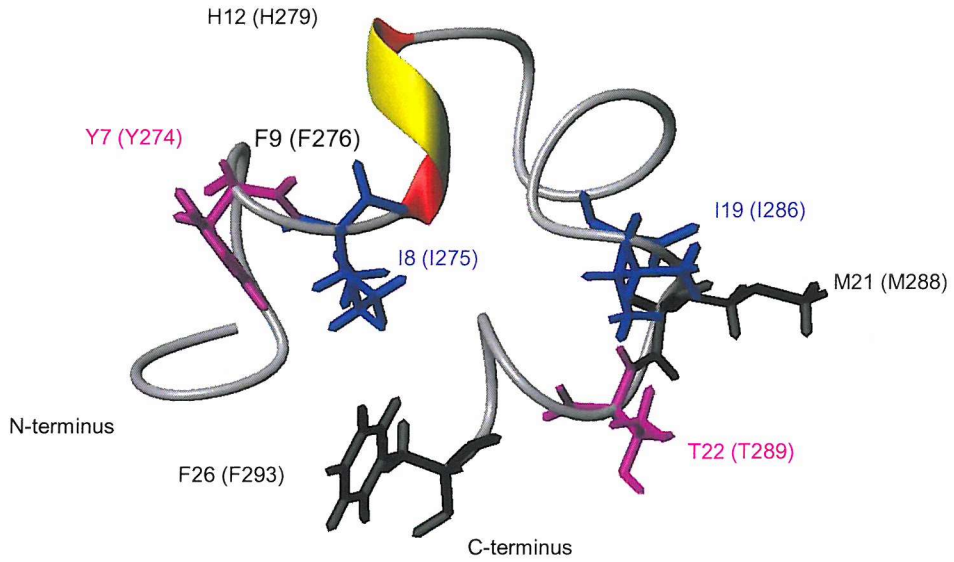
Blue: Sequence synthesised into a peptide.

Red: Proposed extracellular loop.

Yellow: Sequence used in Yeagle's peptide.

Figure 2.24 Alignment of CCR3 against bovine rhodopsin showing the region between L240 and K310 on CCR3 [26, 115, 118].

Analysis of Yeagle's structure (Figure 2.25) was again not particularly insightful. It was possible to imagine a helix from residues M21 to F26 on the C-terminal end however the N-terminal end showed nothing convincing until residues F9 to H12 which are already in the extracellular region.



Black: Other residues mentioned in the text.

Blue: Last residues of the loop as taken from the crystal structure.

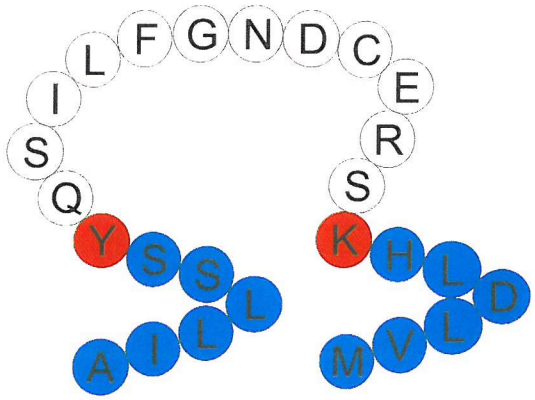
Purple: The end of the helices as defined by Yeagle.

Figure 2.25 Structure of Yeagle's peptide corresponding to extracellular loop 3 of bovine rhodopsin (Y268-F293) [115] (PDB ID: 1EDW). Yeagle states this as a defined structure and whilst mentioning that a peptide corresponding to Y7 to T22 (Y274-T289) has the same turn, but not the helices, the paper never actually states where the helices are. The modelling program defines the N-terminus as forming bends and turns to a helix which is from F9 to H12. The rest of the peptide is bends and turns except the last two residues which are random coiled.

The sequence (Figure 2.26) was synthesised by Southampton Polypeptides using solid phase t-BOC chemistry. The C-terminus was produced as the amide as this mimics the peptide bond and the N-terminus was unprotected. The NMR spectra showed the peptide to be the correct sequence and pure.

A

B



Blue: Residues believed to be from the transmembrane helix.

Red: Residues believed to be the membrane anchors.

Figure 2.26 A. The amino acid sequence of the peptide synthesised to mimic extracellular loop 3. The numbering refers to the positions of the amino acids on CCR3. B. A graphical representation of the proposed structure of the loop.

2.5 Solubilization Trials of the Redesigned Loops

Before the NMR experiments could be initiated, the loops were tested for their solubility in various common solvents (Table 2.1). Disappointingly none of the loops were water soluble and whilst loops 1 and 3 seemed to be soluble in several systems, loop 2 was only solubilized by DMSO. DMSO had previously been shown to be an acceptable solvent for the study of membrane loops [98] and as the loops where not water soluble, DMSO was chosen as the solvent for the initial experiments.

Solvent	Loop 1	Loop 2	Loop 3
100mM Sodium Acetate pH5.5	No	No	No
100% H_20	No	No	No
100% TFE	Yes	No	Yes
10% TFE 90% H_20	Yes	No	Yes
100% Methanol	Yes	No	Yes
10% Methanol 90% H_20	Yes	No	No
100% DMSO	Yes	Yes	Yes

Table 2.1 The solubility of the peptides was tested in various solvents. The results showed that none of the peptides were soluble in the aqueous phase and the only common solvent was DMSO.

2.6 NMR Spectroscopy of the Redesigned Loops in DMSO

2.6.1 Loop 1

The TOCSY and NOESY spectra obtained at 25°C in DMSO showed that there were impurities in the sample, including a water streak despite the use of fresh DMSO-d_6 from vials. As the impurities with the exception of water were not present when a sample of DMSO-d_6 was examined using the same pulse sequences, it was assumed that these impurities were from the peptide synthesis. One of the impurities was proposed to be *m*-cresol from the chemical shift data, which was used as a scavenger during the cleavage of the peptide from its solid support. The assignment was completed (Figure 2.29) but due to the impurities present, the NOEs (particularly the non-sequential ones) were confirmed using spectra recorded at a different temperature (20°C). Where the non-sequential NOEs moved with the corresponding intraresidue NOEs they were assigned as real.

The NOESY spectrum had one long range NOE (F15 to A20), 60 short range NOEs, 156 sequential NOEs as well as the intraresidue NOEs. A summary of the sequential and short range NOEs is shown in Figure 2.27.

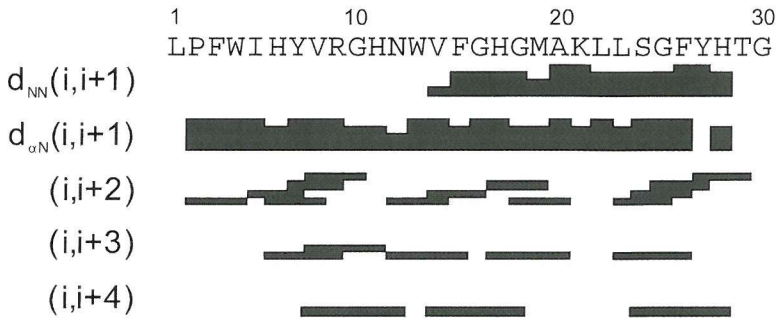


Figure 2.27 Summary of the sequential and short range NOEs for loop 1 in DMSO. The amino acid sequence and numbering are shown at the top. Sequential N-N and α-N NOEs are indicated by black bars, the thickness of the bar represents the strength of the observed NOE. The presence of short-range NOEs is indicated by a solid line.

150 structures were generated using the NOEs (Figure 2.28) however these showed that whilst there is a curved nature to the ensemble, it is almost identical to the structures generated for the sequence without any NOE restraints.

Figure 2.28 Ensemble of 150 structures generated for loop 1 in DMSO. These structures are almost identical to those which could be generated for the sequence with no NOE restraints.

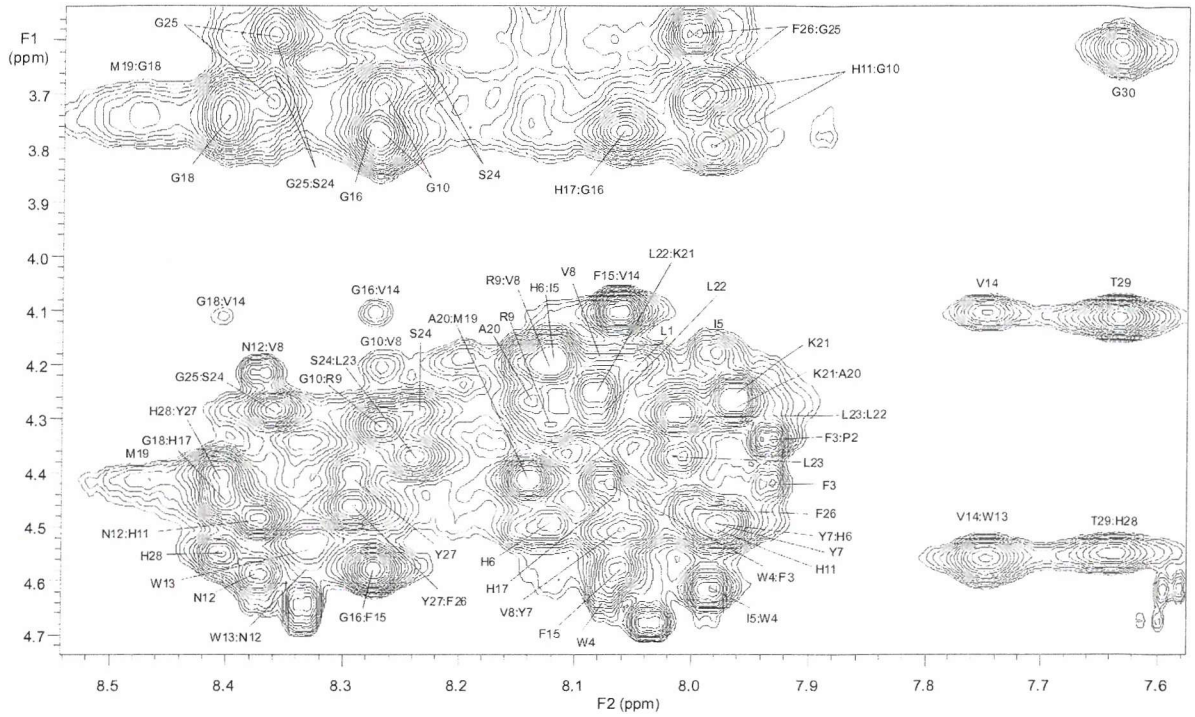


Figure 2.29 The NOESY spectrum (mixing time of 0.4s) of loop 1 in DMSO at 25°C showing the region from 7.58 to 8.55ppm and 3.54 to 4.74ppm. The spectrum was zero filled from 3240 to 8192 points in F2 and from 512 to 2048 points in F1. The Varian window functions were LB (-5), GF (0.06), LB1(0) and GF1(0.02).

2.6.2 Loop 2

The spectra obtained for loop 2 were not assignable (Figure 2.30). The propagation of the NOE in the amide region was very poor. In an attempt to improve this the NOESY (mixing time 0.4s) was repeated with different mixing times (0.15 and 0.25s) but these failed to give better spectra and the assignment of loop 2 proved impossible. The cause of this problem was not obvious but might be due to the fact that loop 2 is believed to rest between the loops 1 and 3 (as in Rhodopsin) as well as the membrane and could be missing many of its usual interactions. Further experimentation would be required to enable the study of this peptide.

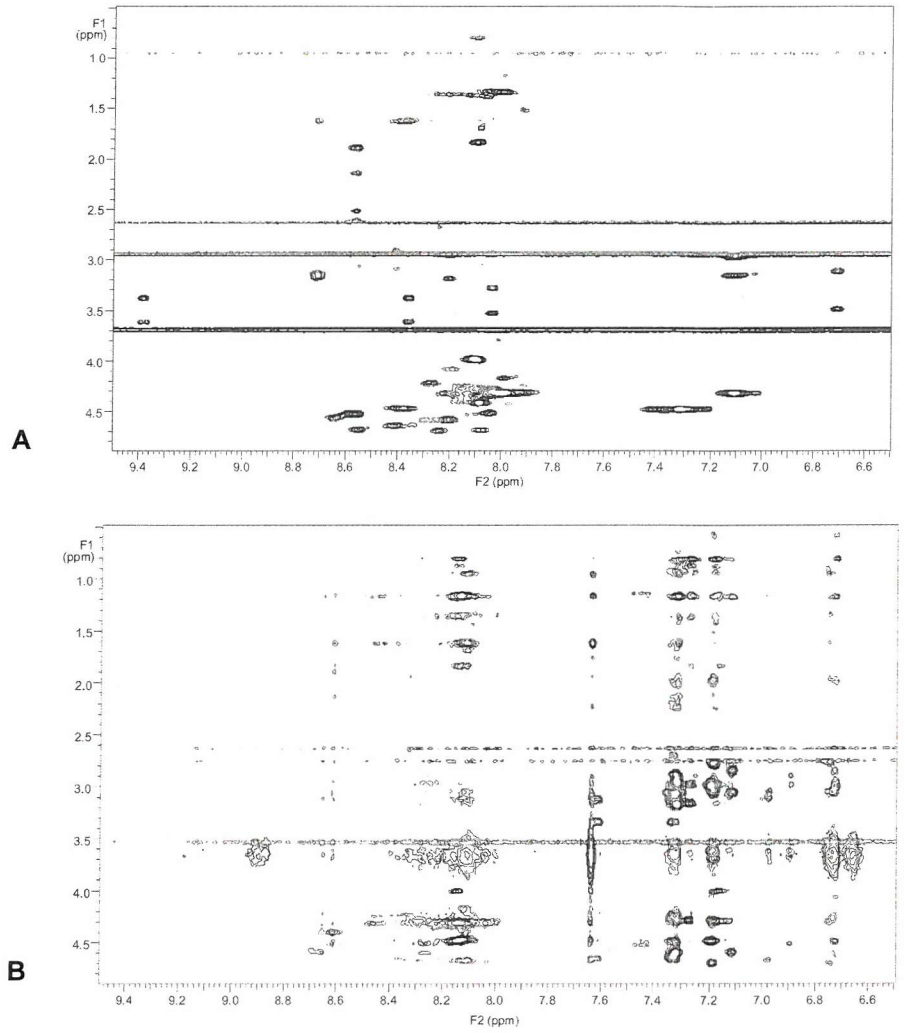


Figure 2.30 A. The TOCSY spectrum for loop 2 in DMSO-d_6 at 25°C showing the region between 6.5 to 9.5ppm and 0 to 5ppm. The spectrum had 4096 points in F2 and was zero filled from 512 to 2048 points in F1. The Varian window functions were LB (-2.2), GF (0.077), LB1(-10) and GF1(0.023). B. The NOESY spectrum (mixing time of 0.4s) for loop 2 in DMSO-d_6 at 25°C showing the region between 6.5 to 9.5ppm and 0 to 5ppm. The spectrum had 4096 points in F2 and was zero filled from 512 to 2048 points in F1. The Varian window functions were LB (-4.5), GF (0.064), LB1(-10.3) and GF1(0.022).

2.6.3 Loop 3

The NOESY and TOCSY spectra recorded at 25°C in DMSO enabled the assignment of the peptide (Figure 2.33). As with loop 1 (section 2.6.1) there were impurities which required a second NOESY spectrum, at a different temperature (20°C) to confirm the assignment of the non-sequential NOEs.

The NOESY spectrum had only 15 short range NOEs and 73 sequential NOEs in addition to the intraresidue NOEs. A summary of the sequential and short range NOEs is shown in Figure 2.31.

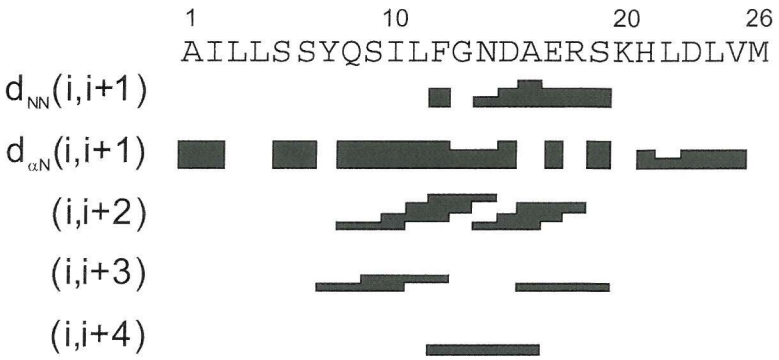


Figure 2.31 Summary of the sequential and short range NOEs for loop 3 in DMSO. The amino acid sequence and numbering are shown at the top. Sequential N-N and α-N NOEs are indicated by black bars, the thickness of the bar represents the strength of the observed NOE. The presence of short-range NOEs is indicated by a solid line.

150 structures were generated using the NOEs (Figure 2.32) however again these proved almost identical to the structures generated for the sequence without any NOE restraints.

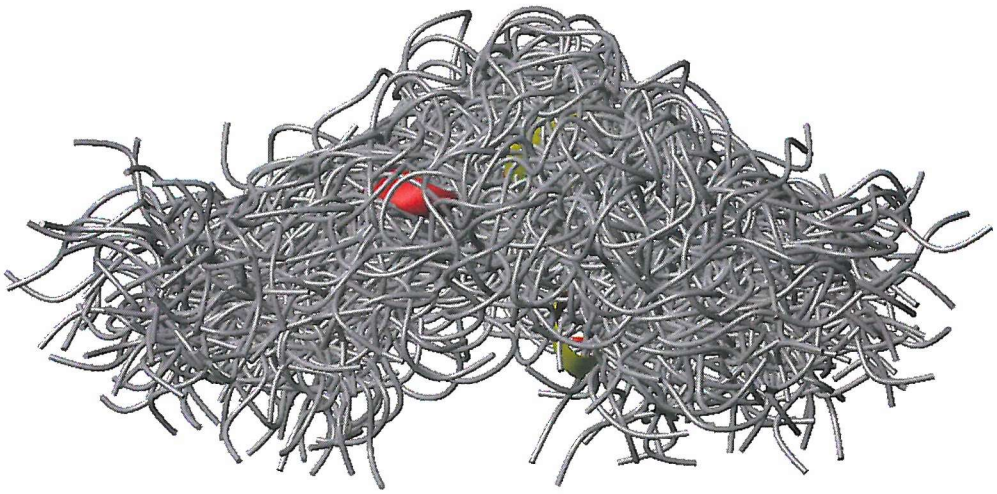


Figure 2.32 Ensemble of 150 structures generated for loop 3 in DMSO. These structures are almost identical to those which could be generated for the sequence with no NOE restraints.

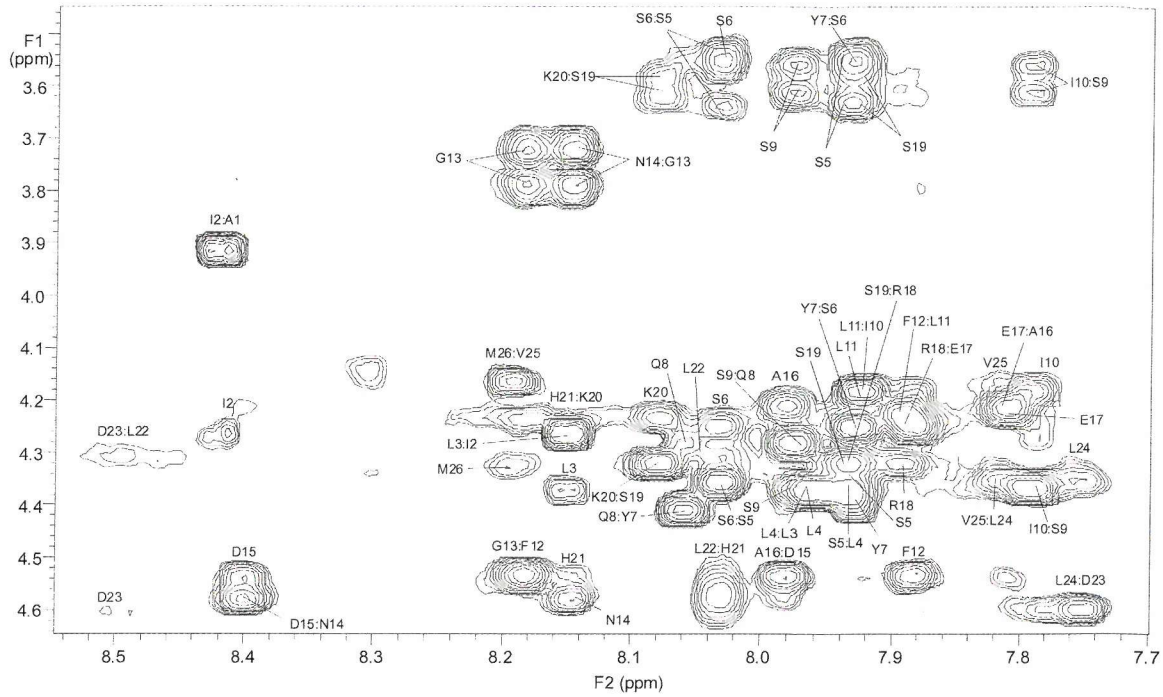


Figure 2.33 The NOESY spectrum (mixing time of 0.4s) of loop 3 in DMSO at 25°C showing the region between 7.7 to 8.55ppm and 3.45 to 4.65ppm. The spectrum was zero filled from 3246 to 8192 points in F2 and from 512 to 2048 points in F1. The Varian window functions were LB (-5), GF (0.064), LB1(-10.3) and GF1(0.029).

2.7 Circular Dichroism in DPC

Whilst the experiments in DMSO were acceptable for the assignment it was believed that an aqueous system would be preferable for the titration of the loops with the chemokines. With none of the loops being soluble in aqueous buffers it was thought a membrane mimic might aid dissolution by stabilising the transmembrane regions of the peptide with hydrophobic interactions. It was further hoped that the use of a membrane mimic might enable the recording of an assignable set of spectra for loop 2. n-Dodecylphosphocholine (DPC) is commercially available in a fully deuterated form (d_{38}) and has been used as a membrane mimic in previous NMR studies [100, 101]. In order to solubilize the loops in DPC, a common solvent (DMSO) was used so that the loops formed a homogeneous solution with the DPC. The solvent was then removed by lyophilization, reconstituted in the required buffer and then sonicated to ensure the insertion of the peptide into the DPC and the homogenous nature of the sample. Any particulate material was removed by centrifugation before the samples were studied by circular dichroism (CD) and NMR spectroscopy.

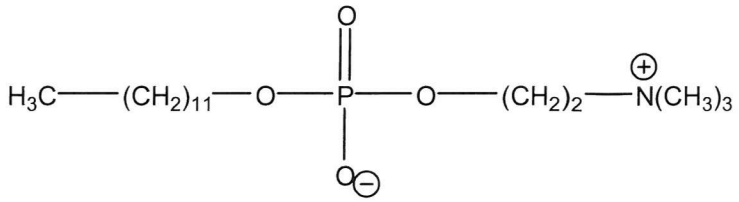
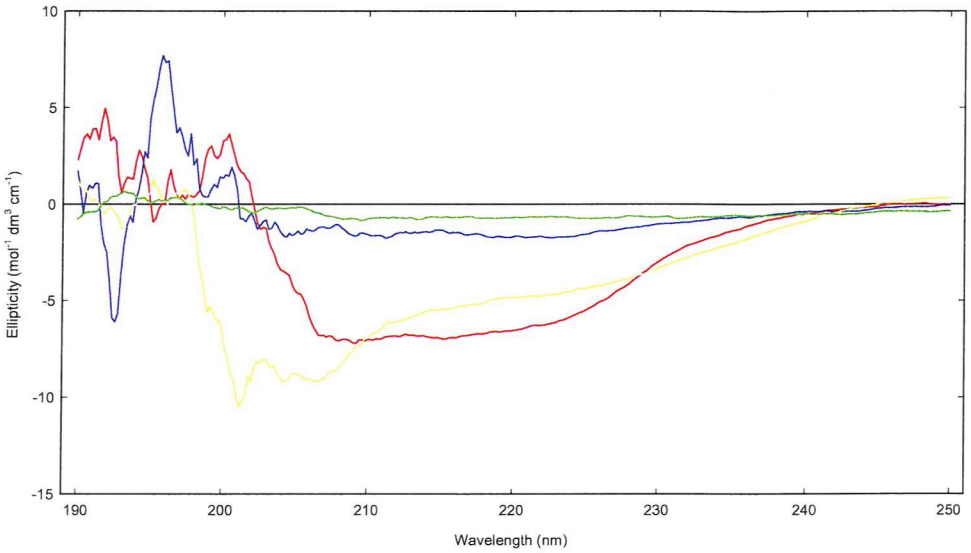


Figure 2.34 The chemical structure of n-dodecylphosphocholine (DPC).



Blue: Loop 2.

Green: DPC blank.

Red: Loop 3.

Yellow: Loop 1.

Figure 2.35 Circular Dichroism spectra for the loops in 100mM sodium acetate-d_6 pH5.5 and 150mM DPC-d_{38}. The data below 200nm is invalid due to the limitations of the system and the excessive voltage applied to the photomultiplier.

CD spectra were obtained for all three loops in DPC. Comparison of the spectra with some standard data sets [139] suggested that loops 1 and 3 were alpha helical in nature. Loop 2 could be alpha helical but its signal was not much stronger than that of DPC suggesting that there was not much of the peptide in solution.

2.8 NMR Spectroscopy of the Redesigned Loops in DPC

There was a problem in recording the TOCSY for all three loops in DPC; whilst the side chains showed propagation the amide hydrogens did not, with the fingerprint region being almost devoid of peaks despite changes in mixing time. In contrast the NOESY data were of a better quality and the assignments were done using the NOESY spectra alone. Performing an assignment in this way is a time intensive, complex task and would

probably have proved impossible due to the small chemical shift dispersion in the amide region, if not for the side chain assignments from the TOCSY which enabled the correct assignment of the alpha hydrogens.

2.8.1 Loop 1

The NOESY for loop1 had many overlapping peaks with most of the amide peaks being within 0.4ppm of each other from 7.8 to 8.2ppm (Figure 2.37). The aromatic region from which NOEs can normally easily be assigned contained only 13 resonances to account for the 26 resonances that should be seen. The assignment of each residues spin system was completed with extreme difficultly and structure generation proved even harder as almost all inter-residue NOEs were impossible to assign until a basic structure had been established from those few unique assignments. The structure was then worked out in a recursive manner with the spectrum. Many dead ends were sampled before the majority of NOEs could be accounted for. It was believed that over 95% of the clearly visible NOEs had been assigned and these were split into 19 long range, 39 short range, 66 sequential and 210 intraresidue NOEs. Figure 2.36 shows a summary of the sequential and short range NOEs.

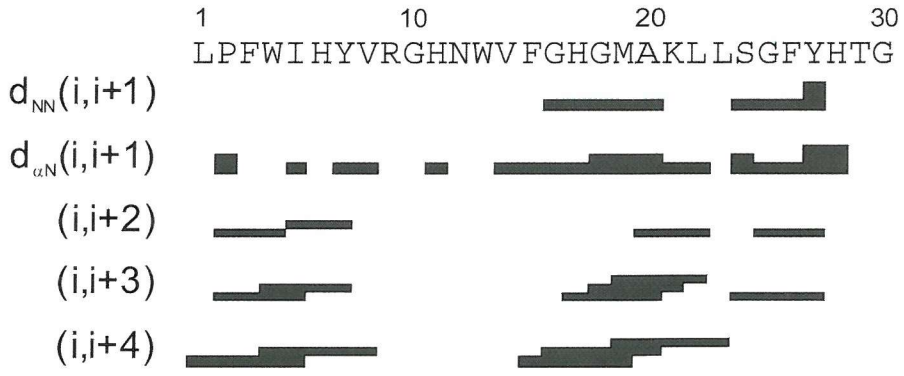


Figure 2.36 Summary of the sequential and short range NOEs for loop 1 in DPC. The amino acid sequence and numbering are shown at the top. Sequential N-N and α-N NOEs are indicated by black bars, the thickness of the bar represents the strength of the observed NOE. The presence of short-range NOEs is indicated by a solid line.

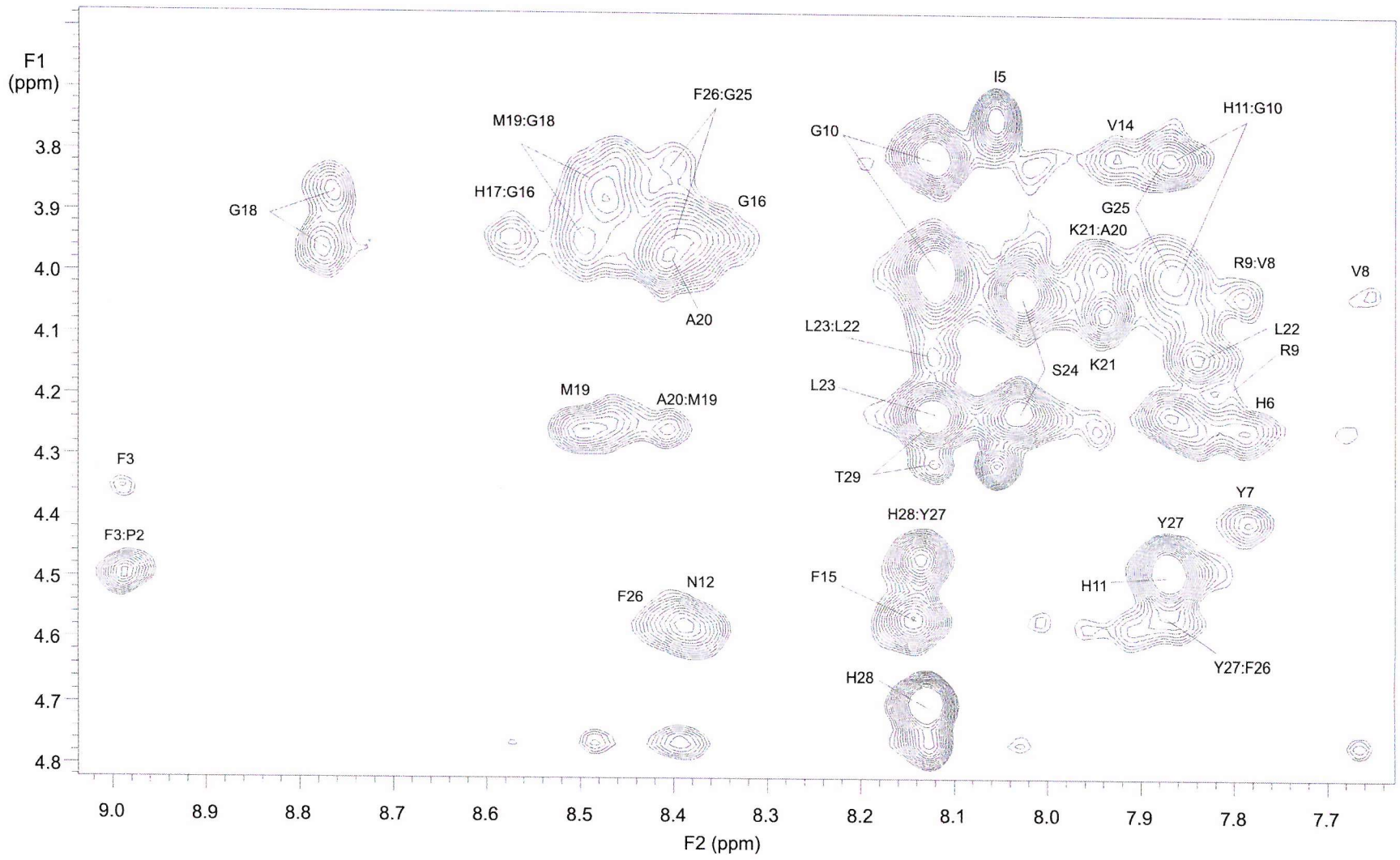


Figure 2.37 The NOESY spectrum (mixing time of 0.15s) for loop 1 in DPC at 25°C showing the region from 7.6 to 9.2ppm and 3.58 to 4.82ppm. The spectrum was zero filled from 4096 to 8192 points in F2 and was zero filled from 384 to 2048 points in F1. The Varian window functions were LB (0), GF (0.037), LB1(0) and GF1(0.018).

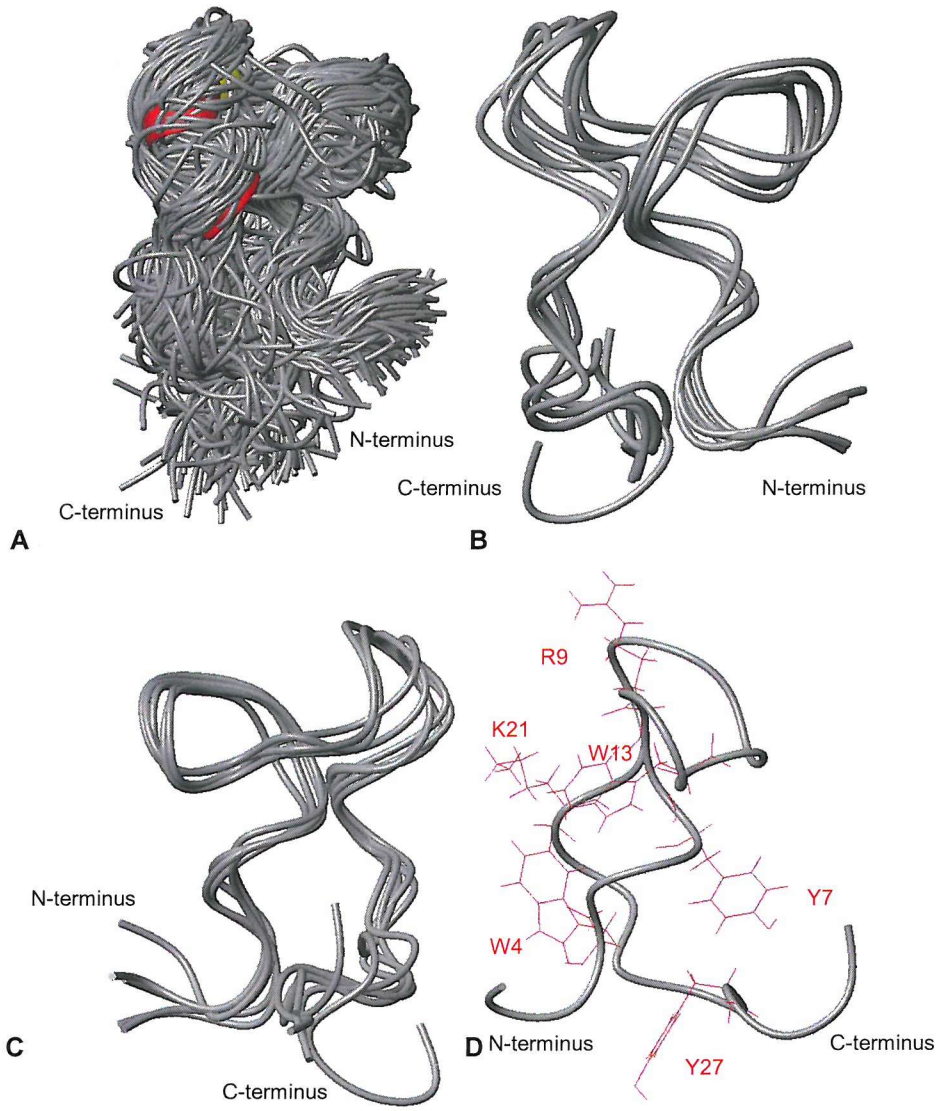


Figure 2.38 Loop 1 structure. A. Ensemble of 150 structures generated from the NOE table superimposed over all residues. All 150 structures appear to sample a similar area of space. B. The lowest five structures in energy that had no NOE violation greater than 0.1 and satisfied the Ramachandran plot, superimposed over all residues (RMSD to mean 1.49Å) C. Figure B rotated 180° and superimposed from P2 to H28 (RMSD to mean 0.97Å). D. Single structure with the residues that can act as membrane anchors highlighted (W4, Y7, R9, W13, K21 and Y27).

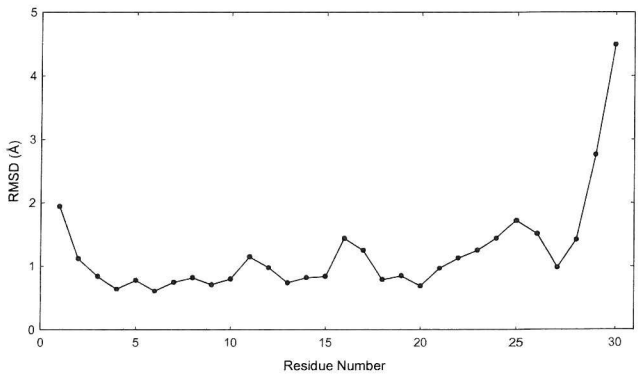


Figure 2.39 Atomic RMSD from the geometric mean structure for each residue when the eight lowest energy structures where superimposed using the N, Ca and C of all the residues. These showed the flexibility to be at the C-terminus.

An ensemble of 150 structures was generated from the NOE data and the structures all sampled a similar conformational space (Figure 2.38A). When the lowest five energy structures with no NOE violations greater than 0.1 and which satisfied the Ramachandran plot were compared (RMSD to mean 1.49Å) it was obvious that they had a very similar fold although there appeared to be a lot of flexibility in the structure (Figure 2.38B). This is confirmed by the plot of the RMSD for each residue with the two ends showing quite a degree of flexibility (Figure 2.39). When the structures where superimposed using the residues P2 to H28 (RMSD to mean 0.97Å) the end flexibility could clearly be seen (Figure 2.38C).

The final structure generated was a non-classical fold, although it did have a loop like structure with the N-terminus and C-terminus been close in space. The predicted structure of two loops formed from a tryptophan zipper appears to be incorrect, although W13 and W4 are close in space. W13 appears to be causing a turn in the loop structure possibly where its side chain is folding back to interact with the membrane.

Figure 2.38D shows those residues that are believed to act as membrane anchors displayed upon one of the structures. It is possible that the membrane ends in the region of W13 and K21, although no firm conclusion could be drawn.

Whilst most of the structures generated did not violate the NOEs, nearly all had residues in the forbidden region of the Ramachandran plot (Figure 2.40B). Whilst attempts have been made to rectify this it would appear that this is an intrinsic feature of the structure generation, probably due to the lack of NOEs. It is noted that most of the published papers on peptide structures do not mention Ramachandran data and this might well be a common problem. The structural statistics and RMSDs for the lowest energy non-misfolded loop 1 structures is shown in Figure 2.40.

RMS deviations from experimental			<16 loop 1>
Distance restraints (Å)[a]	All NOE		0.0130 ± 0.0088
	sequential (66)		0.0059 ± 0.0045
	short (28)		0.0139 ± 0.0087
	long (10)		0.0124 ± 0.0093
	short (R-6 averaged) (11)		0.0185 ± 0.0086
	long (R-6 averaged) (9)		0.0141 ± 0.0129
energies (kcal mol^{-1})[b]	E_{NOE}		1.34 ± 0.8208
	E_{DIHE}		n/a
	E_{REPEL}		15.85 ± 4.0915
Deviations from idealised geometry[c]	bonds (Å)		0.0032 ± 0.0004
	angles (deg.)		0.5282 ± 0.0398
	improper (deg.)		0.2505 ± 0.0301
RMSD (Å)[d]	backbone atoms	(res 2-26)	1.62 ± 0.38
	heavy atoms	(res2-26)	2.59 ± 0.45
(ϕ/ψ) in Ramachandran plot (%)[e]			
residues 1-30	core region		32.6
	additionally allowed regions		41.6
	generously allowed regions		17.9
	forbidden regions		7.9

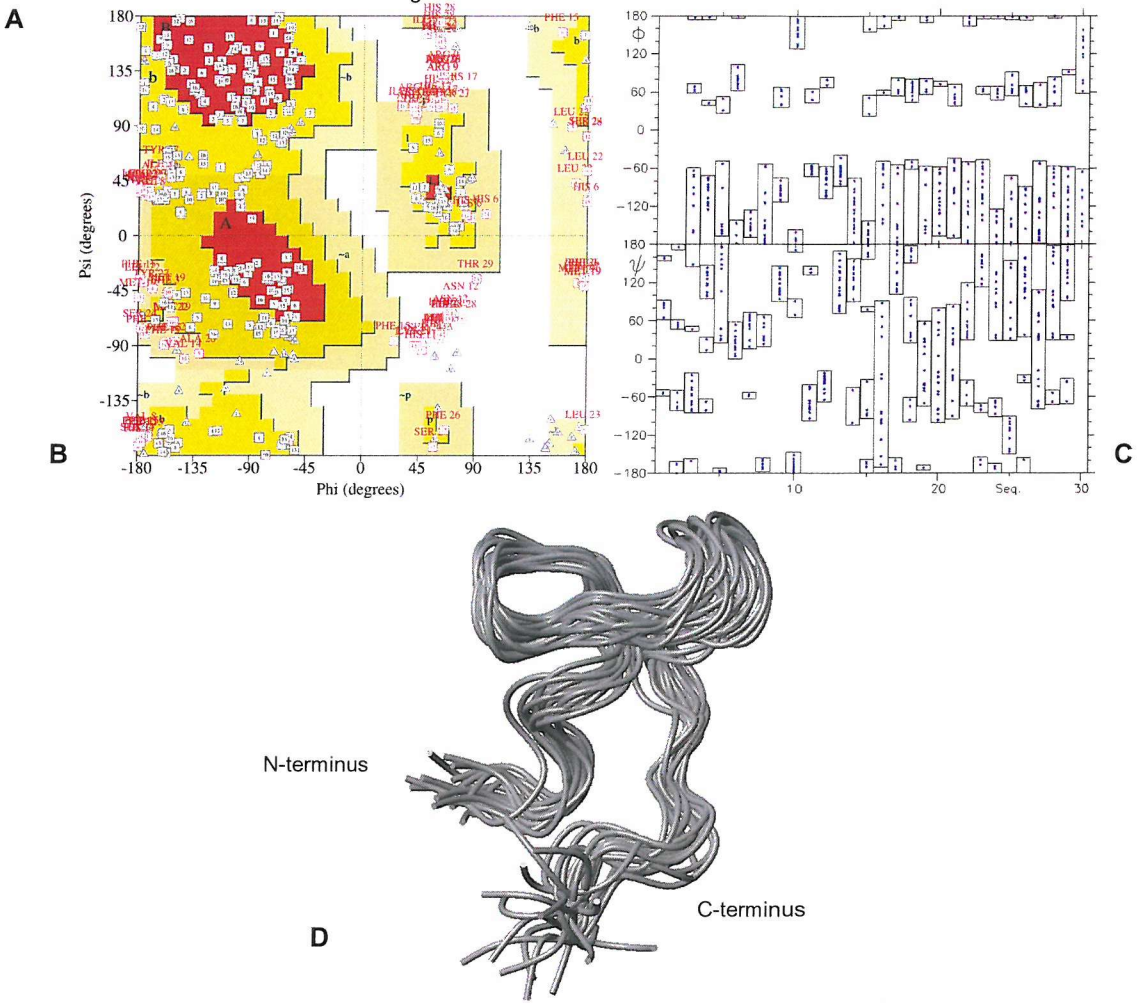
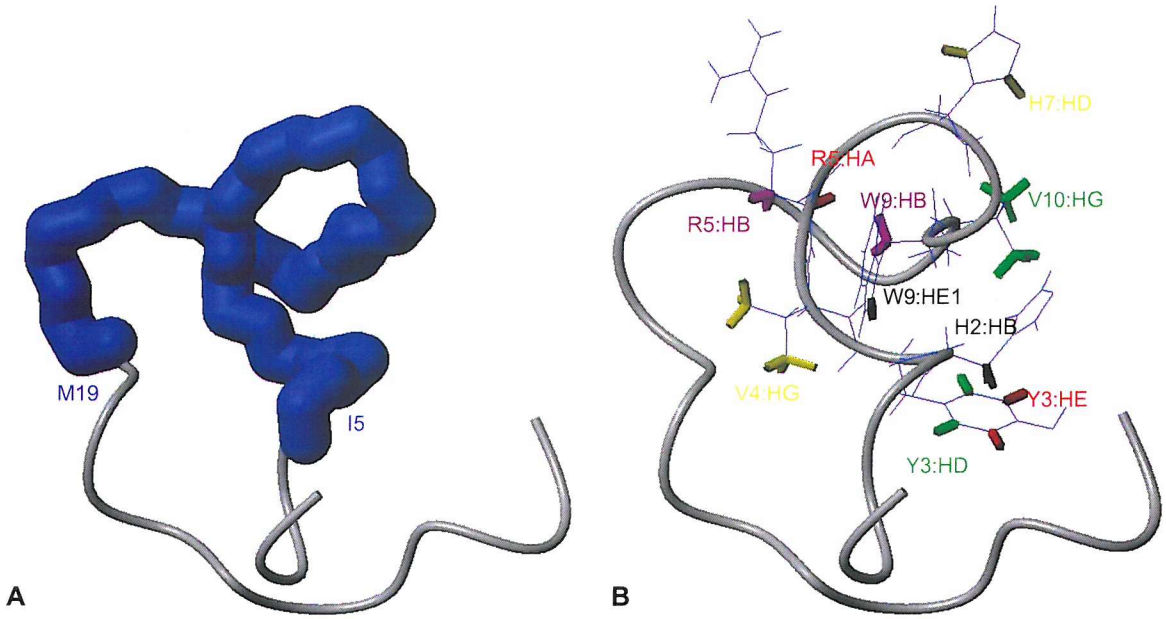


Figure 2.40 Analysis of loop 1 structure. A. Structural statistics. [a] The RMS deviation of the experimental restraints is calculated with respect to the upper and lower limits of the input restraints. [b] The values for E_{NOE} and E_{DIHE} are calculated from a square well potential with a force constant of 50 kcal mol^{-1} Å^2 and 200 kcal mol^{-1} rad^{-2}. E_{REPEL} is calculated with a force constant of 4 kcal mol^{-1} Å^{-4}, and the final van der Waals radii were set to 0.80 times the value used in the CHARMM force field. [c] The values for bonds, angles, and impropers show the deviation from ideal values based on perfect stereochemistry. [d] Root-mean-square deviations to the average structure. [e] As determined by the program PROCHECK [140]. B. Ramachandran plot. C. Angle bars. D. Ensemble of the sixteen structures used to calculate the statistics.

With the structure established for loop 1 it is interesting to examine the NOEs observed in the first generation loop 1 to see whether they match this fold (Figure 2.41). Several of the NOEs seem unlikely, for example H7:HD#:V4HG#, however the most convincing NOE W9:HB#:R5HB# does seem possible. Whether the first generation loop was sampling the same conformation is unclear.



Residues colour coded to show the NOEs:
Black: W9HE1:H2HB#.
Green: Y3HD#:V10HG#.
Purple: W9HB#:R5HB#.
Red: Y3HE#:R5HA.
Yellow: H7HD:V4HG#.

Figure 2.41 Comparison with the original loop 1. A. The residues from the original loop 1 corresponding to I5 to M19 shown on the redesigned loop 1 structure. B. Those residues that where believed to be showing NOEs to each other.

2.8.2 Loop 2

Attempts to study loop 2 in DPC once again proved problematic. Signals appeared weaker than expected and there were particular problems with the signal in the amide region with no assignable NOESY being recorded despite a change in the mixing time (0.15s).

As the recording of the CD spectra also relies upon the backbone hydrogens it is possible that whatever is causing the problem is related to amide hydrogens. This work was abandoned at this stage.

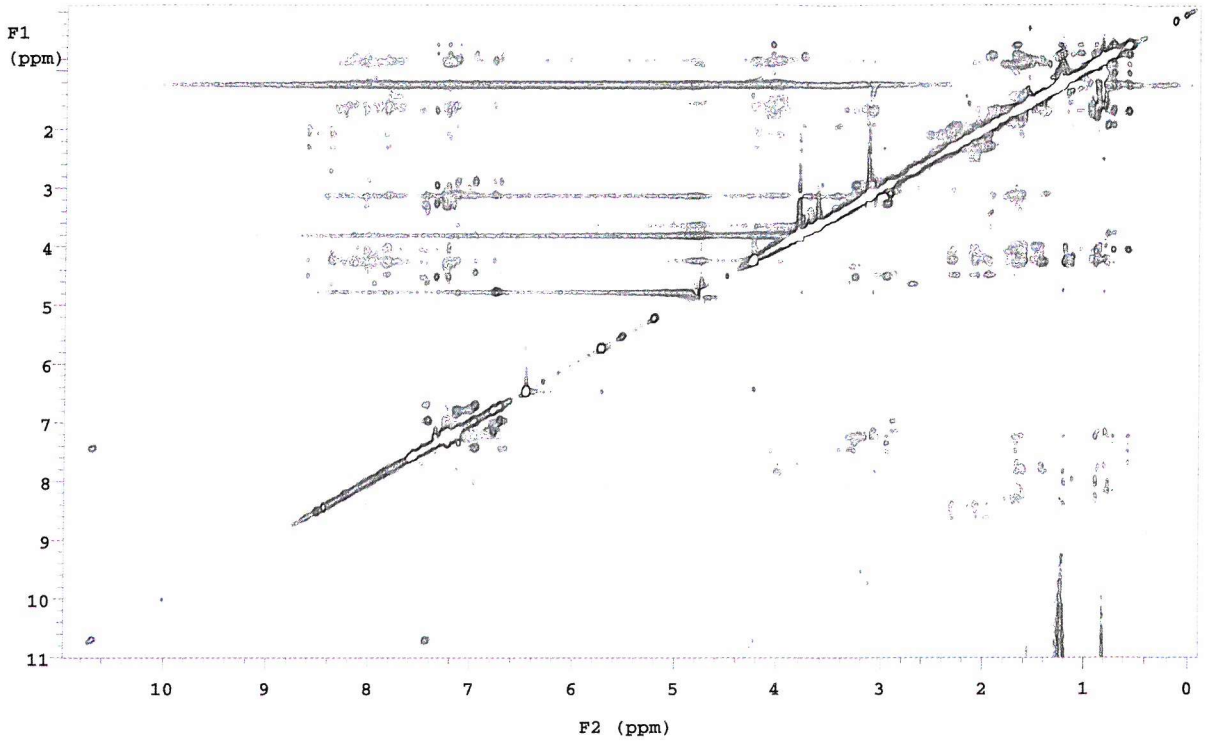


Figure 2.42 The NOESY spectrum (mixing time of 0.4s) for loop 2 in DPC at 25°C showing the region from 0 to 11ppm in both dimensions. The spectrum was zero filled from 4096 to 8192 points in F2 and was zero filled from 256 to 2048 points in F1. The Varian window functions were LB (-5), GF (0.032), LB1(-9.5) and GF1(0.012).

2.8.3 Loop 3

The assignment of loop 3 when compared to loop 1 was much easier, but extremely difficult in its own right. The amide region was slightly more dispersed with most resonances occurring within 0.8ppm of each other and the two aromatic side chains gave almost unique chemical shifts. The structure was again generated from the unique NOEs before the non-unique ones where added, in a recursive manner until over 95% of the NOEs were accounted for. In total there were 19 long range NOEs, 23 short range and 25 sequential as well as the intraresidue ones. The sequential and short range NOEs are summarised in Figure 2.44.

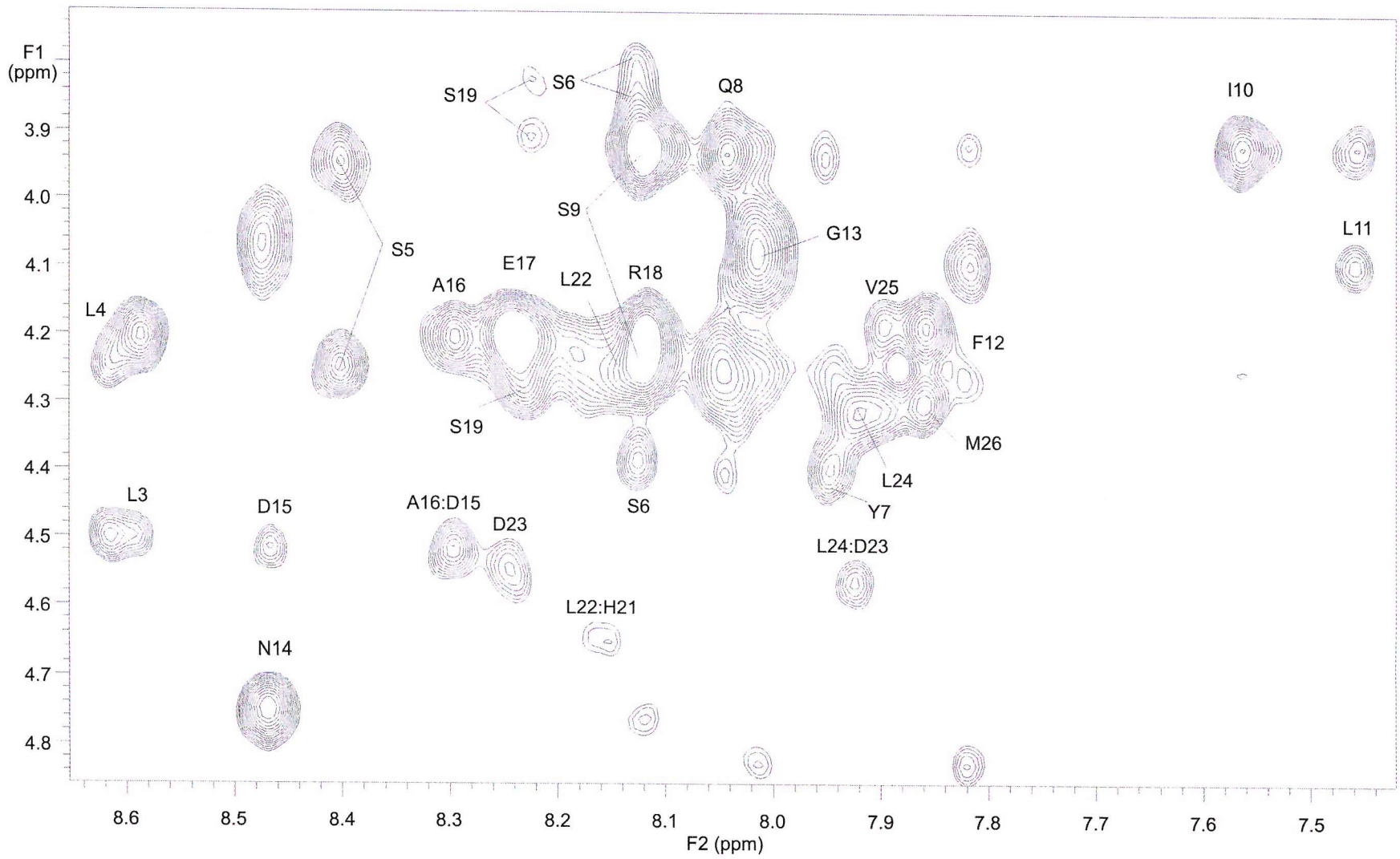


Figure 2.43 The NOESY spectrum (mixing time of 0.4s) for loop 3 in DPC at 25°C showing the region from 7.4 to 8.65ppm and 3.72 to 4.85ppm. The spectrum was zero filled from 4096 to 8192 points in F2 and was zero filled from 256 to 2048 points in F1. The Varian window functions were LB (-6.1), GF (0.038), LB1(-10.3) and GF1(0.012).

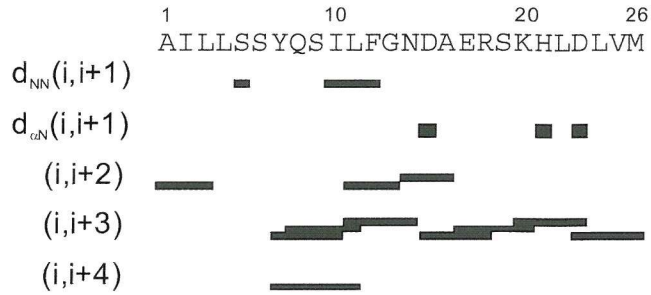


Figure 2.44 Summary of the sequential and short range NOEs for loop 3 in DPC. The amino acid sequence and numbering are shown at the top. Sequential N-N and α-N NOEs are indicated by black bars, the thickness of the bar represents the strength of the observed NOE. The presence of short-range NOEs is indicated by a solid line.

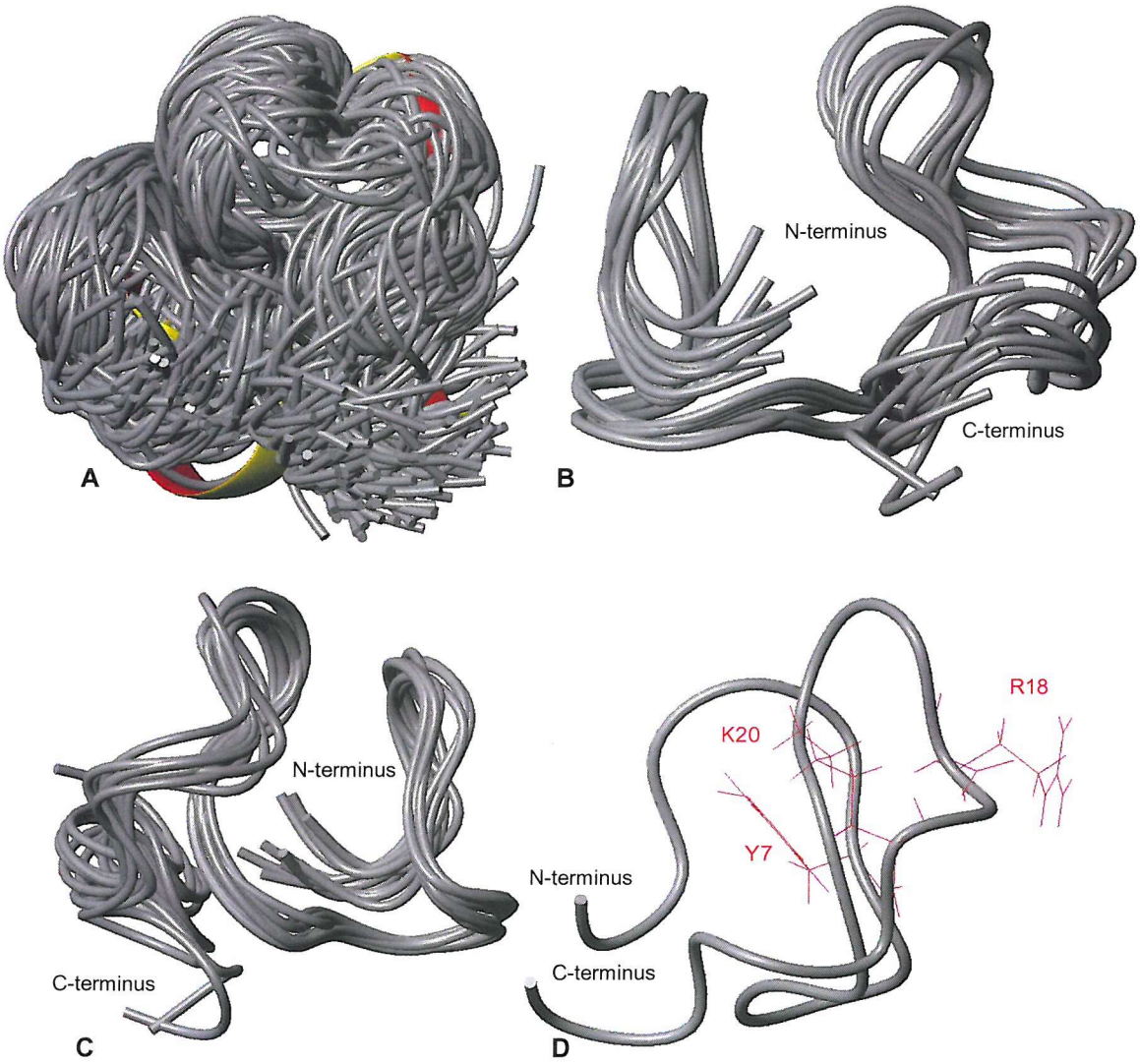


Figure 2.45 Loop 3 structure in DPC. A. Ensemble of 150 structures generated from the NOE table superimposed using all the residues. All 150 structures appear to sample a similar area of space. B. The lowest ten structures in energy that had no NOE violation greater than 0.1. Superimposed using all residues (RMSD to mean 1.02Å) C. Figure B rotated 180° and superimposed from L3 to D23 (RMSD to mean 0.69Å). D. Single structure with the residues that can act as membrane anchors highlighted (Y7, R18 and K20).

As with loop 1 in DPC an ensemble of 150 structures was generated from the NOE table and the structures were all sampling similar conformational space (Figure 2.45A). The lowest ten energy structures with no NOE violations greater than 0.1 had a very similar fold and the plot of the RMSD for each residue (Figure 2.46) again showed the two ends to be fairly flexible, especially the C-terminus. When the structures were superimposed using the residues L3 to D23 (RMSD to mean 0.69Å) this flexibility became clearer (Figure 2.45C).

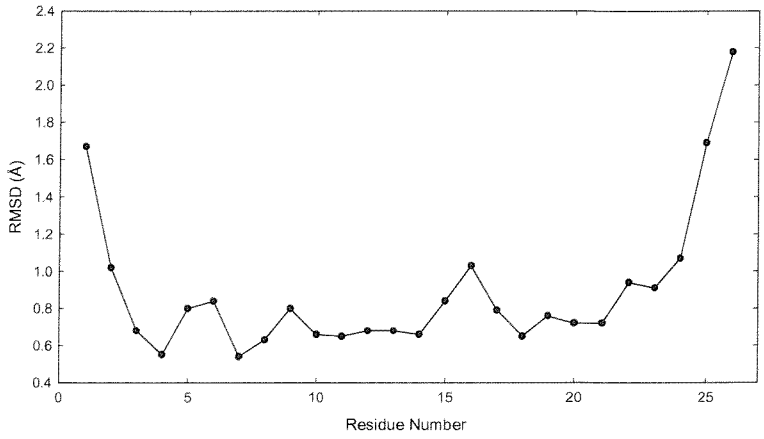


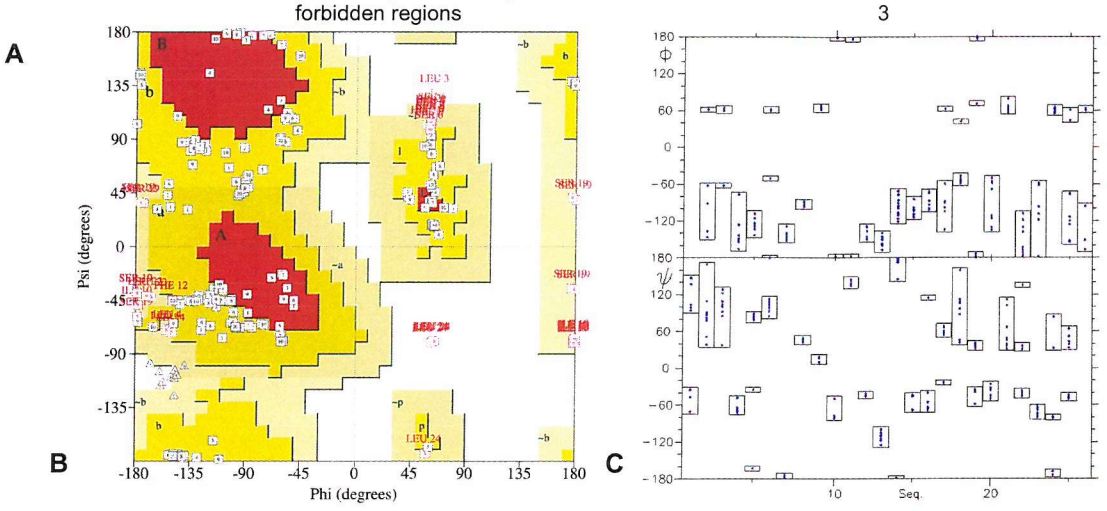
Figure 2.46 Atomic RMSD from the geometric mean structure for each residue when the three lowest energy structures were superimposed using the N, Ca and C of all the residues. This showed only the region from L3 to D23 is well-defined.

The final structure generated was again a non-classical fold, with loop like features. This loop does seem to consist of two separate lobes, which are almost at 90° to each other.

Figure 2.45D shows those residues that are believed to act as membrane anchors displayed upon one of the structures. These can all be placed in a similar plane which suggests this might be where the membrane ends.

As with loop 1 most of the structures generated did not violate the NOEs, but nearly all had residues in the forbidden region of the Ramachandran plot (Figure 2.47B). A family of 10 structures and their calculated structural statistics and RMSDs is shown in Figure 2.47.

	RMS deviations from experimental		<10 loop 3>
Distance restraints (Å)[a]	All NOE		0.0107 ± 0.0035
	intra (70)		0.0001 ± 0.0003
	sequential (25)		0
	short (16)		0.0219 ± 0.0035
	long (15)		0.0136 ± 0.0028
	short (R-6 averaged) (7)		0.0122 ± 0.0043
	long (R-6 averaged) (4)		0.0166 ± 0.0098
energies (kcal mol^{-1})[b]	E_{NOE}		0.7881 ± 0.2466
	E_{DIHE}		n/a
	E_{REPEL}		2.3475 ± 0.2799
Deviations from idealised geometry[c]	bonds (Å)		0.0024 ± 0.0001
	angles (deg.)		0.4732 ± 0.0080
	improper (deg.)		0.2704 ± 0.0158
RMSD (Å)[d]	backbone atoms	(res 3-23)	0.95 ± 0.31
	heavy atoms	(res 3-23)	1.58 ± 0.33
	(ϕ/ψ) in Ramachandran plot (%)[e]		
residues 2-26	core region		13
	additionally allowed regions		66.1
	generously allowed regions		17.8
	forbidden regions		3



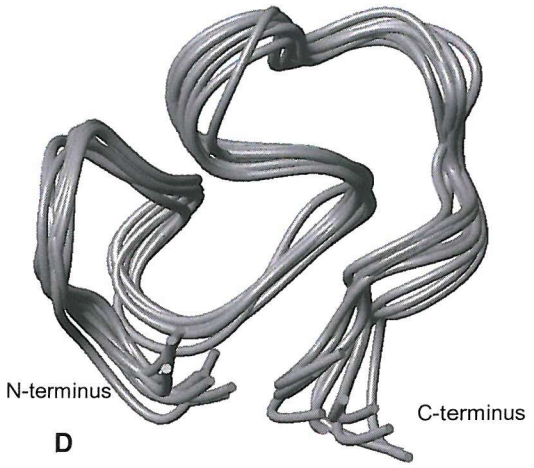


Figure 2.47 Analysis of loop 3 structure in DPC. A. Structural statistics. [a] The RMS deviation of the experimental restraints is calculated with respect to the upper and lower limits of the input restraints. [b] The values for E_{NOE} and E_{DIHE} are calculated from a square well potential with a force constant of 50 kcal mol^{-1} Å^2 and 200 kcal mol^{-1} rad^{-2}. E_{REPEL} is calculated with a force constant of 4 kcal mol^{-1} Å^{-4}, and the final van der Waals radii were set to 0.80 times the value used in the CHARMM force field. [c] The values for bonds, angles, and impropers show the deviation from ideal values based on perfect stereochemistry. [d] Root-mean-square deviations to the average structure. [e] As determined by the program PROCHECK [140]. B. Ramachandran plot. C. Angle bars. D. Ensemble of the ten structures used to calculate the statistics superimposed from L3 to D23.

2.9 Conclusion

There were obvious differences between the spectra recorded in DMSO and those in DPC, with the spectra in DMSO showing far fewer longer-range NOEs. Consequently the structures generated from the DMSO assignments were lacking in secondary and tertiary structure, whereas the DPC assignments resulted in loops with the N and C-termini being close in space, as might be expected in the membrane mimic and indeed the full receptor.

An ensemble of sixteen structures of the 30 amino acid loop 1 were computed by simulated annealing with the inclusion of 334 experiment restraints. Atomic RMSDs about the mean structure was 1.62Å for the backbone atoms and 2.59Å for all the heavy atoms (both values calculated over residues 2-26). The structure generated was a non-classical fold with a loop like structure.

Similarly an ensemble of ten structures of the 26 amino acid loop 3 were computed with 137 experiment restraints. Atomic RMSDs about the mean structure was 0.95Å for the backbone atoms and 1.58Å for all the heavy atoms (both values calculated over residues 3-23). The structure generated was also a non-classical fold which consisted of two loop like structures.

It is concluded that DMSO is not a suitable solvent for the solving the structure of these extracellular loops, although loops could be generated along a similar line to Yeagles by the selective use of the NOE strengths to generate helices at either end of the peptide and a flexible middle region. It is possible that Yeagle might have unwittingly done so in his previously published work, however his loops do match those of the crystal structure. It would be interesting to repeat this work for rhodopsin using a membrane mimic.

It is proposed that structures formed by loops 1 and 3 in DPC are representative of those that would be found in the native transmembrane receptor. Based on this assumption DPC solubilized loops 1 and 3 were used in titration experiments with the chemokine in an attempt to observe their interactions.

3 Human CCL24: Cloning, Expression, Purification and Titration with Loops 1 and 3 Suspended in DPC

3.1 Overview

Plasmids containing various lengths of DNA coding for CCL24 were constructed and used for the overexpression of CCL24 in microbial cell cultures. One of these constructs coding for CCL24 residues G25-P98 was grown and purified from *E.coli* strain BL21(DE3). Once a protocol for the purification was established, ^{15}N labelled CCL24 (G25-P98) was expressed and purified similarly.

Loops 1 and 3 were insoluble in aqueous buffer but could be solubilized using DPC micelles. This allowed a structure to be determined by 2D ^{1}H NMR for both loops in a pseudo membrane environment (chapter 2). ^{15}N labelled CCL24 was titrated into the loop/micelle preparations to identify their respective binding sites. A model for the binding of CCL24 to its receptor CCR3 has been proposed on the basis of these results.

3.2 Introduction

CCL24 cDNA was originally identified in the database of Human Genome Sciences Inc. (Rockville MD) on the basis of the CC motif and the homology to known CC chemokines. The cDNA was isolated from a library derived from activated human monocytes and the mature protein was expressed using Sf9 insect cells [141-143]. Several groups have reported the mature protein with the same N-terminus, but with differing C-termini lengths. From the gene sequence CCL24 would be expected to be a 93 amino acid protein, but there have been four reported C-terminal truncated Variants of 73, 75, 76 and 78 residues respectively [141].

The correct sequence length is uncertain as CCL24 has never been purified from source and consequently the signalling sequence may not have been successfully identified and the C-terminal truncations may have resulted from degradation caused during the purification process. There is also the possibility of post-translational modifications such as acetylation.

A study by Grzegorzewski showed that the selective deletion of the N-terminal amino acids of CCL24 gave rise to different effects. This includes one mutant of CCL24, P4 to R73 (P30 to R99 on the DNA sequence), which was converted from an agonist to an antagonist of CCL11, CCL24 and CCL13 (MCP-4) functional responses in eosinophil calcium flux and chemotaxis assays [144].

Interestingly CCL24 binds not only to CCR3 but also to an uncharacterized receptor on bone marrow cells. In haematopoietic assays CCL24 strongly suppressed colony formation by the high proliferative potential colony-forming cell (HPP-CFC) [142]. The mutant P4 to R73 (P30 to R99 on the DNA sequence) retained this ability [144]. HPP-CFCs have been defined by their ability to form very large colonies (>5mm in diameter) in bone marrow cultures *in vitro*. These colonies include progenitor and mature haematopoietic cells of the granulocyte, macrophage and megakaryocyte. The HPP-CFC is not a true pluripotential haematopoietic stem cell, but more akin to their slightly differentiated progeny, the primitive haematopoietic progenitor cell.

3.3 Results and Discussion

3.3.1 Plasmids

The cDNA for CCL24 (Table 3.1) was obtained by a collaborator from a human cDNA library [145]. This cDNA was amplified by PCR to produce three different lengths of CCL24, which were then ligated into vectors for overexpression in *E.coli*.

For the first construct, the full length of the cDNA (M1 to C119) was amplified by PCR with primers that added a 5' Nde-1 restriction enzyme cut site (Table 3.2A) and a 3' BamH1 restriction enzyme cut site (Table 3.2B). The resulting PCR fragment was then ligated into pET-15b at those cut sites (Figure 3.1) which produced a vector that coded for a hexa-His tag, a thrombin cleavage site and the full length of the cDNA. When the vector was transformed into BL21(DE3) and a small scale growth performed there was no visible induction band, as shown by the SDS-PAGE analysis of the crude cell lysates (Figure 3.2 lanes 21 to 28). This is probably due to the signalling sequence on the cDNA.

Residue Number on Protein						...	...	...	...	...	...	...	...	...	...
Residue						Met	Ala	Gly	Leu	Met	Thr	Ile	Val	Thr	Ser
Residue number on Codon						1	2	3	4	5	6	7	8	9	10
Codon						ATG	GCA	GGC	CTG	ATG	ACC	ATA	GTA	ACC	AGC
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Leu	Leu	Phe	Leu	Gly	Val	Cys	Ala	His	His	Ile	Ile	Pro	Thr	Gly	Ser
11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
CTG	CTG	TTC	CTG	GGA	GTC	TGC	GCT	CAT	CAC	ATC	ATT	CCA	ACA	GGA	AGT
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Val	Val	Ile	Pro	Ser	Pro	Cys	Cys	Met	Phe	Phe	Val	Ser	Lys	Arg	Ile
27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42
GTG	GTC	ATT	CCA	AGC	CCA	TGC	TGC	ATG	TTC	TTT	GTC	AGC	AAG	CGC	ATC
17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
Pro	Glu	Asn	Arg	Val	Val	Ser	Tyr	Gln	Leu	Ser	Ser	Arg	Ser	Thr	Cys
43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58
CCA	GAG	AAC	CGC	GTG	GTC	AGC	TAC	CAG	CTG	AGC	TCC	CGC	AGC	ACC	TGT
33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
Leu	Lys	Ala	Gly	Val	Ile	Phe	Thr	Thr	Lys	Lys	Gly	Gln	Gln	Phe	Cys
59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74
CTG	AAG	GCT	GGA	GTC	ATC	TTC	ACC	ACC	AAG	AAG	GGA	CAG	CAG	TTC	TGT
49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64
Gly	Asp	Pro	Lys	Gln	Glu	Trp	Val	Gln	Arg	Tyr	Met	Lys	Asn	Leu	Asp
75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90
GGA	GAT	CCA	AAG	CAG	GAG	TGG	GTC	CAG	CGC	TAC	ATG	AAA	AAT	CTG	GAT
65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
Ala	Lys	Gln	Lys	Lys	Ala	Ser	Pro	Arg	Ala	Arg	Ala	Val	Ala	Val	Lys
91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106
GCC	AAG	CAG	AAG	AAG	GCC	TCC	CCA	AGG	GCA	CGC	GCA	GTG	GCT	GTG	AAG
81	82	83	84	85	86	87	88	89	90	91	92	93			
Gly	Pro	Val	Gln	Arg	Tyr	Pro	Gly	Asn	Gln	Thr	Thr	Cys			
107	108	109	110	111	112	113	114	115	116	117	118	119			
GGC	CCA	GTC	CAG	CGC	TAT	CCT	GGC	AAC	CAA	ACC	ACC	TGC	TAA		

Table 3.1 The cDNA sequence of CCL24. The first line of numbers in red is the position of the amino acid on the native protein. The second line is the three letter abbreviation of the amino acid coded for by the codon in line four. The third line in blue is the position of the amino acid on the cDNA and the fourth line is the codon in the cDNA.

A

	Overhang	Nde1 Cut site	Ala	Gly	Leu	Met	Thr	Ile	(Val)	
5′	**GGGAATTC**	**CATATG**	**GCA**	**GGC**	**CTG**	**ATG**	**ACC**	**ATA**	**G**	3′
3′			CGT	CCG	GAC	TAC	TGG	TAT	C	5′

B

	(Gly)	Asn	Gln	Thr	Thr	Cys	END	BamH1 Cut Site	Overhang	
5′	GC	AAC	CAA	ACC	ACC	TGC	TAA			3′
3′	**CG**	**TTG**	**GTT**	**TGG**	**TGG**	**ACG**	**ATT**	**CCTAGG**	**CGCGGG**	5′

Table 3.2 The primers used for M1 to C119. In both A and B the first line shows what was coded for, with the residues in brackets signifying only partial codons. The second line shows the bases 5' to 3' and the third line contains the antisense. The primer is in bold. A. The 5' primer. B. The 3' primer.

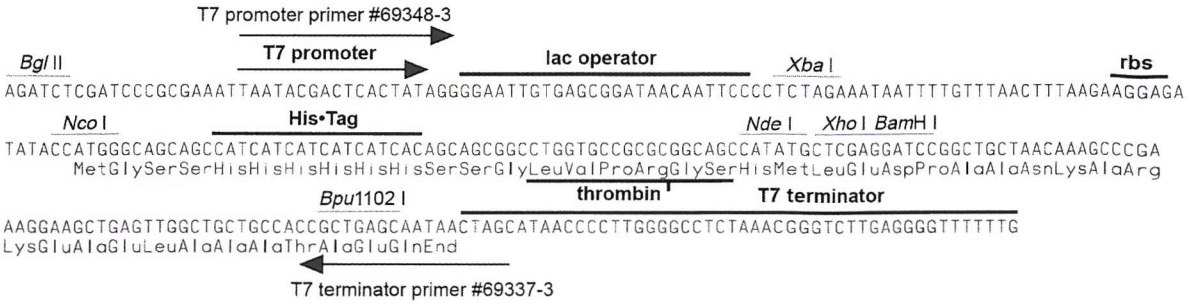


Figure 3.1 The pET-15b vector map showing the cloning and expression region sequence. This shows the position of the restriction enzyme cut sites in relation to the inbuilt hexa-His Tag and the thrombin cleavage site.

For the second construct the cDNA was amplified by PCR from the published N-terminus of V27 to the C-terminus of C119S. C119 was mutated to serine to reduce dimerisation. In the vector (Figure 3.1) there was a Nde1 site directly after the thrombin cleavage site; if this site was used to insert a DNA sequence, then the translated protein produced after thrombin cleavage would have an N-terminal extension of GSHM. The amino acids that formed the thrombin cleavage site were LVPRGS, which thrombin would cleave between arginine and glycine. Examination of the cDNA (Table 3.1) showed that the two residues proceeding V27 were G25 and S26, hence it was decided to overlap these two residues with the end of the thrombin cleavage site. The protein produced by this method (after thrombin cleavage) would have only been extended by GS rather than GSHM. To achieve this the vector was cut at Nco1, thereby removing the hexa-His tag and thrombin cleavage site coded for in the vector. The hexa-His tag was replaced with a deca-His tag, which would theoretically bind the nickel column with a higher affinity. The 5' primer was designed to include a Nco1 restriction enzyme cut site, a deca-His tag and a thrombin cleavage site which overlapped with G25 and S26 of the cDNA (Table 3.3A). The 3' primer included a BamH1 restriction enzyme cut site (Table 3.3B). This was then ligated into pET-15b at the Nco1 and BamH1 cut sites.

A

		Overhang		Nco1 Cut site			(Gly)		Ser	Ser	His	His	His	His	
	5′	**GGCATG**		**CCATGG**			**GC**		**AGC**	**AGC**	**CAT**	**CAT**	**CAT**	**CAT**	3′
	3′						CG		TCG	TCG	GTA	GTA	GTA	GTA	5′
		His	His		His	His	His	His	Ser	Ser	Gly	Leu	Val	Pro	
	5′	**CAT**	**CAT**		**CAT**	**CAT**	**CAT**	**CAT**	**AGC**	**AGC**	**GGC**	**CTG**	**GTG**	**CCG**	3′
	3′	GTA	GTA		GTA	GTA	GTA	GTA	TCG	TCG	CCG	GAC	CAC	GGC	5′
		Arg	Gly		Ser	Val	Val	Ile	Pro	Ser	Pro	Cys			
	5′	**CGC**	**GGA**		**AGT**	**GTG**	**GTC**	**ATT**	**CCA**	**AGC**	**CCA**	**TGC**	3′		
	3′	GCG	CCT		TCA	CAC	CAG	TAA	GGT	TCG	GGT	ACG	5′		

B

	Pro	Gly	Asn	Gln	Thr	Thr	Ser	END	BamH1 Cut Site	Overhang	
5′	CCT	GGC	AAC	CAA	ACC	ACC	TCC	TAA			3′
3′	**GGA**	**CCG**	**TTG**	**GTT**	**TGG**	**TGG**	**AGG**	**ATT**	**CCTAGG**	**CGCGGG**	5′

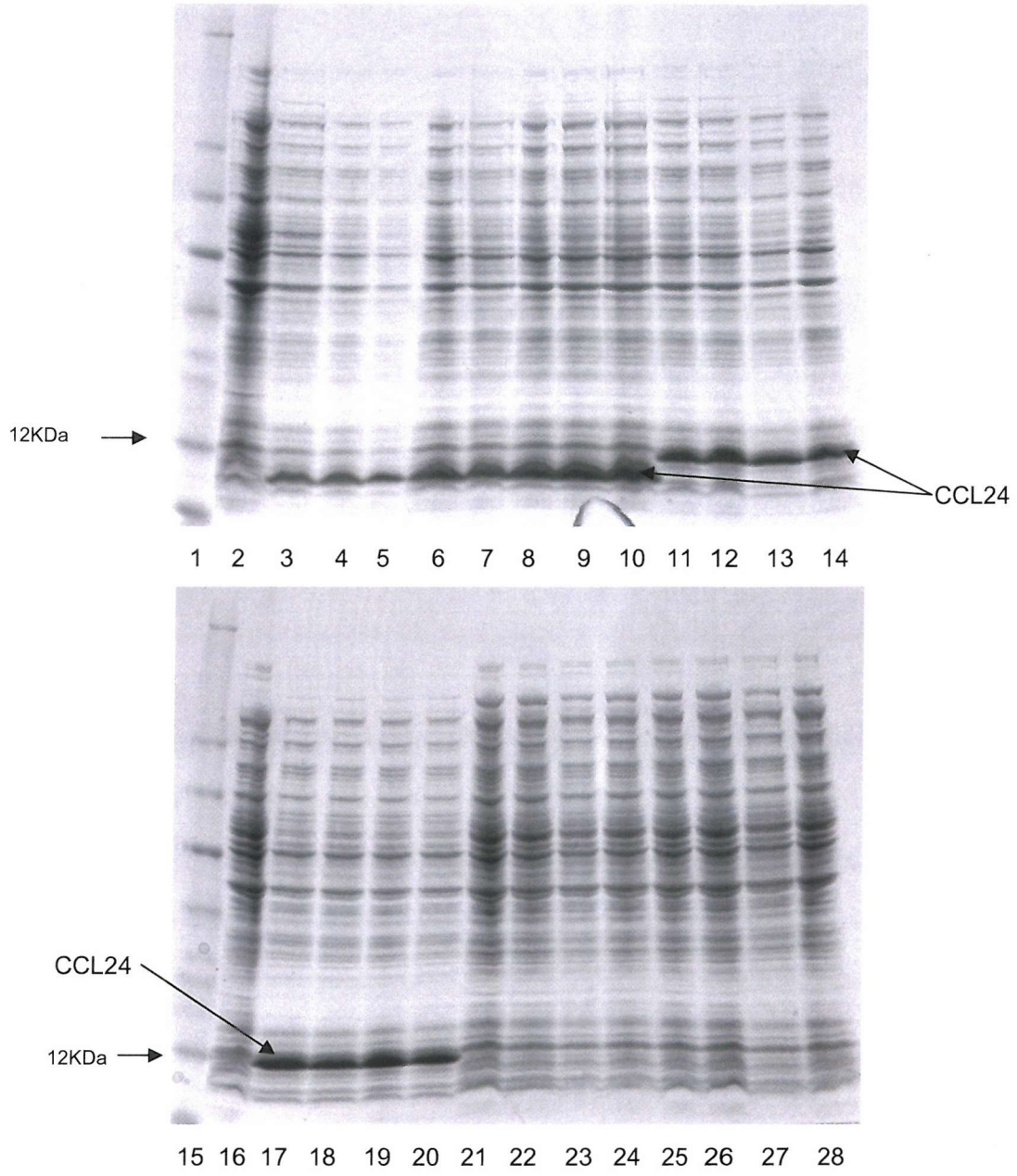
Table 3.3 The primers used for G25 to C119S. In both A and B the first line shows what was coded for, with the residues in brackets signifying only partial codons. The second line shows the bases 5' to 3' and the third line contains the antisense. The primer is in bold. A. The 5' primer. B. The 3' primer.

The third construct was based upon the same N-terminus as the second construct and used the same primer coding for a Nco1 restriction enzyme cut site, a deca-His tag and a thrombin cleavage site which overlapped with G25 and S26 (Table 3.3A). The C-terminus was moved to P98, which is before any of the reported truncations. The 3' primer included a BamH1 restriction enzyme cut site (Table 3.4). This was then ligated into pET-15b at the Nco1 and BamH1 cut sites (Figure 3.1).

	(Lys)	Gln	Lys	Lys	Ala	Ser	Pro	END	BamH1 Cut Site	Overhang	
5′	G	CAG	AAG	AAG	GCC	TCC	CCA	TAA			3′
3′	**C**	**GTC**	**TTC**	**TTC**	**CGG**	**AGG**	**GGT**	**ATT**	**CCTAGG**	**CGCGGG**	**5′**

Table 3.4 The 3' primer used for G25 to P98. The first line shows what was coded for, with the residues in brackets signifying only partial codons. The second line shows the bases 5' to 3' and the third line contains the antisense. The primer is in bold.

Both the G25 to C119S and the G25 to P98 plasmids overexpressed protein in the *E.coli.* strain BL21(DE3). Unfortunately the protein was produced as an insoluble inclusion body at 37°C. Inclusion bodies are normally formed when a protein is being produced too quickly for the bacteria to fold it correctly [146]. Slowing down the rate of production can give the bacteria time to fold it correctly, and thus solublize the protein *in vivo*. Changing the growth temperature to 30°C resulted in the majority of the protein being soluble. This was shown by SDS-PAGE analysis of crude cell lysates (Figure 3.2), with lanes 3 to 10 showing alternately the total protein extract and the soluble protein extract produced by the G25-P98 construct. Lanes 11 to 14 and 17 to 20 show the same result for the G25 to C119S construct.



Lane	Contents
1, 15	Markers.
2, 16	Uninduced.
3, 5, 7, 9	Total protein for the G25-P98 construct grown at 30°C.
4, 6, 8, 10	Soluble protein for the G25-P98 construct grown at 30°C.
11, 13, 17, 19	Total protein for the G25-C119S construct grown at 30°C.
12, 14, 18, 20	Soluble protein for the G25-C119S construct grown at 30°C.
21 to 28	Cell extracts for the M1-C119 construct grown at 37°C showing no induction band.

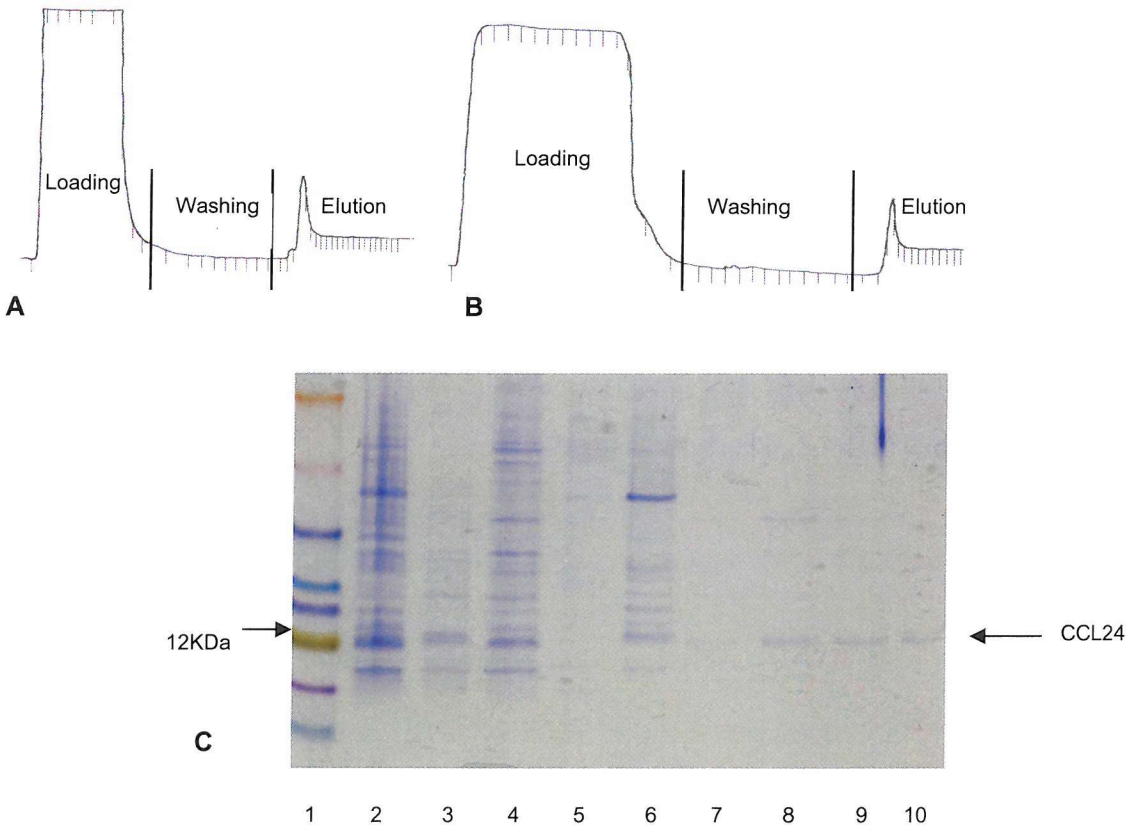
Figure 3.2 SDS-PAGE analysis of crude extracts grown at 30°C. Examination of lanes 3 and 4 showed the expression bands to be of similar intensity in both the total and soluble protein fractions, indicating that the G25-P98 construct was expressing soluble protein. Similarly lanes 11 and 12 showed the expression bands to be of similar intensity in both the total and soluble protein fractions suggesting that the G25-C119S construct was expressing soluble protein. Examination of lanes 21 and 22 showed no visible induction band, which suggested that M1-C119 did not express.

3.3.2 Initial Attempts at Protein Purification of the G25-P98 Construct

As those plasmids which expressed protein contained a deca-His tag, the initial purifications were attempted via application of the crude soluble protein to a nickel affinity column. A nickel affinity column should bind a His tagged protein, which can then be eluted using imidazole. FPLC was used to load, wash and elute from the column. Figure 3.3A shows a typical trace of the resulting absorbance at 280nm, whilst Figure 3.3B shows a control trace where the experiment had been run with no CCL24 present. The two traces appear similar and this was confirmed by SDS-PAGE analysis which showed that the majority of the protein had failed to bind (Figure 3.3C). Western blot analysis showed that the His tag was expressed, despite the failure of the protein to bind the nickel column.

Various attempts were made to increase the affinity between the protein and the column. One of the most common causes for His tag proteins failing to bind to their affinity columns is the His tag being "hidden" within the structure of the native protein [147]. There was a small increase in the amount of protein binding when the lysate was stirred with the nickel column media, but the majority of the His tagged protein once again did not bind to the column.

The deca-His tag was added to increase the strength of the interaction between the nickel column and the protein. This could possibly have caused the protein to bind so tightly that the standard elution condition for the hexa-His tag would not be strong enough to break the interaction between the protein and column. There were two different methods that could be used to avoid this problem: either the concentration of the elutant could be increased or the pH could be lowered below the pKa of the histidine side chain. When no more protein eluted upon using 2M imidazole or after passing buffer at pH5.5 through the column, it was concluded that the His tag was not interacting with the nickel column.



Lane	Contents
1	Markers.
2	Induced CCL24.
3	Supernatant after sonication.
4, 5	Protein that did not bind the nickel column.
6, 7	Protein that washed off the nickel column.
8, 9, 10	Protein that eluted off the nickel column after concentration.

Figure 3.3 FPLC traces (absorbance at 280nm) of two experiments run using the nickel column. The soluble cell extract was loaded and run through until a stable baseline was achieved before the column was washed with buffer containing 10mM imidazole. The protein was then eluted with 250mM imidazole. The shift in the baseline upon elution is due to the presence of high concentrations of imidazole which also absorbs at 280nm. A. FPLC trace from the purification of CCL24 . B. Standard FPLC trace when the nickel column was run with no CCL24 present. The SDS-PAGE analysis of the elutant (not shown) revealed only impurities which were probably non-specifically binding the column. C. SDS-PAGE analysis corresponding to Figure A showing that the majority of the protein did not bind.

The matrix to which the nickel was attached could also be charged with cobalt, zinc or copper instead of nickel [148]. It had been shown that different metal ions can bind different His tagged proteins to varying degrees. When none of these columns bound the protein with significantly greater affinity or specificity than the nickel column, it was concluded that the metal ion was not the cause of the lack of binding.

There was a possibility that the protein might be polymerising via its disulfide bonds, which could be causing problems. The introduction of a reducing agent would have prevented this, but the nickel column matrix was incompatible with the use of a reducing agent. The Talon cobalt matrix claimed to be stable under mild reducing conditions, so this matrix was tested by stirring with a crude lysate. There was no significant rise in the amount of protein binding.

It was proposed that if the column was run under denaturing conditions and there was still no significant rise in the amount of protein binding that the problem was with the His tag. 6M guanidine hydrochloride was added to all the buffers and the purification repeated. This experiment failed to significantly increase the amount of protein binding and it was concluded that a different method of purification was required.

A possible explanation for the column purification problems was highlighted when a sample of pure CCL24 with the His tag still attached was run on a strong cationic exchanger. At pH8 the protein eluted at numerous salt concentrations which suggested a polycharge distribution. Upon lowering the pH below 6 the distribution was largely removed, suggesting that histidines (side chain pKa 6), could be the cause of the problem (Figure 3.4).

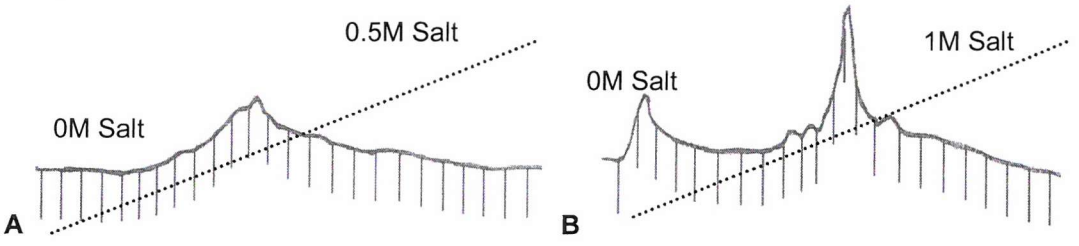


Figure 3.4 FPLC traces (absorbance at 280nm) for the same sample of CCL24 when it was eluted from a strong cationic exchanger using different pH buffers. A. The strong cationic exchanger run with pH8 buffer and the CCL24 eluted at numerous salt fractions which suggested a polycharge distribution. B. The strong cationic exchanger run using pH5.5 buffer where the CCL24 eluted over far fewer salt fractions, which suggested that histidines were the problem as the pKa of the histidine side chain is 6.

3.3.2.1 Purification by Exchange Chromatography

Mayer *et al* had shown that refolded CCL24 bound to a strong cationic exchanger at pH8 and would be eluted with 0.4M sodium chloride [21]. Figure 3.5 shows the FPLC trace for this experiment when performed using crude cell lysate; unexpectedly, CCL24 did not bind even after the use of ammonium sulphate cuts. The experiment was repeated using a strong anionic exchanger. In this case CCL24 eluted over most of the salt gradient as

shown by the FPLC trace (Figure 3.6) and corresponding SDS-PAGE analysis (Figure 3.6B). These results suggested that the protein sample had a polycharge distribution. As CCL24 had not bound to the strong cationic exchanger, it was reasoned that this might have been caused by some form of interaction between CCL24 and an unidentified component or components of the crude lysate, which were changing the overall charge on CCL24 by varying amounts.

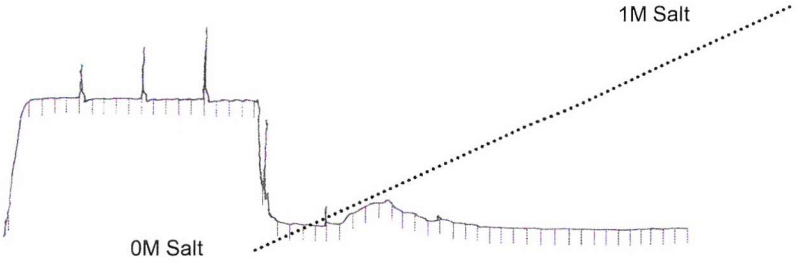


Figure 3.5 FPLC trace (absorbance at 280nm) of the crude cell extract eluted from a strong cationic exchanger using a 0-1M salt gradient run. The SDS-PAGE analysis (not shown) indicated CCL24 did not bind to the column and eluted in the void volume.

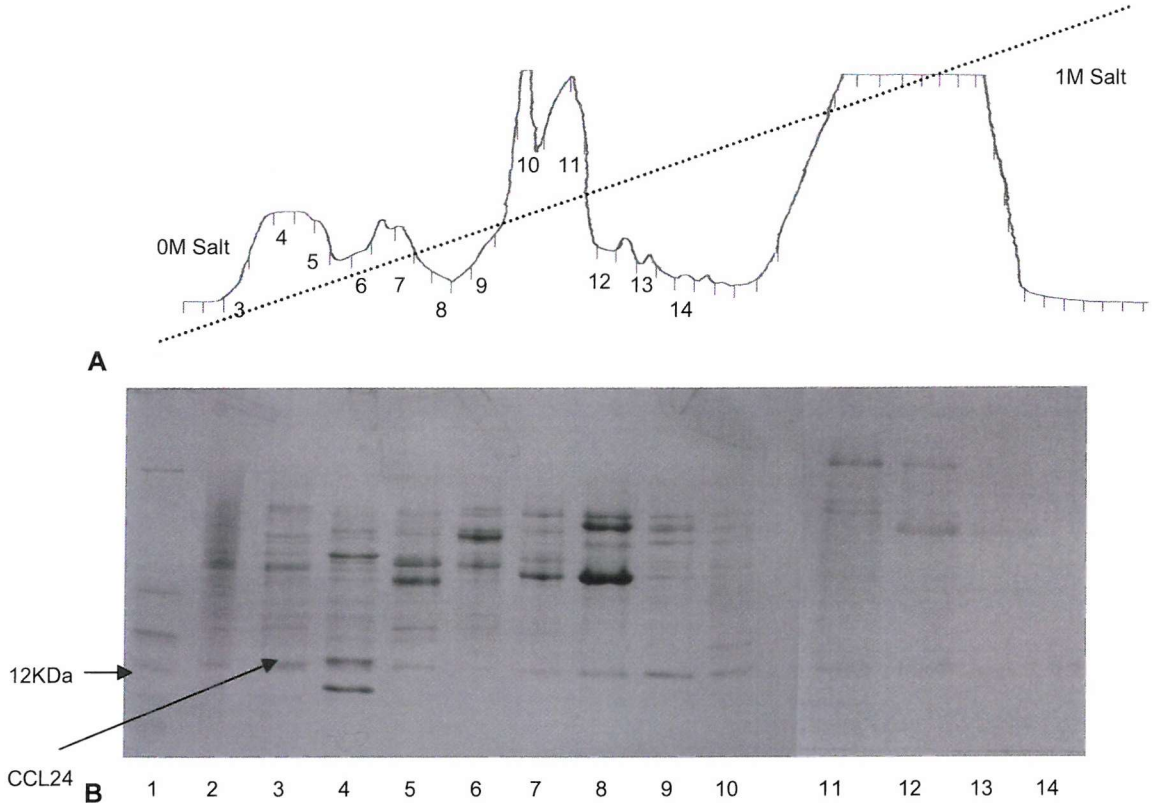
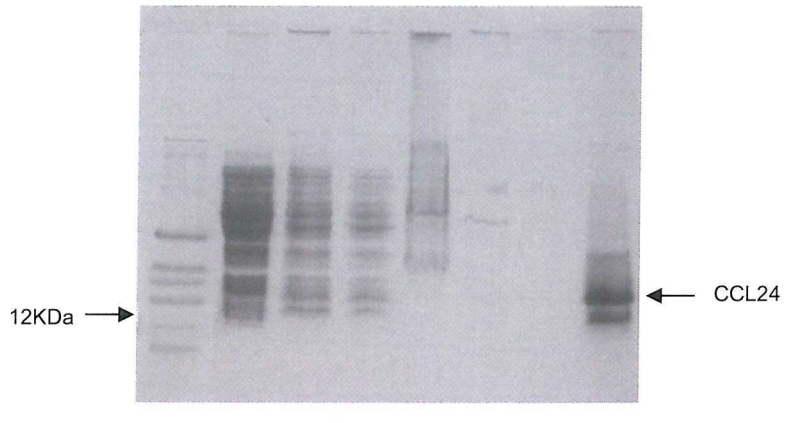


Figure 3.6 A. FPLC trace (absorbance at 280nm) of the crude cell extract eluted from a strong anionic exchanger using a 0-1M salt gradient. CCL24 eluted in numerous salt fractions. The fraction labels correspond to the lanes in Figure B. B. SDS-PAGE analysis of Figure A. Lane 1 shows the markers, Lane 2 the loaded supernatant and the other lanes are alternate 10mL fractions as marked on the trace in Figure A.

3.3.3 Successful Purification from Inclusion Bodies

With the other methods failing it was decided to attempt the purification from the insoluble material produced when the BL21(DE3) cells had been grown at 37°C. Whilst inclusion bodies could harbour relatively pure protein in large amounts, they do require the refolding of the protein. Mayer *et al* had successfully refolded a small amount of CCL24 using a dialysis methodology [21]. An attempt to follow this method resulted on only 1% recovery of correctly refolded material and as such a new methodology was devised.

The purity of the inclusion body before attempting to refold a protein has a direct effect on the yield of the final product [149]. The major impurities of an inclusion body have been shown to be peptidoglycans, lipids, nucleic acids, lipopolysaccharides and membrane proteins [150]. A protocol adapted from Bohmann *et al* [151] minimised these impurities, particularly membrane proteins, by repetitively disrupting the cell pellet whilst it was suspended in different buffers. Figure 3.7 shows a SDS-PAGE analysis of the soluble fractions removed from the insoluble pellet following these successive disruptions. The final pellet was resuspended in clean buffer and recentrifuged, to remove the residual supernatant. This methodology produced a pellet with an estimated purity of 95% by SDS-PAGE analysis.



Lane	Contents
1	Markers.
2	Supernatant from the first sonication (50mM KCl, 1mM DTT, 1mM PMSF, 35mM Tris pH8.5, 12% (w/v) sucrose, 20mM EDTA, 0.5mg/mL lysozyme, 0.5M LiCl, 0.25% (v/v) NP-40).
3	Supernatant from the second sonication (10mM Tris-HCl pH8.5 0.1mM EDTA, 0.5M LiCl, 0.5% (v/v) NP-40, 1mM DTT, 1mM PMSF).
4	Supernatant from the third sonication (10mM Tris-HCl pH8.5 0.1mM EDTA, 0.5M LiCl, 0.5% (v/v) NP-40, 1mM DTT, 1mM PMSF).
5	Supernatant from the fourth sonication (10mM Tris-HCl pH8.5, 0.1mM EDTA, 0.5% (v/v) NP-40, 1mM DTT, 1mM PMSF).
6	Supernatant from the fifth sonication (10mM Tris-HCl pH8.5, 0.1mM EDTA, 0.5% (v/v) NP-40, 1mM DTT, 1mM PMSF).
7	Supernatant from washing the pellet (10mM Tris-HCl pH8.5, 0.1mM EDTA, 0.5% (v/v) NP-40, 1mM DTT, 1mM PMSF).
8	Final pellet containing CCL24.

Figure 3.7 SDS-PAGE analysis showing the soluble fractions from successive sonications where the pellet, containing the inclusion body, was resuspended in different buffers to remove as many impurities as possible.

3.3.3.1 Refolding Methodologies

An unfolded protein cannot search all possible conformational states within the time scale of folding; hence it is generally accepted that proteins follow a finite number of pathways [152]. Whilst a protein is refolding it will go through these various intermediate states until it finds its correctly folded structure. These intermediates tend to form undesirable aggregates via hydrophobic areas of the protein [153-155]. It is at intermediate denaturant concentrations (2-4M GdmCl) that most aggregation takes place [156].

In the kinetic competition between the folded and aggregated states, the folding process is first order whilst the aggregating process is of a higher order as it involves at least two molecules; hence aggregation increases faster with respect to the concentration [157, 158]. Yields of correctly refolded protein are strongly dependent on the initial concentration of the denatured protein [153].

The most effective way to reduce aggregation is to refold at low protein concentration and low temperature. For many proteins this concentration would be between 10 and 50μg/mL which leads to large volume handling problems. The introduction of certain additives into the renaturing buffer has been shown to help prevent aggregation; examples of these additives are L-arginine [159], sugars [160] and low concentrations of denaturant (less than 2M guanidine hydrochloride) [161]. The observation that low concentrations of denaturant helps refolding but intermediate concentrations aid aggregation, means that dialysis is not an ideal methodology as it leaves the protein exposed to intermediate concentrations of the denaturant for long periods of time. In comparison, the direct dilution of the denatured protein into the refolding buffer leads to a much shorter exposure period [162, 163].

It has been suggested that there is a relationship between the concentration of the denaturant, the concentration of the protein, and the time needed to refold the protein (Figure 3.8) [164].

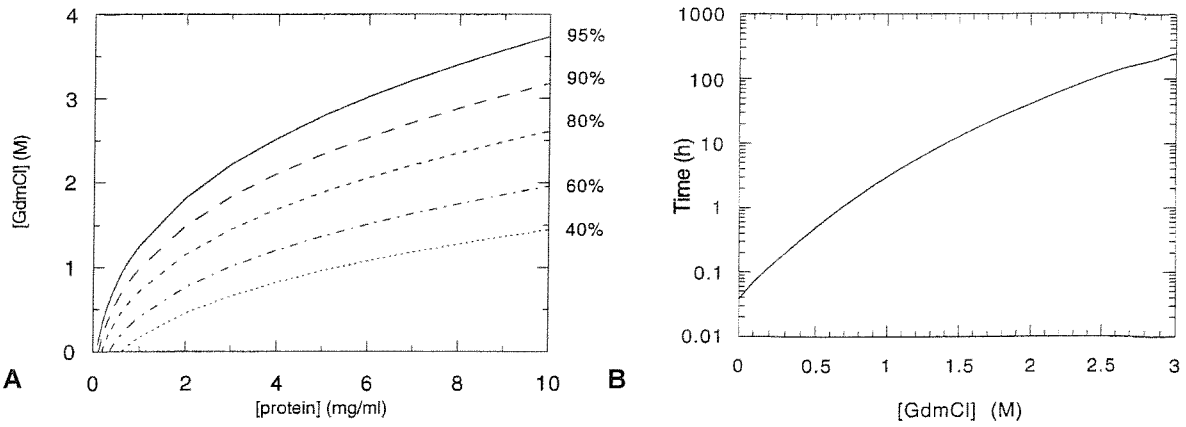


Figure 3.8 The relationship between the concentration of protein, the concentration of denaturant and the time required to refold the protein. A. Concentration of the denaturant against the concentration of the protein with the lines indicating the highest possible percentage yield for the reaction conditions. B. Time needed to refold the protein against the concentration of the denaturant to give a 98% yield of refolded protein.

Based on this information a protocol was devised that refolded the protein at 1mg/mL in 1.5M guanidine hydrochloride, by adding the protein drop wise to the refolding buffer at 4°C and leaving the protein at room temperature for eighteen hours. This resulted in 80% of the protein still remaining in the soluble fraction.

3.3.3.2 Purification of the Partially Refolded Material

The protein was then purified using a cationic exchanger before reforming the disulfide bonds. A typical trace for the cationic exchanger can be seen in Figure 3.4B. The disulfides in chemokines have been reformed by various methods, with the most popular being either air oxidation or oxido-shuffling via reduced and oxidized glutathione (GSH/GSSH). Both methods have their own pros and cons, with air oxidation being significantly cheaper and easier to perform, whereas the redox nature of the oxido-shuffling enables the formation of a much higher yield of the correctly disulfide bonded protein. Both air oxidation and oxido shuffling were attempted. There was a significant improvement in the yield when using the oxidized glutathione. Separation of the oxidised from the reduced forms was performed using a C^{18} column on a HPLC with an acetonitrile gradient (Figure 3.9). The oxidised bonded form eluted first, at 30% acetonitrile, as shown by the mass spectrometry results (Figure 3.10).

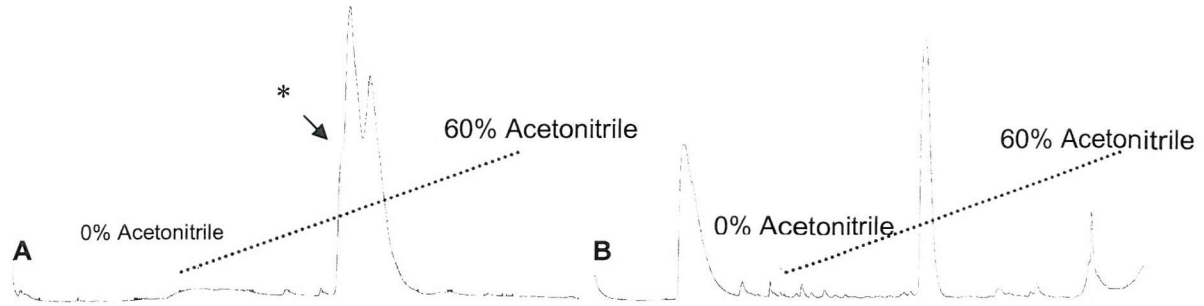


Figure 3.9 HPLC traces (absorbance at 216nm) showing the separation of the oxidised protein from the reduced protein using an 0-60% acetonitrile gradient. A. HPLC trace when air oxidation had been used to reform the disulfides and a mixture of oxidised and reduced forms were present. The first peak that eluted was the oxidised form and is marked with an asterisk. B. HPLC trace when oxido-shuffling had been used to reform the disulfides; only the oxidised form was present. The peak at the start of the trace was due to an injection artifact, which was often observed when using the HPLC.

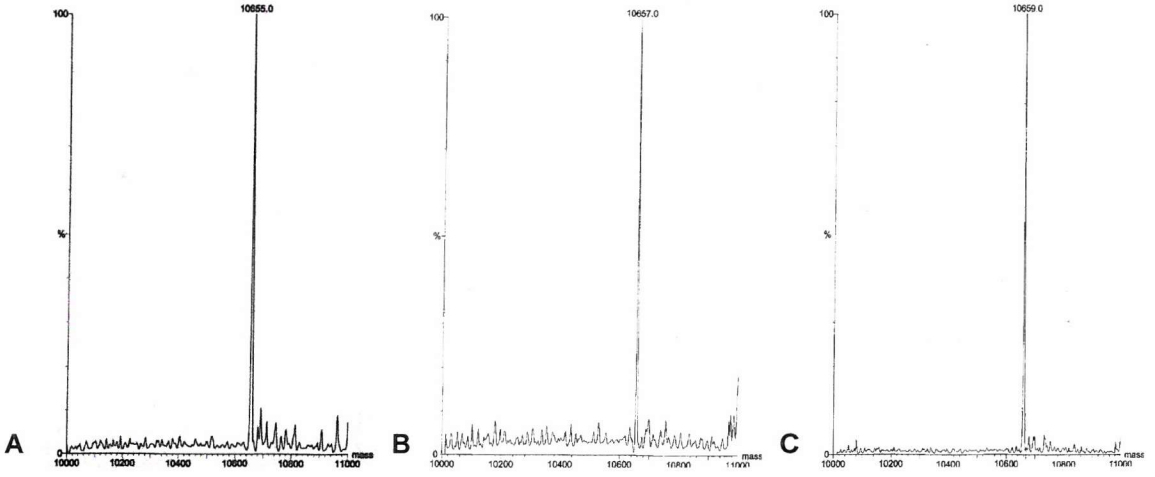
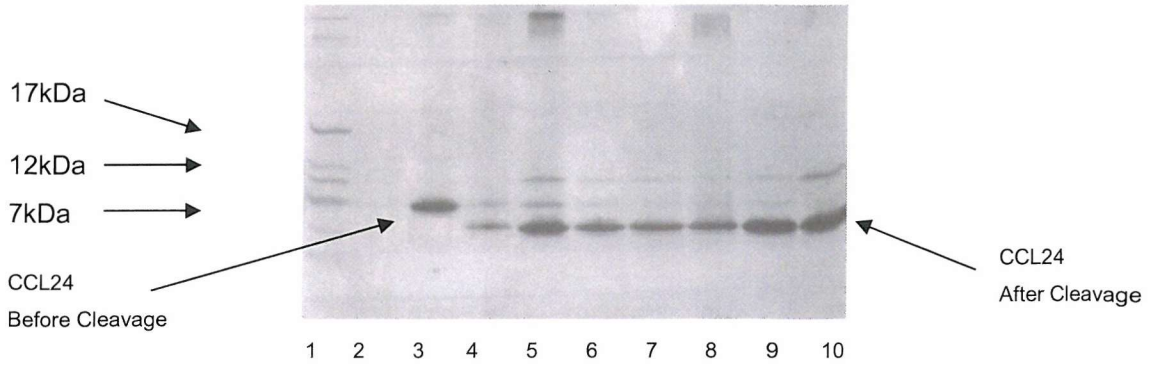


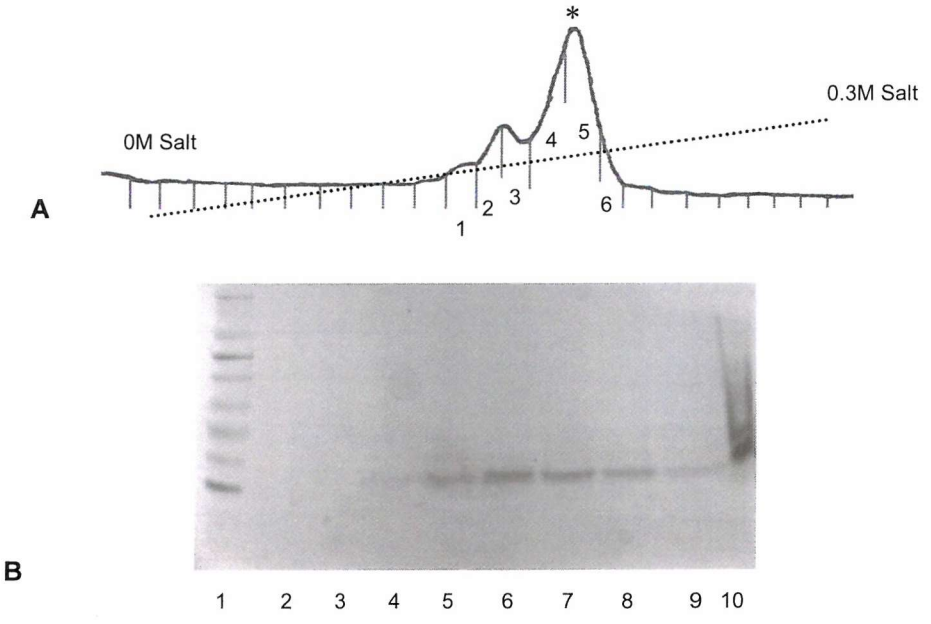
Figure 3.10 The Maxent of the mass envelope of quadrupolar electrospray mass spectrometry. A. Mass of the oxidised protein (10655±0.5Da). B. Mass of a protein with one disulfide formed (10657±0.5Da). C. Mass of the reduced protein (10659±0.5Da).

The His tag was then cleaved using thrombin. This produced a mixture containing several products (Figure 3.11), most of which could be separated using a shallow gradient on a cationic exchanger (Figure 3.12). The final mass of fraction five was 8355±0.5Da.



Lane	Contents
1	Markers.
2	Empty.
3	CCL24 after 0 hours with thrombin.
4	CCL24 after 1 hour with thrombin.
5	CCL24 after 2 hours with thrombin.
6	CCL24 after 3 hours with thrombin.
7	CCL24 after 4 hours with thrombin.
8	CCL24 after 5 hours with thrombin.
9	CCL24 after 6 hours with thrombin.
10	CCL24 after 7 hours with thrombin.

Figure 3.11 SDS-PAGE analysis of the effect of thrombin on the protein with time.



Lane	Contents
1	Markers.
2	Void volume.
3	Void volume.
4	Fraction 1.
5	Fraction 2.
6	Fraction 3.
7	Fraction 4.
8	Fraction 5.
9	Fraction 6.
10	Sample loaded onto column.

Figure 3.12 A. FPLC trace (absorbance at 280nm) of the separation of the cleavage products on a strong cationic exchanger using a 0-0.3M salt gradient. The separation of CCL24, marked with an asterisk, from the other cleavage products is achieved by the use of a shallow salt gradient. B. SDS-PAGE analysis corresponding to Figure A.

3.3.4 Production of 15N Labelled CCL24

M9 minimal growth media using ^{15}N ammonium chloride as the sole nitrogen source was used to introduce the NMR-active isotope of nitrogen. Unfortunately the initial attempts at using M9 minimal media resulted in the production of soluble protein. It was proposed that there were two possible reasons for this. One: that the drop in nutrients was retarding the production of the protein enough to stop it going into inclusion bodies, in a similar way that dropping the temperature had done. Two: that the change in media had caused a shock to the *E.coli* and slowed down the initial rate of production of CCL24. The growth was repeated, but this time upon changing the media the bacteria were left to grow for thirty minutes, to allow the bacteria time to adjust to the media before induction. This produced CCL24 mostly in the insoluble form. The protein was then refolded and purified using the above techniques giving a yield of 4mg per litre of culture, before thrombin cleavage.

3.4 Titration Experiments with the Extracellular Loops

The protein had previously been assigned and the chemical shifts published (Figure 3.13). As these differed slightly from the chemical shifts observed for the native protein in acetate buffer, the assignment was checked using a 3D NOESY-HSQC. With the assignment verified, a standard was run of the protein in DPC before any titrations with the loops were carried out (Figure 3.14). The titrations were conducted with the peptide mimic in the NMR tube. The protein was titrated into the NMR tube increasing its concentration with each titration step. This means that in the initial state the peptide is in a large excess compared to the low concentration of protein, which in turn means that the protein is in its bound form and that the peptide is in its unbound form. As the titration continues it passes through a region where both the peptide and the protein are in their bound forms, to end in a region where the protein is in a large excess compared to the lower concentration of peptide. Now the peptide is in its bound form and the protein is in its unbound form. Assuming that enough protein was added to take the titration to the desired end point, then the first point HSQC showed the bound form of the protein and the final HSQC showed the unbound form of the protein. The ^{15}N-filtered experiments had to be run at the end of the titration in order to observe the bound form of the peptide.

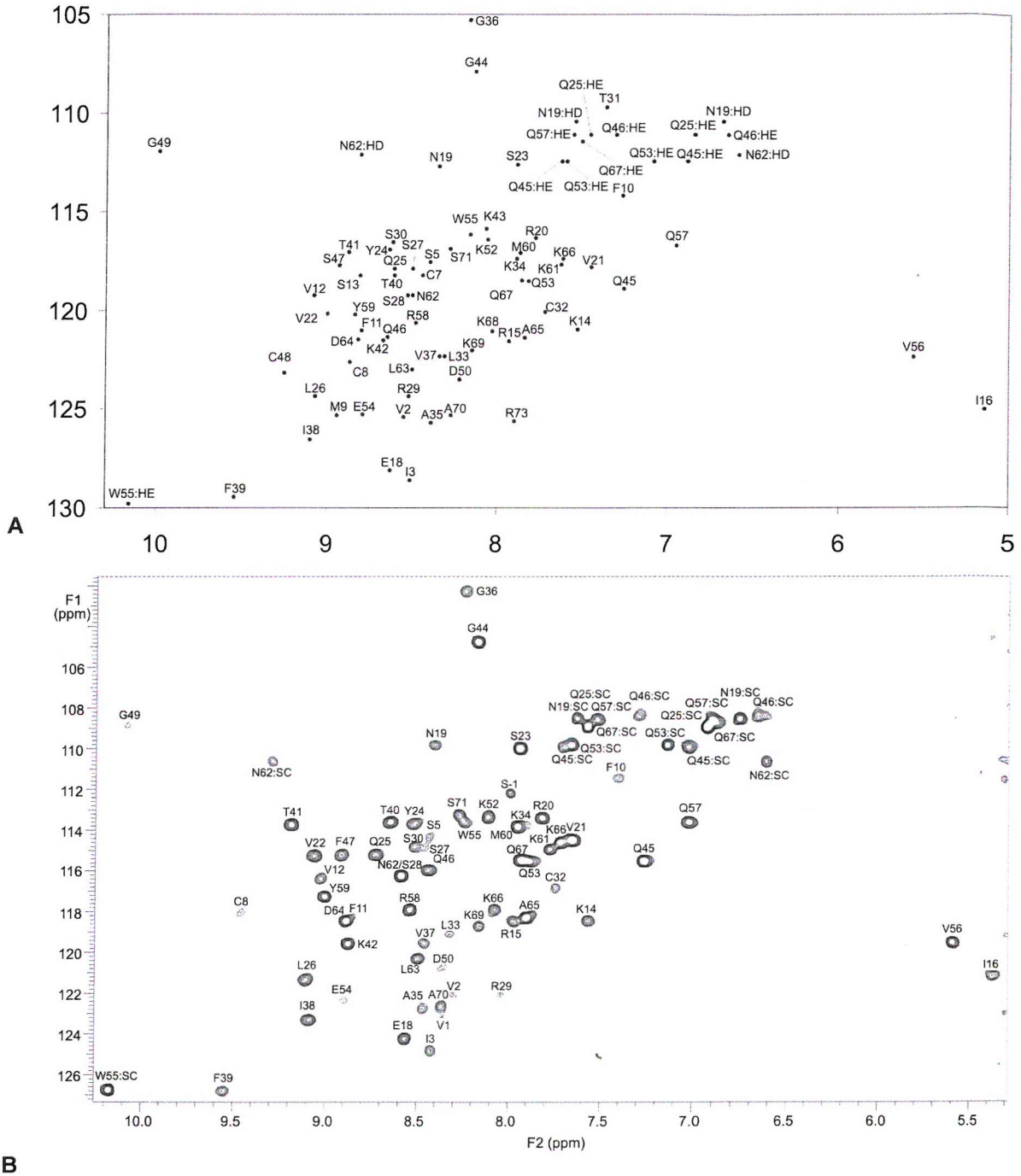


Figure 3.13 HSQCs of CCL24. A. Representation of the HSQC from published values. B. HSQC of refolded CCL24 in DPC showing the region between 5.3 to 10.25ppm and 102 to 127ppm. The spectrum was zero filled from 1024 to 2048 points in F2 and was zero filled from 64 to 512 points in F1. The Varian window functions were SB (-0.031), SBS (-0.02), SB1(-0.036) and SBS1(-0.032).

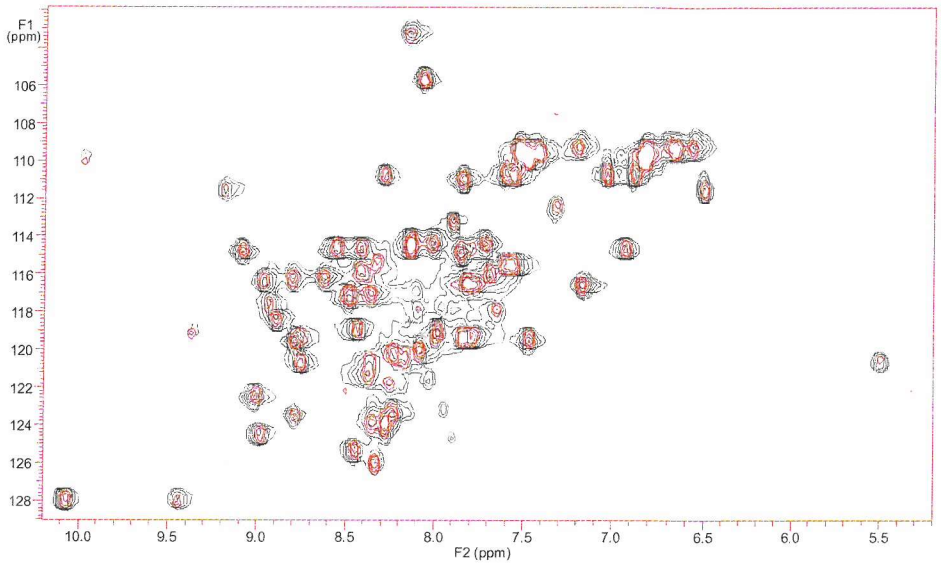


Figure 3.14 Standard titration of CCL24 into DPC micelles showing the first point (0.1mM) in red and the last point (1mM) in black. The vertical scales have been normalised to account for the change in concentration.

3.4.1 CCL24 Titration with Loop 1

CCL24 was titrated into a sample of 0.5mM peptide mimic of loop 1 suspended in 100mM acetate, 150mM DPC pH5.5. Five HSQCs were recorded at CCL24 concentrations of 0.05mM, 0.1mM, 0.25mM, 0.5mM and 1mM. A 2D-NOESY was recorded before the start of the experiment and a ^{15}N-filtered NOESY recorded at 1mM CCL24. An overlay of the first and last HSQCs recorded is shown in Figure 3.15 with the vertical scales adjusted in an attempt to normalize the peak intensities. The data showed that there were both chemical shift and intensity changes of the peaks during the titrations. Subsequent analysis of this data was performed in several stages.

It was necessary that the chemical shift deviations were concentration dependent to give a consensus K_d. Initially for the amide resonances the weighted chemical shift deviations between the first and last points were examined (Figure 3.16), but several peaks appeared to move in the first and second HSQCs only. It was possible that these were the result of the long acquisition time, the poor signal to noise ratio and the difficulty in measuring an accurate chemical shift on a noise distorted crosspeak in these very weak samples (0.05mM and 0.1mM CCL24). To account for this each residues amide hydrogen chemical shift deviation was examined through the course of the titration and those that had significant chemical shift deviation (V1, V2, F10, V12, R15, N19, Q25, S27, C32, Q46, S47, Q53, E54, K61, L63, K69 and S71) were further examined by fitting their

binding isotherms (Figure 3.17A) to reveal those that gave a consensus K_d (V2, V12, R15, Q25, K69 and S71). When the side chain resonances were examined in the same manner, there were several that showed significant chemical shift deviation (Q25, Q45, Q57, N62, Q57) but only the binding isotherm of Q45 matched those of the backbone amide resonances. The average value taken from the binding isotherms calculated the K_d to be 30±30μM.

In addition to the chemical shift change some peaks showed changes in intensity during the titration. To quantify this each residues peak intensity was examined through the course of the titration, after accounting for the changes in concentration and the number of scans, whilst assuming the window functions effect upon the peaks intensities was consistent through out the titration. Those that might have had a significant change could be split into three different classes (Figure 3.17B, C and D). Class 1: residues that showed a large change in peak intensity and comparatively small error bars which corresponded to a significant change (V2, I3, L33, W55, K69, A70). Class 2: residues that showed a small change in peak intensity and comparatively large error bars; as such might or might not have corresponded to a direct interaction (N62 and the side chain of Q45). Class 3: residues that could have a horizontal line drawn between the error bars of their first and last points, but there was a deviation somewhere else in the course of the titration (T41, E54, K68 and the side chains of W55, Q57 and Q67). Whether this is a problem arising due to the long acquisition time and the poor signal to noise ratio of the first and second HSQCs or whether this is a real effect is unknown.

As some peaks undergo intensity changes whilst others are shifting in position it is possible that some of the peaks are undergoing exchange on a slow time scale whilst others are undergoing exchange on a fast time scale. It is probable that they can be accounted for by a single dissociation rate and that the behavior arises from the dependence of the line shape on the chemical shift difference between the bound and unbound states. Larger chemical shift differences lead to the loss of intensity whilst smaller chemical shift deviations lead to a gradual shifting of the peak [165].

When all the information was collated and applied to the CCL24 structure (Figure 3.18) there did not appear to be a defined binding site, with those residues affected by the loop

titration spread across the protein in three major clusters. One cluster was on the C-terminus, another was on the N-terminus whilst the third was at the top of the α-helix.

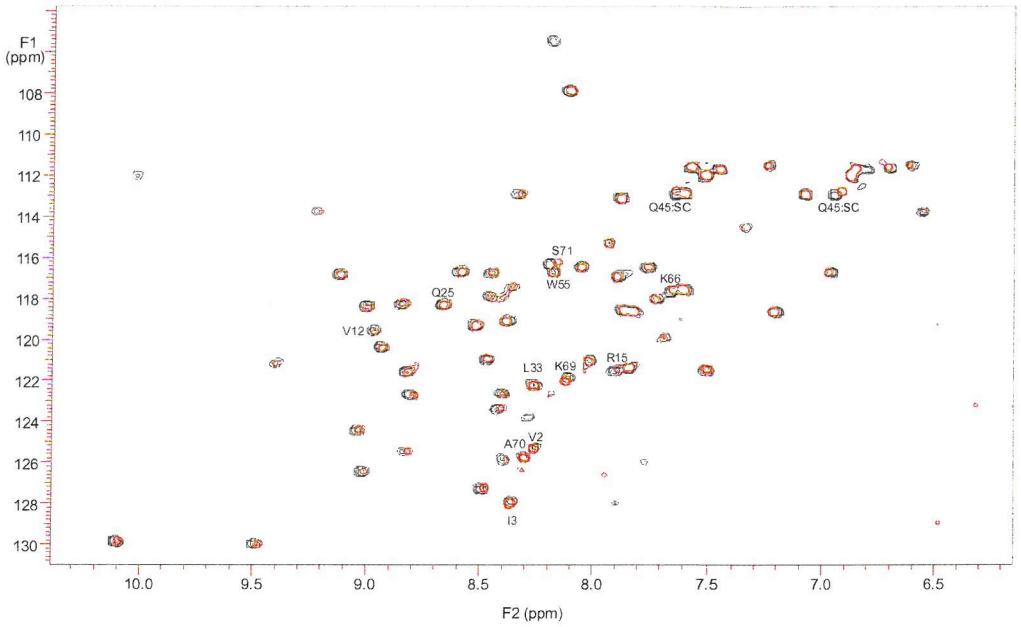


Figure 3.15 HSQC spectra overlay for the titration of CCL24 into loop 1 where an attempt has been made to normalise the vertical scale. The first point is in red (bound form) and the final point is in black (unbound form) showing the region between 6.15 to 10.4ppm and 104 to 131ppm (V56 and I16 are off scale). The marked residues are those that underwent either a concentration dependent chemical shift to give a consensus K_d, or a concentration dependent intensity change. Both spectra were zero filled from 1024 to 2048 points in F2 and were zero filled from 64 to 512 points in F1. The Varian window functions were SB (-0.034), SBS (-0.02), SB1(-0.023) and SBS1(-0.009).

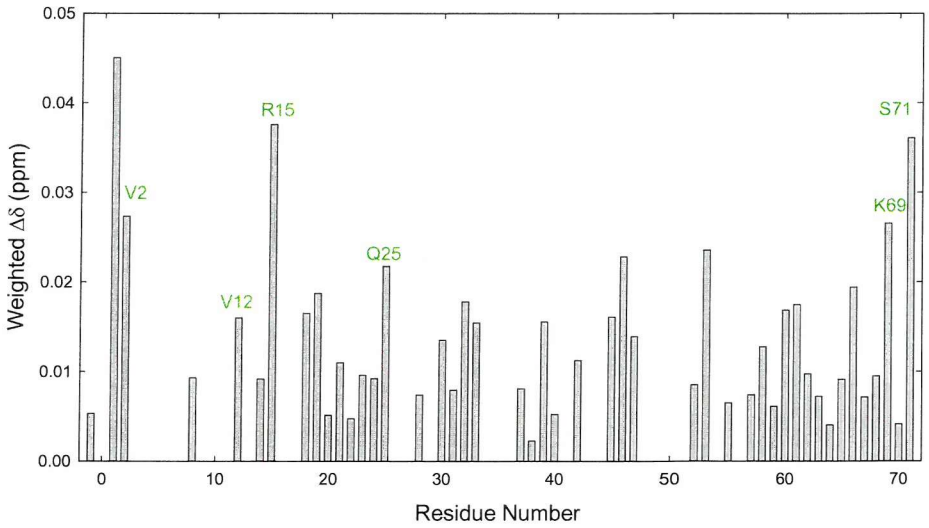


Figure 3.16 The weighted chemical shift difference between the first and last points for the backbone resonances of CCL24 from its titration into loop 1. The residues marked are those that showed concentration dependent chemical shift change and whose binding isotherms gave a consensus K_d (30±30μM). The residues were V2, V12, R15, Q25, K69 and S71.

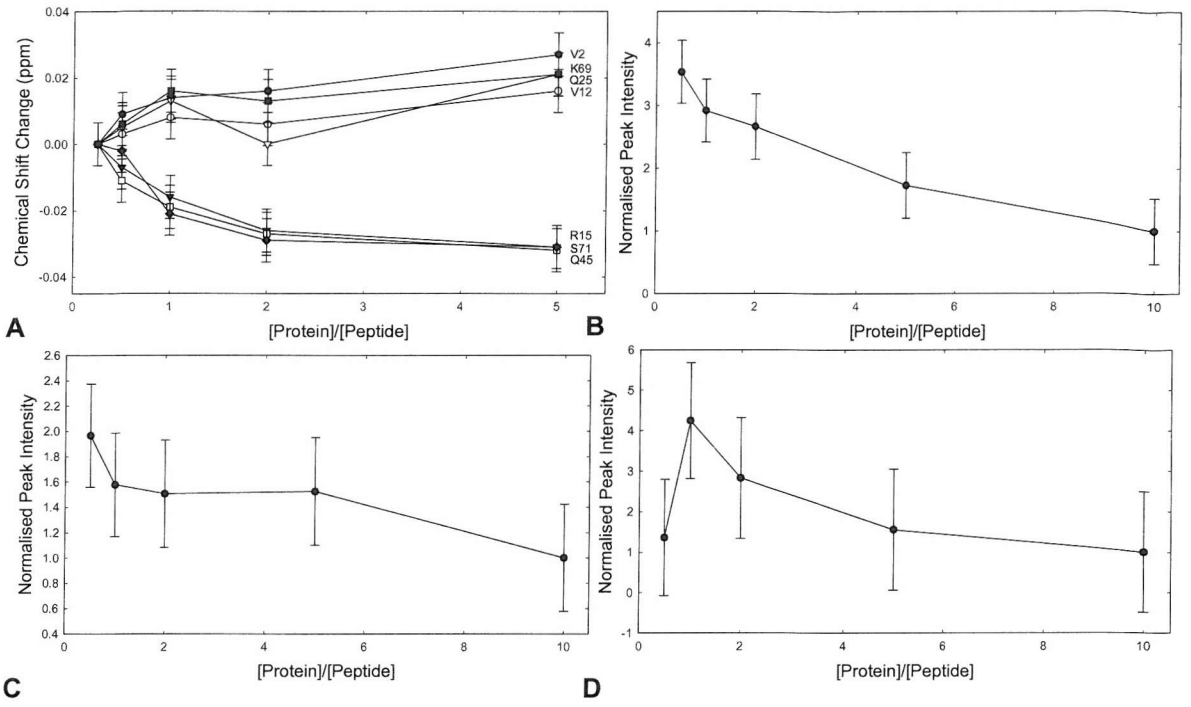
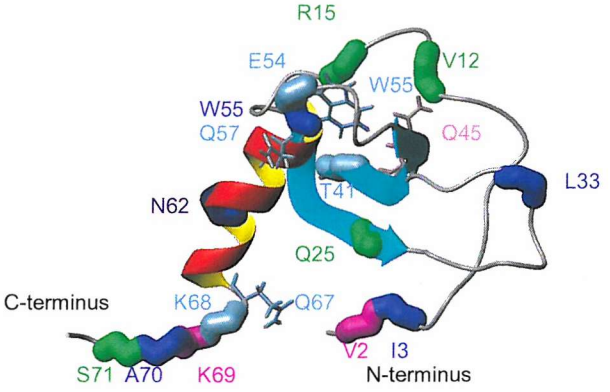


Figure 3.17 The intensity and chemical shift changes through the titration plotted as a function of protein/peptide ratio. A. Binding isotherms obtained upon titration of CCL24 with loop 1 for the amide protons of V2 (filled circles), V12 (open circles), R15 (filled triangles), Q25 (open triangles), K69 (filled squares), S71 (open squares) and the side chain of Q45 (filled diamond). Error bars correspond to the ^{1}H digital resolution of the HSQC spectra. B. Class 1 shown by W55 (from the loop 1 titration) where the 95% error bars are small and there is a large difference between them. C. Class 2 shown by N62 (from the loop 1 titration) where the 95% error bars are large and there is a small difference between them. D. Class 3 shown by E54 (from the loop 1 titration) where a horizontal line could be put through the 95% error bars of the first and last points but there is deviation elsewhere.



Residues colour coded to show those involved in:

Bright blue: Class 1 intensity change (I3, L33, W55 and A70).

Bright green: Chemical shift deviation (V12, R15, Q25 and S71).

Dark blue: Class 2 intensity change (N62).

Light blue: Class 3 intensity change (T41, E54, K68 and the side chains of W55, Q57 and Q67).

Light pink: Both a chemical shift deviation and a Class 2 or 3 intensity change (side chain of Q45).

Purple: Both a chemical shift deviation and a Class 1 intensity change (V2 and K69).

Figure 3.18 CCL24 with the residues affected by the loop 1 titration colour coded. There are three major clusters: one cluster is on the C-terminus, the second is on the N-terminus, whilst the third is at the top of the α-helix.

An overlay of the spectra for the peptide (Figure 3.21) showed the end point spectrum as being largely denuded of peaks with there being no cross-peak to the amide hydrogen of residues F3, W4, I5, R9, W13, V14, F15, G16, H17, G18, M19, G25, F26, H28 and T29. Of the peaks that were observed those corresponding to V8, N12 and A20 had a significant change in chemical shift whilst there was no significant observed shift in the peaks corresponding to H6, Y7, G10, H11, K21, L22, L23, S24 and Y27 (Figure 3.19). Whilst the disappearance of all these peaks could be linked to the binding of CCL24, it is strange that there are so many residues affected on the peptide when compared with so few visible changes on the CCL24 spectra. It is possible that the poorer water suppression, which is characteristic of the ^{15}N-filtered experiments, had resulted in weak peaks remaining in the noise and also in the saturation of some peaks and their nearest neighbours.

When the information was collated onto the loop 1 structure it indicated that there were two sites affected; one site consists of the N-terminus and the C-terminus whilst the second site consists of residues at the top of the loop. It is possible that both sites were involved in the binding, but more likely that the intensity changes in the N-terminus and the C-terminus were due to the poor signal to noise ratio and poor water suppression. As examples of this the W4 peak was an extremely weak originally and the H28 peak lies very close to the water resonance. It was also noted that the aromatic side chains of F3 and W4 are unaffected whilst the side chain of W13 seemed to have disappeared.

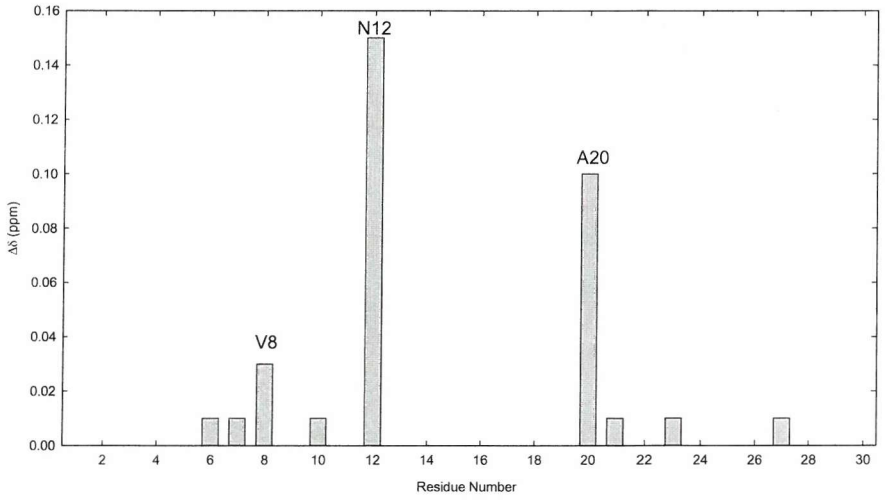
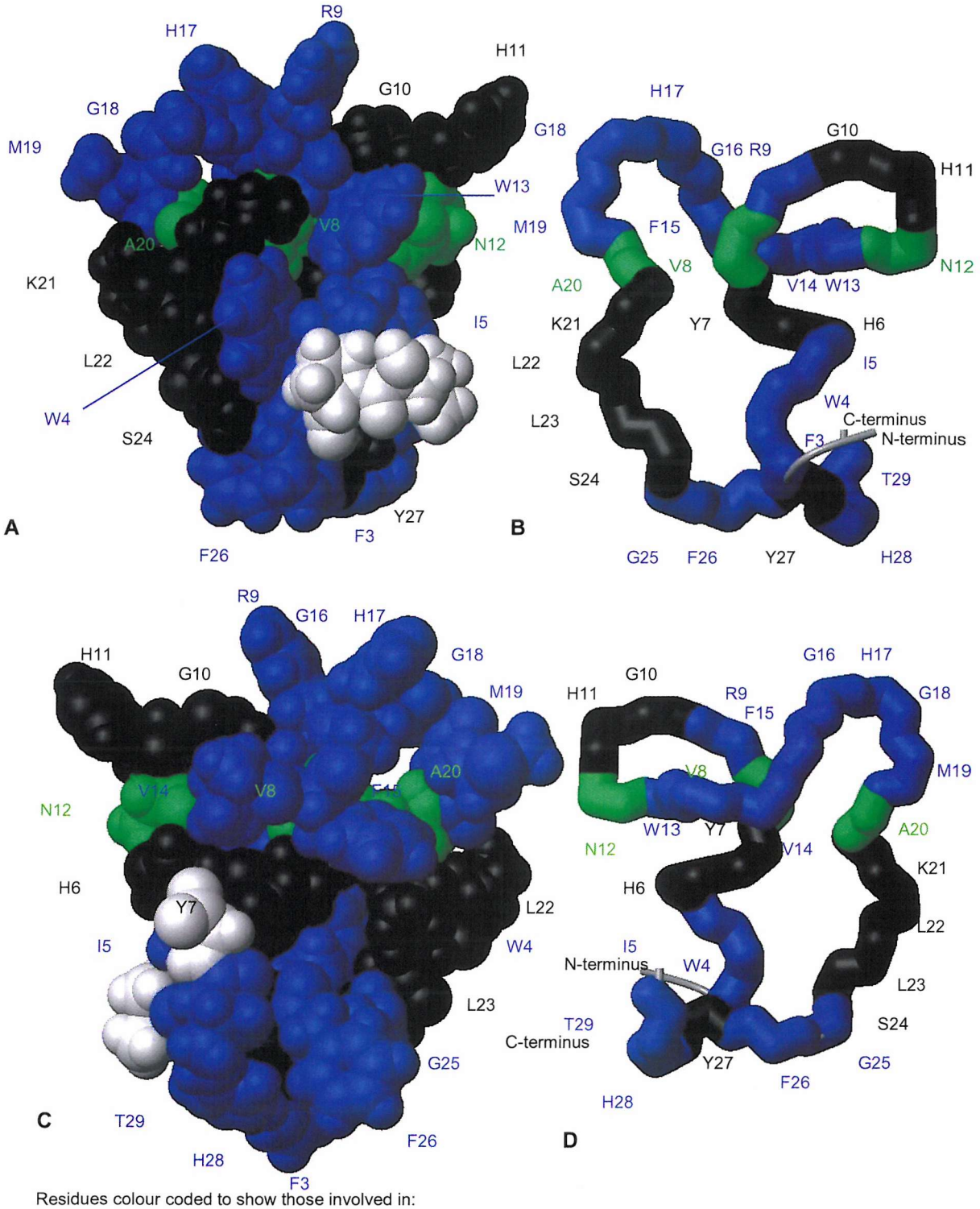


Figure 3.19 The chemical shift deviation of the amide hydrogen between a standard of the loop 1 peptide and a ^{15}N-filtered NOESY run at the end of the titration. Only V8, N12 and A20 showed a significant deviation in their chemical shifts. However there was no observed amide peak in the bound form for F3, W4, I5, R9, W13, V14, F15, G16, H17, G18, M19, G25, F26, H28 and T29.



Residues colour coded to show those involved in:

Black: Unaffected by the titration (H6, Y7, G10, H11, K21, L22, L23, S24 and Y27)

Blue: Intensity change (F3, W4, I5, R9, W13, V14, F15, G16, H17, G18, M19, G25, F26, H28 and T29).

Green: Chemical shift deviation (V8, N12 and A20).

Figure 3.20 The structure of the peptide mimic of loop 1 with those residues that were affected by the titration colour coded. It would appear there are two sites affected by the titration. One site consists of the N-terminus and the C-terminus. The second site consists of residues at the top of the loop. A. Space filled model of loop 1 showing the side that is probably facing inwards towards CCL24. B. Same as Figure A except loop 1 is now in ribbon. C. Same as Figure A except rotated by 180°. D. Same as Figure C except loop 1 is now in ribbon.

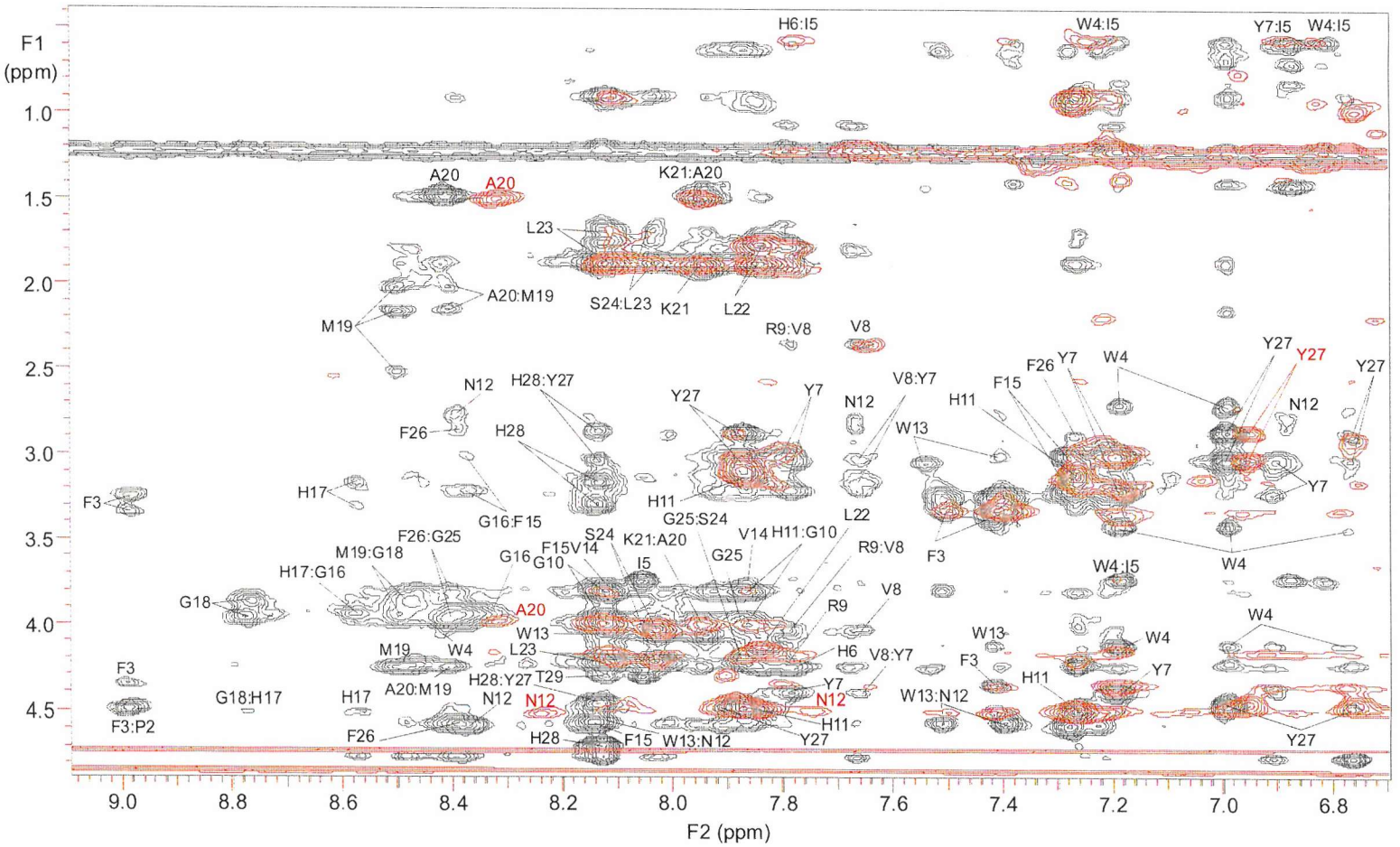


Figure 3.21 NOESY spectra overlay for loop 1 showing the region from 6.7 to 9.85ppm and 0.4 to 4.9ppm. The spectrum taken before the titration is in black (unbound form) and after the addition of CCL24 is in red (bound form). The unbound spectrum is fully labelled in black, whereas in the bound spectrum only those peaks that are significantly perturbed from their original positions that are labelled in red. The unbound spectrum was zero filled from 4096 to 8192 points in F2 and was zero filled from 384 to 2048 points in F1. The Varian window functions were LB (0), GF (0.037), LB1(0) and GF1(0.018). The bound spectrum was zero filled from 4096 to 8192 points in F2 and was zero filled from 256 to 2048 points in F1. The Varian window functions were LB (-5), GF (0.02), LB1(-20) and GF1(0.010).

3.4.2 Loop 3

CCL24 was titrated into 0.5mM peptide mimic of loop 3 suspended in 100mM acetate, 150mM DPC pH5.5. Six HSQCs were recorded at CCL24 concentrations of 0.05mM, 0.1mM, 0.25mM, 0.5mM, 1mM and 2mM. A 2D-NOESY was recorded before the start of the experiment and ^{15}N-filtered NOESY spectra recorded at 1mM and 2mM CCL24. An overlay of the first and last HSQCs recorded is shown in Figure 3.22 with the vertical scales altered in an attempt to normalize the peak intensities. As with loop 1, the data showed there were some chemical shift and intensity changes of the peaks during the titrations and the analysis was performed in several stages.

The amide resonances of the weighted chemical shifts between the first and last points were examined (Figure 3.23), but once again several peaks appeared to move in the first and second HSQCs only. So those that had significant chemical shift deviation (V2, R15, I16, N19, S27, K42, Q46, Q53, L63, Q67, K69 and S71) were further examined by fitting their binding isotherms (Figure 3.24) to reveal those that gave a consensus K_d (V2, R15, Q46, Q53, L63 and K69). For the side chain resonances a couple showed significant chemical shift deviation (Q45 and N62) but neither of their binding isotherms matched those of the backbone amide resonances. The average value taken from the binding isotherms calculated the K_d to be 130±80μM.

The intensity of each residues peak was again examined through the course of the titration, after accounting for the changes in concentration and the number of scans. As with the loop 1 titration, those that might have had a significant change could be split into three different classes. Class 1: residues that corresponded to a significant change (V1, V2, I3, S5, N19, L33, Q46, E54, W55, K68, K69, A70 and the side chains of Q45 and Q25). Class 2: residues that might or might not have corresponded to a direct interaction (K14, I16, S30, C32, D50, Q53, N62 and the side chains of Q46, Q53 and Q67). Class 3: these residues had a deviation somewhere in the titration (E18, Q25, K52 and Q67).

As with loop 1, when all the information was collated and applied to the CCL24 structure (Figure 3.25) there did not appear to be a defined binding site, but there were the same three clusters: on the C-terminus, the N-terminus and at the top of the α-helix.

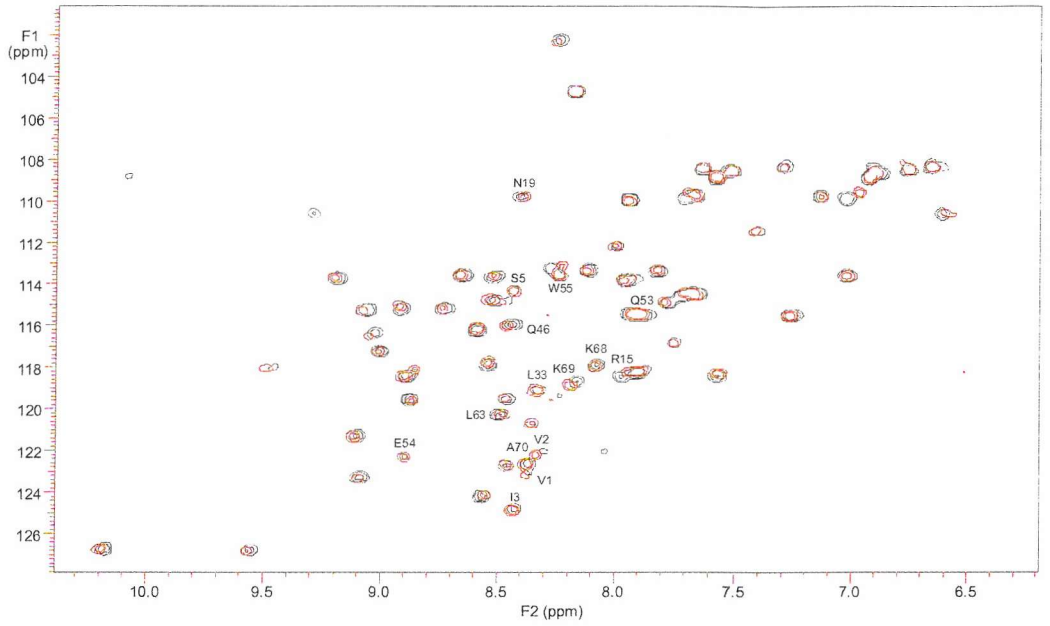


Figure 3.22 HSQC spectra overlay for the titration of CCL24 into loop 3 where an attempt has been made to normalise the vertical scale. The first point is in red (bound form) and the final point is in black (unbound form) showing the region from 6.2 to 10.4ppm and 100.6 to 128ppm (V56 and I16 are off scale). The marked residues are those that underwent either a concentration dependent chemical shift to give a consensus K_d, or a concentration dependent intensity change. Both spectra were zero filled from 1024 to 2048 points in F2 and were zero filled from 64 to 512 points in F1. The Varian window functions were SB (-0.03), SBS (-0.021), SB1(-0.028) and SBS1(-0.017).

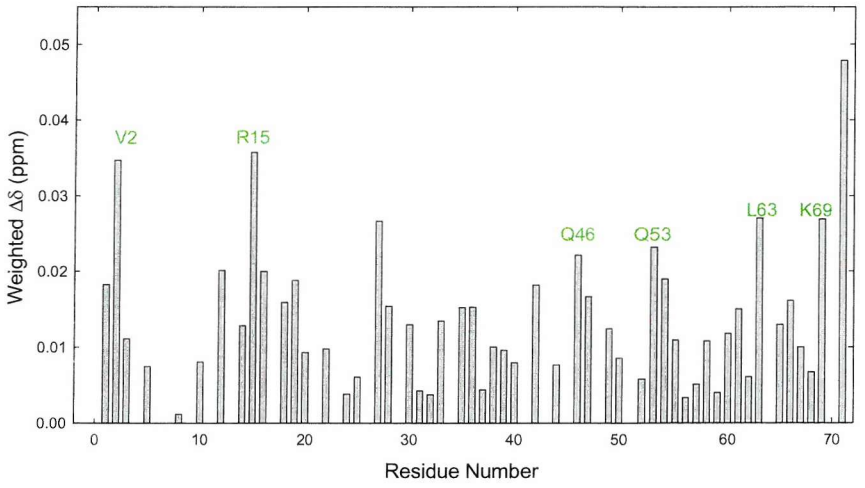


Figure 3.23 The weighted chemical shift difference between the first and last points for the backbone resonances of CCL24 from its titration into loop 3. The residues marked are those that showed concentration dependent chemical shift change and whose binding isotherms gave a consensus K_d (130±80μM). The residues were V2, R15, Q46, Q53, L63 and K69.

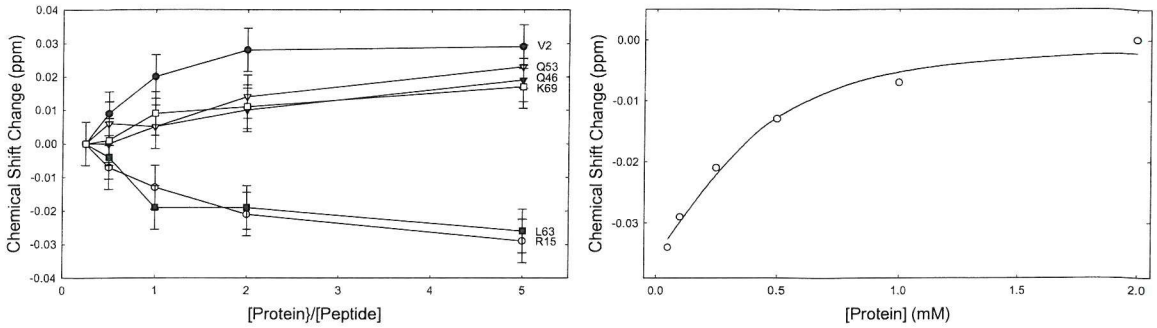
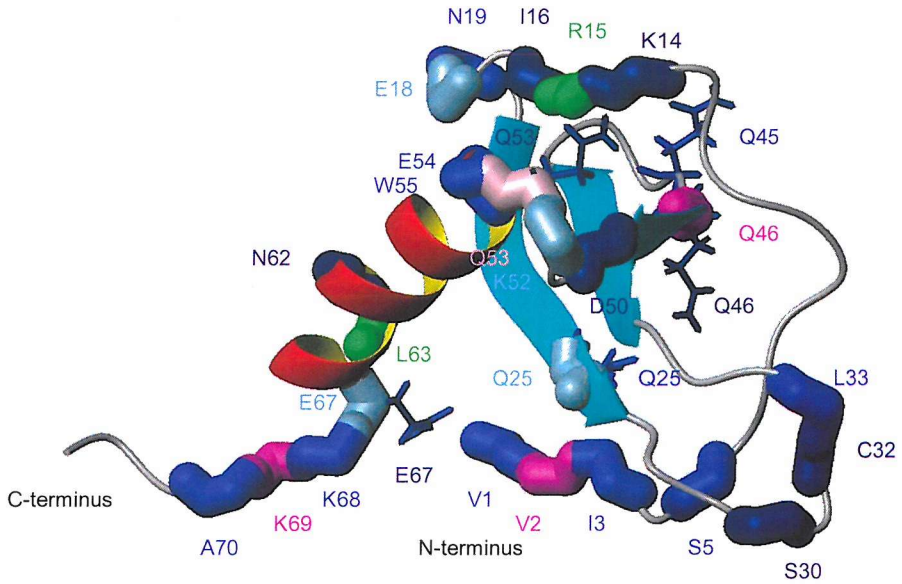


Figure 3.24 A. Binding isotherms obtained upon titration of CCL24 with Loop 3. The chemical shift change is plotted as a function of peptide/protein ratio for the amide protons of V2 (filled circles), R15 (open circles), Q46 (filled triangles), Q53 (open triangles), L63 (filled squares) and K69 (open squares). Error bars correspond to the ^{1}H digital resolution of the HSQC spectra. B. The measured chemical shift deviation of R15 (open circles) with the calculated chemical shift deviation shown as a line.



Residues colour coded to show those involved in:

Bright blue: Class 1 intensity change (V1,I3,S5, N19, L33, E54, W55, K68, A70 and the side chains of Q25 and Q45).

Dark blue: Class 2 intensity change (K14, I16, S30, C32, D50, N62 and the side chains of Q46, Q53 and E67).

Light blue: Class 3 intensity change (E18, Q25, K52 and E67).

Bright green: Chemical shift deviation (R15 and L63).

Light pink: Both a weak binding chemical shift deviation and a Class 2 or 3 intensity change (Q53).

Purple: Both a strong binding chemical shift deviation and a Class 1 intensity change (V2, Q46 and K69).

Figure 3.25 CCL24 with the residues affected by the loop 3 titration colour coded. There are three major clusters: one cluster is on the C-terminus, the second is on the N-terminus, whilst the third is at the top of the α-helix.

The ^{15}N-filtered NOESY recorded at 2mM CCL24 showed significant break through from the protein, so the spectrum recorded at 1mM CCL24 was used instead. The overlay of the

spectra for the peptide (Figure 3.28) showed that the end point spectrum had cross-peaks to the amide hydrogen for residues S5, S6, N14, D15, S19 and K20 disappearing from the spectrum (S6, S19 and K20), or undergoing changes in intensity. Of the peaks that were observed those corresponding to L4, A16, R18, H21, L22, D23, L24, V25 and M26 had a significant change in chemical shift whilst there was no significant observed shift in the peaks corresponding to L3, Y7, Q8, S9, I10, L11, F12, G13 and E17 (Figure 3.26).

When the information was collated onto the structure of loop 3 (Figure 3.27) it would seem that the site affected consisted largely of the C-terminus with a few neighboring residues in the N-terminus also affected. Surprisingly, the top of the loop was unaffected.

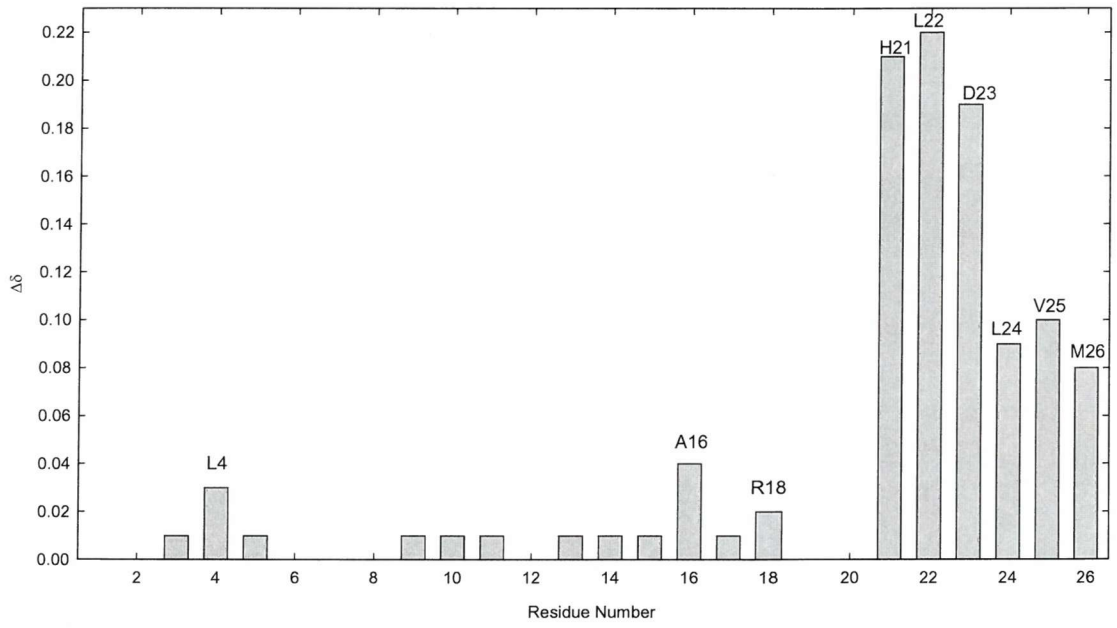
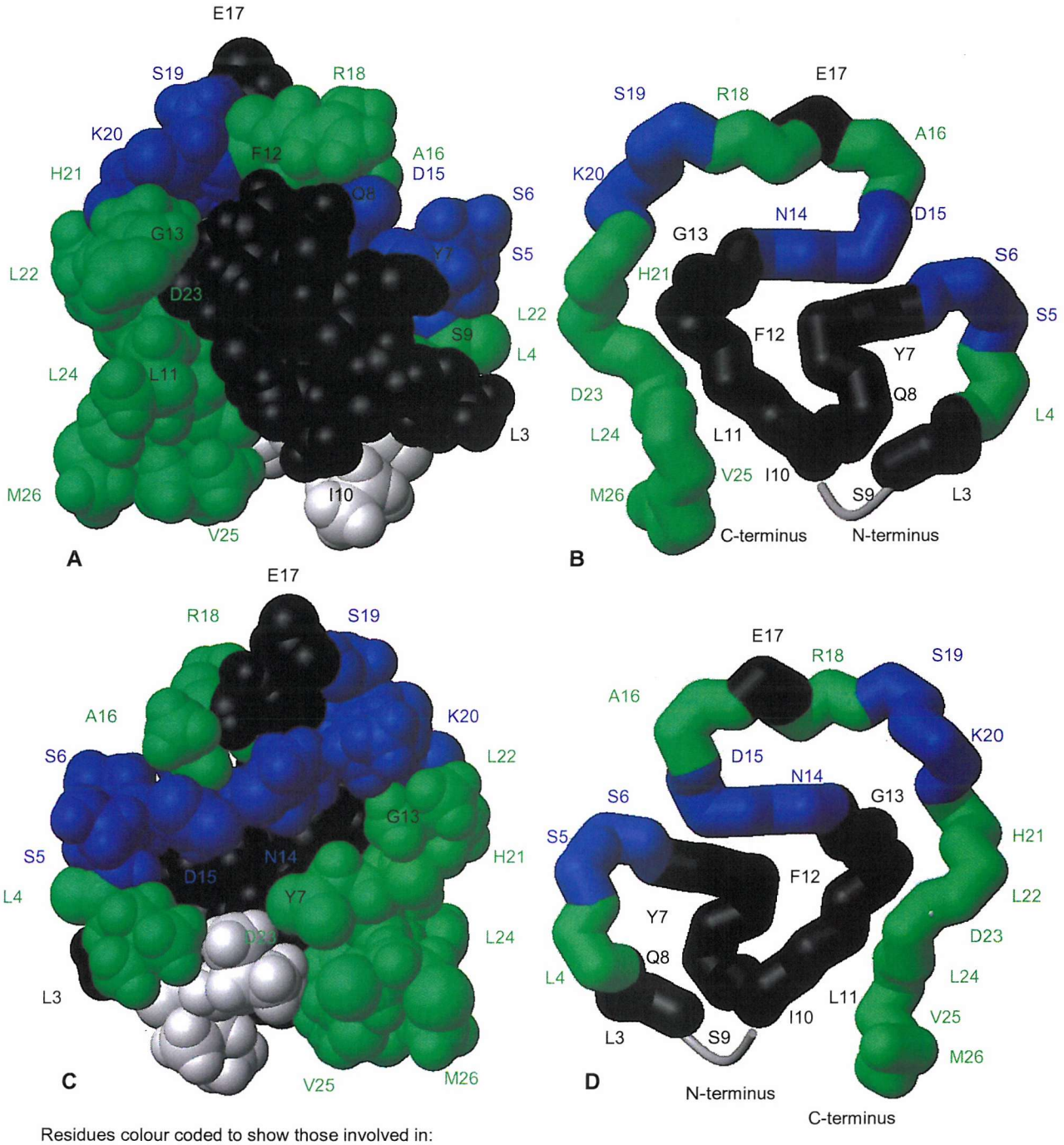


Figure 3.26 The chemical shift deviation of the amide hydrogen between a standard of the loop 3 peptide and the ^{15}N-filtered NOESY run at 1mM CCL24. The peaks undergoing significant chemical shift changes are L4, A16, R18, H21, L22, D23, L24, V25 and M26. There was no amide cross-peak observed for S6, S19 and K20.



Residues colour coded to show those involved in:

Black: Unaffected by the titration (L3, Y7, Q8, S9, I10, L11, F12, G13 and E17)

Blue: Intensity change (S5, S6, N14, D15, S19 and K20).

Green: Chemical shift deviation (L4, A16, R18, H21, L22, D23, L24, V25 and M26).

Figure 3.27 The structure of the peptide mimic of loop 3 with those residues that were affected by the titration colour coded. It would appear that the residues affected consisted largely of the C-terminus with a few neighboring residues in the N-terminus also affected. A. Space filled model of loop 3 showing the side that is probably facing inwards towards CCL24. B. Same as Figure A except loop 3 is now in ribbon. C. Same as Figure A except rotated by 180°. D. Same as Figure C except loop 3 is now in ribbon.

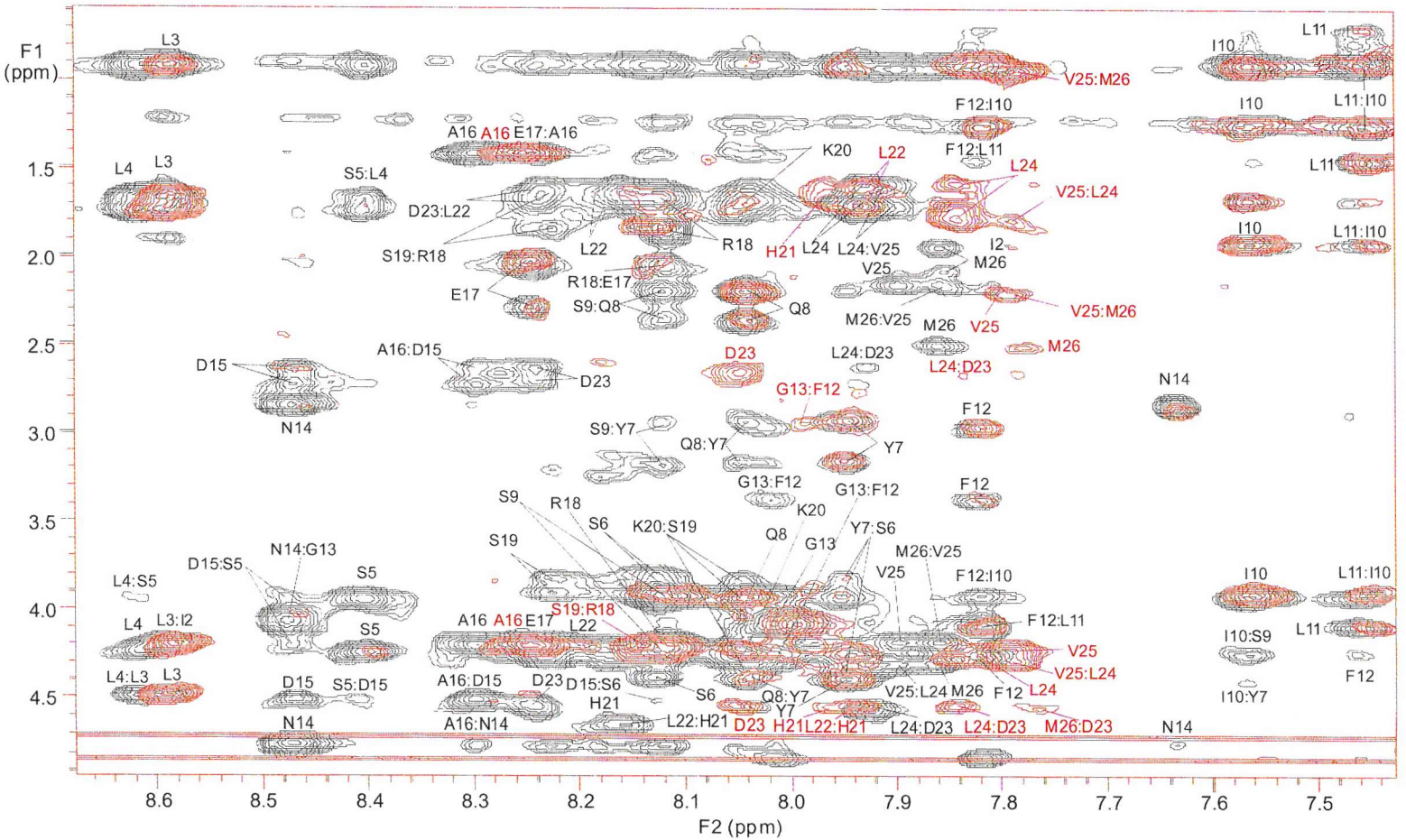
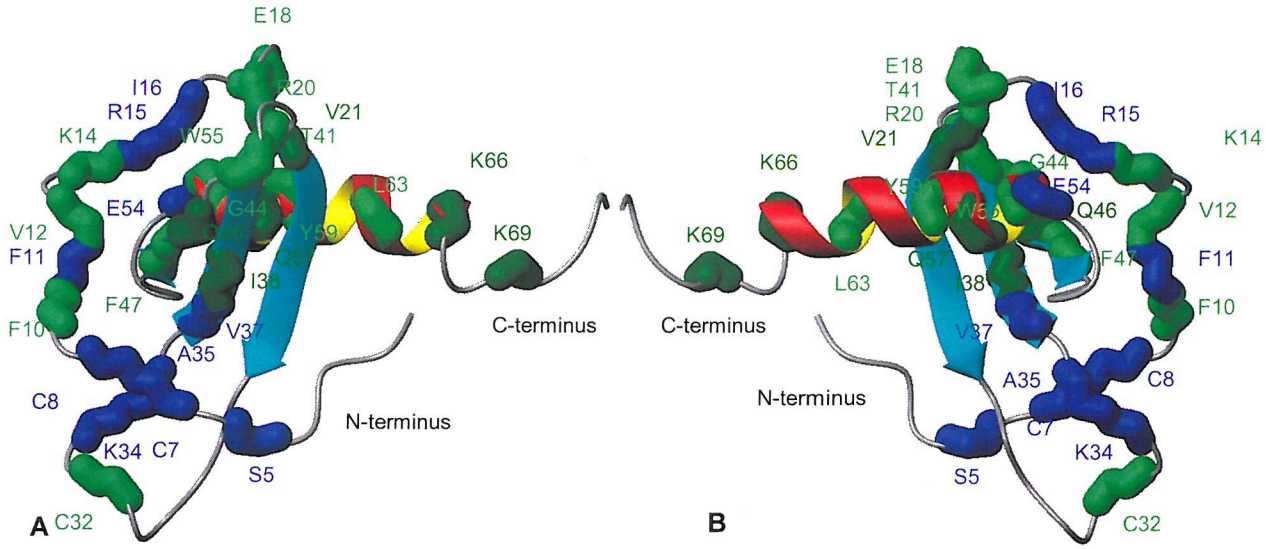


Figure 3.28 NOESY spectra overlay for loop 3 showing the region from 6.7 to 9.45ppm and 0.85 to 4.9ppm. The spectrum taken before the titration is in black (unbound form) and after the addition of CCL24 is in red (bound form). The unbound spectrum is fully labelled in black, whereas in the bound spectrum only those peaks that are significantly perturbed from their original positions are labelled in red. The unbound spectrum was zero filled from 4096 to 8192 points in F2 and was zero filled from 256 to 2048 points in F1. The Varian window functions were LB (-6.1), GF (0.038), LB1(-10.3) and GF1(0.012). The bound spectrum was zero filled from 4096 to 8192 points in F2 and was zero filled from 256 to 2048 points in F1. The Varian window functions were LB (-2.25), GF (0.062), LB1(0) and GF1(0.012).

3.5 Previous NMR Studies on CCR3 and CCL24

A NMR study of CCL24 with a peptide corresponding to the first thirty-five residues of the N-terminus of CCR3 established a dissociation constant of 45μM (90% confidence range 3-148μM) for the interaction [21]. Eight amide resonances of CCL24 (S5, C7, C8, F11, I16, K34, A35 and V37) underwent noticeable changes in lineshape and two peaks (R15 and E54) disappeared from the spectrum, possibly due to extreme broadening. The chemical shift deviations of CCL24 were placed into two categories: those that had large chemical shift deviations (F10, V12, K14, E18, R20, C32, T41, G44, F47, W55, Y59 and L63) and those that had small chemical shift deviations (V21, I38, Q46, Q57, K66 and K69). When these were displayed upon CCL24 (Figure 3.29) they appeared to be spread all over the protein [21]. No intermolecular data was collected, so no experimentally determined models are available, but the authors interpreted their results with the aid of previously determined binding sites [111, 166-168] to suggest that the peptide ran in a groove from C7 and C8 up to R15 and I16.



Residues colour coded to show those involved in:

Bright blue: Intensity change (S5, C7, C8, F11, R15, I16, K34, A35, V37 and E54).

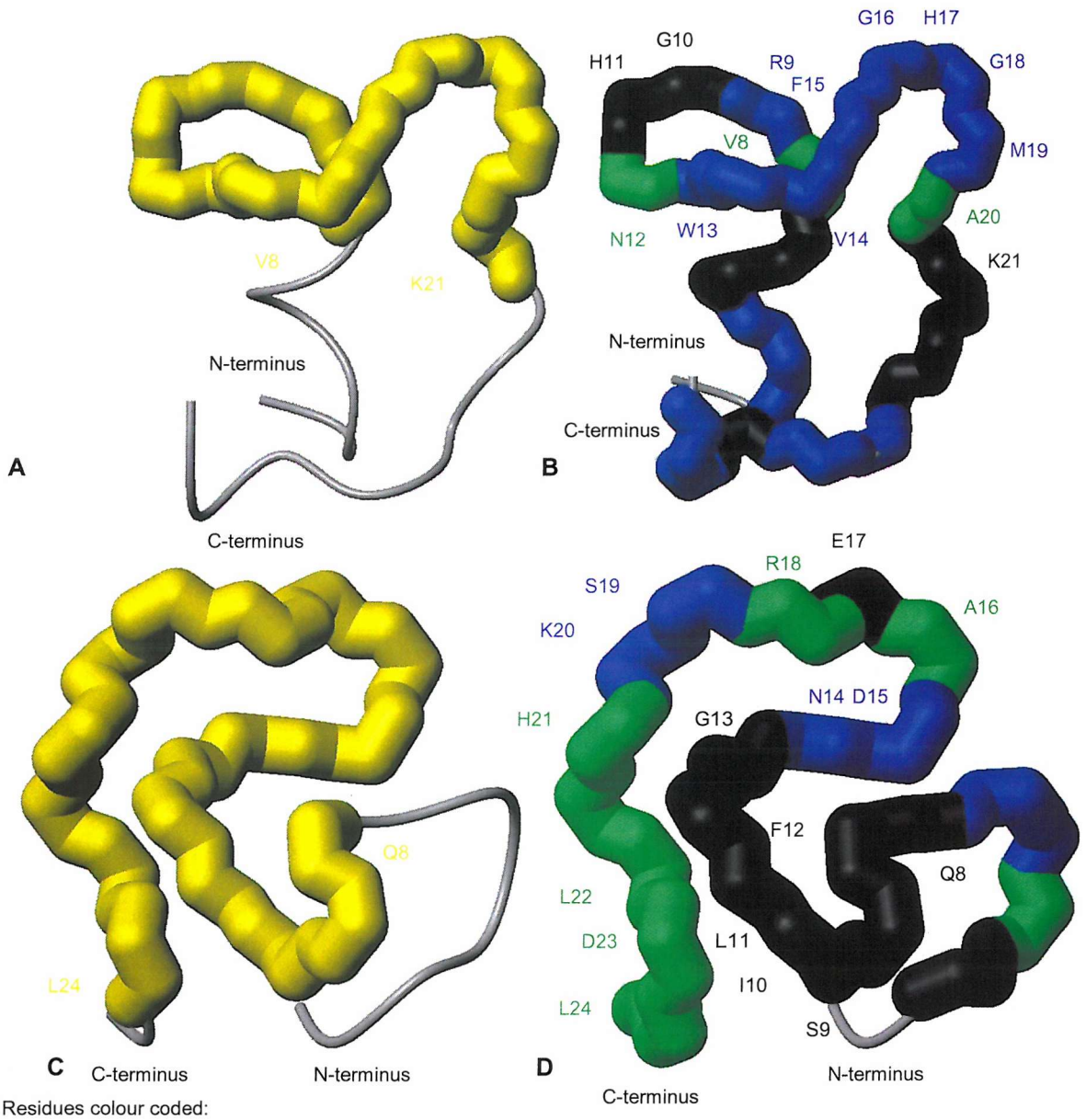
Bright green: Large chemical shift deviation (F10, V12, K14, E18, R20, C32, T41, G44, F47, W55, Y59 and L63).

Dark Green: Small chemical shift deviation (V21, I38, Q46, Q57, K66 and K69)

Figure 3.29 CCL24 with those residues that showed changes in their NMR spectra when a peptide corresponding to the first thirty-five residues of CCR3 was titrated in. The residues are colour coded and spread over most of the protein. A. Ribbon model of CCL24. B Same as Figure A except rotated by 180°.

There is a previously published attempt to mimic the extracellular loops 1 and 3 of CCR3 and study their interaction with a chemokine (CCL11), but this failed to show any interaction [111]. Figure 3.30 compares the loops used in the failed study by mapping their sequence onto the structures of the ones successfully used here. Assuming that CCL11 undergoes similar interactions to CCL24 then the previous loop 1 mimic had all the residues that are believed to be showing a real interaction in this case. This means that either the N-terminus and C-terminus of the loop 1 mimic are involved in the binding, or that the failed peptide did not adopt the correct fold, or that the observers missed the small deviations that were observed with this loop. The failed loop 3 mimic was missing part of the interaction site which could be the cause of it failing; it is also highly probable that the failed loop 3 mimic was not adopting the correct structure, as the interaction S5 to D15 is one of the important restraints [111].

A successful mimic of CCR3 designed with both the N-terminus and loop 3 (Q8 to L24) attached to another protein gave a K_d of 135±48μM for the interaction of CCL11 to loop 3 alone, compared with the K_d of 130±80μM for the interaction of CCL24 observed here [62]. This mimic is missing some of the residues that are affected in the current titration, which suggests either CCL11 and CCL24 are not binding in exactly the same way, or that residues are experiencing secondary effects, or they are not contributing significantly to the overall interaction site.



Residues colour coded:

Black: Unaffected by the titration.

Bright blue: Intensity change.

Bright green: Chemical shift deviation.

Yellow: Failed peptides

Figure 3.30 The failed peptide mimic sequences mapped to the successful peptides structures. A. Failed loop 1 mimic which corresponds to residues V8 to K21 of the successful mimic. B. Successful loop 1 mimic with the residues that showed deviation upon titration. C. Failed loop 3 mimic which corresponds to residues Q8 to L24 of the successful mimic. This sequence was also used in a dual mimic with the N-terminus. D. Successful loop 3 mimic with the residues that showed deviation upon titration.

3.6 Binding Orientation of CCR3 to CCL24

When the information from the titrations performed with loops 1 and 3 was combined with published information for the N-terminus (Figure 3.31) there was unfortunately no obvious binding site that places all the residues involved in contact with the receptor. The most probable cause for this is that some of the residues are showing secondary effects, but it is impossible to tell from this data alone which, if any, are acting in this fashion.

It is normally the case that the N-terminus and C-terminus of a chemokine are flexible, but in contrast to this the published NMR solution structure of CCL24 has a flexible C-terminus whilst the N-terminus has long range NOEs that fix its position in relation to the main body of the protein. Figure 3.31 shows the two extremes for the N-terminus and C-terminus as taken from the published ensemble for CCL24. If this information is used along with the theory that the N-terminus of the chemokine activates the receptor by contacting its loop 2, then the range of possible orientations between CCL24 and CCR3 is much more limited.

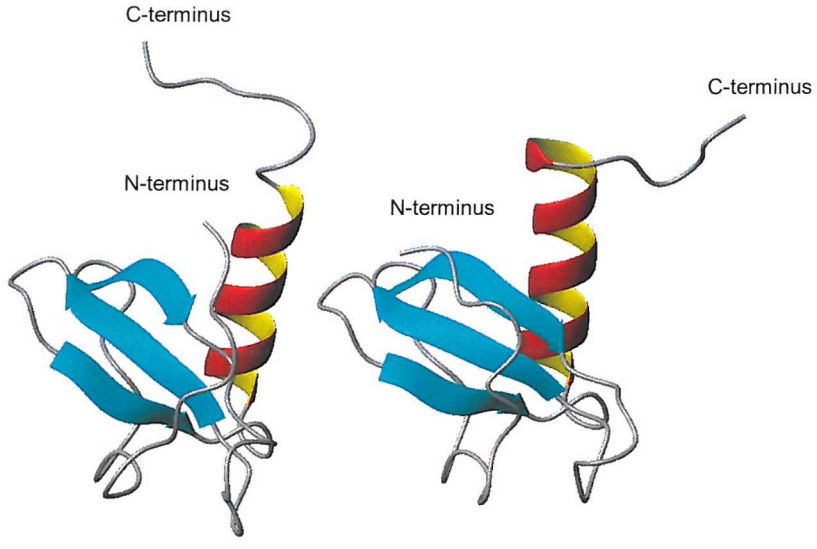
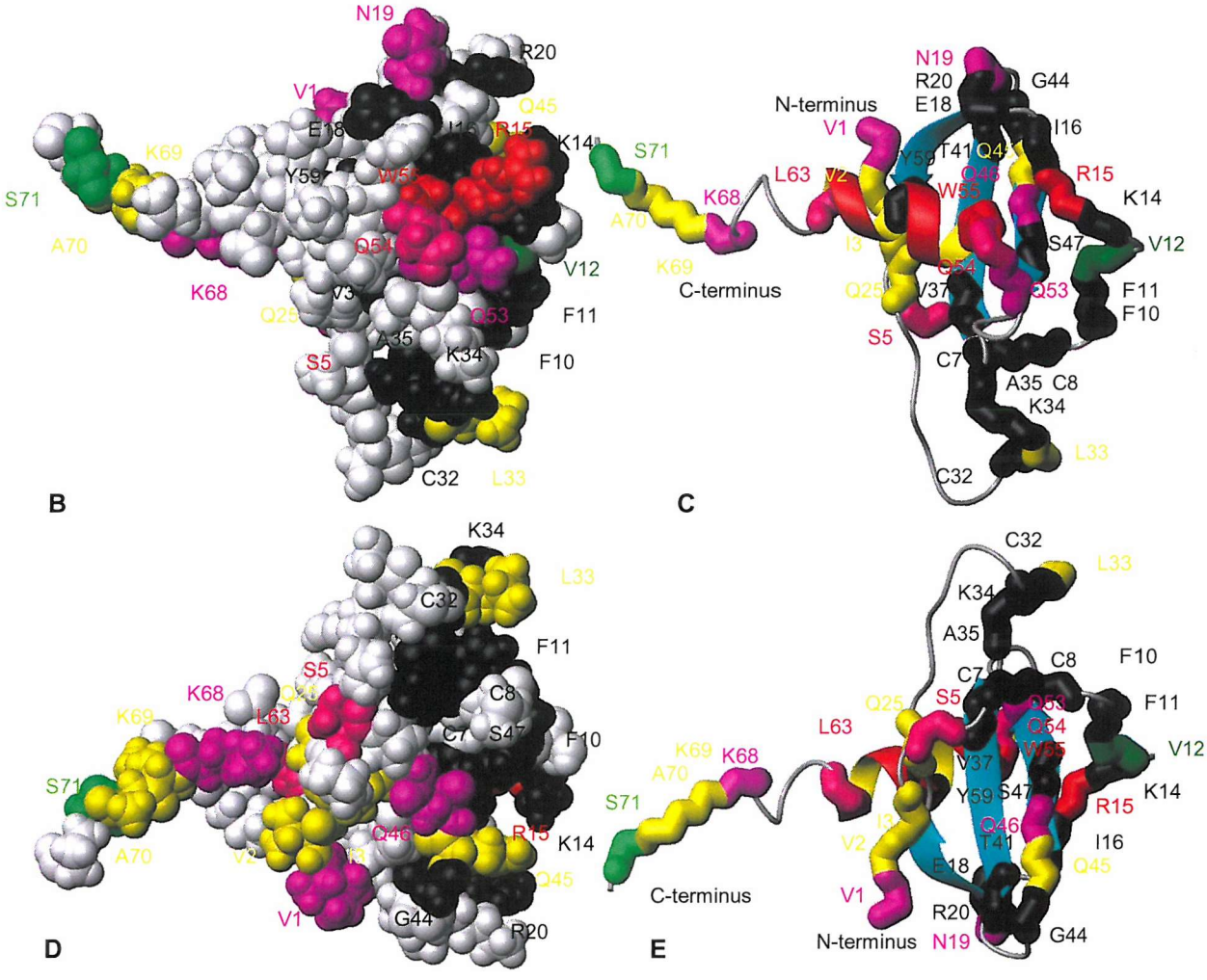


Figure 3.31 The two extremes of the N-terminus and C-terminus of CCL24 as taken from the published ensemble.

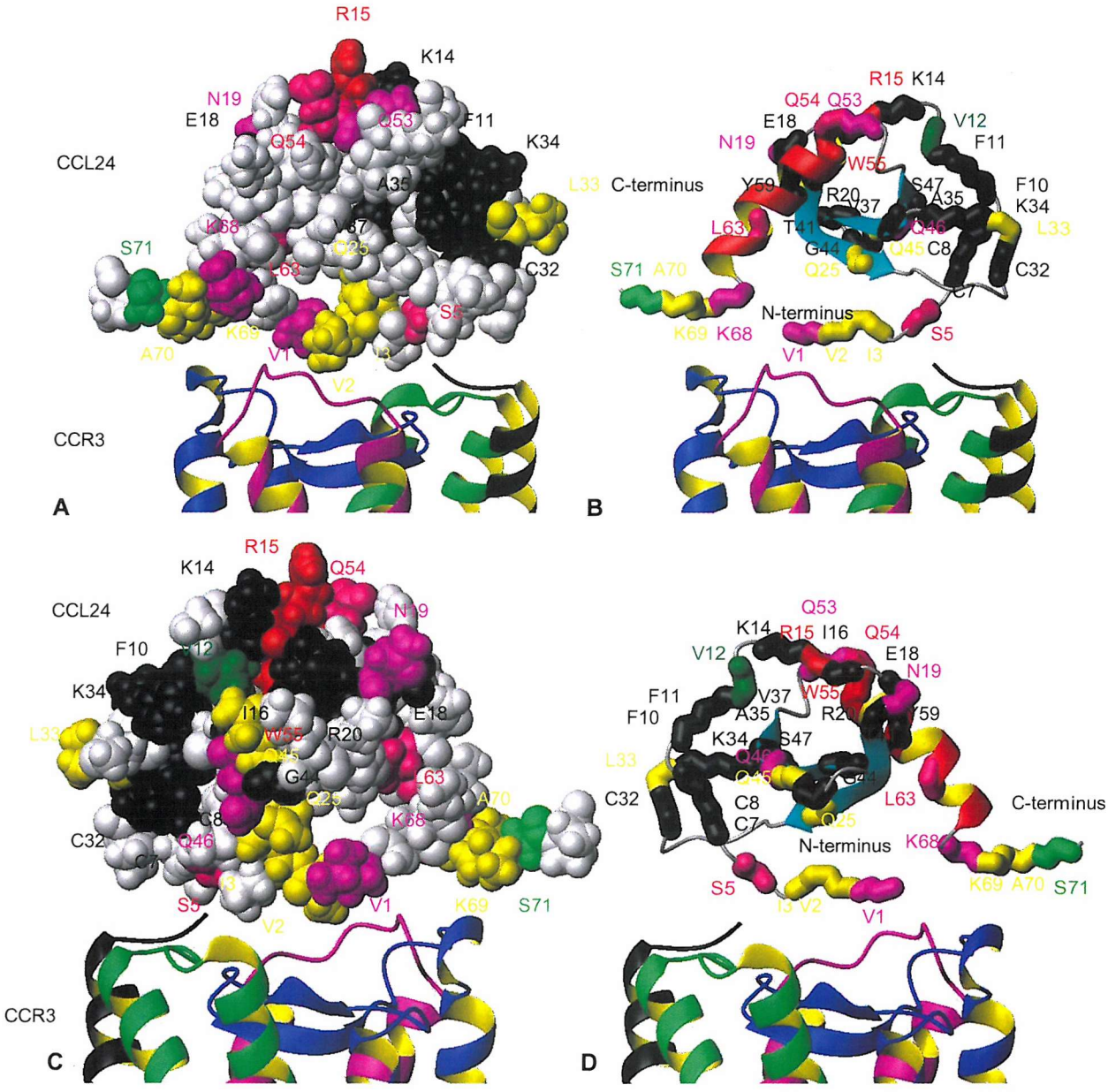
Loop 1 Titration Class 1 Intensity Changes	V2, I3, L33, W55, K69, A70
Loop 1 Titration Chemical Shifts	V2, V12, R15, Q25, K69, S71 and the side chain of Q45
Loop 3 Titration Class 1 Intensity Changes	V1, V2, I3, S5, N19, L33, Q46, Q54, W55, K68, K69, A70 and the side chains of Q45 and Q25
Loop 3 Titration Chemical Shifts	V2, R15, Q46, Q53, L63, K69
N-Terminus Titration Intensity Changes	S5, C7, C8, F11, R15, I16, K34, A35, V37, Q54
N-Terminus Titration Strong Chemical Shifts	F10, V12, K14, E18, R20, C32, T41, G44, S47, W55, Y59, L63

A



CCL24 residues colour coded to show those involved in binding:
Black: N-terminus only (C7, C8, F10, F11, K14, I16, E18, R20, C32, K34, A35, V37, T41, G44, S47 and Y59).
Green: Loop 1 only (S71).
Dark green: Both the N-terminus and loop 1 (V12).
Dark pink: Both the N-terminus and loop 3 (S5, Q54 and L63).
Purple: Loop 3 only (V1, N19, Q46, Q53 and K68).
Red: Thrice N-terminus, loop 1 and loop 3 (R15 and W55)
Yellow: Both loop 1 and loop 3 (V2, I3, Q25, L33, Q45, K69 and A70).

Figure 3.32 Residues believed to be significantly perturbed upon the three titrations. A. Table of the residues. B. Space filled model of CCL24 with all residues believed to be involved colour coded. C. Same as Figure B except CCL24 is now in ribbon. D. Same as Figure B except rotated by 180°. E. Same as Figure D except CCL24 is now in ribbon.

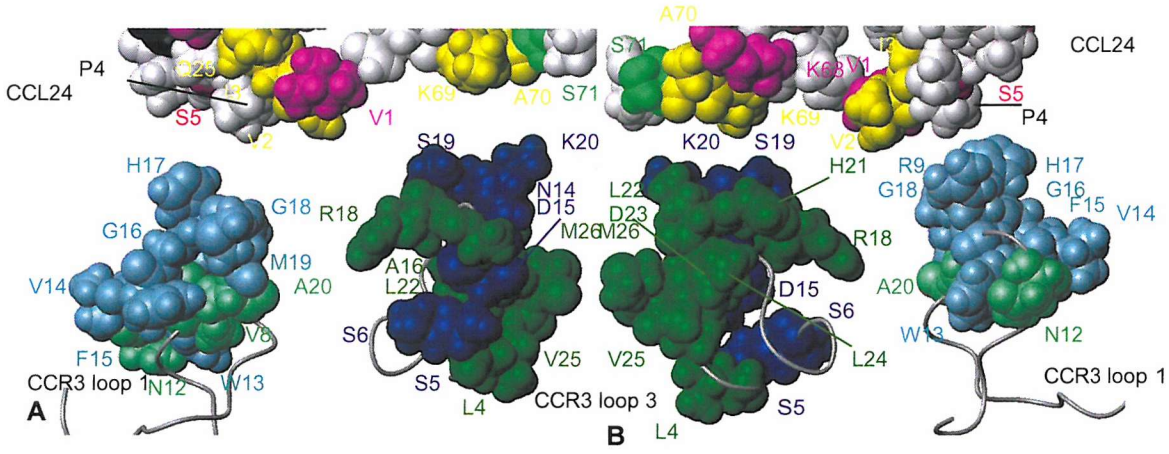


CCR3 colour coded:
Black N-terminus and helix 1.
Blue: Loop 2 and helices 4 and 5.
Green: Loop 1 and helices 2 and 3.
Purple: Loop 3 and helices 6 and 7

CCL24 residues colour coded to show those involved in binding:
Black: N-terminus only (C7, C8, F10, F11, K14, I16, E18, R20, C32, K34, A35, V37, T41, G44, S47 and Y59).
Green: Loop 1 only (S71).
Dark green: Both the N-terminus and loop 1 (V12).
Dark pink: Both the N-terminus and loop 3 (S5, Q54 and L63).
Purple: Loop 3 only (V1, N19, Q46, Q53 and K68).
Red: Thrice N-terminus, loop 1 and loop 3 (R15 and W55)
Yellow: Both loop 1 and loop 3 (V2, V3, Q25, L33, Q45, K69 and A70).

Figure 3.33 A proposed orientation of CCL24 for binding to CCR3 based upon the published ensemble. A. Space filled model of CCL24 with the top of CCR3 in ribbon. The N-terminus of CCR3 has been truncated for clarity, but is assumed to be flexible and pointing outwards away from the membrane. B. Same as Figure A except CCL24 is now in ribbon. C. Same as Figure A except rotated by 180°. D. Same as Figure C except CCL24 is now in ribbon.

If the published ensemble of CCL24 is a true representation of the space that the N-terminus can sample and if the N-terminus of CCL24 has to contact loop 2 on CCR3 and the model of CCR3 shows the correct position of loop 2 then Figure 3.33 shows the only major orientation that accounts for the data. In this conformation the N-terminus and C-terminus of CCL24 would be the interacting with both loops whilst the behaviour of those residues at the top of the α-helix would be a secondary effect probably caused by the binding of the C-terminus, which is at the other end of the α-helix. If the peptide structures are aligned with the helices on the model of CCR3 then a representation like Figure 3.34 could indicate the residues that might be interacting.



CCR3 loop 1 colour coded to show residues undergoing:
Light blue: Intensity change upon binding (R9, W13, V14, F15, G16, H17, G18 and M19).
Light green: Chemical shift change upon binding (V8, N12 and A20).

CCR3 loop 3 colour coded to show residues undergoing:
Dark blue: Intensity change upon binding (S5, S6, N14, D15, S19 and K20).
Dark green: Chemical shift change upon binding (L4, A16, R18, H21, L22, D23, L24, V25 and M26).

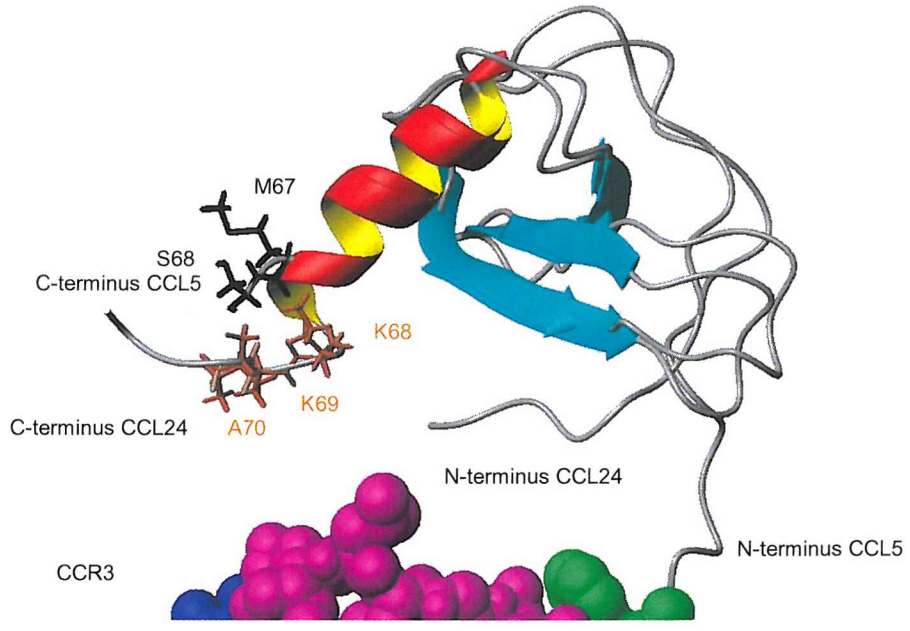
CCL24 residues colour coded to show those involved in binding:
Green: Loop 1 only (S71).
Dark pink: Both the N-terminus and loop 3 (S5).
Purple: Loop 3 only (V1 and K68).
Yellow: Both loop 1 and loop 3 (V2, I3, Q25 and K69).

Figure 3.34 The observed peptide loop mimic structures in an expanded model of the orientation of CCL24 for its binding to CCR3. A. Space filled model of CCL24 with the residues on CCR3 that are believed to be involved in binding in space fill. B. Same as Figure A except rotated by 180°.

There is only one study specific to the binding of CCL24 to CCR3 which was performed using various truncations of the N-terminus of CCL24 [144]. This study showed that the CCL24 mutant P4 to R73 bound with almost the same affinity as the wild type, but did not give rise to a signal. In contrast, the CCL24 mutant P6 to R73 was a much weaker binder and still did not signal. The proposed orientation of CCL24 to CCR3 could explain this result if P4 binds to the loop 1 whilst S5 binds to loop 3 and the N-terminus. The binding of P4 to the loop 1 would also help explain why the loop 1 mimic is binding CCL24 more

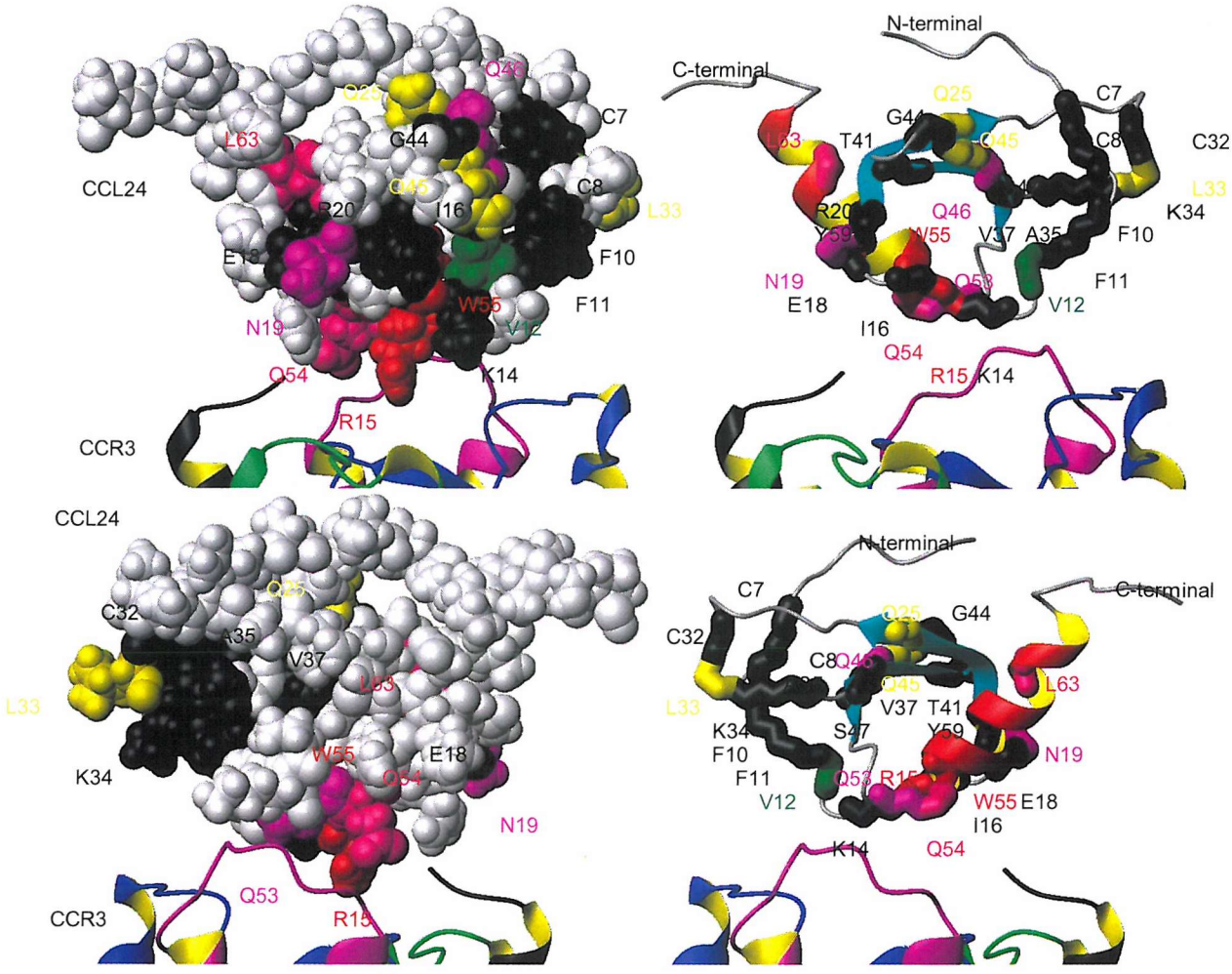
strongly than the loop 3 mimic, whilst not causing as many perturbations of the HSQC. Prolines do not appear in the HSQC, but they could be observed if a ^{13}C labelled sample of CCL24 was produced.

The model suggests that the C-terminus of CCL24 is binding to CCR3. Previous work on identification of the binding site for CCL5 to CCR5 had suggested this possibility, with a dodecamer peptide corresponding to the last twelve residues of CCL5 binding to CCR5, whilst a dodecamer peptide shifted by two amino acids (towards the N-terminus) did not. When the solved structure of CCL5 is superimposed upon the solved structure of CCL24 using the beta sheet (RMSD 1.2Å), it oriented the C-terminal ends of the alpha helices to within close proximity of each other. The residues A70 and K68 of CCL24 were placed about 3 and 5Å away from S68 on CCL5 respectively. This shows that these residues might be involved in the formation of a very similar binding site.



Black: Residues upon CCL5 (M67 and S68).
Blue: loop 2.
Brown: Residues upon CCL24 (K68, K68 and A70).
Green: loop 1.
Purple: loop 3.

Figure 3.35 Overlay of the mean averaged structures of CCL5 (I24-Y29, V39-T43 and Q48-C50) and CCL24 (V21-L26, I37-T41 and F47-G49) showing the C-termini to be about 5Å from each other. The residues in the C-terminus possibly involved in binding CCL5 with CCR5 are shown in black and the residue in the C-terminus of CCL24 believed to be involved in binding CCR3 is shown in brown.



CCR3 colour coded:
Black N-terminus and helix 1.
Blue: Loop 2 and helices 4 and 5.
Green: Loop 1 and helices 2 and 3.
Purple: Loop 3 and helices 6 and 7

CCL24 residues colour coded to show those involved in binding:
Black: N-terminus only (C7, C8, F10, F11, K14, I16, E18, R20, C32, K34, A35, V37, T41, G44, S47 and Y59).
Dark green: Both the N-terminus and loop 1 (V12).
Dark pink: Both the N-terminus and loop 3 (S5, Q54 and L63).
Purple: Loop 3 only (V1, N19, Q46 and Q53).
Red: Thrice N-terminus, loop 1 and loop 3 (R15 and W55)
Yellow: Both loop 1 and loop 3 (V2, V3, Q25, L33 and Q45).

Figure 3.36 A proposed orientation of CCL24 for binding to CCR3 using the residues on the less flexible regions of CCL24. A. Space filled model of CCL24 with the top of CCR3 in ribbon. The N-terminus of CCR3 has been truncated for clarity, but is assumed to be flexible and pointing outwards away from the membrane. B. Same as Figure A except CCL24 is now in ribbon. C. Same as Figure A except rotated by 180°. D. Same as Figure C except CCL24 is now in ribbon.

If the published ensemble is incorrect and the N-terminus and C-terminus are flexible, then another possible orientation for the binding is shown in Figure 3.36 where the top end of the alpha helix is orientated towards the membrane. This would require the N-terminus and C-terminus to be flexible enough to account for their interactions, or for these to be secondary effects.

3.7 Overall Conclusion

The initial results obtained here are encouraging, but further work is required. Loop 3 should be extended by at least three residues on both the N and C-termini as the structure is very compact with both termini close to the loop itself. The extra residues might extend the termini into the lipid. In addition the C-terminus underwent a chemical shift deviation during the titration and such an extension might result in more conclusive information.

Both loops require further work to enable characterisation of which residues are buried in the membrane. This could potentially be done using saturation transfer from the deuterated micelle or by the addition of paramagnetic ions such as manganese into the solution.

The interaction between CCL24 and the peptides could be further studied by NMR spectroscopy if ^{13}C labelled samples were prepared. This would allow the observation of side chain-side chain interactions using ^{13}C –filtered experiments. The system could also be expanded to the other chemokines that bind CCR3 and also to the viral mimics (e.g. vMIPII).

With the system showing promise it is possible that other techniques could be used, for example isothermal titration calorimetry (ITC) and fluorescence. A lack of dependence upon deuterated detergents, of which only a few are available commercially, could enable the study of the effects that the lipid has upon the interaction.

Overall this work illustrates that peptides corresponding to the extracellular loops of membrane proteins when suspended in membrane mimics are suitable for the study of the interaction between membrane proteins and their ligands by NMR spectroscopy. The work is still at a very early stage on this system and its more general applicability is unknown. It should be possible to use the alignments and discussion in chapter 3 to design similar peptides for all the chemokine receptors and hence study their interactions with the chemokines. The work could also theoretically be expanded upon to make peptide mimics of other G-coupled proteins. If a similar system could be established for loop 2 then the work is probably only limited by time, money, the production of pure peptides and labelled proteins.

4 Materials and Methods

4.1 Materials

The majority of chemicals were obtained either from Sigma-Aldrich Chemical Company Ltd, Dorset, or Fisher Scientific UK Ltd, Bishop Meadow Road, Loughborough, Leicestershire. More specialised reagents and kits were obtained from the following sources:

Southampton Polypeptides, University of Southampton, Boldrewood, Bassett Crescent East, Southampton, SO16 7PX.
Peptides.

Amersham Pharmacia Biotech, Pollards Wood, Nightingales Lane, Chalfont, St. Giles, Bucks, HP8 4SP.
8-25% SDS gradient phast gels.
SP Sepharose high load 26/10.
Q Sepharose high load 26/10.
Superdex 75 high load 16/60.
Sephadex G-25 M PD10 Cloumns.

Amicon Division, WR Grace & Co., Beverley, Mass., USA.
50mL Amicon concentrator.

Applied Biosystems, Lingley House, 120 Birchwood Boulevard, Warrington, WA3 7QH.
Amplitaq DNA polymerase.
PCR buffer.

Bio-Rad Laboratories Ltd., Bio-Rad House, Maylands Avenue, Hemel Hempstead, Hertfordshire, HP2 7TD.
Bio-Rad protein assay.

Cambridge Isotopes, Goss Scientific Instruments Ltd, 100 Vicarage Lane, Great Baddow, Essex, CM2 8JB.
^{15}N-Ammonium chloride.

Clontech, 1020 East Meadow Circle, Palo Alto, CA.
TALON metal affinity resin.

Difco Laboratories LTD, BD Biosciences, 21 Between Towns Road, Cowley, Oxford, OX4 3LY.
Bacto typtone.
Yeast.

Goss Scientific Instruments LTD, 100 Vicarage Lane, Essex, CM2 8JB.
DPC-d38

Invitrogen Ltd, 3 Fountain Drive, Inchinnan Business Park, Paisley.
TOPO TA cloning kit.
SDS-PAGE gel markers.

Melford Laboratories Ltd, Bildeston Road, Chelsworth, Ipswich, Suffolk.
Isopropyl-β-D-thiogalactosidase (IPTG).
Ampicillin sodium.

Millipore (U.K.) Ltd., Units 3&5 The Courtyards, Hatters Lane, Watford, WD18 8YH.
Ultrafiltration membranes, regenerated cellulose.
0.45μm Cellulose nitrate membrane filters.

Novagen, Merck Biosciences Ltd, Boulevard Industrial Park, Padge Road, Beeston, Nottingham, NG9 2JR.
BL21(DE3).
BL21(DE3)pLysS.
AD494 (DE3).
Rosetta (DE3).
pET-15b vector.

Qiagen Ltd, Boundary Court, Gatwick Road, Crawley, West Sussex.
QIAquick gel extraction kit.
QIAquick PCR purification kit.
QIAprep spin miniprep kit.
QIAfilter plasmid midi kit.
Ni-NTA superflow resin.

Promega, Delta House, Chilworth Science Park, Southampton, SO16 7NS.
DNA Ladders.
DNA Loading Dye.

Roche Products Ltd., 40 Broadwater Road, P.O. Box 8, Welwyn Garden City, Hertfordshire, AL7 3AY.
Restriction enzymes.
Calf intestinal alkaline phosphatase.

Stratagene 1834 State Highway 71 West, Cedar Creek, USA.
XL1 Blues.

4.2 Microbiological Techniques

4.2.1 Sterilization

All pipette tips, media, stock solutions, microfuge tubes and cryotubes were sterilised in a British Steriliser autoclave at 120°C, pressurized to 1 bar for twenty minutes. The tips and tubes were then dried in a 65°C oven for at least twenty four hours. Stocks of heat labile solutions such as ampicillin and isopropyl-β-D-thiogalactosidase (IPTG) were sterilised by filtration through disposable 0.2μM Millipore filters.

4.2.2 Bacterial Strains

Strain	Genotype	Origin	Reference
BL21(DE3)	F⁻ *ompT hsdS*$_B$ (r_B^- m_B^-) *gal dcm* (DE3)	Novagen	[169]
BL21(DE3)pLysS	F⁻ *ompT hsdS*$_B$ (r_B^- m_B^-) *gal dcm* (DE3) pLysS (Cm^R)	Novagen	[169]
AD494 (DE3)	*ara-leu7697 ΔlacX74 ΔphoAPvull phoR ΔphoR ΔmalF3* F'[*lac*$^+$(*lacI*^{q})*pro*] *trxB::kan* (DE3)	Novagen	[170]
Rosetta (DE3)	F⁻ *ompT hsdS*$_B$ (r_B^- m_B^-) *gal dcm* (DE3) pRARE (Cam^R)	Novagen	
XL1 Blues	1 *rec*A1 *end*A1 *gyr*A96 *thi*-1 *hsd*R17 *sup*E44 *rel*A1 *lac* [F′ *pro*AB *lacI*^{q}ZΔM15 Tn10 (Tet^r)]	Stratagene	[171]

4.2.3 Cloning Vectors

Vector	Relevant Features	Origin	Reference
pET-15b	T7*lac* promoter. *N*-terminal 6 His Tag and thrombin cleavage site. Ampicillin resistance.	Novagen	[169, 172, 173]
TOPO TA	Shuttle vector for direct insertion of PCR products. Screened using X-gal	Invitrogen	[174]

4.2.4 Culture Media

4.2.4.1 Luria-Bertani Broth (LB)

Medium Component	Grams / Litre distilled water
Bacto Tryptone	10
Yeast Extract	5
Sodium Chloride	5
pH7	

The chemicals were dissolved in distilled water and then autoclaved. Antibiotics were added where necessary. The antibiotic was added aseptically from stock solutions to the required final concentration of 100μg/mL ampicillin.

4.2.4.2 M9 Minimal Media

Medium Component	Grams / Litre distilled water
di-Sodium hydrogen orthophosphate anhydrous	6
Potassium di-hydrogen orthophosphate	3
Sodium Chloride	0.5
^{15}N-Ammonium Chloride	1
Magnesium Sulfate	0.5
Calcium Chloride di-hydrate	0.015
Glucose	4
pH7.4	

The chemicals, with the exception of glucose, were dissolved in distilled water and then autoclaved. The glucose was dissolved in 10mL of distilled water and added aseptically. The antibiotic was also added aseptically from stock solutions to the required final concentration of 100μg/mL ampicillin.

4.2.4.3 Agar Plates

LB was aliquoted into 300mL bottles, half filling them, agar (2.25g) was then added prior to autoclaving. This mixture was stored for up to three months.

The bottle lid was removed and the agar carefully heated in a microwave (600w) for cycles of 30s, at half power, followed by a shake of the bottle, until all the agar had melted. The solution was then allowed to cool to 40°C before the appropriate antibiotic was added aseptically to the required final concentration and the mixture poured into agar plates to a depth of 0.5cm. Once the plates were set they were inverted and sealed with parafilm before storage at 4°C. Before use each plate was dried off in a warm oven for about 30 minutes to remove condensation.

4.2.5 Strain Storage

Bacterial strains were stored for long term purposes as glycerol stocks. Strains were grown in LB overnight and then 0.5mL added to cryotubes that had been sterilized containing 0.5mL glycerol and stored at -70°C. Purchased competent cells were aliquoted and used as specified in the manufacturers documentation. Bacterial strains containing plasmid were stored in the short term on LB-agar plates containing the appropriate antibiotic at 4°C and destroyed after three weeks.

4.2.6 Preparation of Competent Cells

Where competent cells were not purchased the calcium chloride method [175] was used to render the bacterial cells competent for subsequent transformation. 15μL of host strain in glycerol stock was used to inoculate 20mL of fresh LB media and grown overnight for fourteen hours. This was then used to inoculate 20mL fresh LB media (1:100 dilution). The cells were incubated at 37°C with shaking until they reached log phase (OD_{600} of roughly 0.3). The cells were then harvested at 3000g for fifteen minutes at 4°C and resuspended in 20mL of sterile prechilled 50mM $CaCl_2$. After incubating on ice for one hour the cells were repelleted as before and resuspended in 1500μL of sterile prechilled 50mM $CaCl_2$. The competent cells were kept on ice, either until required or for a maximum of twenty-four hours whichever was sooner.

4.2.7 Transformation of Competent Cells with Plasmid

Purchased cells were used as directed by the manufacturers documentation. For home-made competent cells, plasmid DNA (typically 20-200ng) was added to 200μL of competent cells and incubated on ice for at least thirty minutes. The cells were then heat shocked at 42°C for thirty seconds and returned to the ice for five minutes before 800μL LB was added and the cells incubated at 37°C for one hour. Finally, the cells were plated out onto a LB plate containing the appropriate antibiotic required for selection and incubated at 37°C overnight.

4.2.8 Polymerase Chain Reaction (PCR)

In PCR two oligonucleotides are designed to match the ends of the template (*e.g.* cDNA) and can be used to add on extra bases. There should be approximately an 18 base pair matching overlap between the oligonucleotide and the template if only a few extra bases are being added *via* the oligonucleotide (*e.g.* a restriction site). This matching overlap should be increased as the number of extra bases increases. PCR works by using a heat stable polymerase. A standard reaction set is listed below:

For a 50μL reaction	μL
10 * PCR Buffer	5
10 * 2.5 mM dNTPS	2.5
Oligonucleotide for 5' to 3'	2.5
Oligonucleotide for 3' to 5'	2.5
Taq polymerase	1
Template	0.5
H_2O	36

Run using a Biometra PCR Machine with 30 cycles of: one minute at 94°C to denature the DNA. One minute at 55°C to anneal the DNA. One minute at 72°C to elongate the DNA. This was followed by fifteen minutes at 72°C at the end of the cycles to enable the polymerase to fully complete the reactions.

4.2.9 Agarose Gel Electrophoresis

Submerged horizontal agarose mini-gels were used for identification, sizing and purification of DNA fragments. The percentage of agarose used in the gel varied between

1% and 2% w/v depending on the size of the DNA being investigated. The agarose was dissolved in 1 times TAE buffer by boiling, in a microwave, until the solution was clear. The temperature was then allowed to cool to 50°C before ethidium bromide was added to a final concentration of 0.3μg/mL. The gel solution was then poured onto the gel plate to sufficient depth to cover the teeth of a comb, which was then inserted. The gel was allowed to set before submerging in an electrophoresis tank containing 1 times TAE buffer and optionally 0.2μg/mL ethidium bromide before the comb was removed. Samples were prepared by the addition of Promega blue/orange 6 times gel loading dye and water as necessary to enable the loading of a minimum 5ng. The gel was run on a Bio Rad power pac3000 at a constant voltage of 100V until sufficient separation had occurred. The fluorescent DNA bands were visualised with a UV transluminator.

50 times TAE buffer stock was made as follows: 242g Tris base, 57.1mL glacial acetic acid, 100mL 0.5M EDTA pH8.0, distilled water to 1litre.

4.2.10 Purification of DNA from an Agarose Gel

DNA fragments from 70bp to 10kbp could be purified using a QIAquick Gel Extraction Kit. This dissolves the agar and the DNA is subsequently purified using a QIAquick spin column, which binds DNA and requires centrifugal force (10000g) to pull fluids through. The DNA band was cut from the gel using a fresh scalpel blade and placed in an eppendorf. The QIAquick Gel Extraction Kit was used as recommended by the manufacturers instructions. The DNA was eluted in water.

4.2.11 Restriction Digestion of DNA

The use of restriction enzymes to digest DNA was performed according to the manufacturers recommendations using the buffers purchased.

4.2.12 Dephosphorylation of DNA

The 5' phosphate group was removed from vector DNA to reduce self ligation. Calf intestinal alkaline phosphatase (CIP) was added to the vector restriction digests, for the last fifteen minutes in the ratio of 0.5 units : 1μg.

4.2.13 Purification of DNA after Restriction Digestion

When large concentrated amounts of DNA were to be purified a QIAquick PCR Purification Kit was used. It works using a QIAquick column which binds DNA and requires centrifugal force (10000g) to pull fluids through. The QIAquick PCR Purification Kit was used as in the manufacturers instructions. The DNA was eluted in water.

For small dilute amounts of DNA a phenol : chloroform extraction followed by ethanol precipitation was used.

4.2.14 Phenol : Chloroform Extraction of DNA

Phenol extraction was used to denature and remove any contaminating protein from a DNA solution. An equal volume of saturated phenol was added to the DNA solution and vortexed until an emulsion had formed. The sample was then centrifuged (1min, 12000g). The upper aqueous layer containing the DNA was removed and placed in a new tube. Care was taken not to transfer any of the protein found at the interface. Chloroform extraction was then used to remove traces of phenol. An equal volume of chloroform was added to the DNA solution and vortexed to form an emulsion. The sample was then centrifuged at (1min, 12000g). The upper aqueous layer containing the DNA was removed and placed in a fresh tube until required.

4.2.15 Ethanol Precipitation of DNA

DNA solutions were concentrated by addition of 0.1 relative volumes of 3M sodium acetate pH5.2 and 2.5 relative volumes of ice cold 95% ethanol. The solutions were mixed and then centrifuged (20min, 12000g, 4°C). The supernatant was then removed. The pellet washed with ice cold 70% ethanol and centrifuged (1min, 12000g, 4°C). The pellet was then air dried for five minutes before being resuspended in the required volume of water.

4.2.16 Ligation of DNA

Sticky-end and blunt-end ligation was performed using a Rapid DNA Ligation Kit. This is a commercially available kit with buffers that have been optimised for this process. The cut vector and insert were diluted in DNA dilution buffer so that in a total volume of 10μL the ratio was 1:3. 10μL of T4 DNA ligation buffer was thoroughly mixed before 1μL T4

DNA ligase was added. The mixture was allowed to incubate for five minutes at room temperature before transformation into non expression bacteria.

4.2.17 Screening of Colonies using PCR

PCR screening uses oligonucleotides that match the plasmid on either side of the insertion site making it possible to see if the size of the insert is correct. Plasmids from ligation were transformed into XL1 blues and grown on plates. Individual colonies were picked with a sterile loop and transferred initially to a 20μL PCR mixture, where only a few cells need to detach and lyse to release enough DNA to work with. The same loop was then used to inoculate a 3mL LB and relevant antibiotic growth mixture which was incubated at 37°C whilst shaking at 100rpm.

The PCR was run on the same program as described in section 4.2.8. The product was run on an agarose gel and the colonies showing the correct size insert were used to inoculate LB media ready for plasmid DNA preparation.

4.2.18 Plasmid DNA Preparation

Small amounts (up to 20μg) of plasmid DNA were prepared using a QIAprep Spin Miniprep Kit which was used as described in the manufacturers instructions. The DNA was eluted in water.

Large amounts (up to 100μg) of plasmid DNA were prepared using a QIAfilter Plasmid Midi Kit which was used as described in the manufacturers instructions. The DNA was eluted in water.

4.2.19 Screening of DNA using Restriction Digests

This uses restriction enzymes to cut both in the insert and the plasmid to see if the correct length of DNA can be seen on a gel. The restriction sites can be found by using internet programs (*e.g.* webcutter) and the best possible choice is a restriction enzyme that will cut both insert and plasmid once only, to give two products of different length.

4.2.20 Measurement of DNA Concentration

The optical density at 260 nm allowed the DNA concentration to be estimated.

For double stranded DNA 1AU at A_{260} =50μg/mL

For single stranded DNA 1AU at A_{260} =40μg/mL

4.2.21 Sequencing of DNA

Sequencing was done using a capillary sequencer. This sequencing uses oligonucleotides that match the plasmid on either side of the insertion site. It works by only having one oligonucleotide in a PCR mixture and allowing a polymerase to start there and run along the plasmid until it falls off (about 700bp). In the mixture are nucleotides tagged with fluorescent dyes, so that each individual base can be labelled and observed. The process is then repeated using an oligonucleotide that starts at the other side of the insert.

A typical sequence reaction would consist of the following:

7.5 μL sequencing reaction	μL
DNA	1
Oligonucleotide (5pmoles/mL)	1
Sequence Buffer	2
Sequence Reaction Mixture	1
Water	2.5

This would be run with a PCR Sequence of 30 seconds at 96°C to warm up the machine, followed by a loop of 24 cycles of: ten seconds at 96°C to denature the DNA. Five seconds at 50°C to anneal the DNA. Four minutes at 60°C to elongate the DNA.

4.3 Protein Techniques

4.3.1 SDS-PAGE Analysis

4.3.1.1 Sample Preparation

Protein samples were boiled with one times gel loading buffer (Tris pH6.8 560mM, Dithiothreitol 100mM, SDS 2%, Glycerol 10% Bromophenol Blue R 0.1%) and loaded at 20ng of protein per well.

4.3.1.2 Self Poured SDS-PAGE Gels

The gel apparatus (Bio Rad mini-protean II electrophoresis cell) was setup according to the manufacturers recommendations. Solutions of both stacking and running gels were treated with the required volume of fresh ammonium persulfate immediately before setting with

TEMED. The gels were run with a Bio Rad power pac3000 using separate inside cathode (100mM Tris, 100mM Tricine 0.1% SDS pH8.2) and outside anode (200mM Tris pH8.9) buffers. The gels were run for fifty minutes at 150V.

Gel Component	Stacking Gel	Separating Gel
Polyacrylamide (30% w/v)	0.66mL	3.33mL
Analytical Water	2.9mL	2.03mL
Glycerol	-	1.3mL
3M Tris, O.3% SDS, pH8.45	1.24mL	3.33mL
Ammonium persulfate (100mg/mL)	25µL	50µL
TEMED	10µL	10µL

4.3.1.3 Phastgels

8-20% SDS-PAGE gels were run on an Amersham Pharmacia Biotech Phastgel system according to manufacturers recommendations.

4.3.1.4 Staining and Destaining

SDS-PAGE gels were stained (Bromophenol Blue R 0.5g, Methanol 45mL, H_2O 45mL, Glacial acetic acid 10mL) for at least thirty minutes at room temperature before destaining (Methanol 45mL, H_2O 45mL, Glacial acetic acid 10mL). The destain was refreshed as necessary.

4.3.2 Determination of Protein Concentration

Protein concentration was determined either by using the absorption at 280 or using the Bio-Rad protein assay. A 1mL Bio-Rad protein assay consisted of 10µL of protein solution 790µL of distilled water and 200µL of Bio-Rad reagent. Bio-Rad reagent is based on the Bradford Assay[176]. The protein binds to the dye Coomassie brilliant blue G-250 in a concentration dependent manner. This assay was routinely used for protein samples between 200 and 1400µg/mL. The colour was allowed to develop for five minutes at room

temperature and the absorbance was measured at 595nm in a 1mL cuvette. The concentration was calculated using the conversion AU 0.1=1.95μg of protein.

4.3.3 FPLC

Samples were run on a Amersham Pharmacia Biotech FPLC system according to the manufacturers recommendations.

4.3.4 UV

UV measurements were made using a Hitachi U-2000 spectrometer according to the manufacturers recommendations.

4.3.5 CD

CD measurements were made using a Jassco J-720 spectropolarimeter according to the manufacturers recommendations. The cuvettes path length was 0.1mm.

4.3.6 Mass Spectrometry

Mass spectrometry was done using a Micromass Quatro II, triple quadropole electrospray. The samples were run in 50:50 water:acetonitrile.

4.3.7 Sonication

All the protein sonications were performed using a Sanyo MSE soniprep 150 ultrasonic disintegrator. For eppendorf gel samples a titanium exponential microprobe, end diameter 3mm, was used for 30 seconds with a 10 micron amplitude. For batch purifications the sample was contained within a 50mL sonication vessel upon an ice bath. A solid probe, end diameter 9.5mm, was used for 20 cycles of 20s on, 40s off with a 10 micron amplitude.

4.3.8 Protein Databases and General Programs

All databases were accessed via ExPASy [177] with the protein sequences taken from SWISS-PROT/TREMBL [178] and analysed using ProtParam. All structures were downloaded from the Protein Data Bank [179] except for CCR3 which was generated using SWISS-Model. All models were made using MOLMOL [180] and all sequence alignments done using ClustalX [181].

4.4 Protocols for CCL24

4.4.1 Plasmid for the Sequence G25 to P98

CCL24 cDNA was amplified by PCR, using the primers shown in Figure 4.1 and a band of approximately 300bp isolated from an agarose gel. The DNA was then cut with Nco 1 and BamH1 and purified using the QIAquick PCR purification method. The same enzymes were used to cut pET-15b (Figure 4.3), which was then purified by phenol: chloroform extraction followed by ethanol precipitation. The plasmid and the insert were then ligated and transformed into XL1 blues. The resulting colonies were screened by PCR, using oligonucleotides that corresponded to the T7 promoter and termination sequences Figure 4.2, before plasmid DNA preparation. This DNA was then screened using the restriction enzymes Cla 1 and Sac 1, before capillary sequencing.

5' oligonucleotide for G25 to P98

GGCATGCCATGGGCAGCAGCCATCATCATCATCATCATCATCATCATCATAG

CAGCGGCCTGGTGCCGCGCGGAAGTGTGGTCATTCCAAGCCCATGC

3' oligonucleotide for G25 to P98

GGGCGCGGATCCTTATGGGGAGGCCTTCTTCTGC

Figure 4.1 The primers for amplification of G25 to P98.

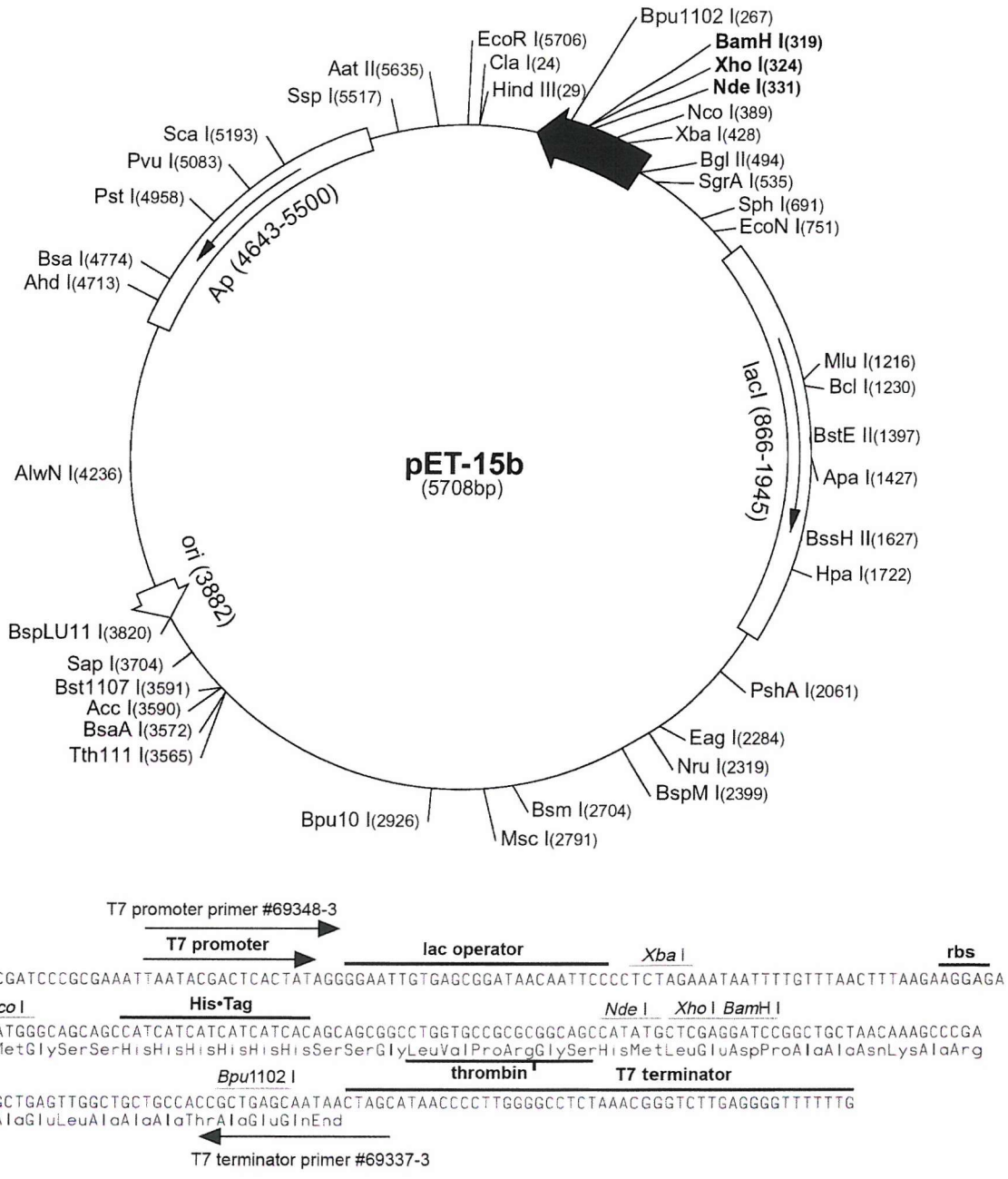
T7 promoter sequence

TAATACGACTCACTATAGGG

T7 termination sequence

GCTAGTTATTGCTCAGCGG

Figure 4.2 The primers corresponding to the T7 promoter and termination sequences.



B

Figure 4.3 The pET-15b vector map. A. The circle map representation of the vector which shows the unique restriction enzyme cut sites. B. The cloning and expression region sequence. This shows the position of the restriction enzyme cut sites in relation to the inbuilt hexa-His Tag and the thrombin cleavage site.

4.4.2 Plasmid for the Sequence G25 to C119S

CCL24 cDNA was amplified by PCR, using the primers shown in Figure 4.4 and a band of approximately 400bp isolated from an agarose gel. The DNA was then cut with Nco 1 and BamH 1 and purified using the QIAquick PCR purification method. The same enzymes were used to cut pET-15b (Figure 4.3), which was then purified by phenol: chloroform extraction followed by ethanol precipitation. The plasmid and the insert were then ligated

together and transformed into XL1 blues. The resulting colonies were screened by PCR, using oligonucleotides that corresponded to the T7 promoter and termination Figure 4.2, before plasmid DNA preparation. This DNA was then screened using the restriction enzymes Cla 1 and Sac 1, before capillary sequencing.

5' oligonucleotide for G25 to C119S

GGCATGCCATGGGCAGCAGCCATCATCATCATCATCATCATCATCATCATAG

CAGCGGCCTGGTGCCGCGCGGAAGTGTGGTCATTCCAAGCCCATGC

3' oligonucleotide for G25 to C119S

GGGCGCGGATCCTTAGGAGGTGGTTTGGTTGCCAGG

Figure 4.4 The primers for amplification of G25 to C199S.

4.4.3 Plasmid for the Sequence M1 to C119

CCL24 cDNA was amplified by PCR, using the primers shown in Figure 4.5 and a band of approximately 400bp isolated from an agarose gel. The DNA was then cut with Nde 1 and BamH 1 and purified using the QIAquick PCR purification method. The same enzymes were used to cut pET-15b (Figure 4.3), which was then purified by phenol: chloroform extraction followed by ethanol precipitation. The plasmid and the insert were then ligated together and transformed into XL1 blues. The resulting colonies were screened by PCR, using oligonucleotides that corresponded to the T7 promoter and termination sequence Figure 4.2, before plasmid DNA preparation. This DNA was then screened using the restriction enzymes Cla 1 and Sac 1, before capillary sequencing.

5' oligonucleotide for M1 to C119

GGGAATTCCATATGGCAGGCCTGATGACCATAG

3' oligonucleotide for M1 to C119

GGGCGCGGATCCTTAGCAGGTGGTTTGGTTGC

Figure 4.5 The primers for amplification of M1 to C119.

4.4.4 Expression of the CCL24 Vectors

The plasmid was transformed into BL21(DE3) and selected for, using ampicillin, on an agar plate. A 5mL LB sample was inoculated and grown to a mid log phase at 37°C,

induced for three hours at 37°C and the total cell extract run on an SDS-PAGE gel. With this positive result a further 5mL LB growth was done, but this time the cell pellet was sonicated and a SDS-PAGE sample removed to show the total protein composition of the cells. The sonicate was then centrifuged (15min, 15000g) and a further SDS-PAGE sample prepared from the supernatant, to quantify the amount of the protein in a soluble form. The growth was repeated, but after growing to mid log phase at 37°C, the sample was cooled to 30°C before induction for three hours. Another corresponding growth was performed with all steps occurring at 30°C.

4.4.5 His Tag Purification of G25 to P98

Two litres of cell pellet were resuspended in lysis buffer (50mM sodium hydrogen phosphate, 300mM sodium chloride pH8) containing a Roche EDTA free antiprotease tablet, sonicated, centrifuged (20min, 25000g) and the supernatant loaded onto a nickel affinity column, attached to an FPLC, that had previously been equilibrated in lysis buffer. Once a stable baseline had been achieved, the column was washed with five column volumes of wash buffer (50mM sodium hydrogen phosphate, 300mM sodium chloride, 20mM imidazole pH8) before elution with elution buffer (50mM sodium hydrogen phosphate, 300mM sodium chloride, 250mM imidazole pH8). For elution by changing the pH the elution buffer was changed to 50mM sodium hydrogen phosphate, 300mM sodium chloride pH5.5.

4.4.6 Purification using Traditional Methodologies

Two litres of cell pellet where resuspended (50mM sodium hydrogen phosphate, 100mM sodium chloride, 2mM DDT, 2mM PMSF pH8), sonicated, centrifuged (20min, 25000g). The resulting supernatant was diluted 1:4 (50mM sodium hydrogen phosphate, 2mM DDT, pH8) and then loaded onto a cationic exchanger (SP Sepherose) column, attached to an FPLC, which had been pre-equilibrated (50mM sodium hydrogen phosphate, 2mM DDT, pH8). A gradient from 0M to 2M sodium chloride was run.

The preparation was repeated, but this time the cationic exchanger was swapped for an anionic exchanger (Q Sepherose) and a gradient from 0M to 1M sodium chloride run.

4.4.7 Purification from Inclusion Bodies

For 2 litres of culture volume, the pellet was resuspended in 15mL of buffer A (100mM KCl, 2mM DTT, 2mM PMSF, 10mM Tris pH8.5, 25% (w/v) sucrose) and 4mL of buffer B (300mM Tris pH8.5 100mM EDTA, 4mg/mL lysozyme) and left on ice for thirty minutes. 20mL of buffer C (1M LiCl, 20mM EDTA, 0.5% (v/v) NP-40) was added before sonication. The sonicate was centrifuged (20min, 30000g) and the pellet resuspended in 50mL of buffer D (10mM Tris-HCl pH8.5 0.1mM EDTA, 0.5M LiCl, 0.5% (v/v) NP-40, 1mM DTT, 1mM PMSF) and resonicated and recentrifuged. The pellet was resuspended for a second time in buffer D and the process repeated. After centrifugation the pellet was resuspended in buffer E (10mM Tris-HCl pH8.5, 0.1mM EDTA, 0.5% (v/v) NP-40, 1mM DTT, 1mM PMSF) and resonicated and recentrifuged. The pellet was suspended in buffer E again and the process repeated. Finally the pellet was resuspended in buffer E and centrifuged.

4.4.7.1 Refolding Methodologies

The inclusion body was suspended in denaturant (6M GuHCl, 25mM DTT, 50mM sodium acetate pH5.5) and stirred for two hours at room temperature before been centrifuged (20min, 25000g). The supernatant was diluted dropwise at 4°C such that the final protein concentration was 1.5mg/mL, and the buffer contained 1.5M guanidine hydrochloride, 25mM DTT, 50mM sodium acetate pH5.5. This was left overnight at room temperature.

4.4.7.2 Protein Purification

The refolded protein was diluted 1 in 15 (25mM DTT, 50mM sodium acetate pH5.5) and the resulting solution was centrifuged (20min, 3000g) before loading onto a cationic exchange (SP Sepherose) column that had been pre-equilibrated (50mM sodium acetate, 5mM DTT pH5.5) and a 0 to 2M salt gradient run.

4.4.7.3 Reforming the Disulfide Bonds

Air oxidation was performed by gentle stirring the protein (1mg/mL) in 50mM sodium acetate pH5.5 overnight at room temperature. Oxido shuffling via reduced and oxidized glutathione (GSH/GSSH) was performed as above but with the addition of 12mM reduced

glutathione and 1.2mM oxidized glutathione. Separation of the oxidised from the reduced form was performed using a C18 column on a HPLC with an acetonitrile gradient.

4.4.7.4 Removal of His Tag with Thrombin Cleavage

The tag was cleaved either using 2 units of thrombin per mg of protein at 4°C overnight in thrombin cleavage buffer (50mM sodium acetate, 2.5mM calcium chloride, 100mM sodium chloride) or with 5units per mg of protein at room temperature for 4hrs in thrombin cleavage buffer.

4.4.8 Production of ^{15}N Labelled CCL24

The plasmid was transformed into BL21(DE3) and the ampicillin resistant colonies selected for using agar plates. The inoculation cultures were grown overnight and used to inoculate ampicillin resistant LB media at a ratio of 1:40. The cultures where grown until an OD_{595} of 1 was achieved, before being centrifuged (20min, 3000g, 30°C). The pellet was resuspended in an equivalent volume of M9 media, to that of LB media and left to grow for thirty minutes at 37°C before inducing for four hours.

4.5 Protocols for CCL11

4.5.1 Growth

A Novagen pET-15b plasmid (Figure 4.3) containing the DNA sequence coding for CCL11 was used. The plasmid was transformed into BL21(DE3) and the ampicillin resistant colonies were selected for via agar plates. The inoculation cultures were grown to an OD_{595} of 0.7 and then centrifuged (15min, 2000g, 4°C). The pellets were then resuspended in fresh ampicillin LB, and stored at 4°C overnight. Ampicillin resistant LB media was inoculated with these cultures at a ratio of 1:100 and grown at 37°C to an OD_{595} of 0.7. The flasks where then induced with IPTG for between three to five hours before centrifugation (20min, 15000g, 4°C) and the pellet stored at -20°C until used in purification.

4.5.2 Initial Method of His Tag Purification using a Nickel Column

The cells were allowed to reach room temperature and resuspended in a minimal amount of lysis buffer (50mM NaH_2PO_4, 300mM NaCl, 10mM imidazole pH8) containing Roche

EDTA free protease inhibitor. The cells were then sonicated before been pelleted using an ultracentrifuge (30min, 40000g) and the supernatant loaded onto a nickel column. The column was washed with wash buffer (50mM NaH_2PO_4, 300mM NaCl, 20mM imidazole pH8), before the protein was eluted with elution buffer (50mM NaH_2PO_4, 300mM NaCl, 250mM imidazole pH8). The fractions containing protein were combined and concentrated to less than 2mL using a vivaspin 5000, before being loaded onto a Superdex 75 Hiload 16/60 column. The column was run (50mM NaH_2PO_4, 300mM NaCl) and the fractions containing CCL11 pooled and lyophilised before desalting using a PD10.

4.5.3 Final Method of His Tag Purification using a Nickel Column

The cells were allowed to reach room temperature and resuspended in a minimal amount of lysis buffer (50mM NaH_2PO_4, 300mM NaCl pH8) containing Roche EDTA free protease inhibitor. The cells were then sonicated before being pelleted (2Omin, 20000g) the supernatant loaded onto a nickel column. The pellet was resuspended in a minimal amount of lysis buffer (50mM NaH_2PO_4, 300mM NaCl pH8) containing Roche EDTA free protease inhibitor, sonicated before being pelleted (2Omin, 20000g). The supernatant was also loaded onto the nickel column. The column was washed with wash buffer (50mM NaH_2PO_4, 300mM NaCl, 20mM imidazole pH8), before the protein was eluted with elution buffer (50mM NaH_2PO_4, 300mM NaCl, 250mM imidazole pH8). The fractions containing protein were combined, concentrated to half its volume using an Amicon (1Kda) at 4°C, before dilution back to its original volume (50mM NaH_2PO_4 pH8). This was repeated twice to remove the imidazole before finally being concentrated to less than 10mL. The sample was then diluted to 50mL before being loaded onto a cationic exchanger (SP Sepharose), attached to an FPLC, which had been pre-equilibrated (50mM sodium hydrogen phosphate pH8). A gradient from 0M to 2M sodium chloride was run. Those fractions that by SDS-PAGE analysis contained pure CCL11 were pooled and lyophilised before being desalted using a PD10.

4.5.4 Production of ^{15}N Labelled CCL11

The plasmid was transformed into the rosetta strain and the ampicillin resistant colonies selected for using agar plates. The inoculation cultures were grown overnight and used to inoculate ampicillin resistant LB media at a ratio of 1:40. The cultures where grown until an OD_{595} of 1 was achieved, before being centrifuged (20min, 3000g, 30°C). The pellet

was resuspended in an equivalent volume of M9 media, to that of LB media and left to grow for thirty minutes at 37°C before inducing for four hours.

4.6 Suspension of the Peptides in DPC

150mM DPC in a 20mL scintillation vial was suspended in fresh vial d_6-DMSO (600μL). 0.5mM of the peptide in an eppendorf was dissolved in DMSO (150μL). The DPC was then very gentle heated using a water bath at 60°C, with the compound never spending more than one second at any time in direct contact with the water. Once the DPC had dissolved the pre-dissolved peptide was added and immediately the vial was swirled to mix the two solutions and plunged into ice. The sample was lyophilized for two days in the dark. The sample was stored in this form until required at which point it was resuspended in the NMR buffer (20mM sodium acetate-d_6, 1mM sodium azide, 10% D_2O pH5.5) and sonicated. The peptide sonications were done using a Ultrawave U100 (Ultrawave Ltd. Cardiff) and were performed in 10min stages until the sample cleared or three cycles which ever came first. The solution was then transferred to an eppendorf and centrifuged (15min, 13000g) before use.

4.7 Nuclear Magnetic Resonance Spectroscopy

Nuclear magnetic resonance experiments were performed on a Varian INOVA 600 MHz spectrometer. The data sets were viewed in three different ways: 2D-homonuclear experiments were viewed using VNMR, as supplied with the spectrometer. 2D-heteronuclear experiments were initially viewed with VNMR before data analysis using nmrview. 3D data sets were processed using nmrPipe [182] and viewed with nmrDraw and PIPP [183].

4.7.1 *ACT*-ACP NMR Spectroscopy

4.7.1.1 Mutant R72A

3mg of R72A mutant was dissolved in 600μL aqueous buffer (20mM sodium acetate, 1mM sodium azide, 10% D_2O pH5.05). Standard pulse sequences were employed to obtain NOESY (mixing times 0.15s, spectral width 6500Hz in both dimensions, 512

complex t_1 points and 32 transients) and TOCSY (mixing time 50ms, spectral width of 6500Hz in both dimensions, 512 complex t_1 points and 32 transients) at 25°C.

4.7.1.2 Mutant Y56A

3mg of Y56A was dissolved in 600µL aqueous buffer (2mM di-potassium hydrogen orthophosphate, 1mM sodium azide, 10% D_2O pH6). Standard pulse sequences were employed to obtain NOESY (mixing times 0.15s, spectral width 7000Hz in both dimensions, 400 complex t_1 points and 40 transients) and TOCSY (mixing time 50ms, spectral width of 7000Hz in both dimensions, 400 complex t_1 points and 40 transients) at 25°C.

4.7.2 Peptide Loop NMR Spectroscopy

4.7.2.1 Initial Loop 1

6mg peptide was dissolved in 600µL aqueous buffer (20mM sodium acetate, 1mM sodium azide, 10% D_2O pH5.09). Standard pulse sequences were employed to obtain DFQ-COSY, NOESY (mixing times 0.15, 0.3, 0.4, and 0.5s, spectral width 8000Hz in both dimensions, 128 complex t_1 points and 48 transients), TOCSY (mixing time 50ms, spectral width of 8000Hz in both dimensions, 256 complex t_1 points and 48 transients) and ROESY (mixing time 75ms, spectral width 8000Hz in both dimensions, 256 complex t_1 points and 64 transients) spectra at 30°C. At 8°C, a TOCSY (mixing time 50ms, spectral width of 8000Hz in both dimensions, 256 complex t_1 points and 48 transients) and a NOESY (mixing time 0.4ms, spectral width 8000Hz in both dimensions, 128 complex t_1 points and 256 transients) were collected. The spectra were referenced to the water peak which had previously been referenced to TMS.

4.7.2.2 N-Terminus

7mg peptide was dissolved in 600µL aqueous buffer (20mM sodium acetate, 1mM sodium azide, 10% D2O pH5.09). Standard pulse sequences were employed to obtain DFQ-COSY, NOESY (mixing time 0.4s, spectral width of 6000Hz in both dimensions, 256 complex t_1 points and 64 transients) and TOCSY (mixing time 50ms, spectral width 6000Hz in both dimensions, 256 complex t_1 points and 32 transients) spectra at both 30°C

and 8°C. The spectra were referenced to the water peak which had previously been referenced to TMS.

4.7.2.3 Redesigned Loop 1

5mg of peptide was dissolved in 600μL of DMSO-d_6. Standard pulse sequences were employed to obtain TOCSY (mixing time 50ms, spectral width 8500Hz in both dimensions, 256 complex t_1 points and 48 transients) and NOESY (mixing time 0.4s, spectral width 8500Hz in both dimensions, 384 complex t_1 points and 64 transients) spectra at 25°C and 20°C. The spectra were referenced to a standard of TMS in DMSO and assigned.

The resuspended peptide in DPC had standard pulse sequences employed to obtain TOCSY (mixing time 40 and 50ms, spectral width 8500Hz in both dimensions, 512 complex t_1 points and 16 transients) and NOESY (mixing time 0.15 and 0.4s, spectral width 8500Hz in both dimensions, 512 complex t_1 points and 16 transients) spectra at 25°C. The spectra were referenced to the water peak which had previously been referenced to TMS.

4.7.2.4 Redesigned Loop 2

5mg of peptide was dissolved in 600μL of DMSO-d_6. Standard pulse sequences were employed to obtain TOCSY (mixing time 50ms, spectral width 8200Hz in both dimensions, 512 complex t_1 points and 64 transients) and NOESY (mixing times 0.15, 0.25 and 0.4s, spectral width 8200Hz in both dimensions, 512 complex t_1 points and 32 transients) spectra at 25°C. The spectra were referenced to a standard of TMS in DMSO.

The resuspended peptide in DPC had standard pulse sequences employed to obtain TOCSY (mixing time 50ms, spectral width 8500Hz in both dimensions, 256 complex t_1 points and 160 transients) and NOESY (mixing time 0.15 and 0.4s, spectral width 8500Hz in both dimensions, 256 complex t_1 points and 160 transients) spectra at 25°C. The spectra were referenced to the water peak which had previously been referenced to TMS.

4.7.2.5 Redesigned Loop 3

5mg of peptide was dissolved in 600μL of DMSO-d_6. Standard pulse sequences were employed to obtain TOCSY (mixing time 50ms, spectral width 6500Hz in both dimensions, 512 complex t_1 points and 32 transients) and NOESY (mixing time 0.4s, spectral width 6500Hz in both dimensions, 512 complex t_1 points and 32 transients) spectra at 25°C and 20°C. The spectra were referenced to a standard of TMS in DMSO and assigned.

The resuspended peptide in DPC had standard pulse sequences employed to obtain TOCSY (mixing time 50ms, spectral width 8500Hz in both dimensions, 256 complex t_1 points and 64 transients) and NOESY (mixing time 0.4s, spectral width 8500Hz in both dimensions, 256 complex t_1 points and 64 transients) spectra at 25°C. The spectra were referenced to the water peak which had previously been referenced to TMS.

4.7.3 Structural Calculations.

Structural calculations were performed using the hybrid distance geometry simulated annealing protocol [184] within the program XPLOR version 3.8 [185]. One hundred and fifty structures were generated for each structure. The final family was chosen based on low XPLOR energies, agreement with experimental constraints and the quality of the structures as determined using the program AQUA/Procheck [140].

4.7.4 Titrations

4.7.4.1 CCL24 Titration with Redesigned Loop 1 in DPC

0.5mM peptide mimic of loop 1 was dissolved in 600μl of 100mM acetate, 150mM DPC pH5.5. 1mM CCL24 was dissolved in 20μl of 100mM acetate, 150mM DPC pH5.5. CCL24 was titrated into the peptide mimic and five HSQCs were recorded at CCL24 concentrations of 0.05mM, 0.1mM, 0.25mM, 0.5mM and 1mM. All the HSQCs where recorded with a spectra width of 8000Hz in both dimensions and 64 complex t_1 points. The number of transients varied with the concentration of CCL24: 512 (0.05mM), 512 (0.1mM), 256 (0.25mM), 96 (0.5mM) and 64 (1mM). A 2D-NOESY was recorded before the start of the experiment (mixing time of 0.4s, spectral width of 8500Hz in both dimensions, 384 complex t_1 points and 64 transients). A ^{15}N-filtered NOESY was

recorded at 1mM CCL24 (mixing time of 0.4s, spectral width of 8500Hz in both dimensions, 256 complex t_1 points and 128 transients).

4.7.4.2 CCL24 Titration with Redesigned Loop 3 in DPC

CCL24 was titrated into 0.5mM peptide mimic of loop 3 suspended in 100mM acetate, 150mM DPC pH5.5. Six HSQCs were recorded at CCL24 concentrations of 0.05mM, 0.1mM, 0.25mM, 0.5mM, 1mM and 2mM. All the HSQCs where recorded with a spectra width of 8000Hz in both dimensions and 64 complex t_1 points. The number of transients varied with the concentration of CCL24 768 (0.05mM), 768 (0.1mM), 320 (0.25mM), 80 (0.5mM), 80 (1mM) and 32 (2mM). A 2D-NOESY was recorded before the start of the experiment (mixing time of 0.4s, spectral width of 8500Hz in both dimensions, 256 complex t_1 points and 64 transients). ^{15}N-filtered NOESY were recorded at 1mM and 2mM CCL24 (mixing time of 0.4s, spectral width of 8500Hz in both dimensions, 256 complex t_1 points and 128 transients).

4.7.5 Interpretation of the Chemical Shift Data

The weighted chemical shifts where calculated using:

$$\mathrm{Sqrt}\{(\delta_N/5)^2+\delta_H^2\} \quad (1)$$

Where δ_N and δ_H are the chemical shift differences in the nitrogen and hydrogen dimensions respectively.

For a standard titration where the protein concentration is kept constant and the peptide concentration increased, the chemical shifts observed for the protein in the presence of the peptide are the population weighted averages of the chemical shifts of free and bound CCL24 [111]. Therefore, the observed chemical shift change ($\Delta\delta_{obs}$)is given by:

$$\Delta\delta_{obs}=\Delta\delta_{max}(C/C_{eot}) \quad (2)$$

In which $\Delta\delta_{max}$ is the chemical shift difference between the bound and free CCL24. C_{eot} is the total concentration of CCL24 and C is the concentration of bound CCL24 which is given by:

$$C=0.5[(K_d+C_{eot}+C_{pep})-\mathrm{sqrt}\{(K_d+C_{eot}+C_{pep})^2-4C_{eot}C_{pep}\}] \quad (3)$$

Where C_{pep} is the total concentration of the peptide and K_d is the dissociation constant of the CCL24-peptide complex.

However in this case the peptide concentration was kept constant and the protein concentration increased, which means that whilst the chemical shifts observed in the presence of the peptide are the population weighted averages of the chemical shifts of free and bound CCL24, the ratio of the concentration of CCL24 to the peptide must be considered. Therefore, the observed chemical shift change is given by:

$$\Delta\delta_{obs}=\Delta\delta_{max}(1-[C/C_{pep}]) \quad (4)$$

In which $\Delta\delta_{max}$ is the chemical shift difference between the bound and free CCL24. C_{pep} is the total concentration of the peptide and C is the concentration of bound CCL24 which is given by equation 2.

For each residue with significant chemical shift deviation a plot of $\Delta\delta_{obs}$ against C_{pep} had the parameters $\Delta\delta_{max}$ and K_d simultaneously fitted using SigmaPlot Regression Wizard.

4.7.6 Interpretation of the Peak Intensities

Due to the changes in the concentration and the number of scans during the titration the measured intensities were not directly comparable and with no literature standard it was necessary to devise a novel method of interpretation. The intensities where divided by the number of scans because the intensity should have increased linearly with the number of scans. This value was then multiplied by the ratio of the concentration to the end concentration of the titration. These figures where then normalized such that the end point was 1. An error bar was devised based upon the noise value for each spectrum which was divided by the square root of the number of scans, because the noise was random and not linear with the number of scans; the noise should not be variable with the concentration of the sample. If the peaks had a Gaussian distribution then the 95% error bar would be twice the noise. This information was plotted for each residue. Where a horizontal line could be drawn through all the error bars of the data points then the residue was assumed not to be showing any significant change in intensity.

Appendix I Chemokine Nomenclature

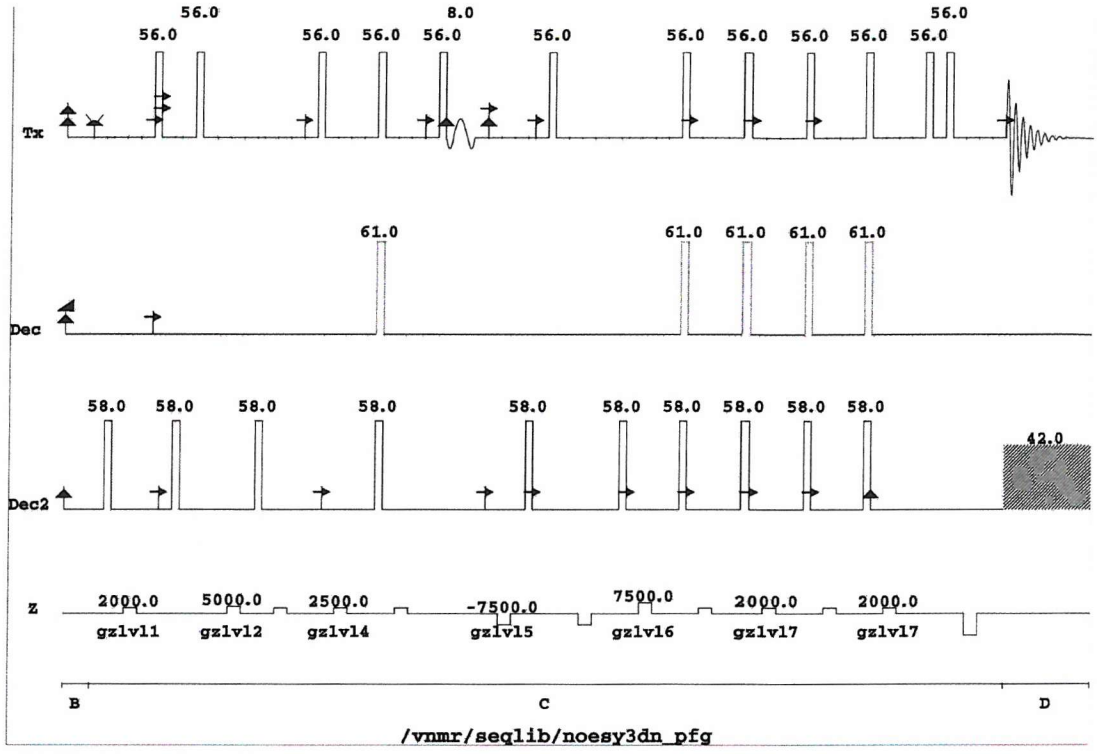
The nomenclature of the chemokine receptors.

Name	Previous Name	Selective Chemokine
CXC Chemokine receptors		
CXCR1	IL8RA; IL-8R-I; IL-8Rα	None
CXCR2	IL8RB; IL-8R-II; IL-8Rβ	CXCL1; CXCL5; CXCL7
CXCR3	IP10/Mig R, GPR9	CXCL9; CXCL10; CXCL11
CXCR4	HUMSTER; LESTR; fusin; HM89; LCR1; NPYR;	CXCL12
CXCR5	BLR-1; MDR15	CXCL13
CXCR6		CXCL16
CC Chemokine receptors		
CCR1	CKR1; CC CKR1; MIP-1α/RANTES; CMKBR1	CCL14
CCR2	CKR2; CC CK2; CC CKR2; MCP-1; CMKBR2	CCL2
CCR3	CKR3; CC CKR3; eotaxin receptor; CMKBR3	CCL11; CCL24
CCR4	CKR4; CC CKR4; K5-5; CMKBR4; CHEMR1	CCL17; CCL22
CCR5	CKR5; CC CKR5; ChemR13; CMKBR5	CCL4
CCR6	GPR-CY4; CKR-L3; STRL22; DRY-6; DCR2; BN-1;	CCL20
CCR7	EBI-1; BLR-2; CMKBR7	CCL19; CCL21
CCR8	TER1; CKR-L1; GPR-CY6; ChemR1; CMKBR8	CCL1
CCR9	GPR 9-6	CCL25
CCR10	GPR2	CCL27
C Chemokine receptor		
XCR1	GPR5	XCL1; XCL2
CX3C Chemokine receptor		
CX3CR1	GPR13; V28;CMKBRL1	CX3CL1
Chemokine-binding proteins		
Duffy	DARC; glycoprotein D	None
D6	CCR9; CCR10	None

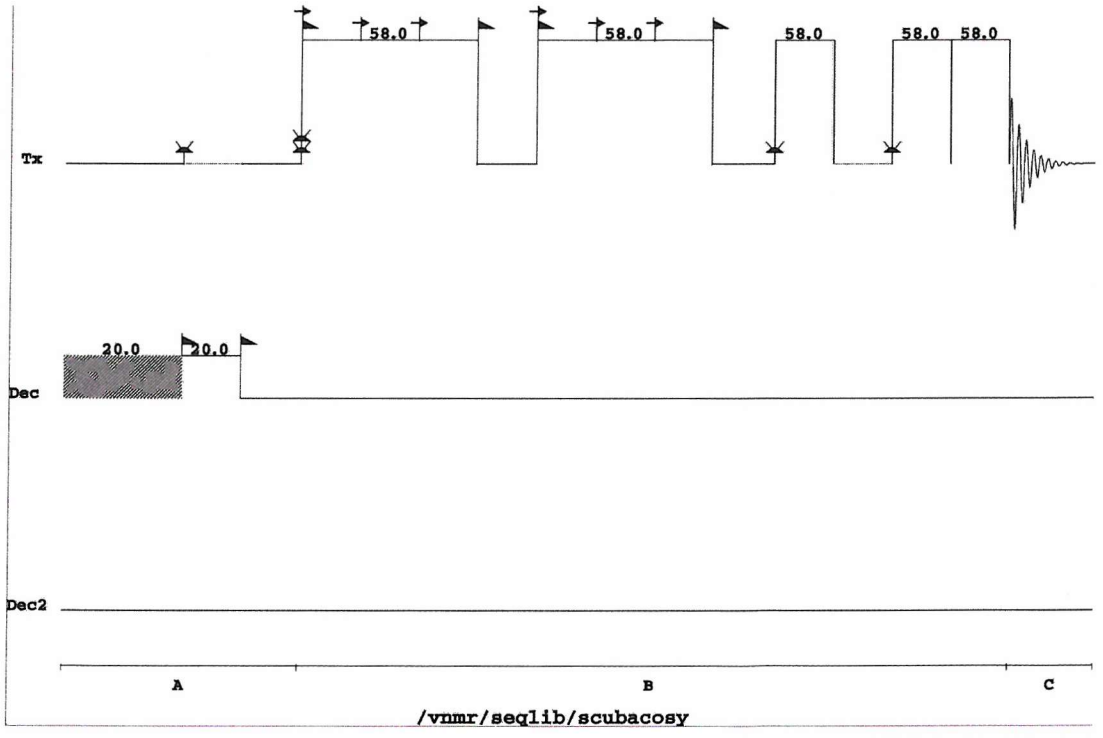
The nomenclature of the chemokine ligands.

Name	Common Synonyms	Subclass
CXC (α) Chemokines		
CXCL1	GROα; MGSA; SCYB1; NAP-3; GRO1 oncogene	ELR+
CXCL2	Groβ MIP-2α; SCYB2;GRO2 oncogene	ELR+
CXCL3	Groχ; MIP-2β; SCYB3 GRO3 oncogene	ELR+
CXCL4	Platelet factor-4; SCYB4	ELR-
CXCL5	ENA-78; SCYB5	ELR+
CXCL6	GCP-2; SCYB6	ELR+
CXCL7	PBP; CTAP-III; β- TG; NAP-2; SCYB7; low-affinity platelet factor 4	ELR+
CXCL8	IL8; SCYB8; MDNCF; NAP-1; LYNAP; NAF; GCP-1	ELR+
CXCL9	Mig; SCYB9	ELR-
CXCL10	γIP-10; SCYB10	ELR-
CXCL11	I-TAC; β-R1; IP9; H174; SCYB11	ELR-
CXCL12	SDF-1α; SDF-1β; PBSF; SCYB12; TPAR1; TLSF	ELR-
CXCL13	BCA-1; BLC; SCYB13	ELR-
CXCL14	BRAK; bolekine; SCYB14	ELR-
CXCL15 (reserved)		
CXCL16		
CC (β) Chemokines		
CCL1	I309	4 Cys
CCL2	MCP-1; MCAF; HC11	4 Cys
CCL3	MIP-1α;MIP-1αS; LD78α; GOS19-1; PAT 464.1; TY-5; SISα	4 Cys
CCL3L1	LD78β; MIP-1αP; GOS19-2; PAT 464.2	4 Cys
CCL4	MIP-1β; ACT-2; PAT 744; H400; SIS-χ; LAG-1; HC21; G-26; MAD-5	4 Cys
CCL5	RANTES; SIS-δ	4 Cys
CCL6 (reserved)		
CCL7	MCP-3; NC28; FIC	6 Cys
CCL8	MCP-2; HC14	4 Cys
CCL9 (reserved)		
CCL10 (reserved)		
CCL11	Eotaxin	4 Cys
CCL12 (reserved)		
CCL13	MCP-4; Ckβ10; NCC-1	4 Cys
CCL14	CC-1; HCC-1; NCC-2; CCCK-1/CCCK-3; Ckβ1; MCIF	4 Cys
CCL15	HCC-2; leukotactin-1 (Lkn-1); MIP-5; CC-2; Ncc-3; MIP-1δ	6 Cys
CCL16	HCC-4; LEC; NCC-4; LMC; monotactin-1 (Mtn-1); LCC-1; ILINCK	4 Cys
CCL17	TARC; STCP-1	4 Cys
CCL18	DC-CK-1; PARC; MIP-4; AMAC-1; ckβ7	4 Cys
CCL19	MIP-3β; ELC; exodus-3; ckβ11	4 Cys
CCL20	MIP-3α; LARC; exodus-1	4 Cys
CCL21	6Ckine; SLC; exodus-2; TCA4; ckβ9	6 Cys
CCL22	MDC; STCP-1	4 Cys
CCL23	MPIF-1; MIP-3; ckβ8-1	6 Cys
CCL24	MPIF-2; eotaxin-2; ckβ6	4 Cys
CCL25	TECK, ckβ15	4 Cys
CCL26	Eotaxin-3; MIP-4α	4 Cys
CCL27	Eskine; CTACK; ILC; skinkine	4 Cys
CCL28	MEC	4 Cys
C (χ) Chemokines		
XCL1	Lymphotactin α; SCM-1α; ATAC	2 Cys
XCL2	Lymphotactin β; SCM-1β; ATAC	2 Cys
CX3C (δ) Chemokine		
CX3CL1	Fractalkine; CX3C ligand	4 Cys

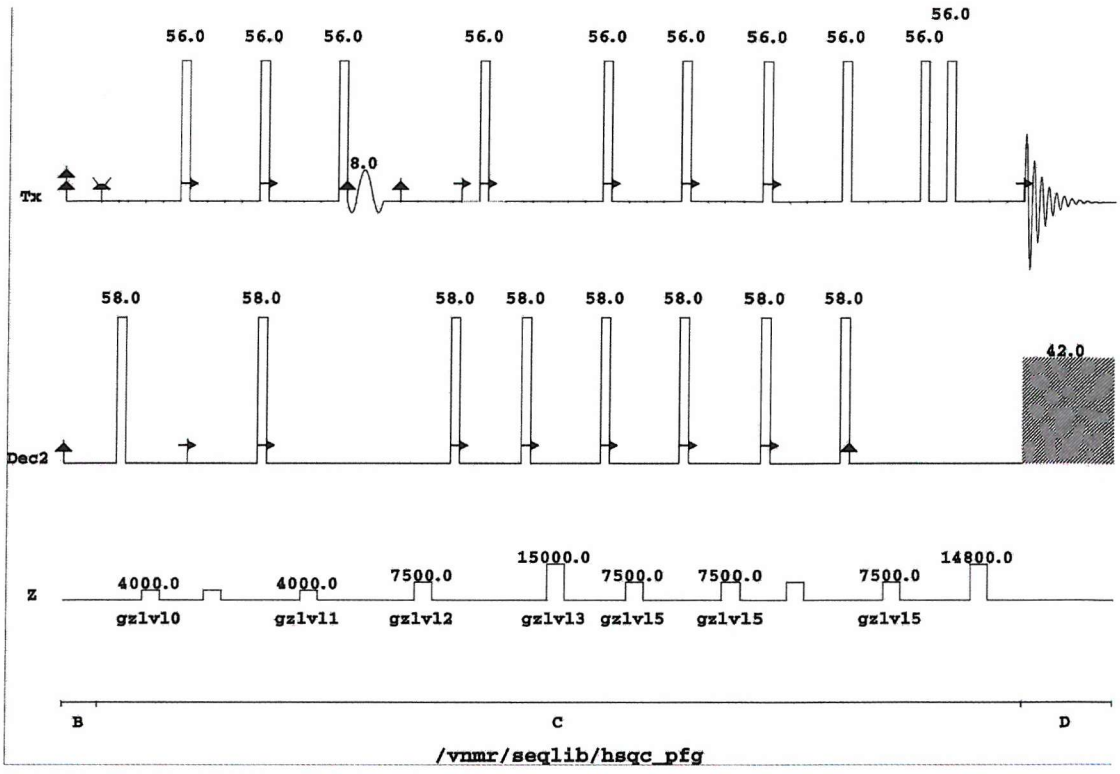
Appendix II Pulse Sequences



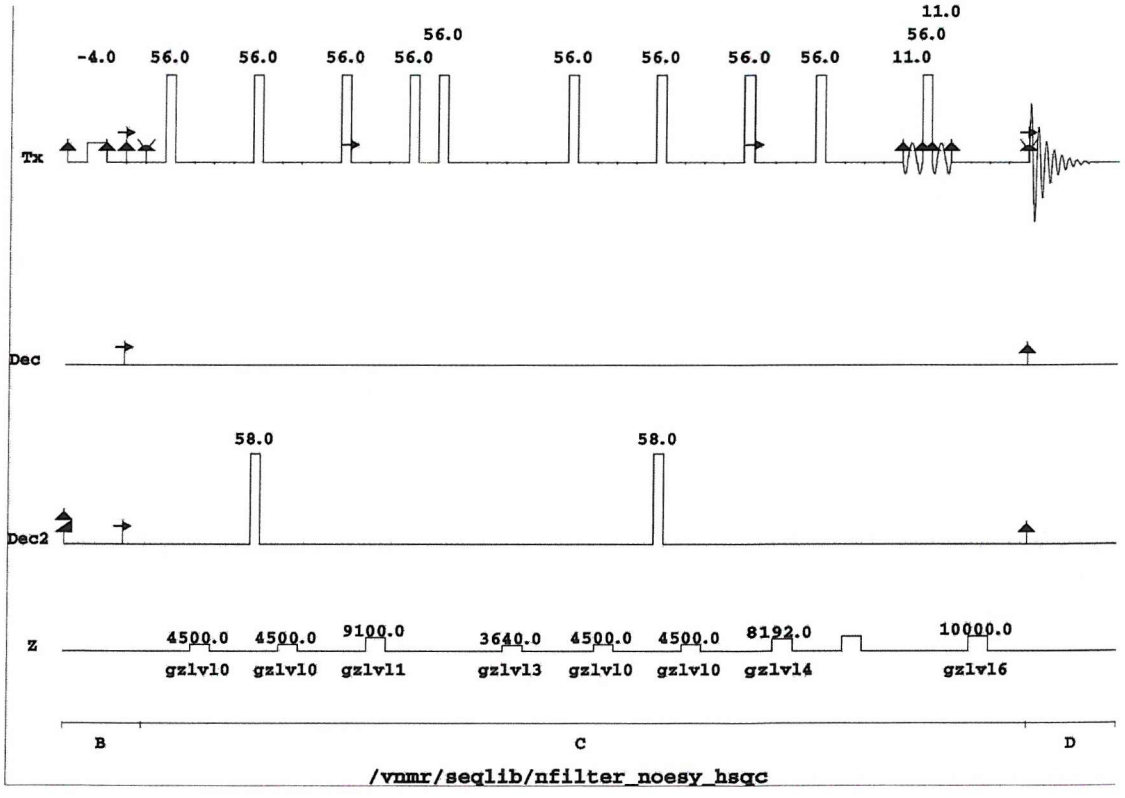
3D-NOESY HSQC



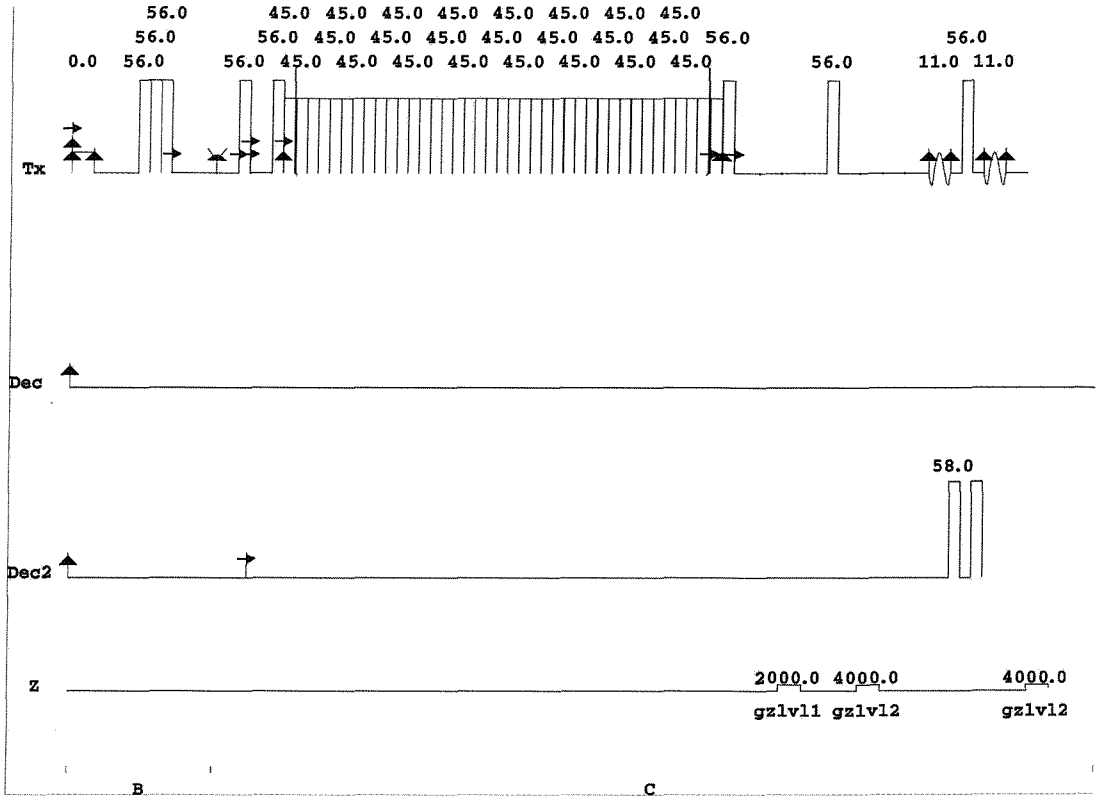
COSY



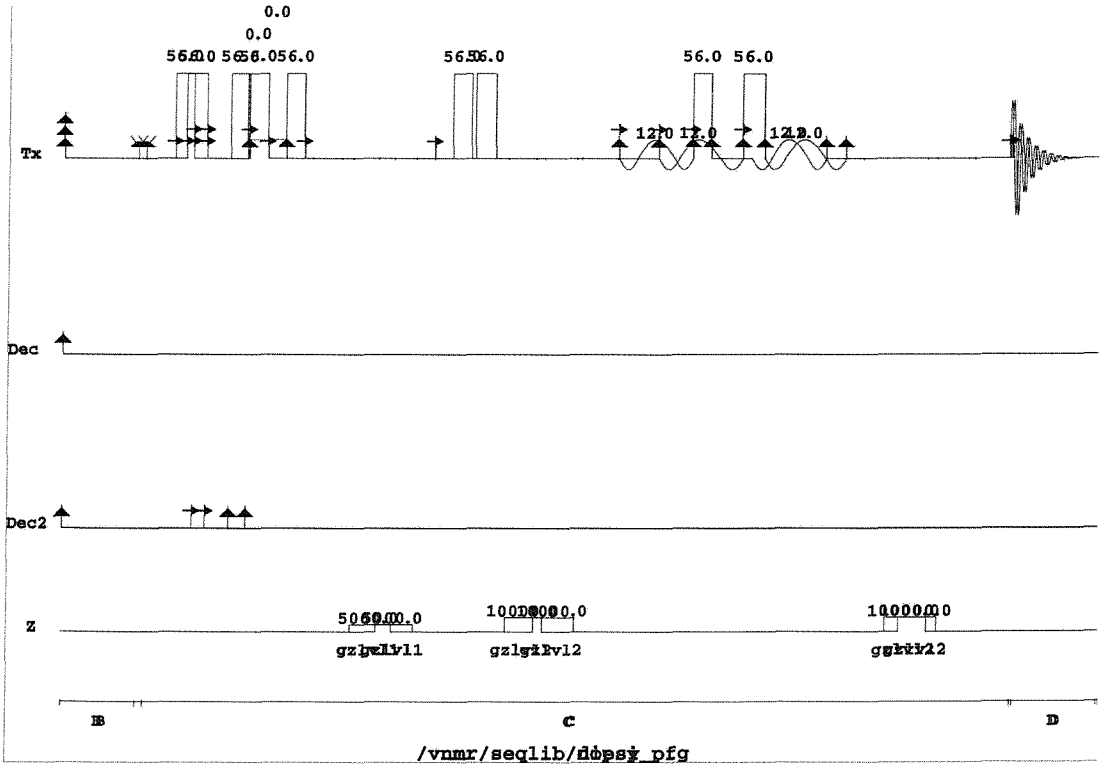
HSQC



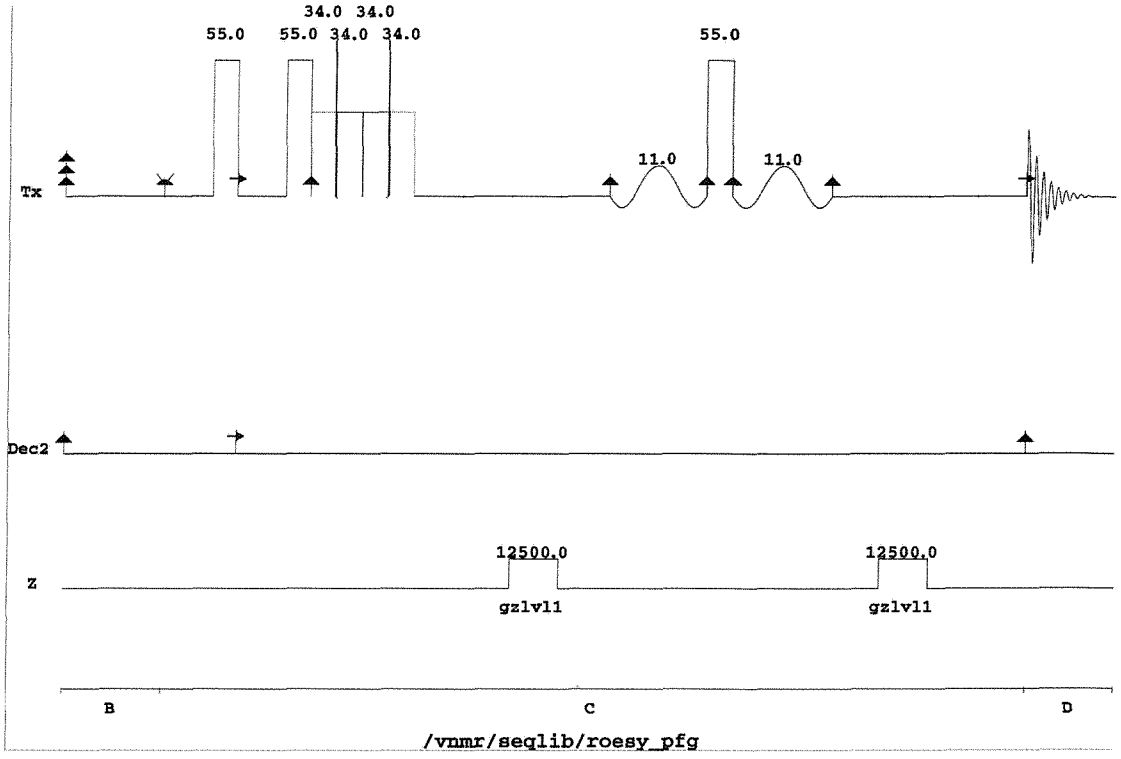
^{15}N-Filter NOESY



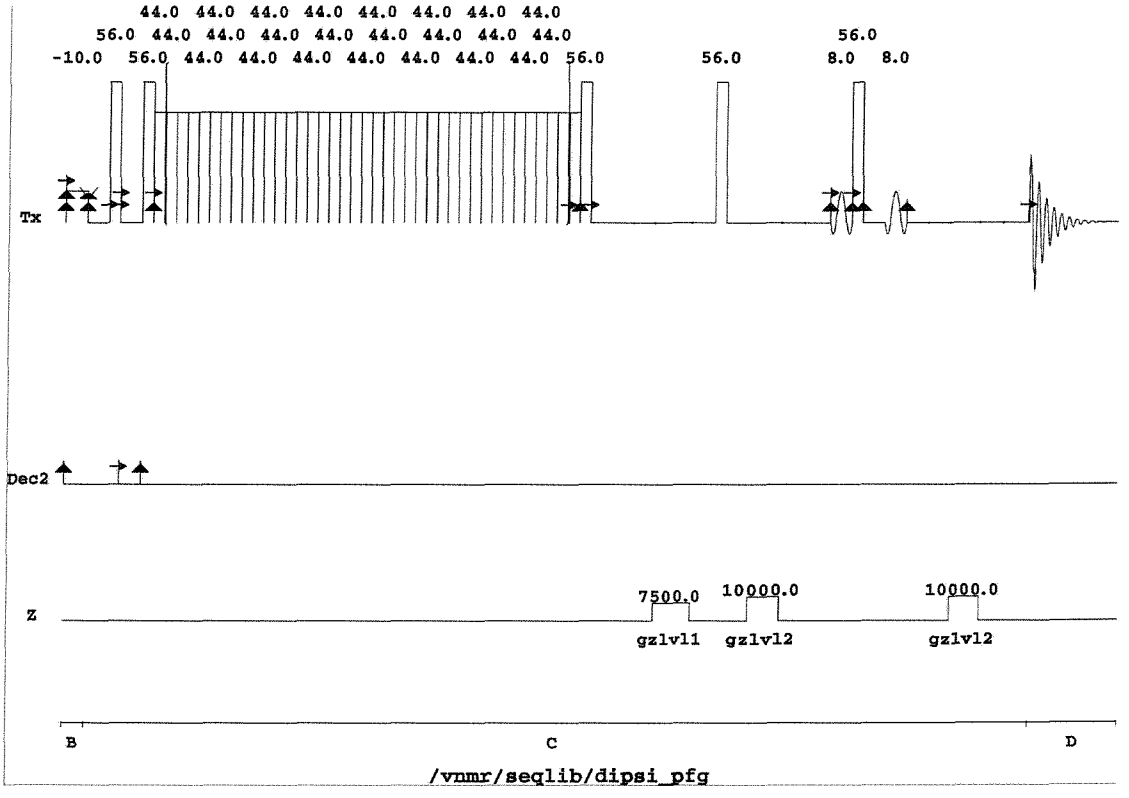
^{15}N-filtered TOCSY



NOESY


34.0 34.0
55.0
55.0 34.0 34.0
55.0
11.0
11.0
Tx
Dec2
12500.0
12500.0
Z
gzlvl1
gzlvl1
B
C
D
/vnmr/seqlib/roesy_pfg

ROESY


44.0 44.0 44.0 44.0 44.0 44.0 44.0 44.0
56.0 44.0 44.0 44.0 44.0 44.0 44.0 44.0 44.0
-10.0 56.0 44.0 44.0 44.0 44.0 44.0 44.0 44.0 56.0
56.0
56.0
8.0 8.0
Tx
Dec2
7500.0
10000.0
10000.0
Z
gzlvl1
gzlvl2
gzlvl2
B
C
D
/vnmr/seqlib/dipsi_pfg

TOCSY

Appendix III Chemical Shifts

R72A *act*-ACP ^{1}H Resonance Assignment Table.

Residue	NH	CαH	CβH	CγH	Others
A2		4.29			
T3	8.68	4.40	4.18	1.37	
L4	8.19	4.38	1.50	1.62	0.88, 0.96
L5	9.44	4.56	1.58	1.43	0.61, 0.65
T6	9.07	4.79	4.86	1.41	
T7	8.85	3.90	4.34	1.38	
D8	8.13	4.61	2.73		
D9	7.83	4.54	2.70, 3.05		
L10	8.04	4.22	1.60, 2.28	1.85	1.04, 1.11
R11	8.82	3.90	2.04	1.57	3.35, 3.47, 7.53
R12	7.99	4.10	1.94	1.73	3.32, 3.38, 7.52
A13	7.75	4.42	1.67		
L14	8.43	4.36	1.48, 2.00	1.85	0.72, 0.88
V15	8.14	4.03	2.28	1.06, 1.15	
E16	8.23	4.23	2.15, 2.24	2.63, 2.66	
C17S	7.96	4.56	4.02, 4.13		
A18	8.03	4.45	1.56		
G19	8.30	4.06, 4.13			
E20	8.43	4.51	2.09, 2.24	2.45	
T21	8.31	4.52	4.38	1.23	
D22	8.51	4.74	2.85		
G23	8.52	4.07			
T24	8.06	4.36	4.22	1.23	
D25	8.55	4.73	2.72, 2.86		
L26	8.35	4.30	1.21	1.50	0.56, 0.77
S27	8.24	4.56	3.96, 4.06		
G28	8.58	3.97, 4.20			
D29	8.65	4.86	2.71, 2.89		
F30	7.93	4.66	3.11, 3.36		7.21, 7.27, 7.34
L31	7.58	3.47	1.11, 1.67	1.33	
D32	7.88	5.07	2.40, 3.06		
L33	7.29	4.38	1.81	1.47	0.88, 0.96
R34	8.44	4.75	1.87, 1.66	1.96, 2.22	3.36, 3.46, 7.92
F35	8.48	4.37	3.10, 3.58		6.92, 6.98, 7.16
E36	9.58	4.37	2.16. 2.25	2.52	
D37	7.60	4.81	3.00, 3.10		
I38	7.74	4.61	2.41	1.49, 1.84	0.89, 1.06
G39	7.71	3.85, 4.27			
Y40	8.03	4.66	2.80, 2.98		6.79, 6.99
D41	8.01	4.78	2.90, 3.34		

S42	8.92	4.27	4.21		
L43	7.80	4.29	1.86	1.74	0.98, 1.07
A44	8.19	4.30	1.57		
L45	8.43	3.84	1.48, 1.63	1.35	0.28, 0.37
M46	8.23	4.27	2.30, 2.36	2.76, 2.86	
E47	8.55	4.24	2.21, 2.36	2.49, 2.63	
T48	8.33	3.98	4.40	1.13	
A49	8.78	3.96	1.54		
A50	8.19	4.28	1.65		
R51	8.11	4.25	2.05, 2.21	1.74, 1.97	3.32, 3.46, 7.57
L52	8.45	4.31	1.46		0.87, 1.04
E53	8.95	4.11	2.37, 2.50	2.79	
S54	8.05	4.38	4.12		
R55	8.12	4.10	1.61, 1.89	1.35, 0.98	3.02, 3.09, 7.20
Y56	8.12	4.72	2.80, 3.47		6.83, 7.44
G57	7.98	4.11			
V58	7.53	4.69	2.15	0.93, 0.98	
S59	8.50	4.84	3.77, 3.81		
I60	8.22	4.67	1.81	1.12, 1.50	0.85, 1.04
P61		4.54	3.59, 4.02	2.14, 2.55	
D62	8.85	4.47	2.76, 2.84		
D63	8.64	4.56	2.77, 2.84		
V64	7.42	3.83	2.24	0.98, 1.03	
A65	8.26	4.02	1.40		
G66	8.07	3.91			
R67	7.54	4.63	1.76, 1.87	2.09	3.24, 3.29, 7.43
V68	7.39	4.47	2.42	1.14, 1.21	
D69	9.05	5.30	2.88, 3.03		
T70	7.47	4.97			
P71					
R72A	8.99	3.66	1.64		
E73	7.88	4.26	2.28, 2.47	2.54, 2.66	
L74	7.80	4.33	1.71, 2.44	0.99	
L75	8.73	3.91	1.59, 2.17	1.71	0.80, 0.88
D76	8.70	4.50	2.73, 2.80		
L77	8.01	4.29	1.95		
I78	8.30	3.75	1.95	1.28, 1.75	0.81, 0.88
N79	8.92	4.91	2.78, 3.08		7.35, 7.71
G80	8.42	4.04			
A81	7.62	4.46	1.63		
L82	8.31	4.22	2.02, 1.85	1.58	0.88, 0.98
A83	7.99	4.29	1.61		
E84	7.76	4.41	2.15, 2.29	2.49, 2.56	
A85	7.84	4.49	1.58		
A86	7.99				

Y56A *act*-ACP ^{1}H Resonance Assignment Table.

Residue	NH	CαH	CβH	CγH	Others
A2					
T3	8.68	4.33	4.06	1.24	
L4	8.10	4.21	1.43	1.56	0.81 0.89
L5	9.39	4.65	1.89	1.37	0.81
T6	9.63	4.80		1.29	
T7	8.81	3.85	4.25	1.28	
D8	8.03	4.56	2.60		
D9	7.70	4.37	2.84, 3.07		
L10	7.83	4.10	1.72, 2.46	1.89	1.13
R11	8.73	3.53	2.82, 2.72		3.32
R12	7.70	3.94			
A13	7.70	4.24	1.67		
L14	8.29	4.29	1.67, 1.99	1.67	
V15	8.39	3.78	2.30	1.06, 1.14	
E16	8.37		2.17, 2.27	2.40	
C17S	7.72	4.48	4.05, 3.97		
A18	7.64	4.23	1.57		
G19	8.160	3.95			
E20	8.41	4.37	2.09, 2.26	2.40	
T21	8.12	4.42	4.27		
D22	8.40	4.62	2.72		
G23	8.40	3.97			
T24	8.01	4.23	4.11	1.27	
D25	8.49	4.63	2.77, 2.63		
L26	8.37	4.26	1.30	1.51	0.59 0.80
S27	8.30	4.29	3.87, 3.98		
G28	8.46	3.76, 4.21			
D29	8.52	4.74	2.55, 2.76		
F30	7.79	4.58	2.98, 3.15		7.15, 7.19
L31	7.20	3.24	1.07, 1.49	0.97	0.41, 0.65
D32	7.87	4.96	2.26, 2.95		
L33	7.13	4.29	1.70	1.35	0.97, 1.05
R34	8.37	4.63	1.68, 1.94	1.99	3.39, 3.50, 8.01
F35	8.26	4.14	3.03, 3.37		6.84, 6.88, 7.01
E36	9.34	4.24	2.01, 2.15	2.35	
D37	7.41	4.69	3.88, 2.98		
I38	7.59	4.54	2.42	1.51, 1.81	0.79, 1.04
G39	7.58	3.79, 4.16			
Y40	8.00	4.53	2.68, 3.04		6.69, 6.90
D41	8.00	4.64	2.78, 3.22		
S42	8.77	4.14	4.24		
L43	7.61	4.16	1.85	1.78	1.03, 1.09
A44	8.10	4.22	1.56		

L45	8.49	3.64	0.94, 1.61	1.17	0.20, 0.36
M46	8.00	4.21	2.28, 2.36	2.80, 2.80	
E47	8.36	4.13	2.11, 2.22	2.48	
T48	8.17	3.73	4.35	1.02	
A49	8.70	3.83	1.40		
A50	8.22	4.08	1.55		
R51					
L52	8.34	4.19	2.03, 2.17		
E53	9.11	3.82	2.17, 2.54	2.71	
S54	8.03	4.32	4.04		
R55	7.66	4.14	1.80, 2.09	1.82	3.28
Y56A	8.24	4.38	1.61		
G57	8.05	3.90, 4.05			
V58	7.57	4.65	2.03	0.94	
S59	8.23	4.80	3.64, 3.68		
I60	7.84	4.63	2.05		1.05
P61		4.56	2.58	2.23, 2.14	
D62	8.75	4.36	2.62, 2.71		
D63	8.58	4.35	2.64		
V64	7.29	3.73	2.06	0.90, 0.95	
A65	7.89	3.83	1.24		
G66	7.92	3.74, 3.94			
R67	7.22	4.51	1.64, 1.71	1.77, 1.99	
V68	7.11	4.32	2.34	1.04, 1.14	
D69	9.00	5.19	2.76, 2.92		
T70	7.37	4.93			
P71					
R72	8.82	3.51	1.62, 1.90	0.54, 0.77	
E73	7.77	4.10	2.10, 2.28	2.49	
L74	7.71	4.21	2.49		
L75	8.29	4.16	2.20		
D76	8.72	4.38	2.60, 2.69		
L77	8.00	4.16	2.01		
I78	7.92	3.60	1.82	1.13	0.77
N79	8.80	4.84	2.67, 3.00		7.79, 7.83
G80	8.30	3.92			
A81	7.50	4.35	1.51		
L82	7.93	4.11	2.06	1.67	
A83	7.83	4.19	1.50		
E84	7.77	4.27	2.10, 2.28	2.45, 2.51	
A85	7.81	4.35	1.45		
A86	7.86	4.12	1.37		

First Generation Loop 1 ^{1}H Resonance Assignment Table at 8°C.

Residue	NH	CαH	CβH	CγH	Others
I1	8.17	3.96	1.64	1.06, 1.32	0.79, 0.65
H2	8.57	4.73	3.03, 3.13		7.16
Y3	8.31	4.53	2.81, 2.90		7.01, 6.74
V4	8.26	3.98	1.9	0.85	
R5	8.49	4.15	1.72	1.47, 1.55	2.99, 7.09, 6.78, 7.07
G6	8.49	3.86, 3.92			
H7	8.44	4.52	2.97		7.04, 8.43
N8	8.55	4.64	2.67, 2.76		6.93, 7.61
W9	8.11	4.58	3.17		10.07, 7.08, 7.38, 7.17, 7.09, 7.53
V10	7.92	3.93	1.81	0.74	
F11	8.22	4.38	2.97, 3.03		7.24, 7.35, 7.28
G12	8.26	3.69, 3.88			
H13	8.36	4.59	3.04, 3.23		7.16
G14	8.52	3.82, 3.90			
M15	8.31	4.42	1.94, 2.09	2.46, 2.57	

First Generation Loop 1 ^{1}H Resonance Assignment Table at 30°C.

Residue	NH	CαH	CβH	CγH	Others
I1	8.02	3.99	1.66	1.08, 1.32	0.79, 0.67
H2	8.42	4.71	3.04, 3.14		7.16, 8.49
Y3	8.15	4.56	2.83, 2.94		7.02, 6.75
V4	8.11	4.01	1.93	0.85	
R5	8.33	4.19	1.72, 1.78	1.51, 1.58	3.07, 7.08
G6	8.33	3.90			
H7	8.31	4.53	2.98		7.06, 8.43
N8	8.41	4.63	2.67, 2.74		
W9	8.02	4.62	3.17		10.01, 7.11, 7.39, 7.17, 7.09, 7.55
V10	7.80	3.95	1.84	0.72	
F11	8.04	4.46	3.05		7.23, 7.34, 7.27
G12	8.15	3.75, 3.88			
H13	8.28	4.62	3.08, 3.25		7.19, 8.46
G14	8.41	3.89			
M15	8.19	4.43	1.96, 2.08	2.48, 2.57	

N-Terminus ^{1}H Resonance Assignment Table at 8°C.

Residue	NH	CαH	CβH	CγH	Others
M1	8.05	4.19	2.45, 2.34	1.91, 1.84	
T2	8.61	4.30	4.00	1.05	
T3	8.30	4.22	4.05	1.01	
S4	8.33	4.26	3.68		
L5	8.28	4.17	1.24		0.71, 0.66
D6	8.14	4.42	2.51, 2.42		
T7	7.90	4.14	4.03	0.98	
V8	8.03	3.89	1.87	0.72	
E9	8.38	4.10	2.02, 1.96	1.73, 1.69	
T10	8.02	4.06	3.91	0.90	
F11	8.26	4.38	2.96, 2.82		7.12, 7.03, 7.06
G12	8.20	3.74			
T13	7.95	4.25	4.07	1.00	
T14	8.14	4.16	3.96	0.94	
S15	8.20	4.24	3.55		
Y16	8.05	4.26	2.67		6.75, 6.56
Y17	7.87	4.29	2.63, 2.79		6.87, 6.60
D18	7.99	4.33	2.45		
D19	8.03	4.35	2.53, 2.46		
V20	7.94	3.87	1.99	0.73	
G21	8.29	3.71			
L22	7.93	4.11	1.46	1.41	0.71, 0.66
L23	8.10	4.12	1.49	1.38	0.71, 0.66
S24	8.05	4.20	3.65		
E25	8.20	4.09	2.13, 2.06	1.88, 1.78	
K26	8.13	4.08	1.66, 1.56	1.44, 1.25	2.78, 7.39
A27	8.17	4.05	1.19		
D28	8.27	4.44	2.58, 2.49		
T29	8.02	4.12	4.02	1.02	
R30	8.11	4.00	1.66	1.47	3.00, 7.26
A31	8.03	4.06	1.21		
L32	7.91	4.07	1.48	1.39	0.66, 0.71
M33	8.08	4.05	2.27, 1.95	1.80	
A34	8.18	4.08	1.21		
Q35	7.82	4.07			

N-Terminus ^{1}H Resonance Assignment Table at 30°C.

Residue	NH	CαH	CβH	CγH	Others
M1	8.10	4.34	2.56, 2.47	2.03, 1.95	
T2	8.67	4.44	4.14	1.17	
T3	8.35	4.35	4.18	1.14	
S4	8.37	4.40	3.81		
L5	8.32	4.32	1.56		0.84, 0.79
D6	8.22	4.57	2.65, 2.56		
T7	7.96	4.27	4.15	1.11	
V8	8.06	4.03	2.00	0.85	
E9	8.43	4.25	2.16, 2.09	1.87, 1.81	
T10	8.05	4.20	4.04	1.03	
F11	8.28	4.54	3.09, 2.96		7.16, 7.20, 7.25
G12	8.26	3.88			
T13	8.01	4.37	4.19	1.13	
T14	8.19	4.29	4.10	1.07	
S15	8.25	4.37	3.67		
Y16	8.09	4.40	2.81		6.89, 6.69
Y17	7.92	4.43	2.95, 2.76		7.00, 6.75
D18	8.05	4.47	2.58		
D19	8.11	4.50	2.65, 2.59		
V20	8.01	4.01	2.11	0.85	
G21	8.38	3.85			
L22	7.97	4.24	1.59	1.51	0.79, 0.84
L23	8.15	4.26	1.54	1.62	0.79, 0.84
S24	8.09	4.33	3.80		
E25	8.27	4.24	2.26, 2.20	2.00, 1.89	
K26	8.19	4.22	1.80, 1.69	1.62, 1.37	2.91
A27	8.21	4.19	1.32		
D28	8.32	4.58	2.71, 2.61		
T29	8.07	4.24	4.16	1.14	
R30	8.19	4.12	1.78	1.60	3.13, 7.35
A31	8.09	4.16	1.37		
L32	7.96	4.19	1.59	1.52	0.84, .079
M33	8.14	4.19	2.32, 2.08	1.93	
A34	8.24	4.20	1.33		
Q35	7.88	4.21			

Redesigned Loop 1 ^{1}H Resonance Assignment Table in DMSO.

Residue	NH	CαH	CβH	CγH	Others
L1	8.04	4.20	1.70	2.10	0.73
P2		4.34	1.78, 1.90	3.36	3.59
F3	7.93	4.42	2.83, 3.00		7.18
W4	8.07	4.62	3.00, 3.15		6.96, 7.05, 7.13, 7.34, 7.59, 10.90
I5	7.98	4.17	1.36, 1.69	0.73, 1.05	0.79
H6	8.12	4.49	2.91, 2.83		6.78, 7.53
Y7	7.97	4.51	2.74, 2.94		6.65, 7.00
V8	8.07	4.21	2.00	0.84	
R9	8.12	4.31	1.59, 1.72	1.52	3.10
G10	8.27	3.80, 3.69			
H11	7.99	4.49	2.87, 2.91		6.81, 7.53
N12	8.37	4.59	2.42, 2.58		6.99, 7.49
W13	8.34	4.57	3.02, 3.15		6.95, 7.05, 7.14, 7.33, 7.55, 10.90
V14	7.75	4.10	1.91	0.70	
F15	8.06	4.58	2.82, 3.06		7.24, 7.26, 7.27
G16	8.27	3.77			
H17	8.06	4.45	2.91, 2.97		6.87, 7.56
G18	8.40	3.74			
M19	8.47	4.41	1.85, 1.94	2.43, 2.50	
A20	8.14	4.27	1.23		
K21	7.96	4.25	1.68	1.32	1.53, 2.74
L22	8.08	4.29	1.48	1.59	0.83, 0.87
L23	8.01	4.37	1.50	1.61	0.83, 0.87
S24	8.24	4.28	3.60, 3.67		
G25	8.36	3.58, 3.72			
F26	8.00	4.46	2.74, 2.97		7.17, 7.18, 7.23
Y27	8.29	4.40	2.78, 2.96		6.66, 7.05
H28	8.41	4.55	2.98		6.87, 7.50
T29	7.64	4.11		0.96	
G30	7.63	3.62, 3.43			

Redesigned Loop 1 ^{1}H Resonance Assignment Table in DPC.

Residue	NH	CαH	CβH	CγH	Others
L1		4.14			
P2		4.50	2.17	1.94	3.02
F3	9.00	4.36	3.35, 3.27		7.41, 7.50
W4	8.43	4.13	2.71, 3.42		6.81, 6.99, 7.18, 7.38, 10.96
I5	8.06	3.73		0.69, 0.87, 1.42	0.58
H6	7.78	4.25	3.06		7.54
Y7	7.79	4.41	3.02, 3.29		6.91, 7.20
V8	7.67	4.03	2.36	1.08, 1.19	
R9	7.80	4.24	1.92	1.65, 1.80	3.15
G10	8.12	3.81, 3.99			
H11	7.87	4.51	3.12		7.26
N12	8.39	4.57	2.79		6.88, 7.67
W13	8.12	4.01	3.01		6.97, 7.02, 7.40, 7.51, 10.96
V14	7.92	3.82	1.87	0.62	
F15	8.15	4.57	3.03, 3.21		7.11, 7.22, 7.27
G16	8.37	3.95			
H17	8.58	4.51	3.18, 3.30		
G18	8.77	3.88, 3.97			
M19	8.50	4.27	2.01, 2.16	2.54	
A20	8.41	4.00	1.49		
K21	7.94	4.06	1.89	1.48	1.71, 2.97
L22	7.84	4.15	1.88	1.76	0.94, 0.98
L23	8.12	4.23	1.89	1.67	0.89
S24	8.03	4.22	4.03		
G25	7.87	3.82, 3.98			
F26	8.40	4.58	2.88		7.27
Y27	7.88	4.47	2.88, 3.03		6.77, 7.00
H28	8.13	4.69	3.17, 3.30		
T29	8.13	4.32	4.24	1.19	
G30	8.06				

Redesigned Loop 3 ^{1}H Resonance Assignment Table in DMSO.

Residue	NH	CαH	CβH	CγH	Others
A1	8.05	3.91	1.31		
I2	8.42	4.27	1.73	1.09, 1.49	0.86
L3	8.15	4.37	1.45	1.59	0.83, 0.87
L4	7.96	4.37	1.46	1.63	0.82, 0.87
S5	7.94	4.36	3.54, 3.62		
S6	8.03	4.25	3.53, 3.55		
Y7	7.93	4.41	2.69, 2.96		6.65, 7.02
Q8	8.06	4.29	1.79, 1.91	2.14	6.82, 7.29
S9	7.98	4.37	3.56, 3.63		
I10	7.79	4.19	1.74	1.09, 1.40	0.79
L11	7.93	4.25	1.35	1.39	0.80, 0.86
F12	7.89	4.54	2.85, 3.05		7.26
G13	8.18	3.72, 3.79			
N14	8.15	4.58	2.51, 2.61		7.04, 7.52
D15	8.40	4.54	2.58, 2.74		
A16	7.98	4.21	1.27		
E17	7.81	4.23	1.80, 1.96	2.28	
R18	7.89	4.32	1.72	1.53	3.12
S19	7.94	4.32	3.56, 3.63		
K20	8.08	4.23	1.66	1.32	1.52, 2.76
H21	8.15	4.57	2.94, 3.07		
L22	8.04	4.31	1.45	1.59	0.86
D23	8.51	4.60	2.74		
L24	7.76	4.36	1.45	1.58	0.86
V25	7.82	4.16	1.99	0.89	
M26	8.19	4.32	1.87, 1.99	2.44, 2.50	

Redesigned Loop 3 ^{1}H Resonance Assignment Table in DPC.

Residue	NH	CαH	CβH	CγH	Others
A1		4.20	1.53		
I2	7.81	4.27	2.07	1.94	
L3	8.59	4.51	1.66	1.73	0.92, 0.96
L4	8.62	4.26	1.67	1.74	
S5	8.40	4.26	3.95		
S6	8.13	4.39	3.82		
Y7	7.95	4.43	2.96, 3.18		6.80, 7.08
Q8	8.04	3.94	2.21	2.38	6.85, 7.40
S9	8.12	4.27	3.93		
I10	7.56	3.94	1.93	1.26	0.92
L11	7.46	4.11	1.48		0.72, 0.80
F12	7.82	4.27	2.99, 3.40		7.17, 7.28, 7.39
G13	8.01	4.06, 4.13			
N14	8.47	4.76	2.85		6.94, 7.64
D15	8.47	4.53	2.67, 2.73		
A16	8.29	4.21	1.43		
E17	8.25	4.20	2.03, 2.06	2.28, 2.30	
R18	8.12	4.23	1.87	1.67	3.19
S19	8.22	4.30	3.83, 3.92		
K20	8.04	4.27	1.73	1.38	1.64, 2.94
H21	8.18	4.67	3.13, 3.26		7.26
L22	8.16	4.25	1.64	1.68	
D23	8.24	4.57	2.65, 2.71		
L24	7.93	4.33	1.61	1.72	0.88, 0.93
V25	7.90	4.20	2.17	0.94	
M26	7.86	4.32	1.95, 2.09	2.51	

CCL24 HSQC Backbone Assignment.

Residue	NH	N
S-1	7.95	114.96
V1	8.30	125.86
V2	8.30	125.60
I3	8.41	128.28
P4		
S5	8.42	117.86
P6		
C7		
C8	9.41	120.81
M9	8.96	125.18
F10	7.35	114.76
F11	8.80	122.59
V12	9.01	119.51
S13	8.84	118.85
K14	7.54	121.56
R15	7.90	121.57
I16	5.37	124.26
P17		
E18	8.51	127.29
N19	8.34	112.99
R20	7.79	116.71
V21	7.63	117.81
V22	9.04	118.44
S23	7.92	113.20
Y24	8.49	116.87
Q25	8.70	118.38
L26	9.08	124.50
S27	8.48	117.90
S28	8.51	119.42
R29	7.99	124.79
S30	8.56	118.44
T31	7.40	109.97
C32	7.74	120.39
L33	8.35	122.62
K34		
A35	8.41	125.80
G36	8.21	105.36
V37	8.44	122.70
I38	9.06	126.40
F39	9.52	129.90
T40	8.63	116.84
T41	9.15	116.86
K42	8.80	122.59
K43	7.95	115.70
G44	8.14	107.97
Q45	7.25	118.85
Q46	8.45	119.36
S47	8.93	118.39
C48	9.40	121.04
G49	10.05	112.14
D50	8.30	123.47
P51		
K52	8.08	116.38
Q53	7.85	118.56
E54	8.84	125.49
W55	8.17	116.23
V56	5.58	122.60
Q57	7.01	116.84
R58	8.51	121.01
Y59	8.96	120.50
M60	7.93	116.96
K61	7.77	117.95
N62	8.50	119.37
L63	8.43	123.16
D64	8.85	121.56
A65	7.86	121.49
K66	7.67	117.46
Q67	7.89	118.52
K68	8.06	121.14
K69	8.11	121.67
A70	8.32	125.73
S71	8.22	115.98
P72		

CCL24 HSQC Side Chain Assignment.

Residue	NH	NH	N
N19	6.73	7.60	111.62
Q25	6.88	7.49	111.65
Q45	6.95	7.67	113.12
Q46	6.63	7.26	111.56
Q53	7.11	7.63	113.15
W55	10.15		129.90
Q57	6.87	7.4	111.32
N62	6.59	9.20	113.72
Q67	6.90	7.55	112.41

Appendix IV NOE Tables

```
! Loop 1 in DMSO NOE Table
! LPFWIHYVRG HNWVFGHGMA KLLSGFYHTG
clas long
assign (resid 15 and name HD# )(resid 20 and name HB# ) 4.5 2.7 2.5

clas aveshort
assign (resid 2 and name HB# )(resid 3 and name HD# ) 4.5 2.7 2.5
assign (resid 6 and name HN )(resid 7 and name HD# ) 4.5 2.7 2.5
assign (resid 6 and name HN )(resid 7 and name HD# ) 4.5 2.7 2.5
assign (resid 6 and name HE# )(resid 7 and name HE# ) 4.5 2.7 2.5
assign (resid 7 and name HD# )(resid 8 and name HA ) 4.5 2.7 2.5
assign (resid 7 and name HD# )(resid 8 and name HB# ) 4.5 2.7 2.5
assign (resid 7 and name HE# )(resid 8 and name HA ) 4.5 2.7 2.5
assign (resid 7 and name HD# )(resid 9 and name HB1 ) 4.5 2.7 2.5
assign (resid 7 and name HD# )(resid 9 and name HB2 ) 4.5 2.7 2.5
assign (resid 7 and name HD# )(resid 9 and name HG# ) 4.5 2.7 2.5
assign (resid 7 and name HE# )(resid 9 and name HA ) 4.5 2.7 2.5
assign (resid 7 and name HE# )(resid 9 and name HB1 ) 4.5 2.7 2.5
assign (resid 7 and name HE# )(resid 9 and name HB2 ) 4.5 2.7 2.5
assign (resid 7 and name HE# )(resid 9 and name HG# ) 4.5 2.7 2.5
assign (resid 14 and name HA )(resid 15 and name HD# ) 4.5 2.7 2.5
assign (resid 14 and name HB# )(resid 15 and name HD# ) 3.5 1.7 2.5
assign (resid 15 and name HD# )(resid 16 and name HA# ) 4.5 2.7 2.5
assign (resid 23 and name HG# )(resid 26 and name HD# ) 4.5 2.7 2.5
assign (resid 23 and name HB# )(resid 26 and name HD# ) 4.5 2.7 2.5
assign (resid 24 and name HB1 )(resid 26 and name HD# ) 4.5 2.7 2.5
assign (resid 24 and name HB2 )(resid 26 and name HD# ) 4.5 2.7 2.5
assign (resid 25 and name HA# )(resid 27 and name HD# ) 4.5 2.7 2.5
assign (resid 25 and name HA# )(resid 27 and name HE# ) 4.5 2.7 2.5
assign (resid 26 and name HN )(resid 27 and name HD# ) 4.5 2.7 2.5
assign (resid 26 and name HD# )(resid 27 and name HN ) 4.5 2.7 2.5
assign (resid 27 and name HN )(resid 28 and name HD# ) 4.5 2.7 2.5
assign (resid 27 and name HD# )(resid 28 and name HN ) 3.5 1.7 2.5
assign (resid 27 and name HD# )(resid 29 and name HA ) 4.5 2.7 2.5
assign (resid 27 and name HD# )(resid 29 and name HG# ) 4.5 2.7 2.5
assign (resid 27 and name HE# )(resid 29 and name HG# ) 4.5 2.7 2.5

clas seq
!HN(i)-HN(i+1)

assign (resid 14 and name HN )(resid 15 and name HN ) 4.5 2.7 0.5
assign (resid 15 and name HN )(resid 16 and name HN ) 3.0 1.2 0.5
assign (resid 16 and name HN )(resid 17 and name HN ) 3.0 1.2 0.5
assign (resid 17 and name HN )(resid 18 and name HN ) 3.0 1.2 0.5
assign (resid 18 and name HN )(resid 19 and name HN ) 3.0 1.2 0.5
assign (resid 19 and name HN )(resid 20 and name HN ) 3.5 1.7 0.5
assign (resid 20 and name HN )(resid 21 and name HN ) 2.3 0.5 0.5
assign (resid 21 and name HN )(resid 22 and name HN ) 2.3 0.5 0.5
assign (resid 22 and name HN )(resid 23 and name HN ) 3.0 1.2 0.5
assign (resid 23 and name HN )(resid 24 and name HN ) 3.0 1.2 0.5
assign (resid 24 and name HN )(resid 25 and name HN ) 3.0 1.2 0.5
assign (resid 25 and name HN )(resid 26 and name HN ) 3.0 1.2 0.5
assign (resid 26 and name HN )(resid 27 and name HN ) 2.3 0.5 0.5
assign (resid 27 and name HN )(resid 28 and name HN ) 2.3 0.5 0.5
assign (resid 28 and name HN )(resid 29 and name HN ) 3.0 1.2 0.5

!HA(i)-HN(i+1)

assign (resid 2 and name HA )(resid 3 and name HN ) 2.3 0.5 0.5
assign (resid 3 and name HA )(resid 4 and name HN ) 2.3 0.5 0.5
assign (resid 4 and name HA )(resid 5 and name HN ) 2.3 0.5 0.5
assign (resid 5 and name HA )(resid 6 and name HN ) 2.3 0.5 0.5
assign (resid 6 and name HA )(resid 7 and name HN ) 3 1.2 0.5
assign (resid 7 and name HA )(resid 8 and name HN ) 2.3 0.5 0.5
assign (resid 8 and name HA )(resid 9 and name HN ) 2.3 0.5 0.5
assign (resid 9 and name HA )(resid 10 and name HN ) 2.3 0.5 0.5
assign (resid 10 and name HA1 )(resid 11 and name HN ) 3 1.2 0.5
assign (resid 10 and name HA2 )(resid 11 and name HN ) 3 1.2 0.5
assign (resid 11 and name HA )(resid 12 and name HN ) 3 1.2 0.5
assign (resid 12 and name HA )(resid 13 and name HN ) 3.5 1.7 0.5
assign (resid 13 and name HA )(resid 14 and name HN ) 2.3 0.5 0.5
assign (resid 14 and name HA )(resid 15 and name HN ) 2.3 0.5 0.5
assign (resid 15 and name HA )(resid 16 and name HN ) 3 1.2 0.5
assign (resid 16 and name HA# )(resid 17 and name HN ) 2.3 0.5 1.5
assign (resid 17 and name HA )(resid 18 and name HN ) 2.3 0.5 0.5
assign (resid 18 and name HA# )(resid 19 and name HN ) 3.5 1.7 1.5
assign (resid 19 and name HA )(resid 20 and name HN ) 3.5 1.7 0.5
assign (resid 20 and name HA )(resid 21 and name HN ) 2.3 0.5 0.5
assign (resid 21 and name HA )(resid 22 and name HN ) 3.0 1.2 0.5
assign (resid 22 and name HA )(resid 23 and name HN ) 2.3 0.5 0.5
assign (resid 23 and name HA )(resid 24 and name HN ) 3.0 1.2 0.5
assign (resid 24 and name HA )(resid 25 and name HN ) 2.3 0.5 0.5
assign (resid 25 and name HA# )(resid 26 and name HN ) 2.3 0.5 1.5
assign (resid 25 and name HA# )(resid 26 and name HN ) 2.3 0.5 1.5
assign (resid 26 and name HA )(resid 27 and name HN ) 2.3 0.5 0.5
assign (resid 28 and name HA )(resid 29 and name HN ) 3.0 1.2 0.5

!HB(i)-HN(i+1)

assign (resid 2 and name HB# )(resid 3 and name HN ) 4.5 2.7 1.5
assign (resid 4 and name HB1 )(resid 5 and name HN ) 3.0 1.2 1.5
assign (resid 4 and name HB2 )(resid 5 and name HN ) 3.0 1.2 1.5
assign (resid 5 and name HB# )(resid 6 and name HN ) 3.0 1.2 1.5
assign (resid 6 and name HB1 )(resid 7 and name HN ) 3 1.2 1.5
assign (resid 6 and name HB2 )(resid 7 and name HN ) 2.3 0.5 1.5
assign (resid 7 and name HB1 )(resid 8 and name HN ) 2.3 0.5 1.5
assign (resid 7 and name HB2 )(resid 8 and name HN ) 2.3 0.5 1.5
assign (resid 8 and name HB# )(resid 9 and name HN ) 3.0 1.2 1.5
assign (resid 9 and name HB# )(resid 10 and name HN ) 3.5 1.7 1.5
assign (resid 11 and name HB1 )(resid 12 and name HN ) 3 1.2 1.5
assign (resid 11 and name HB2 )(resid 12 and name HN ) 3 1.2 1.5
assign (resid 13 and name HB1 )(resid 14 and name HN ) 3.5 1.7 1.5
assign (resid 13 and name HB2 )(resid 14 and name HN ) 3.5 1.7 1.5
assign (resid 14 and name HB# )(resid 15 and name HN ) 3.0 1.2 1.5
assign (resid 15 and name HB1 )(resid 16 and name HN ) 3.0 1.2 1.5
assign (resid 15 and name HB2 )(resid 16 and name HN ) 3.0 1.2 1.5
assign (resid 17 and name HB# )(resid 18 and name HN ) 3.0 1.2 1.5
assign (resid 19 and name HB1 )(resid 20 and name HN ) 4.5 2.7 1.5
assign (resid 19 and name HB2 )(resid 20 and name HN ) 4.5 2.7 1.5
assign (resid 20 and name HB# )(resid 21 and name HN ) 3.0 1.2 1.5
assign (resid 21 and name HB# )(resid 22 and name HN ) 3.0 1.2 1.5
assign (resid 22 and name HB# )(resid 23 and name HN ) 3 1.2 1.5
assign (resid 23 and name HB# )(resid 24 and name HN ) 3.0 1.2 1.5
assign (resid 24 and name HB# )(resid 25 and name HN ) 3.0 1.2 1.5
assign (resid 26 and name HB# )(resid 27 and name HN ) 3.0 1.2 1.5

!HG(i)-HN(i+1)

assign (resid 5 and name HG# )(resid 6 and name HN ) 3.0 1.2 2.5
assign (resid 9 and name HG# )(resid 10 and name HN ) 3.0 1.2 2.5
assign (resid 14 and name HG# )(resid 15 and name HN ) 2.3 0.5 2.5
assign (resid 21 and name HG# )(resid 22 and name HN ) 3.0 1.2 2.5
assign (resid 22 and name HG# )(resid 23 and name HN ) 3.0 1.2 2.5
assign (resid 29 and name HG# )(resid 30 and name HN ) 3.5 1.7 2.5

!HD(i)-HN(i+1)

assign (resid 6 and name HD# )(resid 7 and name HN ) 3.5 1.7 2.5
```

assign (resid 11 and name HD#)(resid 12 and name HN) 4.5 2.7 2.5
assign (resid 21 and name HD#)(resid 22 and name HN) 3.5 1.7 2.5
assign (resid 22 and name HD#)(resid 23 and name HN) 3.5 1.7 2.5
assign (resid 22 and name HD#)(resid 23 and name HN) 3.5 1.7 2.5
assign (resid 23 and name HD#)(resid 24 and name HN) 3.5 1.7 2.5
assign (resid 28 and name HD#)(resid 29 and name HN) 4.5 2.7 2.5

!HN(i)-OTHER(i+1)

assign (resid 10 and name HN)(resid 11 and name HD#) 4.5 2.7 2.5
assign (resid 18 and name HN)(resid 19 and name HB1) 4.5 2.7 1.5
assign (resid 18 and name HN)(resid 19 and name HB2) 4.5 2.7 1.5
assign (resid 27 and name HN)(resid 28 and name HD#) 4.5 2.7 2.5

!HA(i)-OTHER(i+1)

assign (resid 1 and name HA)(resid 2 and name HB#) 4.5 2.7 2.5
assign (resid 4 and name HA)(resid 5 and name HG2#) 4.5 2.7 2.5
assign (resid 4 and name HA)(resid 5 and name HG12) 3.5 1.7 2.5
assign (resid 4 and name HA)(resid 5 and name HG1#) 3.5 1.7 2.5
assign (resid 4 and name HA)(resid 5 and name HB#) 4.5 2.7 1.5
assign (resid 5 and name HA)(resid 6 and name HD#) 4.5 2.7 2.5
assign (resid 7 and name HA)(resid 8 and name HB#) 4.5 2.7 1.5
assign (resid 7 and name HA)(resid 8 and name HG#) 3.5 1.7 2.5
assign (resid 9 and name HA)(resid 10 and name HA#) 4.5 2.7 1.5
assign (resid 13 and name HA)(resid 14 and name HA) 4.5 2.7 0.5
assign (resid 13 and name HA)(resid 14 and name HB#) 4.5 2.7 1.5
assign (resid 13 and name HA)(resid 14 and name HG#) 4.5 2.7 2.5
assign (resid 14 and name HA)(resid 15 and name HZ) 4.5 2.7 2.5
assign (resid 15 and name HA)(resid 16 and name HA#) 4.5 2.7 1.5
assign (resid 16 and name HA#)(resid 17 and name HD#) 4.5 2.7 2.5
assign (resid 16 and name HA#)(resid 17 and name HE#) 4.5 2.7 2.5
assign (resid 17 and name HA)(resid 18 and name HA#) 3.5 1.7 1.5
assign (resid 18 and name HA#)(resid 19 and name HB1) 4.5 2.7 2.5
assign (resid 18 and name HA#)(resid 19 and name HB2) 4.5 2.7 2.5
assign (resid 19 and name HA)(resid 20 and name HB#) 4.5 2.7 1.5
assign (resid 23 and name HA)(resid 24 and name HB1) 4.5 2.7 1.5
assign (resid 23 and name HA)(resid 24 and name HB2) 4.5 2.7 1.5
assign (resid 28 and name HA)(resid 29 and name HG#) 4.5 2.7 2.5
assign (resid 29 and name HA)(resid 30 and name HA#) 4.5 2.7 1.5

!HB(i)-OTHER(i+1)

assign (resid 7 and name HB1)(resid 8 and name HG#) 4.5 2.7 2.5
assign (resid 7 and name HB2)(resid 8 and name HG#) 4.5 2.7 2.5
assign (resid 8 and name HB#)(resid 9 and name HA) 4.5 2.7 1.5
assign (resid 9 and name HB1)(resid 10 and name HA1) 4.5 2.7 1.5
assign (resid 9 and name HB1)(resid 10 and name HA2) 4.5 2.7 1.5
assign (resid 9 and name HB2)(resid 10 and name HA1) 4.5 2.7 2.5
assign (resid 9 and name HB2)(resid 10 and name HA2) 4.5 2.7 2.5
assign (resid 12 and name HB1)(resid 13 and name HD#) 4.5 2.7 2.5
assign (resid 12 and name HB2)(resid 13 and name HD#) 4.5 2.7 2.5
assign (resid 13 and name HB1)(resid 14 and name HG#) 4.5 2.7 2.5
assign (resid 13 and name HB2)(resid 14 and name HG#) 4.5 2.7 2.5
assign (resid 19 and name HB1)(resid 20 and name HA) 4.5 2.7 1.5
assign (resid 19 and name HB2)(resid 20 and name HA) 4.5 2.7 1.5
assign (resid 19 and name HB1)(resid 20 and name HB#) 4.5 2.7 1.5
assign (resid 19 and name HB2)(resid 20 and name HB#) 4.5 2.7 1.5
assign (resid 23 and name HB#)(resid 24 and name HB#) 4.5 2.7 1.5

!HG(i)-OTHER(i+1)

assign (resid 1 and name HG#)(resid 2 and name HB#) 4.5 2.7 2.5
assign (resid 5 and name HG2#)(resid 6 and name HD#) 4.5 2.7 2.5
assign (resid 5 and name HG2#)(resid 6 and name HE#) 4.5 2.7 2.5
assign (resid 5 and name HG2#)(resid 6 and name HB#) 4.5 2.7 2.5
assign (resid 9 and name HG#)(resid 10 and name HA#) 4.5 2.7 2.5
assign (resid 14 and name HG#)(resid 15 and name HD#) 4.5 2.7 2.5
assign (resid 14 and name HG#)(resid 15 and name HB1) 4.5 2.7 2.5
assign (resid 14 and name HG#)(resid 15 and name HB2) 4.5 2.7 2.5
assign (resid 19 and name HG1)(resid 20 and name HB#) 4.5 2.7 2.5
assign (resid 19 and name HG2)(resid 20 and name HB#) 4.5 2.7 2.5
assign (resid 21 and name HE#)(resid 22 and name HN) 3.5 1.7 2.5
assign (resid 23 and name HG#)(resid 24 and name HB#) 4.5 2.7 2.5
assign (resid 23 and name HD#)(resid 24 and name HB#) 4.5 2.7 2.5
assign (resid 29 and name HG#)(resid 30 and name HA#) 4.5 2.7 2.5

!AROMATICS-OTHER(i+1)

assign (resid 4 and name HD#)(resid 5 and name HG#) 4.5 2.7 2.5
assign (resid 4 and name HE3)(resid 5 and name HA) 4.5 2.7 2.5
assign (resid 4 and name HD#)(resid 5 and name HA) 4.5 2.7 2.5
assign (resid 11 and name HD#)(resid 10 and name HA#) 4.5 2.7 2.5
assign (resid 11 and name HD#)(resid 10 and name HA#) 4.5 2.7 2.5
assign (resid 13 and name HD#)(resid 14 and name HA) 4.5 2.7 2.5
assign (resid 13 and name HD#)(resid 14 and name HB#) 4.5 2.7 2.5
assign (resid 13 and name HE3)(resid 14 and name HA) 4.5 2.7 2.5
assign (resid 28 and name HD#)(resid 29 and name HA) 4.5 2.7 2.5
assign (resid 28 and name HD#)(resid 29 and name HG#) 4.5 2.7 2.5
assign (resid 28 and name HE#)(resid 29 and name HA) 4.5 2.7 2.5
assign (resid 28 and name HE#)(resid 29 and name HG#) 4.5 2.7 2.5

clas short

!HN(i)-OTHER(i+2)

assign (resid 2 and name HB#)(resid 4 and name HN) 4.5 2.7 2.5
assign (resid 5 and name HB#)(resid 7 and name HA) 4.5 2.7 1.5
assign (resid 5 and name HG2#)(resid 7 and name HA) 4.5 2.7 2.5
assign (resid 5 and name HG12)(resid 7 and name HA) 4.5 2.7 2.5
assign (resid 5 and name HG1#)(resid 7 and name HA) 4.5 2.7 2.5
assign (resid 6 and name HE#)(resid 8 and name HB#) 4.5 2.7 2.5
assign (resid 7 and name HA)(resid 9 and name HB#) 4.5 2.7 2.5
assign (resid 7 and name HA)(resid 9 and name HG2) 4.5 2.7 2.5
assign (resid 8 and name HA)(resid 10 and name HN) 3.5 1.7 0.5
assign (resid 8 and name HB#)(resid 10 and name HN) 4.5 2.7 2.5
assign (resid 8 and name HG#)(resid 10 and name HA#) 4.5 2.7 2.5
assign (resid 12 and name HB1)(resid 14 and name HN) 4.5 2.7 2.5
assign (resid 12 and name HB2)(resid 14 and name HN) 4.5 2.7 2.5
assign (resid 12 and name HB1)(resid 14 and name HG#) 4.5 2.7 2.5
assign (resid 12 and name HB2)(resid 14 and name HG#) 4.5 2.7 2.5
assign (resid 14 and name HB#)(resid 16 and name HN) 4.5 2.7 2.5
assign (resid 14 and name HA)(resid 16 and name HN) 3.5 1.7 0.5
assign (resid 14 and name HN)(resid 16 and name HN) 4.5 2.7 2.5
assign (resid 14 and name HG#)(resid 16 and name HA#) 4.5 2.7 2.5
assign (resid 14 and name HG#)(resid 16 and name HN) 3.5 1.7 2.5
assign (resid 17 and name HE#)(resid 19 and name HB#) 4.5 2.7 2.5
assign (resid 17 and name HE#)(resid 19 and name HG#) 4.5 2.7 2.5
assign (resid 17 and name HE#)(resid 19 and name HG#) 4.5 2.7 2.5
assign (resid 17 and name HN)(resid 19 and name HG#) 4.5 2.7 2.5
assign (resid 18 and name HA#)(resid 20 and name HB#) 4.5 2.7 2.5
assign (resid 23 and name HB#)(resid 25 and name HN) 3.5 1.7 2.5
assign (resid 24 and name HN)(resid 26 and name HN) 4.5 2.7 2.5
assign (resid 27 and name HN)(resid 29 and name HG#) 4.5 2.7 2.5
assign (resid 27 and name HB#)(resid 29 and name HG#) 4.5 2.7 2.5

!OTHER(i)-OTHER(i+3)

assign (resid 6 and name HE#)(resid 9 and name HB1) 4.5 2.7 2.5
assign (resid 6 and name HE#)(resid 9 and name HB2) 4.5 2.7 2.5
assign (resid 6 and name HE#)(resid 9 and name HG#) 4.5 2.7 2.5
assign (resid 8 and name HB#)(resid 11 and name HD#) 4.5 2.7 2.5
assign (resid 8 and name HG#)(resid 11 and name HD#) 4.5 2.7 2.5
assign (resid 12 and name HB#)(resid 15 and name HD#) 4.5 2.7 2.5
assign (resid 17 and name HD#)(resid 20 and name HB#) 4.5 2.7 2.5
assign (resid 17 and name HE#)(resid 20 and name HB#) 4.5 2.7 2.5
assign (resid 23 and name HB#)(resid 26 and name HZ) 4.5 2.7 2.5
assign (resid 23 and name HG#)(resid 26 and name HZ) 4.5 2.7 2.5

!OTHER(i)-OTHER(i+4)

assign (resid 8 and name HA)(resid 12 and name HN) 3.0 1.2 0.5
assign (resid 14 and name HN)(resid 18 and name HA#) 4.5 2.7 0.5
assign (resid 24 and name HB#)(resid 28 and name HD#) 4.5 2.7 2.5
assign (resid 24 and name HB1)(resid 28 and name HE#) 4.5 2.7 2.5
assign (resid 24 and name HB2)(resid 28 and name HE#) 4.5 2.7 2.5

clas intra
!HN(i)-HA(i)

assign (resid 1 and name HN)(resid 1 and name HA) 4.5 2.7 0.5
assign (resid 3 and name HN)(resid 3 and name HA) 3.5 1.7 0.5
assign (resid 4 and name HN)(resid 4 and name HA) 3.5 1.7 0.5
assign (resid 5 and name HN)(resid 5 and name HA) 3.5 1.7 0.5
assign (resid 6 and name HN)(resid 6 and name HA) 2.3 0.5 0.5
assign (resid 7 and name HN)(resid 7 and name HA) 2.3 0.5 0.5
assign (resid 8 and name HN)(resid 8 and name HA) 3.0 1.2 0.5
assign (resid 9 and name HN)(resid 9 and name HA) 3 1.2 0.5
assign (resid 10 and name HN)(resid 10 and name HA1) 3 1.2 0.5
assign (resid 10 and name HN)(resid 10 and name HA2) 2.3 0.5 0.5
assign (resid 11 and name HN)(resid 11 and name HA) 3 1.2 0.5
assign (resid 12 and name HN)(resid 12 and name HA) 2.3 0.5 0.5
assign (resid 13 and name HN)(resid 13 and name HA) 3.5 1.7 0.5
assign (resid 14 and name HN)(resid 14 and name HA) 2.3 0.5 0.5
assign (resid 15 and name HN)(resid 15 and name HA) 3.5 1.7 0.5
assign (resid 16 and name HN)(resid 16 and name HA#) 2.3 0.5 0.5
assign (resid 17 and name HN)(resid 17 and name HA) 2.3 0.5 0.5
assign (resid 18 and name HN)(resid 18 and name HA#) 2.3 0.5 0.5
assign (resid 19 and name HN)(resid 19 and name HA) 3.5 1.7 0.5
assign (resid 20 and name HN)(resid 20 and name HA) 2.3 0.5 0.5
assign (resid 21 and name HN)(resid 21 and name HA) 3.0 1.2 0.5
assign (resid 22 and name HN)(resid 22 and name HA) 2.3 0.5 0.5
assign (resid 23 and name HN)(resid 23 and name HA) 3.0 1.2 0.5
assign (resid 24 and name HN)(resid 24 and name HA) 2.3 0.5 0.5
assign (resid 25 and name HN)(resid 25 and name HA#) 2.3 0.5 0.5
assign (resid 25 and name HN)(resid 25 and name HA#) 2.3 0.5 0.5
assign (resid 26 and name HN)(resid 26 and name HA) 3.0 1.2 0.5
assign (resid 27 and name HN)(resid 27 and name HA) 2.3 0.5 0.5
assign (resid 28 and name HN)(resid 28 and name HA) 3.0 1.2 0.5
assign (resid 29 and name HN)(resid 29 and name HA) 2.3 0.5 0.5
assign (resid 30 and name HN)(resid 30 and name HA#) 3.0 1.2 0.5

```
! Loop 1 in DPC NOE Table
! LPFWIHYVRG HNWVFGHGMA KLLSGFYHTG
clas long

assign (resid 4 and name HZ2 )(resid 22 and name HB# ) 4.5 2.7 2.5
assign (resid 5 and name HG1#)(resid 12 and name HD2# ) 3.5 1.7 2.5
assign (resid 7 and name HN )(resid 15 and name HD# ) 4.5 2.7 2.5
assign (resid 7 and name HN )(resid 23 and name HB# ) 4.5 2.7 2.5
assign (resid 7 and name HB# )(resid 24 and name HN ) 4.5 2.7 2.5
assign (resid 9 and name HB# )(resid 18 and name HA# ) 4.5 2.7 2.5
assign (resid 9 and name HG# )(resid 18 and name HA# ) 4.5 2.7 2.5
assign (resid 9 and name HD# )(resid 18 and name HA# ) 4.5 2.7 2.5
assign (resid 15 and name HB# )(resid 21 and name HN ) 4.5 2.7 2.5
assign (resid 24 and name HA )(resid 4 and name HD1 ) 4.5 2.7 2.5

clas avelong

assign (resid 3 and name HA )(resid 27 and name HE# ) 4.5 2.7 2.5
assign (resid 4 and name HA )(resid 27 and name HE# ) 4.5 2.7 2.5
assign (resid 4 and name HB# )(resid 27 and name HE# ) 4.5 2.7 2.5
assign (resid 5 and name HD# )(resid 27 and name HD#) 4.5 2.7 2.5
assign (resid 5 and name HG1#)(resid 27 and name HD# ) 4.5 2.7 2.5
assign (resid 7 and name HE# )(resid 15 and name HA ) 4.5 2.7 2.5
assign (resid 7 and name HD# )(resid 12 and name HD2# ) 4.5 2.7 2.5
assign (resid 7 and name HD# )(resid 14 and name HG#) 4.5 2.7 2.5
assign (resid 8 and name HG# )(resid 15 and name HE#) 4.5 2.7 2.5

clas aveshort

assign (resid 4 and name HA )(resid 7 and name HE#) 4.5 2.7 2.5
assign (resid 4 and name HB# )(resid 7 and name HE#) 4.5 2.7 2.5
assign (resid 5 and name HG1#)(resid 7 and name HE# ) 4.5 2.7 2.5
assign (resid 5 and name HD# )(resid 7 and name HE# ) 4.5 2.7 2.5
assign (resid 5 and name HD# )(resid 7 and name HD# ) 4.5 2.7 2.5
assign (resid 5 and name HG1#)(resid 7 and name HD# ) 4.5 2.7 2.5
assign (resid 6 and name HA )(resid 7 and name HD# ) 4.5 2.7 2.5
assign (resid 6 and name HA )(resid 7 and name HE# ) 4.5 2.7 2.5
assign (resid 7 and name HN )(resid 7 and name HE#) 4.5 2.7 2.5
assign (resid 7 and name HN )(resid 7 and name HD#) 4.5 2.7 2.5
assign (resid 14 and name HG# )(resid 15 and name HD# ) 4.5 2.7 2.5
assign (resid 14 and name HA )(resid 15 and name HD# ) 4.5 2.7 2.5
assign (resid 14 and name HN )(resid 15 and name HD# ) 4.5 2.7 2.5
assign (resid 15 and name HN )(resid 15 and name HD#) 3.5 1.7 2.5
assign (resid 15 and name HD# )(resid 16 and name HN ) 4.5 2.7 2.5
assign (resid 15 and name HD# )(resid 19 and name HB# ) 4.5 2.7 2.5
assign (resid 24 and name HA )(resid 27 and name HD#) 4.5 2.7 2.5
assign (resid 24 and name HB# )(resid 27 and name HD#) 4.5 2.7 2.5
assign (resid 24 and name HB# )(resid 27 and name HE#) 4.5 2.7 2.5
assign (resid 25 and name HA# )(resid 27 and name HD# ) 4.5 2.7 2.5
assign (resid 27 and name HN )(resid 27 and name HE#) 4.5 2.7 2.5
assign (resid 27 and name HN )(resid 27 and name HD#) 3.5 1.7 2.5
assign (resid 27 and name HD# )(resid 28 and name HN ) 4.5 2.7 2.5
assign (resid 27 and name HD# )(resid 28 and name HA ) 4.5 2.7 2.5
assign (resid 27 and name HE# )(resid 28 and name HN ) 4.5 2.7 2.5

clas seq
!HN(i)-HN(i+1)

assign (resid 16 and name HN )(resid 17 and name HN ) 4.5 2.7 0.5
assign (resid 17 and name HN )(resid 18 and name HN ) 4.5 2.7 0.5
assign (resid 18 and name HN )(resid 19 and name HN ) 4.5 2.7 0.5
assign (resid 20 and name HN )(resid 21 and name HN ) 4.5 2.7 0.5
assign (resid 24 and name HN )(resid 25 and name HN ) 4.5 2.7 0.5
assign (resid 25 and name HN )(resid 26 and name HN ) 4.5 2.7 0.5
assign (resid 26 and name HN )(resid 27 and name HN ) 4.5 2.7 0.5
assign (resid 27 and name HN )(resid 28 and name HN ) 3 1.2 0.5

!HA(i)-HN(i+1)

assign (resid 2 and name HA )(resid 3 and name HN ) 3.5 1.7 0.5
assign (resid 5 and name HA )(resid 6 and name HN ) 4.5 2.7 0.5
assign (resid 7 and name HA )(resid 8 and name HN ) 4.5 2.7 0.5
assign (resid 8 and name HA )(resid 9 and name HN ) 4.5 2.7 0.5
assign (resid 11 and name HA )(resid 12 and name HN ) 4.5 2.7 0.5
assign (resid 14 and name HA )(resid 15 and name HN ) 4.5 2.7 0.5
assign (resid 15 and name HA )(resid 16 and name HN ) 4.5 2.7 0.5
assign (resid 16 and name HA# )(resid 17 and name HN ) 4.5 2.7 1.5
assign (resid 17 and name HA )(resid 18 and name HN ) 4.5 2.7 0.5
assign (resid 18 and name HA1 )(resid 19 and name HN ) 3.5 1.7 0.5
assign (resid 18 and name HA2 )(resid 19 and name HN ) 3.5 1.7 0.5
assign (resid 19 and name HA )(resid 20 and name HN ) 3.5 1.7 0.5
assign (resid 20 and name HA )(resid 21 and name HN ) 3.5 1.7 0.5
assign (resid 21 and name HA )(resid 22 and name HN ) 4.5 2.7 0.5
assign (resid 22 and name HA )(resid 23 and name HN ) 4.5 2.7 0.5
assign (resid 24 and name HA )(resid 25 and name HN ) 3.5 1.7 0.5
assign (resid 25 and name HA# )(resid 26 and name HN ) 4.5 2.7 1.5
assign (resid 26 and name HA )(resid 27 and name HN ) 4.5 2.7 0.5
assign (resid 27 and name HA )(resid 28 and name HN ) 3.0 1.2 0.5
assign (resid 29 and name HA )(resid 30 and name HN ) 3.0 1.2 0.5

!HB(i)-HN(i+1)

assign (resid 7 and name HB# )(resid 8 and name HN ) 4.5 2.7 1.5
assign (resid 8 and name HB# )(resid 6 and name HN ) 4.5 2.7 1.5
assign (resid 15 and name HB# )(resid 16 and name HN ) 4.5 2.7 1.5
assign (resid 17 and name HB# )(resid 18 and name HN ) 4.5 2.7 1.5
assign (resid 19 and name HB# )(resid 20 and name HN ) 4.5 2.7 1.5
assign (resid 20 and name HB# )(resid 21 and name HN ) 3.5 1.7 1.5
assign (resid 23 and name HB# )(resid 24 and name HN ) 4.5 2.7 1.5
assign (resid 27 and name HB# )(resid 28 and name HN ) 3.5 1.7 1.5

!HG(i)-HN(i+1)

assign (resid 8 and name HG# )(resid 6 and name HN ) 4.5 2.7 2.5
assign (resid 19 and name HG# )(resid 20 and name HN ) 4.5 2.7 2.5
assign (resid 21 and name HG# )(resid 22 and name HN ) 4.5 2.7 2.5
assign (resid 23 and name HG# )(resid 24 and name HN ) 4.5 2.7 2.5

!HN(i)-OTHER(i+1)

assign (resid 19 and name HN )(resid 20 and name HB# ) 4.5 2.7 1.5

!HA(i)-OTHER(i+1)

assign (resid 1 and name HA )(resid 2 and name HD# ) 4.5 2.7 2.5
assign (resid 5 and name HA )(resid 6 and name HD# ) 4.5 2.7 2.5
assign (resid 19 and name HA )(resid 20 and name HB# ) 4.5 2.7 1.5

!HB(i)-OTHER(i+1)

assign (resid 19 and name HB# )(resid 20 and name HA ) 4.5 2.7 1.5

!HG(i)-OTHER(i+1)

assign (resid 14 and name HG# )(resid 15 and name HB# ) 4.5 2.7 2.5
assign (resid 19 and name HG# )(resid 20 and name HB# ) 4.5 2.7 2.5
assign (resid 14 and name HG# )(resid 15 and name HA ) 4.5 2.7 2.5

!AROMATICS-(i+1)

assign (resid 4 and name HH2 )(resid 5 and name HG2# ) 4.5 2.7 2.5
assign (resid 5 and name HA )(resid 13 and name HZ2 ) 4.5 2.7 2.5
assign (resid 10 and name HA# )(resid 13 and name HD1 ) 4.5 2.7 2.5
assign (resid 10 and name HA# )(resid 13 and name HZ3 ) 4.5 2.7 2.5
assign (resid 10 and name HA# )(resid 13 and name HH2 ) 4.5 2.7 2.5
assign (resid 11 and name HN ) (resid 13 and name HZ3 ) 4.5 2.7 2.5
assign (resid 12 and name HB# )(resid 13 and name HD1 ) 4.5 2.7 2.5
assign (resid 15 and name HZ )(resid 16 and name HA# ) 4.5 2.7 2.5
assign (resid 15 and name HZ )(resid 16 and name HN# ) 4.5 2.7 2.5
```

```
!OTHER(i)-OTHER(i+1)

assign (resid 5 and name HD# )(resid 6 and name HA ) 4.5 2.7 2.5
assign (resid 5 and name HD# )(resid 6 and name HN ) 4.5 2.7 2.5

clas short

! (i)-OTHER(i+2)

assign (resid 2 and name HB# )(resid 4 and name HD1 ) 4.5 2.7 2.5
assign (resid 2 and name HG# )(resid 4 and name HD1 ) 4.5 2.7 2.5
assign (resid 2 and name HB# )(resid 4 and name HE1 ) 4.5 2.7 2.5
assign (resid 2 and name HB# )(resid 4 and name HZ2 ) 4.5 2.7 2.5
assign (resid 2 and name HG# )(resid 4 and name HZ2 ) 4.5 2.7 2.5
assign (resid 20 and name HA )(resid 22 and name HD#) 4.5 2.7 2.5

!OTHER(i)-OTHER(i+3)

assign (resid 2 and name HA )(resid 5 and name HG1# ) 4.5 2.7 2.5
assign (resid 2 and name HA )(resid 5 and name HG2# ) 4.5 2.7 2.5
assign (resid 2 and name HB# )(resid 5 and name HG1# ) 4.5 2.7 2.5
assign (resid 2 and name HB# )(resid 5 and name HG2# ) 4.5 2.7 2.5
assign (resid 2 and name HD# )(resid 5 and name HD#) 4.5 2.7 2.5
assign (resid 2 and name HD# )(resid 5 and name HG1# ) 4.5 2.7 2.5
assign (resid 2 and name HD# )(resid 5 and name HG2# ) 4.5 2.7 2.5
assign (resid 5 and name HA )(resid 8 and name HB#) 4.5 2.7 2.5
assign (resid 5 and name HA )(resid 8 and name HG#) 4.5 2.7 2.5
assign (resid 17 and name HA )(resid 20 and name HB#) 3.5 1.7 2.5
assign (resid 18 and name HA#)(resid 21 and name HB# ) 4.5 2.7 2.5
assign (resid 18 and name HA#)(resid 21 and name HD# ) 4.5 2.7 2.5
assign (resid 19 and name HN )(resid 22 and name HG# ) 4.5 2.7 2.5
assign (resid 24 and name HA )(resid 27 and name HB# ) 4.5 2.7 2.5

!OTHER(i)-OTHER(i+4)

assign (resid 1 and name HA )(resid 5 and name HG# ) 4.5 2.7 2.5
assign (resid 1 and name HA )(resid 5 and name HD# ) 4.5 2.7 2.5
assign (resid 4 and name HH2 )(resid 8 and name HG#) 4.5 2.7 2.5
assign (resid 4 and name HZ3 )(resid 8 and name HB# ) 4.5 2.7 2.5
assign (resid 4 and name HZ3 )(resid 8 and name HG# ) 4.5 2.7 2.5
assign (resid 15 and name HZ )(resid 19 and name HB# ) 4.5 2.7 2.5
assign (resid 16 and name HN )(resid 20 and name HB# ) 4.5 2.7 2.5
assign (resid 19 and name HG# )(resid 23 and name HD#) 4.5 2.7 2.5

clas intra
!HN(i)-HA(i)

assign (resid 3 and name HN )(resid 3 and name HA ) 4.5 2.7 0.5
assign (resid 5 and name HN )(resid 5 and name HA ) 3.0 1.2 0.5
assign (resid 6 and name HN )(resid 6 and name HA ) 4.5 2.7 0.5
assign (resid 7 and name HN )(resid 7 and name HA ) 3.5 1.7 0.5
assign (resid 8 and name HN )(resid 8 and name HA ) 4.5 2.7 0.5
assign (resid 10 and name HN )(resid 10 and name HA# ) 3.0 1.2 1.5
assign (resid 13 and name HN )(resid 13 and name HA ) 3 1.2 0.5
assign (resid 14 and name HN )(resid 14 and name HA ) 4.5 2.7 0.5
assign (resid 16 and name HN )(resid 16 and name HA# ) 4.5 2.7 0.5
assign (resid 17 and name HN )(resid 17 and name HA ) 4.5 2.7 0.5
assign (resid 18 and name HN )(resid 18 and name HA# ) 3.5 1.7 0.5
assign (resid 18 and name HN )(resid 18 and name HA# ) 3.5 1.7 0.5
assign (resid 19 and name HN )(resid 19 and name HA ) 3.5 1.7 0.5
assign (resid 20 and name HN )(resid 20 and name HA ) 4.5 2.7 0.5
assign (resid 21 and name HN )(resid 21 and name HA ) 3.5 1.7 0.5
assign (resid 22 and name HN )(resid 22 and name HA ) 3.5 1.7 0.5
assign (resid 24 and name HN )(resid 24 and name HA ) 3.0 1.2 0.5
assign (resid 25 and name HN )(resid 25 and name HA# ) 3.0 1.2 0.5
assign (resid 28 and name HN )(resid 28 and name HA ) 3.5 1.7 0.5
assign (resid 29 and name HN )(resid 29 and name HA ) 4.5 2.7 0.5

!HN(i)-HB(i)

assign (resid 8 and name HN )(resid 8 and name HB#) 4.5 2.7 1.5
assign (resid 15 and name HN )(resid 15 and name HB#) 3.5 1.7 1.5
assign (resid 20 and name HN )(resid 20 and name HB#) 3.0 1.2 1.5
assign (resid 22 and name HN )(resid 22 and name HB#) 3.0 1.2 1.5
assign (resid 23 and name HN )(resid 23 and name HB#) 3.0 1.2 1.5
assign (resid 24 and name HN )(resid 24 and name HB#) 3.0 1.2 1.5
assign (resid 3 and name HN )(resid 3 and name HB# ) 4.5 2.7 1.5
assign (resid 6 and name HN )(resid 6 and name HB# ) 4.5 2.7 1.5
assign (resid 7 and name HN )(resid 7 and name HB# ) 4.5 2.7 1.5
assign (resid 11 and name HN )(resid 11 and name HB# ) 3.0 1.2 1.5
assign (resid 17 and name HN )(resid 17 and name HB# ) 4.5 2.7 1.5
assign (resid 19 and name HN )(resid 19 and name HB# ) 3.5 1.7 1.5
assign (resid 21 and name HN )(resid 21 and name HB# ) 3.0 1.2 1.5
assign (resid 26 and name HN )(resid 26 and name HB# ) 4.5 2.7 1.5
assign (resid 27 and name HN )(resid 27 and name HB# ) 3.0 1.2 1.5
assign (resid 28 and name HN )(resid 28 and name HB# ) 3.5 1.7 1.5
assign (resid 29 and name HN )(resid 29 and name HB# ) 4.5 1.7 1.5

!HN(i)-HG(i)

assign (resid 8 and name HN )(resid 8 and name HG# ) 4.5 2.7 2.5
assign (resid 14 and name HN )(resid 14 and name HG# ) 4.5 2.7 2.5
assign (resid 19 and name HN )(resid 19 and name HG# ) 4.5 2.7 2.5
assign (resid 22 and name HN )(resid 22 and name HG# ) 4.5 2.7 2.5
assign (resid 23 and name HN )(resid 23 and name HG# ) 4.5 2.7 2.5

!HN(i)-HD(i)

assign (resid 6 and name HN )(resid 6 and name HD#) 4.5 2.7 2.5
assign (resid 21 and name HN )(resid 21 and name HD# ) 4.5 2.7 2.5
assign (resid 22 and name HN )(resid 22 and name HD# ) 4.5 2.7 2.5
assign (resid 13 and name HN )(resid 13 and name HD1 ) 4.5 2.7 2.5
```

```
! Loop 3 in DMSO NOE Table
! AILLSSYQSILFGNDAERSKHLDLVM
clas long

clas aveshort

assign (resid 6 and name HA )(resid 7 and name HD# ) 4.5 2.7 2.5
assign (resid 6 and name HB# )(resid 7 and name HD# ) 4.5 2.7 2.5
assign (resid 7 and name HD# )(resid 10 and name HD# ) 4.5 2.7 2.5
assign (resid 7 and name HN )(resid 7 and name HD# ) 4.5 2.7 2.5
assign (resid 7 and name HN )(resid 7 and name HE# ) 4.5 2.7 2.5
assign (resid 9 and name HA )(resid 12 and name HD# ) 4.5 2.7 2.5
assign (resid 9 and name HB# )(resid 12 and name HD# ) 4.5 2.7 2.5
assign (resid 10 and name HA )(resid 12 and name HD# ) 4.5 2.7 2.5
assign (resid 11 and name HB# )(resid 12 and name HD# ) 4.5 2.7 2.5
assign (resid 11 and name HD# )(resid 12 and name HD# ) 4.5 2.7 2.5
assign (resid 11 and name HD# )(resid 12 and name HD# ) 4.5 2.7 2.5
assign (resid 11 and name HG# )(resid 12 and name HD# ) 4.5 2.7 2.5
assign (resid 12 and name HN )(resid 12 and name HD# ) 4.5 2.7 2.5
assign (resid 12 and name HD# )(resid 16 and name HB# ) 4.5 2.7 2.5
assign (resid 11 and name HA )(resid 12 and name HD# ) 4.5 2.7 2.5
assign (resid 12 and name HD# )(resid 13 and name HA# ) 4.5 2.7 2.5

clas seq

!HN(i)-HN(i+1)
assign (resid 12 and name HN )(resid 13 and name HN ) 3.5 1.7 0.5
assign (resid 14 and name HN )(resid 15 and name HN ) 4.5 2.7 0.5
assign (resid 15 and name HN )(resid 16 and name HN ) 3.5 1.7 0.5
assign (resid 16 and name HN )(resid 17 and name HN ) 3 1.2 0.5
assign (resid 17 and name HN )(resid 18 and name HN ) 3.5 1.7 0.5
assign (resid 18 and name HN )(resid 19 and name HN ) 3.5 1.7 0.5
assign (resid 19 and name HN )(resid 20 and name HN ) 3.5 1.7 0.5

!HA(i)-HN(i+1)

assign (resid 1 and name HA )(resid 2 and name HN ) 3 1.2 0.5
assign (resid 2 and name HA )(resid 3 and name HN ) 3 1.2 0.5
assign (resid 5 and name HA )(resid 6 and name HN ) 3 1.2 0.5
assign (resid 6 and name HA )(resid 7 and name HN ) 3 1.2 0.5
assign (resid 8 and name HA )(resid 9 and name HN ) 3 1.2 0.5
assign (resid 9 and name HA )(resid 10 and name HN ) 3 1.2 0.5
assign (resid 10 and name HA )(resid 11 and name HN ) 3 1.2 0.5
assign (resid 11 and name HA )(resid 12 and name HN ) 3 1.2 0.5
assign (resid 12 and name HA )(resid 13 and name HN ) 3 1.2 0.5
assign (resid 13 and name HA# )(resid 14 and name HN ) 3.5 1.7 0.5
assign (resid 13 and name HA# )(resid 14 and name HN ) 3.5 1.7 0.5
assign (resid 14 and name HA )(resid 15 and name HN ) 3.5 1.7 0.5
assign (resid 15 and name HA )(resid 16 and name HN ) 3.0 1.2 0.5
assign (resid 17 and name HA )(resid 18 and name HN ) 3 1.2 0.5
assign (resid 19 and name HA )(resid 20 and name HN ) 3 1.2 0.5
assign (resid 21 and name HA )(resid 22 and name HN ) 3.5 1.7 0.5
assign (resid 22 and name HA )(resid 23 and name HN ) 4.5 2.7 0.5
assign (resid 23 and name HA )(resid 24 and name HN ) 3.5 1.7 0.5
assign (resid 24 and name HA )(resid 25 and name HN ) 3.5 1.7 0.5
assign (resid 25 and name HA )(resid 26 and name HN ) 3.5 1.7 0.5

!HB(i)-HN(i+1)

assign (resid 1 and name HB# )(resid 2 and name HN ) 3.5 1.7 1.5
assign (resid 2 and name HB# )(resid 3 and name HN ) 4.5 2.7 2.0
assign (resid 5 and name HB# )(resid 6 and name HN ) 4.5 2.7 2.0
assign (resid 6 and name HB# )(resid 7 and name HN ) 3.5 1.7 2.0
assign (resid 7 and name HB# )(resid 8 and name HN ) 4.5 2.7 2.0
assign (resid 7 and name HB# )(resid 8 and name HN ) 4.5 2.7 2
assign (resid 8 and name HB# )(resid 9 and name HN ) 3.5 1.7 2
assign (resid 8 and name HB# )(resid 9 and name HN ) 3.5 1.7 2
assign (resid 9 and name HB# )(resid 10 and name HN ) 3.5 1.7 0.5
assign (resid 9 and name HB# )(resid 10 and name HN ) 3.5 1.7 0.5
assign (resid 11 and name HB# )(resid 12 and name HN ) 3.5 1.7 2.0
assign (resid 12 and name HB# )(resid 13 and name HN ) 4.5 2.7 2.0
assign (resid 12 and name HB# )(resid 13 and name HN ) 4.5 2.7 2.0
assign (resid 14 and name HB# )(resid 15 and name HN ) 3.5 1.7 2.0
assign (resid 14 and name HB# )(resid 15 and name HN ) 3.5 1.7 2
assign (resid 15 and name HB# )(resid 16 and name HN ) 4.5 2.7 2
assign (resid 15 and name HB# )(resid 16 and name HN ) 4.5 2.7 2
assign (resid 16 and name HB# )(resid 17 and name HN ) 3 1.2 1.5
assign (resid 17 and name HB# )(resid 18 and name HN ) 4.5 2.7 2
assign (resid 17 and name HB# )(resid 18 and name HN ) 4.5 2.7 2
assign (resid 18 and name HB# )(resid 19 and name HN ) 3.5 1.7 2
assign (resid 19 and name HB# )(resid 20 and name HN ) 3.5 1.7 2
assign (resid 19 and name HB# )(resid 20 and name HN ) 3.5 1.7 2
assign (resid 21 and name HB# )(resid 22 and name HN ) 4.5 2.7 2.
assign (resid 21 and name HB# )(resid 22 and name HN ) 4.5 2.7 2.
assign (resid 22 and name HB# )(resid 23 and name HN ) 4.5 2.7 2
assign (resid 23 and name HB# )(resid 24 and name HN ) 4.5 2.7 2
assign (resid 24 and name HB# )(resid 25 and name HN ) 4.5 2.7 2

!HG(i)-HN(i+1)

assign (resid 8 and name HG# )(resid 9 and name HN ) 4.5 2.7 2
assign (resid 10 and name HG1# )(resid 11 and name HN ) 4.5 2.7 2.0
assign (resid 10 and name HG1# )(resid 11 and name HN ) 4.5 2.7 2.0
assign (resid 11 and name HG# )(resid 12 and name HN ) 4.5 2.7 2.0
assign (resid 17 and name HG# )(resid 18 and name HN ) 4.5 2.7 2
assign (resid 18 and name HG# )(resid 19 and name HN ) 3.5 1.7 2

!HD(i)-HN(i+1)

assign (resid 10 and name HD# )(resid 11 and name HN ) 3.5 1.7 2.5
assign (resid 11 and name HD# )(resid 12 and name HN ) 3.5 1.7 2

assign (resid 14 and name HD2# )(resid 15 and name HN ) 4.5 2.7 2.5
assign (resid 18 and name HD# )(resid 19 and name HN ) 4.5 2.7 2.0
assign (resid 24 and name HD# )(resid 25 and name HN ) 4.5 2.7 2

! Others
assign (resid 2 and name HN )(resid 3 and name HB# ) 4.5 2.7 2.0
assign (resid 4 and name HD# )(resid 5 and name HN ) 4.5 2.7 2.5
assign (resid 4 and name HD# )(resid 5 and name HN ) 4.5 2.7 2.5
assign (resid 7 and name HN )(resid 8 and name HG# ) 4.5 2.7 2.0
assign (resid 7 and name HN )(resid 8 and name HB# ) 4.5 2.7 2.0
assign (resid 7 and name HN )(resid 8 and name HB# ) 4.5 2.7 2.0
assign (resid 10 and name HN )(resid 11 and name HG# ) 4.5 2.7 2.0
assign (resid 12 and name HN )(resid 13 and name HA# ) 4.5 2.7 0.5
assign (resid 17 and name HN )(resid 18 and name HG# ) 4.5 2.7 2
assign (resid 15 and name HN )(resid 16 and name HB# ) 4.5 2.7 1.5

clas short

! (i)-H(i+2)

assign (resid 8 and name HG# )(resid 10 and name HN ) 4.5 2.7. 2.0
assign (resid 10 and name HA )(resid 12 and name HN ) 3.5 1.7 0.5
assign (resid 11 and name HA )(resid 13 and name HN ) 3.5 1.7 0.5
assign (resid 12 and name HA )(resid 14 and name HN ) 4.5 2.7 0.5
assign (resid 14 and name HA )(resid 16 and name HN ) 4.5 2.7 0.5
assign (resid 14 and name HD# )(resid 16 and name HB# ) 4.5 2.7 2.5
assign (resid 14 and name HD# )(resid 16 and name HB# ) 4.5 2.7 2.5
assign (resid 15 and name HA )(resid 17 and name HN ) 4.5 2.7 0.5
assign (resid 16 and name HB# )(resid 18 and name HN ) 4.5 2.7 1.5

!(i)-H(i+3)

assign (resid 16 and name HB# )(resid 19 and name HN ) 4.5 2.7 1.5

clas intra

assign (resid 2 and name HN )(resid 2 and name HA ) 4.5 2.7 0.5
assign (resid 2 and name HN )(resid 2 and name HB ) 4.5 2.7 2
```

assign (resid 2 and name HN)(resid 2 and name HD#) 4.5 2.7 2.5
assign (resid 2 and name HN)(resid 2 and name HG1#) 4.5 2.7 2
assign (resid 3 and name HN)(resid 3 and name HA) 4.5 2.7 0.5
assign (resid 3 and name HN)(resid 3 and name HB#) 3.5 1.7 2
assign (resid 3 and name HN)(resid 3 and name HD#) 4.5 2.7 2.5
assign (resid 3 and name HN)(resid 3 and name HG#) 4.5 2.7 2.0
assign (resid 4 and name HN)(resid 4 and name HB#) 3.5 1.7 2
assign (resid 4 and name HN)(resid 4 and name HD#) 4.5 2.7 2.5
assign (resid 4 and name HN)(resid 4 and name HD#) 4.5 2.7 2.5
assign (resid 4 and name HN)(resid 4 and name HG#) 4.5 2.7 2
assign (resid 5 and name HN)(resid 5 and name HB#) 3 1.2 2
assign (resid 6 and name HN)(resid 6 and name HA) 3.5 1.7 0.5
assign (resid 6 and name HN)(resid 6 and name HB#) 3.5 1.7 2
assign (resid 7 and name HN)(resid 7 and name HA) 3.5 1.7 0.5
assign (resid 7 and name HN)(resid 7 and name HB#) 3.5 1.7 2
assign (resid 7 and name HN)(resid 7 and name HB#) 3.5 1.7 2
assign (resid 8 and name HN)(resid 8 and name HA) 3.5 1.7 0.5
assign (resid 8 and name HN)(resid 8 and name HB#) 3.5 1.7 2
assign (resid 8 and name HN)(resid 8 and name HB#) 3.5 1.7 2
assign (resid 8 and name HN)(resid 8 and name HG#) 3.5 1.7 2
assign (resid 9 and name HN)(resid 9 and name HB#) 3.5 1.7 2
assign (resid 9 and name HN)(resid 9 and name HB#) 3.5 1.7 2
assign (resid 10 and name HN)(resid 10 and name HA) 3.5 1.7 0.5
assign (resid 10 and name HN)(resid 10 and name HB) 3.5 1.7 2
assign (resid 10 and name HN)(resid 10 and name HD#) 3.5 1.7 2.5
assign (resid 10 and name HN)(resid 10 and name HG1#) 4.5 2.7 2
assign (resid 10 and name HN)(resid 10 and name HG1#) 4.5 2.7 2
assign (resid 11 and name HN)(resid 11 and name HB#) 3.5 1.7 2
assign (resid 11 and name HN)(resid 11 and name HD#) 4.5 2.7 2.5
assign (resid 11 and name HN)(resid 11 and name HD#) 4.5 2.7 2.5
assign (resid 11 and name HN)(resid 11 and name HG#) 4.5 2.7 2
assign (resid 12 and name HN)(resid 12 and name HA) 3.5 1.7 0.5
assign (resid 12 and name HN)(resid 12 and name HB#) 3.5 1.7 2
assign (resid 12 and name HN)(resid 12 and name HB#) 3.5 1.7 2
assign (resid 13 and name HN)(resid 13 and name HA#) 3.5 1.7 0.5
assign (resid 13 and name HN)(resid 13 and name HA#) 3.5 1.7 0.5
assign (resid 14 and name HN)(resid 14 and name HA) 3.5 1.7 0.5
assign (resid 14 and name HN)(resid 14 and name HB#) 3.5 1.7 2
assign (resid 14 and name HN)(resid 14 and name HB#) 3.5 1.7 2
assign (resid 14 and name HN)(resid 14 and name HD2#) 4.5 2.7 2.5
assign (resid 14 and name HN)(resid 14 and name HD2#) 4.5 2.7 2.5
assign (resid 15 and name HN)(resid 15 and name HA) 3.5 1.7 0.5
assign (resid 15 and name HN)(resid 15 and name HB#) 3.5 1.7 2
assign (resid 15 and name HN)(resid 15 and name HB#) 3.5 1.7 2
assign (resid 16 and name HN)(resid 16 and name HA) 3.5 1.7 0.5
assign (resid 16 and name HN)(resid 16 and name HB#) 3 1.2 1.5
assign (resid 17 and name HN)(resid 17 and name HB#) 3.5 1.7 2
assign (resid 17 and name HN)(resid 17 and name HB#) 3.5 1.7 2
assign (resid 17 and name HN)(resid 17 and name HG#) 4.5 2.7 2
assign (resid 18 and name HN)(resid 18 and name HA) 3.5 1.7 0.5
assign (resid 18 and name HN)(resid 18 and name HB#) 3.5 1.7 2
assign (resid 18 and name HN)(resid 18 and name HD#) 4.5 2.7 2.5
assign (resid 18 and name HN)(resid 18 and name HG#) 4.5 2.7 2
assign (resid 19 and name HN)(resid 19 and name HB#) 3.5 1.7 2
assign (resid 19 and name HN)(resid 19 and name HB#) 3.5 1.7 2
assign (resid 20 and name HN)(resid 20 and name HA) 3.5 1.7 0.5
assign (resid 20 and name HN)(resid 20 and name HB#) 3.5 1.7 2
assign (resid 20 and name HN)(resid 20 and name HD#) 4.5 2.7 2
assign (resid 20 and name HN)(resid 20 and name HG#) 3.5 1.7 2
assign (resid 21 and name HN)(resid 21 and name HB#) 4.5 2.7 2
assign (resid 21 and name HN)(resid 21 and name HB#) 4.5 2.7 2
assign (resid 22 and name HN)(resid 22 and name HA) 3.5 1.7 0.5
assign (resid 22 and name HN)(resid 22 and name HB#) 3.5 1.7 2
assign (resid 22 and name HN)(resid 22 and name HD#) 4.5 2.7 2.5
assign (resid 22 and name HN)(resid 22 and name HG#) 3.5 1.7 2.0
assign (resid 23 and name HN)(resid 23 and name HA) 4.5 2.7 0.5
assign (resid 24 and name HN)(resid 24 and name HA) 3.5 1.7 0.5
assign (resid 24 and name HN)(resid 24 and name HB#) 3.5 1.7 2
assign (resid 24 and name HN)(resid 24 and name HD#) 4.5 2.7 2.5
assign (resid 24 and name HN)(resid 24 and name HG#) 4.5 2.7 2.0
assign (resid 25 and name HN)(resid 25 and name HA) 4.5 2.7 0.5
assign (resid 25 and name HN)(resid 25 and name HB) 4.5 2.7 2
assign (resid 26 and name HN)(resid 26 and name HA) 4.5 2.7 0.5
assign (resid 26 and name HN)(resid 26 and name HB#) 4.5 2.7 2
assign (resid 26 and name HN)(resid 26 and name HB#) 4.5 2.7 2

```
! Loop 3 DPC
! AILLSSYQSILFGNDAERSKHLDLVM
clas long

assign (resid 2 and name HN )(resid 26 and name HG1# ) 4.5 2.7 2.0
assign (resid 5 and name HA )(resid 15 and name HN ) 4.5 2.7 0.5
assign (resid 18 and name HB# )(resid 13 and name HN ) 4.5 2.7 2.0
assign (resid 22 and name HG# )(resid 12 and name HZ ) 4.5 2.7 2.5
assign (resid 26 and name HG1# )(resid 3 and name HN ) 4.5 2.7 1

clas avelong
assign (resid 7 and name HD# )(resid 12 and name HD# ) 4.5 2.7 2.5

clas aveshort

assign (resid 5 and name HA )(resid 7 and name HD# ) 3.5 1.7 2.5
assign (resid 5 and name HA )(resid 7 and name HE# ) 3.5 1.7 2.5
assign (resid 5 and name HB# )(resid 7 and name HE# ) 4.5 2.7 2.5
assign (resid 6 and name HB# )(resid 7 and name HD# ) 4.5 2.7 2.5
assign (resid 6 and name HN )(resid 7 and name HD# ) 4.5 2.7 2.5
assign (resid 6 and name HN )(resid 7 and name HE# ) 4.5 2.7 2.5
assign (resid 7 and name HN )(resid 7 and name HD# ) 3.5 1.7 2.5
assign (resid 7 and name HN )(resid 7 and name HE# ) 4.5 2.7 2.5
assign (resid 7 and name HE# )(resid 8 and name HN ) 4.5 2.7 2.5
assign (resid 7 and name HD# )(resid 8 and name HN ) 4.5 2.7 2.5
assign (resid 7 and name HD# )(resid 10 and name HD# ) 3.5 1.7 2.5
assign (resid 7 and name HD# )(resid 10 and name HB# ) 4.5 2.7 2.5
assign (resid 7 and name HE# )(resid 10 and name HD# ) 3.5 1.7 2.5
assign (resid 7 and name HE# )(resid 10 and name HB# ) 4.5 2.7 2.5
assign (resid 7 and name HD# )(resid 10 and name HN ) 4.5 2.7 2.5
assign (resid 10 and name HD# )(resid 12 and name HD# ) 4.5 2.7 2.5
assign (resid 10 and name HD# )(resid 12 and name HE# ) 4.5 2.7 2.5
assign (resid 11 and name HD# )(resid 12 and name HD# ) 4.5 2.7 2.5
assign (resid 11 and name HD# )(resid 12 and name HE# ) 4.5 2.7 2.5
assign (resid 12 and name HN )(resid 12 and name HD# ) 4.5 2.7 2.5
assign (resid 12 and name HD# )(resid 13 and name HA# ) 4.5 2.7 2.5
assign (resid 12 and name HE# )(resid 14 and name HD2# ) 4.5 2.7 2.5

clas seq

!HN(i)-HN(i+1)

assign (resid 4 and name HN )(resid 5 and name HN ) 4.5 2.7 0.5
assign (resid 5 and name HN )(resid 6 and name HN ) 4.5 2.7 0.5
assign (resid 6 and name HN )(resid 7 and name HN ) 4.5 2.7 0.5
assign (resid 7 and name HN )(resid 8 and name HN ) 4.5 2.7 0.5
assign (resid 8 and name HN )(resid 9 and name HN ) 4.5 2.7 0.5
assign (resid 9 and name HN )(resid 10 and name HN ) 4.5 2.7 0.5
assign (resid 10 and name HN )(resid 11 and name HN ) 4.5 2.7 0.5
assign (resid 11 and name HN )(resid 12 and name HN ) 4.5 2.7 0.5
assign (resid 12 and name HN )(resid 13 and name HN ) 4.5 2.7 0.5
assign (resid 13 and name HN )(resid 14 and name HN ) 4.5 2.7 0.5
assign (resid 15 and name HN )(resid 16 and name HN ) 4.5 2.7 0.5
assign (resid 17 and name HN )(resid 18 and name HN ) 4.5 2.7 0.5
assign (resid 18 and name HN )(resid 19 and name HN ) 4.5 2.7 0.5
assign (resid 19 and name HN )(resid 20 and name HN ) 4.5 2.7 0.5
assign (resid 23 and name HN )(resid 24 and name HN ) 4.5 2.7 0.5

!HA(i)-HN(i+1)

assign (resid 2 and name HA )(resid 3 and name HN ) 3.5 1.7 0.5
assign (resid 3 and name HA )(resid 4 and name HN ) 3.5 1.7 0.5
assign (resid 4 and name HA )(resid 5 and name HN ) 4.5 2.7 0.5
assign (resid 7 and name HA )(resid 8 and name HN ) 4.5 2.7 0.5
assign (resid 9 and name HA )(resid 10 and name HN ) 4.5 2.7 0.5
assign (resid 10 and name HA )(resid 11 and name HN ) 4.5 2.7 0.5
assign (resid 11 and name HA )(resid 12 and name HN ) 3.5 1.7 0.5
assign (resid 12 and name HA )(resid 13 and name HN ) 3.5 1.7 0.5
assign (resid 13 and name HA# )(resid 14 and name HN ) 3.5 1.7 0.5
assign (resid 13 and name HA# )(resid 14 and name HN ) 3.5 1.7 0.5
assign (resid 15 and name HA )(resid 16 and name HN ) 3.5 1.7 0.5
assign (resid 18 and name HA )(resid 19 and name HN ) 3.5 1.7 0.5
assign (resid 19 and name HA )(resid 20 and name HN ) 3.5 1.7 0.5
assign (resid 21 and name HA )(resid 22 and name HN ) 4.5 2.7 0.5
assign (resid 23 and name HA )(resid 24 and name HN ) 4.5 2.7 0.5
assign (resid 24 and name HA )(resid 25 and name HN ) 3.5 1.7 0.5
assign (resid 25 and name HA )(resid 26 and name HN ) 3.5 1.7 0.5

!HB(i)-HN(i+1)

assign (resid 4 and name HB# )(resid 5 and name HN ) 3.5 1.7 2.0
assign (resid 6 and name HB# )(resid 7 and name HN ) 4.5 2.7 2.0
assign (resid 6 and name HB# )(resid 7 and name HN ) 4.5 2.7 2.0
assign (resid 7 and name HB# )(resid 8 and name HN ) 4.5 2.7 2.0
assign (resid 7 and name HB# )(resid 8 and name HN ) 4.5 2.7 2.0
assign (resid 8 and name HB# )(resid 9 and name HN ) 4.5 2.7 2.0
assign (resid 10 and name HB# )(resid 11 and name HN ) 4.5 2.7 2.0
assign (resid 12 and name HB# )(resid 13 and name HN ) 4.5 2.7 2.0
assign (resid 12 and name HB# )(resid 13 and name HN ) 4.5 2.7 2.0
assign (resid 15 and name HB# )(resid 16 and name HN ) 4.5 2.7 2.0
assign (resid 15 and name HB# )(resid 16 and name HN ) 4.5 2.7 2.0
assign (resid 16 and name HB# )(resid 17 and name HN ) 3.5 1.7 2.5
assign (resid 17 and name HB# )(resid 18 and name HN ) 4.5 2.7 2.0
assign (resid 18 and name HB# )(resid 19 and name HN ) 4.5 2.7 2
assign (resid 19 and name HB# )(resid 20 and name HN ) 4.5 2.7 2.0
assign (resid 19 and name HB# )(resid 20 and name HN ) 4.5 2.7 2.0
assign (resid 21 and name HB# )(resid 22 and name HN ) 4.5 2.7 2.0
assign (resid 21 and name HB# )(resid 22 and name HN ) 4.5 2.7 2.0
assign (resid 22 and name HB# )(resid 23 and name HN ) 3.5 1.7 2
assign (resid 23 and name HB# )(resid 24 and name HN ) 4.5 2.7 2.0
assign (resid 23 and name HB# )(resid 24 and name HN ) 4.5 2.7 2.0
assign (resid 24 and name HB# )(resid 25 and name HN ) 4.5 2.7 2.0
assign (resid 25 and name HB# )(resid 26 and name HN ) 4.5 2.7 2.0

!HG(i)-HN(i+1)

assign (resid 4 and name HG# )(resid 5 and name HN ) 4.5 2.7 2.5
assign (resid 8 and name HG# )(resid 9 and name HN ) 4.5 2.7 2.5
assign (resid 10 and name HG1# )(resid 11 and name HN ) 3.5 1.7 1
assign (resid 11 and name HG# )(resid 12 and name HN ) 4.5 2.7 2.5
assign (resid 18 and name HG# )(resid 19 and name HN ) 4.5 2.7 2
assign (resid 20 and name HG# )(resid 21 and name HN ) 4.5 2.7 2.0
assign (resid 22 and name HG# )(resid 21 and name HA ) 4.5 2.7 2.0
assign (resid 22 and name HG# )(resid 21 and name HA ) 4.5 2.7 2.0
assign (resid 22 and name HG# )(resid 23 and name HN ) 3.5 1.7 2
assign (resid 24 and name HG# )(resid 25 and name HN ) 4.5 2.7 2.0
assign (resid 25 and name HG# )(resid 26 and name HN ) 3.5 1.7 2.5

!HD(i)-HN(i+1)

assign (resid 10 and name HD# )(resid 11 and name HN ) 3.5 1.7 2.5
assign (resid 11 and name HD# )(resid 12 and name HN ) 4.5 2.7 2.5

! other
assign (resid 4 and name HN )(resid 5 and name HB# ) 4.5 2.7 2
assign (resid 6 and name HN )(resid 7 and name HB# ) 4.5 2.7 2.0
assign (resid 7 and name HB# )(resid 8 and name HA ) 4.5 2.7 2.0
assign (resid 7 and name HB# )(resid 8 and name HA ) 4.5 2.7 2.0
assign (resid 7 and name HN )(resid 8 and name HB# ) 4.5 2.7 2.0
assign (resid 10 and name HN )(resid 11 and name HG# ) 4.5 2.7 2.5
assign (resid 10 and name HB# )(resid 11 and name HD# ) 4.5 2.7 2.5
assign (resid 10 and name HA )(resid 11 and name HD# ) 4.5 2.7 2.5
assign (resid 10 and name HA )(resid 11 and name HG# ) 4.5 2.7 2.5
assign (resid 11 and name HD# )(resid 12 and name HZ ) 3.5 1.7 2.5
assign (resid 12 and name HZ )(resid 13 and name HN ) 4.5 2.7 2.5
assign (resid 13 and name HN )(resid 14 and name HA ) 4.5 2.7 0.5
assign (resid 15 and name HB# )(resid 16 and name HA ) 4.5 2.7 2.0
assign (resid 15 and name HA )(resid 16 and name HB# ) 4.5 2.7 2.5
assign (resid 17 and name HN )(resid 18 and name HB# ) 4.5 2.7 2.0
assign (resid 19 and name HN )(resid 18 and name HD# ) 4.5 2.7 2.0
```

clas short

!Other(i)-(i+2)

assign (resid 11 and name HN)(resid 13 and name HN) 4.5 2.7 0.5

!Other(i)-(i+3)

assign (resid 20 and name HN)(resid 23 and name HA) 4.5 2.7 0.5
assign (resid 23 and name HA)(resid 26 and name HN) 4.5 2.7 0.5

!Other(i)-(i+4)

assign (resid 7 and name HB#)(resid 11 and name HD#) 4.5 2.7 2.5
assign (resid 16 and name HB#)(resid 20 and name HE#) 4.5 2.7 2.5

clas intra

assign (resid 2 and name HN)(resid 2 and name HA) 4.5 2.7 0.5
assign (resid 3 and name HN)(resid 3 and name HA) 3.5 1.7 0.5
assign (resid 3 and name HN)(resid 3 and name HG#) 3.5 1.7 2.0
assign (resid 3 and name HN)(resid 3 and name HB#) 3.5 1.7 2.0
assign (resid 4 and name HN)(resid 4 and name HA) 3.5 1.7 0.5
assign (resid 5 and name HN)(resid 5 and name HB#) 3.5 1.7 2
assign (resid 6 and name HN)(resid 6 and name HA) 3.5 1.7 0.5
assign (resid 6 and name HN)(resid 6 and name HB#) 3.5 1.7 2.0
assign (resid 6 and name HN)(resid 6 and name HB#) 3.5 1.7 2.0
assign (resid 7 and name HN)(resid 7 and name HA) 4.5 2.7 0.5
assign (resid 7 and name HN)(resid 7 and name HB#) 4.5 2.7 2.0
assign (resid 7 and name HN)(resid 7 and name HB#) 4.5 2.7 2
assign (resid 8 and name HN)(resid 8 and name HG#) 3.5 1.7 2
assign (resid 8 and name HN)(resid 8 and name HB#) 3.5 1.7 2.0
assign (resid 9 and name HN)(resid 9 and name HB#) 3 1.2 2.0
assign (resid 9 and name HN)(resid 9 and name HA) 3.5 1.7 0.5
assign (resid 10 and name HN)(resid 10 and name HA) 3.5 1.7 0.5
assign (resid 10 and name HN)(resid 10 and name HB#) 4.5 2.7 2.0
assign (resid 10 and name HN)(resid 10 and name HG1#) 4.5 2.7 2
assign (resid 10 and name HN)(resid 10 and name HD#) 3.5 1.7 2.5
assign (resid 11 and name HN)(resid 11 and name HA) 4.5 2.7 0.5
assign (resid 11 and name HN)(resid 11 and name HG#) 4.5 2.7 2
assign (resid 11 and name HN)(resid 11 and name HD#) 4.5 2.7 2.5
assign (resid 11 and name HN)(resid 11 and name HD#) 4.5 2.7 2.5
assign (resid 12 and name HN)(resid 12 and name HA) 4.5 2.7 0.5
assign (resid 12 and name HN)(resid 12 and name HB#) 4.5 2.7 2.0
assign (resid 12 and name HN)(resid 12 and name HB#) 4.5 2.7 2.0
assign (resid 12 and name HN)(resid 12 and name HZ) 3.5 1.7 2.5
assign (resid 13 and name HN)(resid 13 and name HA#) 3.5 1.7 0.5
assign (resid 13 and name HN)(resid 13 and name HA#) 3.5 1.7 0.5
assign (resid 14 and name HN)(resid 14 and name HD2#) 4.5 2.7 2.5
assign (resid 14 and name HN)(resid 14 and name HB#) 3.5 1.7 2
assign (resid 14 and name HN)(resid 14 and name HD2#) 4.5 2.7 2.5
assign (resid 15 and name HN)(resid 15 and name HA) 4.5 2.7 0.5
assign (resid 15 and name HN)(resid 15 and name HB#) 4.5 2.7 2.0
assign (resid 15 and name HN)(resid 15 and name HB#) 4.5 2.7 2.0
assign (resid 16 and name HN)(resid 16 and name HA) 3.5 1.7 0.5
assign (resid 16 and name HN)(resid 16 and name HB#) 3 1.2 1.5
assign (resid 17 and name HN)(resid 17 and name HA) 3.5 1.7 0.5
assign (resid 17 and name HN)(resid 17 and name HB#) 3.5 1.7 2.0
assign (resid 17 and name HN)(resid 17 and name HG#) 4.5 2.7 2.5
assign (resid 18 and name HN)(resid 18 and name HG#) 3.5 1.7 2.0
assign (resid 18 and name HN)(resid 18 and name HA) 3.5 1.7 0.5
assign (resid 18 and name HN)(resid 18 and name HB#) 3.5 1.7 2.0
assign (resid 19 and name HN)(resid 19 and name HA) 4.5 2.7 0.5
assign (resid 19 and name HN)(resid 19 and name HB#) 4.5 2.7 2.0
assign (resid 19 and name HN)(resid 19 and name HB#) 4.5 2.7 2.0
assign (resid 20 and name HN)(resid 20 and name HD#) 4.5 2.7 2.0
assign (resid 20 and name HN)(resid 20 and name HB#) 3.5 1.7 2.0
assign (resid 20 and name HN)(resid 20 and name HG#) 4.5 2.7 2.0
assign (resid 20 and name HN)(resid 20 and name HA) 3.5 1.7 0.5
assign (resid 21 and name HN)(resid 21 and name HA) 4.5 2.7 0.5
assign (resid 21 and name HN)(resid 21 and name HB#) 4.5 2.7 2.0
assign (resid 21 and name HN)(resid 21 and name HB#) 4.5 2.7 2.0
assign (resid 22 and name HN)(resid 22 and name HA) 4.5 2.7 0.5
assign (resid 22 and name HN)(resid 22 and name HG#) 3.5 1.7 2.0
assign (resid 22 and name HN)(resid 22 and name HB#) 3.5 1.7 2.0
assign (resid 23 and name HN)(resid 23 and name HA) 4.5 2.7 0.5
assign (resid 23 and name HN)(resid 23 and name HB#) 4.5 2.7 2.0
assign (resid 23 and name HN)(resid 23 and name HB#) 4.5 2.7 2.0
assign (resid 24 and name HN)(resid 24 and name HG#) 4.5 2.7 2.0
assign (resid 24 and name HN)(resid 24 and name HB#) 3.5 1.7 0.5
assign (resid 25 and name HN)(resid 25 and name HA) 3.5 1.7 0.5
assign (resid 25 and name HN)(resid 25 and name HB#) 4.5 2.7 2.0
assign (resid 25 and name HN)(resid 25 and name HG#) 3.5 1.7 2.5
assign (resid 26 and name HN)(resid 26 and name HG#) 4.5 2.7 2.0
assign (resid 26 and name HN)(resid 26 and name HB#) 4.5 2.7 2.0
assign (resid 26 and name HN)(resid 26 and name HB#) 4.5 2.7 2.0
assign (resid 26 and name HN)(resid 26 and name HA) 3.5 1.7 0.5

Appendix V CCL11 Purification

Overview

CCL11 was isolated via standard protein purification techniques. Once the protocol had been established the amount of protein produced was too low to enable the incorporation of the ^{15}N label. Reinserting the DNA coding for CCL11 into fresh plasmid and changing the cell line resulted in increased yield which allowed ^{15}N labelling: however previously unseen protein purification problems resulted in the production of too little material to enable further analysis in the time available.

Introduction

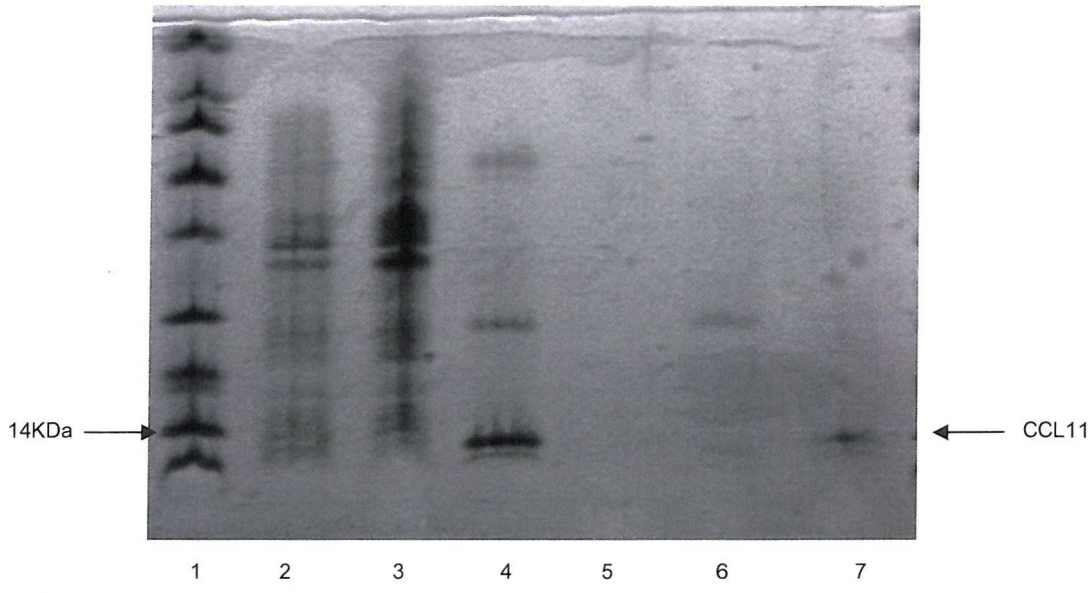
CCL11 was identified as an attractant for eosinophils in the bronchoalveolar lavage fluid of actively sensitised guinea-pigs after aerosol allergen challenge [186]. This led to a cDNA probe to find the human version of CCL11 [63]. CCL11 is considered to be a very important chemokine in allergic and eosinophilic inflammation, with particular interest shown in its' regulation of asthma [187]. The solution structure of CCL11 has been solved and the backbone dynamics studied by NMR spectroscopy [20, 188].

CCL11 was considered to be a selective ligand of CCR3 and to have antagonist properties to CXCR3 [189], but there are conflicting reports which suggested CCL11 acts as an antagonist for CCR2 and an agonist for CCR5 [190], whilst other have reported CCL11 as a partial agonist to CCR2 and as having no effect on CCR5 [191].

Growth and Purification

The plasmid used was Novagen pET-15b with the DNA sequence coding for CCL11 inserted between the Nde-1 and BamH1 cut sites. This plasmid had previously been shown to yield 7 mg/L of unlabelled protein [20]. The protein was overexpressed in *E.coli* BL21(DE3) and purification was attempted by applying the crude soluble protein to a nickel affinity column because the plasmid coded for a hexa-His tag. A nickel affinity column should bind a His tagged protein, which could then be eluted using imidazole. An FPLC was used to load, wash and elute from the column. The fractions containing the eluted protein were combined and concentrated before being loaded onto a Superdex 75 gel filtration column. Figure V.1 shows the SDS-PAGE analysis of the major steps of the purification. Unfortunately the yield was only 0.5mg/L even after attempting similar steps

to those taken in chapter 4. This yield was too low for the minimal media growth process required for the production of ^{15}N-labelled CCL11.



Lane	Contents
1	Markers
2	Uninduced total protein.
3	Induced total protein.
4	Elution from the nickel column (concentrated)
5	First peak from the gel filtration column
6	Second peak from the gel filtration column
7	Third peak from the gel filtration column
8	Markers

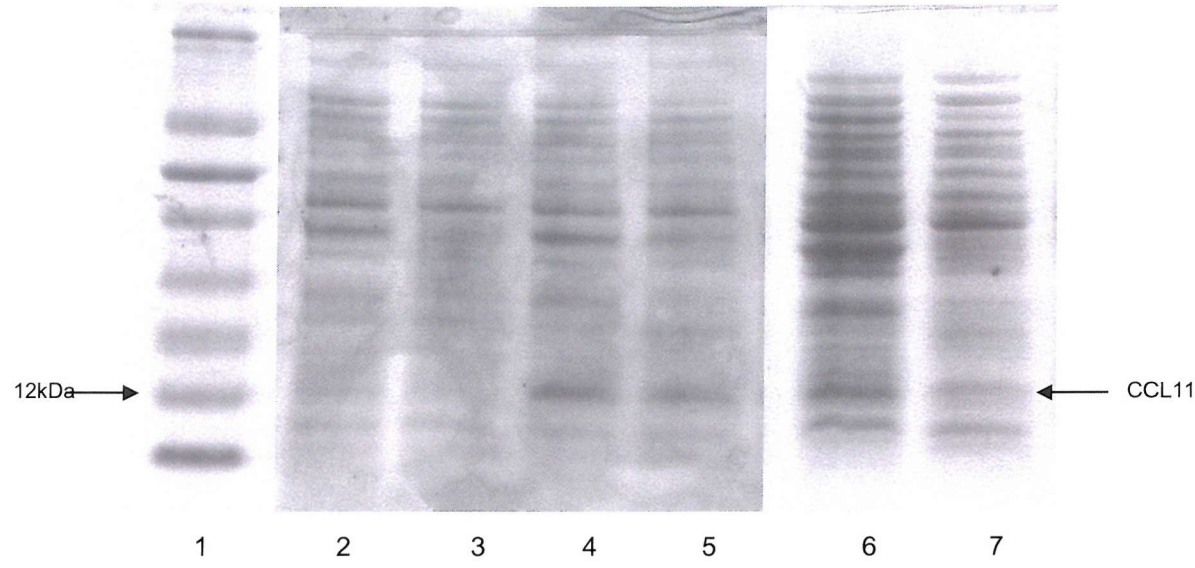
Figure V.1 SDS-PAGE analysis showing the purification stages of CCL11.

Changing the Plasmid and Cell Line

The low yield of CCL11 compared with the previously published results from this plasmid [188] suggested that there might be a problem with the promoter region of the plasmid; as such the DNA coding for CCL11 was cut out and reinserted into fresh plasmid. The overexpression band was greatly improved, but it was noted that the codon usage of the CCL11 sequence was not suited to *E.Coli* (table V.1). The rosetta strain has been modified for the rare codon usage of arginine (AGG, AGA, CGA), leucine (CTA), isoleucine (ATA) and proline (CCC) and when used in this case resulted in a significantly improved overexpression band, therefore this strain was used for subsequent growths (Figure V.2).

Number		1	2	3	4	5	6	7	8	9	10	11	12
Residue		G	P	A	S	V	P	T	T	C	C	F	N
Codon		GGG	CCA	GCT	TCT	GTC	CCA	ACC	ACC	TGC	TGC	TTT	AAC
13	14	15	16	17	18	19	20	21	22	23	24	25	26
L	A	N	R	K	I	P	L	Q	R	L	E	S	Y
CTG	GCC	AAT	**AGG**	AAG	**ATA**	**CCC**	CTT	CAG	**CGA**	**CTA**	GAC	AGC	TAC
27	28	29	30	31	32	33	34	35	36	37	38	39	40
R	R	I	T	S	G	K	C	P	Q	K	A	V	I
AGG	**AGA**	ATC	ACC	AGT	GGC	AAA	TGT	**CCC**	AAA	CAG	GCT	GTG	ATC
41	42	43	44	45	46	47	48	49	50	51	52	53	54
F	K	T	K	L	A	K	D	I	C	A	D	P	K
TTC	AAG	ACC	AAA	CTG	GCC	AAG	GAT	ATC	TGT	GCC	GAC	**CCC**	AAG
55	56	57	58	59	60	61	62	63	64	65	66	67	68
K	K	W	V	Q	D	S	M	K	Y	L	D	Q	K
AAG	**AGG**	TGG	GTG	CAG	GAT	TCC	ATG	AAG	TAT	CTG	GAC	CAA	AAA
69	70	71	72	73	74								
S	P	T	P	K	P								
TCT	CCA	ACT	CCA	AAG	CCA	TAA							

Table V.1 DNA sequence coding for CCL11 showing the rare codon usage in red.



Lane	Contents
1	Markers.
2	Uninduced total protein.
3	Uninduced soluble protein.
4	Induced total protein for the rosetta strain.
5	Induced soluble protein for the rosetta strain.
6	Induced total protein for the BL21(DE3).
7	Induced soluble protein for the BL21(DE3).

Figure V.2 SDS-PAGE analysis of the induction band from the rosetta strain versus BL21(DE3).

New Preparation

An unlabelled growth of CCL11 was purified following the methodology published by Ye et al [111] which used a nickel chelating column, followed by a strong cationic exchanger. This resulted in a yield of about 10mg/mL which was suitable for expression in minimal media. The protein was ^{15}N-labelled using the standard techniques and 10mg of CCL11 was purified and stored lyophilized at 4°C whilst awaiting study. Unfortunately there was a long delay and when the sample was finally examined using NMR spectroscopy significant levels of various impurities were detected. A new growth was initiated as it was suspected that these impurities were a degradation product resulting from the long storage. The pellet obtained from 4L of growth media was purified in two batches, back to back, on the nickel chelating column. The first batch was as expected, but the second batch showed a marked decrease in yield, with a significant amount not binding to the column. As these two samples were from the same growth it was reasoned that the problem must lie with the nickel chelating column. When a 1M imidazole wash failed to yield any more protein, an EDTA strip of the column was undertaken. The SDS PAGE analysis of this showed the presence of a large quantity of protein. The nickel affinity column was recharged with nickel and the void volume from the previous run was reloaded before the column was restripped with EDTA. The combined material from all these runs was concentrated and desalted before being loaded onto the strong cationic exchanger. The SDS-PAGE analysis (figure V.3) clearly shows the co-elution of CCL11 with impurities, supporting the theory that CCL24 could not be purified from the solution phase due to the interaction between it and unknown components of the lysate. The CCL11 peak that corresponded to the published salt elution value was pure by SDS-PAGE analysis, so it was desalted and examined by NMR spectroscopy. It was found to contain the same impurities as the original sample.

All the material obtained from the strong cationic exchanger was concentrated, desalted and loaded onto a HPLC which showed both the oxidised and reduced forms of CCL11 were present, so all the material was concentrated by lyposisation and subject to oxido refolding before desalting, lyposation and examination by NMR spectroscopy. Whilst the material was free of impurities, it was in the fully reduced form due to the oxidized-glutathione being off. The oxido refolding was repeated and the sample desalted and lyposized. Only 2mg of material was obtained which was not enough for any of the titrations.

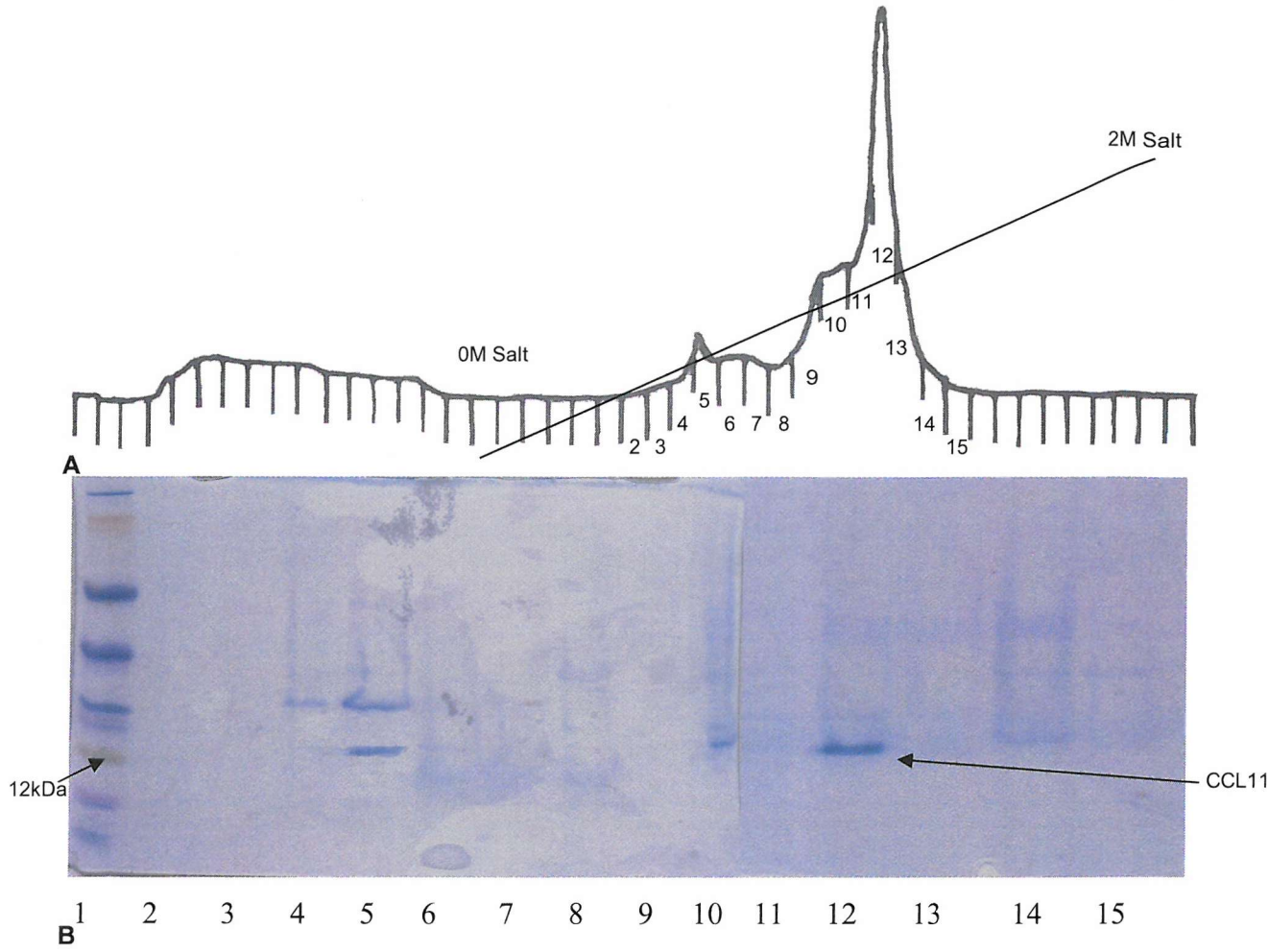


Figure V.3 A. FPLC trace of the absorbance at 280nm for a strong cationic exchanger, when the concentrated factions from the nickel column were loaded and a salt gradient run. CCL11 appears to co-elute with the other contaminating proteins. The fraction labels correspond to the lanes in Figure B. B. SDS-PAGE analysis of Figure A. Lane 1 shows the markers.

Appendix VI Assignment of *act*-ACP Mutants NMR Spectra

Polyketides are found in a variety of organisms and are a chemically diverse range of compounds despite being made from the same basic repeating unit. Polyketides are produced via a series of decarboxylation condensations, followed by cycles of different reactions that vary depending upon the final product.

There are three major types of polyketide synthases. Type I systems are found in higher animals and fungi and consist of large multifunctional enzymes. Each enzymatic reaction has its own distinct domain and these individual domains are joined together by peptide linkers [192]. Type II systems are found in prokaryotes and consist of discrete enzymes [193, 194]. Type III systems are similar to type I systems but are found in plants and lack an acyl carrier protein [195].

One of the most studied type II polyketide systems is that of the antibiotic benzoisochromanequinone actinorhodin (*act*) from *Streptomyces coelicolor* [193] (Figure VI.1A). One of the proteins involved in the early stages of *act* biosynthesis is the acyl carrier protein (ACP). During polyketide biosynthesis the growing polypeptide chain is tethered via a thioester linkage to the sulfhydryl group of 4'-phosphopantetheine cofactor that is covalently linked to a conserved serine (S42 for *act*-ACP). *Act*-ACP is converted from the inactive apo form to its holo form by holo-acyl carrier protein synthase (ACPS) (Figure VI.1B) [196, 197]. Polyketide (and fatty acid) biosynthesis is primed by malonyl transfer to the ACP in preparation for ketosynthase catalysed decarboxylative condensation with an acetate starter unit.

In fatty acid biosynthesis a specific enzyme, malonyl CoA:holo-acyl carrier protein transacylases (MCAT) catalyses the transfer of malonyl CoA to the ACP to form malonyl-ACP (Figure VI.1C). However, type II polyketide synthases lack an obvious MCAT gene in their biosynthetic cluster giving rise to the notion of "crosstalk" where the fatty acid synthase MCAT (FAS-MCAT) attaches malonate to the holo-*act*-ACP *in vivo*. However, it has been observed that holo-*act*-ACP can undergo an autocatalytic "self-malonylation" without the need for an MCAT enzyme. Self-malonylation is a contentious issue with some claiming it is due to the co-purification of MCAT [198].

The holo-*act*-ACP has also been shown to be able to transfer malonate to holo-FAS-ACP, whilst the apo-*act*-ACP was unable to perform this suggesting that this is not related to a co-purified MCAT but is actually a property of holo-*act*-ACP [199].

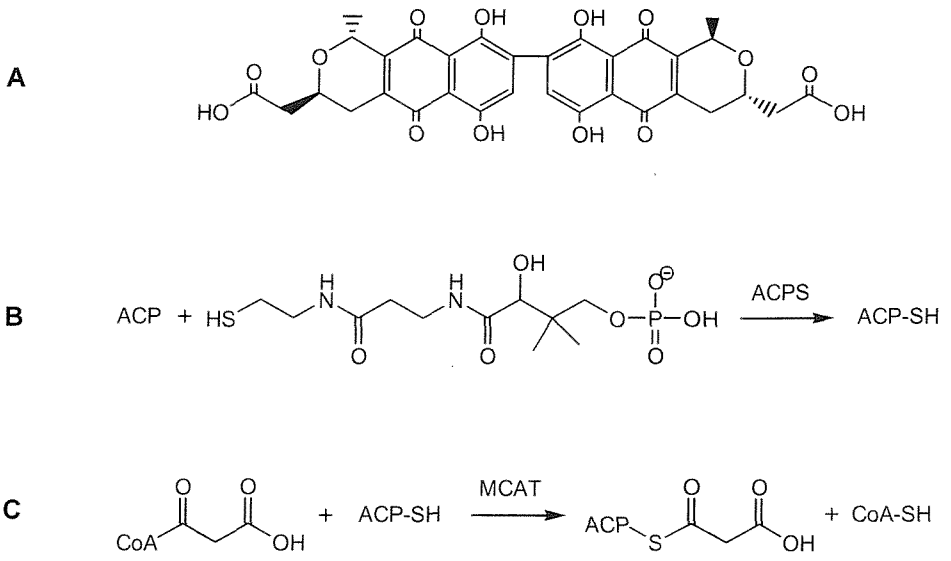
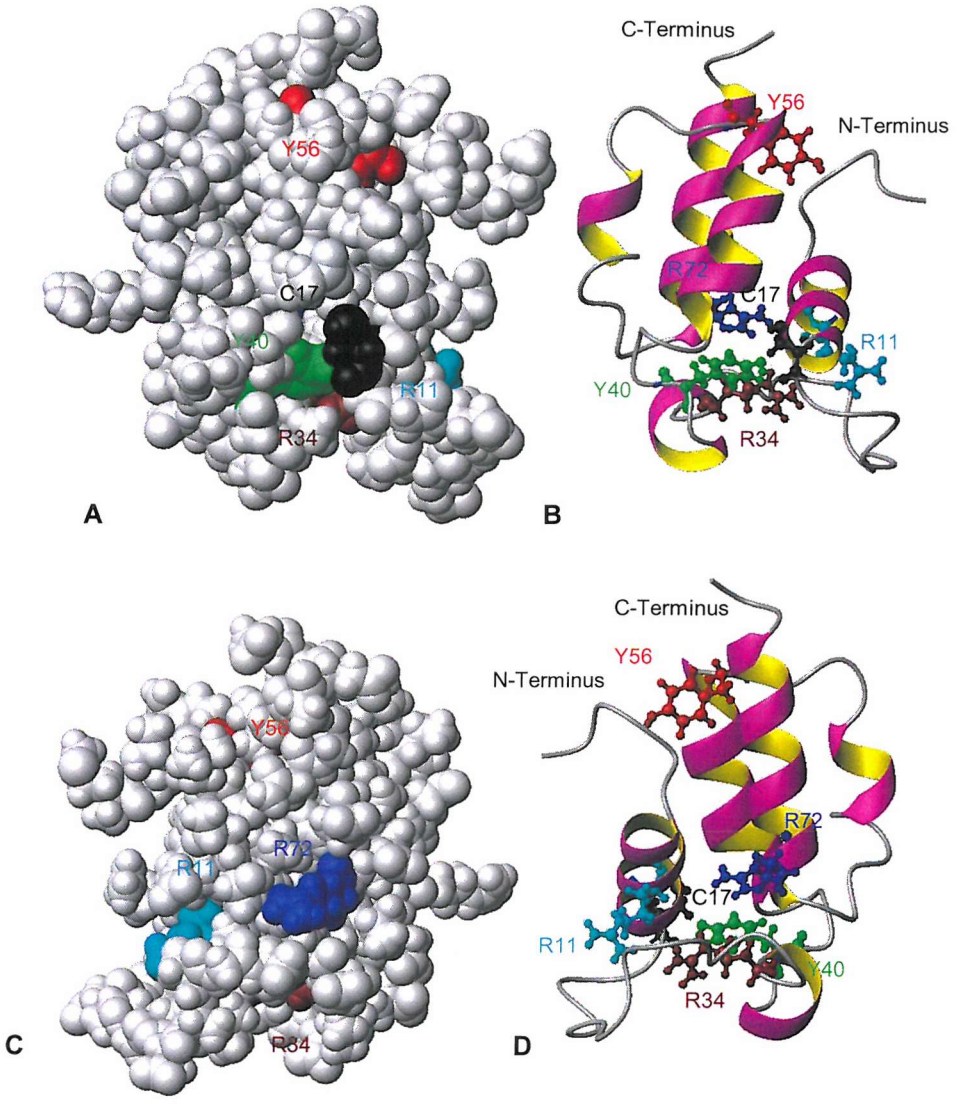


Figure VI.1 A. Structure of benzoisochromanequinone actinorhodin. B. The conversion of apo-ACP to holo-ACP via the addition of a phosphopantetheine arm by ACPS. C. The transfer of malonyl CoA by MCAT to the phosphopantetheine arm on holo-ACP.

To further study these issues a series of *act*-ACP mutants were made and their activities measured. The residues mutated were R11, R34, Y40, Y56 and R72 (Figure VI.2). Y40 and Y56 were chosen because highly conserved tyrosines are used by acyl CoA binding proteins (ACBPs) in the interaction with the CoA portion of acyl CoA thioesters [200]. A sequence alignment showed Y40 and Y56 to be highly conserved in the type II polyketide ACPs, whilst in type II FAS-ACPs Y40 is missing and Y56 is replaced with phenylalanine or lysine. R11, R34 and R72 were chosen because an arginine residue had been shown to participate in the substrate selection of the MCAT domain for rat-FAS and intermediate binding and stabilization of CoA in other enzymes [201]. A sequence alignment showed R11, R34 and R72 to be highly conserved in the type II polyketide ACPs and absent in type II FAS-ACPs.



Black: C17.

Blue: R72.

Brown: R34.

Cyan: R11.

Green:Y40.

Red: Y56.

Figure VI.2 *act*-ACP showing the residues mutated in the study. B. Figure A in ribbon. C. Figure A rotated 180°. D. Figure C in ribbon.

These five residues were individually mutated to alanine and their activity assayed for both self-malonylation and transfer to holo-*Streptomyces coelicolor*-FAS-ACP (holo-Sc-FAS-ACP) [199]. Figure VI.3 shows the results achieved and whilst the lack of error bars makes it difficult to draw conclusions Figure VI.3A and Figure VI.3B would suggest that the mutations R72A and Y56A severely affected the transfer to holo-Sc-FAS-ACP, although the graph suggests that the reaction was still proceeding albeit at a much reduced rate. Figure VI.3C and Figure VI.3D seem to show that R72A had a greater effect on self-

malonylation than Y56A, with the rate appearing to have slowed in both cases. The transfer to holo-Sc-FAS-ACP appears to be more severely affected than self-malonylation although the assay for the transfer to holo-Sc-FAS-ACP could require the holo-*act*-ACP to first self-malonylate.

These experiments seem to suggest that self-malonylation is a real effect, although it is worth noting that the plasmids and bacterial strains were not the same in all cases and therefore the results do not rule out an impurity being co-purified.

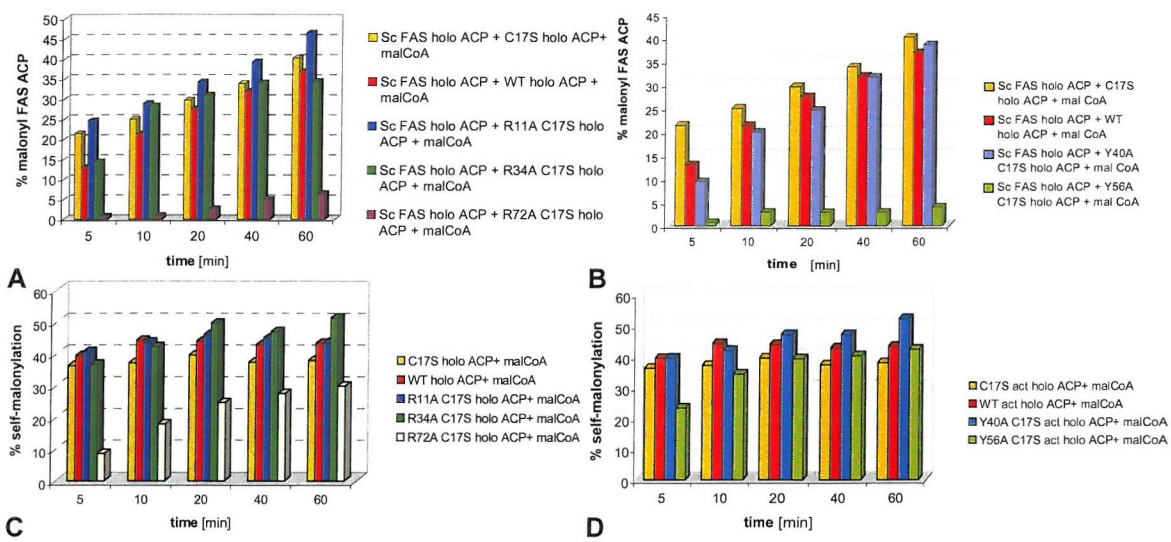


Figure VI.3 A. Percentage malonyl transfer to holo-Sc-FAS-ACP using malonyl CoA and a variety of holo-act-ACPs. The graph shows wild type, C17S and the double mutants R11A C17S, R34A C17S and R72A C17S. B. Same as A but showing wild type, C17S and the double mutants Y40A C17S and Y56A C17S. C. Percentage self-malonylation using malonyl CoA and a variety of holo-act-ACPs. The graph shows wild type, C17S and the double mutants R11A C17S, R34A C17S and R72A C17S. D. Same as B but showing wild type, C17S and the double mutants Y40A C17S and Y56A C17S. The graphs were taken from reference [199].

The wild type structure had already been determined by NMR spectroscopy [202, 203] and the mutants R72A and Y56A were studied by NMR spectroscopy to see whether their change in function was related to a structural change.

Results and Discussion

TOCSY and NOESY (mixing time of 0.15s) spectra were recorded for both the R72A and Y56A mutant. The spectra for both were of comparable quality. The R72A spectra correlated well with the native data, however the Y56A spectra had a smaller shift dispersion that made the assignment problematic, as there were many overlapping peaks. This repeatedly created ambiguous assignments which meant that large sections of the protein had to be tentatively assigned before the assignment could be proved. This is in contrast to the R72A spectra where the sequential method worked in more of a text book manner and proved to be relatively straightforward to assign.

The assigned amide regions of the TOCSY for both R72A and Y56A are shown in Figure VI.5 and Figure VI.6 respectively and the chemical shift data of the two proteins is given in the appendix.

When a protein structure has been solved by NMR spectroscopy there are normally several significant NOEs that define the overall fold of the protein. Rather than generate structures for each mutant it was possible to examine their NOESY spectra for the significant NOEs of the wild type. This gave an indication of the probability of the mutants having the same protein structure as the wild type (Figure VI.4). R72A (Figure VI.4B) produced a nearly identical NOE pattern compared to the wild type apo-*act*-ACP and therefore it was likely to be adopting the same structure. Y56A (Figure VI.4C) was missing some key long range NOEs due to the mutation itself. However with the long range NOE from L31 to R72 still apparent it seemed that the global folding was still intact, but no firm conclusion could be drawn.

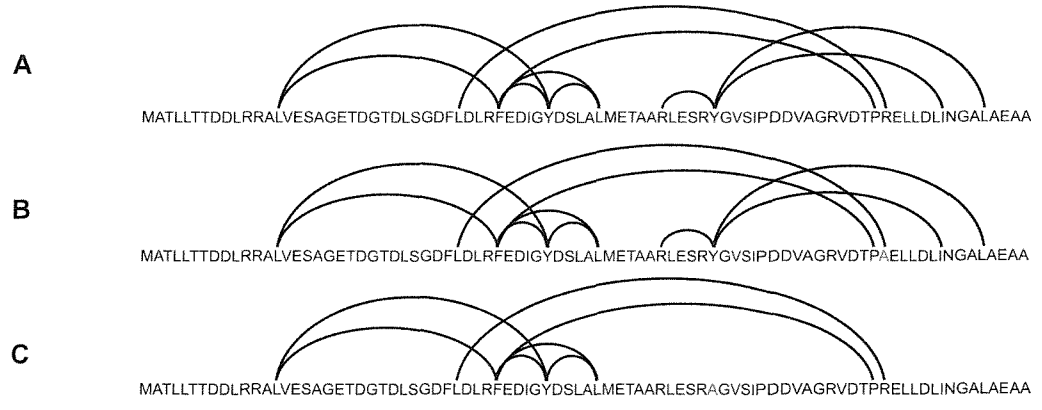


Figure VI.4 Schematic of the NOEs that are important for the overall structure of the apo-act-ACP. A. The wild type. B. R72A (blue). C. Y56A (red).

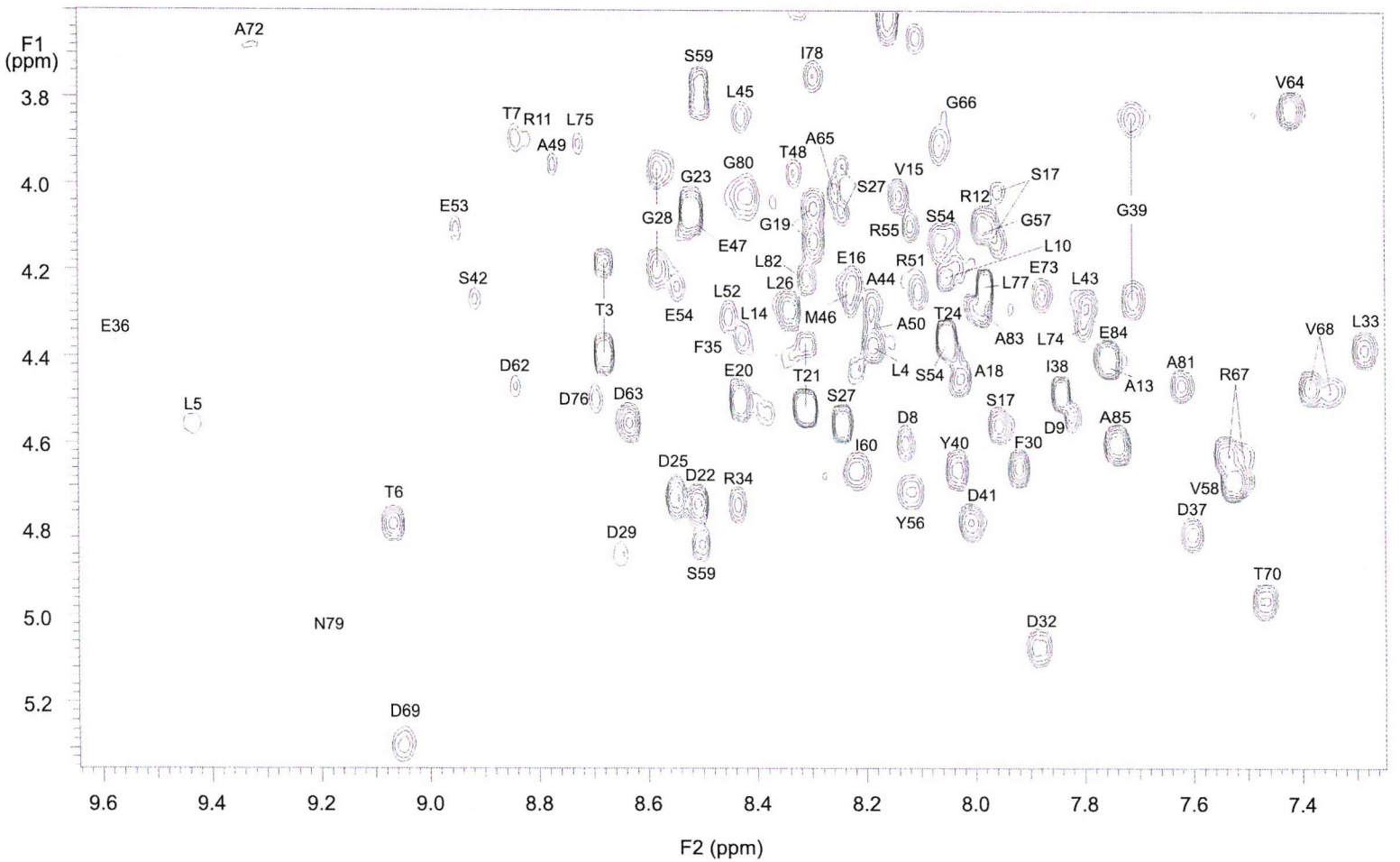


Figure VI.5 The TOCSY spectrum for R72A at 25°C showing the region between 7.25 to 9.65ppm and 3.5 to 5.35ppm. R67 and V68 both seem to be sampling two different confirmations. The spectrum was zero filled from 4096 to 8192 points in F2 and zero filled from 512 to 2048 points in F1. The Varian window functions were LB (-3.45), GF (0.055), LB1(-10) and GF1(0.027).

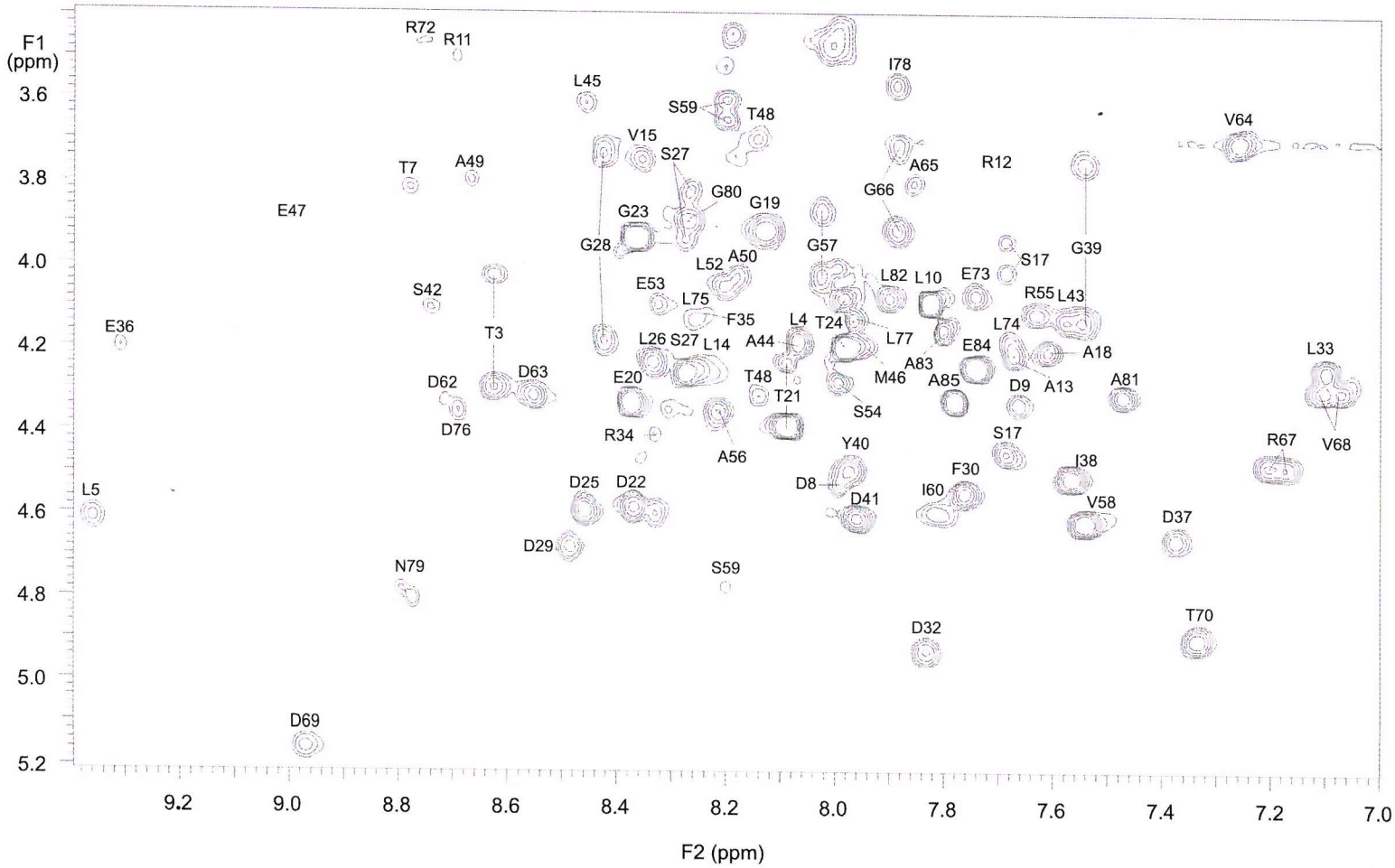


Figure VI.6 The TOCSY spectrum for Y56A at 25°C showing the region between 7 to 9.4ppm and 3.5 to 5.2ppm. R67 and V68 both seem to be sampling two different confirmations. The spectrum was zero filled from 4096 to 8192 points in F2 and zero filled from 400 to 2048 points in F1. The Varian window functions were LB (-3.37), GF (0.083), LB1(-10) and GF1(0.020).

The chemical shift index of the wild type (Figure VI.7A) showed that this technique was not foolproof with helix 2 not being predicted, helix 3 being predicted to be too long, and helix 5 too short. The chemical shift index of R72A (Figure VI.7B) was identical to the wild type suggesting that they were adopting a very similar if not identical structure. Comparison of the wild type to Y56A (Figure VI.7C) showed some differences, particularly around the C-terminal ends of helices 1 and 5 suggesting there might have been structural changes at these sites which are in close proximity to Y56A.

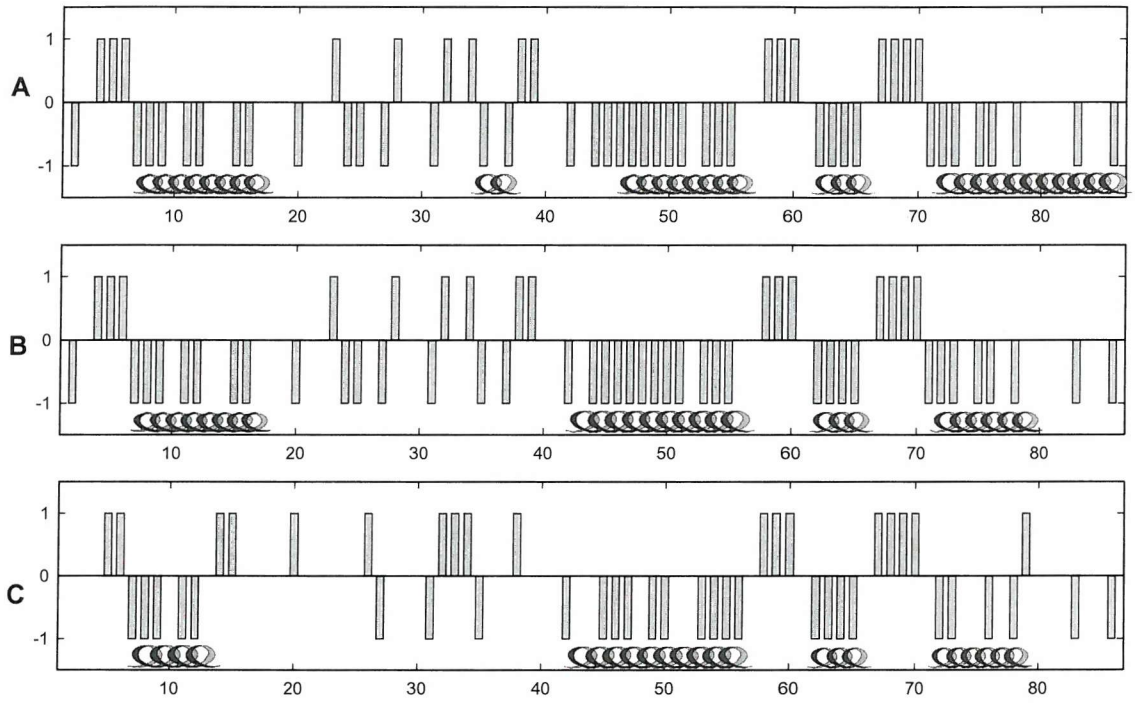


Figure VI.7 The chemical shift index of: A. Wild type. B. R72A C. Y56A. The alpha helices as determined from the published structure are shown on the wild type: T7-E16, F35-I38, L45-R55, P61-V64 and P71-A83. The helices shown on R72A and Y56A are those that would be suggested by the chemical shift index.

For both mutants the chemical shifts of the amide and alpha hydrogens were compared to the wild type (Figures VI.8 and VI.9). Whilst most residues were within 0.1ppm of the wild type, several showed greater deviation and those that had deviations greater than 0.2ppm are displayed upon the protein (Figures VI.10 and VI.11). As both mutations have similar effects but are at opposite ends of the protein structure it is possible that those residues affected by both mutations are important contributors to the observed effects. Figure VI.12 shows that three out of the four residues cluster, suggesting that the activity might be mediated by R11, E16 and T48. As R11 was one of the mutants made and did not have any effect on the activity it is likely that either E16 or T48 (or both) are mediating

the behaviour. It is noted at this point that both mutants have two chemical shifts for R67 and V68 which might be the cause of the behaviour.

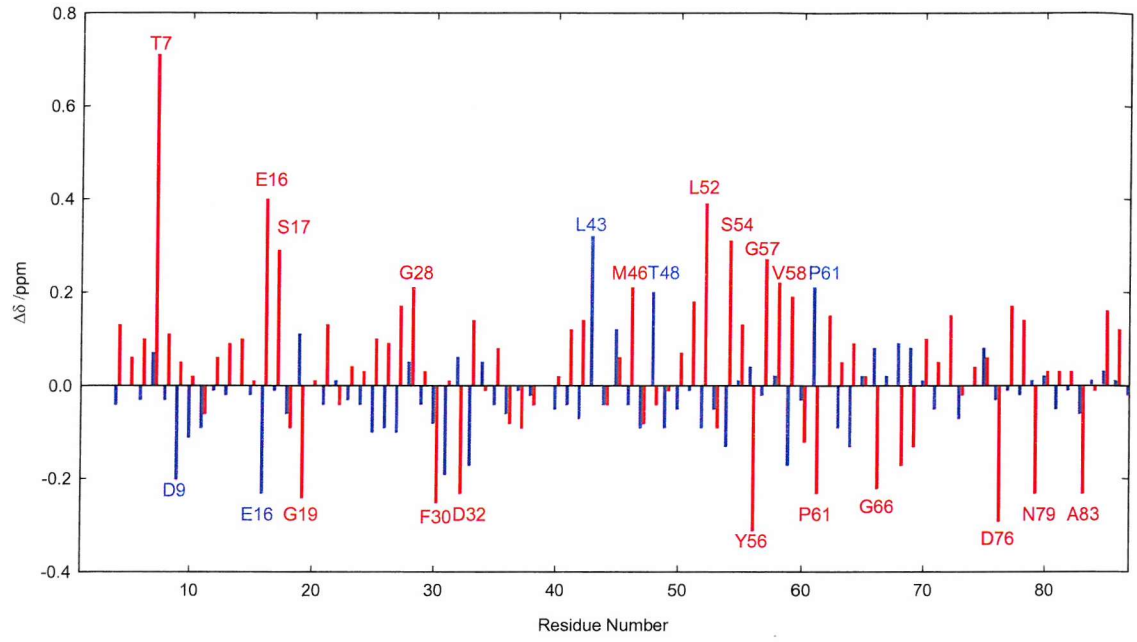


Figure VI.8 The chemical shift deviation (Δδ) of the amide hydrogen from the wild type for R72A in blue and Y56A in red. The marked residues are those with a deviation greater than 0.2ppm.

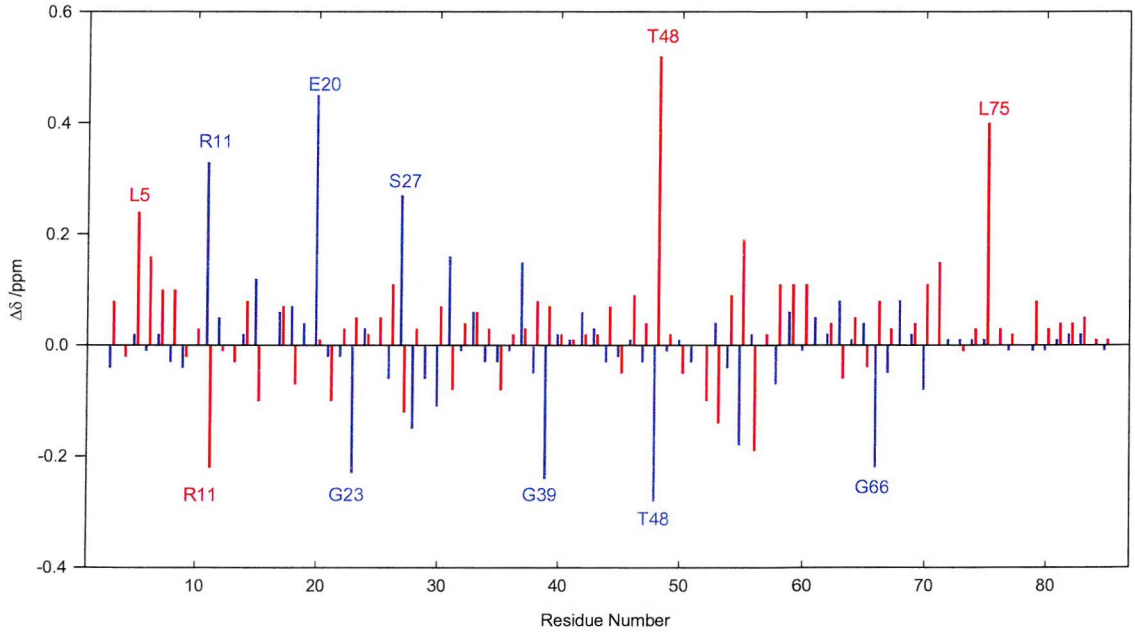
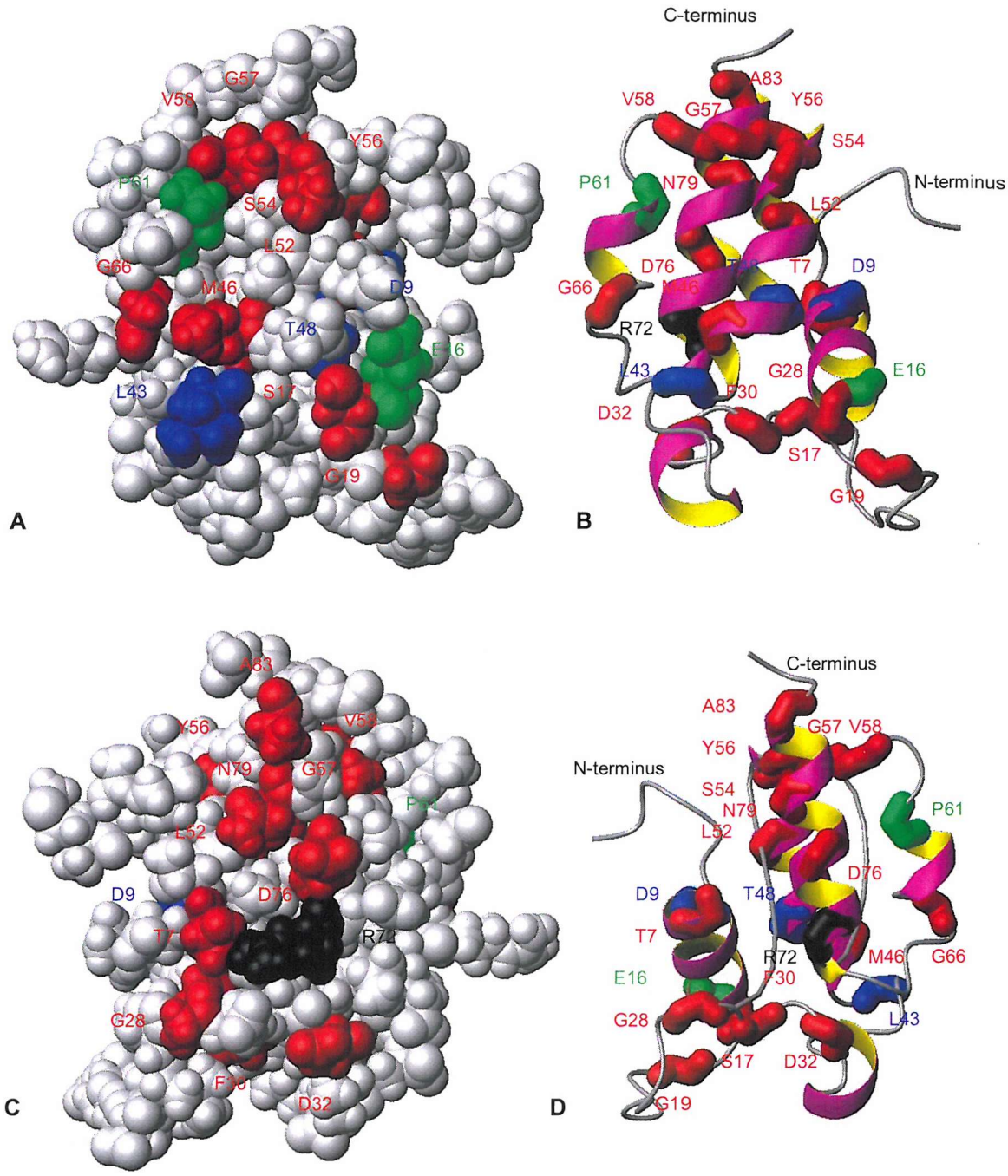


Figure VI.9 The chemical shift deviation (Δδ) of the alpha hydrogen from the wild type for R72A in blue and Y56A in red. The marked residues are those with a deviation greater than 0.2ppm.



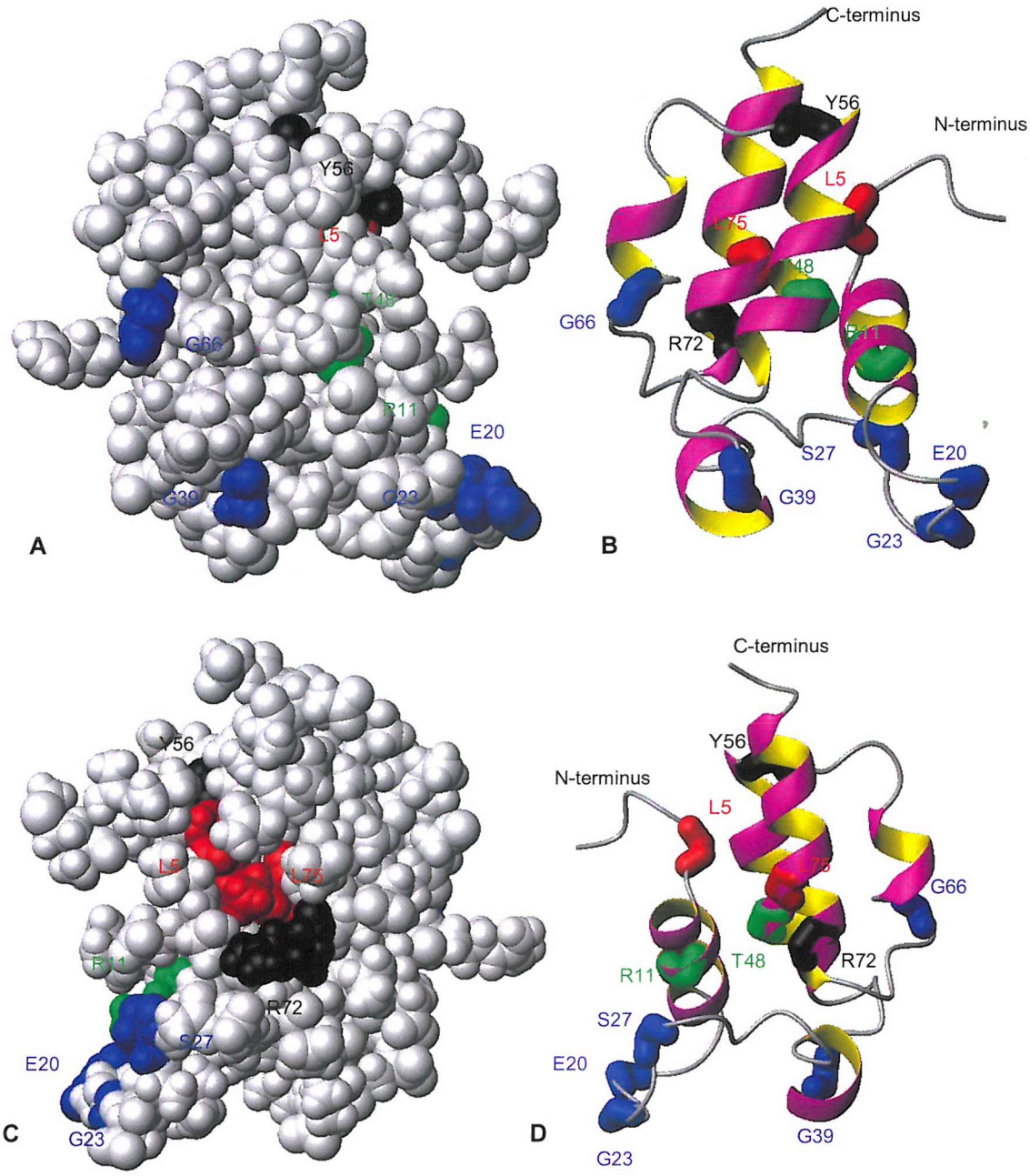
Black: R72 for structural reference.

Blue: Residues affected in R72A only (D9, L43 and T48).

Green: Residues affected in both mutants (E16 and P61).

Red: Residues affected in Y56A only (T7, S17, G19, G28, F3O, D32, M46, L52, S54, Y56, G57, V58, G66, D76, N79 and A83).

Figure VI.10 Residues whose amide hydrogen showed a significant chemical shift compared with wild-type act ACP. B. Figure A in ribbon. C. Figure A rotated by 180°. D. Figure C in ribbon.



Black: Y56 and R72 for structural reference.

Blue: Residues affected in R72A only (E20, G23, S27, G39 and G66).

Green: Residues affected in both mutants (R11 and T48).

Red: Residues affected in Y56A only (L5 and L75).

Figure VI.11 A. Residues whose alpha hydrogen showed a significant chemical shift compared with wild-type act ACP. B. Figure A in ribbon. C. Figure A rotated by 180°. D. Figure C in ribbon.

To try to understand this result a sequence alignment of all known and putative ACPs was made, after the removal of several sequences that did not align the active site serine (to which the phosphopantetheine arm attaches) which were assumed to be misclassified. Most of the known sequences are from the fatty acid biosynthesis pathway, so the low-scoring segments in the polyketide ACPs were those that were unique to the polyketides and therefore might be responsible for the unique properties of polyketide ACPs (Figure VI.13). These residues did not include E18, but did included T48 and are shown for apo-*act*-ACP (identified as ACPX_STRCO on the sequence alignment) in Figure VI.15. A realignment of all those sequences (Figure VI.14) that lay between the most outlaying known polyketide ACPs (identified as ACPX_STRAW and ACPX_STRHA on the sequence alignment) showed that the threonine was not fully conserved although there was normally a threonine or serine within two residues but the threonine is part of an alpha-helix so this could point the side chain in a totally different direction.

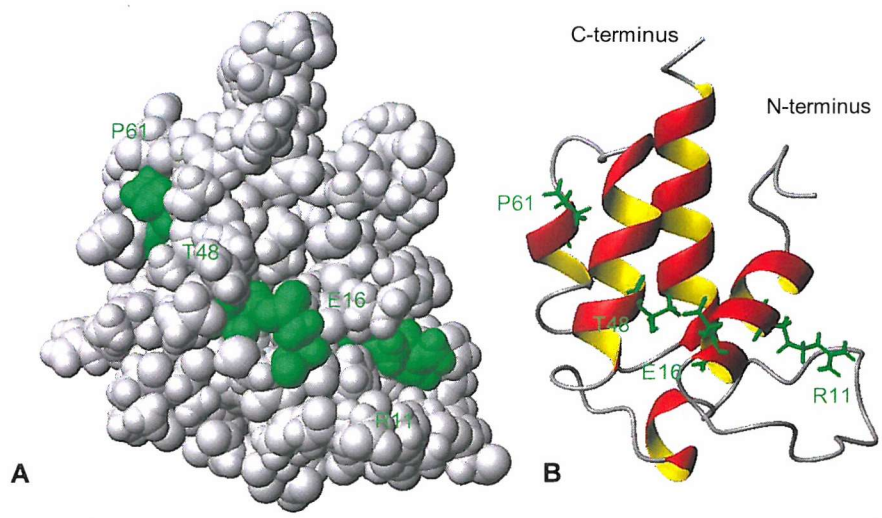


Figure VI.12 Residues affected by both R72A and Y56A. A. Space Filled. B. Ribbon.

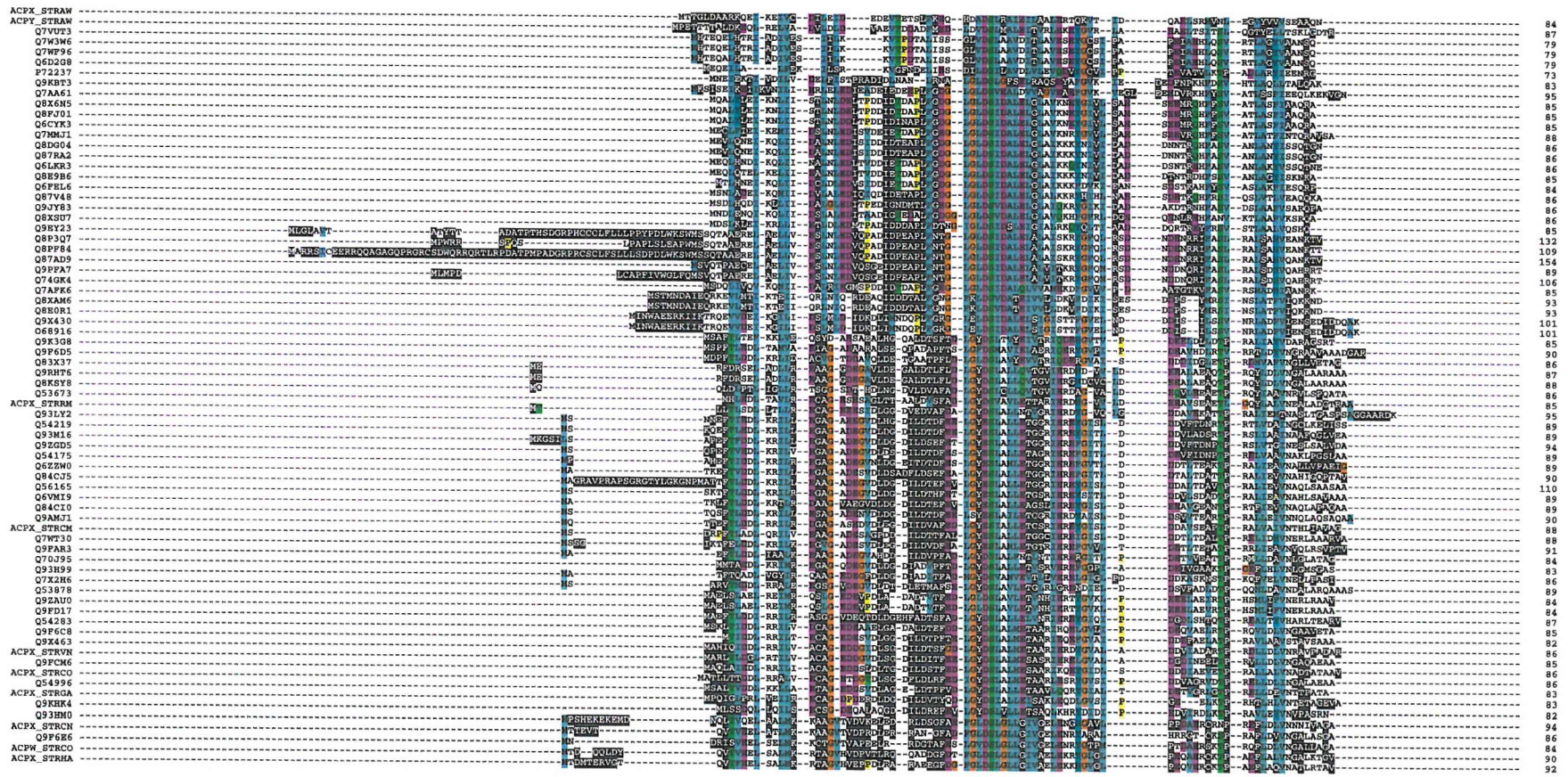


Figure VI.13 Sequence alignment showing all the polyketide ACP sequences when aligned against the FAS-ACP sequences. The areas highlighted in black are those that are low scoring segments.

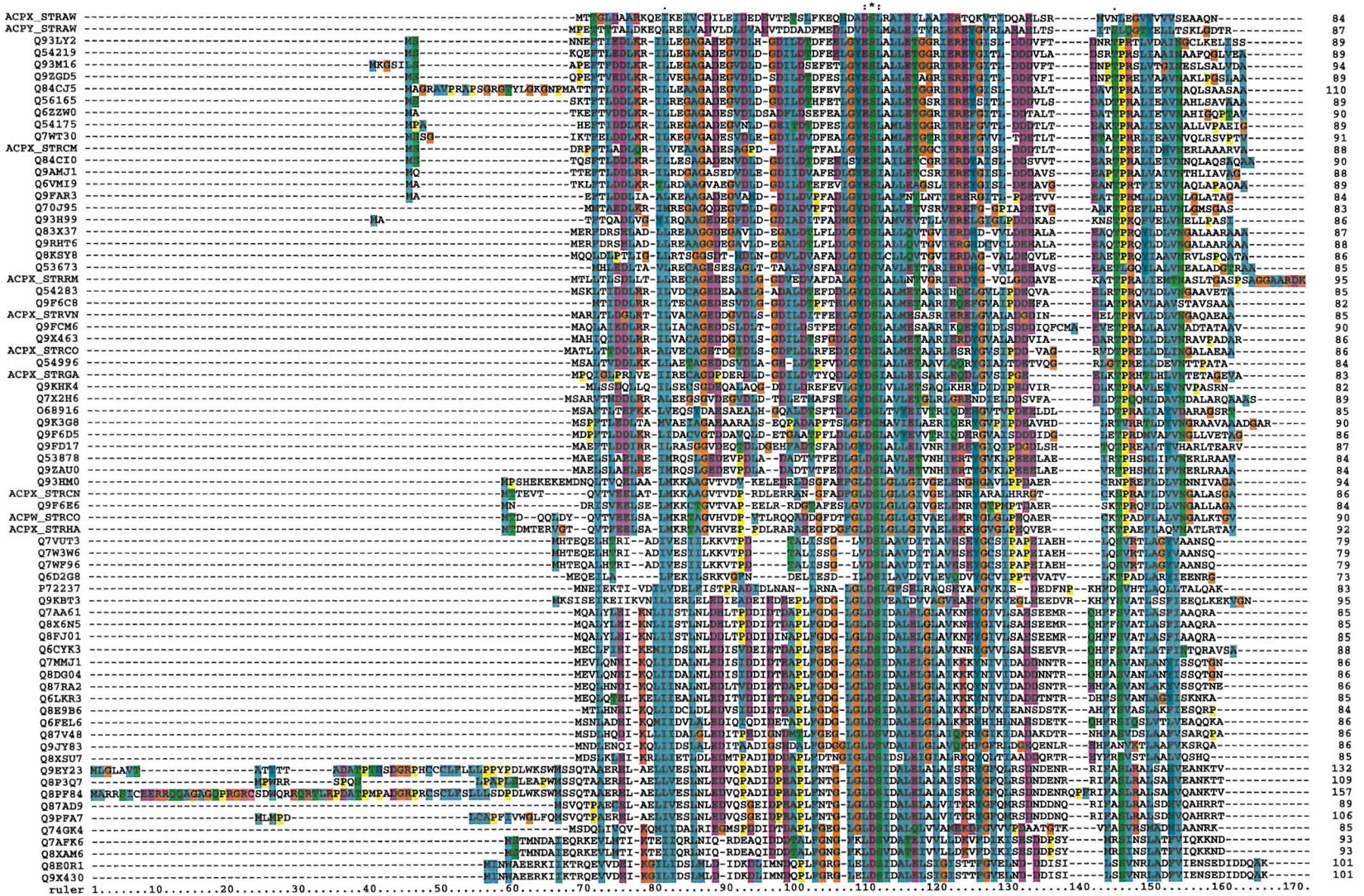


Figure VI.14 Sequence alignment showing all the polyketide ACP sequences when realigned against themselves.

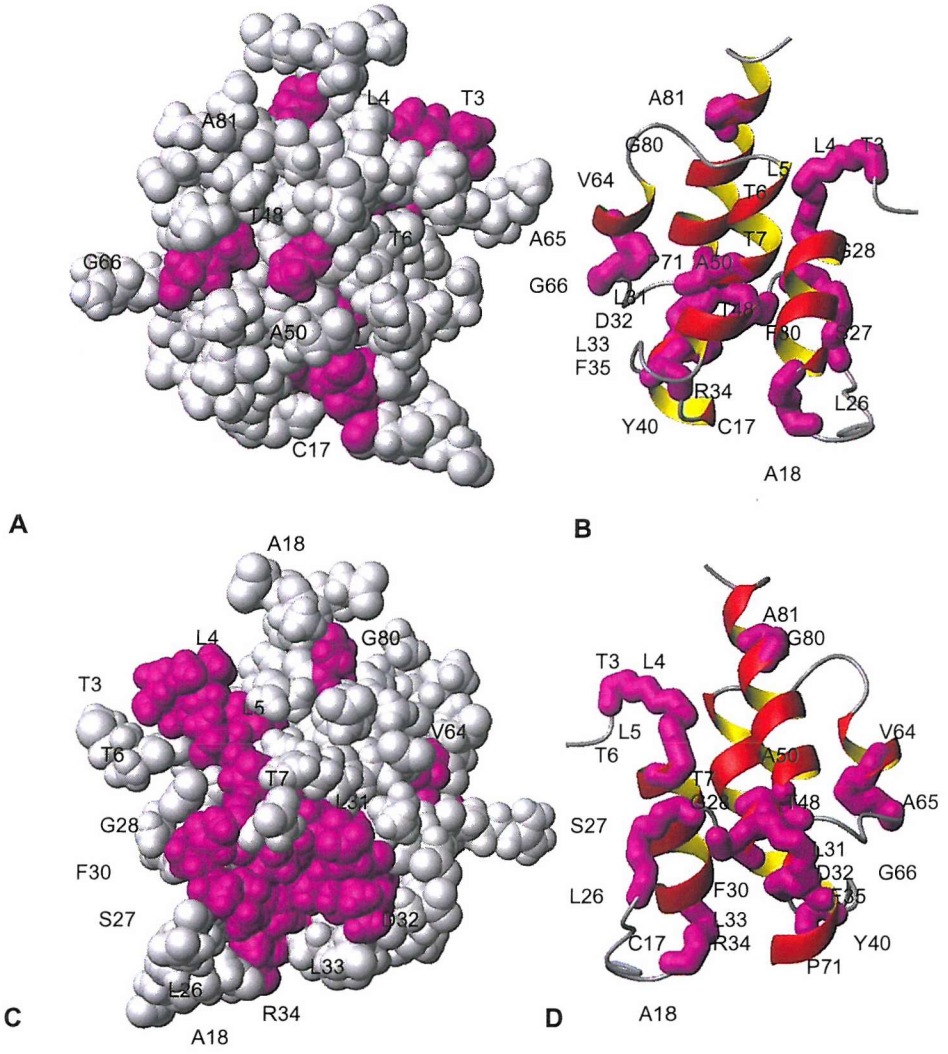


Figure VI.15 A. Residues that show low sequence scores for act-ACP when compared with the fatty acid polyketides (T3, L4, L5, T6, T7, C17S, A18, L26, S27, G28, F30, L31, D32, L33, R34, F35, Y40, T48, A49, A50, V64, A65, G66, P71, G80 and A81). B. Figure A in ribbon. C. Figure A rotated by 180°. D. Figure C in ribbon.

It is possible that the two mutants affected different residues but are causing the same result, which means that any of the other residues could be involved. If the two mutants affect the same residue then this is most probably T48, which could be acting as the MCAT active site, but if this was the case then T48 would be expected to be fully conserved, which it is not. The possibility that the tertiary structure allows for this cannot be discounted. Unfortunately the other two known polyketide-ACP structures are Q54996 (Fren) and ACPX_STRRM which, when the phylogenetic tree was examined (Figure VI.16), were very closely related to *act* and aligned the threonine. It is also possible that not all polyketide ACPs have the self-malyonation and transfer ability.

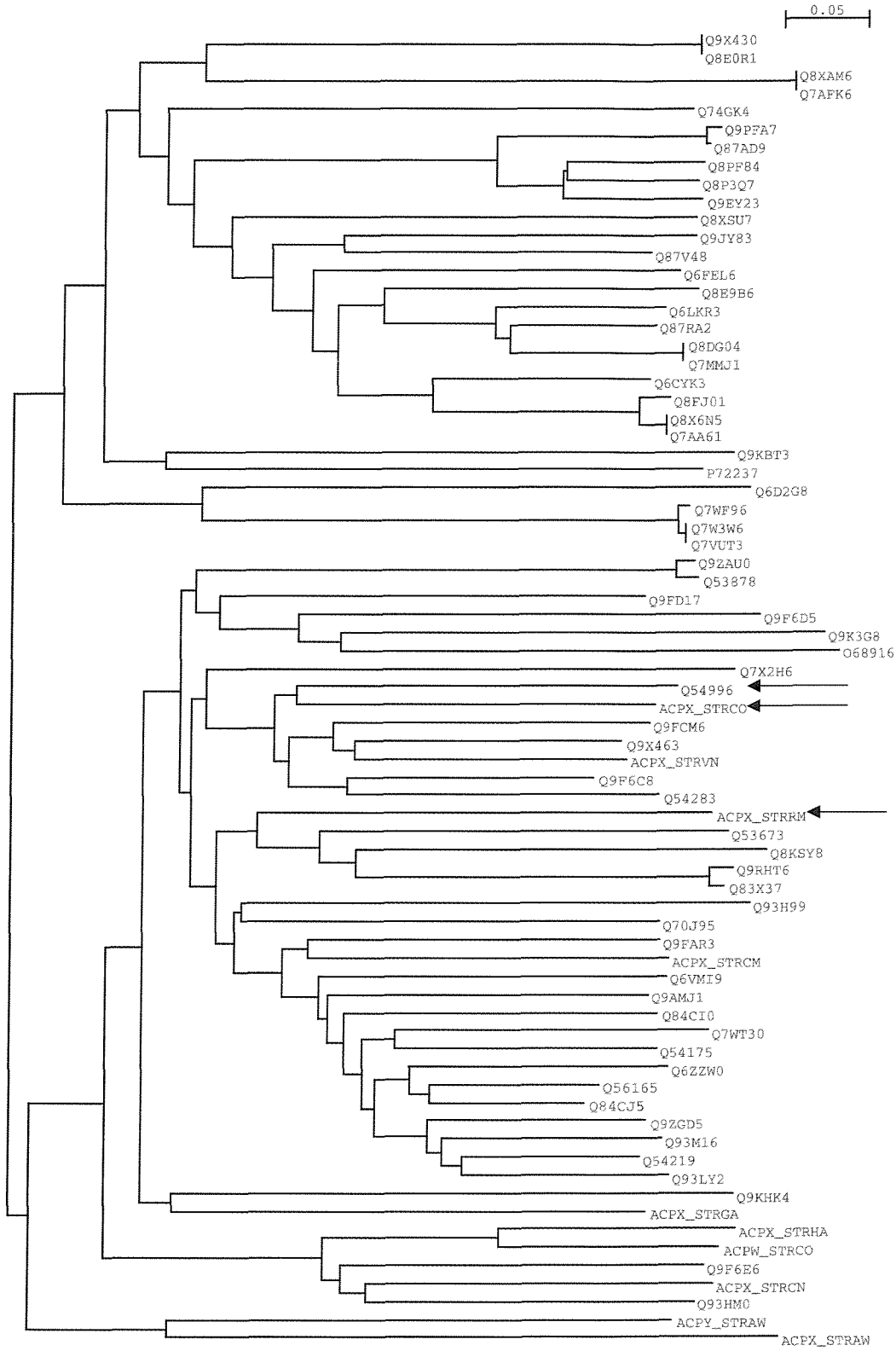


Figure VI.16 Phylogenetic tree of the sequence re-alignment of Figure VI.14 showing the three known structures of polyketide type II ACPs (Q54996, ACPX_STRCO and ACP_STRRM) are closely related.

5 References

1. Roitt, I., Brostoff, J., and Male, D. (1998) *Immunology*, Mosby International Limited, London

2. Schwarz, E. (1914) Die Lehre von der allgemeinen und ortlichen 'Eosinophilie'. *Ergeb Allge Pathol Pathol Anat* 17, 137-789

3. Speirs, R. S. (1958) Advances in the knowledge of the eosinophil in relation to antibody formation. *Ann NY Acad Sci* 73, 283

4. Austen, K. F. (1978) Homeostasis of effector systems which can also be recruited for immunologic reactions. *J Immunol* 121, 793-805

5. Baker, J. R., Bassett, E. G., and de Souza, P. (1976) Eosinophils in healing dermal wounds. *J Anat* 121, 401a

6. Bassett, E. G. (1962) Infiltration of eosinophils into the modified connective tissue of oestrous and pregnant animals. *Nature* 194, 1259-1261

7. Bassett, E. G. B., J. R. de Souza P (1977) A light microscopical study of healing incised dermal wounds in rats, with special reference to eosinophil leucocytes and to the collagenous fibres of the periwound areas. *Br J Exp Pathol* 58, 581-605

8. Archer, G. T. (1968) The function of the eosinophil. *Bibl Hemat* 29, 71-85

9. Spry, C. J. F. (1988) *Eosinophils: A Comprehensive Review and Guide to the Scientific and Medical Literature*, Oxford University Press, Oxford

10. Bass, D. A. (1982) Eosinophil behaviour during host defense reactions. In *'Advances in Host Defense Mechanisms'* (Gallin J I, F. A. S., ed) Vol. 1 pp. 211-241, Raven Press, New York

11. Walz, D. A., Wu, V. Y., de Lamo, R., Dene, H., and McCoy, L. E. (1977) Primary structure of human platelet factor 4. *Thromb Res* 11, 893-898

12. Murphy, P. M., Baggiolini, M., Charo, I. F., Hebert, C. A., Horuk, R., Matsushima, K., Miller, L. H., Oppenheim, J. J., and Power, C. A. (2000) International union of pharmacology. XXII. Nomenclature for chemokine receptors. *Pharmacol Rev* 52, 145-176

13. Wolpe, S. D., and Cerami, A. (1989) Macrophage inflammatory proteins 1 and 2: members of a novel superfamily of cytokines. *FASEB J* 3, 2565-2573

14. Schall, T. J. (1991) Biology of the RANTES/SIS cytokine family. *Cytokine* 3, 165-183

15. Oppenheim, J. J., Zachariae, C. O., Mukaida, N., and Matsushima, K. (1991) Properties of the novel proinflammatory supergene "intercrine" cytokine family. *Annu Rev Immunol* 9, 617-648

16. Lindley, I. J. D., Westwick J., Kunkel S. L. (1993) Nomenclature announcement - the chemokines. *Immunol Today* 14, 24

17. Zlotnik, A., and Yoshie, O. (2000) Chemokines: a new classification system and their role in immunity. *Immunity* 12, 121-127

18. Murphy, P. M. (2002) International Union of Pharmacology. XXX. Update on chemokine receptor nomenclature. *Pharmacol Rev* 54, 227-229

19. Thorpe, R. (2003) Chemokine/chemokine receptor nomenclature. *Cytokine* 21, 48-49

20. Crump, M. P., Rajarathnam, K., Kim, K. S., Clark-Lewis, I., and Sykes, B. D. (1998) Solution structure of eotaxin, a chemokine that selectively recruits eosinophils in allergic inflammation. *J Biol Chem* 273, 22471-22479

21. Mayer, K. L., and Stone, M. J. (2000) NMR solution structure and receptor peptide binding of the CC chemokine eotaxin-2. *Biochemistry* 39, 8382-8395

22. Ye, J., Mayer, K. L., Mayer, M. R., and Stone, M. J. (2001) NMR solution structure and backbone dynamics of the CC chemokine eotaxin-3. *Biochemistry* 40, 7820-7831

23. Baldwin, J. M. (1993) The probable arrangement of the helices in G protein-coupled receptors. *EMBO J* 12, 1693-1703

24. Unger, V. M., Hargrave, P. A., Baldwin, J. M., and Schertler, G. F. (1997) Arrangement of rhodopsin transmembrane alpha-helices. *Nature* 389, 203-206

25. Lomize, A. L., Pogozheva, I. D., and Mosberg, H. I. (1999) Structural organization of G-protein-coupled receptors. *J Comput Aided Mol Des* 13, 325-353

26. Palczewski, K., Kumasaka, T., Hori, T., Behnke, C. A., Motoshima, H., Fox, B. A., Le Trong, I., Teller, D. C., Okada, T., Stenkamp, R. E., Yamamoto, M., and Miyano, M. (2000) Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* 289, 739-745

27. Ai, L. S., and Liao, F. (2002) Mutating the four extracellular cysteines in the chemokine receptor CCR6 reveals their differing roles in receptor trafficking, ligand binding, and signaling. *Biochemistry* 41, 8332-8341

28. Samanta, A. K., Dutta, S., and Ali, E. (1993) Modification of sulfhydryl groups of interleukin-8 (IL-8) receptor impairs binding of IL-8 and IL-8-mediated chemotactic response of human polymorphonuclear neutrophils. *J Biol Chem* 268, 6147-6153

29. Dutta, S., Ali, E., and Samanta, A. K. (1993) Functions of interleukin-8 are mediated through thiol group(s) of IL-8 receptor in human polymorphonuclear neutrophils. Effects of 5,5'-dithio-bis(2-nitrobenzoic acid) on IL-8 receptor. *FEBS Lett* 325, 262-266

30. Benkirane, M., Jin, D. Y., Chun, R. F., Koup, R. A., and Jeang, K. T. (1997) Mechanism of transdominant inhibition of CCR5-mediated HIV-1 infection by ccr5delta32. *J Biol Chem* 272, 30603-30606

31. Lapham, C. K., Zaitseva, M. B., Lee, S., Romanstseva, T., and Golding, H. (1999) Fusion of monocytes and macrophages with HIV-1 correlates with biochemical properties of CXCR4 and CCR5. *Nat Med* 5, 303-308

32. Rodriguez-Frade, J. M., Vila-Coro, A. J., de Ana, A. M., Albar, J. P., Martinez, A. C., and Mellado, M. (1999) The chemokine monocyte chemoattractant protein-1 induces functional responses through dimerization of its receptor CCR2. *Proc Natl Acad Sci U S A* 96, 3628-3633

33. Blanpain, C., Doranz, B. J., Vakili, J., Rucker, J., Govaerts, C., Baik, S. S., Lorthioir, O., Migeotte, I., Libert, F., Baleux, F., Vassart, G., Doms, R. W., and Parmentier, M. (1999) Multiple charged and aromatic residues in CCR5 amino-terminal domain are involved in high affinity binding of both chemokines and HIV-1 Env protein. *J Biol Chem* 274, 34719-34727

34. Govaerts, C., Blanpain, C., Deupi, X., Ballet, S., Ballesteros, J. A., Wodak, S. J., Vassart, G., Pardo, L., and Parmentier, M. (2001) The TXP motif in the second transmembrane helix of CCR5. A structural determinant of chemokine-induced activation. *J Biol Chem* 276, 13217-13225

35. Navenot, J. M., Wang, Z. X., Trent, J. O., Murray, J. L., Hu, Q. X., DeLeeuw, L., Moore, P. S., Chang, Y., and Peiper, S. C. (2001) Molecular anatomy of CCR5 engagement by physiologic and viral chemokines and HIV-1 envelope glycoproteins: differences in primary structural requirements for RANTES, MIP-1 alpha, and vMIP-II Binding. *J Mol Biol* 313, 1181-1193

36. Lee, B., Sharron, M., Blanpain, C., Doranz, B. J., Vakili, J., Setoh, P., Berg, E., Liu, G., Guy, H. R., Durell, S. R., Parmentier, M., Chang, C. N., Price, K., Tsang, M., and Doms, R. W. (1999) Epitope mapping of CCR5 reveals multiple conformational states and distinct but overlapping structures involved in chemokine and coreceptor function. *J Biol Chem* 274, 9617-9626

37. Youn, B. S., Yu, K. Y., Alkhatib, G., and Kwon, B. S. (2001) The seventh transmembrane domain of cc chemokine receptor 5 is critical for MIP-1beta binding and receptor activation: role of MET 287. *Biochem Biophys Res Commun* 281, 627-633

38. Ahuja, S. K., and Murphy, P. M. (1996) The CXC chemokines growth-regulated oncogene (GRO) alpha, GRObeta, GROgamma, neutrophil-activating peptide-2, and epithelial cell-derived neutrophil-activating peptide-78 are potent agonists for

the type B, but not the type A, human interleukin-8 receptor. *J Biol Chem* 271, 20545-20550

39. Monteclaro, F. S., and Charo, I. F. (1996) The amino-terminal extracellular domain of the MCP-1 receptor, but not the RANTES/MIP-1alpha receptor, confers chemokine selectivity. Evidence for a two-step mechanism for MCP-1 receptor activation. *J Biol Chem* 271, 19084-19092

40. Crump, M. P., Gong, J. H., Loetscher, P., Rajarathnam, K., Amara, A., Arenzana-Seisdedos, F., Virelizier, J. L., Baggiolini, M., Sykes, B. D., and Clark-Lewis, I. (1997) Solution structure and basis for functional activity of stromal cell-derived factor-1; dissociation of CXCR4 activation from binding and inhibition of HIV-1. *EMBO J* 16, 6996-7007

41. Monteclaro, F. S., and Charo, I. F. (1997) The amino-terminal domain of CCR2 is both necessary and sufficient for high affinity binding of monocyte chemoattractant protein 1. Receptor activation by a pseudo-tethered ligand. *J Biol Chem* 272, 23186-23190

42. Blanpain, C., Lee, B., Tackoen, M., Puffer, B., Boom, A., Libert, F., Sharron, M., Wittamer, V., Vassart, G., Doms, R. W., and Parmentier, M. (2000) Multiple nonfunctional alleles of CCR5 are frequent in various human populations. *Blood* 96, 1638-1645

43. Alkhatib, G., Combadiere, C., Broder, C. C., Feng, Y., Kennedy, P. E., Murphy, P. M., and Berger, E. A. (1996) CC CKR5: a RANTES, MIP-1alpha, MIP-1beta receptor as a fusion cofactor for macrophage-tropic HIV-1. *Science* 272, 1955-1958

44. Yang, C., Li, M., Limpakarnjanarat, K., Young, N. L., Hodge, T., Butera, S. T., McNicholl, J. M., Mastro, T. D., and Lal, R. B. (2003) Polymorphisms in the CCR5 coding and noncoding regions among HIV type 1-exposed, persistently seronegative female sex-workers from Thailand. *AIDS Res Hum Retroviruses* 19, 661-665

45. Han, K. H., Green, S. R., Tangirala, R. K., Tanaka, S., and Quehenberger, O. (1999) Role of the first extracellular loop in the functional activation of CCR2. The first extracellular loop contains distinct domains necessary for both agonist binding and transmembrane signaling. *J Biol Chem* 274, 32055-32062

46. Arias, D. A., Navenot, J. M., Zhang, W. B., Broach, J., and Peiper, S. C. (2003) Constitutive activation of CCR5 and CCR2 induced by conformational changes in the conserved TXP motif in transmembrane helix 2. *J Biol Chem* 278, 36513-36521

47. Blanpain, C., Doranz, B. J., Bondue, A., Govaerts, C., De Leener, A., Vassart, G., Doms, R. W., Proudfoot, A., and Parmentier, M. (2003) The core domain of chemokines binds CCR5 extracellular domains while their amino terminus interacts with the transmembrane helix bundle. *J Biol Chem* 278, 5179-5187

48. Govaerts, C., Bondue, A., Springael, J. Y., Olivella, M., Deupi, X., Le Poul, E., Wodak, S. J., Parmentier, M., Pardo, L., and Blanpain, C. (2003) Activation of

CCR5 by chemokines involves an aromatic cluster between transmembrane helices 2 and 3. *J Biol Chem* 278, 1892-1903

49. Shinkai, A., Komuta-Kunitomo, M., Sato-Nakamura, N., and Anazawa, H. (2002) N-terminal domain of eotaxin-3 is important for activation of CC chemokine receptor 3. *Protein Eng* 15, 923-929

50. Guan, E., Wang, J., Roderiquez, G., and Norcross, M. A. (2002) Natural truncation of the chemokine MIP-1 beta /CCL4 affects receptor specificity but not anti-HIV-1 activity. *J Biol Chem* 277, 32348-32352

51. Proudfoot, A. E., Power, C. A., Hoogewerf, A. J., Montjovent, M. O., Borlat, F., Offord, R. E., and Wells, T. N. (1996) Extension of recombinant human RANTES by the retention of the initiating methionine produces a potent antagonist. *J Biol Chem* 271, 2599-2603

52. Townson, J. R., Graham, G. J., Landau, N. R., Rasala, B., and Nibbs, R. J. (2000) Aminooxypentane addition to the chemokine macrophage inflammatory protein-1alpha P increases receptor affinities and HIV inhibition. *J Biol Chem* 275, 39254-39261

53. Zoffmann, S., Turcatti, G., Galzi, J., Dahl, M., and Chollet, A. (2001) Synthesis and characterization of fluorescent and photoactivatable MIP-1alpha ligands and interactions with chemokine receptors CCR1 and CCR5. *J Med Chem* 44, 215-222

54. Pakianathan, D. R., Kuta, E. G., Artis, D. R., Skelton, N. J., and Hebert, C. A. (1997) Distinct but overlapping epitopes for the interaction of a CC-chemokine with CCR1, CCR3 and CCR5. *Biochemistry* 36, 9642-9648

55. Hemmerich, S., Paavola, C., Bloom, A., Bhakta, S., Freedman, R., Grunberger, D., Krstenansky, J., Lee, S., McCarley, D., Mulkins, M., Wong, B., Pease, J., Mizoue, L., Mirzadegan, T., Polsky, I., Thompson, K., Handel, T. M., and Jarnagin, K. (1999) Identification of residues in the monocyte chemotactic protein-1 that contact the MCP-1 receptor, CCR2. *Biochemistry* 38, 13013-13025

56. Beck, C. G., Studer, C., Zuber, J. F., Demange, B. J., Manning, U., and Urfer, R. (2001) The viral CC chemokine-binding protein vCCI inhibits monocyte chemoattractant protein-1 activity by masking its CCR2B-binding site. *J Biol Chem* 276, 43270-43276

57. Laurence, J. S., Blanpain, C., Burgner, J. W., Parmentier, M., and LiWang, P. J. (2000) CC chemokine MIP-1 beta can function as a monomer and depends on Phe13 for receptor binding. *Biochemistry* 39, 3401-3409

58. Laurence, J. S., Blanpain, C., De Leener, A., Parmentier, M., and LiWang, P. J. (2001) Importance of basic residues and quaternary structure in the function of MIP-1 beta: CCR5 binding and cell surface sugar interactions. *Biochemistry* 40, 4990-4999

59. Bondue, A., Jao, S. C., Blanpain, C., Parmentier, M., and LiWang, P. J. (2002) Characterization of the role of the N-loop of MIP-1 beta in CCR5 binding. *Biochemistry* 41, 13548-13555

60. Proudfoot, A. E., Fritchley, S., Borlat, F., Shaw, J. P., Vilbois, F., Zwahlen, C., Trkola, A., Marchant, D., Clapham, P. R., and Wells, T. N. (2001) The BBXB motif of RANTES is the principal site for heparin binding and controls receptor selectivity. *J Biol Chem* 276, 10620-10626

61. Pease, J. E., Wang, J., Ponath, P. D., and Murphy, P. M. (1998) The N-terminal extracellular segments of the chemokine receptors CCR1 and CCR3 are determinants for MIP-1alpha and eotaxin binding, respectively, but a second domain is essential for efficient receptor activation. *J Biol Chem* 273, 19972-19976

62. Datta, A., and Stone, M. J. (2003) Soluble mimics of a chemokine receptor: chemokine binding by receptor elements juxtaposed on a soluble scaffold. *Protein Sci* 12, 2482-2491

63. Ponath, P. D., Qin, S., Ringler, D. J., Clark-Lewis, I., Wang, J., Kassam, N., Smith, H., Shi, X., Gonzalo, J. A., Newman, W., Gutierrez-Ramos, J. C., and Mackay, C. R. (1996) Cloning of the human eosinophil chemoattractant, eotaxin. Expression, receptor binding, and functional properties suggest a mechanism for the selective recruitment of eosinophils. *J Clin Invest* 97, 604-612

64. Bannert, N., Craig, S., Farzan, M., Sogah, D., Santo, N. V., Choe, H., and Sodroski, J. (2001) Sialylated O-glycans and sulfated tyrosines in the NH2-terminal domain of CC chemokine receptor 5 contribute to high affinity binding of chemokines. *J Exp Med* 194, 1661-1673

65. Preobrazhensky, A. A., Dragan, S., Kawano, T., Gavrilin, M. A., Gulina, I. V., Chakravarty, L., and Kolattukudy, P. E. (2000) Monocyte chemotactic protein-1 receptor CCR2B is a glycoprotein that has tyrosine sulfation in a conserved extracellular N-terminal region. *J Immunol* 165, 5295-5303

66. Seibert, C., Cadene, M., Sanfiz, A., Chait, B. T., and Sakmar, T. P. (2002) Tyrosine sulfation of CCR5 N-terminal peptide by tyrosylprotein sulfotransferases 1 and 2 follows a discrete pattern and temporal sequence. *Proc Natl Acad Sci U S A* 99, 11031-11036

67. Purcell, E. M., Torrey, H. C., and Pound, R. V. (1946) Resonance absorption by nuclear magnetic moments in a solid. *Phys. Rev.* 69, 37-38

68. Bloch, F., Hansen, and Packard, M. (1946) Nuclear Induction. *Phys. Rev.* 69, 127

69. Williamson, M. P., Marion, D., and Wuthrich, K. (1984) Secondary structure in the solution conformation of the proteinase inhibitor IIA from bull seminal plasma by nuclear magnetic resonance. *J Mol Biol* 173, 341-359

70. Kendrew, J. C., Dickerson, R. E., Strandberg, B. E., Hart, R. G., Davies, D. R., Philips, D. C., and Shore, V. C. (1960) Structure of myoglobin: A three dimensional fourier synthesis at 2 Å resolution. *Nature* 185, 422-427

71. Kline, A. D., Braun, W., and Wuthrich, K. (1988) Determination of the complete three-dimensional structure of the alpha-amylase inhibitor tendamistat in aqueous solution by nuclear magnetic resonance and distance geometry. *J Mol Biol* 204, 675-724

72. Rattle, H. (1995) *An NMR Primer for Life Scientists*, Partnership Press, Fareham, Hants

73. Claridge, T. D. W. (1999) *High-resolution NMR techniques in organic chemistry*, Pergamon, Amsterdam ; Oxford

74. Cavanagh, J. (1996) *Protein NMR spectroscopy : principles and practice*, Academic Press, San Diego, Calif.

75. Levitt, M. H. (2001) *Spin dynamics : basics of nuclear magnetic resonance*, John Wiley & Sons, Chichester

76. Rosman, K. J. R., and Taylor, P. D. P. (1998) Isotopic composition of the elements. *Pure Appl. Chem.* 70, 217-235

77. Harris, R. K., Becker, E. D., Cabral de Menezes, S. M., Goodfellow, R., and Granger, P. (2001) NMR nomenclature. Nuclear spin properties and conventions for chemical shifts (IUPAC Recommendations 2001). *Pure Appl. Chem.* 73, 1795-1818

78. Bothner-By, A. A., Stephens, R. L., and Lee, J. (1984) Structure Determination of a Tetrasaccharide: Transient Nuclear Overhauser Effects in the Rotating Frame. *J. Am. Chem. Soc.* 106, 811-813

79. Williamson Michael, P., and Neuhaus, D. (2000) *The nuclear Overhauser effect in structural and conformational analysis*, John Wiley, New York ; Chichester

80. Oezguen, N., Adamian, L., Xu, Y., Rajarathnam, K., and Braun, W. (2002) Automated assignment and 3D structure calculations using combinations of 2D homonuclear and 3D heteronuclear NOESY spectra. *J Biomol NMR* 22, 249-263

81. Wuthrick, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.

82. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J Mol Biol* 222, 311-333

83. Merutka, G., Dyson, H. J., and Wright, P. E. (1995) 'Random coil' 1H chemical shifts obtained as a function of temperature and trifluoroethanol concentration for the peptide series GGXGG. *J Biomol NMR* 5, 14-24

84. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) 1H, 13C and 15N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J Biomol NMR* 5, 67-81

85. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry* 31, 1647-1651

86. Andersen, N. H., Neidigh, J. W., Harris, S. M., Lee, G. M., Liu, Z. H., and Tong, H. (1997) Extracting information from the temperature gradients of polypeptide NH chemical shifts .1. The importance of conformational averaging. *Journal of the American chemical society* 119, 8547-8561

87. Clarkson, J., and Campbell, I. D. (2003) Studies of protein-ligand interactions by NMR. *Biochem Soc Trans* 31, 1006-1009

88. Hore, P. J. (1995) *Nuclear magnetic resonance*, Oxford University Press, Oxford

89. Zuiderweg, E. R. (2002) Mapping protein-protein interactions in solution by NMR spectroscopy. *Biochemistry* 41, 1-7

90. Thomas, G. H. (2001) New routes to membrane protein structures. Practical course: current methods in membrane protein research. *EMBO Rep* 2, 187-191

91. Arora, A., and Tamm, L. K. (2001) Biophysical approaches to membrane protein structure determination. *Curr Opin Struct Biol* 11, 540-547

92. Baleja, J. D. (2001) Structure determination of membrane-associated proteins from nuclear magnetic resonance data. *Anal Biochem* 288, 1-15

93. MacArthur, M. W., and Thornton, J. M. (1991) Influence of proline residues on protein conformation. *J Mol Biol* 218, 397-412

94. Woolfson, D. N., and Williams, D. H. (1990) The influence of proline residues on alpha-helical structure. *FEBS Lett* 277, 185-188

95. Killian, J. A., and von Heijne, G. (2000) How proteins adapt to a membrane-water interface. *Trends Biochem Sci* 25, 429-434

96. White, S. H., and Wimley, W. C. (1999) Membrane protein folding and stability: physical principles. *Annu Rev Biophys Biomol Struct* 28, 319-365

97. Popot, J. L., and Engelman, D. M. (2000) Helical membrane protein folding, stability, and evolution. *Annu Rev Biochem* 69, 881-922

98. Katragadda, M., Alderfer, J. L., and Yeagle, P. L. (2000) Solution structure of the loops of bacteriorhodopsin closely resembles the crystal structure. *Biochim Biophys Acta* 1466, 1-6

99. Yeagle, P. L., Alderfer, J. L., Salloum, A. C., Ali, L., and Albert, A. D. (1997) The first and second cytoplasmic loops of the G-protein receptor, rhodopsin, independently form beta-turns. *Biochemistry* 36, 3864-3869

100. Lauterwein, J., Bosch, C., Brown, L. R., and Wuthrich, K. (1979) Physicochemical studies of the protein-lipid interactions in melittin-containing micelles. *Biochim Biophys Acta* 556, 244-264

101. Brown, L. R., Bosch, C., and Wuthrich, K. (1981) Location and orientation relative to the micelle surface for glucagon in mixed micelles with dodecylphosphocholine: EPR and NMR studies. *Biochim Biophys Acta* 642, 296-312

102. Fernandez, C., and Wuthrich, K. (2003) NMR solution structure determination of membrane proteins reconstituted in detergent micelles. *FEBS Lett* 555, 144-150

103. Pellegrini, M., Bisello, A., Rosenblatt, M., Chorev, M., and Mierke, D. F. (1998) Binding domain of human parathyroid hormone receptor: from conformation to function. *Biochemistry* 37, 12737-12743

104. Piserchio, A., Bisello, A., Rosenblatt, M., Chorev, M., and Mierke, D. F. (2000) Characterization of parathyroid hormone/receptor interactions: structure of the first extracellular loop. *Biochemistry* 39, 8153-8160

105. Giragossian, C., and Mierke, D. F. (2001) Intermolecular interactions between cholecystokinin-8 and the third extracellular loop of the cholecystokinin A receptor. *Biochemistry* 40, 3804-3809

106. Zhang, L., DeHaven, R. N., and Goodman, M. (2002) NMR and modeling studies of a synthetic extracellular loop II of the kappa opioid receptor in a DPC micelle. *Biochemistry* 41, 61-68

107. Giragossian, C., and Mierke, D. F. (2002) Intermolecular interactions between cholecystokinin-8 and the third extracellular loop of the cholecystokinin-2 receptor. *Biochemistry* 41, 4560-4566

108. Ulfers, A. L., Piserchio, A., and Mierke, D. F. (2002) Extracellular domains of the neurokinin-1 receptor: structural characterization and interactions with substance P. *Biopolymers* 66, 339-349

109. Giragossian, C., Sugg, E. E., Szewczyk, J. R., and Mierke, D. F. (2003) Intermolecular interactions between peptidic and nonpeptidic agonists and the third extracellular loop of the cholecystokinin 1 receptor. *J Med Chem* 46, 3476-3482

110. Giragossian, C., Schaschke, N., Moroder, L., and Mierke, D. F. (2004) Conformational and molecular modeling studies of beta-cyclodextrin-heptagastrin and the third extracellular loop of the cholecystokinin 2 receptor. *Biochemistry* 43, 2724-2731

111. Ye, J. P., Kohli, L. L., and Stone, M. J. (2000) Characterization of binding between the chemokine eotaxin and peptides derived from the chemokine receptor CCR3. *J Biol Chem* 275, 27250-27257

112. Mayer, M. R., and Stone, M. J. (2001) Identification of receptor binding and activation determinants in the N-terminal and N-loop regions of the CC chemokine eotaxin. *J Biol Chem* 276, 13911-13916

113. Okada, T., Fujiyoshi, Y., Silow, M., Navarro, J., Landau, E. M., and Shichida, Y. (2002) Functional role of internal water molecules in rhodopsin revealed by X-ray crystallography. *Proc Natl Acad Sci U S A* 99, 5982-5987

114. Luecke, H., Schobert, B., Richter, H. T., Cartailler, J. P., and Lanyi, J. K. (1999) Structure of bacteriorhodopsin at 1.55 A resolution. *J Mol Biol* 291, 899-911

115. Yeagle, P. L., Salloum, A., Chopra, A., Bhawsar, N., Ali, L., Kuzmanovski, G., Alderfer, J. L., and Albert, A. D. (2000) Structures of the intradiskal loops and amino terminus of the G-protein receptor, rhodopsin. *J Pept Res* 55, 455-465

116. Combadiere, C., Ahuja, S. K., and Murphy, P. M. (1995) Cloning and functional expression of a human eosinophil CC chemokine receptor. *J Biol Chem* 270, 16491-16494

117. Daugherty, B. L., Siciliano, S. J., DeMartino, J. A., Malkowitz, L., Sirotina, A., and Springer, M. S. (1996) Cloning, expression, and characterization of the human eosinophil eotaxin receptor. *J Exp Med* 183, 2349-2354

118. Ponath, P. D., Qin, S., Post, T. W., Wang, J., Wu, L., Gerard, N. P., Newman, W., Gerard, C., and Mackay, C. R. (1996) Molecular cloning and characterization of a human eotaxin receptor expressed selectively on eosinophils. *J Exp Med* 183, 2437-2448

119. Berners-Lee, T., Cailliau, R., JF., G., and Pollerman, B. (1992) World Wide Web: The Information Universe. Electronic Networking. *Research, Applications and Policy* 2, 52-58

120. Hofmann, K. S., W. (1993) TMBASE - A database of membrane spanning protein segments. *Biol. Chem. Hoppe-Seyler* 374, 166

121. Mitaku, S., and Hirokawa, T. (1999) Physicochemical factors for discriminating between soluble and membrane proteins: hydrophobicity of helical segments and protein length. *Protein Eng* 12, 953-957

122. Mitaku, S., Hirokawa, T., and Tsuji, T. (2002) Amphiphilicity index of polar amino acids as an aid in the characterization of amino acid preference at membrane-water interfaces. *Bioinformatics* 18, 608-616

123. Mitaku, S., Ono, M., Hirokawa, T., Boon-Chieng, S., and Sonoyama, M. (1999) Proportion of membrane proteins in proteomes of 15 single-cell organisms analyzed by the SOSUI prediction system. *Biophys Chem* 82, 165-171

124. Hirokawa, T., Boon-Chieng, S., and Mitaku, S. (1998) SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics* 14, 378-379

125. Tusnady, G. E., and Simon, I. (1998) Principles governing amino acid composition of integral membrane proteins: application to topology prediction. *J Mol Biol* 283, 489-506

126. Persson, B., and Argos, P. (1994) Prediction of transmembrane segments in proteins utilising multiple sequence alignments. *J Mol Biol* 237, 182-192

127. Persson, B., and Argos, P. (1996) Topology prediction of membrane proteins. *Protein Sci* 5, 363-371

128. Neote, K., DiGregorio, D., Mak, J. Y., Horuk, R., and Schall, T. J. (1993) Molecular cloning, functional expression, and signaling characteristics of a C-C chemokine receptor. *Cell* 72, 415-425

129. Charo, I. F., Myers, S. J., Herman, A., Franci, C., Connolly, A. J., and Coughlin, S. R. (1994) Molecular cloning and functional expression of two monocyte chemoattractant protein 1 receptors reveals alternative splicing of the carboxyl-terminal tails. *Proc Natl Acad Sci U S A* 91, 2752-2756

130. Power, C. A., Meyer, A., Nemeth, K., Bacon, K. B., Hoogewerf, A. J., Proudfoot, A. E., and Wells, T. N. (1995) Molecular cloning and functional expression of a novel CC chemokine receptor cDNA from a human basophilic cell line. *J Biol Chem* 270, 19495-19500

131. Samson, M., Labbe, O., Mollereau, C., Vassart, G., and Parmentier, M. (1996) Molecular cloning and functional expression of a new human CC-chemokine receptor gene. *Biochemistry* 35, 3362-3367

132. Baba, M., Imai, T., Nishimura, M., Kakizaki, M., Takagi, S., Hieshima, K., Nomiyama, H., and Yoshie, O. (1997) Identification of CCR6, the specific receptor for a novel lymphocyte-directed CC chemokine LARC. *J Biol Chem* 272, 14893-14898

133. Birkenbach, M., Josefsen, K., Yalamanchili, R., Lenoir, G., and Kieff, E. (1993) Epstein-Barr virus-induced genes: first lymphocyte-specific G protein-coupled peptide receptors. *J Virol* 67, 2209-2220

134. Tiffany, H. L., Lautens, L. L., Gao, J. L., Pease, J., Locati, M., Combadiere, C., Modi, W., Bonner, T. I., and Murphy, P. M. (1997) Identification of CCR8: a human monocyte and thymus receptor for the CC chemokine I-309. *J Exp Med* 186, 165-170

135. Zaballos, A., Gutierrez, J., Varona, R., Ardavin, C., and Marquez, G. (1999) Cutting edge: identification of the orphan chemokine receptor GPR-9-6 as CCR9, the receptor for the chemokine TECK. *J Immunol* 162, 5671-5675

136. Jarmin, D. I., Rits, M., Bota, D., Gerard, N. P., Graham, G. J., Clark-Lewis, I., and Gerard, C. (2000) Cutting edge: identification of the orphan receptor G-protein-coupled receptor 2 as CCR10, a specific receptor for the chemokine ESkine. *J Immunol* 164, 3460-3464

137. Schweickart, V. L., Epp, A., Raport, C. J., and Gray, P. W. (2000) CCR11 is a functional receptor for the monocyte chemoattractant protein family of chemokines. *J Biol Chem* 275, 9550-9556

138. Cochran, A. G., Skelton, N. J., and Starovasnik, M. A. (2001) Tryptophan zippers: stable, monomeric beta -hairpins. *Proc Natl Acad Sci U S A* 98, 5578-5583

139. Gore, M. G. (2000) *Spectrophotometry and spectrofluorimetry : a practical approach*, Oxford University Press, Oxford

140. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* 8, 477-486

141. Forssmann, U., Uguccioni, M., Loetscher, P., Dahinden, C. A., Langen, H., Thelen, M., and Baggiolini, M. (1997) Eotaxin-2, a novel CC chemokine that is selective for the chemokine receptor CCR3, and acts like eotaxin on human eosinophil and basophil leukocytes. *J Exp Med* 185, 2171-2176

142. Patel, V. P., Kreider, B. L., Li, Y., Li, H., Leung, K., Salcedo, T., Nardelli, B., Pippalla, V., Gentz, S., Thotakura, R., Parmelee, D., Gentz, R., and Garotta, G. (1997) Molecular and functional characterization of two novel human C-C chemokines as inhibitors of two distinct classes of myeloid progenitors. *J Exp Med* 185, 1163-1172

143. White, J. R., Imburgia, C., Dul, E., Appelbaum, E., O'Donnell, K., O'Shannessy, D. J., Brawner, M., Fornwald, J., Adamou, J., Elshourbagy, N. A., Kaiser, K., Foley, J. J., Schmidt, D. B., Johanson, K., Macphee, C., Moores, K., McNulty, D., Scott, G. F., Schleimer, R. P., and Sarau, H. M. (1997) Cloning and functional characterization of a novel human CC chemokine that binds to the CCR3 receptor and activates human eosinophils. *J Leukoc Biol* 62, 667-675

144. Grzegorzewski, K. J., Yao, X. T., Kreider, B., Olsen, H. S., Morris, T. S., Zhang, L., Sanyal, I., Nardelli, B., Zukauskas, D., Brewer, L., Bong, G. W., Kim, Y., Garotta, G., and Salcedo, T. W. (2001) Analysis of eosinophils and myeloid progenitor responses to modified forms of MPIF-2. *Cytokine* 13, 209-219

145. Sloane, P. (1999) *cDNA of CCL24*, Celltech, 208 Bath Road, Slough, Berks, SL1 3WE

146. Kane, J. F., and Hartley, D. L. (1988) Formation of recombinant protein inclusion bodies in Escherichia coli. *Trends Biotechnol* 6, 95-101

147. Qiagen (1997) *The QIAexpressionist, A handbook for high-level expression and purification of 6xHis-tagged proteins.*, Qiagen, Crawley

148. Amersham-Biosciences (2002) *Purification of (His)6-tagged proteins using HiTrap Chelating HP columns charged with different metal ions*, Amersham-Biosciences, Little Chalfont

149. Shire, S. J., Bock, L., Ogez, J., Builder, S., Kleid, D., and Moore, D. M. (1984) Purification and immunogenicity of fusion VP1 protein of food and mouth disease virus. *Biochemistry* 23, 6474-6480

150. Thatcher, D. R. (1990) Recovery of Therapeutic proteins from inclusion bodies: problems and process strategies. *Biochem Soc Trans* 18, 234-235

151. Bohmann, D., and Tjian, R. (1989) Biochemical analysis of transcriptional activation by Jun: differential activity of c- and v-Jun. *Cell* 59, 709-717

152. Levinthal, C. (1968) Are there pathways for protein folding. *J. Chem. Phys* 65, 44-45

153. Goldberg, M. E., Rudolph, R., and Jaenicke, R. (1991) A kinetic study of the competition between renaturation and aggregation during the refolding of denatured-reduced egg white lysozyme. *Biochemistry* 30, 2790-2797

154. Wetzel, R. (1994) Mutations and off-pathway aggregation of proteins. *Trends Biotechnol* 12, 193-198

155. De Bernardez Clark, E., Hevehan, D., Szela, S., and Maachupalli-Reddy, J. (1998) Oxidative renaturation of hen egg-white lysozyme. Folding vs aggregation. *Biotechnol Prog* 14, 47-54

156. London, J., Skrzynia, C., and Goldberg, M. E. (1974) Renaturation of Escherichia coli tryptophanase after exposure to 8 M urea. Evidence for the existence of nucleation centers. *Eur J Biochem* 47, 409-415

157. Kiefhaber, T., Rudolph, R., Kohler, H. H., and Buchner, J. (1991) Protein aggregation in vitro and in vivo: a quantitative model of the kinetic competition between folding and aggregation. *Biotechnology (N Y)* 9, 825-829

158. Zettlmeissl, G., Rudolph, R., and Jaenicke, R. (1979) Reconstitution of lactic dehydrogenase. Noncovalent aggregation vs. reactivation. 1. Physical properties and kinetics of aggregation. *Biochemistry* 18, 5567-5571

159. Buchner, J., and Rudolph, R. (1991) Renaturation, purification and characterization of recombinant Fab-fragments produced in Escherichia coli. *Biotechnology (N Y)* 9, 157-162

160. Maeda, Y., Yamada, H., Ueda, T., and Imoto, T. (1996) Effect of additives on the renaturation of reduced lysozyme in the presence of 4M urea. *Protein Eng* 9, 461-465

161. Orsini, G., and Goldberg, M. E. (1978) The renaturation of reduced chymotrypsinogen A in guanidine HCl. Refolding versus aggregation. *J Biol Chem* 253, 3453-3458

162. Rudolph, R., and Lilie, H. (1996) In vitro folding of inclusion body proteins. *FASEB J* 10, 49-56

163. De Bernardez Clark, E. (1998) Refolding of recombinant proteins. *Curr Opin Struct Biol* 9, 157-163

164. Hevehan, D. L., and Clark, E. D. (1997) Oxidative renaturation of lysozyme at high concentrations. *Biotechnology and Bioengineering* 54, 221-230

165. Booth, V., Keizer, D. W., Kamphuis, M. B., Clark-Lewis, I., and Sykes, B. D. (2002) The CXCR3 binding chemokine IP-10/CXCL10: structure and receptor interactions. *Biochemistry* 41, 10418-10425

166. Clubb, R. T., Omichinski, J. G., Clore, G. M., and Gronenborn, A. M. (1994) Mapping the binding surface of interleukin-8 complexes with an N-terminal fragment of the type 1 human interleukin-8 receptor. *FEBS Lett* 338, 93-97

167. Skelton, N. J., Quan, C., Reilly, D., and Lowman, H. (1999) Structure of a CXC chemokine-receptor fragment in complex with interleukin-8. *Structure Fold Des* 7, 157-168

168. Mizoue, L. S., Bazan, J. F., Johnson, E. C., and Handel, T. M. (1999) Solution structure and dynamics of the CX3C chemokine domain of fractalkine and its interaction with an N-terminal fragment of CX3CR1. *Biochemistry* 38, 1402-1414

169. Studier, F. W., and Moffatt, B. A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* 189, 113-130

170. Derman, A. I., Prinz, W. A., Belin, D., and Beckwith, J. (1993) Mutations that allow disulfide bond formation in the cytoplasm of Escherichia coli. *Science* 262, 1744-1747

171. Bullock, W. O., Fernandez, J. M., and Short, J. M. (1987) XL1-Blue - A high-efficiency plasmid transforming recA *Escherichia-coli* strain with beta-galactosidase selection. *Biotechniques* 5, 376

172. Rosenberg, A. H., Lade, B. N., Chui, D. S., Lin, S. W., Dunn, J. J., and Studier, F. W. (1987) Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene* 56, 125-135

173. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol* 185, 60-89

174. Shuman, S. (1994) Novel approach to molecular cloning and polynucleotide synthesis using vaccinia DNA topoisomerase. *J Biol Chem* 269, 32678-32684

175. Cohen, S. N., Chang, A. C., and Hsu, L. (1972) Nonchromosomal antibiotic resistance in bacteria: genetic transformation of Escherichia coli by R-factor DNA. *Proc Natl Acad Sci U S A* 69, 2110-2114

176. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72, 248-254

177. Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., and Bairoch, A. (2003) ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* 31, 3784-3788

178. Apweiler, R., Gateau, A., Contrino, S., Martin, M. J., Junker, V., O'Donovan, C., Lang, F., Mitaritonna, N., Kappus, S., and Bairoch, A. (1997) Protein sequence annotation in the genome era: the annotation concept of SWISS-PROT+TREMBL. *Proc Int Conf Intell Syst Mol Biol* 5, 33-43

179. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Res* 28, 235-242

180. Koradi, R., Billeter, M., and Wuthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph* 14, 51-55, 29-32

181. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25, 4876-4882

182. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6, 277-293

183. Garrett, D. S., Powers, R., Gronenborn, A. M., and Clore, G. M. (1995) A common sense approach to peak picking in two-, three- and four-dimensional spectra using automatic computer analysis of contour diagrams. *J. Magn. Reson.* 95, 214-220

184. Nilges, M., Clore, G. M., and Gronenborn, A. M. (1988) Determination of three-dimensional structures of proteins from interproton distance data by hybrid distance geometry-dynamical simulated annealing calculations. *FEBS Lett* 229, 317-324

185. Brünger Axel, T. (1992) *X-PLOR version 3.1 : a system for X-ray crystallography and NMR*, Yale University Press, New Haven, [Conn.]

186. Griffiths-Johnson, D. A., Collins, P. D., Rossi, A. G., Jose, P. J., and Williams, T. J. (1993) The chemokine, eotaxin, activates guinea-pig eosinophils in vitro and causes

their accumulation into the lung in vivo. *Biochem Biophys Res Commun* 197, 1167-1172

187. Baggiolini, M. (1996) Eotaxin: a VIC (very important chemokine) of allergic inflammation? *J Clin Invest* 97, 587

188. Crump, M. P., Spyracopoulos, L., Lavigne, P., Kim, K. S., Clark-lewis, I., and Sykes, B. D. (1999) Backbone dynamics of the human CC chemokine eotaxin: fast motions, slow motions, and implications for receptor binding. *Protein Sci* 8, 2041-2054

189. Weng, Y., Siciliano, S. J., Waldburger, K. E., Sirotina-Meisher, A., Staruch, M. J., Daugherty, B. L., Gould, S. L., Springer, M. S., and DeMartino, J. A. (1998) Binding and functional properties of recombinant and endogenous CXCR3 chemokine receptors. *J Biol Chem* 273, 18288-18291

190. Ogilvie, P., Bardi, G., Clark-Lewis, I., Baggiolini, M., and Uguccioni, M. (2001) Eotaxin is a natural antagonist for CCR2 and an agonist for CCR5. *Blood* 97, 1920-1924

191. Martinelli, R., Sabroe, I., LaRosa, G., Williams, T. J., and Pease, J. E. (2001) The CC chemokine eotaxin (CCL11) is a partial agonist of CC chemokine receptor 2b. *J Biol Chem* 276, 42957-42964

192. Cortes, J., Haydock, S. F., Roberts, G. A., Bevitt, D. J., and Leadlay, P. F. (1990) An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of Saccharopolyspora erythraea. *Nature* 348, 176-178

193. Malpartida, F., and Hopwood, D. A. (1984) Molecular cloning of the whole biosynthetic pathway of a Streptomyces antibiotic and its expression in a heterologous host. *Nature* 309, 462-464

194. Tuan, J. S., Weber, J. M., Staver, M. J., Leung, J. O., Donadio, S., and Katz, L. (1990) Cloning of genes involved in erythromycin biosynthesis from Saccharopolyspora erythraea using a novel actinomycete-Escherichia coli cosmid. *Gene* 90, 21-29

195. Schroder, J. (1999) Probing plant polyketide biosynthesis. *Nat Struct Biol* 6, 714-716

196. Magnuson, K., Jackowski, S., Rock, C. O., and Cronan, J. E., Jr. (1993) Regulation of fatty acid biosynthesis in Escherichia coli. *Microbiol Rev* 57, 522-542

197. Vanden Boom, T., and Cronan, J. E., Jr. (1989) Genetics and regulation of bacterial lipid metabolism. *Annu Rev Microbiol* 43, 317-343

198. Florova, G., Kazanina, G., and Reynolds, K. A. (2002) Enzymes involved in fatty acid and polyketide biosynthesis in Streptomyces glaucescens: role of FabH and FabD and their acyl carrier protein specificity. *Biochemistry* 41, 10462-10471

199. Szafranska Anna, E. (2001) *Studies of malonyl transfer in type II polyketide synthases*, Chemistry, University of Bristol, Bristol

200. Kragelund, B. B., Andersen, K. V., Madsen, J. C., Knudsen, J., and Poulsen, F. M. (1993) Three-dimensional structure of the complex between acyl-coenzyme A binding protein and palmitoyl-coenzyme A. *J Mol Biol* 230, 1260-1277

201. Rangan, V. S., and Smith, S. (1997) Alteration of the substrate specificity of the malonyl-CoA/acetyl-CoA:acyl carrier protein S-acyltransferase domain of the multifunctional fatty acid synthase by mutation of a single arginine residue. *J Biol Chem* 272, 11975-11978

202. Crump, M. P., Crosby, J., Dempsey, C. E., Murray, M., Hopwood, D. A., and Simpson, T. J. (1996) Conserved secondary structure in the actinorhodin polyketide synthase acyl carrier protein from Streptomyces coelicolor A3(2) and the fatty acid synthase acyl carrier protein from Escherichia coli. *FEBS Lett* 391, 302-306

203. Crump, M. P., Crosby, J., Dempsey, C. E., Parkinson, J. A., Murray, M., Hopwood, D. A., and Simpson, T. J. (1997) Solution structure of the actinorhodin polyketide synthase acyl carrier protein from Streptomyces coelicolor A3(2). *Biochemistry* 36, 6000-6008