Universidade de Aveiro Departamento de Química
2015

Susana Seabra Aveiro

A proteína que liga ao hemo, p22HBP: um estudo por RMN da dinâmica e das interações hemo-proteína

The p22HBP heme binding protein: an NMR study of the dynamics and heme-protein interactions

Universidade de Aveiro Departamento de Química
2015

Susana Seabra Aveiro

A proteína que liga ao hemo, p22HBP: um estudo por RMN da dinâmica e das interações hemo-proteína

The p22HBP heme binding protein: an NMR study of the dynamics and heme-protein interactions

Tese apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Bioquímica, realizada sob a orientação científica do Professor Brian James Goodfellow do Departamento de Química da Universidade de Aveiro e da Professora Glória Cruz Ferreira, College of Medicine, Universidade do Sul da Flórida.

Apoio financeiro da Fundação para a Ciência e Tecnologia (FCT)- bolsa de Investigação FCT SFRH/BD/64519/2009 e projeto PTDC/QUI/64203/2006 financiado pelo Fundo Europeu de Desenvolvimento Regional FEDER e pelo programa Operacional de Fatores de Competitividade (COMPETE); da Universidade de Aveiro – Centro de Investigação em Materiais Cerâmicos e Compósitos (CICECO). Agradecemos à Rede Nacional de RMN (RNRMN), suportada por fundos da FCT.

To my precious, Alexandre and Beatriz

o júri

presidente

Professor Fernando Joaquim Fernandes Tavares Rocha
Professor Catedrático do Departamento de Geociências da Universidade de Aveiro

Professor Carlos Frederico de Gusmão Geraldes
Professor Catedrático da Faculdade de Ciências e Tecnologia da Universidade de Coimbra

Doutora Maria dos Anjos Lopez Macedo
Professora Auxiliar da Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa

Doutora Rita Maria Pinho Ferreira
Professora Auxiliar do Departamento de Química da Universidade de Aveiro

Doutor Jorge Silva Dias
Investigador Auxiliar da Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa

Doutor Brian James Goodfellow
Professor Auxiliar do Departamento de Química da Universidade de Aveiro

agradecimentos

Ao meu orientador, professor Brian, o meu sincero obrigado pela disponibilidade demonstrada ao longo destes anos, pela paciência, pelos ensinamentos, desafios e críticas. Obrigado também pela companhia e solidariedade no galão descafeinado!

À professora Glória, co-orientadora, agradeço imenso a oportunidade que me deu de trabalhar no seu laboratório na Universidade do Sul da Flórida. Agradeço também todos os esclarecimentos e ensinamentos ao longo destes anos.

Agradeço ao professor Jean Marc pela breve, mas muito enriquecedora, estadia no seu laboratório em Grenoble, pela hospitalidade e disponibilidade.

À rede nacional de RMN, agradeço a oportunidade de usar os espectrómetros podendo assim enriquecer a minha tese de doutoramento. Ao CICECO e à FCT pela atribuição da bolsa de doutoramento, sem a qual não teria a possibilidade de conduzir este trabalho.

Ao grupo de espectrometria de massa do departamento de Química, principalmente ao Rui, Rita e professora Rosário, por me terem adotado estes anos todos, sempre disponíveis a colaborar! Uma palavra muito especial para a minha Cristina, o meu ombro amigo, a minha conselheira, a que sempre fez de tudo para eu conseguir avançar, sem desmotivar! Que sempre me fez ver que o mais importante é mesmo a “nossa menina”! Cristina o meu sincero e sentido obrigado!

Ao laboratório do Professor Manuel Santos bem como da Professora Odete Silva agradeço a partilha de conhecimentos e de equipamentos.

Aos meus colegas e amigos do departamento de Química que ao longo destes anos me acompanharam nesta luta! Pelo companheirismo, pelas conversas, noitadas (poucas mas de qualidade!) o meu obrigado!

Às Anitas, que nesta última fase foram incansáveis! Momentos difíceis que foram ultrapassados, momentos maravilhosos que jamais esquecerei!

A todos os meus amigos, um grande bem-haja por toda a amizade, desabafos, palavras de encorajamento principalmente quando o desânimo dominava.

À minha madrinha Bia pelo reconhecimento, pelo orgulho que tem em mim e por todo o seu amor!

À minha adorada e maravilhosa família que, apesar de ser enorme, vive das alegrias de todos e sofre com a tristeza de todos! Carol, Caty, Priiimas...não tenho palavras para agradecer...tenho-vos a todas no meu coração!

Ao meu anjo da Guarda que me protege todos os dias e que me conforta nos dias menos bons.

Aos meus pais o meu sincero obrigado por acreditarem em mim, por todo o apoio e ajuda ao longo destes anos, muitas vezes sem ter como, e pelo amor incondicional.

Ao Alexandre, quero agradecer todos estes anos que tem estado ao meu lado, pelo amor e companheirismo, pela paciência que nem sempre foi suficiente! Pelas nossas conquistas estes últimos anos e por aquilo que agora vivemos. Foste e sempre serás o meu suporte! Este triunfo também é teu!

À minha princesa Beatriz, obrigada pelos milhares de beijinhos e abraços que me deu e pelas inúmeras vezes que me disse “Mamã és tão linda, gosto muito de ti” sempre que me sentia mais cansada e ansiosa. Obrigada por ter aceite a minha ausência por diversas vezes. Obrigada por compreender que a mamã nem sempre tinha tempo ou paciência para brincar! Foi por ti que segui em frente! E é por ti que vou continuar a lutar!

De coração cheio, o meu muito obrigado!

palavras-chave

p22HBP, hemina, protoporfirina IX, interação proteína-hemo, sobreexpressão de proteínas, marcação isotópica, extinção de fluorescência, constantes de dissociação, espectroscopia de ressonância magnética nuclear (RMN), mapeamento dos desvios químicos, dinâmica, relaxação transversal, relaxação longitudinal, NOE heteronuclear, tempo de correlação, tensor, mutagénese, quimérica.

resumo

O trabalho apresentado nesta Tese focou-se na dinâmica e nas interações moleculares da p22HBP e do complexo p22HBP-tetrapirrol, nomeadamente nos resíduos chave envolvidos nesta interação. Estudos prévios de modelação molecular identificaram três possíveis resíduos chave R56, K64 e K177 como sendo importantes na interação com os tetrapirróis, através de interações eletrostáticas com os grupos propionato do tetrapirrol. Foram desenhados e construídos variantes da p22HBP murina e foram desenvolvidos estudos de extinção de fluorescência e RMN para avaliar a integridade dos variantes e a sua interação com os tetrapirróis. Os mesmos estudos de modelação molecular identificaram ainda uma zona flexível (Y171-R180) na p22HBP que diminui a mobilidade com a interação do tetrapirrol. Para confirmar esta alteração de mobilidade, foram realizados estudos de dinâmica, baseados em RMN. Por fim, com o intuito de obter uma versão não funcional da p22HBP humana, foi planeada e construída uma versão quimérica da p22HBP humana. No futuro, esta nova versão da p22HBP quimérica, será importante para os estudos de *knockdown* envolvendo siRNA.

O capítulo um introduz uma revisão dos aspetos biológicos da p22HBP nomeadamente os estudos estruturais e as possíveis funções que foram identificadas. Os principais objetivos da tese são também apresentados neste capítulo. No capítulo dois é apresentada uma descrição detalhada dos diferentes vectores de sobreexpressão (pNJ2 e pet28-a) e dos métodos de sobreexpressão e purificação da p22HBP murina e respectivos variantes, bem como da p22HBP humana. Todos os sistemas de sobreexpressão e purificação utilizados obtiveram bons rendimentos e permitiram a marcação isotópica das proteínas. No capítulo 3 são apresentados os resultados de extinção de fluorescência para a interação da p22HBP murina e humana com hemina através das constantes de dissociação determinadas na ordem dos nanomolar. Os mesmos estudos foram realizados para os variantes da p22HBP murina, com alterações hidrofóbicas e de polaridade nos resíduos R56, K64 e K177. Em alguns casos, as constantes de dissociação determinadas são mais elevadas, embora não se tenham verificado alterações significativas na força da interação proteína-hemo. As interações tetrapirrólicas com a p22HBP foram também estudadas por espectroscopia de RMN, onde foram mapeadas as diferenças nos desvios químicos para identificar a localização da zona de interação. A localização da zona de interação dos variantes da p22HBP e a p2HBP humana mantém-se igual à p22HBP murina. No capítulo 4 encontram-se os resultados das experiências 2D e 3D realizadas na p22HBP humana, isotopicamente marcada com $^{15}N/^{13}C$, para identificar as ressonâncias da cadeia principal. 82% dos sistemas de spin da cadeia principal foram identificados através da comparação com a p22HBP murina, das titulações com PPIX e de cálculos teóricos baseados nos desvios químicos (Talos+). No capítulo 5 são apresentados os resultados das experiências de relaxação, usados para comprovarem a dinâmica do *loop* na p22HBP aquando da interação com o tetrapirrol. A proteína no seu todo move-se de uma forma isotrópica na forma livre e ligada. No entanto os resultados para comprovar as alterações de mobilidade no *loop* 171-180 na presença de hemo, foram inconclusivos uma vez que só a um resíduo foi atribuído um sistema de spin, e não foi indicativo da perda significativa de mobilidade. O último capítulo descreve o planeamento e a construção da p22HBP quimérica. Para tal, a sequência que codifica a hélix alfa 1 da p22HBP humana, no plasmídeo phHBP1, foi substituída pela sequência homóloga da SOUL humana, uma proteína com uma estrutura 3D semelhante mas não liga ao hemo. Os resultados no entanto demonstraram que ou a sequência não foi introduzida corretamente no plasmídeo ou o sistema de purificação não foi adequado.

keywords

p22HBP, hemin, PPIX, heme-binding, overexpression, isotopic labeling, fluorescence quenching, dissociation constants, nuclear magnetic resonance spectroscopy (NMR), chemical shift mapping, dynamics, transverse relaxation, longitudinal relaxation, heteronuclear NOE, correlation time, diffusion tensor, mutagenesis, chimeric.

abstract

The work presented in this Thesis investigates the dynamics and molecular interactions of p22HBP and the p22HBP-tetrapyrrole complex. Specifically, the key residues involved when a tetrapyrrole binds to p22HBP were sought. Previous molecular modelling studies identified three possible charged residues R56, K64 and K177 as possibly being important in tetrapyrrole binding via electrostatic interactions with the propionate groups of the tetrapyrrole. A number of variants of murine p22HBP were therefore prepared and fluorescence quenching and NMR used to verify the integrity of the variants and their interaction with tetrapyrrole. The same molecular modelling studies identified a mobile loop Y171-R180 in p22HBP that decreased in mobility on tetrapyrrole binding, therefore to confirm this mobility change dynamics studies based on NMR relaxation experiments were carried out. Finally in order to obtain a non heme-binding form of human p22HBP a chimeric p22HBP was designed and constructed. This construct, and the resulting protein, will be important for future siRNA knockdown studies where rescue or recovery of function experiments are required to prove the knockdown results.

Chapter one discusses the current state of the art in terms of the biological, structural and functional aspects of p22HBP. The main objectives of the Thesis are also introduced here. Chapter two presents a detailed description of the different expression vectors (pNJ2 and pet28-a) and procedures used for overexpression and purification of murine p22HBP and its variants and human p22HBP. All expression and purification systems used gave good yields and allowed isotopic labeling to be carried out. The fluorescence quenching results for tetrapyrrole binding to murine p22HBP and variants are presented in chapter three along with the dissociation constants that were found to be in the nanomolar range for wild type murine and human p22HBP. The same studies were performed for murine p22HBP variants, with hydrophobic and polar changes being introduced at R56, K64 and K177. The dissociation constants were found to double in some cases but no significant changes in the strength of hemin-protein interactions were observed. The tetrapyrrole interaction with p22HBP was also followed by NMR spectroscopy, where chemical shift mapping was used to identify binding pocket location. All the variants and wild type human p22HBP were found to bind at the same location. Chapter 4 contains the data from 2D and 3D experiments carried out on $^{15}N/^{13}C$ labelled human p22HBP that was used to obtain backbone assignments. Comparison with wild type murine p22HBP assignments, PPIX titrations and theoretical calculations based on chemical shifts (Talos+) allowed 82% of the backbone resonances to be assigned. The results from the relaxation experiments used to probe the dynamics of the mobile loop in p22HBP on binding to tetrapyrrole are presented in chapter 5. The overall protein was found to tumble isotropically in the free and bound forms however the results to probe mobility changes in the 171-180 loop on tetrapyrrole binding proved inconclusive as only residue could be assigned and this did not seem to become significantly less mobile. The final chapter describes the design and construction of a chimeric p22HBP. For these purpose, the alfa1-helix sequence of human p22HBP in the phHBP1 plasmid was replaced by its homologous sequence in hSOUL, a non heme-binding protein with identical 3D structure. The results however indicated that either the incorrect sequence was introduced into the plasmid or the purification procedure was inadequate.

Index

Abbreviations and symbols

°C	Celsius degrees
B^*_{tot}	Total bound ligand
BSA	Bovine serum albumin
C(t)	Correlation function
ckSOUL	Chicken SOUL
COSY	Correlation spectroscopy
CSA	Chemical shift anysotropy
CSI	Chemical Shift Index
DD	Dipole-dipole
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotide triphosphates
DTT	Dithiothreitol
E.coli	Escherichia coli
EDTA	Ethylenediamine tetraacetic acid
FECH	Ferrochelatase
FKBP	FK506-binding protein
FMF	Fast Model Free
FPRL2	Formil pepide receptor like 2
FT	Fourier transform
GF	Gel filtration
hHBP	Human Heme Binding protein
HIC	Hydrophobic interaction chromatography
HSA	Human Serum Albumin
hSOUL	Human SOUL
HSQC	Heteronuclear Single Quantum Coherence
IMAC	Immobilized metal ion affinity chromatography
IPTG	Isopropyl-ß-D-thio-galactoside
Ka	Affinity constant
Kd	Dissociation constant
K_{off}	Dissociation rate
K_{on}	Association rate
LB	Luria Broth media

MES	2-(N-morpholino)ethanesulfonic acid
mHBP	murine Heme Binding Protein
mRNA	Messenger Ribonucleic Acid
mSOUL	Murine SOUL
NADH	Reduced nicotinam dinucleotide
Ni-NTA	Nickel-nitrilotriacetic acid
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
NOESY	Nuclear Overhauser Effect Spectroscopy
OD	Optical Density
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
phoA	Alkaline phosphatase
PPIX	Protoporphyrin IX
PTNMR	Portuguese Nuclear Magnetic Resonance Network
R_1	Longitudinal relaxation rate
R_2	Transverse relaxation rate
R_{ex}	Conformational exchange contribution
r_H	Effective hydrodynamic radius
RMSD	Root Mean Square Deviation
RNA	Ribonucleic acid
rpm	Rotation per minute
RSS	Residual sum square
R_{tot}	Total receptor
S^2	Generalized order parameter
siRNA	Small Interfering Ribonucleic acid
SDS-PAGE	Sodium dodecyl sulfate- polyacrylamide gel electrophoresis
SOC	Super Optimal broth with Catabolic repressor
T_1	Longitudinal relaxation time
T_2	Transverse relaxation time
TB	Terrific Broth Media
TBE	Tris/Borate/EDTA buffer
TMS	Tetramethylsilane

TOCSY	Total Correlation Spectroscopy
TRBP	Transactivating Response RNA-binding protein
TROSY	Transverse Relaxation Optimized Spectroscopy
TSP	3-(trimethylsilyl)propionate
U	enzyme unit
γ	Magnetogyric ratio
η_w	Viscosity of the solvent
κ_B	Boltzmann constant
τ_c	Correlation time
J_ω	Spectral density
ν_0	Larmor frequency

1 Introduction

1.1 Heme

In the human body, 65% to 75% of the total iron is present as heme iron in red blood cells. Heme and its active forms play important physiological roles such as electron transfer, O_2 transport and storage. Because of the toxicity and low solubility of heme, the intracellular level of uncommitted heme is maintained at a low concentration ($<10^{-9}$ M) [1]. In heme proteins, heme is involved in many aspects of oxidative metabolism and function both as an electron carrier and a catalyst for redox reactions [2]. The heme taken up by cells plays roles in their proliferation and differentiation, in mediating gene expression at the level of transcription, and by working as a regulatory molecule [1]. Despite being ubiquitous, free heme has inherent peroxidase activity and can intercalate and disrupt lipid bilayers of cell membranes, resulting in cytotoxicity[3]. It can also cause oxidative stress by generating reactive oxygen species[4], produce DNA damage, lipid peroxidation and protein denaturation [5].

Heme is composed of an iron molecule tetra-coordinated at the center of a large heterocyclic organic ring called a porphyrin. The insertion of ferrous iron into the tetrapyrrole macrocycle of Protoporphyrin IX (Figure 1.1) is catalyzed by ferrochelatase, an enzyme which resides in the mitochondrial matrix [6].

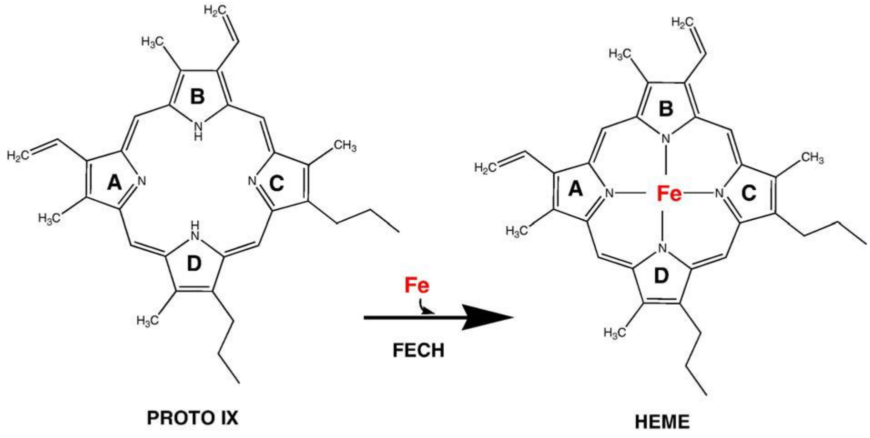


Figure 1.1. Heme synthesis. The final step in heme synthesis is catalyzed by the mitochondrial enzyme ferrochelatase (FECH). The enzyme catalyzes the insertion of one atom of ferrous iron (red) into the Protoporphyrin IX macrocycle [2].

The heme biosynthetic pathway can be broken down into four basic processes: formation of the pyrrole; assembly of the tetrapyrrole; modification of the tetrapyrrole side chains; oxidation of protoporphyrinogen IX to Protoporphyrin IX and finally the insertion of iron

[2]. These processes require eight enzymes, four acting in the mitochondria and the remaining four acting in the cytosol (Figure 1.2) [7]. The first enzyme of the heme biosynthetic pathway is aminolevulinate synthase, which catalyzes the condensation of glycine and succinyl-CoA to form 5-aminolevulinic acid [8]. At least in non-erythroid cells, this reaction is the rate limiting step in heme production [9]. The first and the last three steps of heme biosynthesis occur in the mitochondria, whereas all remaining steps occur in the cytosol. The intermediates formed during biosynthesis must therefore cross the mitochondrial membrane (Figure 1.2) [3].

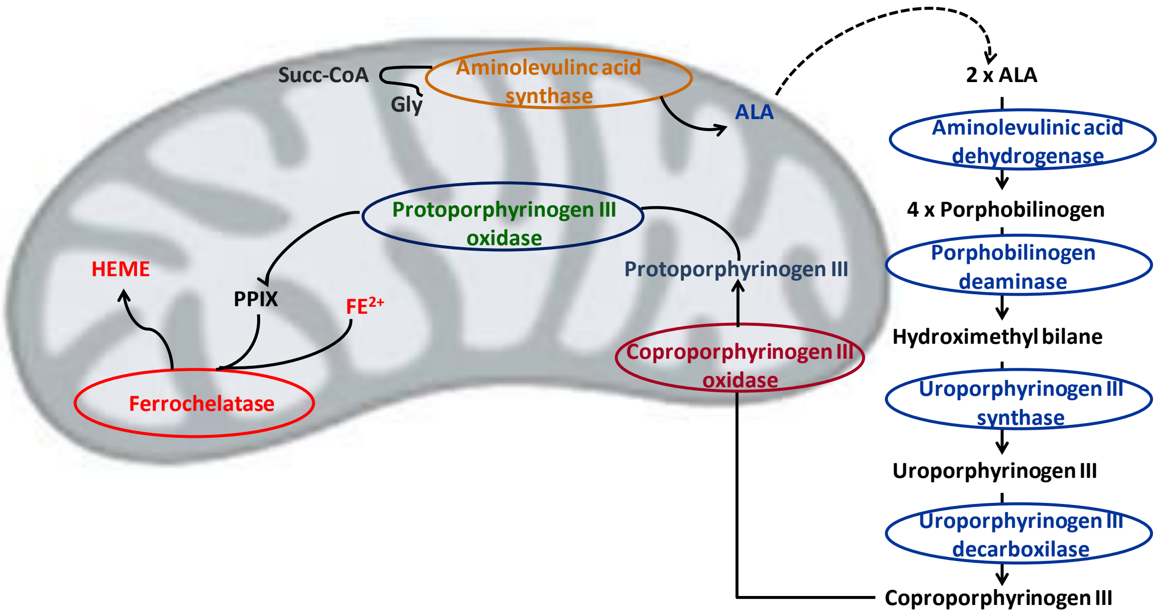


Figure 1.2. Schematic representation of the heme biosynthetic pathway in mammals. The different enzymes are compartmentalized between the mitochondria (represented in gray) and the cytosol. Adapted from reference [10] .

Porphobilinogen synthase is the second enzyme of the heme biosynthetic pathway (Figure 1.2) and it catalyzes the asymmetric condensation of two molecules of 5-aminolevulinic acid to yield the monopyrrole, porphobilinogen. The third and fourth enzymes of the heme biosynthetic pathway, porphobilinogen deaminase and uroporphyrinogen III synthase, catalyze the conversion of four molecules of porphobilinogen into uroporphyrinogen III. Uroporphyrinogen decarboxylase catalyzes the decarboxylation of uroporphyrinogen III to coproporphyrinogen III. Coproporphyrinogen oxidase, the antepenultimate enzyme of the heme biosynthetic pathway, catalyzes the conversion of two propionate groups at positions two and four of coproporphyrinogen III to two vinyl groups. The product of the coproporphyrinogen oxidase-catalyzed reaction is, therefore, protoporphyrinogen III. The six-electron oxidation of protoporphyrinogen III into Protoporphyrin IX is the next step and it is catalyzed by Protoporphyrinogen Oxidase. Finally ferrochelatase catalyzes the insertion of ferrous iron into the Protoporphyrin IX to yield Heme [1]–[3], [7].

Heme synthesis takes place mostly in developing red blood cells in the bone marrow but about 15% of the daily production occurs in the liver for the formation of heme-containing enzymes [2]. Because both iron overload and iron deficiency are incompatible with normal body physiology, mammals regulate their iron levels at both the systemic and cellular levels [3].

The mechanisms of heme uptake, catabolism and trafficking in the cells are linked to cytosolic heme binding proteins. These are responsible for the intracellular transient transport of heme from the place of its enzymatic synthesis to the site of hemeprotein synthesis and from the site of degradation of hemeproteins to the site of enzymatic heme degradation [1]. A number of heme binding proteins have been isolated and characterized based on their ability to bind heme such as FABP, GSTs, HBP23 [11] and p22HBP [12]. However, since there are few studies of the functions of these proteins, their participation in cellular regulation by heme and the intracellular transport of heme still remains poorly understood [13].

1.2 p22HBP

p22HBP (Heme Binding Protein) is a 22 kDa protein expressed constitutively in various tissues with highest mRNA levels seen in liver, spleen and kidney [12], [14]. p22HBP was initially purified from mouse liver cell extracts and was characterized as a cytosolic heme-binding protein due to its high affinity for hemin. In addition to hemin, p22HBP can also bind intermediaries from heme biosynthesis, such as Protoporphyrin IX and coproporphyrinogen, other porphyrins, bilirubin and fatty acids [12], [14]. In mice, the gene that codifies p22HBP is located on chromosome 6 and encodes for a 190 aminoacid protein while in humans the gene is located on chromosome 12 and encodes for a 189 aminoacid sequence. These two proteins have an homology of 87% (Figure 1.3) [15].

p22HBP belongs to an evolutionary conserved heme binding protein family, with a number of putative members in animal, plant and bacterial species. The SOUL protein, or heme binding protein 2, a p22HBP homologous protein, has been identified in chicken (ckSOUL), murine (mSOUL) and in humans (hSOUL). Figure 1.3 shows the protein sequence alignment between these proteins. mSOUL has 27% identity to p22HBP from the same organism. The heme-binding properties and coordination structure of SOUL are distinct from those of p22HBP [16]. hSOUL was initially identified as a heme binding protein, and biochemical studies performed by Sato *et al* in 2004 revealed that the protein specifically binds one heme per monomer [16]. Comparison of the SOUL-encoding sequence, with those of human and murine p22HBP led Lathrop *et al.* [17] and Shin *et al.* [18], to conclude that possibly SOUL does not have any heme-binding motif, although it has some hydrophobic amino acid-rich segments. More recently, Freire *(2012)* [19] confirmed that hSOUL does not bind heme using chemical shift mapping by NMR spectroscopy. Finally it was found that SOUL protein can promote mitochondrial permeability transition and facilitate necrotic cell death under different types of stress conditions [20].

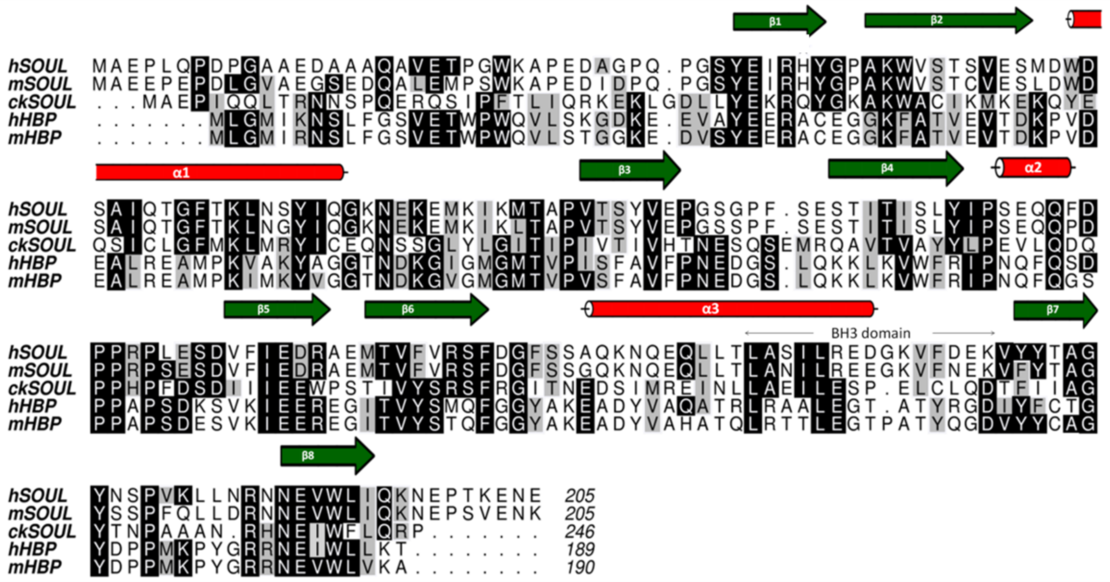


Figure 1.3. Protein sequence alignment between murine and human p22HBP, ckSOUL and mammalian SOUL using CLUSTALW. Identical amino acids are shaded in black and similar residues are shaded in grey. Adapted from [16], [21].

The three-dimensional structure of murine p22HBP (Figure 1.4), the first for a protein from HBP/SOUL family was determined, in 2006 by NMR methods. The protein consists of a 9-stranded distorted β-barrel flanked by two long α-helices. Each α helix packs against a four-stranded sheet in an equivalent way, such that the β_2-β_3-α_A-β_4-β_5 subdomain (residues 21-105) is equivalent to the β_6-β_7-α_B-β_8-β_9 subdomain (residues 114-190) [10]. This type of conformation suggests symmetry evocative of an ancestral gene-duplication event (Figure 1.4). The work of Dias *et al.* in 2006 concluded that heme binding to murine p22HBP was via a hydrophobic pocket on the surface of the protein with the centre of the heme ring located near M63 (Figure 1.5).

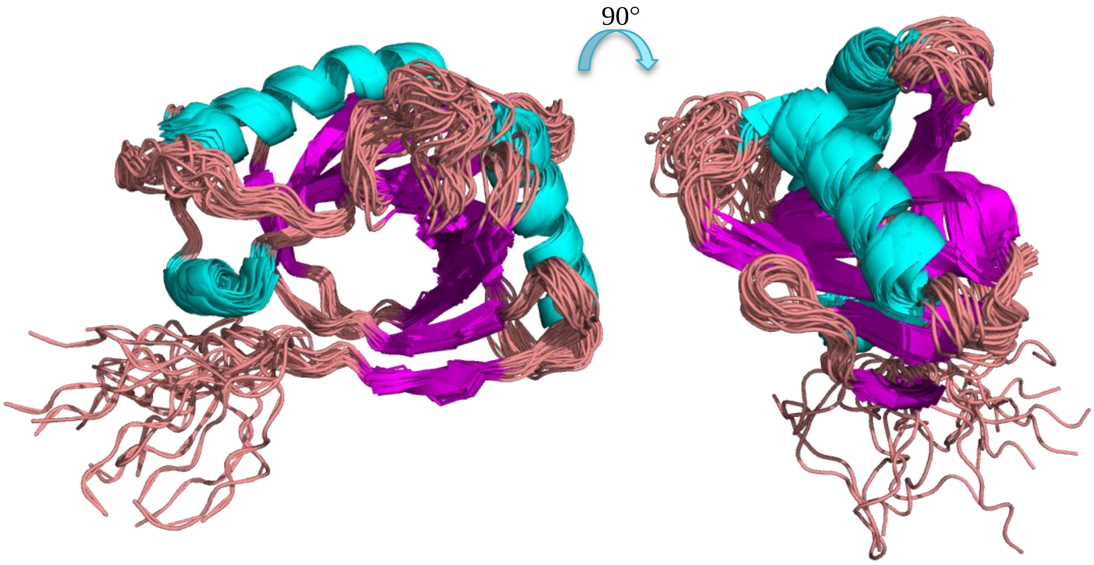


Figure 1.4**.** Two views of the 20 murine p22HBP conformers (pdb 2GOV).

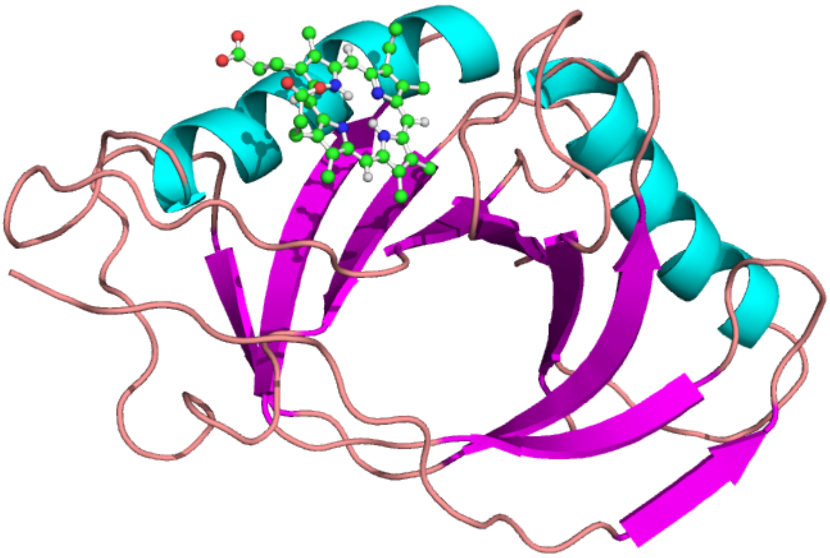


Figure 1.5. Model of the p22HBP-hemin complex. The binding location was determined by minimizing the differences between experimental chemical shift differences and calculated PPIX ring current shifts [10].

The study used ring current shifts and chemical shift mapping to localise the tetrapyrrole. However, due to the symmetrical nature of tetrapyrrole ring current shifts the position of the propionate groups could not be determined. In order to characterize further the molecular recognition process that takes place when heme binds to p22HBP a molecular modelling study has been carried out by Micaelo *et al.* Here heme and heme intermediates involved in heme synthesis were studied. These results confirmed that the p22HBP binding pocket is essentially composed of nonpolar residues that create a hydrophobic binding region exposed to the solvent and that this binding pocket is conserved for both the murine and human proteins [22]. This study also found that hemin and PPIX have identical binding orientations (Figure 1.6 and Figure 1.7), in which the stabilization of the propionate side chains is mainly achieved by electrostatic interactions with R56, K64 and K177 (K176 in human pp2HBP) located at the edges of the protein-binding pocket. The sequentially (and structurally) conserved lysine and arginine residues seem to play an identical role in stabilizing the tetrapyrrole propionate side chains in both proteins [22].

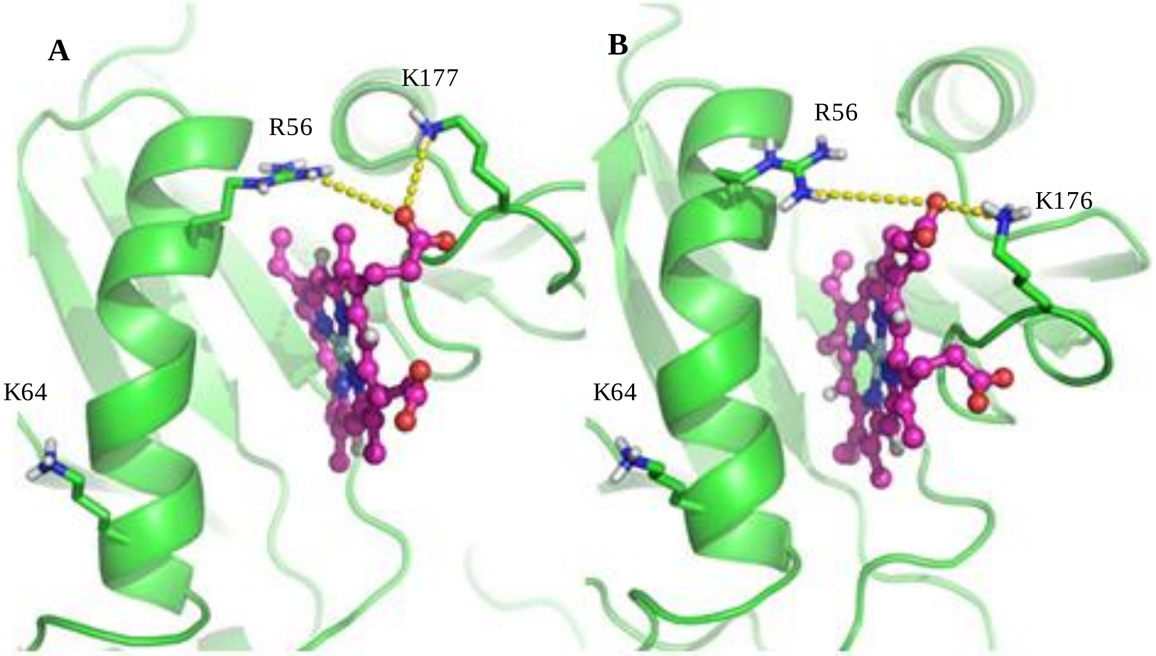


Figure 1.6**.** Representative structure of the Hemin-mHBP (A) and Hemin-hHBP (B) complexes. The binding site of each complex is shown with Hemin rendered in ball and stick. The protein is rendered in cartoon. Key side chain residues are rendered in sticks. Adapted from [22].

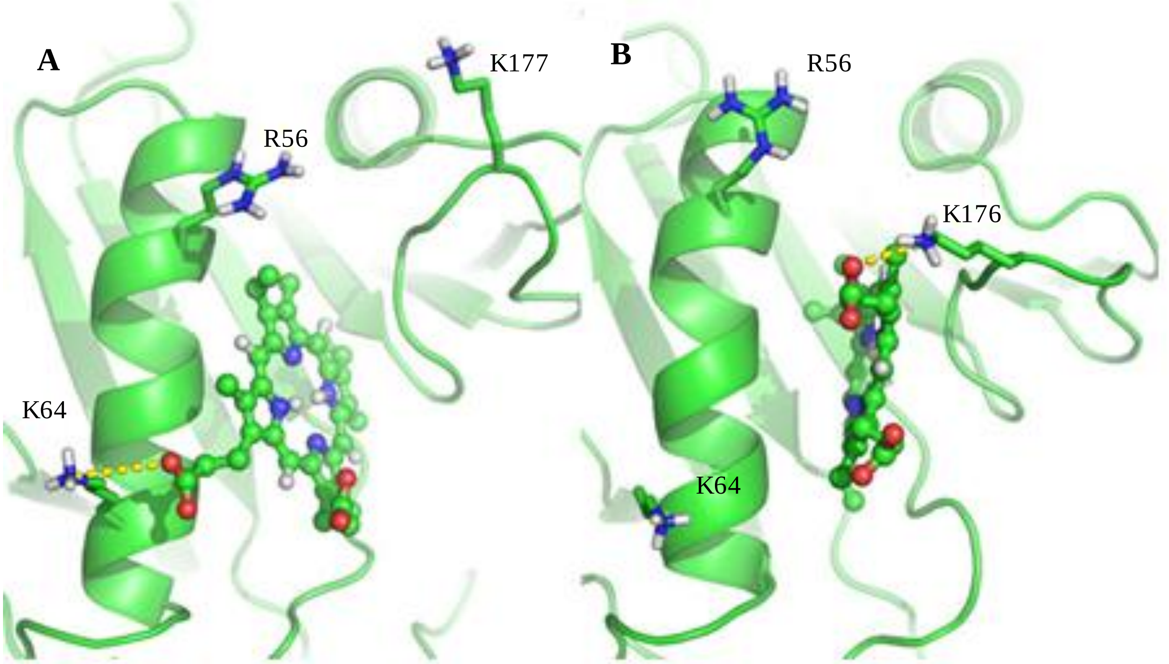


Figure 1.7. Representative structure of the PPIX-mHBP (A) and PPIX-hHBP (B) complexes. The binding site of each complex is shown with PPIX rendered in ball and stick. The protein is rendered in cartoon. Key side chain residues are rendered in sticks. Adapted from [22].

A preliminary 3D structure of hSOUL was proposed by Freire *et al.* [23] where molecular replacement, using the structures of murine p22HBP (pdb 2GOV) as search models, allowed a preliminary model to be obtained (pdb 4AYZ). More recently (2011) Ambrosi *et al* [24] solved the crystal structure of human SOUL BH3 domain in complex with Bcl-xL by X-ray crystallography (pdb 3R85). It was found that although the 3D structures of hSOUL and p22HBP are very similar (RMSD 3.26) hSOUL has no hydrophobic patch near the a1-helix and therefore does not bind heme.

Although a structure is available for p22HBP its function remains unknown. In 2002 Blackmon *et al.* reported K_d values of the order of µM, for p22HBP upon binding heme and other tetrapyrroles (hemin and Protoporphyrin XI) [14]. Babusiak *et al.* carried out a mass spectrometry based proteomic study, using erythrocyte precursor cells labelled with ^{59}Fe-hemin. They demonstrated that p22HBP was a component of one of the four multiproteic complexes identified in hemoglobin synthesis. It was suggested that p22HBP can represent an heme transporter or a chaperone for the insertion of heme into hemoglobin or, in addition, a regulator of coproporphyrinogen transport into mitochondria [25]. More recent studies have suggested that p22HBP has potent chemoattractant activity, related to infection and apoptosis. In these studies, an acetylated N-terminal fragment (residues 1-21) of p22HBP was purified from porcine spleen extract, named F2L, and was subsequently found to selectively recruit leukocytes by activating a G-protein coupled receptor, the formyl peptide receptor like 2 (FPRL2), expressed specifically on monocyte and dendritic cells[26], [27].

Preliminary functional studies of p22HBP were undertaken in collaboration with Jean-Marc Moulis (data not published). siRNA (small interfering RNA) experiments were used. Here small RNA oligonucleotides (21-25 bp) that are chosen to specifically bind to the complementary sequence of an mRNA which results in repression of the translation of the mRNA target [28], [29]. Briefly, the technique works by hijacking an endogenous process involving RNAse III (Dicer) and its co-factor TRBP (transactivating response RNA-binding protein) that act with the Argonaut protein family. The complex is called the RNA-Induced silencing complex (RISC). Dicer cleaves long double-stranded RNA (dsRNA) molecules into short double-stranded fragments with two unpaired nucleotides at each 3' end, giving siRNA. Each siRNA is converted into two-single stranded RNAs: a passenger strand, which is eliminated, and a guide strand which is incorporated into RISC. When the guide strands pairs with a complementary sequence on an mRNA molecule, cleavage is induced by Argonaut and the target gene is silenced [30]. Using this method K562 (human erythroleukemia cells) and HepG2 (human hepatocarcinoma) cells were transfected with control siRNA (no targeting of p22HBP) and siRNA targeting 3 different siRNA sequences for p22HBP. RT-PCR (Real Time Polymerase Chain Reaction) indicated that siRNA knock-down of p22HBP was occurring. These cells were used to study if the first step of heme biosynthesis was sensitive to the loss of p22HBP by analyzing ALAS2 expression. No changes for ALAS2 mRNA band were found and no back regulation on heme biosynthesis was observed. However, the expression of Heme oxygenase-1 was also studied and it was found that HMOX-1 was strongly upregulated in p22HBP knockdown cells. This is a strong indication that p22HBP is involved in heme transport or regulation.

1.3 Objectives

The overall aim of the research carried out for this thesis was to probe, in more detail than has been carried until now, the dynamics and molecular interactions (including any key residues involved in binding) involved in tetrapyrrole binding to p22HBP in order to identify key residues involved in their interaction with Hemin and PPIX. The main techniques used to carry these studies include molecular biology, NMR spectroscopy and Fluorescence Quenching.

Using results from molecular modelling studies a number of basic amino acids were identified as having interactions with the propionate groups of heme/PPIX and therefore are possible targets for site-specific mutagenesis studies. A number of variants of murine p22HBP were prepared using molecular biological techniques in collaboration with the USF, Tampa, USA. These variants were studied by NMR, to assign peaks in the HSQC spectra, and by fluorescence quenching to study tetrapyrrole-protein interactions. The human form of p22HBP was also prepared and studied by NMR in order to assign the backbone resonances and to subsequently study tetrapyrrole binding. The dynamics of the protein backbone in solution and when bound to PPIX was also analysed using NMR relaxation studies. As functional studies involving siRNA knockdown normally require a functional and non-functional form of the knocked-down protein to perform a recovery of function or rescue experiment a chimeric human p22HBP was also constructed. Chimeric proteins have found wide application for the study of protein folding and protein structure stability [31]. Chimeric proteins are created through the joining of two or more function domains, which originally coded for domains in separate proteins, using recombinant DNA technology. Translation of this fusion gene results in a single polypeptide with functional properties derived from each of the original proteins. This approach was used to attempt to produce a non heme-binding version of human p22HBP. The α1-helix that interacts with heme/PPIX in p22HBP was replaced with the corresponding α1-helix found in hSOUL, a protein with an identical 3D structure to p22HBP but does not bind heme/PPIX.

2 p22HBP cloning, overexpression and purification

2.1 Introduction

Due to the inherent insensitivity of NMR spectroscopy, the first challenge of any dynamic and structural studies by NMR is the production of enough amount of pure and correctly folded protein. The minimum concentration of a protein for NMR studies is approximately 0.5 mM. In order to obtain this amount of protein cloning and protein overexpression is carried out. Typically a bacterium such as *E. coli* is used to perform this overexpression, as these systems are cheap and have been utilized extensively over the last 20 or 30 years. Therefore the first goal was to optimize the overexpression of p22HBP for both the murine and human forms and also for the variants.

The primary choices for initial screening of recombinant expression systems are bacterial cells due to the cost-effectiveness, well characterized genetics and the large range of many different bacterial expression strains. *E.coli* growth is performed in chemically defined media where it can reach high densities. The widely used rich medium called Luria-Bertani broth (LB) is popular with bacteriologists because it permits fast growth and good growth yields for many species and supports *E.coli* growth to an optical density at 600 nm (OD_{600}) of 7 [32]. Moreover, *E.coli* accepts a wide variety of plasmids with a diversity of copy number (both low and high); and can be obtained with a wide range of plasmid systems that facilitate high-throughput cloning and expression [33]. Despite these advantages, there are some drawbacks of using *E. coli* for recombinant protein production such as no secretion systems for efficient release of proteins to the growth medium and the inability to perform extensive disulfide-bond formation and other posttranslational modifications [34]. Furthermore, proteins expressed in large amounts can precipitate into insoluble aggregates called inclusion bodies; and initial lysis steps to recover cytoplasmatic proteins often results in the release of endotoxins, which must be removed from the final product[35].

Several vectors such as lambda phages, plasmids and cosmids, can achieve the insertion of DNA into host cells. Plasmids are widely used as they ensure easy insertion of the recombinant DNA into the host cells. The plasmid vectors must possess specific features to perform overexpression: they must have at least one DNA sequence which acts as an *origin of replication*, so they are able to multiply within the cell and they normally contain an antibiotic resistance gene often used as a *selection marker* to ensure that only the bacteria with the plasmid of interest can grow in a culture. A *promoter* also has to be present in the vector to induce protein expression and to reach suitable yields.

One of the most popular commercial vector systems for *E.coli* is the Novagen pET system based on the T7 promoter. In pET systems, the plasmid containing the target gene is transferred to an *E. coli* strain containing a chromosomal copy of the gene for T7 RNA polymerase. There are a large number of cell strains that can be used in order to express proteins in *E. coli*. The most widely used cell strain is BL21 which does not encode the Ion, ompT, and ompP proteases, thus minimizing protein degradation during purification. Furthermore, there are different strains within BL21. For example, BL21(DE3) contains a T7 RNA polymerase gene that is integrated into the genome and that is under control of the lacUV5 promoter, which is a lactose analog IPTG inducible system. This means that the addiction of IPTG to a medium containing these cells will induce T7 RNA polymerase production which will ultimately transcribe the target gene in the plasmid (see Figure 2.1).

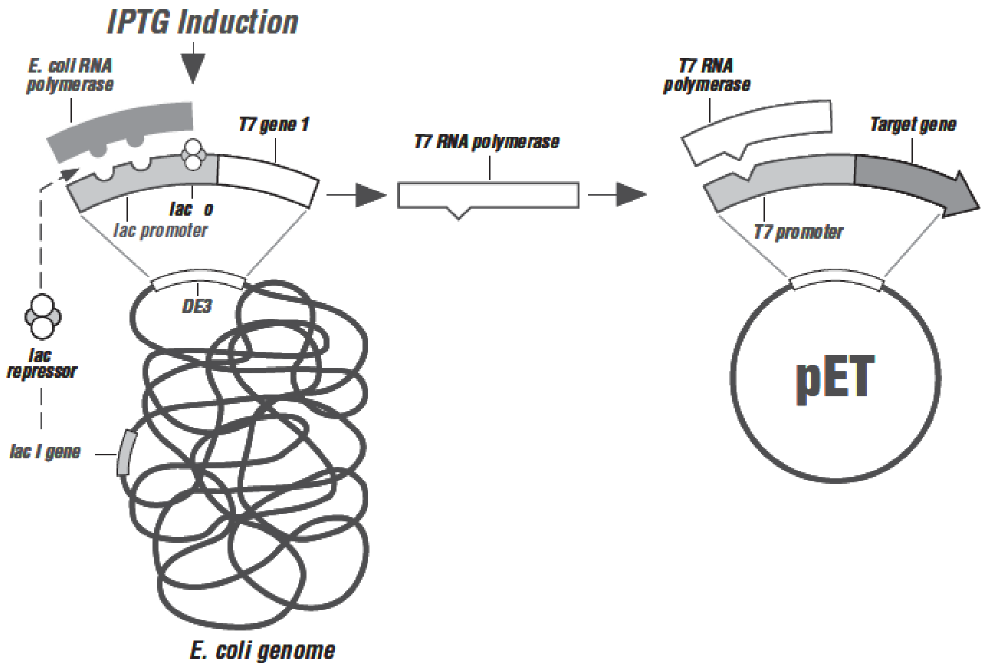


Figure 2.1. Induction of the lac gene by IPTG, in the pET system[36].

After the overexpression step, the purification of proteins from native or recombinant sources, in their biologically active form follows and represents a step of major importance in molecular biology. It must be remembered that proteins are easily degraded and modified by microorganisms, very sensitive to temperature, pH, salts, metals and it is of extreme importance to plan precisely the purification procedure in terms of buffer composition, temperature, sequence and duration of the steps. Chromatographic methods are used extensively for the purification of proteins. There are a range of techniques available including anion and cation exchange, which can be carried out at different pHs, hydrophobic interaction chromatography, gel filtration and affinity chromatography [37]. Normally a stationary phase (matrix) is packed in a column through which a mobile phase, which contains the protein, is eluted. The proteins in the mobile phase are adsorbed on the stationary phase to be later removed by elution with a suitable mobile phase resulting in separation of the different molecules. Depending on the column, proteins can be separated by means of charge (ionic change chromatography), hydrophobicity (HIC – chromatography of hydrophobic interaction), size (GF – gel filtration) or capacity to link specific chemical groups (affinity chromatography – Ni-Sepharose) [38].

Immobilized metal ion affinity chromatography (IMAC) is one of the most useful methods for purifying recombinant proteins as this tag can be easily introduced at the N- or C-termini of cloned proteins. As histidine can form a complex with divalent metals around neutral pH a polyhistidine-tag binds strongly to a stationary phase that contains Ni^{2+}. Therefore separation of the histidine-tagged proteins from most untagged proteins even under denaturing conditions can be carried out. The binding can be disrupted by eluting the tagged sample from the resin by reducing the pH and increasing the ionic strength of the buffer or by including EDTA (Ethylenediamine tetraacetic acid) or imidazole in the buffer [39]. In order to utilize this purification method all murine and human p22HBP variants were cloned with N-terminal hexahistidine-tags.

2.2 Murine p22HBP

A plasmid, pNJ2, containing the gene encoding murine p22HBP lacking the first 6 amino terminal residues was obtained with a hexahistidine tag at the N-terminus resulting in a N-terminal of: MKQSTHHHHHHN. The pNJ2 expression plasmid was generated at the USF, Tampa, by subcloning a PCR-amplified fragment for heme-binding protein p22HBP – from Taketani *et al.* (1998) [12] into pGF23 [64] (a derivative of CASS-3, [65]) previously digested with Sal I and Bam HI. This pNJ2 plasmid (Figure 2.2) was then transformed into BL21 (DE3) cells and grown in a modified version of MOPS minimal medium (phosphate limiting) with ampicilin as selection. Details can be found in Dias *et al* [40].

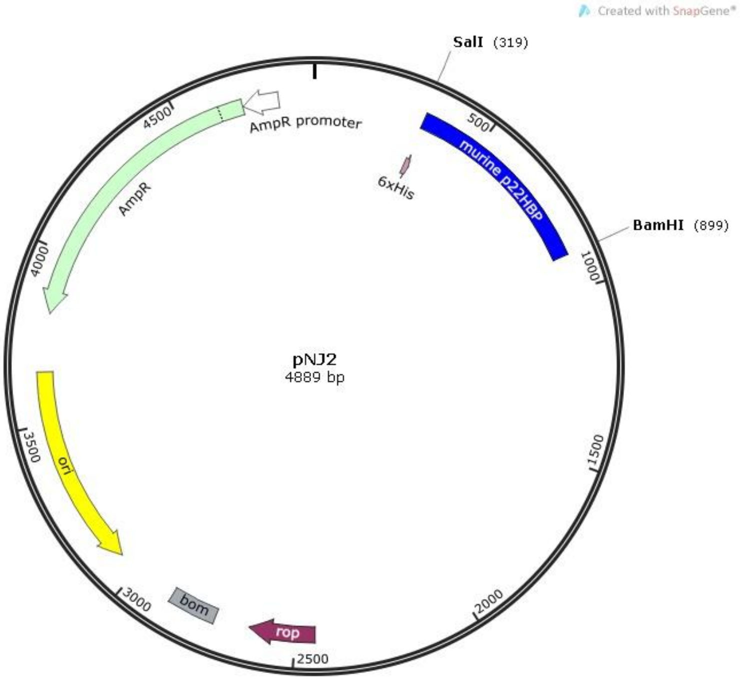


Figure 2.2. Schematic representation of pNJ2 plasmid. Constructed by Naomi Jeanty and Professor Glória Cruz Ferreira, University of South Florida.

The cells were grown using a standard procedure in which LB medium is used to achieve good cell density and a MOPS minimum medium is subsequently use for induction. For overexpression purposes, starter cultures were prepared by inoculating 20 µL of cells taken from glycerol stocks in 20 mL of LB media (Nzytech) previously autoclaved and supplemented with ampicilin 50 mg/mL, and incubated at 37ºC with shaking at 180 rpm for 12 to 16 hours (overnight). The culture was then inoculated into 1 L LB medium, enriched with 1 mL ampicilin 50 mg/mL and incubated for 5 hours at 37 ºC, 180 rpm. The cells were then harvested by centrifugation at 8000 rpm for 10 min at 4ºC, ressuspended into an autoclaved MOPS medium supplemented with 1 mL ampicilin 50 mg/mL, 2 mL O solution, 100 µl P solution, 1mL S solution, 0.5 Thiamine/HCl 0,2%, 10 mL $NaHCO_3$ 1M, 4 g glucose and 1 g NH_4Cl. See appendix 9.1 for MOPS and O, P, S solutions receipts. For isotopic labelling, 1 g of ^{15}N-NH_4Cl was added to MOPS media, and this was the unique source of Nitrogen for *E.coli* growth. Ressuspended cells were incubated at 30ºC, 180 rpm, for 12-16 hours and enriched MOPS media was the inductor of murine p22HBP overexpression. The cells were recovered by centrifugation at 8000 rpm for 10 minutes and ressuspended in a 50 mM KH_2PO_4/ 300 mM NaCl buffer at pH 8.0. Incubation for 1 hour with Lysozyme 10 mg/mL, with gently mixing, was performed to start cell lysis. Cell lysate was sonicated for 6 minutes (59 seconds with 59 second intervals) while the suspension was kept on ice to prevent heating. Centrifugation at 20000 rpm at 4ºC for 1 hour was applied to separate the homogenate components into cytoplasmatic fraction (supernatant) and cell pellet. The supernatant was then applied to a Ni-NTA (Nickel-nitrilotriacetic acid) column with Sepharose matrix functionalized with chelating groups charged with Ni^{2+} for purification. The system was equilibrated with 50 mM KH_2PO_4, 300 mM NaCl buffer, pH 8, and the purification procedure was performed under native conditions. His-tagged protein was eluted using an imidazole gradient (10-500 mM). The purest fractions containing murine p22HBP were pooled together and concentrated using 10 kDa centricons (Millipore) to 2.5 mL. This volume was applied to a PD10 Desalting column (Amersham Biosciences), previously equilibrated with 20 mL of 50 mM phosphate buffer and eluted with this buffer to a final volume of 3.5 mL. SDS-PAGE (Sodium dodecyl sulfate- polyacrylamide gel electrophoresis) 15% (Mini-Protean Biorad) was used to examine purity of the protein separation as well as to determine its molecular weight.

Purified protein samples were quantified by UV-Vis spectrophotometric method, according to the Lambert-Beer law, by measuring the absorption at 280 nm. The experimental value of molar absortivity for murine p22HBP variants was found to be 31574 $M^{-1}cm^{-1}$.

In Figure 2.3, the protein expression profile after purification with an imidazole gradient in a Ni-NTA column is shown. The band for murine p22HBP appears at approximately 24 kDa as expected. The gel also reveals that the imidazole gradient used was appropriate to obtain the purest fraction of murine p22HBP (75 mM imidazole). On the right side of

Figure 2.3, a UV-Vis spectra of murine p22HBP (diluted 1:50) is shown. An absorbance of 0.165 was obtained at 280 nM which corresponds to 0.26 mM murine p22HBP.

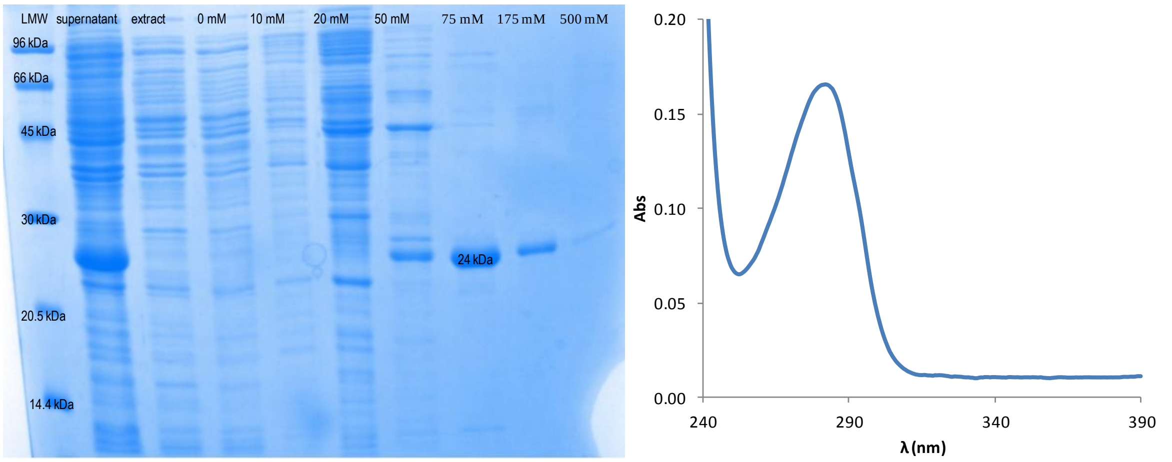


Figure 2.3. Left: SDS-Page analysis of the different fractions obtained from a Ni-NTA Agarose column for murine p22HBP. Right: UV/Vis spectrum of murine p22HBP 0.26 mM.

2.3 Murine p22HBP variants

Murine p22HBP site specific variants were prepared based on the results of molecular modelling studies carried out by Micaelo *et al.* in 2010. The basic aminoacids arginine 56, lysine 64 and lysine 177, thought to be important for heme-binding, were the initial targets for the mutagenesis studies. This work was carried out at the College of Medicine, University of South Florida under the supervision of Professor Glória Ferreira.

The strategy used to construct these variants was based on whole plasmid PCR. In this approach, a two-step polymerase chain reaction (PCR) is used to first create a megaprimer with the desired mutation and then to insert this mutation into a plasmid (Figure 2.4).

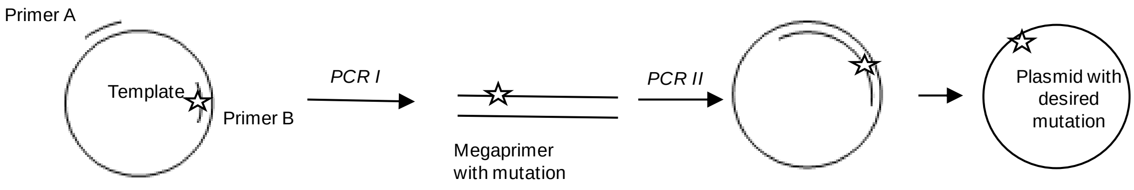


Figure 2.4. Schematic representation of megaprimer and whole plasmid principle. Adapted from [63].

All enzymes used were purchased from New England Biolabs. Primers with desired mutations (Table 2.1) used for the PCR amplifications were purchased from *Integrated DNA Technologies*, and were as follows:

Primer	**Variant**	**Sequence**
HBP7	K177A	GACCCTCCCATG**GCG**CCCTATGGACGC
HBP8	K64A	GCCCAAGATCATG**GCG**TATGTGGGTGGCACC
HBP9	K177E	GACCCTCCCATG**GAG**CCCTATGGACGCCGT
HBP10	R56E	GTGGATGAGGCTCTC**GAG**GAAGCGATGCCCAAG
HBP11	K64E	GCCCAAGATCATG**GAG**TATGTGGGTGGCACC
HBP12	R56A/K64A	GTGGATGAGGCTCTC**GCG**GAAGCGATGCCCAAGATCATG**GCG**TATGTGGGTGGCACC

Table 2.1. *Primer* sequences that codify mutations (bold) for different murine p22HBP variants.

For Megaprimer construction, plasmid DNA template pNJ2 (20 ng), 7.5 µL of each 5' primer [20 µM] (table 3); 2,5 µL of each 3' primer [20 µM] (table 3); 2.5 U of Vent Polimerase, 10 µL Vent reaction buffer 10x [200 mM Tris-HCl, 100 mM $(NH_4)_2SO_4$, 100 mM KCl, 2 mM $MgSO_4$, 1 % Triton X-100]; 16 µL dNTPs 1,25 mM, were mixed in different PCR tubes, in a final volume of 100 µL and subjected to PCR (MJ mini gradient thermal cycler from BioRad) as follows: initial denaturation at 94°C for 4 minutes; 28 cycles at 94°C for 40 seconds, 55°C for 1 minute and 72ºC for 1 minute, followed by final extension at 72°C for 5 minutes [63]. The megaprimers were then purified by a QIAquick® PCR purification kit [66], analyzed and quantified by agarose gel 1%.

Variant	***Primer* 5'**	***Primer* 3'**
K177A	HBP-7	PBRevo2
K64A	HBP-8	HBP-R2
K177E	HBP-9	PBRevo2
R56E	HBP-10	HBP-R2
K64E	HBP-11	HBP-R2
R56A/K64A	HBP-12	HBP-R2

Table 2.2. Murine p22HBP variants, its mutations and primers used in PCR. HBP-R2 5'-TTATAAGGATCCTCATGCCTTCACAAGCCAGACCTCGT-3'; pBRevo2 -5' TACGAG TTGCATGATAA -3'.

In a second round of PCR, plasmid DNA template pNJ2 (20 ng), 20 µL of each megaprimer (50 ng/µL) generated in the first round of PCR, 2.5 U of Vent Polymerase, 10 µL Vent reaction buffer 10x, 6 µL $MgSO_4$ 10x, 2.5 µL dNTPs 20 mM were mixed, in different PCR tubes, in a final volume of 100 mL and subjected to PCR under the same conditions as the first round. The samples were treated with *Dpn I* at 37ºC for 45 minutes. This restriction enzyme is specific for dam- methylated DNA fragments and allows the elimination of the original, methylated, plasmid in contrast to the un-methylated plasmid generated by Whole Plasmid PCR containing the desired mutation. *Dpn I* treated DNA samples were purified as previously by QIAquick® PCR purification kit [66] and in order to evaluate *Dpn I* digest efficacy, samples were analyzed on a 1% agarose gel, using λ-DNA-BstE *digest* from *New England Biolabs* as a marker.

E.coli DH5α competent cells were then transformed with the plasmids obtained after *DpnI* treatment. The transformation was obtained by electroporation: 1 µL of each sample (plasmid with desired mutation) was mixed in 40 µL of competent cells, transferred to electroporation cuvettes and submitted to a pulse of 1.8 Volts, 3 seconds. After electroporation, cells were immediately ressuspended in 1 mL of LB media, enriched by 20 % of sterile glucose, and incubated at 37ºC, 150 rpm. After recovering the cells, 300 µL of each culture were spread on LB plates containing by Ampicilin. These plates were incubated for 12 hours, and the presence of colonies indicated successful transformation.

Single colonies were picked from plates, and used to inoculate LB media for further growth at 37ºC, overnight. Plasmid DNA was purified with GeneJET® Plasmid Miniprep kit of Fermentas [69]. The DNA preps were sent for sequencing at the University of Florida, USA. HBP-7 and HBP-10 were sequenced using pBRSTOP as primer while the remaining templates were sequenced using pBRevo1. Sequencing results indicated that the desired mutations had been introduced into the original pNJ2 plasmid (data not shown). Finally a triple R56A/K64A/K177A variant was constructed using the same whole plasmid PCR method by Jerome Clayton and Professor Glória Ferreira (College of Medicine, University of South Florida).

Murine p22HBP variants were overexpressed in the same hosts cells as murine p22HBP wild type with identical overexpression and purification protocols as described in section 2.2. The fractions obtained from Ni-NTA column were analyzed by SDS-PAGE, and Figure 2.5 represents a SDS-PAGE image of the R56A/K64A variant. On the gel, the target protein band appears in a position expected for murine p22HBP (around 22 kDa). The gel also shows that the imidazole gradient used (10-500 mM) with the purest fractions collected at 75 and 50 mM imidazole. These fractions were then gathered, concentrated and desalted. Figure 2.5 also shows the UV visible spectra of R56A/K64A, where an A_{280} of 0.267 (considering the dilution factor, 1:25) and a ε of 31574 cm^{-1} M^{-1}, gave a protein concentration of 0.21 mM. In Table 2.3, protein yields obtained after overexpression and purification of all murine p22HBP variants are shown.

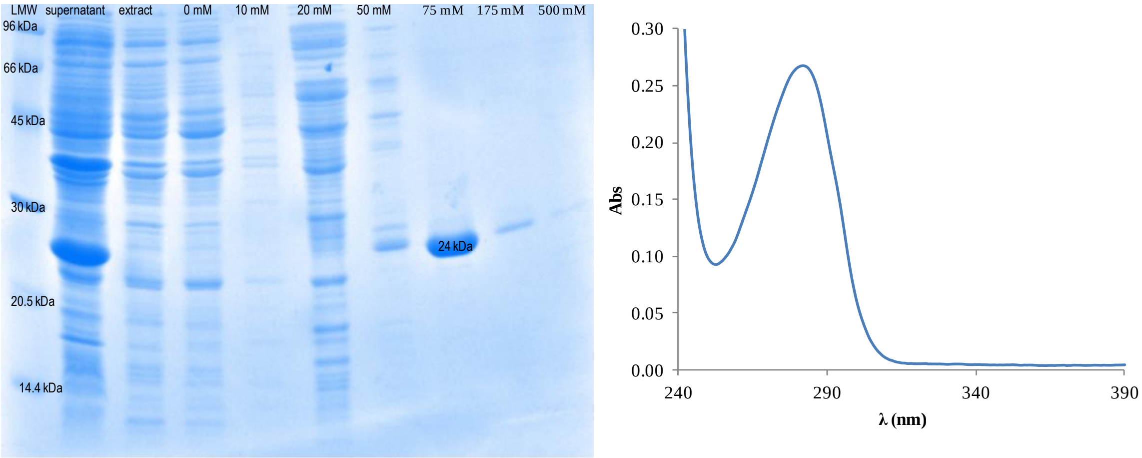


Figure 2.5. Left: SDS-Page analysis of the different fractions obtained from Ni-NTA Agarose column in murine p22HBP-R56A/K64A purification. Right: UV-Visible spectra of the concentrated fractions of p22HBP variant 0.21 mM.

Murine p22HBP variants	**Concentration [mM]**	**Yield (mg protein/L culture)**
R181A	0.20	16
K177E	0.18	14
K64A	0.47	36
R56A/K64A/K177A	0.15	12
K64E	0.22	18
R56A/K64A	0.21	19
R56E	0.19	17

Table 2.3. Protein yields obtained in overexpression and purification of murine p22HBP variants. The final volume obtained for each variant was 3.5 mL.

2.4 Human p22HBP

The gene encoding human p22HBP, with a His-tag located at N-terminal, was introduced into pET28a expression vector, flanked by NcoI and XhoI sites, after the T7 promoter and the His-tag encoding sequence (Figure 2.7). This plasmid was used to transform *E.coli* BL21 (DE3) competent cells, a step carried out by Nzytech genes & enzymes Ltd.

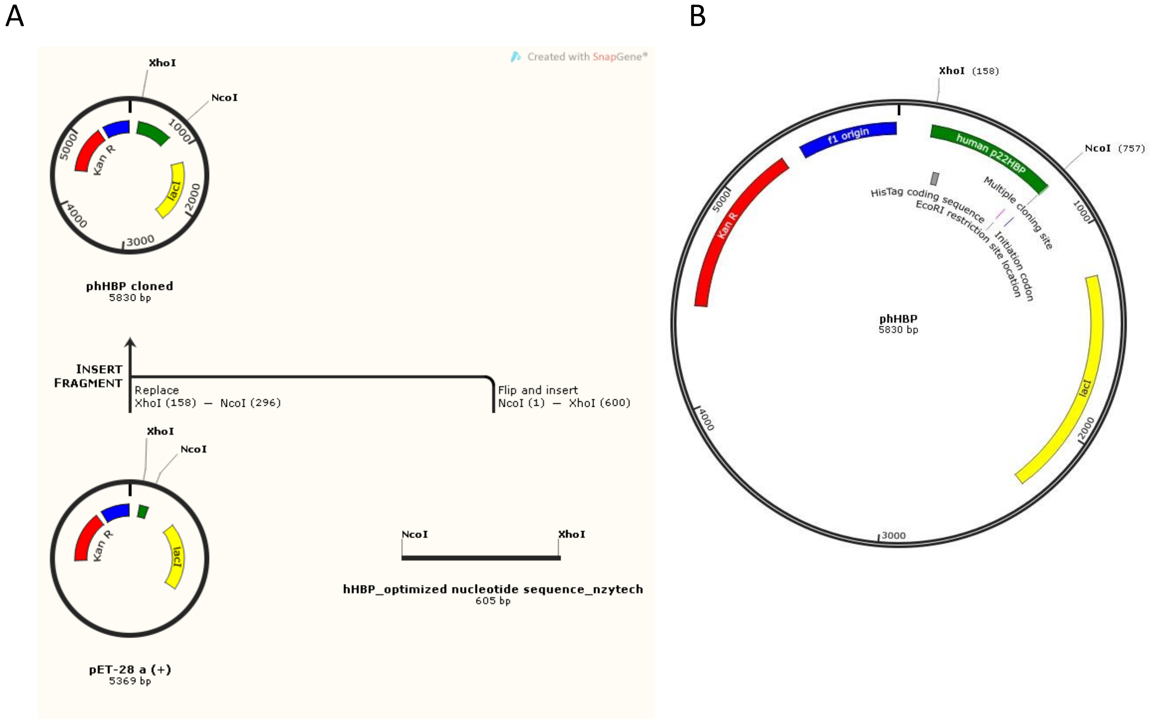


Figure 2.6. A) Human p22HBP cloning procedure. B) The pet28-a plasmid map for human p22HBP with then encoding sequence flanked by XhoI and NcoI restriction sites.

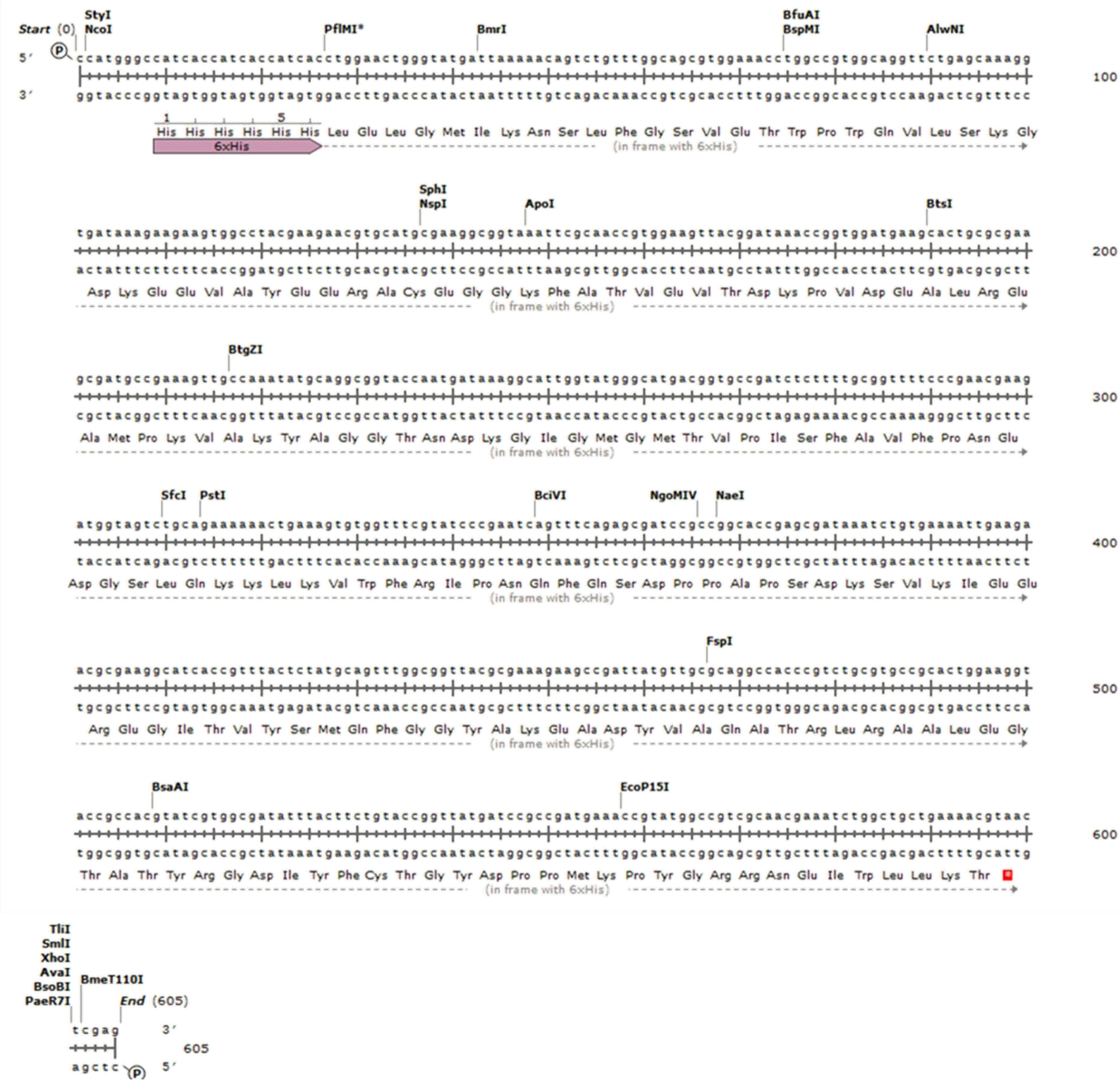


Figure 2.7. Optimized gene encoding sequence for human p22HBP overexpression in *E.coli* BL21 strains.

Glycerol stocks were prepared in order to store these *E.coli* strains. After an overnight culture in LB enriched with Kanamycin 50 mg/mL, 375 µL of growth media were collected and mixed with 125 µL 80% glycerol (previously autoclaved). Different glycerol stocks were prepared and stored at -80ºC.

Human p22HBP overexpression differs from murine p22HBP in antibiotic resistance and induction media. Human p22HBP is overexpressed in M9 media enriched with kanamycin 50 mg/mL and not in MOPS media enriched with ampicilin 50 mg/mL as described for murine p22HBP.

Starter cultures were prepared by inoculating 20 μL of cells taken from glycerol stocks in 20 mL of LB media previously autoclaved and supplemented with kanamycin 50 mg/mL, and incubated at 37ºC with shaking at 180 rpm, overnight. The culture was then inoculated into 1 L of LB medium and incubated for 5 hours at 37 ºC, 180 rpm. The cells were then harvested by centrifugation at 8000 rpm for 10 min at 4ºC and resuspended into an autoclaved M9 medium (1 L H_2O enriched with M9 salts (see receipt in appendix 9.1): 6.4 g $Na_2HPO_4.7H_2O$, 1.5 g KH_2PO_4 and 0.25 g NaCl) supplemented with 1 mL kanamycin 50 mg/mL, 0.5 mL $MgSO_4$ 1M, 0.5 mL $CaCl_2$ 0.1M, 0.25ml Thiamine-HCl 0.2%, 0.5ml $FeSO_4$ 0.1M, 1 g NH_4Cl and 4 g glucose. For isotopic labelling, NH_4Cl and glucose were replaced by $^{15}N-NH_4Cl$ and ^{13}C-glucose.

Resuspended cells in M9 medium were left at 30ºC, 150 rpm, 2 hours for adaptation to the new medium. After 2 hours, overexpression was induced with 1 mL of IPTG 0.5M and left overnight at 30ºC, 150 rpm. Cell lysis and purification was similar to that described for murine p22HBP in section 2.2. The experimental value of molar absortivity for human p22HBP was found to be 33205 $M^{-1}cm^{-1}$.

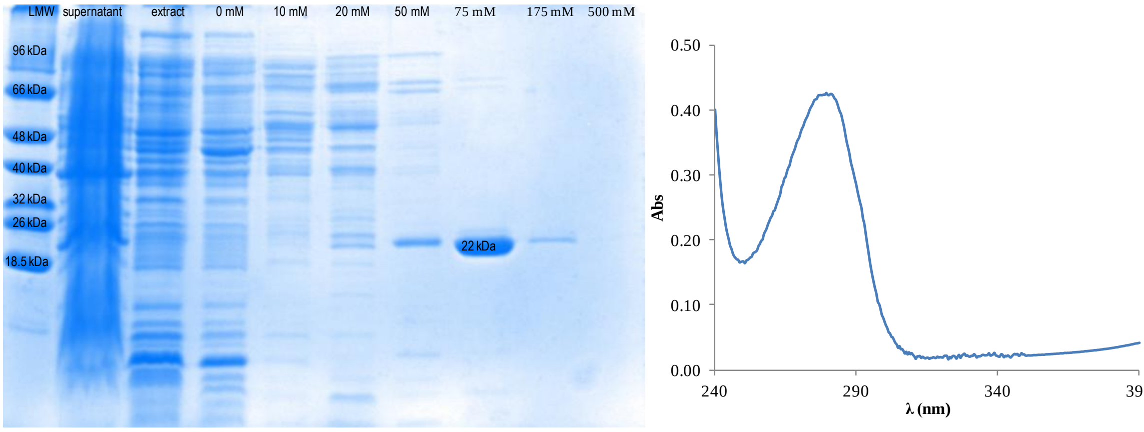


Figure 2.8. Left: SDS-Page analysis of the different fractions obtained from Ni-NTA Agarose column in human p22HBP purification. Right: UV-Visible spectra of concentrated fractions of human p22HBP 0.25 mM.

As shown in Figure 2.8, pure human p22HBP was obtained from the 75 mM imidazole fraction. Considering the protein concentration, an absorbance of 0.422 was obtained at 280 nm for a 1:20 diluted sample, which corresponds to a final concentration of 0.25 mM. A final volume of 3.5 mL of human p22HBP was obtained which corresponds to a yield of 20 mg per Litre.

3 p22HBP-Heme binding studies by Fluorescence Quenching

3.1 Introduction

Fluorescence spectroscopy is a well developed technique and is very useful when applied to biological systems and in particular it is used to study protein-ligand interactions [41]. Many biological molecules display fluorescence, particularly those that contain aromatic systems, such as reduced nicotinamide dinucleotide (NADH), oxidized flavins, chlorophyll and proteins [41]. Almost all proteins have natural fluorophores, tyrosine and tryptophan, which allow changes in protein conformation to be studied, and in case of absence of these residues, site-specific labelling with external fluorophores can be made by mutagenesis and chemical modifications [42], [43]. Combined with biophysical analysis of structure, these methods can reveal the complete and complex nature of protein ligand structure and dynamics [44]. Despite the recent developments in fluorescence applications, as well as the continuous improvement of instrumentation and technology, the principles on which is based this phenomenon remain the same [41].

Fluorescence is a special case of photoluminescence, a phenomenon in which light emission occurs from the excited electronic states of a molecule or an atom. Electronic states can be classified into two categories, singlet states and triplet states. When electrons in a molecule have their spins paired, the electronic state is known as singlet state. Triplet states are those in which electrons are unpaired (Figure 3.1). Depending on the nature of the excited state, photoluminescence can be classified as fluorescence or phosphorescence.

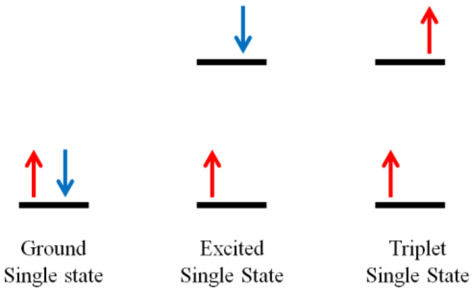


Figure 3.1. Possible electronic states of 2 electrons in a molecule or atom.

The partial energy diagram for a photoluminescence system, the Jablonski diagram, illustrates these electronic states and their respective electronic transitions (Figure 3.2).

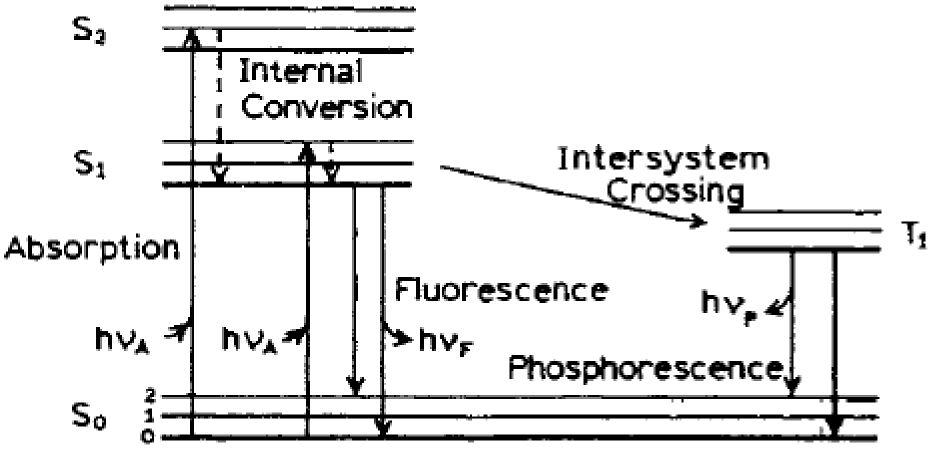


Figure 3.2. Jablonski diagram [45].

Each of the ground or excited electronic states has different vibrational levels corresponding to possible changes in vibrational modes. The energy of a photon required to generate a particular excited state is the difference in energy between the excited state and the ground state, represented by:

$$E = \frac{hc}{\lambda}$$

where λ is the wavelength of the light, h is Planck's constant (6.626×10^{-34} Js) and c is the speed of light (2.998×10^{8} m/s). Once a molecule absorbs energy and is excited into a singlet state, it can return immediately to the ground state, spin allowed, by emission of energy: this is fluorescence. Relaxation rates are typically 10^{8} s^{-1}, with a fluorescence mean time of 10^{9} s^{-1}.

Phosphorescence, another possible way to return to the ground state, corresponds to energy emission from a triplet state (see Figure 3.2). The transitions from triplet states occur more slowly (10^{-3} to 10^{2} s^{-1}) with a mean relaxation time milliseconds to seconds.[46] Triplet states populated by direct absorption from the ground state are insignificant and the most efficient way to populate triplet states is by intersystem crossing. This process is a spin-dependent internal conversion.

A special case of Fluorescence occurs when there is an intensity decrease in fluorescence emission [41]. This fluorescence quenching can occur by mechanisms such as molecular collisions or during molecular contact (static quenching). In both cases molecular contact between the fluorophore and quencher is required. In static quenching, the quencher makes a stable ground-state complex with the fluorophore and prevents the fluorophore from entering the excited state. In dynamic (collisional) quenching, the quencher transiently contacts the excited state fluorophore and provides a route for the excited state fluorophore to lose energy without emitting a photon [44].

Fluorescence quenching experiments are normally carried out by measuring the fluorescence of a dye in the presence of increasing concentrations of quencher. In the case of ligand-protein interactions, the intrinsic fluorescence of the protein is used and measured as either the ligand (quencher) concentration is varied [42].

3.2 Principles of receptor binding experiments

In order to understand the results obtained from FQ studies a basic understanding of the theory receptor-ligand interactions is needed in order to choose the best model to fit the experimental results.

There are three main types of receptor binding experiments [47]:

- kinetic experiments, where the binding of one or more concentrations of a ligand to a receptor (fixed concentration) is measured with an incrementing series of time points. These data are used to estimate association (K_{on}) and dissociation (K_{off}) rate constants.
- saturation experiments, where binding of different concentrations of a ligand with a receptor (fixed concentration) is measured at equilibrium and analysed to determine the binding constant (affinity constant, K, or dissociation constant K_d).
- competition/modulation experiments, where the binding of one or more fixed concentrations of a ligand with a receptor (fixed concentration) is measured at equilibrium in the presence of an increasing concentration of a competing ligand.

3.3 Equilibrium binding model used to model FQ data

Considering a simple two state equilibrium-binding model:

$$P + L \underset{K_{off}}{\overset{K_{on}}{\rightleftarrows}} PL$$

Equation 3.1

The equilibrium affinity constant, K, with units of Molarity^{-1} is defined as

$$K = \frac{[PL]}{[P][L]} = \frac{K_{on}}{K_{off}}$$

Equation 3.2

As K increases so does the concentration of the protein-ligand complex, with a consequent reduction in the free species. The decrease in free ligand concentration as a consequence of protein binding is called ligand depletion.

Alternatively K_d, the dissociation constant, is defined as:

$$K_d = \frac{[P][L]}{[PL]} = \frac{K_{off}}{K_{on}}$$

Equation 3.3

The dissociation constant defines the tendency of the protein-ligand complex to dissociate. Swillens in 1995 [48] defined a binding model that accounts for ligand depletion at high receptor concentrations. He considered that, when the receptor concentration is too high and the added ligand concentration is in the bound form, the typical equilibrium binding experiments cannot be described using a standard binding model. In the Swillens model the ligand binds to a single receptor site, although nonspecific binding is also catered for. Michaelis Menten Kinetics where K_d and R_{total} (total receptor concentration) are characterized.

Considering Equation 3.3, and assuming that, at equilibrium, the total receptor concentration [R$_{tot}$] is defined as,

$$R_{tot} = B^*_{tot} + L^*$$

Equation 3.4

and the total bound ligand concentration, B^*_{tot}, can be defined as a function of free ligand concentration, L^*:

$$[B^*_{tot}] = \frac{1}{K_d}([L] - [B^*_{tot}])[R_{tot}] \Leftrightarrow$$

$$[B^*_{tot}] = \frac{[R_{tot}]}{K_d}([L] - [B^*_{tot}]) \Leftrightarrow$$

$$K_d[B^*_{tot}] = [R_{tot}][L] - [R_{tot}]\,[B^*_{tot}] \Leftrightarrow$$

$$K_d[B^*_{tot}] = [R_{tot}][L] - [R_{tot}]\,([L] - [L^*]) \Leftrightarrow$$

$$K_d[B^*_{tot}] = [R_{tot}][L^*] \Leftrightarrow$$

$$K_d[B^*_{tot}] = [L^*]([R_{tot}] - [B^*_{tot}]) \Leftrightarrow$$

$$K_d[B^*_{tot}] + [L^*][B^*_{tot}] = [L^*][R_{tot}] \Leftrightarrow$$

$$[B^*_{tot}] = \frac{[L^*][R_{tot}]}{K_d + [L^*]}$$

Equation 3.5

Nonspecific binding, characterized by α, is defined as the ratio of nonspecifically bound ligand to free ligand, and depends on ligand concentration. This linearity was defined by Swillens [48] as:

$$[B^*_{tot}] = \frac{[L^*][R_{tot}]}{K_d + [L^*]} + \alpha L^*$$

Equation 3.6

$$b^*_{tot} = \frac{r_{tot}L^*}{K_d + L^*} + \alpha\beta L^*$$

Equation 3.7

At equilibrium, it follows that a quadratic equation (Equation 3.8) governs the relationship between the concentration of the receptor-ligand complex, the total ligand concentration, the total receptor concentration and K$_d$ [48].

$$b_{tot}^{*2}(1+\alpha) - b_{tot}^{*}[\beta(K_d + L_{tot}^{*})(1+\alpha) + r_{tot} + \alpha\beta L_{tot}^{*}] + \beta L_{tot}^{*}[\alpha\beta(K_d + L_{tot}^{*}) + r_{tot}] = 0$$

Equation 3.8

In the case of fluorescence quenching data, dissociation constants can be obtained accounting for ligand depletion, by nonlinear fitting of the emission maxima (y) as a function of tetrapyrrole concentration (x) (Equation 3.9)[15].

$$y = \frac{I_0 - (I_0 - I_{inf})}{2[hbp]}\left[(k_d + [hbp] + x) - \sqrt{(k_d + [hbp] + x)^2 - 4[hbp]x}\right]$$

Equation 3.9. Equation used for k_d values determination, by plotting the emission maxima, y, as a function of tetrapyrrole concentration, x, where I_0 and I_{inf} are emission intensities at 0 and saturating concentrations of tetrapyrrole, respectively, and [hbp] is the protein concentration [15].

3.4 Ligand-protein interactions revealed by intrinsic fluorescence quenching

Fluorescence quenching provides a means of probing the accessibility of aromatic residues to small molecules and thus can reveal information about the structural environment surrounding the small molecule. This technique involves the quantification of protein fluorescence in the presence of increasing amounts of quencher, followed by fitting of the data to quantify the interaction of the quencher with the protein.

Intrinsic fluorescence of proteins is a result of a contribution of three aromatic aminoacid residues (Tryptophan, Tyrosine and Phenylalanine). These aromatic residues can absorb radiation around 280nm and become excited from their ground state (S0) to an higher energy electronic state (S1) (see Figure 3.2). The energy is rapidly lost via vibrational energy to the surroundings as thermal energy. During the return of the system to the ground state emission of radiation of lower energy or non-radiative exchange such as quenching can occur.

However, the observed quantum yields (fluorescence efficiency) for tryptophan, tyrosine and phenylalanine of 0.2, 0.1 and 0.04 respectively coupled with their relative absorption coefficients, (Tryptophan 5540 M^{-1} cm^{-1}, Tyrosine 1480 M^{-1} at 280 nm and 195cm^{-1}/M at 257.5nm for Phenylalanine) make tryptophan mainly responsible for protein fluorescent emission. It should be noted that the emission maximum for Trp can shift between 330 nm in nonpolar environments to 360 nm in polar environments [44]. Phenylalanine does not absorb above 275 nm and its weak fluorescence is not normally observed and the fluorescence intensity of free tyrosine is approximately one-fifth of tryptophan and in proteins it is usually much weaker. This is due to a combination of interactions, such as the presence of neighbouring charged groups, hydrogen bonding to peptide carbonyl groups and non-radiative energy transfer to tryptophan or disulfide bonds, resulting in tyrosine fluorescence being diminished or quenched [46].

The intensity of fluorescence observed for both tyrosine and tryptophan depends on interactions with ligands that may quench the fluorescence emission from the native conformation. The degree of quenching varies from protein to protein and provides an indicator of solvent accessibility to tryptophan residues and the dynamics of protein conformation. The presence of specific extrinsic quenchers in physical contact with an excited protein can lead to sharing or transfer of the excitation energy, leading to reduction in the quantum yield and consequent decrease of fluorescence intensity. The yield of this quenching depends on physical access of the quencher to the protein [49]. Thus for a ligand-protein complex, when the ligand binds to its receptor protein, there is a change in the fluorescence of the protein if an aromatic group is near to the binding site. This makes fluorescence a useful tool to study the equilibrium and kinetics of a wide range of proteins and ligands, enzymes and substrates.

In conclusion, fluorescence quenching is an important tool to characterize protein binding to specific ligands and therefore fluorescence quenching was used to study the heme-p22HBP interaction, and to evaluate how binding changes with the introduction of point mutations near the p22HBP binding site.

3.4.1 Aromatic residues location

As mentioned previously, Tryptophan, Tyrosine and Phenylalanine are the main amino acids responsible for intrinsic fluorescence of proteins. In an attempt to understand which residues will be more important in intrinsic fluorescence decrease with heme interaction, it is crucial to locate them in protein structure.

Considering murine p22HBP, the aromatic residues present in the sequence are Tryptophan 16, 18, 101 and 186; Tyrosine 31, 66, 131, 138, 144, 162, 167, 168, 172 and 179; Phenylalanine 10, 41, 84, 87, 102, 108 and 135 (Figure 3.3, left). The aromatic residues present in human p22HBP are Tryptophan 16, 18, 101 and 185; Tyrosine 31, 65, 131, 138, 144, 161, 166, 171 and 178; and Phenylalanine 10, 41, 84, 87, 102, 108, 135 and 167 (Figure 3.3, right).

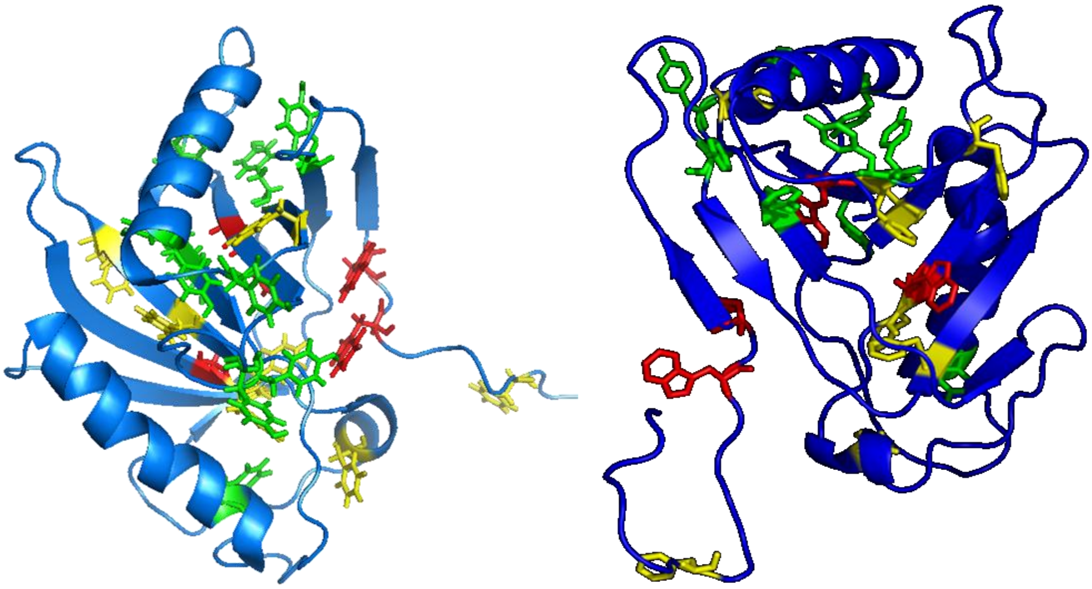


Figure 3.3. The location of the aromatic residues (Phe-yellow, Trp-red, Tyr-green) in murine (left: pdb 2GOV) and human p22HBP (right: modeller structure: see chapter 4).

3.5 Material and methods

3.5.1 Sample preparation

Protein and ligand solutions should be prepared in buffers with low absorbance (less than 0.1) in the regions of excitation and emission and fluorescence close to zero. When the protein is added to the buffer solution the observed fluorescence should arise exclusively from the protein. Therefore, the buffer solution used to dissolve the protein should be used as fluorescence blank.

The protein sample should not be too concentrated otherwise absorption will be too high. (Figure 3.4 a). Another parameter that should be considered is the purity of sample (Figure 3.4 b). Protein solutions must be freed from turbidity resulting from dust and aggregated protein. If the sample contains any fluorescent impurities, the fluorescence emission will be distorted by the impurity fluorescence emission. [41] These impurities can be removed by filtration or centrifugation. Many proteins can be filtered using a 0.22 µm filter fitted to a syringe. On the other hand, if the filters adsorb protein, the solution can be clarified by centrifuging at 4ºC to avoid protein denaturation.

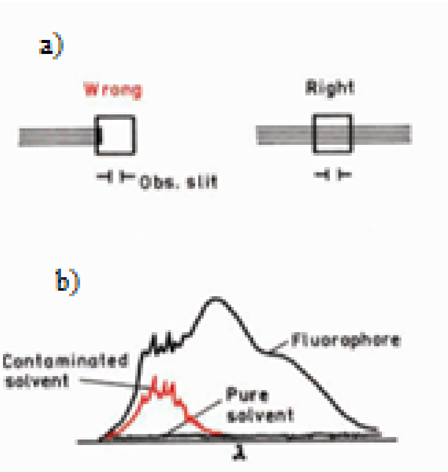


Figure 3.4. Common errors in sample preparation. a) Fluorophore concentration too high; b) Contaminated sample or cuvette. Adapted from [41].

3.5.2 Hemin solutions preparation

Hemin solutions used in fluorescence quenching studies were freshly made for each titration session. Approximately 1 mg of Hemin (Sigma-Aldrich) was weighed and dissolved in 50 µL of 25% ammonia followed by dilution in 1 mL of ddH$_2$O. 150 µL of

Tween 80 (1.5% v/v) was added and mixed vigorously followed by 7 mL of ddH_2O. The pH was adjusted to 8.0 with NaH_2PO_4. The total volume was adjusted to 10 mL with ddH_2O [15], giving a solution with final concentration of approximately 1.54 x 10^{-4} M. The final hemin concentration was estimated by measuring the UV-Visible absorption at 400 nm. The experimental value for the molar absortivity of hemin was 32257 $M^{-1}cm^{-1}$.

3.5.3 Fluorescence measurements

All fluorescence measurements were performed using a Fluoromax fluorescence spectrophotometer. The fluorescence cuvettes used were the standard 10 x 10 mm fluorescence quartz cuvettes (with all four faces polished) from Hellma. The protein samples used for fluorescence quenching measurements were prepared by dilution from a stock solution using a 50 mM phosphate buffer at pH 8.0. The protein concentration used in the titrations was estimated by UV spectroscopy (ε_{280} = 33920 $M^{-1}cm^{-1}$) to be 100 nM. Increasing volumes of Hemin [1.5 mM], were added (total final hemin volume 0.5 mL), to 2 mL of protein solution. After each addition, sample homogenization and equilibration was carried out (1 minute) followed by an emission scan from 300 to 400 nm with excitation at 295 nm. Each titration was repeated twice with new protein and porphyrin solutions, from the same stock, to avoid errors associated with sample preparation. Each titration gave an emission maximum as a function of tetrapyrrole concentration and these values were used to determine dissociation constants using Equation 3.9. Intensity values were corrected for dilution factors.

3.6 Results and discussion

Experimental results for intrinsic Fluorescence Quenching of murine p22HBP are shown in Table 3.1.

Exp. No	**Int$_{340nm}$ (x10^6)**	**V$_{Hemin}$ ad. (µL)**	**V$_{Hemin}$ ad. (µL)Total**	**dilution factor**	**Int$_{340nm}$ (x10^6)corr**	**n$_{adic}$ (x10^{-10}mol)**	**[Hemin] (µM)**	**[p22hbp] (nM)**
1	2.81	0	0	1.00	2.81	0	0.00	100.00
2	2.20	3	3	1.00	2.21	3	0.01	99.90
3	1.58	5	8	1.00	1.58	8	0.04	99.60
4	1.15	10	18	1.01	1.16	18	0.09	99.10
5	0.92	20	38	1.02	0.94	38	0.19	98.10
6	0.76	50	88	1.04	0.79	88	0.42	95.80
7	0.62	100	188	1.09	0.67	188	0.86	91.40
8	0.53	200	388	1.19	0.63	387	1.62	83.80
9	0.46	350	738	1.37	0.63	737	2.69	73.00
10	0.44	500	1240	1.62	0.72	1240	3.82	61.80

Table 3.1. Experimental fluorescence quenching data for the titration of murine p22HBP with hemin.

The decrease in maximum intensity emission values at 340 nm shown in Table 3.1, is a result of the intrinsic murine p22HBP fluorescence being quenched by increasing amounts of hemin interacting with the protein. Dilution factors were calculated by dividing the total solution volume present in fluorescence cell (murine p22HBP + Hemin) by the initial volume (2 mL of protein). The maximum intensity corresponds to the intrinsic fluorescence of the aromatic residues present in the protein. When hemin is titrated with p22HBP, protein-ligand interactions occur and as a result the fluorescence is quenched via a loss of excitation energy due to molecular collisions between the fluorophore (aromatic residues of murine p22HBP) and quencher (hemin). The experimental data were analyzed and p22HBP emission maxima as a function of Hemin concentration (Figure 3.5) was plotted. Using Equation 3.9, K_d values were determined by fitting the protein emission maxima as a function of hemin concentration. OriginPro software was used for fitting and independent parameters as protein concentration were used for fitting each curve.

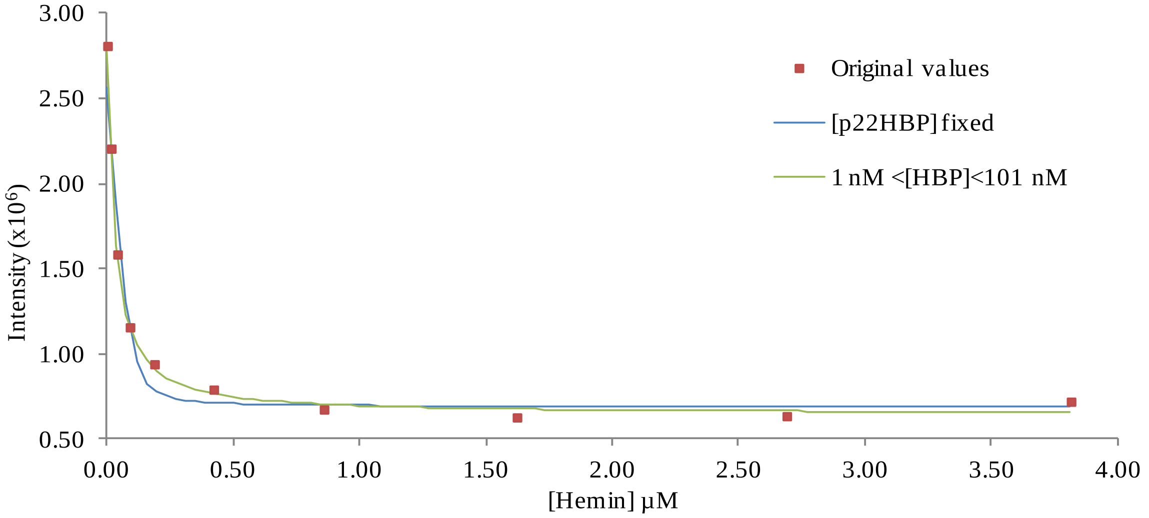


Figure 3.5. Intrinsic tryptophan fluorescence of murine p22HBP at maximum emission (340 nm) as a function of hemin concentration.

Figure 3.5 shows two different curves that resulted from fitting the same data points. The protein concentration was considered fixed [100 nM] (blue curve) or allowed to vary between 1 nM and 101 nM (green curve). The K_d values determined with these different approaches are shown in Table 3.2.

p22HBP	**approach**	**K_d (nM)**	**[p22HBP] (nM)**	**Reduced χ^2 (x10^{10})**	**Adjusted R^2**
murine p22HBP	[p22HBP] fixed	4.81 ± 5.42	100.00 ± 0.00	2.73	0.95
	[p22HBP] variable	23.70 ± 4.00	18.20 ± 8.43	0.18	1.00
human p22HBP	[p22HBP]fixed	59.14 ± 33.20	100.00 ± 0.00	1.60	0.97
	[p22HBP] variable	79.74 ± 10.54	101.00 ± 51.19	0.50	0.99
K64A	[p22HBP]fixed	8.47 ± 5.10	100.00 ± 0.00	0.86	0.98
	[p22HBP] variable	27.13 ± 4.54	31.18 ± 9.85	0.09	1.00
K177A	[p22HBP]fixed	8.29 ± 6.35	100.00 ± 0.00	1.43	0.96
	[p22HBP] variable	27.60 ± 7.38	23.70 ± 15.60	0.28	0.99
R56A/k64A	[p22HBP]fixed	-5.82 ± 9.46	100.00 ± 0.00	2.12	0.97
	[p22HBP] variable	46.82 ± 20.85	1.00 ± 33.08	0.41	0.99
R56E	[p22HBP]fixed	48.35 ± 18.31	100.00 ± 0.00	0.98	0.97
	[p22HBP] variable	76.84 ± 24.02	1.00 ± 43.37	0.34	0.99

p22HBP	**approach**	**K_d (nM)**	**[p22HBP] (nM)**	**Reduced χ^2 (x10^{10})**	**Adjusted R^2**
K64E	[p22HBP]fixed	17.47 ± 9.32	100.00 ± 0.00	2.20	0.97
	[p22HBP] variable	55.56 ± 11.42	1.00 ± 19.45	0.30	1.00
K177E	[p22HBP]fixed	22.75 ± 11.39	100.00 ± 0.00	2.52	0.96
	[p22HBP] variable	53.53 ± 11.60	1.00 ± 21.48	0.42	0.99
R56A/K64A/K177A	[p22HBP]fixed	34.55 ± 15.12	100.00 ± 0.00	1.95	0.96
	[p22HBP] variable	62.63 ± 16.24	1.00 ± 29.75	0.44	0.99

Table 3.2. Dissociation constants obtained by non-linear fitting of the emission maxima as a function of Hemin concentration for human p22HBP, murine p22HBP and respective variants. [p22HBP] variable corresponds to a range between1 and 101 nM.

As shown in Table 3.2, the best fit was obtained when the protein concentration was allowed to vary between 1 and 101 nM. This makes sense as, during the titration, the p22HBP concentration decreases with increasing amounts of Hemin. Moreover, it is a fundamental principle in biophysics that dissociation constants can only be measured precisely for protein concentrations of the order of K_d [50] [51].

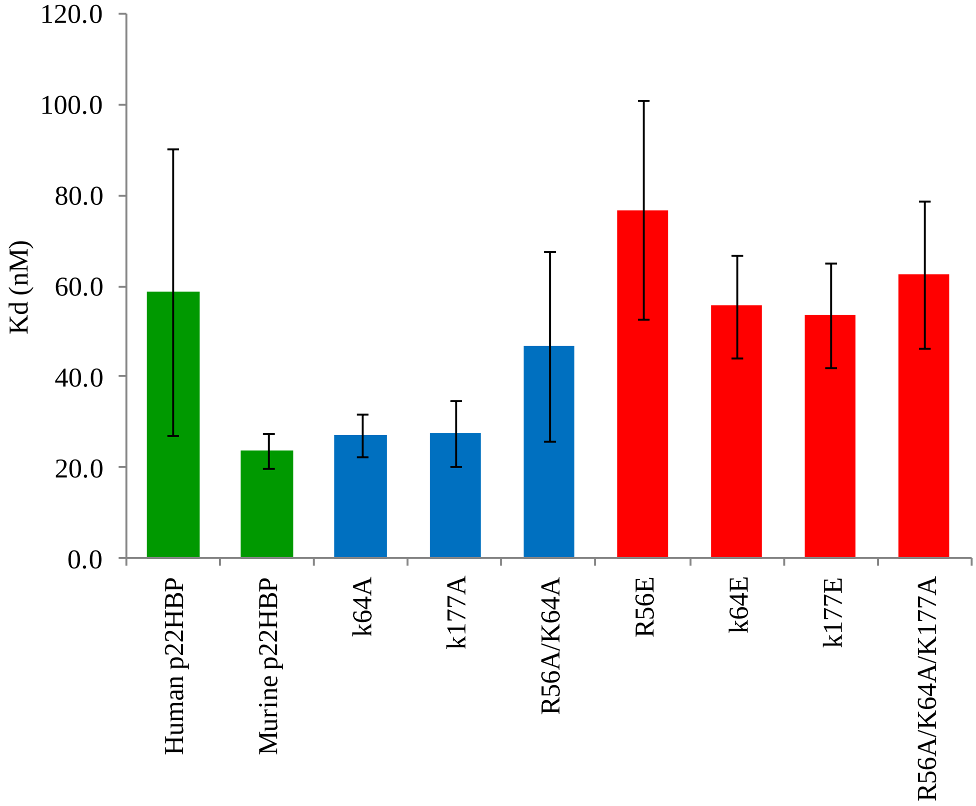


Figure 3.6. p22HBP dissociation constants obtained by non-linear fitting of the emission maxima as a function of Hemin concentration.

Considering first the murine and human wild type proteins; although human p22HBP has a much larger associated error, the K_d value is ca. 3 times larger than that for the murine protein. However, the values for the dissociation constants obtained for these Hemin complexes were of the same order of magnitude (nM) and are comparable to those from Dias *et al.* [40] , Delgado (2011) [52] and Freire (2012) [19] as the same technique was used for all of these studies. Dias *et al.* obtained K_d values of 3 nM for the hemin complexes. Dias *et a.l* reported Kd for murine p22HBP of 0.4 nM with PPIX and 3.0 nM with Hemin. Freire (2012) reported 2.6 nM for murine p22HBP Kd with PPIX and 11.1 nM with Hemin. In case of human p22HBP, Freire (2012) and Delgado (2011) [52] reported a Kd of 6.4 nM with PPIX and 20.4 nM with Hemin. These values can also be compared to those initially reported by Taketani *et al.* [1] although a different methodology was used for determination of K_d. However Blackmon *et al.* [14] determined K_d values with a discrepancy of 10^{-3} (µM versus nM). In this case, receptor concentrations were higher, when compared to concentrations used in this work (nM), which may have unintentionally influenced the resulting K_d values.

The difference between human and murine p22HBP binding found here may result from the differences in primary sequence (human-murine: E28D, A30S, V62I, A63M, I74V, I82V, S110G D111S, K118E, M133T, 147QH, R150Q, A153T, A154T, R162-Q163, I165-V166, F167-Y168, T169-A170, I184-V185, L187-V188, T189-A190) or the fact that human p22HBP has the complete N-terminus while the murine form starts at N7. In any case the difference is small and both forms of the protein show tight binding especially as measured by NMR where peaks are seen for the free and bound forms of the proteins in slow exchange (vide infra). In order to probe in more detail heme-p22HBP binding, and due to that fact the chemical shift mapping and ring current shift studies by Dias *et al.* could not distinguish the orientation of the tetrapyrrole ring when bound to p22HBP, a selected set of amino acids were chosen to be mutated based on molecular modelling studies (Micaelo *et al.*) [22]. If tetrapyrrole binding is stabilized by interactions between the propionates of the porphyrin ring and conserved positively charged residues located at the edge of the binding site: arginine 56, lysine 64 and lysine 177 (176 in human), the replacement of these positively charged residues by a negatively charged side chain such as glutamic acid should destabilize binding due to electrostatic repulsion. The same effect should also be seen by replacement with a neutral side chain such as alanine, although to a lesser extent. Figure 3.7 shows the position of the mutated residues in relation to the bound tetrapyrrole ring [22].

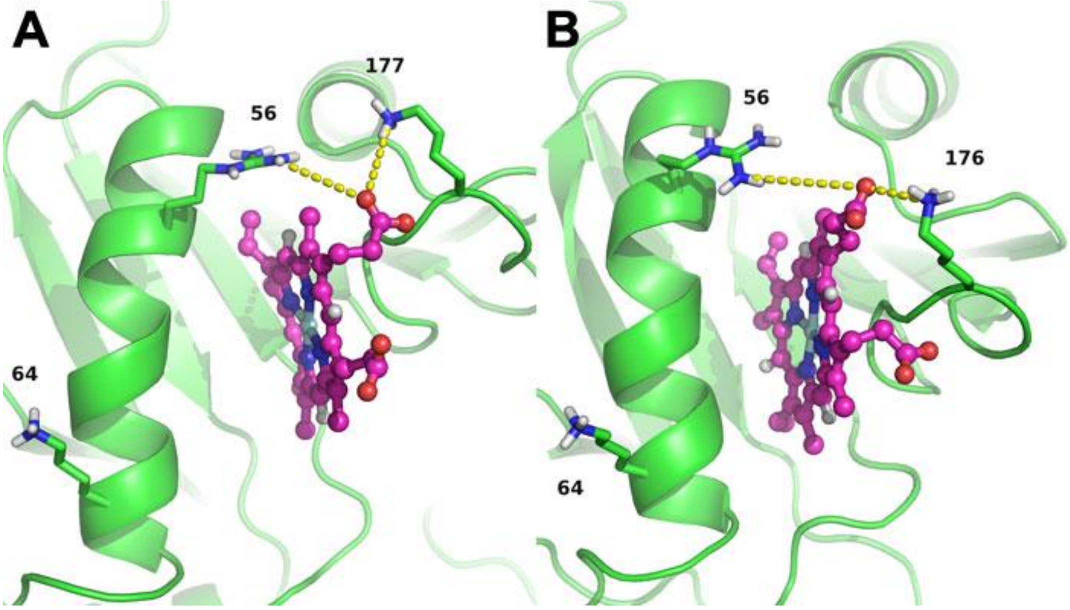


Figure 3.7: Representative structures of the hemin-murine p22HBP (A) and hemin-human p22HBP (B) complexes. The heme-binding site of each complex is shown with hemin rendered in ball and stick. The protein is rendered in cartoon. Key side chain residues are rendered in sticks. Reprinted from [22].

Considering first the R56E variant the K_d value is about 2.5 times larger than murine wt-p22HBP therefore a slight reduction in binding is occurring which indicates that electrostatic interactions may modulating heme binding but only to a small extent. The results for the K64E and K64A variants show that this residue does not have any role in stabilizing heme binding as the K64A variant has the same K_d value as wild type. The higher K_d for K64E (x 2) must result from slight repulsion between the propionates and the negatively charge glutamic acid. The results for K177A (x 1) and K177E (x 2) are very similar to the K64A and K64E variants indicating no role in stabilization for this residue either. The double R56A/K64A and triple R56A/K64A/K177A variants follow the pattern of the previous variants: a slight reduction in binding, confirming that heme binding does not involve electrostatic interactions to any great extent. In summary the fluorescence quenching results indicate that electrostatic interactions between the propionates of the tetrapyrrole ring and conserved charged residues in murine p22HBP are not important for heme binding and that an hydrophobic interaction with the p22HBP hydrophobic pocket identified in Dias *et al*. is the main driving force for binding.

4 Protein NMR spectroscopy of p22HBP

4.1 Introduction

Although NMR was discovered in 1946, its application to biological systems only started in the late 1960s and early 1970s. The application was very limited due to the poor sensitivity and very low resolution offered by the one-dimensional techniques used in those days[53]. Fourier transformation (FT) NMR that permitted rapid recording of NMR signals and 2D NMR spectroscopy that radically increased spectral resolution, in combination with the advance of stable magnets at higher fields led to rapid advances and in the mid 1980s several groups reported the first generation of solution structures of small proteins (< 10 kDa) using 2D NMR methods. The structure of the α-amylase inhibitor Tendamistat determined independently by NMR and crystallography confirmed the success of the NMR methods for structure calculations[54].

In the late 1980s and early 1990s, when multidimensional heteronuclear NMR methods, in conjunction with advances in molecular biological techniques, were developed the molecular size limit of NMR structures jumped to approximately 35 kDa. In June 2013, the number of structures available in the PDB archive determined using Nuclear Magnetic Resonance (NMR) spectroscopy has passed the 10,000 mark. Nowadays, NMR-derived structures account for more than 10% of the PDB archive.

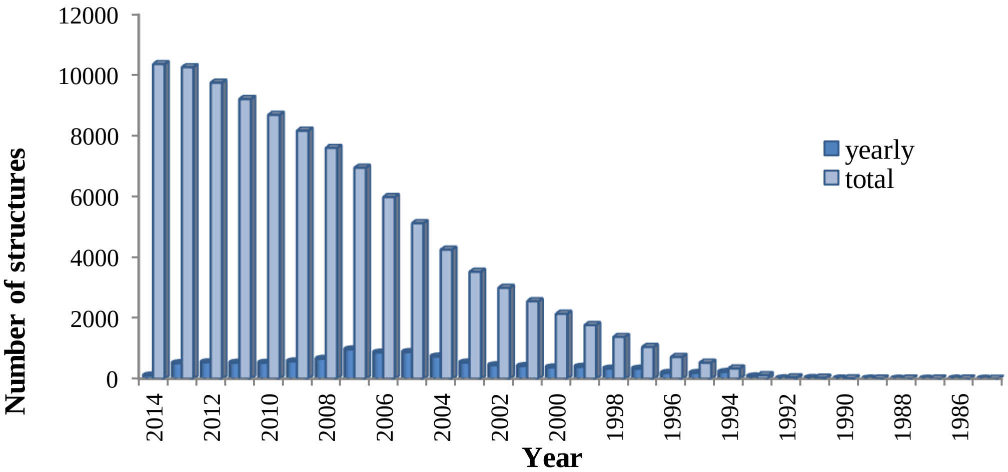
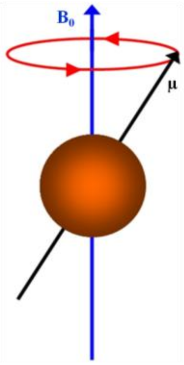


Figure 4.1. Yearly growth of released structures in the PDB solved by NMR. Adapted from Protein Data Bank (http://www.pdb.org/pdb/, March 2014)



Multidimensional Heteronuclear NMR of isotopically labelled proteins has opened the door to studying a wide variety of proteins and protein domains. The technique has been successfully applied in the field of structure determination and dynamic characterization of proteins [55]. Techniques based on NMR spectroscopy are also an important tool to observe and characterize the interactions between proteins and their ligands[56].

4.2 Basic principles

The basis of NMR spectroscopy is the property of an isotope of an element to have a nuclear spin which results in a nuclear magnetic moment [56]. Only nuclei with a non-zero nuclear spin quantum number **I** can be observed in an NMR experiment. Nuclei with an even number of both mass and charge have a spin quantum number of zero and are NMR inactive. By applying an external magnetic field non-degenerate energy states are produced by the interaction between the applied magnetic field $\mathbf{B_0}$ and nuclear angular moment **P**,

$$P = \hbar\sqrt{I(I+1)}$$

Equation 4.1

where $\hbar$ is Planck's constant divided by 2π and **I** the nuclear spin quantum number. The angular moment **P** can be characterized by the z component, **Pz** and is defined as

$$P_z = \hbar m$$

Equation 4.2

where the magnetic quantum number **m** has a total possible values of 2I+1 and defines the orientations of nuclear angular momentum. This definition quantifies in space the number of projections of nuclear angular momentum on the z axis. For example, magnetic nuclei with spin I=1/2 (e.g. ^{1}H, ^{13}C, ^{15}N, ^{19}F) have allowed **m** of 1/2 and -1/2. Thus two spin states are possible: one aligned with the z axis, the α state, and the other aligned against, the β state.

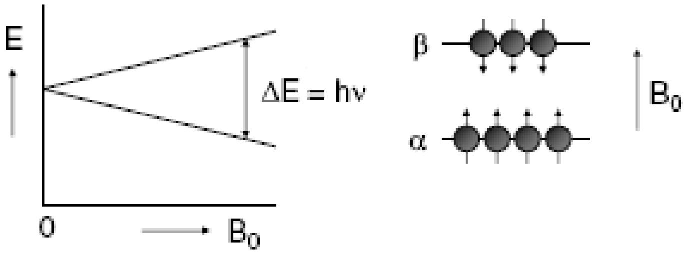


Figure 4.2. The different spin states of a nucleus in a magnetic field [57].

The magnetic moment, **μ**, or nuclear moment of a nucleus is defined as,

$$\mu = \gamma P = \gamma \hbar \sqrt{I(I+1)}$$

Equation 4.3

where γ is the nuclear gyromagnetic ratio, a characteristic constant for a specific nucleus. Thus, the angular moment P is the same for nuclei with the same magnetic quantum number and magnetic moment μ is different for each nuclei. The magnetic moment is used to characterize nuclear spins and is parallel to the angular moment if γ is positive or antiparallel if γ is negative. When nuclei are placed in an external magnetic field B_0, they will rotate about it due to the torque generated by the interaction of the nuclear angular moment P with the magnetic field. For each orientation state, also known as a Zeeman state or spin state, there is energy associated with this continuous rotation which is characterized by the frequency of the precession, the Larmor frequency, ω_0. The energy of a Zeeman state can be described in terms of Larmor frequency as

$$E = -\mu B_0 = -\mu_z B_0 = -m\hbar\gamma B_0 = m\hbar\omega_0$$

Equation 4.4

where B_0 is the external magnetic field strength in Tesla, and $\omega_0 = -\gamma B_0$.

The energy difference between of the allowed transitions (for instance between the quantized α and β states for a spin ½ nucleus) is given by

$$\Delta E = \hbar\gamma B_0$$

Equation 4.5

If ΔE is replaced by $\Delta E = \hbar\omega$, the frequency of the required electromagnetic radiation for the transition is defined by a linear dependence on the magnetic field strength:

$$\omega = \gamma B_0 \ (\mathrm{rad.s^{-1}})$$

Equation 4.6

$$\upsilon = \frac{\omega}{2\pi}\ (\mathrm{Hz})$$

Equation 4.7

The Larmor frequency for the ^{1}H nucleus (a proton) at specific field strength is innfact used to characterize the magnetic field of a spectrometer. The energy difference between two transition states becomes larger with the increasing strength of the magnetic field.

According to Boltzmann's equation, the ratio of the populations in α and β states is defined by

$$\frac{N_\beta}{N_\alpha} = e^{-\Delta E/kT} = e^{-\hbar\gamma B_0/kT} = \frac{1}{e^{\hbar\gamma B_0/kT}}$$

Equation 4.8

where $\boldsymbol{N_\alpha}$ and $\boldsymbol{N_\beta}$ are the population of the α and β states, respectively, **T** temperature and **k** is Boltzmann's constant. This equation indicates that a small fraction of spins will contribute to signal intensity at room temperature due to the small energy difference between α and β states which makes NMR spectroscopy a very insensitive spectroscopic technique. As ΔE is directly related to B_0, a stronger magnetic field will give better sensitivity, as the energy separation and therefore the population difference will increase.

An observable NMR signal result from an ensemble of nuclear spins, in the presence of the magnetic field, where α state nuclear magnetic moments are distributed randomly about a processional cone, parallel to the z-axis, and the β states randomly distributed in an antiparallel manner. The sum of the z components of the nuclear moments gives a net magnetization M_0 aligned along the z axis. The vector M_0 therefore results from the small population difference between the α and β states.

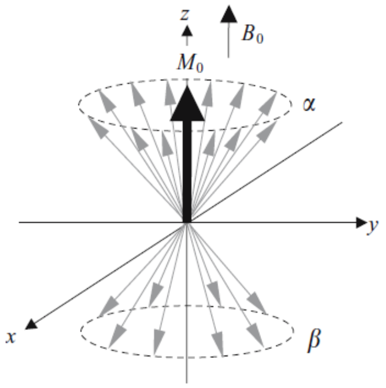


Figure 4.3. Two processional cones for a collection of 1/2 spin nuclei in the α- and β-states [57].

Once at thermal equilibrium, the spins have no phase coherence in the transverse plane and the net longitudinal magnetization is a static vector and the frequencies associated with the nuclear magnetization can only be observed by rotating the net longitudinal magnetization towards or into the transverse plane. This can be accomplished by subjecting the sample to a short pulse (few µs) of radiofrequency irradiation (RF) in MHz range with a magnetic component $\mathbf{B_1}$, to excite all frequencies of a given nucleus at the same time. The initial longitudinal magnetization experiences a torque from the applied $\mathbf{B_1}$ field and the $\mathbf{M_0}$ vector will rotate towards the transverse plane where it can be detected, as illustrated in Figure 4.4.

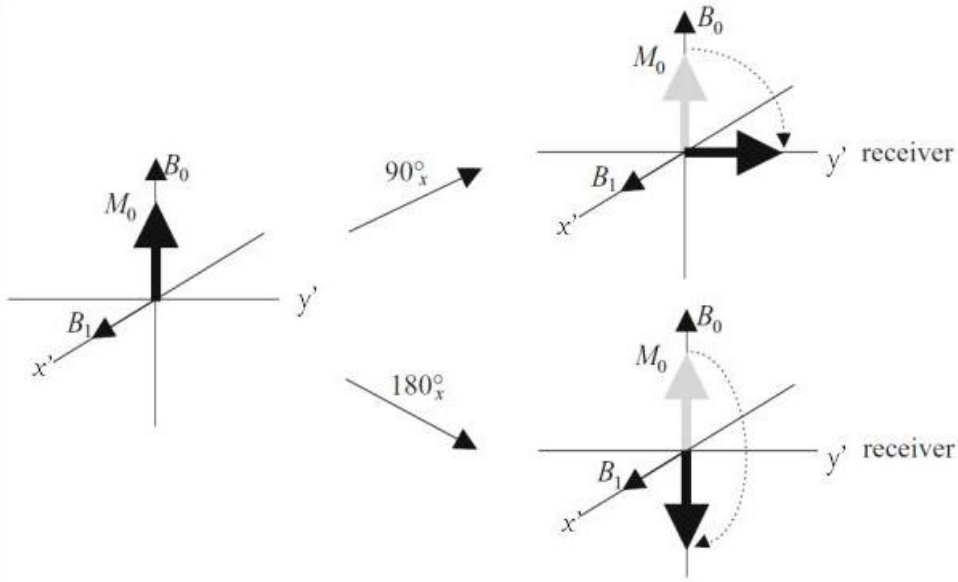


Figure 4.4. Representation of net magnetization M_0 under equilibrium conditions (left) and the effects on the M_0 of a 90º and 180º rf pulses [57].

The amplitude and duration of the pulse will define the angle θ through which magnetization vector M turns. When $\mathbf{B_1}$ field is applied long enough, $\mathbf{M_0}$ can be completely excited onto the transverse plane (called 90º pulse) reaching the maximum signal intensity, or even inverted to the -z axis (called 180º pulse) where no signal is detected since only magnetization in the x,y plane is able to induce a signal in the detection coil.

When the RF pulse is switched off, the system will immediately adjust to re-establish the Boltzmann distribution, and so the transverse magnetization will decay under the interaction of the static magnetic field $\mathbf{B_0}$ while precessing about the z axis and realign along the z axis. This return to equilibrium, referred to as relaxation, causes the NMR signal to decay with time, generating the observed Free Induction Decay (FID). In order to separate the individual signals and display them in terms of their frequencies, the FID (time domain) is converted to frequency spectra by applying a Fourier Transformation.

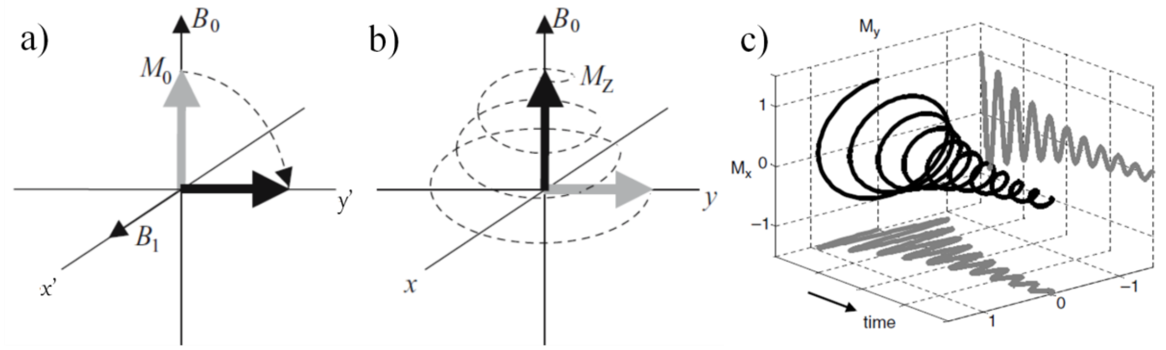


Figure 4.5. NMR experiment: a) after an Rf pulse M lies in the *x'-y'* plane and precesses about the z-axis (b) resulting in a time domain free induction decay (FID) (c) detected after the application of the *RF* pulse [57].

Relaxation is one of the most important phenomena in NMR and by measuring parameter related to relaxation the dynamics of the nuclei under study can be observed. The longitudinal relaxation time or spin lattice relaxation time (T_1) describes the rate at which the magnetization returns to the thermodynamic equilibrium along B_0, after an rf pulse. T_1 is correlated with the overall rotational tumbling of the molecule in solution and may be further affected by intramolecular mobility in flexible structures [58], [59].

The transverse relaxation time or spin-spin relaxation time (T_2) describes the decay of the effective magnetization observed in the x,y plane after a 90° pulse. T_2 is correlated with dynamic processes in the molecule under study; in particular it decreases with increasing molecular size, which presents a limiting factor for high resolution NMR with large proteins. This happens because NMR resonance linewidths in solution are inversely proportional to the T_2 relaxation time, which decreases with increasing molecular size and tumbling time. This line broadening, in addition to the increase in the number of resonances observed, due to the increase in molecular weight, causes chemical shift overlap and loss of spectral sensitivity, making spectral analysis and peak identification more difficult in large molecules [22], [58], [59]. Therefore there is a molecular size limit encountered in NMR of biological systems, large proteins above 500 amino acids cannot be studied by NMR [58].

4.3 Chemical shift

Different useful parameters can be extracted from NMR spectra, providing important information for molecular structure characterization. A key feature is the chemical shift which is sensitive to the local environment of a nucleus [60]. The phenomenon of chemical shift is caused by the shielding of nuclei from the external magnetic field by electrons in molecular orbitals. The effective magnetic field experienced by nucleus, B_{eff}, results from the contributions of B_0 and the local magnetic fields produced by the movements of surrounding electrons and is expressed as

$$B_{eff} = B_0(1 - \sigma)$$

Equation 4.9

where σ is the shielding constant, which reflects the extent to which the electron cloud around the nucleus shields it from external magnetic field. Thus, protons at the various sites in the molecule are magnetically shielded to different extents according with their chemical environment (type of chemical bond and neighbouring atoms). These slight changes in local magnetic field experienced by each nucleus will result in different frequencies in an NMR spectrum.

The shielding constant is influenced by several factors, such as the spherical electronic distribution of s orbital electrons. This type of shielding is known by diamagnetic shielding (σ_{dia}) and refers to the induced field with an opposite direction to the external magnetic field B_0. Electron orbitals other than s, induce a shielding effect from a nonspherical electronic distribution in which the induced local field has the same direction as B_0, known as paramagnetic shielding (σ_{para}).

$$\sigma = \sigma_{dia} + \sigma_{para}$$

Equation 4.10

σ_{dia} and σ_{para} have opposite contributions to the shielding constant: σ_{para} is proportional to $(m^2\Delta E)^{-1}$ where m represents the mass of the nucleus and ΔE the excitation energy to the lowest excited molecular orbital and asymmetry of electronic dispersion. On the other hand, σ_{dia} is proportional to m^{-1} and the symmetry of electronic distribution. These parameters will dictate chemical shifts range for protons and heteronuclei. Chemical shift range for different heteronuclei in proteins is shown in Table 4.1 . For protons, energy gap is large and consequently the σ_{para} is very small resulting in a small shift range, normally 10 ppm whereas in ^{13}C, ΔE is small and σ_{para} assumes an important contribution to the shielding. The bonding environment near the nuclei induces a distortion of the spherical electronic distribution which can significantly affect the respective chemical shift value of nuclei. Thus ^{13}C has a large range of possible chemical shifts (300 ppm) when compared with ^{1}H and this behaviour is usually observed for other heteronuclei. The local magnetic field produced, in opposite direction to B_0, by precession Larmor, makes paramagnetic contributions dominant when compared with diamagnetic.

Nucleus	**NH$_{backbone}$**	**NH$_{sidechain}$**	**CH$_{aromatic}$**	**C$_{\alpha}$H**	**C$_{O}$**	**C$_{\beta}$H**
^{1}H	8-10	6.5-8	6.5-8	3.5-5		1-4
^{13}C			110-140	40-65	170-185	20-75
^{15}N	110-140					

Table 4.1. Chemical shift range in proteins and peptides. ^{1}H and ^{13}C chemical shifts in parts per million are referenced to DSS and ^{15}N in parts per million is referenced to liquid NH_3.

Shielding constants can also be influenced by the ring current effect, an important contribution generated by delocalized electrons of p orbitals in an aromatic ring. When exposed to an external magnetic field, the π electrons circulating above and below the ring produce an additional magnetic field that opposes B_0 at the center of the aromatic ring and adds to B_0 at the edge of the ring. Thus, there is shielding at the center of the ring and protons directly attached to the ring are exposed to a field larger than B_0 due to the addition of the induced field, experiencing deshielding (Figure 4.6).

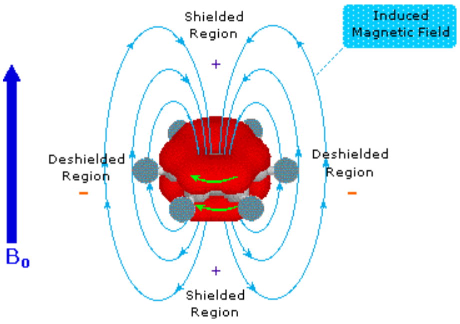


Figure 4.6. Diagram of an aromatic ring current. B_0 is the external magnetic field represented by blue arrow. The green arrow shows the direction of the ring current while light blue arrows represent the direction of the induced magnetic field. Extracted from http://www2.chemistry.msu.edu/faculty/reusch/VirtTxtJml/Spectrpy/nmr/nmr1.htm.

The position of a resonance signal in a NMR spectrum is measured by its resonance frequency although it is not expressed in units of Hertz, since this would make chemical shifts dependent on the magnetic field strength. To overcome this inconvenience the frequency scale of NMR spectrum is expressed in terms of ppm, normalized by using a signal from a reference compound, which for proton is usually tetramethylsilane (TMS) or 3-(trimethylsilyl)propionate (TSP), and is defined as:

$$\delta = \frac{\upsilon - \upsilon_{\mathrm{ref\,(Hz)}}}{\upsilon_{\mathrm{ref}}(\mathrm{Hz})} 10^{6}\ \mathrm{ppm}$$

Equation 4.11

where υ and υ_{ref} are, respectively, the positions, in Hz, observed for the signal of interest and for the reference compound. This dimensionless quantity is defined as chemical shift, **δ**, given in parts per million, and it is independent of the external magnetic field strength of spectrometer. Table 4.2 represents the average chemical shifts of active nucleus (H, C and N) that are present in the 20 naturally of amino acids.

Residue	**H_N**	**H_α**	**$^{13}C_O$**	**$^{13}C_\alpha$**	**$^{13}C_\beta$**	**^{15}N**
ALA	8.19	4.26	177.73	53.15	19	123.22
ARG	8.24	4.3	176.42	56.79	30.68	120.78
ASP	8.31	4.59	176.40	54.69	40.88	120.65
ASN	8.34	4.67	175.27	53.54	38.69	118.93
CYS	8.39	4.66	174.87	58.24	32.66	120.13
GLU	8.33	4.25	176.89	57.35	30	120.66
GLN	8.22	4.27	176.32	56.59	29.18	119.88
GLY	8.33	3.94	173.88	45.36	-	109.65
HIS	8.25	4.61	175.25	56.49	30.24	119.66
ILE	8.27	4.18	175.85	61.63	38.61	121.45
LEU	8.22	4.32	176.99	55.64	42.30	121.83
LYS	8.18	4.26	176.65	56.96	32.77	121.04
MET	8.26	4.41	176.2	56.12	32.99	120.09
PHE	8.36	4.63	175.43	58.11	39.95	120.47
PRO	-	4.4	176.73	63.34	31.85	133.96
SER	8.28	4.48	174.64	58.74	63.79	116.26
THR	8.24	4.46	174.57	62.23	69.72	115.41
TRP	8.29	4.68	176.13	57.68	29.98	121.67
TYR	8.32	4.63	175.4	58.13	39.32	120.52
VAL	8.29	4.18	175.63	62.51	32.72	121.12

Table 4.2. The statistics presented in this table were calculated from the full BMRB (Biological Magnetic Resonance Bank) database. This includes only the diamagnetic proteins. The calculated statistics are derived from a total of 4603403 chemical shifts in the 20 natural amino acids found in proteins [61] .

In principle, the structural and chemical environment of the atoms dictates chemical shifts of NMR-active nuclei such as ^{1}H, ^{15}N and ^{13}C. In proteins chemical shifts of signals from a protein are never used to determine which amino acids are present as this information is has to be known *a priori* when studying a protein by NMR. Apart from the dependence of the chemical shift on chemical structure, neighbouring amino acids, there is also a dependence on secondary structure described by Case *et al.* in the middle 1980s. The deviations of $^{13}C_{\alpha}$ (and to some extent, $^{13}C_{\beta}$) chemical shifts from their random coil values can be well correlated with the α-helix or β-sheet conformations: $^{13}C_{\alpha}$ chemical shifts larger than the random coil values tend to occur for helical residues whereas the opposite is observed for β-sheet residues. A good correlation is also observed for proton Ha shifts with secondary structures: $^{1}H_{\alpha}$ shifts smaller than the random coil values tend to occur for helical residues whereas the opposite is observed for β-sheet residues [53]. Chemical shift is also highly sensitive to the exact environment of the atom, and therefore yields information about whether a small molecule binds to a target protein, what parts of the small molecule are interacting and to which part of the macromolecular target the ligand is bound [62]. Chemical shifts for the free and bound states will, in general, be different because of changes in environment [63].

4.4 Spin coupling constants

J coupling constants are derived from the scalar interactions between atoms and they provide geometric information about torsion angles of the bonds between atoms in molecule (Figure 4.7). The most useful and coupling constants are vicinal scalar coupling constants, ^{3}J, between atoms separated from each other by three covalent bonds. Scalar couplings are used in multidimensional (2D, 3D, 4D) correlation experiments to transfer magnetization from one spin to another in order to identify spin systems, e.g. spins which are connected by up to three chemical bonds [53], [58].

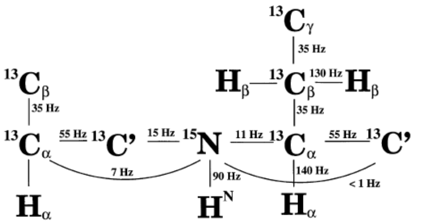


Figure 4.7. Spin system of the peptide backbone and size of the ^{1}J and ^{2}J coupling constants that are used for magnetization transfer in ^{13}C and ^{15}N-labelled proteins.[64]

4.5 High molecular weight protein NMR techniques

The one-bond coupling ^{1}H-^{15}N is the most important starting point for the NMR analysis of proteins. This bond is present in every amino acid residue in a protein, except for the N-terminus and proline residues (Figure 4.8). The experiments used to correlate bound ^{1}H and ^{15}N nuclei is called the ^{1}H-^{15}N HSQC (heteronuclear single quantum correlation). This HSQC experiment exploits the repeating nature of the protein's primary sequence and three-dimensional structure. As there is about one peak per residue, a ^{15}N-^{1}H HSQC spectrum is something of a NMR fingerprint of a protein and is usually the first heteronuclear experiment performed on a newly purified protein. HSQCs spectra are very commonly used to detect ligand binding - if the fingerprint changes (peaks move) it indicates that binding is occurring. Although natural abundance ^{1}H-^{15}N HSQC spectra can be acquired, isotopic labelling is normally required to obtain an HSQC spectrum.

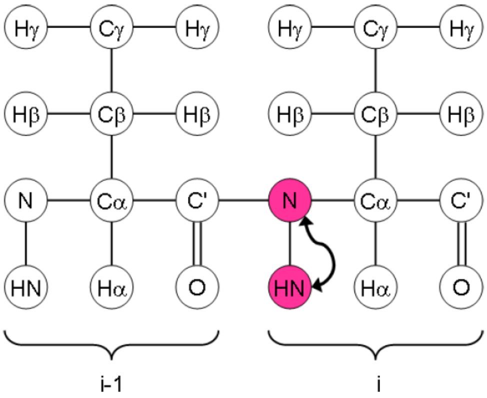


Figure 4.8. Protein backbone. Each aminoacid is connected via a peptide bond between the carbonyl carbon of first aminoacid and the nitrogen of the next aminoacid. The ^{1}H-^{15}N HSQC experiment detects protons directly coupled to nitrogen and the resulting spectra contains one peak for every aminoacid in the protein [30].

Magnetization is transferred from the proton to attached ^{15}N nuclei via the J-coupling. The chemical shift evolves on the nitrogen and the magnetization is then transferred back to the proton for detection. The H-N correlation seen include backbone amide groups, Trp side-chain Nε-Hε; groups and Asn/Gln side-chain Nδ-Hδ2/Nε-Hε2 groups. The Arg Nε-Hε peaks are in principle also visible, but because the Nε chemical shift is outside the region usually recorded, the peaks are folded/aliased (this essentially means that they appear as negative peaks and the Nε chemical shift has to be specially calculated) [65].

HSQC spectra degrade as the molecular weight of a protein increases: line widths increase and intensity decreases. For larger proteins the introduction of the TROSY (Transverse Relaxation Optimized Spectroscopy) experiment and higher magnetic fields (>600MHz) allowed a wide range of new applications of solution NMR, in particular in the emerging field of structural and functional genomics. The TROSY experiment can replace the HSQC experiment for large proteins. This technique has been developed to reduce relaxation losses during the chemical shift evolution of a heteronucleus X (e. g. ^{15}N), the $X \rightarrow {}^{1}H$ magnetization transfer and the acquisition time (Figure 4.9). Transverse relaxation is mainly caused by DD (dipole-dipole) coupling and CSA (chemical shift anisotropy). The DD interaction is independent of the static magnetic field, whereas the CSA increases with higher magnetic fields. The optimal TROSY effect, for an amide proton, is about 23.5 T, corresponding to a proton resonance frequency of 1000 MHz [66]. The technique is especially useful combined with deuteration of the protein.

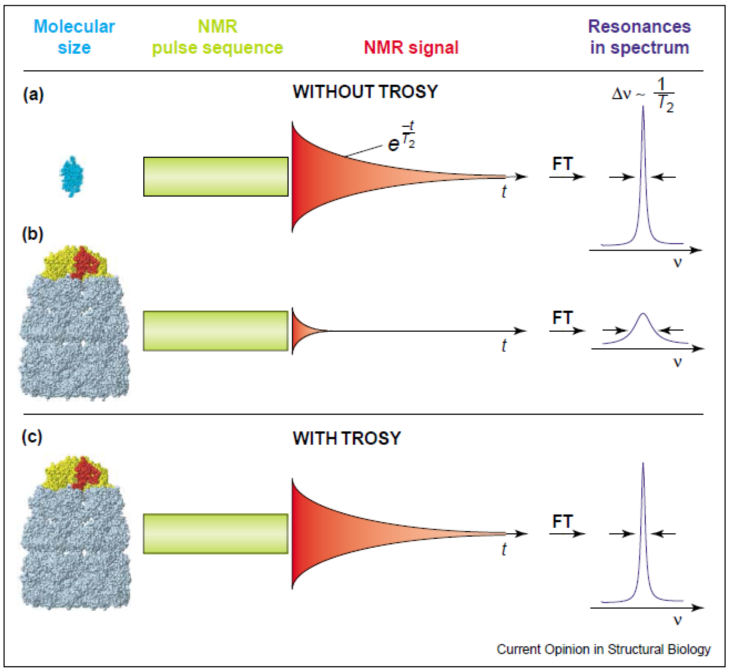


Figure 4.9. NMR spectroscopy with small and large molecules in solution. (a) The NMR signal obtained from small molecules in solution relaxes slowly; it has a long transverse relaxation time (T_2). A large T_2 value translates into narrow line widths in the NMR spectrum after Fourier transformation (FT) of the NMR signal. (b) For larger molecules, the decay of the NMR signal is faster resulting in a smaller T_2. (c) Using TROSY, the transverse relaxation can be considerably reduced, which results in improved spectral resolution and improved sensitivity for large molecules [67].

In detail, considering the components of a doublet of I in a weakly coupled two spin system IS the transverse relaxation rates are different due to addition or subtraction of the influence of DD coupling and the CSA. A narrow and a broad component is the result. In a non-decoupled two-dimensional ^{15}N-HSQC spectrum only one component of the quartet has a narrow line width in both dimensions (the TROSY component) as shown in below (Figure 4.10). TROSY experiments select solely these narrow components and suppress the broad components of the quartets. Note that the signals in a TROSY experiment are shifted in both dimensions by $^{½}J_{NH}$. To summarize, TROSY suppresses transverse relaxation, e. g. in ^{15}N-^{1}H moieties by constructive use of interference between DD coupling and CSA [68].

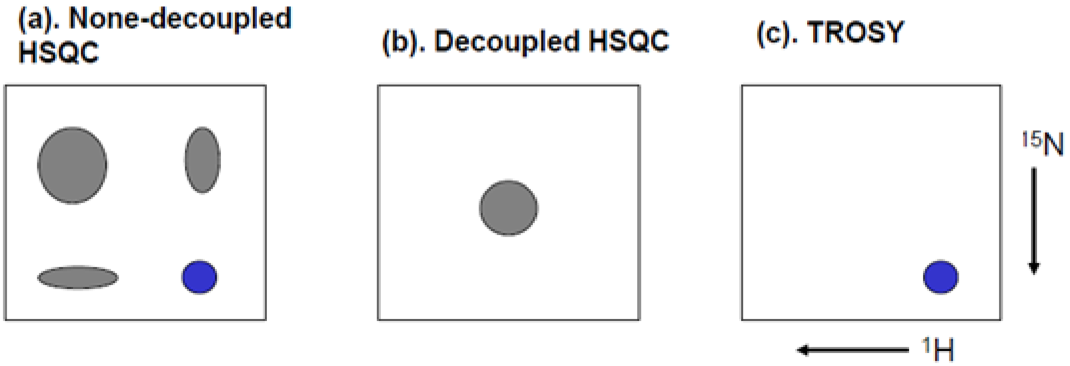


Figure 4.10. Region of ^{15}N-^{1}H correlation spectra. a) None-decoupled HSQC spectrum with different relaxation rates (line width) for each of the four components of ^{15}N-^{1}H correlation. b) Decoupled ^{15}N, ^{1}H HSQC spectrum; c) TROSY- selectively detect only the narrowest component (1 out of 4). Adapted from [69].

With TROSY spectra there is an intrinsic loss in sensitivity due to rejection of the broad components of the quartets (Figure 4.10). However, for large proteins (>20 kDa) at high magnetic fields, the detection of the most slowly relaxing peak compensates for the loss of sensitivity [69].

4.6 Sequential resonance assignment in unlabeled and ^{15}N labeled proteins

For the complete investigation of the structure and dynamics of proteins by NMR a complete resonance assignment is a prerequisite [59], [65]. Sequential assignment is a process by which a particular amino acid spin system identified in the spectrum is assigned to a particular residue in the amino acid sequence.

It is not always feasible to produce a ^{13}C, ^{15}N doubly labelled protein sample. Initially the sequential assignment method was developed by Wüthrich and co-workers, based on the identification of spin systems within individual amino acids using through-bond ^{1}H-^{1}H connectivity measured with correlation spectroscopy (COSY) and later on by total correlation spectroscopy (TOCSY) [70]. In a second step, the sequential connectivity between neighbouring amino acids is established via nuclear Overhauser effect spectroscopy (NOESY). In the two most common types of secondary structure (α-helix and β-sheet) the peptide chain brings ^{1}H of the peptide backbone and the side chains of neighbouring residues close together (< 5Å) so that they are observable by NOE spectroscopy. Small peptide segments of different lengths are thus obtained which are then matched to the primary sequence. In order to complete the resonance assignment these segments are extended and linked [71].

For larger proteins spectral overlap means this method cannot be used. In these cases other nuclei must be introduced and observed (indirectly). Here labelled proteins are required and the assignment is based on through bond correlations based on scalar coupled nuclei. Triple-resonance heteronuclear correlation experiments are used to observe nuclei coupled by efficient magnetization transfer through heteronuclear spins (^{1}H, ^{13}C and ^{15}N). [53] Nowadays many 3D triple resonance experiments are available whose names indicate the nuclei they correlate. Figure 4.11 shows how CBCA(CO)NH and HNCACB experiments are used to link neighboring residues. One ^{1}H-, ^{13}C-, and ^{15}N-heteronuclear three-dimensional NMR spectrum, which records the one bond coupling between $^{1}H_N$ and ^{15}N and the one and two bond coupling between ^{15}N and $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ in one residue. The spectrum is called a HNCACB. This type of experiment also records the coupling across $^{13}C_O$ to the $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ in the preceding residue. [72] The other experiment measures the heteronuclear coupling between $^{1}H_N$ and ^{15}N in one residue and the coupling across $^{13}C_O$ to the $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ in the preceding residue. This spectrum is called a CBCA(CO)NH spectrum. In a combined analysis of these two types of spectra it is possible from each individual $^{1}H_N$ - ^{15}N peak in the $^{1}H_N$ -^{15}N correlation spectrum, to identify the $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ in the same residue and the preceding residue. If the same $^{13}C_{\alpha}$ - $^{13}C_{\beta}$ pair, as shown in the open green frame of figure 13, are seen to couple to two different pairs of $^{1}H_N$ –^{15}N couplings as indicated by the black and red arrows in the two panels, they may be assigned as signals from neighboring residues. As C_{α} and C_{β} chemical shifts are indicative of amino acid type the linked residues can be compared to the primary sequence and an assignment can be made (Figure 4.11).

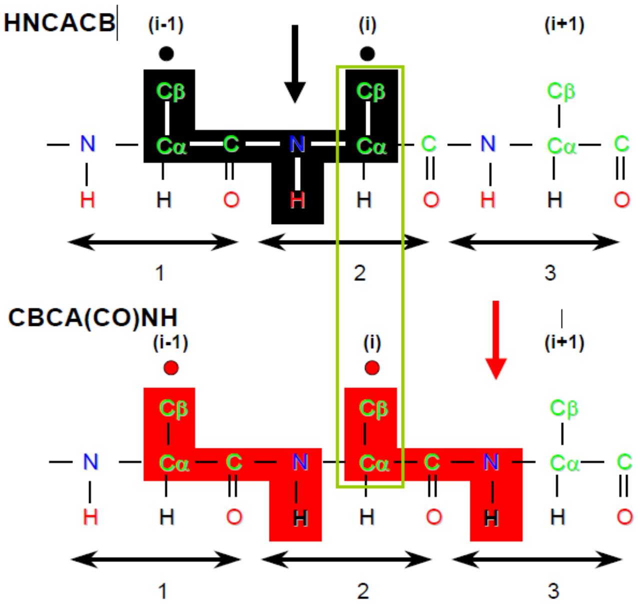


Figure 4.11. The two panels show the scalar coupling correlation, which is measured by the HNCACB (top) and by the CBCA(CO)NH (bottom). In the HNCACB (top) the coupling is mediated through the chemical bonds shown on a black background. The ^1H_N –^{15}N coupling pair of residue (i) is correlated to the 13Cα - 13Cβ pair of residue (i) and (i-1). In the CBCA(CO)NH (bottom) the coupling is mediated through the bonds shown on a red background. Here the ^1H_N – ^{15}N coupling pair of residue (i) is correlated to the 13Cα - 13Cβ in residue (i-1). [71]

To obtain a robust assignment strategy, however, other 3D experiments must be used, for example a set of additional 4 triple-resonance experiments HN(CA)CO, HNCO, HA(CACO)NH and HA(CA)NH) which deliver the CB, CA, C’ and HA frequencies of residues i and of previous residues i-1. HNCO and HN(CA)CO experiments are used to obtain the carbonyl ^{13}C chemical shifts of residue i and i-1. The experiments HA(CACO)NH and HA(CA)NH or HN(CA)HA deliver the HA frequencies of residue i and i-1. Most recent versions and a variety of other triple-resonance experiments are reviewed in the literature and other combinations of triple-resonance experiments are possible, even strategies based on 4D experiments [53], [71], [73].

The segments which have been assigned can be extended by finding matching spin patterns until the protein is completely assigned. Dynamics and exchange can lead to missing signals which prevent complete assignment. A correct backbone assignment is indicated by a full consistency of all data and represents a starting point for studies of the structure, dynamics and binding properties of proteins [74].

4.7 Chemical shift mapping to identify Ligand Binding

Small molecules in the presence of proteins tumble and diffuse much more rapidly when they are free in solution compared with when they are bound. By monitoring changes in NMR spectral properties, it is possible to evaluate how and where the binding between a ligand and a protein occurs [75]. Mapping binding sites in proteins can be achieved by two-dimensional NMR experiments such as HSQC and TROSY [63]. In practice, using ^{1}H NMR spectra can be difficult due to the difficulty of assigning chemically shifted resonances in the presence of a ligand due to overlap[62]. The first application of ^{1}H/^{15}N HSQC experiments to screen ligands for binding activity was demonstrated for the FK506-binding protein (FKBP) that inhibits calcineurim (a serine-threonine phosphatase) and blocks T cell activation when it is complexed to FK506 [76]. Valuable information on intermolecular interactions can be derived from chemical shift mapping. If sequence-specific resonance assignments for the receptor protein are available, they become even more useful [62]. When a ligand binds to a receptor protein, the chemical shifts of both ligand and protein proton resonance signals are affected [75], mainly nuclei located in the protein binding pocket, a reflection of different nuclear environments around the binding site [60].

4.8 Sample preparation

Protein sample preparation for NMR studies can be time-consuming as the protein under study has to be purified and isotopically labelled [77]. The development of a good expression system is normally the first step for protein overexpression and the following problems should be avoided: protein precipitation at high concentration, low stability and low expression levels. As NMR samples require large quantities of isotopically labelled proteins at milimolar concentrations, recombinant technology is widely used as it can provide high concentrations of proteins compared to extraction from natural sources. In addition, these recombinant expression systems can be controlled to produce protein domains or to attach tags for simple purification, which is particularly important for NMR studies that focus on the structures and dynamics of protein domains or domain-domain complexes. Currently, most isotopically labelled recombinant proteins are expressed in a bacterial host such as *E. coli* [78], and many are commercially available with different fusion partners and different features depending on the laboratory source. It is frequently difficult to predict the best expression vector for a particular protein due to the different behaviour of individual proteins in different expression systems [53], [59], [70].

For Heteronuclear multidimensional NMR experiments, the protein is overexpressed in a bacterial host, and labelled isotopically with ^{13}C, ^{15}N, or ^{2}H [49], [72]. Because the cost of ^{13}C, ^{15}N and ^{2}H source compounds is significantly higher than natural abundant sources, the isotopic labelling of the proteins is usually done in minimal growth media using bacterial expression systems, commonly standard or modified versions of M9 minimal media employing ^{13}C glucose for carbon labelling, ^{15}N ammonium sulphate or ^{15}N ammonium chloride for nitrogen labelling, and deuterium oxide for deuteration [79], [80].

Purification of isotope-labelled proteins is the next step of sample preparation and probably the most time-consuming. The procedures are the same for purifying non labelled-proteins and are discussed in chapter 2. If the labelled proteins contain tags, fusion targeted affinity columns are the first step of purification after cell lysis. When proteins are not fused, the chemical structure and physical properties of the proteins are the two key parameters used to develop the most efficient purification protocols. Normally 90% purity is sufficient for Heteronuclear NMR studies.

The last step for NMR sample purification is to choose a good buffer in which the protein is concentrated to approximately 1 mM. Phosphate buffer at pH 5–8 (20–50 mM) with or without salt (e.g., KCl, NaCl) is often used for many NMR samples. High quality NMR tubes should be used for protein samples, which are usually tubes 5 mm in diameter containing 0.5 mL 95% H_2O/5% $^{2}H_2O$ for aqueous samples. If the volume of the sample is limited, microtubes can be chosen for a total sample volume of approximately 200 μL, such as Shigemi micro tubes (Shigemi Inc., Allison Park, PA). The buffer contains from 5-7% $^{2}H_2O$ used for ^{2}H lock. In addition, the samples are usually required to be degassed by blowing high purity argon or nitrogen gas into them to remove oxygen—the paramagnetic property of which will broaden the line shapes of protein resonances [53].

4.9 Murine p22HBP backbone assignments

Previously, Dias *et al.* [10] obtained resonance backbone assignments and determined the solution structure of murine p22HBP by NMR spectroscopy. The researchers used standard methods to obtain chemical shifts assignments. H_N, Hα, N_H, CO, Cα and Cβ resonances were carried out manually by using the 2D ^{1}H-^{15}N HSQC and ^{1}H-^{15}N TROSY and 3D trHNCO, trHN(CO)CA, trHNCA, trHN(CO)CACB, trHNCACB, ^{15}N-edited NOESY-HSQC (mixing time, 60 ms), and ^{15}N-edited TOCSY-HSQC (mixing time, 43.2 ms) spectra. Starting point was peak picking in the 2D TROSY spectrum and transposed to trHNCO where the spin systems were identified. The peaklist generated in trHNCO was loaded onto the trHN(CO)CA, trHNCA, trHN(CO)CACB and trHNCACB spectra, where the alfa and beta carbon resonances were identified for each spin system. With these 3D spectra, sequential connectivities between different spin systems were determined as shown in Figure 4.12.

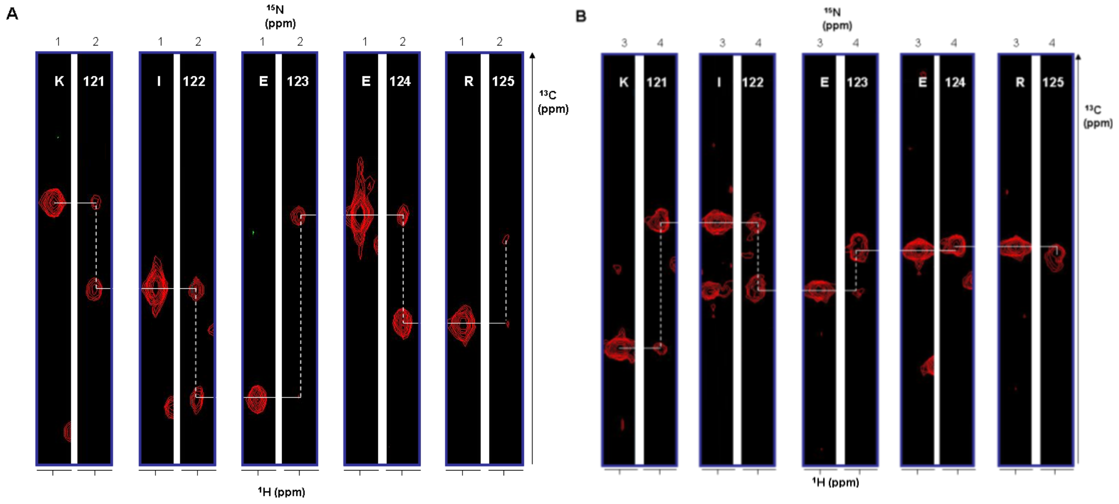


Figure 4.12. Sequential backbone assignments of the Lys 121 to Arg 125 resonances for murine p22HBP. A) trHN(CO)CA (1) and trHNCA (2) spectra acquired in the ^{15}N/^{13}C labelled p22HBP sample; B) trHNCO (3) and HN(CA)CO (4) spectra acquired in the ^{2}H/^{15}N/^{13}C labelled p22HBP sample.

More than 90% of the ^{1}H, ^{15}N and ^{13}C protein resonances were assigned although some solvent exposed amide resonances in the N-terminal (residues 7-17) and loop residues 173-181 were undetected due to exchange broadening, along with residues 80, 90, 136, 152-3, 157 and 164.

Chemical shifts differences were also studied by NMR for murine p22HBP with PPIX and Hemin and several ^{15}N labelling were prepared with increasing amounts of PPIX or Hemin. A 2D ^{1}H-^{15}N TROSY spectrum was collected for each sample and chemical shift changes were monitored upon the titration with either PPIX or Hemin. Researchers observed that in the presence of less amounts of porphyrins, two sets of TROSY signals were registered, corresponding to the free protein and to the PPIX-bound form. With excess of PPIX, the original signals corresponding to free protein were no longer visible. This was explained by the formation of a high affinity 1:1 complex with an off rate corresponding to the slow exchange regime of the chemical shift time scale. A more detailed study of porphyrin titrations was developed: 3D trHNCO and trHNCA for three samples, murine p22HBP, murine p22HBP-PPIX and murine p22HBP-Hemin. PPIX and Hemin induced chemical shift differences were mapped as a function of murine p22HBP sequence in order to identify the tetrapyrrole binding site (Figure 4.13).

Figure 4.13. Experimental ^{1}H chemical shift differences observed for murine p22 HBP upon binding of PPIX or Hemin plotted as a function of the p22HBP amino acid sequence and secondary structural elements [10].

As shown previously, hemin induced some signals to be lost which was attributed to extreme broadening by the paramagnetic Fe(III) center (gray boxes along the residue axis). Despite the difference in the magnitude of the shift perturbations by PPIX and hemin, the patterns of chemical shift mapping were identical suggesting that the two ligands bind the same site in similar orientations. The largest shift differences were observed in a cleft bounded by the α1 helix (around M63) and the β8– β9 loop (residues 171–180) and they were shielding in nature suggesting that the middle of the αA helix must be above or below the plane of the porphyrin ring. In addition, because the porphyrin can induce strong and highly directional ring current shifts, relatively distant parts of the structure were also affected [10].

4.10 Materials and methods

4.10.1 NMR measurements

Data acquisition for human p22HBP backbone assignment

NMR spectra were recorded on a Bruker DRX500 equipped with a four-channel probe and triple axis gradients, and more recently Bruker Avance III HD 700, at 303 K. Bruker Avance II 800 MHz from CERMAX-ITQB. These spectrometers are integrated in Portuguese Nuclear Magnetic Resonance Network (PTNMR). 3mm Shigemi micro tubes (Shigemi Inc., Allison Park, PA) with 250 µL of protein sample were used in all NMR experiments. Human p22HBP experiments for backbone assignment were performed with double labelled protein (^{15}N and ^{13}C) at approximately 1.0 mM concentration, in 50 mM phosphate buffer pH 8.0, with 10 % D_2O (CortecNet). Protein deuteration was obtained by using Deuterium oxide (D_2O) 99.8% (Cortecnet) in minimal media preparation instead of distilled water. All NMR data were processed and analyzed by NMRPipe/NMRDraw [81], Topspin and CARA [82] software. In Table 4.3 is described the list of experiments carried out to human p22HBP backbone assignment.

Experiment	**Pulseprogram**	**FID size**	**Nucleus**	**Spectral width (Hz)**	**Acquisition mode**	**scans**
^{15}N HSQC-NOESY	noesyhsqcf3gpwg3d	2048	^{1}H	12820.513	DQD	ns 256
trHNCO 2H	trhncogp3d	2048	^{1}H	10775.862	DQD	ds 32
		40	^{15}N	2189.445	Echo-antiecho	ns 16
		128	^{13}C	3320.446	States-TPPI	
trHNCA	trhncagp3d2	2048	^{1}H	9615.385	DQD	ds 32
		40	^{15}N	2311.077	Echo-antiecho	ns 16
		128	^{13}C	4829.151	States-TPPI	
trHNCOCA 2H	trhncocagp2h3d	2048	^{1}H	9615.385	DQD	ds 32
		40	^{15}N	2222.219	Echo-antiecho	ns 16
		128	^{13}C	4829.151	States-TPPI	
trHNCOCACB 2H	trhncocacbgp2h3d	2048	^{1}H	9615.385	DQD	ds 32
		40	^{15}N	2222.222	Echo-antiecho	ns 16
		128	^{13}C	11363.637	States-TPPI	
^{1}H^{15}N-TROSY	trosyf3gpphsi19.2	1024	^{1}H	8417.509	DQD	ds 32
		256	^{15}N	2432.713	Echo-antiecho	ns 8

Table 4.3. Human p22HBP NMR experiments. Human p22HBP NMR spectra parameters, including pulseprogram, FID size, number of scans, spectral width and corresponding acquisition mode, acquired for backbone assignment.

For murine p22HBP variants chemical shift mapping, ^{1}H^{15}N-TROSY spectra were acquired, in 50 mM phosphate buffer, pH 8.0, 10 % D_2O, using trosyf2gpst19 as pulseprogram, with 128 ns and 256 ds. For ^{1}H, it was used a FID size of 2048, spectral width of 8012.82 Hz and acquisition mode of DQD. For ^{15}N, it was used a FID size of 256, spectral width of 2027.351 Hz and States-TPPI as acquisition mode.

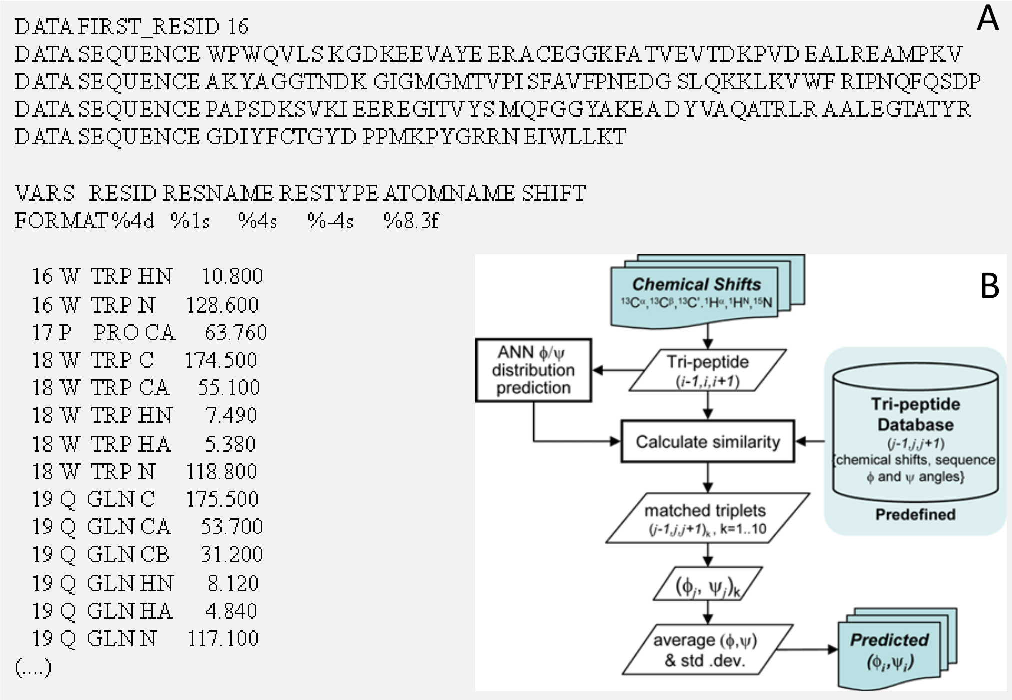
Tetrapyrrole titrations of murine p22HBP, variants and human p22HBP.

For chemical shift mapping; uniformly ^{15}N, ^{13}C-labeled samples of murine p22HBP variants (50 mM phosphate buffer, pH 8.0) in H_2O/D_2O (7%) at a concentration of 1 mM were used. All NMR spectra were acquired at 303 K on a Bruker DRX500 spectrometer equipped with a four-channel probe and triple axis gradients. A series of 2D spectra, ^{1}H, ^{15}N-TROSY were acquired for p22HBP variants: K64A, K177A, K177E, R56A/K64A and R56A/K63A/K177A; in the presence (p22HBP:PPIX, 1:1.4) and absence of Porphyrin. Due to their poor solubility in acidic and neutral pH, tetrapyrroles were initially dissolved in ammonia followed by dilution in water. After addition of a surfactant, Tween 80 (1.5% v/v), the pH was adjusted to 8.0 with KH_2PO_4 and the volume adjusted to 1 mL. Hemin and PPIX concentrations used in titration were estimated by UV spectroscopy. Molar absortivity for PPIX was found to be 97071 $M^{-1}cm^{-1}$ while for Hemin was 32257 $M^{-1}cm^{-1}$.

4.10.2 Talos+ calculations

Assigned residue assignments for human p22HBP were used as input to the program Talos+ to predict protein phi and psi backbone torsion angles [83]. Talos+ is a hybrid system for empirical prediction of protein phi and psi backbone torsion angles using a combination of six kinds of chemical shift assignments (HN, Hα, Cα, Cβ, CO, N) for a given residue sequence [83]. In Figure 4.14 a scheme of how TALOS+ perform the Ψ and Φ predictions is shown.

The Talos+ server was used (http://haddock.chem.uu.nl/enmr/services/TALOS/) and a chemical shift file in Talos format (Figure 4.14 B) was prepared and uploaded.



```
DATA FIRST_RESID 16
DATA SEQUENCE WPWQVLS KGDKEEVAYE ERACEGGKFA TVEVTDKPVD EALREAMPKV
DATA SEQUENCE AKYAGGTNDK GIGMGMTVPI SFAVFPNEDG SLQKKLKVWF RIPNQFQSDP
DATA SEQUENCE PAPSDKSVKI EEREGITVYS MQFGGYAKEA DYVAQATRLR AALEGTATYR
DATA SEQUENCE GDIYFCTGYD PPMKPYGRRN EIWLLKT

VARS   RESID RESNAME RESTYPE ATOMNAME SHIFT
FORMAT %4d   %1s     %4s     %-4s     %8.3f

  16 W  TRP HN     10.800
  16 W  TRP N     128.600
  17 P   PRO CA     63.760
  18 W  TRP C     174.500
  18 W  TRP CA     55.100
  18 W  TRP HN      7.490
  18 W  TRP HA      5.380
  18 W  TRP N     118.800
  19 Q  GLN C     175.500
  19 Q  GLN CA     53.700
  19 Q  GLN CB     31.200
  19 Q  GLN HN      8.120
  19 Q  GLN HA      4.840
  19 Q  GLN N     117.100
(....)
```

Figure 4.14. A- Chemical shift file in TALOS format. B- Flowchart for TALOS+ database search procedure [83].

4.11 Results and discussion

4.11.1 Chemical shifts differences of p22HBP variants

NMR experiments were carried out in order to confirm that no structural changes have been introduced by mutations in p22HBP and were evaluated by chemical shift mapping using ^{1}H-^{15}N TROSY experiments.

A preliminary TROSY spectrum (Figure 4.15) of ^{15}N labelled p22HBP-K64A sample showed good chemical shift dispersion (indicative of folded protein) and contained the expected number of signals for a 190 residue protein. The backbone assignments of p22HBP variants were performed using TROSY spectra superimposed on wild type murine p22HBP spectrum.

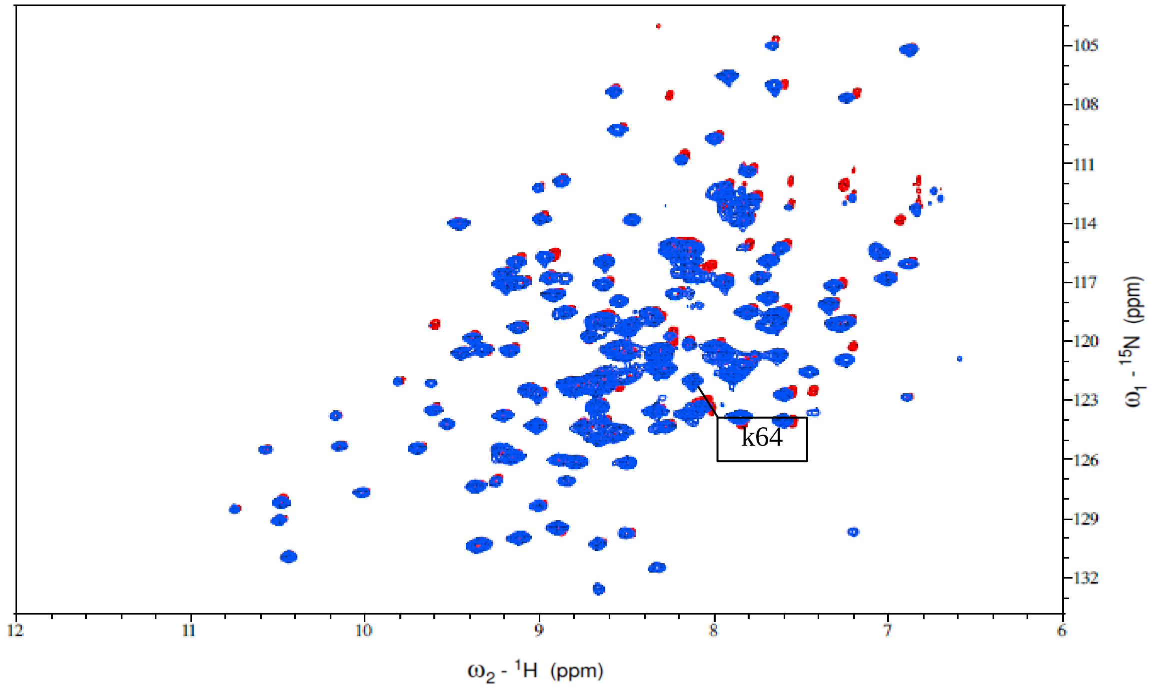
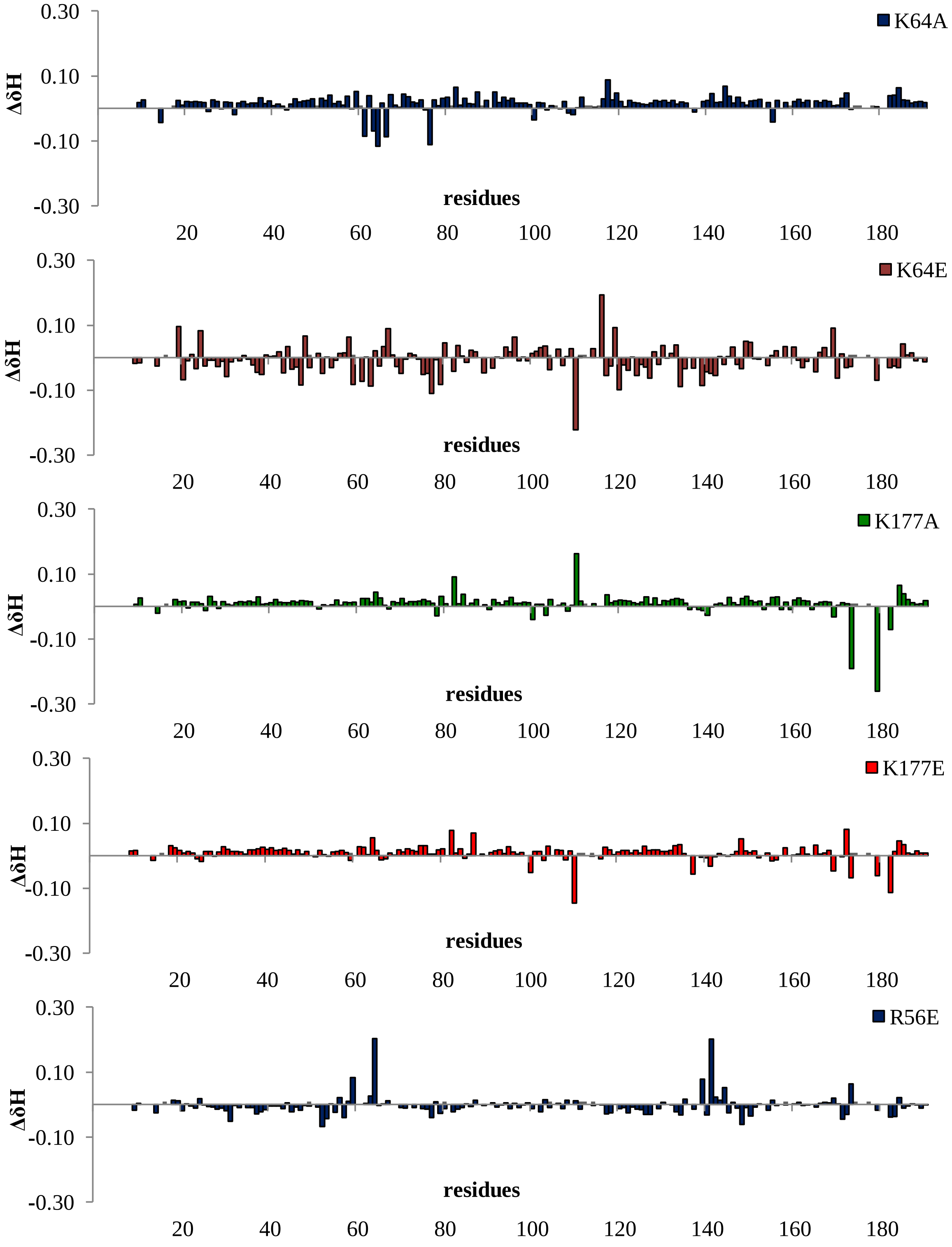


Figure 4.15. ^{1}H,^{15}N TROSY spectra of murine p22HBP (red), and K64A (blue) recorded at 500 MHz. Both spectra were acquired at 303 K in 50 mM phosphate buffer at pH 8.0.

Chemical shift differences were observed for the variants, mainly located near the modified residues. Figure 4.16. shows the main differences between p22HBP variants and wild type murine p22HBP. The small differences indicate similar, correctly folded, structures in solution


K64A
K64E
K177A
K177E
R56E
ΔδH
0.30
0.10
-0.10
-0.30
residues
20
40
60
80
100
120
140
160
180

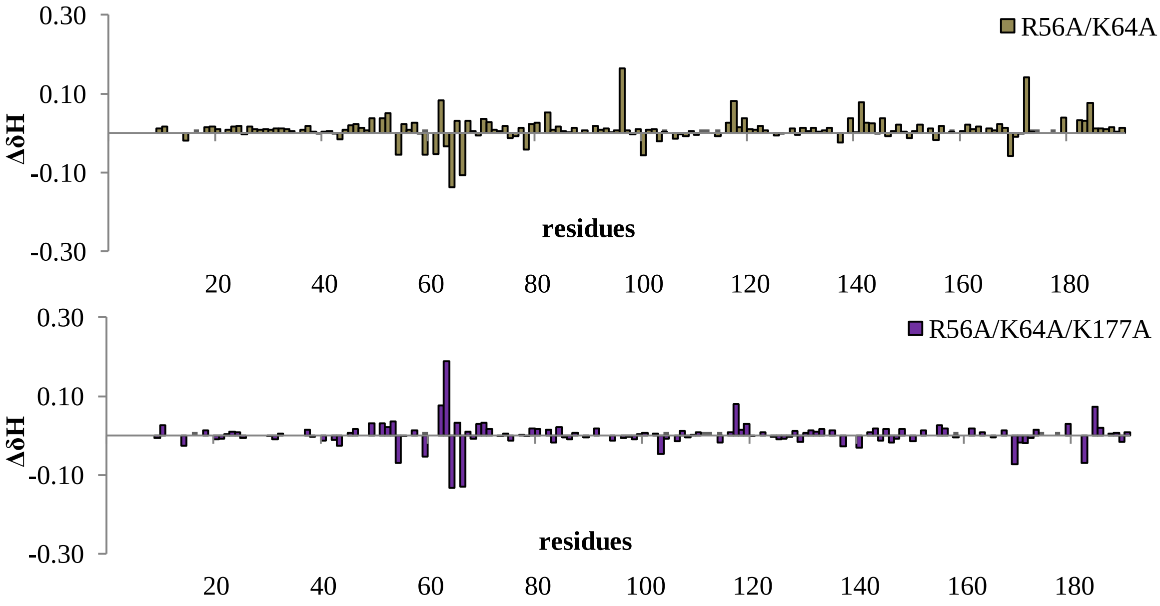


Figure 4.16. Chemical shift mapping for p22HBP variants when compared to murine p22HBP. Grey bars correspond to prolines.

For K64, the main differences were observed around residue K64. When lysine was replaced for an alanine, the shift differences were lower when compared with those observed in variant with lysine replaced by Glutamic acid. In the overall, the proton shifts differences were less than 0.1, and the structure remain as murine p22HBP wild type. Residues G110 and S116 experiment a chemical shift difference of -0.222 and 0.193 respectively in K64E. S116 is far from K64 in sequence but with tertiary structure, are close in space and thus also experiment chemical environment changes.

In both K177 variants, chemical shift differences were observed around residue 177 and also in residue G110. Curiously, although G110 is not close in space with K177, in K177A, the spin system of Glycine 110 moves to the left (with high ^{1}H chemical shift) while in K177E moves to the right, which corresponds to lower ^{1}H chemical shift.

For R56E variant, the main differences were also observed near the mutation site as expected. Residue E141 is close in space with R56 and the presence of a Glutamic acid instead of Arginine, promoted the peak deviation to left which corresponds to high ^{1}H chemical shift.

Double and triple variants, experiment the same behavior as in single variants, and the main chemical shift differences were observed around residues R56 and K64. In

R56A/K64A a new deviation is observed in residue K96 although it was not confirmed in R56 and K64 variants. This residue is located near R56 and could experiment the hydrophobic changes with the replacement of Arginine 56 by an Alanine.

4.11.2 Tetrapyrrole titrations of murine p22HBP, variants and human p22HBP

NMR studies were also used to evaluate the interactions of murine p22HBP variants with PPIX. There is slow exchange (on the NMR chemical shift time scale) between free and bound forms. This allows most peaks in the bound form to be identified relatively easily and therefore allows mapping of the interaction site. In Figure 4.17, an overlay of the TROSY spectrum of ^{15}N labelled p22HBP-K64A with PPIX and p22HBP-K64A (free) is shown.

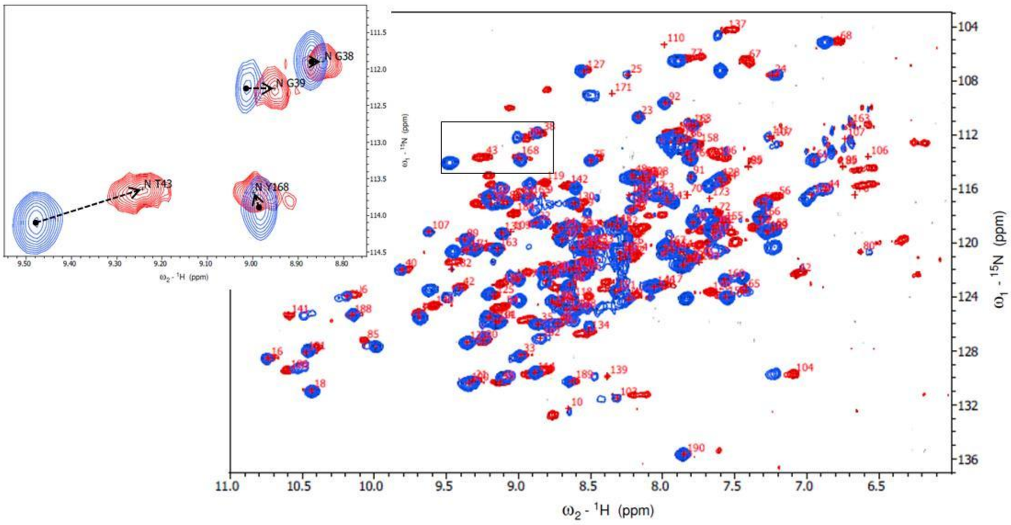
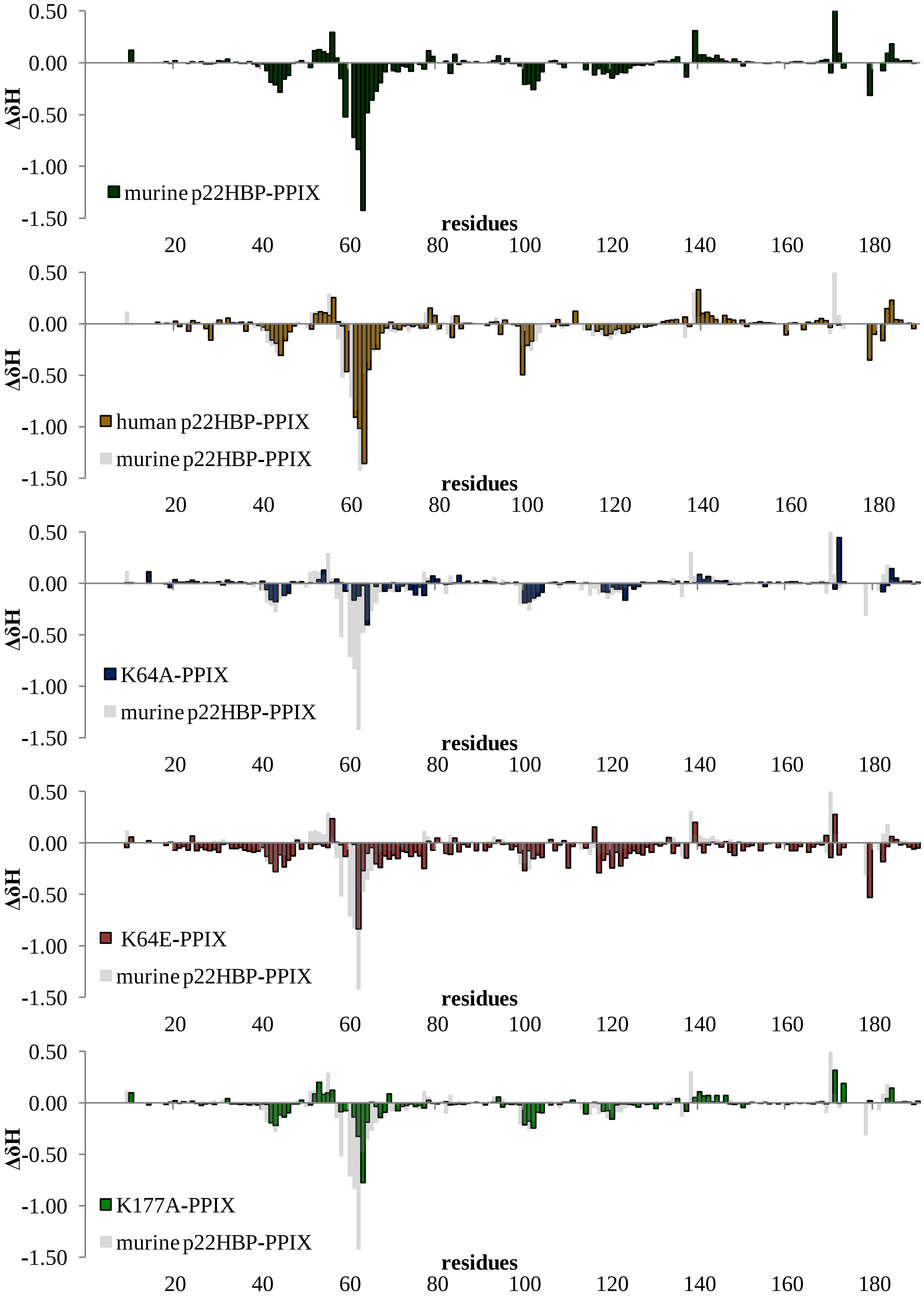


Figure 4.17. ^{1}H,^{15}N TROSY spectra of K64A (blue) and K64A-PPIX (red) recorded at 500 MHz. Both spectra were acquired at 303 K in 50 mM phosphate buffer at pH 8.0.

By analyzing the chemical shift differences observed in the TROSY spectra of all variants with PPIX (spectra not shown) it was possible to determine chemical shifts differences and identify the location of the tetrapyrrole binding site for the variants (Figure 4.18).


0.50
0.00
-0.50
-1.00
-1.50
ΔδH
murine p22HBP-PPIX
residues
20
40
60
80
100
120
140
160
180
human p22HBP-PPIX
murine p22HBP-PPIX
K64A-PPIX
murine p22HBP-PPIX
K64E-PPIX
murine p22HBP-PPIX
K177A-PPIX
murine p22HBP-PPIX

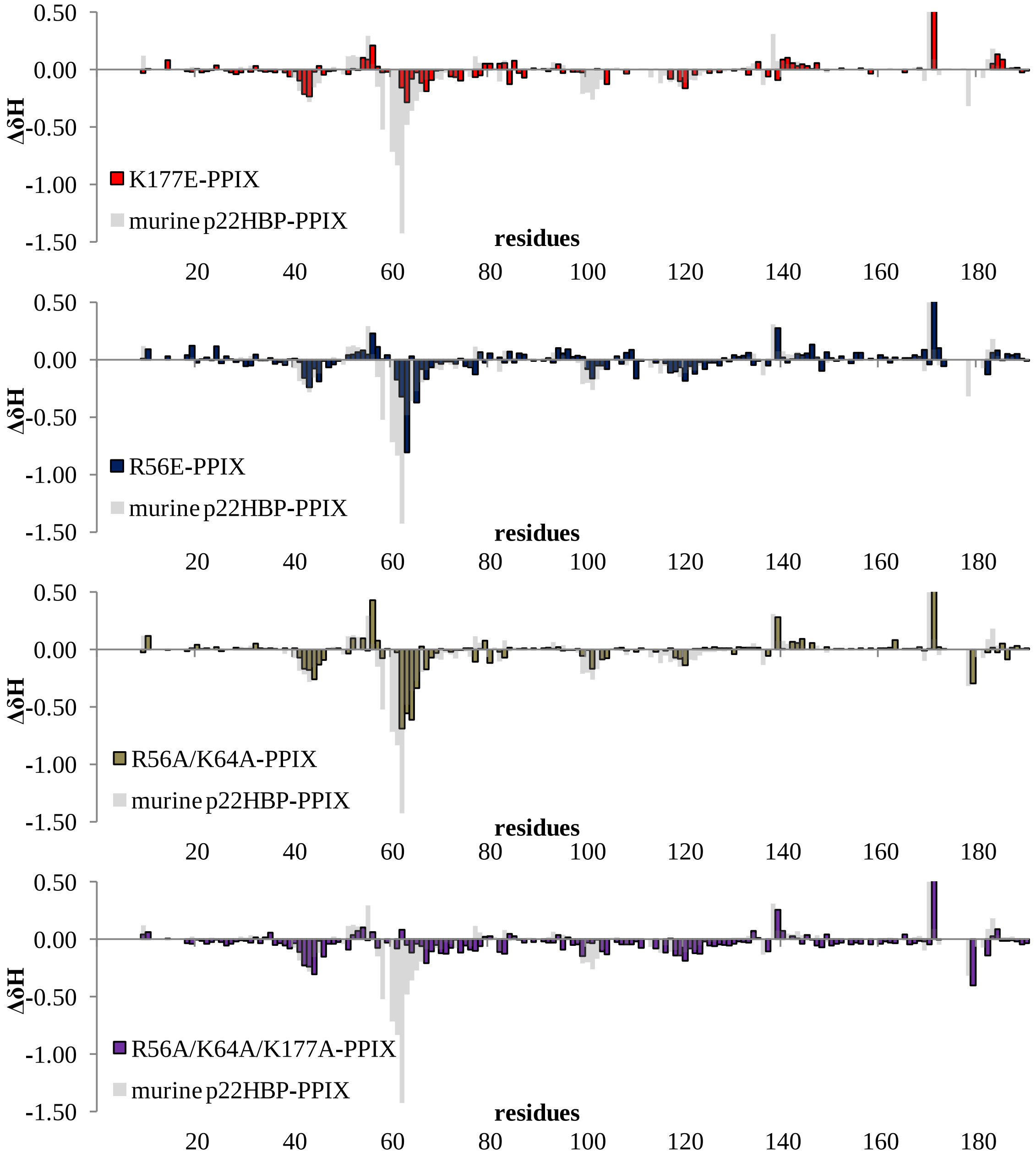


Figure 4.18. ^{1}H chemical-shift deviations plotted by the residue number for human and murine p22HBP and respective variants, upon binding to PPIX.

According to Figure 4.18, the largest chemical shifts differences of p22HBP upon binding with PPIX were observed in murine and human p22HBP. As the same pattern of chemical shift differences is observed for human p22HBP and variants when compared to murine p22HBP, it is possible to conclude that the mode and position of binding is very similar to the wild type form. The largest shifts differences were observed for the α-helix extending from Valine 51 to Glycine 67 and were less pronounced in the region around residues 105 and 171. In human p22HBP and variants, PPIX also induces positive and negative shifts as reported for murine p22HBP [10].The ring currents of porphyrin induce shielding in nuclei above and below the plane of the porphyrin ring and deshielding around the "edge" of the ring.

4.11.3 Backbone Resonance assignments for human p22HBP

Backbone chemical shift assignments for human p22HBP were obtained by a variety of strategies: human/murine p22HBP sequence alignment, triple resonance experiments, talos+ calculations and tetrapyrrole titration experiments. Initially, the assignment of H_N, N_H, peaks was carried out manually by using 2D ^{1}H-^{15}N HSQC and ^{1}H-^{15}N TROSY experiments and by comparing assigned peaks for the murine form of p22HBP with the human form. As the two forms have 87% homology, the resonance signals of conserved residues in both human and murine p22HBP were expected to appear in the same region of the HSQC/TROSY spectrum. This strategy however did not provide complete assignment and therefore triple resonance experiments were carried out: 3D trHNCO, trHN(CO)CA, trHNCA, trHN(CO)CACB, trHNCACB, ^{15}N-edited NOESY-HSQC and ^{15}N-edited TOCSY-HSQC spectra. Spin systems picked in 2D ^{1}H-^{15}N HSQC spectra were correlated with trHNCO peaks to identify NH backbone peaks and to help determine how many peaks were appearing in overlapped regions of the 2D HSQC spectra. Although the TROSY-HNCO is highly sensitive, it provides only one C'_{i-1} peak per residue and therefore does not allow the establishment of connectivity between spin systems. Sequential connectivity was achieved by comparing trHNCA and trHN(CO)CA spectra: trHNCA correlates each NH group with the C_α of its own residue (strongly) and of the preceding residue (weakly) and trHN(CO)CA only correlates the NH group with the preceding C_α chemical shifts. Figure 4.19 shows the sequential assignment of segment D91-Q95 using a 3D trHNCA spectrum.

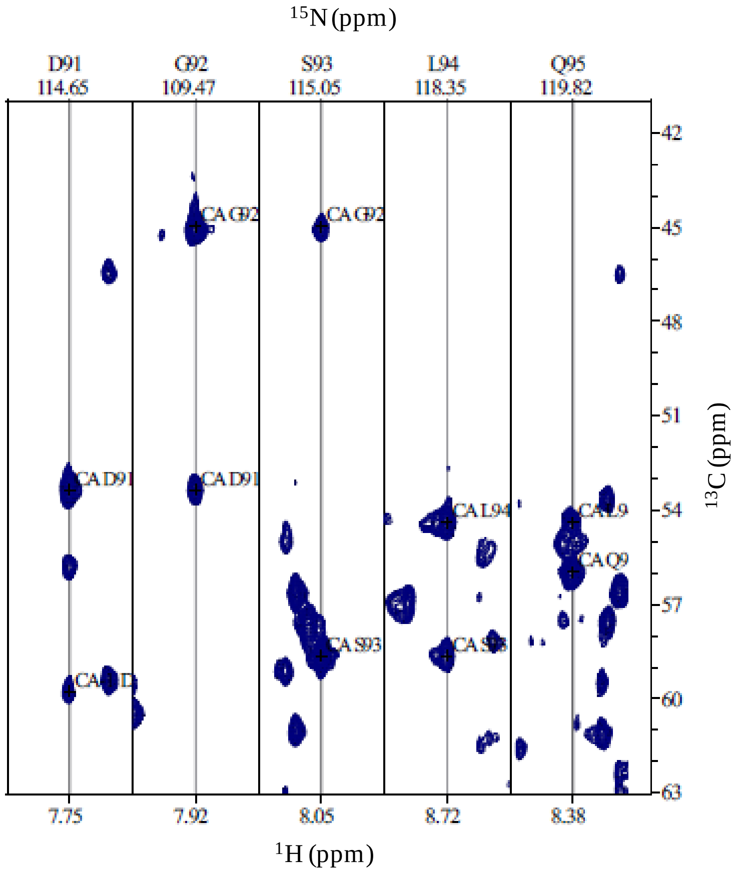


Figure 4.19. Sequential assignments of the resonances from residues D91 to Q95 of human p22HBP using the trHNCA spectrum.

trHN(CO)CACB and trHNCACB spectra were used, in addition to trHNCA and trHNCOCA spectra, to identify alpha and beta carbon resonances for each spin system and to determine sequential connectivity between different spin systems. A set of spectra was also recorded for a deuterated form of human p22HBP at high field (800MHz, ITQB). This allowed a trHN(CA)CO experiment to be obtained and compared to an trHNCO to aid sequential assignment. Data from a PPIX titration was also used to confirm the assignments. Figure 4.20 shows the chemical shifts changes observed on PPIX binding for human p22HBP using the final set of assignments.

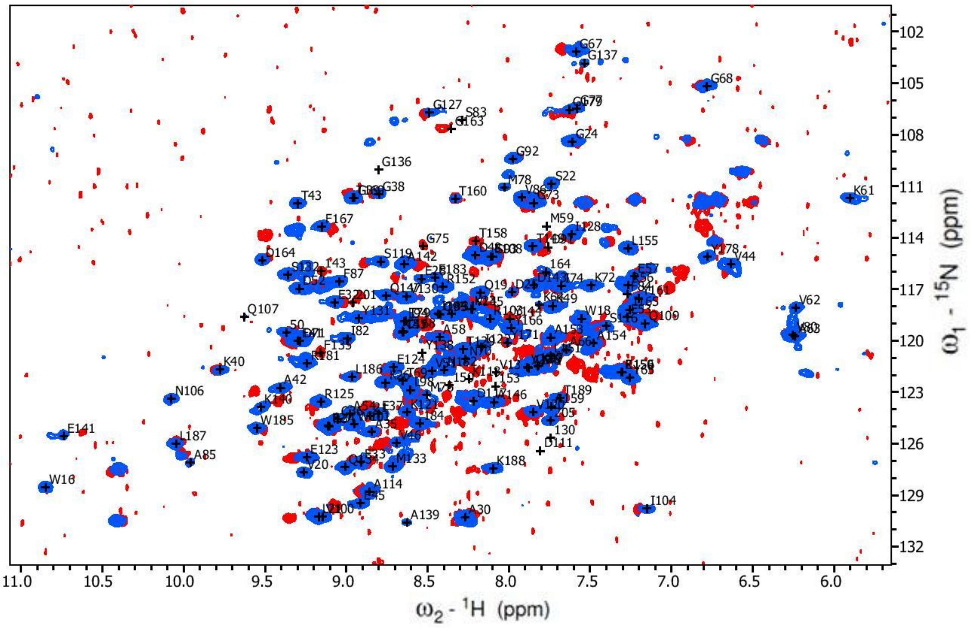


Figure 4.20. ^{1}H, ^{15}N TROSY of human p22HBP titrated with PPIX (1:1.4) at 303K. Red-human p22HBP free, blue-human p22HBP titrated with PPIX.

Assuming a similar structure and similar binding mode for PPIX the very similar pattern seen for the chemical shift changes for the human form compared to the murine form confirms the backbone assignments are correct (Figure 4.18). Finally as a last step a Talos+ calculation was run using the assigned ^{1}H, ^{15}N and ^{13}C resonances. TALOS+ is a hybrid system for empirical prediction of protein phi and psi backbone torsion angles using a combination of six kinds of chemical shift assignments (HN, Hα, Cα, Cβ, CO, N) for a given residue sequence. Assuming the structures of the murine and human forms are very similar the secondary structure predicted by Talos+ was compared to the actual secondary structure seen for murine p22HBP in solution. At this point some assignments that were indicating no secondary structure using Talos+ but had been assigned to a section of protein with an α-helix were modified and the Talos+ calculation rerun. This was repeated until the best match between predicted and “actual” secondary structure was obtained. The complete results from the Talos+ prediction are shown in Appendix 9.1 and it is possible to see that the predicted results for human p22HBP are in agreement with murine p22HBP and the secondary structural elements are well defined. A complete table of chemical shifts for assigned nuclei in human p22HBP based on trHNCA, trHNCOCA, trHNCACB, trHNCOCACB, HN(CA)CO and HNCO spectra and Talos+ calculations at 303 K is shown in appendix 9.2. Figure 4.21 shows a visual summary of the backbone human p22HBP assignments.

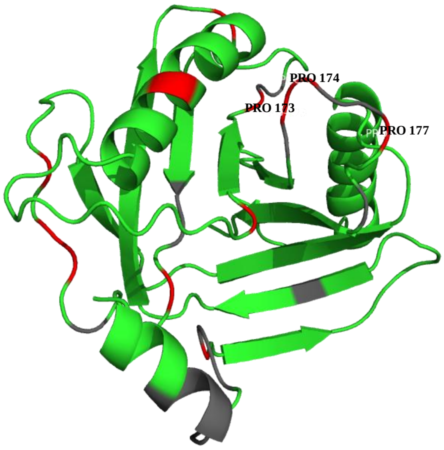


Figure 4.21. Human p22HBP assignments overview. Assigned residues are colored in green and unassigned residues are colored in grey. Prolines are colored in red. (Structure was calculated using the Robetta server).

Approximately 82% of all backbone resonances and 92% of the structured regions of the protein were assigned while for the unstructured loop regions only 63% of observed peaks were assigned. There are still 22 peaks that have not been assigned in the HSQC spectrum and 9 residues are absent in HSQC spectra. Rapid exchange is probably the reason for the lack of peaks in the loop regions.

The assignments were also be checked using the chemical shift index (CSI) used for the identification of protein secondary structure [84]. This method is based on the secondary structure shift, which is the difference between the observed chemical shift and the random coil value (Table 4.2) assigned to this amino acid type in the unfolded conformation. The chemical shift index method is used for assigning secondary structural elements in proteins before structure calculation. The secondary structure is established by an alpha helix if four or more sequential negative $\Delta\delta$HA (classified as -1) and/or positive $\Delta\delta C\alpha$ (classified as +1) are found and a beta-strand if three or more sequential positive $\Delta\delta H\alpha$ (+1) and/or negative $\Delta\delta C\alpha$ (-1) are found. In Figure 4.22, the CSI for human p22HBP is shown and a complete table with CSI results is shown in appendix 9.2.

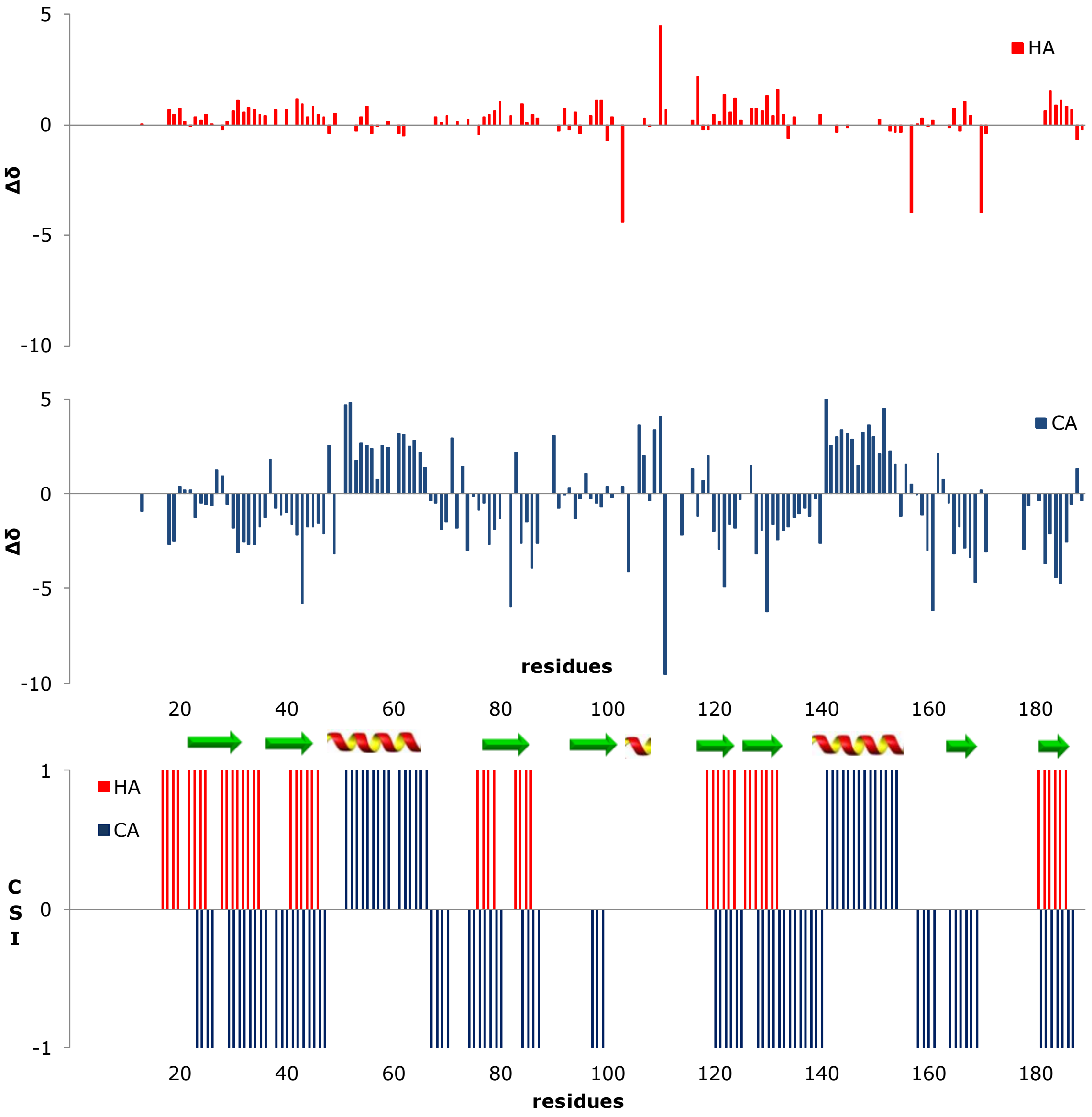


Figure 4.22. Chemical shift index plotted as a function of human p22HBP sequence. Red bars are related to HA chemical shifts variations and blue bars are related to CA chemical shifts variations.

The CSI results indicate a β-sheet for residues 31-36 and 39-48 and an α-helix is seen for 51-67 although HA chemical shifts variations are not well defined. Four more β-sheets are located for residues 82-87, 96-103, 120-125 and 128-135. Another α-helix is seen for residues 141-158 and the C-terminus end with two β-sheets 166-170 and 182-186. These results are in agreement with secondary structures seen in the murine p22HBP structure.

5 Conformational dynamics of human p22HBP

5.1 Introduction

The relationship between structure, dynamics and function in complex molecular systems is one of the fundamental problems in understanding life at the molecular level. Most of biological functions in cells are performed by proteins, which make proteins the molecular machines that control all key events in living cells. Static three-dimensional structures and respective amino acid sequences provide very important information on the organization and on interactions with other molecules. However, proteins are not strictly static and assume a variety of ensembles or conformations that are essential for most protein function. There are a variety of possible protein motions such as, bond vibrations, side-chain rotations, segmental motions and domain movements. These motions occur over a variety of length and time scales, and have been linked to functionally relevant phenomena such as allosteric signaling and enzyme catalysis [85][86]. As dynamic properties intrinsic to a protein structure are extremely complex (a variety of motions occur in the same molecule at the same time) they can be difficult to analyze. However due to the nature of the NMR method, isotopic labelling coupled with multidimensional NMR methods can be used to monitor dynamic behaviour in solution, and can provide a complete picture of protein motions ranging from picoseconds up to seconds. The type of motion that is accessible to NMR is illustrated in Figure 5.1.

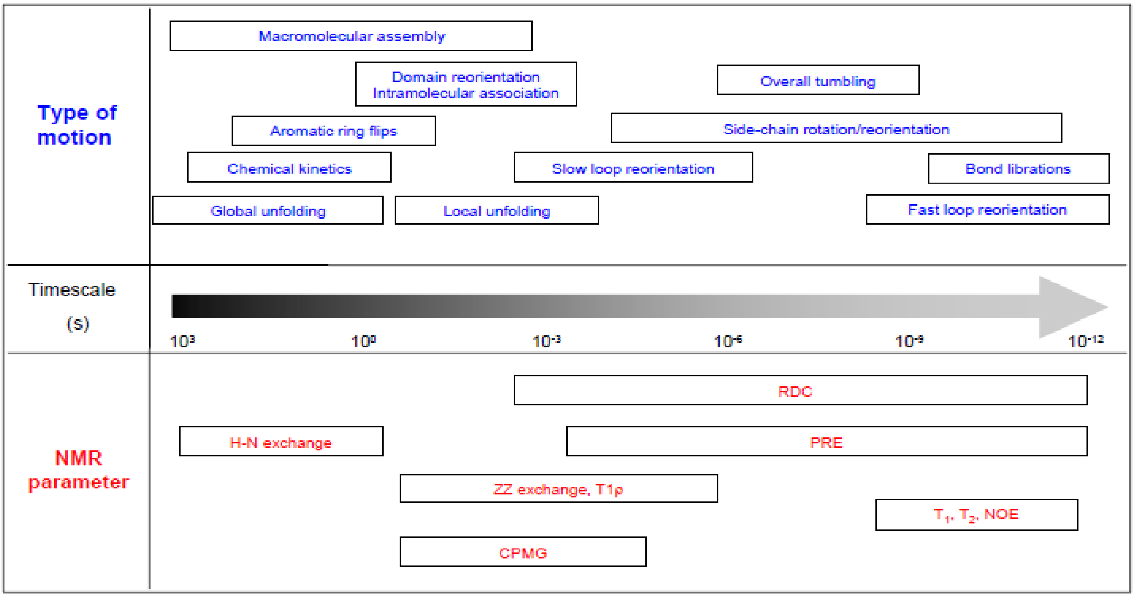


Figure 5.1. Dynamic time scales, types of protein motions and NMR parameters available to study these processes [87].

5.1.1 Correlation times and Spectral Density Functions

In an external magnetic field, molecular tumbling and internal motions cause reorientation of an ^{1}H-^{15}N spin pair (represented with a single bond vector) with respect to this external magnetic field and the local magnetic field at the ^{15}N nucleus fluctuates due to the directly attached ^{1}H magnetic dipole (Figure 5.2). Despite the fact that the local dipolar interaction between the ^{1}H and ^{15}N spins average to zero, because of molecular tumbling, the time-dependent fluctuations in the field at any point in time are not zero and cause a spin system that has been perturbed by radio-frequency pulses to return to thermal equilibrium. This process is called relaxation. Fluctuations of the local magnetic fields are sensitive to overall molecular tumbling and internal motions and the measurement of NMR relaxation rates provides an important source of information about the dynamics of a system.

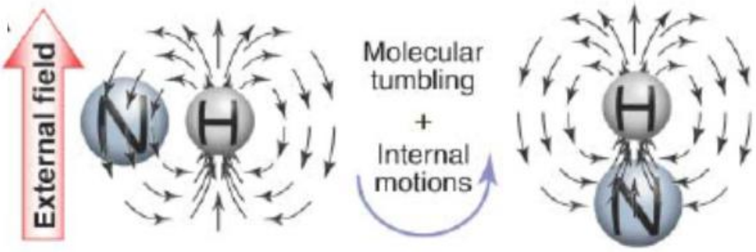


Figure 5.2. A. Orientation-dependent magnetic field experienced by an amide ^{15}N nucleus due to the directly bonded proton [88].

The magnitude of the changes depends on how fast the molecule tumbles. Considering the orientation of the bond vector, with a fixed orientation with respect to a molecular frame of reference, at time t and at a time $t+\delta$, for a large molecule which rotates slowly, the orientation at time $t+\delta$ is very similar to the orientation at time t: both orientations are correlated to a high degree. On the other hand, if the molecule tumbles fast, the bond vector orientations at time t and $t+\delta$ are very different, and are not correlated to each other anymore.

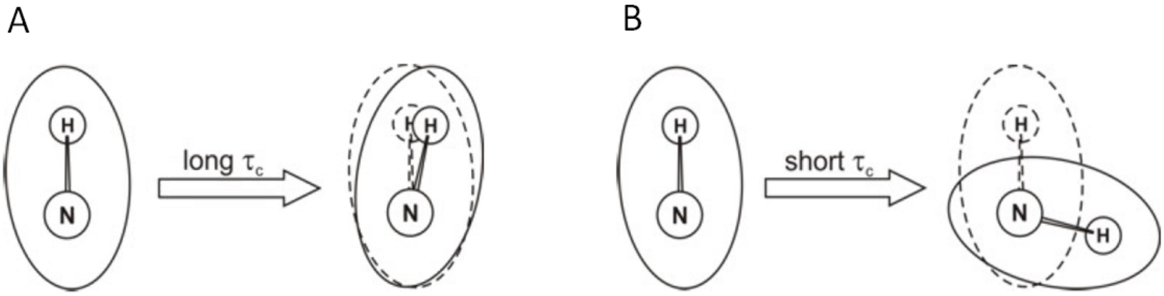


Figure 5.3. ^{15}N–^{1}H bond vector orientation for long correlation time (A) and short correlation time (B) [89].

This correlation can be described by a correlation function C(t) and for isotropic rotational diffusion of a spherical top, it is

$$C(t) = A_c e^{-\frac{t}{\tau_c}}$$

Equation 5.1

where the normalization constant $A_c$equals $\frac{1}{5}$ and τ_c is the rotational correlation time of the molecule. The rotational correlation time can also be related to the hydrodynamic property of a molecule via Stoke's law:

$$\tau_c = \frac{4\pi\eta_w r_H^3}{3\kappa_B T}$$

Equation 5.2

in which η_w is the viscosity of the solvent, $r_H$is the effective hydrodynamic radius of the solute, κ_B is the Boltzmann constant and T is the temperature. According to Equation 5.2, large values of τ_c correspond to slow tumbling (large molecules or low temperatures) and small values of τ_c indicate fast tumbling (small molecules or high temperatures). The Fourier transformation of a correlation function results in the corresponding spectral density function $J(\omega)$:

$$J(\omega) = A_j \frac{\tau_c}{1 + \omega^2 \tau_c^2}$$

Equation 5.3

As the correlation time is dependent on the motional regime (affected by molecular size), the spectral density function is different for different regimes i.e. slow, intermediate or fast [90]. Considering a small molecule with rapid tumbling and consequently short τ_c, there is an almost equal chance of having components at both high and low frequencies, up to ca. $1/\tau_c$ at which point the probability drops rapidly. On the other hand, for large molecules, with slow tumbling and consequently high τ_c, the probability of generating rapidly oscillating fields is very small and thus the corresponding spectral density is concentrated into a smaller frequency window.

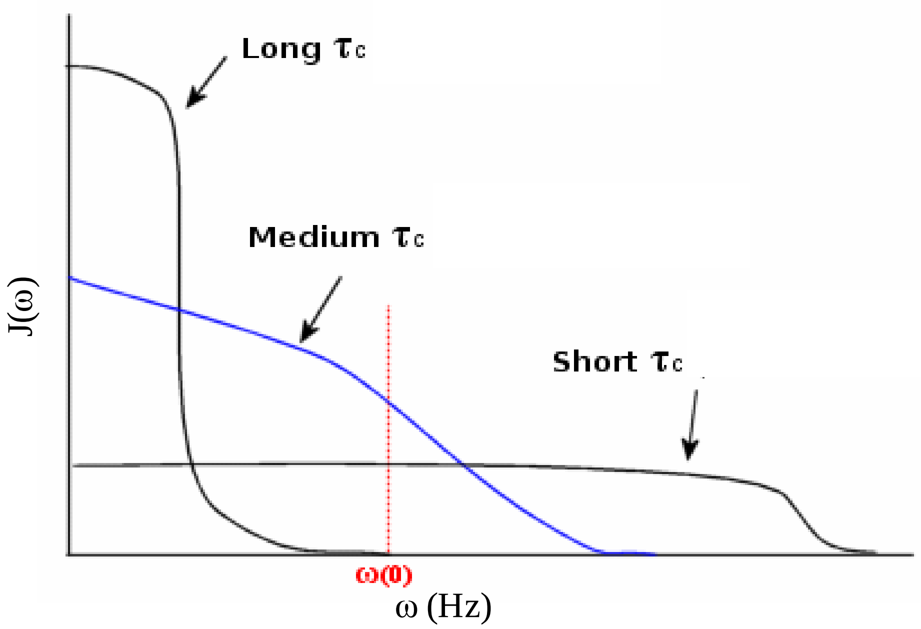


Figure 5.4. Spectral density (J(ω)) variation with the molecular motion as a function of the frequency (ω) [91].

5.1.2 The NOE

One of the most important parameters to evaluate the dynamic behaviour of a protein is obtained from the measurement of the Nuclear Overhauser Effect (NOE), which occurs between protons located close in space (within 5Å), and arises from dipole-dipole cross-relaxation. The NOE can be determined by observing the change in intensity of one resonance when the spin transitions of another spin is perturbed from its equilibrium population [92].

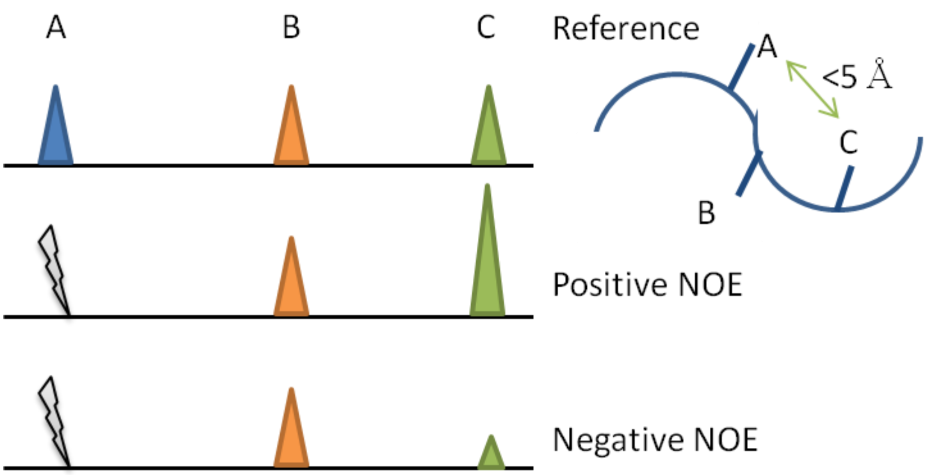


Figure 5.5. Schematic representation of NOE effect. Irradiation of resonance A leads to an increase of peak intensity of the neighboring spin C (positive NOE) or to a decrease of peak intensity (negative NOE).

The NOE is associated with spectral densities, J(ω), which are related via Fourier transformation with their respective correlation functions of reorientional motion, according to Equation 5.4.

$$NOE = 1 - \frac{\left|\frac{\gamma_H}{\gamma_N}\right| d^2[6J(\omega_H + \omega_N) - J(\omega_H - \omega_N)]}{R_1}$$

Equation 5.4

where,

$$d = -\frac{\left[\frac{\mu_0}{4\pi}\right]\gamma_H \gamma_N h}{4\pi r_{HN}^3}$$

Equation 5.5

As the equation for NOE contains a $1/r^6$ dependence on the distance, r, between 2 spins, the effect can be used to extract distance information. This is used to identify protons in a protein that are close in space but distant in terms of protein primary sequence in NMR structure determination [53], [93], [94].

For large molecules and/or high viscosity solvents the zero-quantum relaxation pathway is very efficient (molecular motion is slower than the Larmor precession frequency), and ω_0 >

ω_2. Under these conditions negative NOEs approaching -100% can be observed. It is sometimes worthwhile to manipulate solvent viscosity and temperature to achieve negative NOE's, since these are inherently larger than the positive NOEs seen under conditions of fast molecular motion.

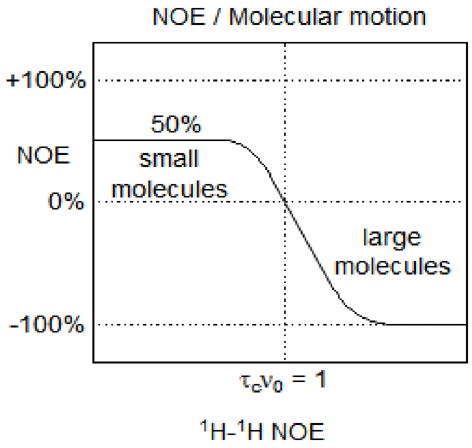


Figure 5.6. Effect of Molecular motion and molecular size on NOE.

The heteronuclear two-dimensional ^{1}H-^{15}N Nuclear Overhauser Effect, or hetNOE, is a widely used NMR experiment to obtain protein dynamics information over fast time scales (pico to nanosecond). Here flexible regions or unstructured parts of the protein and vice-versa can be identified if backbone assignments are available. For instance, ^{15}N-^{1}H HetNOE values lower than 0.65, at 600 MHz, are normally indicative of flexibility on a picosecond timescale for the protein backbone [95]. The HetNOE can be obtained experimentally using the ratio of cross-peak intensities from two experiments, with (I_{NOE}) and without (I_{NONOE}) presaturation of amide ^{1}H nuclear spins, normally known as NOE and NONOE experiments (Equation 5.6). The respective uncertainties, σNOE, are shown in Equation 5.7, where I and δ are the peak intensities and the level of experimental noise, respectively.

$$NOE = \frac{I_{NOE}}{I_{NONOE}}$$

Equation 5.6

$$\sigma NOE = |\mathrm{NOE}|\sqrt{\left(\frac{\delta_{NOE}}{I_{NOE}}\right)^2 + \left(\frac{\delta_{NONOE}}{I_{NONOE}}\right)^2}$$

Equation 5.7

5.1.3 Transverse and Longitudinal Relaxation

^{15}N relaxation experiments are used to probe backbone dynamics in proteins and measure either the recovery of ^{15}N Z-magnetization to its equilibrium value (T_1) or the decay of magnetization orthogonal to the Z axis to zero (T_2) (Figure 5.7).

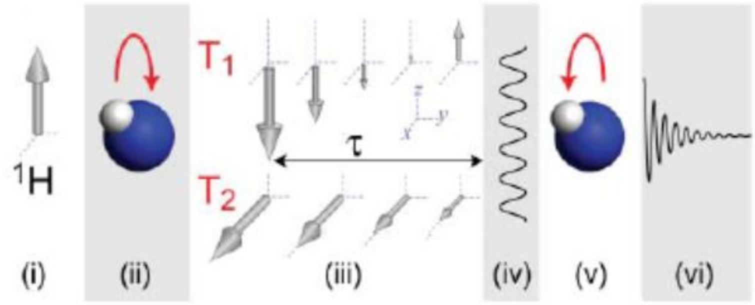


Figure 5.7. Illustration of amide ^{15}N relaxation experiment. (i) equilibrium ^{1}H magnetization, (ii) ^{1}H to ^{15}N transfer, (iii) relaxation delay, (iv) indirect ^{15}N chemical shift detection, (v) ^{15}N to ^{1}H transfer and (vi) direct ^{1}H chemical shift detection [88].

The longitudinal and transverse relaxation rates, R_1 and R_2 are associated with spectral densities, J(ω), which are related via Fourier transformation with the respective correlation functions of reorientional motion. Considering the backbone ^{15}N amide nucleus, the main sources of relaxation are ^{15}N chemical shift anisotropy and dipolar interaction with bound ^{1}H. The relaxation parameters can be defined as:

$$R_1 = 3(d^2 + c^2)J(\omega_N) + d^2[3J(\omega_N) + J(\omega_H - \omega_N) + 6J(\omega_H + \omega_N)]$$

Equation 5.8

$$R_2 = \frac{1}{2}(d^2 + c^2)[4J(0) + 3J(\omega_N)] + \frac{1}{2}\, d^2[J(\omega_H - \omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)] + R_{ex}$$

Equation 5.9

where,

$$d = -\frac{\left[\frac{\mu_0}{4\pi}\right]\gamma_H \gamma_N h}{4\pi r_{HN}^3} \; and \; c = -\omega_N CSA/3$$

Equation 5.10

d represents the contribution from ^{15}N-^{1}H dipolar coupling whereas c represents ^{15}N chemical shift anisotropy. R_{ex} is the conformational exchange contribution to the measured

R_2. These equations are widely used in protein dynamics analysis to estimate spectral densities.

The return of magnetization to the z-axis consequently causes loss of magnetization in the x-y plane, therefore T_2 is always less than or equal to T_1. Thus, all aspects that influence T_1 will also indirectly influence T_2 and all other frequencies acting on the x-y plane will also act on T_2. Equation 5.11 and Equation 5.12 indicates this behavior:

$$\frac{1}{\mathrm{T}_1} = \gamma^2 \overline{H}^2 \frac{\tau_c}{1 + (2\pi \nu_o \tau_c)^2}$$

Equation 5.11

$$\frac{1}{\mathrm{T}_2} = \gamma^2 \overline{H}^2 \frac{\tau_c}{1 + (2\pi \nu_o \tau_c)^2} + \tau_c$$

Equation 5.12

where γ is the magnetogyric ratio, τ_c is the correlation time, ν_o is the Larmor frequency and $\overline{H}^2$ is the mean-square average of the local magnetic fields. Figure 5.8 illustrates how T_2 relaxation decreases with increasing molecular size and tumbling time, τ_c.

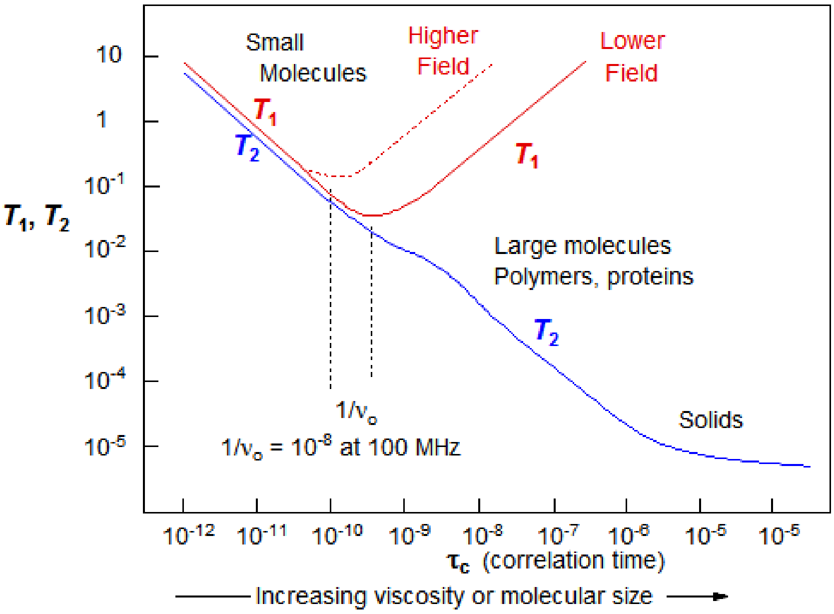


Figure 5.8. T_1 and T_2 behaviour as a function of correlation time. τ = Molecular correlation time: the time it takes the average molecule to rotate one radian [95].

5.1.4 The Model Free Approach for the analysis of relaxation data

The model-free approach was introduced by G.Lipari and A. Szabo in 1982 and extended by G.M. Clore and co-workers and nowadays is the most common way to analyze NMR relaxation data. This approach allows characterization of internal motions on time scales faster than the overall molecular tumbling using the dependence of the longitudinal and transverse relaxation rates R_1 and R_2 and the heteronuclear NOE on the spectral density function $J(\omega)$. The original method introduces two parameters for the study of NMR relaxation data, a generalized order parameter S^2 and an internal correlation time τ_i. Once the spectral density function of this formalism is obtained without invoking a model or any other assumptions on the kind of motions, and S^2 and τ_i are defined in a model independent way, the approach is referred to as “model-free”.

Considering a ^{15}N-^{1}H spin pair in a protein whose overall motion can be described by a single correlation time, the orientation of the bond vector changes due to internal motion, and is not fixed with respect to a molecular frame of reference. Assuming that the overall and internal motions are independent, and this is the fundamental assumption of the Model Free approach, the total correlation function is given as:

$$C(t) = C_o(t)C_i(t)$$

Equation 5.13

where o and i refer to overall and internal motions respectively.

For isotropic overall motion $C_o(t)$ is given by equation Equation 5.1 with $A_c = \frac{1}{5}$. The internal correlation function can be defined as:

$$C_i(t) = S^2 + (1 - S^2)e^{-\frac{t}{\tau_i}}$$

Equation 5.14

where τ_i is the correlation time and S^2 is the squared order parameter of the internal motion . S^2 refers do spatial restriction of the motion ranging between 0 and 1. For S^2 approximately 1, internal motions of the bond vector are said to be restricted and relaxation is defined by global motion. On the other hand, if S^2 is approximately 0, the unrestricted internal motions describe the relaxation.

The squared order parameter allows a simple geometrical interpretation depending on particular motional model. The relationship between model-free parameters and internal motion can be represented as shown in Figure 5.9 with the bond vector, μ, diffusing in a cone with an angle, θ, defined by the diffusion tensor and the equilibrium orientation of the bond vector, which characterizes the angular amplitude of the internal motion [96].

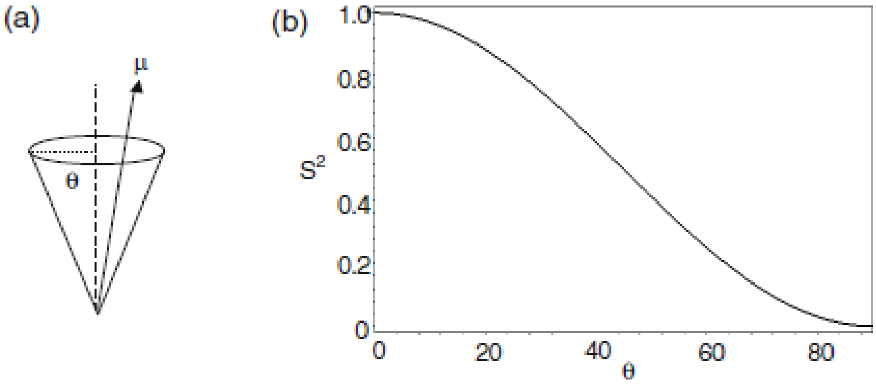


Figure 5.9. Relationship between internal motion and model-free parameters [96].

The quantity S^2 is given by the equation:

$$S^2 = \left[\frac{\cos\theta(1+\cos\theta)}{2}\right]^2$$

Equation 5.15

S^2 is a parameter that characterizes the angular amplitude of the internal motion, reaching the maximum value when θ is equal to zero and the motion of the vector is restricted to the fixed orientations. S^2 decreases rapidly as θ increases and the motion of the vector becomes more flexible. When θ is 75° or higher, the motion becomes completely isotropic, with an S^2 of almost zero.

Inserting Equation 5.1 and Equation 5.14 into Equation 5.14 yields:

$$C(t) = \frac{1}{5} e^{-\frac{t}{\tau_c}} \left[S^2 + (1 - S^2) e^{-\frac{t}{\tau_i}}\right]$$

Equation 5.16

with a Fourier transformation leading to the corresponding spectral density function:

$$J(\omega) = \frac{2}{5}\left[\frac{S^2\tau_c}{1+(\omega\tau_c)^2} + \frac{(1-S^2)\tau'}{1+(\omega\tau')^2}\right]$$

Equation 5.17

where

$$\tau'^{-1} = \tau_c^{-1} + \tau_i^{-1}$$

Equation 5.18

When the internal motion is slow compared to overall molecular tumbling $(\tau_i \gg \tau_c)$, then $\tau' \approx \tau_c$, and the spectral densityis given by $J(\omega)_{global}$. On the other hand, if the internal motion is faster than rotational correlation $(\tau_i \ll \tau_c)$, then $\tau' \approx \tau_i$ and the spectral density function is scaled by S^2: $J(\omega) = S^2 J(\omega)_{global}$.

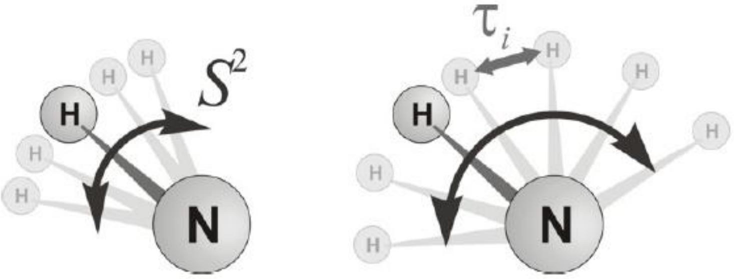


Figure 5.10. S^2 and τ_i illustration. S^2 describes the spatial restriction of the motion, in this case the motion of a ^{15}N-^{1}H bond vector. The time scale of the motion is given by τ_i. Left: highly restricted motions, $S^2 \rightarrow 1$. Right: largely unrestricted motion $S^2 \rightarrow 0$ [89].

In last case, C(t) rapidly decays to a plateau S^2 with a time constant τ_i due to internal motions. As time increases, global motions take over and C decays according to the overall correlation time τ_c(Figure 5.11)

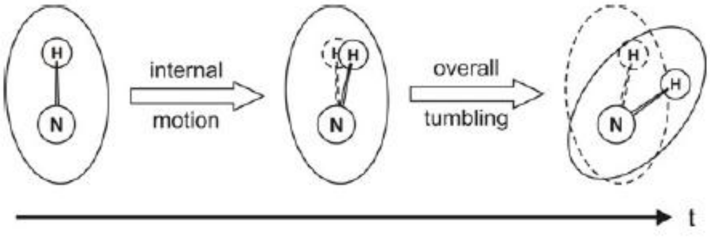


Figure 5.11. ^{15}N-^{1}H bond vector orientation according to internal motion and overall tumbling [89].

This approach was developed to include internal motions both on a fast and slow timescale. The ^{15}N-^{1}H bond vector reorients fast due to restricted internal motion and slow due to overall tumbling (Figure 5.12). The extended Lipari-Szabo formalism introduces an additional motion, and defines the correlation function of the internal motions as:

$$C_i(t) = C_f(t).C_s(t) = S^2 + \left(1 - S_f^2\right)e^{-\frac{t}{\tau_f}} + (S_f^2 - S^2)e^{-\frac{t}{\tau_s}}$$

Equation 5.19

where $S^2 = S_f^2 S_s^2$; S_s^2 and S_f^2 are the squared order parameters of the slow and fast internal motion, respectively; τ_s and τ_f are the corresponding correlation times.

A simple model for the extended Lipari-Szabo formalism is illustrated in Figure 5.12

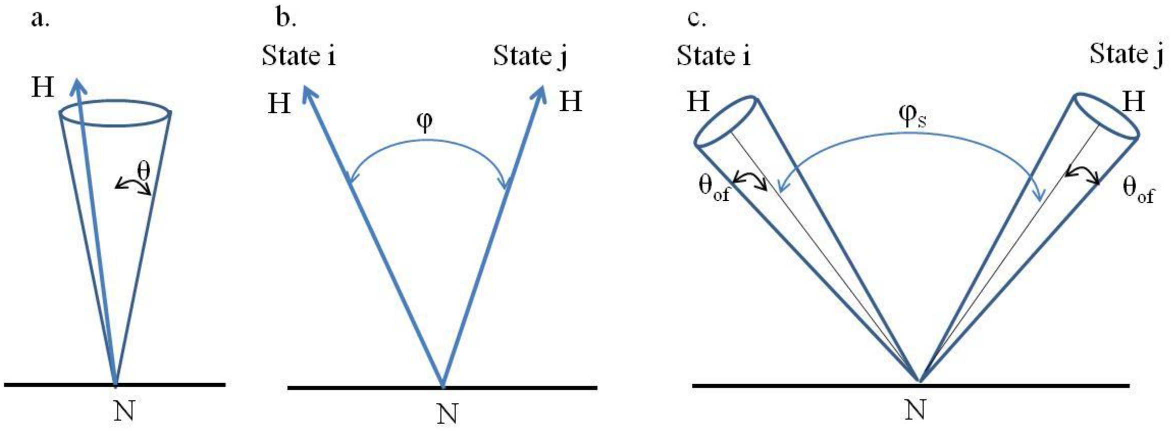


Figure 5.12. Specific motional models for interpretation of model-free order parameters. a. diffusion in a cone motional model (the N-H bond vector is assumed to diffuse freely within a cone defined by the semiangle θ); b. two-site jump model (the N-H bond vector is assume to alternate between two states i and j, separated by an angle φ);c. combined diffusion in a cone and two site jump models for internal bond vector motions (the N-H bond vector is assumed to alternate between two equally sized cones (on the slower timescale) or freely diffuse within each cone (on the faster time scale)). θf is the cone semiangle and φ is the semiangle between the two cones [97].

The slower motion is represented by a jump between two states (i and j) while faster motion is represented as free diffusion within two axially symmetric cones centered about the two I and j states. θ_{of} is the semi angle of the cone and φ_s is the angle between the NH vectors in the two states (i and j).

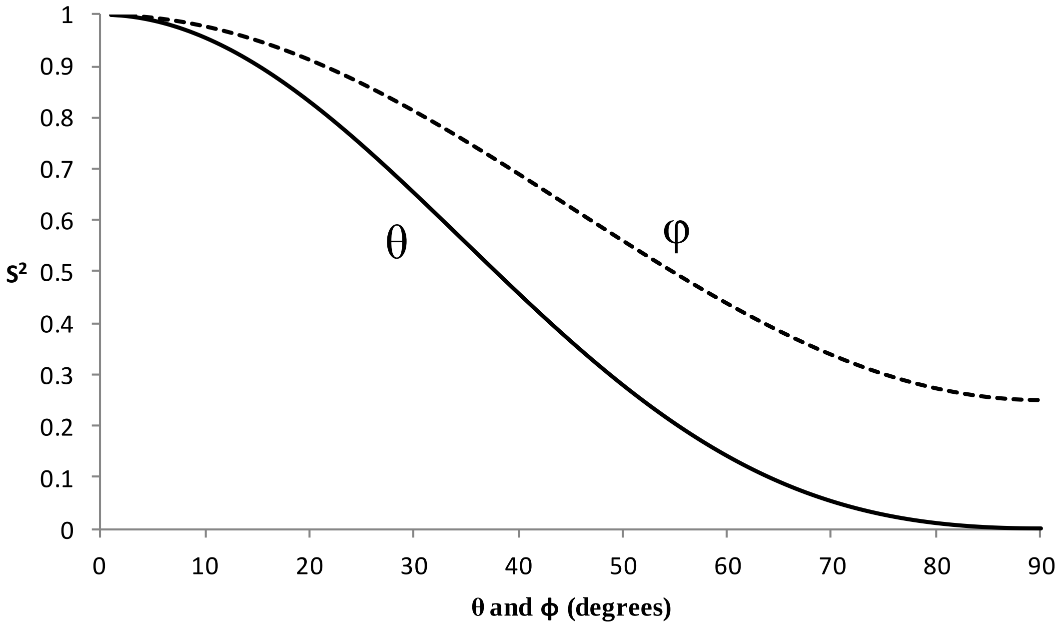


Figure 5.13. Relationships of the model-free order parameter (S^2) to the cone semiangle (θ) and the two-site jump angle (φ).

The full spectral density function of motions described by generalized order parameter, occurring on the ns-ps time scale is defined by:

$$J(\omega) = \frac{2}{5}\left[\frac{S^2\tau_c}{1+\omega^2\tau_c^2} + \frac{(1-S_f^2)\tau_f'}{1+\omega^2\tau_f'^2} + \frac{(S_f^2-S^2)\tau_s'}{1+\omega^2\tau_s'^2}\right]$$

Equation 5.20

where $\tau_k' = \tau_k\tau_c/(\tau_k+\tau_c)$ and $k = f\ or\ s$. If $S_f^2 = 1$, Equation 5.20 is reduced to Equation 5.17, which retrieves a reduced spectral density.

5.1.5 The Diffusion Tensor

Molecular tumbling in solution is an important tool for NMR relaxation. For a large number of proteins studied so far, an isotropic overall rotational diffusion was assumed as they proteins adopt approximately spherical globular shapes. However, it has been emphasized that anisotropic rotational diffusion has strong effects on spin relaxation and thus on the interpretation NMR relaxation data. Accordingly, it is important a detailed study of the rotational diffusion tensor for the analysis of intramolecular motions in non-spherical proteins. The rotational diffusion tensor characterizes how a molecule “behaves” in solution, as a sphere or something different. The tumbling is the same for all directions when the protein as a globular, spherical shape, which is defined as isotropic tumbling and hence isotropic diffusion tensor. In turn, if rotational diffusion is anisotropic, the molecular tumbling is described by three diffusion coefficients: if they all have different magnitudes it represents a completely anisotropic tensor; if two of them have similar size, it characterizes an axially symmetric diffusion tensor. In Figure 5.14 a schematic representation of rotational diffusion tensor is shown.

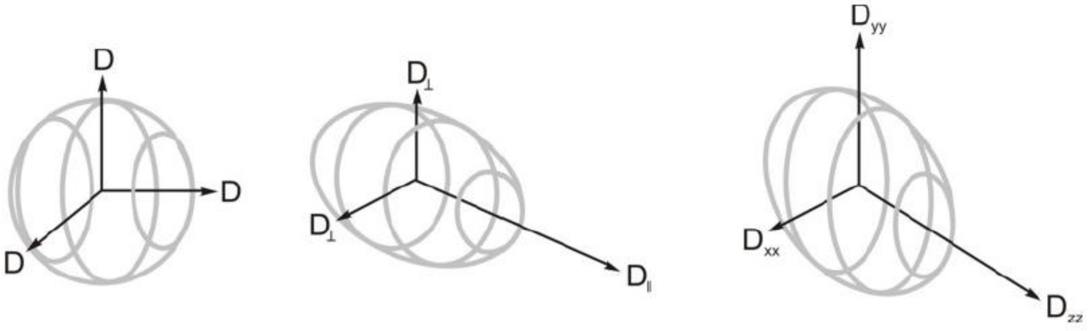


Figure 5.14. Illustration of rotational diffusion tensors. Left: the isotropic tumbling is represented for a globular spheric shape; middle: axially symmetric diffusion tensor; right: asymmetric rotational diffusion tensor [89].

Rotational diffusion tensor has to be estimated before relaxation data fitting to the Lipari-Szabo spectral density functions. This estimation has to be as accurate as possible since all relaxation rates during the fitting process depend on the diffusion tensor. Two methods were developed for determining the diffusion tensor: analysis of local diffusion coefficients using local correlation time or direct fitting of the R_2/R_1 ratios for ^{15}N-^{1}H bond vectors with highly restricted internal motions.

Direct fitting of the R_2/R_1 ratios for ^{15}N-^{1}H bond vectors with highly restricted internal motions is widely used to determine the diffusion tensor. Rotation around the long axis of the tensor is faster than rotation around a perpendicular axis. Therefore, transverse relaxation depends on the orientation of the bond vector in the diffusion frame. ^{15}N-^{1}H vectors aligned parallel to the long axis of the diffusion tensor are not reoriented by rotations around the axis and consequently are characterized by faster transverse relaxation. This statement is illustrated in Figure 5.15 , where an example of a protein with three helices, with axially symmetric rotational diffusion is shown. Helix B is aligned parallel to the long axis thus the bond vectors have a faster transverse relaxation when compared to helix A and C. In this case, rotational diffusion anisotropy is evident from the plot: all vectors oriented approximately parallel to the long axis of the tensor (helix B) have higher R_2/R_1 ratios due to a slower reorientation of their ^{15}N-^{1}H bond vectors.

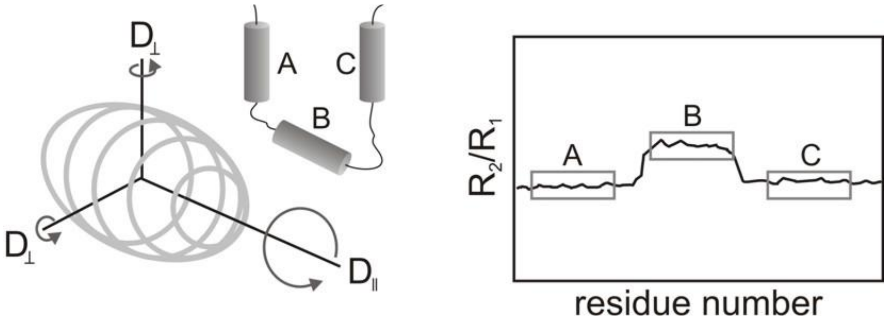


Figure 5.15. Estimation of the diffusion tensor anisotropy using the R_2/R_1 ratio. Residues with ^{15}N–^{1}H bond vectors oriented parallel to D∥ (helix B, Left) are readily identified in a plot of R_2/R_1 as a function of residue number (Right).

5.1.6 Model definitions

In order to obtain the motional parameters described in this chapter, the experimental NMR data have to be fitted against the equations defining the relaxation rates, with the appropriate forms of spectral density. In most cases, only three experimental parameters are available: the longitudinal and transverse relaxation rates and the heteronuclear NOE. The model free approach uses five different models to analyze ^{15}N relaxation data (Table 5.1)

Model 1 and 3: Model 1 only requires one parameter, the squared order parameter S^2, and it is the simplest model of all. In this model, the internal motions are assumed to be very fast, with the correlation time for the internal motion $(\tau_i \ll \tau_c)$. In the case of chemical exchange as an additional source of relaxation, R_{ex} is introduced as second fit parameter in model 3.

Model 2 and 4: Model 2 is also referred as “classical” Lipari-Szabo. In this model, τ_i is relaxation active and the spectral density function is defined by Equation 5.17. As for model 3, R_{ex} is introduced in the case of chemical exchange to characterize model 4.

Model 5: The extended Lipari-Szabo model describes internal motions that take place on two distinct time scales, τ_f and τ_s, which is an extended form of model free spectral density function. In this model, it is considered that the contribution of the fast motion can be discarded. Thus, while the fast motion contributes to the overall S^2, $S^2 = S_f^2 S_s^2$, the term containing the fast effective correlation time τ_f is omitted.

	Model	$J(\omega)$	Parameters
1	Simplified model-free (with isotropic tumbling)	$\frac{2}{5}\left[\frac{S^2\tau_c}{1+(\omega\tau_c)^2}\right]$	S^2
2	Original model-free (slow isotropic tumbling with faster, spatially restricted internal motions)	$\frac{2}{5}\left[\frac{S^2\tau_c}{1+(\omega\tau_c)^2}+\frac{(1-S^2)\tau'}{1+(\omega\tau')^2}\right]$	S^2 $\tau_i = \tau_f$
3	Like model 1 with conformational exchange term, *Rex*	$\frac{2}{5}\left[\frac{S^2\tau_c}{1+(\omega\tau_c)^2}\right]$	S^2 R$_{ex}$
4	Like model 2 with conformational exchange term, *Rex*	$\frac{2}{5}\left[\frac{S^2\tau_c}{1+(\omega\tau_c)^2}+\frac{(1-S^2)\tau'}{1+(\omega\tau')^2}\right]$	S^2 $\tau_i = \tau_f$ R$_{ex}$
5	Extended model-free (two time scales of internal motion with isotropic tumbling)	$\frac{2}{5}\left[\frac{S^2\tau_m}{1+(\omega\tau_m)^2}+\frac{(S_f^2-S^2)\tau_s}{1+(\omega\tau_s)^2}\right]$	S_f^2 S^2 $\tau_i = \tau_s$

Table 5.1. Different models used in a model free analysis of relaxation rates. [97]

5.2 Materials and methods

5.2.1 Sample preparation

For relaxation measurements, 250 µL of ^{15}N labelled human p22HBP 1 mM, were used in 50 mM phosphate buffer, pH 8.0, with 10 % D_2O. D_2O 99.8% was obtained from Eurisotop. Shigemi micro tubes (Shigemi Inc., Allison Park, PA) were used in all relaxation experiments. For relaxation measurements of human p22HBP with PPIX, 125 µL of ^{15}N labelled human p22HBP 1 mM, 10 % D_2O, pH 8.0 were mixed with 125 µL of PPIX 1.4 mM, 10 % D_2O, pH 8.0.

5.2.2 ^{15}N relaxation measurements

Longitudinal and transverse relaxation time (T_1 and T_2 respectively) and ^{15}N- {^1H} heteronuclear NOE values for native human p22HBP were recorded on a Bruker DRX500 and Bruker Avance III HD 700, at 303 K, equipped with triple resonance, TXI probe (500 and 700 MHz), and operating at 500.130 and 700.130 MHz, respectively for ^1H, and at 50.697 and 70.971 MHz, respectively, for ^{15}N. 3mm.

Heteronuclear NOE values were calculated as the ratio of peak intensities in spectra recorded with and without saturation. In ^1H-^{15}N HSQC-NOE without saturation, a total recycle delay, d1, of 10 seconds was used in place of the saturation delay to guarantee the same recycle delay between scans for both experiments. NOE errors were calculated from the uncertainties in the peak intensities measurements by the root mean square noise of each peak in both spectra.

The ^1H-^{15}N steady state NOE experiments were recorded with HSQCNOEF3GPSI pulse program from Bruker library, using Echo/Antiecho-TPPI gradient selection, with decoupling during acquisition. A relaxation delay of 10 seconds was used, with 32 transients in a matrix with 2048 data points in F2 and 128 in F1 with interleaved manner, NOE and NONOE. The interleaved spectra were separated by a Bruker standard macro *split.*

Backbone relaxation parameters, T_1 and T_2, were determined by acquiring pseudo-3D spectra consisting in a series of 2D heteronuclear ^1H-^{15}N-HSQC experiments where the relaxation period varied. For the ^{15}N longitudinal relaxation time (T_1), 10 time points were collected (0.8, 0.6, 0.4, 0.25, 0.025, 1.5, 3.0, 0.25, 0.6 and 0.025 seconds). The spectrum was acquired with 2048 points in ^{1}H dimension and 128 points in the ^{15}N dimension and 24 scans. The spectral width was 8012.820 Hz in the ^{1}H dimension and 2027.352 Hz in the ^{15}N dimension and the relaxation delay was 3s. For the ^{15}N transverse relaxation time (T_2) 9 time points were collected (0.017, 0.034, 0.051, 0.068, 0.085, 0.102, 0.119, 0.136 and 0.153 s), using the pulse program hsqct2etf3gpsi3d. The spectrum was acquired using the same conditions as for T_1.

For T_1 determination, the fit function

$$I(t) = I_0 . e^{-t/T_1}$$

Equation 5.21

was used and for T_2 determination

$$I(t) = I_0 . e^{-t/T_2}$$

Equation 5.22

where I(t) is a decay curve of y values (peak intensities), t is the x-variable of time, I_0 the amplitude at t=0 and T_2 are fitted. The start parameter for I_0 is the y-value at lowest time (automatically chosen by the software) and the start parameter for T_2 is introduced by the user.

All NMR data were processed using Bruker Topspin 3.2 software. For analyzing human p22HBP relaxation data it was used the FASTModelfree software (FMF). Fast Model Free reduces user interaction to a minimum, once every step is performed automatically: creation of like input files, model assignment. The analysis of relaxation data using FMF has at least three steps: initial estimation of the rotational correlation time or diffusion tensor, model selection and a final optimization.

5.2.3 Model Free Analysis

Robetta structures

For human p22HBP structure prediction it was used the Robetta server once a 3D structure was not available for this protein. The Robetta server (http://robetta.bakerlab.org) has automated tools for protein structure prediction and analysis. For structure prediction, sequences submitted to the server are analyzed into putative domains and structural models are generated using either comparative modeling or de novo structure prediction methods. If a confident match to a protein of known structure is found using BLAST, PSI-BLAST, FFAS03 or 3D-Jury, it has to be used as a template for comparative modeling.[98]

For human p22HBP, the murine p22HBP structure was used as template, pdb 2GOV, and the following sequence submitted:

1 MGHHHHHHLE LGMIKNSLFG SVETWPWQVL SKGDKEEVAY EERACEGGKF 50

51 ATVEVTDKPV DEALREAMPK VAKYAGGTND KGIGMGMTVP ISFAVFPNED 100

101 GSLQKKLKVW FRIPNQFQSD PPAPSDKSVK IEEREGITVY SMQFGGYAKE 150

151 ADYVAQATRL RAALEGTATY RGDIYFCTGY DPPMKPYGRR NEIWLLKT 198

Estimation of the Overall Diffusion Tensor

Prior to estimation the rotational correlation time, the R_2/R_1 ratio has to be calculated, and residues either subject to chemical exchange or undergoing fast internal motions have to be excluded. The input file which contains only accepted residues is called R_2R_1.input with the sequence number, values of R_2/R_1 and uncertainties as a list:

Residue	*R_2/R_1*	*uncertainty*
16	*13.317*	*2.141*
18	*12.598*	*1.847*
19	*11.564*	*1.644*

For diffusion tensor estimation, a pdb file of human p22HBP as to be available and the center of mass of the molecular structure should be translated to the coordinate origin using *pdbinertia* script, where the option *–rt* translate and rotate the coordinates of input pdb structure to an *output_pdb.*

```
pdbinertia –rt (input_pdb)(output_pdb)
```

For a complete distinction between axially symmetric and complete anisotropic tumbling, QUADRIC_CONFUSION script was used, where the local effective correlation times are calculated from the R_2/R_1 ratios using R2R1_TM.

The program request for the following information:

```
#R2/R1   Analysis:
#nucleus: 15N
#bond length (A): 1.02
#CSA (ppm): -160
#Fields for R1 and R2 (MHz): 500.13
#residue  tm        dtm
16        13.577          1.149
18        13.170          1.089
19        12.561          0.988
```

Once local effective correlation times are used, no initial guess for τ_c is needed. The control file for quadric diffusion has the following content:

```
0.1       1.4     40
1 'N'
tm500n.out
rob1_nt_rot.pdb
rob1_nt_ax_qdr500n.pdb
rob1_nt_an_qdr500n.pdb
```

Quadric_Confusion is invoked with the command

```
quadric_diffusion quadric.ctrl>quadric.log
```

The results for isotropic, axially symmetric and fully anisotropic model are indicated in the output file.

The FMF.config File

The FMF configuration starts with FMF.config file containing all relevant information and parameters:

```
tensor Axially Symmetric
cutoff 0.95
Fcutoff 0.80
optimize Yes
maxloop 10
almost1 20
S2cutoff 0.7
seed 1985
numsim 300
jobname hhbp
gamma -2.710
rNH 1.02
N15CSA -160
tm 12.12
tmMin 7.8
tmMax 15.5
tmGrid 5
tmConv 0.001
Dratio 1.22
DratioMin 0.5
DratioMax 1.4
DratioGrid 5
DratioConv 0.001
Theta 140
ThetaMin 0
ThetaMax 180
ThetaGrid 10
ThetaConv 0.005
Phi 40
PhiMin 0
PhiMax 180
PhiGrid 10
PhiConv 0.005
model1only No
mpdb /home/bjg/hhbp_fmf_r2r1_700_500/rob1_nt_ax_qdr700.pdb
file{0}{R1} /home/bjg/hhbp_fmf_r2r1_700_500/R1500_hhbp.txt
file{0}{R2} /home/bjg/hhbp_fmf_r2r1_700_500/R2500_hhbp.txt
file{0}{NOE} /home/bjg/hhbp_fmf_r2r1_700_500/NOE500_hhbp.txt
file{0}{field} 500
file{1}{R1} /home/bjg/hhbp_fmf_r2r1_700_500/R1700_hhbp.txt
file{1}{R2} /home/bjg/hhbp_fmf_r2r1_700_500/R2700_hhbp.txt
file{1}{NOE} /home/bjg/hhbp_fmf_r2r1_700_500/NOE700_hhbp.txt
file{1}{field} 700
```

The script starts with the definition of diffusion tensor. Possible options are isotropic and axial symmetric; the fully anisotropic model is not yet implemented in this software. The cutoff of 0.95 corresponds to a confidence limit for the χ^2 test. Any model is fitted to the relaxation data of a given spin acording to residual sum squared error. The meaning of the confidence region can be defined as “There is a 95% chance that the true parameters fall within this region around the fitted ones”. The confidence limit for F-test is defined as 0.80 and it is used for the comparison of models with a different number of parameters, like model 1 and 2 for example. almost1 20 defines a new cutoff for residual sum squared error. The lower limit of S^2 is defined by S^2 cutoff, and the optimization of the diffusion tensor is improved if residues with a larger degree of motional freedom are excluded from the analysis. The maxloop 10 defines the maximum number of runs and the analysis will be stoped if convergence of the diffusion tensor is achieved or maxloop runs have been performed. Seed 1985 and numsim 300 are the integer for the random number generator necessary for Monte-Carlo simulations and the number of Monte-Carlo simulations respectively. The constants for the spin pair for which data is analyzed are defined by gamma, rNH and ^{15}N CSA. The initial guess as well as upper and lower limits for correlation time (tm) is defined by tm, tmMin, tmMax respectively. tmGrid refers to the number of steps for grid searching and tmConv controls the convergence of Tc. Dratio is only used if tensor is ste to axially simetric. The script ends with names including the paths of the coordinate file, R_1, R_2 and hetNOE datafiles. A complete modelfree analysis usually takes 8-10 hours depending on numsim and the optimized values for the rotational diffusion tensor are extracted. The simulations were perperformed on a Linux system (Ubuntu 12.04) with a dual core processor.

The input files with R_1, R_2 and NOE results for the different magnetic fields are listed in appendix 9.5.

5.3 Results and discussion

5.3.1 human p22HBP

^{15}N relaxation parameters

Longitudinal (R_1) and transverse (R_2) relaxation rates as well as ^{1}H-^{15}N steady state NOE values were obtained for free human p22HBP at 303K. Detailed tables of all relaxation parameters are shown in appendix 9.5, 9.6 and 9.7.

In Figure 5.16 and Figure 5.17, NOE values plotted as a function of human p22HBP sequence residues, for both magnetic field used, 500 MHz and 700 MHz, are shown.

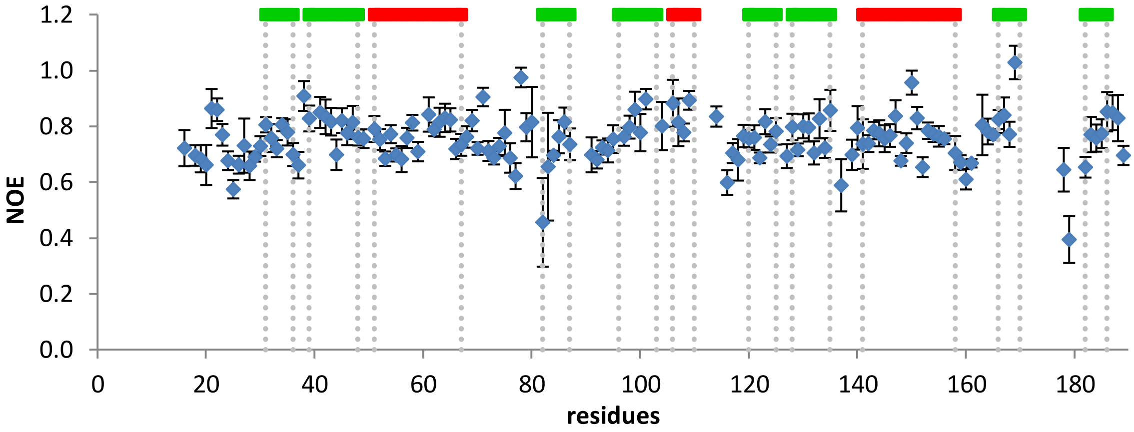


Figure 5.16. Heteronuclear ^{15}N-^{1}H NOE values plotted as a function of human p22HBP sequence at 500 MHz.

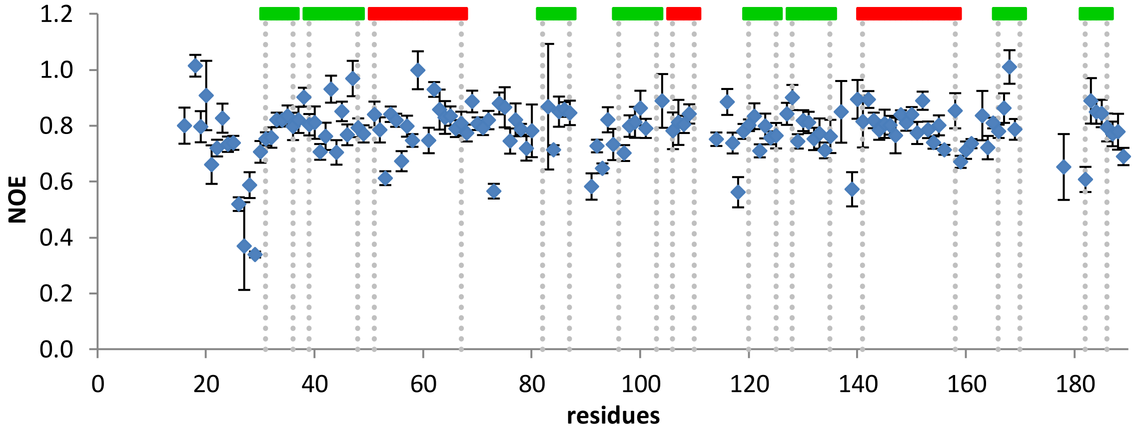


Figure 5.17. Heteronuclear ^{15}N-^{1}H NOE values plotted as a function of human p22HBP sequence at 700 MHz.

Analysing Figure 5.16 and Figure 5.17 it is possible to conclude that human p22HBP structure is well ordered, as indicated by an average value of 0.8 for the hetNOE. However, NOE values from the N-terminus were not available as this region of the protein has an absence of peaks. Proline residues have no NH signal in an HSQC spectra and thus do not retrieve any hetNOE value. Low intensities for the peaks of residues 40, 90, 96, 102, 110, 111, 157, 171, 172 and 180, 181 mean no hetNOE values could be determined. For residues I82 and G179 the low hetNOE values determined could be a consequence of the location of these residues (I82 is located near the binding pocket and could experiment some distortions in its position while G179 is located in the loop region with some flexible motion associated) (Figure 5.18).

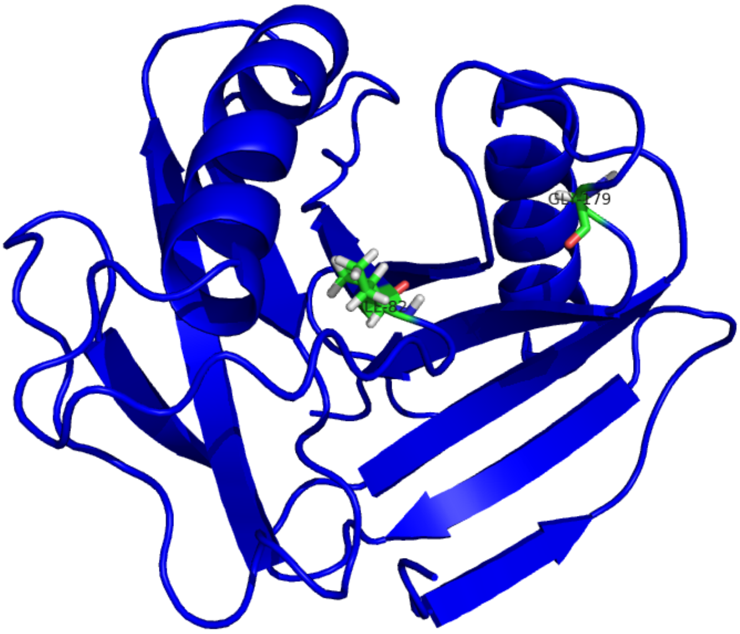


Figure 5.18. Residues I82 and G179 location in human p22HBP structure (Robetta model).

By analyzing the secondary structure present in murine p22HBP (and by inference in the human form), it is possible to conclude that α helix and β sheet NOE values are averaged at 0.775±0.049 (500 MHz) and 0.790±0.057 (700 MHz) while the loop regions average to 0.691±0.051 (500 MHz) and 0.774±0.123 (700 MHz). The lower values observed between secondary structures correspond to disordered regions of human p22HBP. NOE values from region flanked by residues 170-182 could not be measured but a NOE value of 0.644±0.078 for Tyrosine 178 and 0.394±0.084 for Glycine 179 at 500 MHz suggest that this region of the backbone is more dynamically disordered which is in agreement with flexible behaviour of the homologous backbone region of murine p22HBP reported by Dias *et al* and Micaelo *et al* [40] [22]. As reported by Tjandra (1995)[99], HetNOE lower than 0.65 are indicative of a considerable flexibility on a picosecond timescale. In Table 5.2, a detailed analysis of hetNOE values for each section of secondary structure in murine and human p22HBP is shown.

		β1	β2	α1	β3	β4	β5	β6	α2	β7	β8
p22HBP	human	0.761	0.799	0.763	0.687	0.819	0.756	0.777	0.766	0.866	0.761
	murine	0.837	0.844	0.821	0.775	0.793	0.816	0.832	0.805	0.811	0.829

Table 5.2. Average hetNOE values for human and murine p22HBP secondary structures. [p22HBP]= 1mM. at 500 MHz.

With the exception of β4 and β7, hetNOE values are higher in murine than human p22HBP. Overall, the total average hetNOE value is 0.82 for murine and 0.78 for human p22HBP.

R_1 and R_2 values were determined by fitting intensity data according to Equation 5.21 and Equation 5.22. In Figure 5.19 a series of spectra used to determine T_2 are shown, where it is possible to see the peaks decrease in intensity as the mixing time increases.

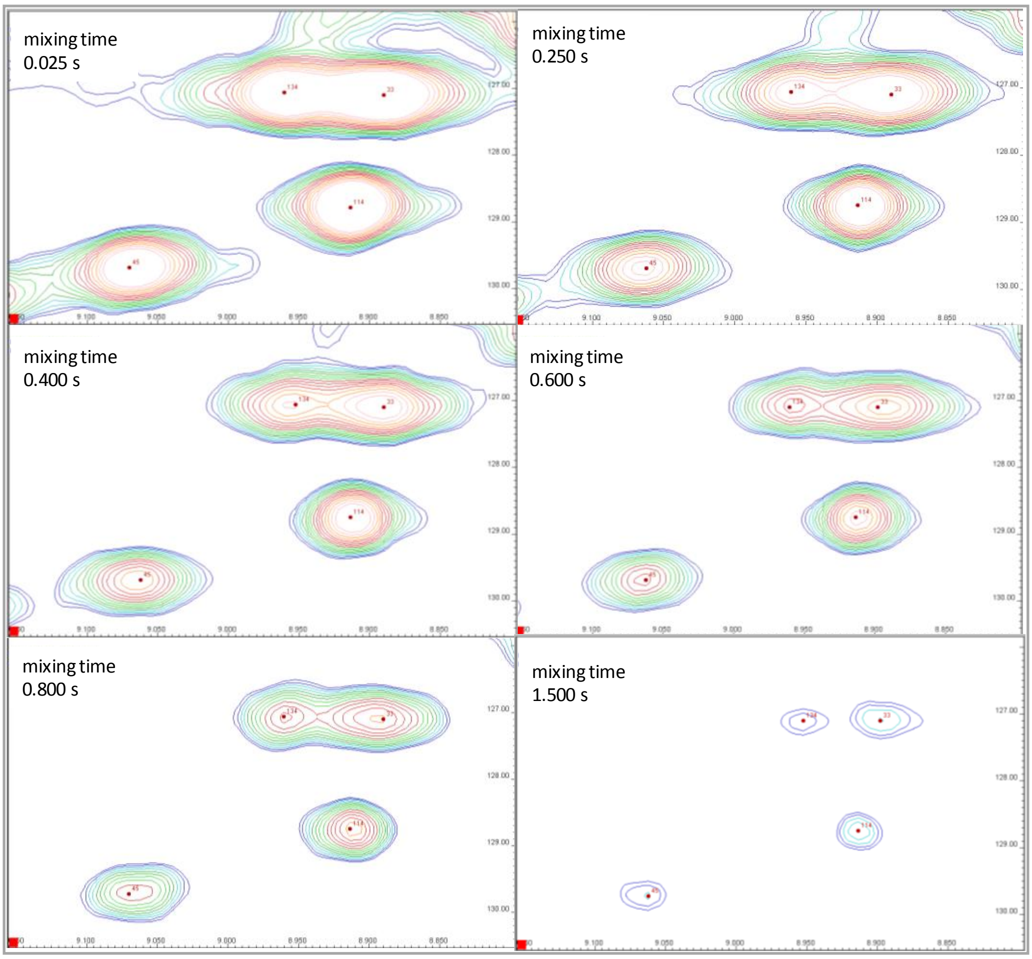


Figure 5.19. T_2 measurements of human p22HBP at 500 MHZ, 303 K.

In Figure 5.20, an example of T_1 fitting value for residue 114 of human p22HBP at 500 MHz is shown.

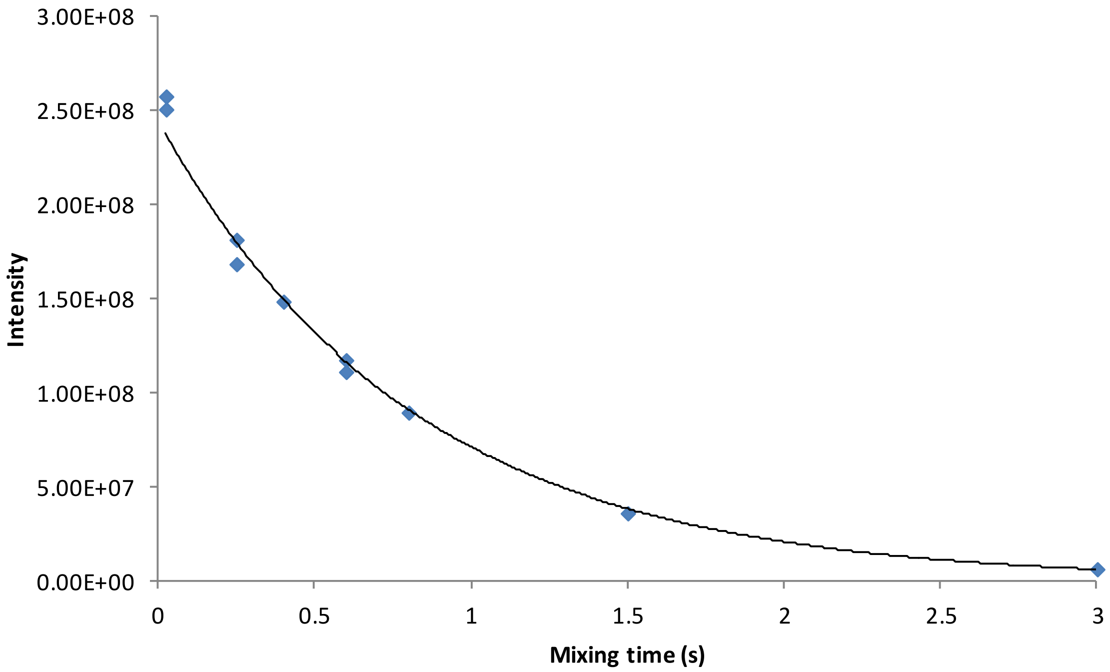


Figure 5.20. Fit of longitudinal relaxation time of residue 114 of human p22HBP against Equation 5.21 using peak intensities of T_1 measurements at 500 MHz, 303K.

In order to estimate the diffusion tensor, R_1 and R_2 values were calculated and R_2/R_1 ratios analyzed as a function of p22HBP sequence (Figure 5.21 and Figure 5.22). Estimation of the diffusion tensor anisotropy using R_2/R_1 ratios in human p22HBP indicates no evidence of anisotropy as all secondary structural elements have approximately the same R_2/R_1 ratio as shown in Figure 5.21 and Figure 5.22.

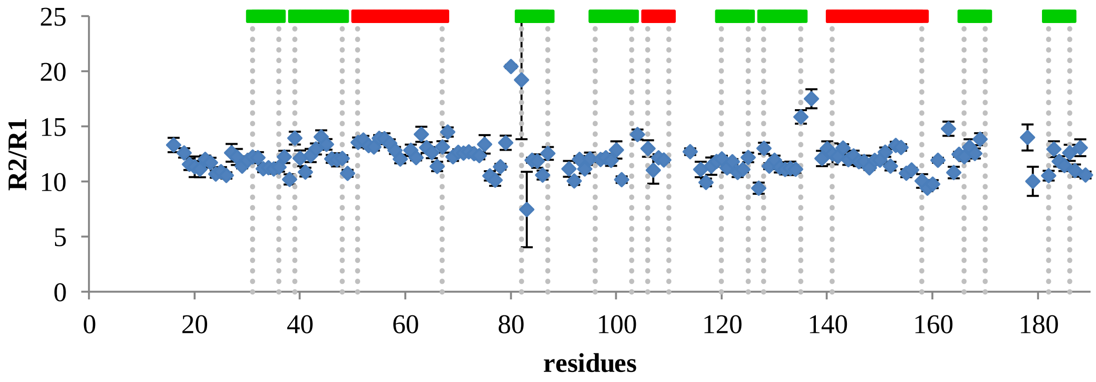


Figure 5.21. R_2/R_1 plot as a function of human p22HBP sequence, at 500 MHz, 303 K.

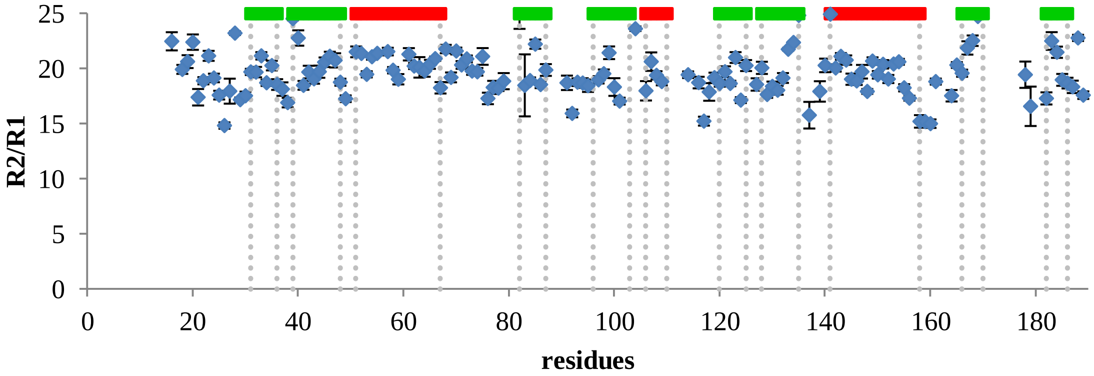


Figure 5.22. R_2/R_1 plot as a function of human p22HBP sequence, at 700 MHz, 303 K.

The R_2/R_1 ratio is higher for residues 80, 82, 135 and 137. For residue V80, this high ratio is a consequence of an high R_2 value (also confirmed at 700 MHz). This residue is located in a loop region and could have some anisotropy when compared with the global structure. Residue I82 has an high error associated and it was not possible to extract a conclusion about is behavior.

Comparison of these results at 500 and 700MHz are in agreement with the dependence of relaxation rates with correlation time and field strength shown in Figure 5.5. For higher fields, T_1 values are lower at lower fields and consequently have higher R_1 values. In contrast T_2 values are constant when considering different field strengths. Therefore the R_2/R_1 ratio is higher for 700 MHz (averaged 19.766±0.490) than for 500 MHz (averaged 12.163±0.528).

Robetta structures

As no structure exists for human p22HBP a three-dimensional structure was predicted using the programs Robetta and Modeller. As sequence homology is very high for human and murine p22HBP, the murine p22HBP structure was used as a template (pdb 2GOV). The structures obtained for human p22HBP using Modeller and Robetta are represented in Figure 5.23. Interestingly, the Robetta model has an α helix between residues L2 and F10 at the N-terminus. Each α helix is separated with a four-stranded β sheet as for murine p22HBP structure.

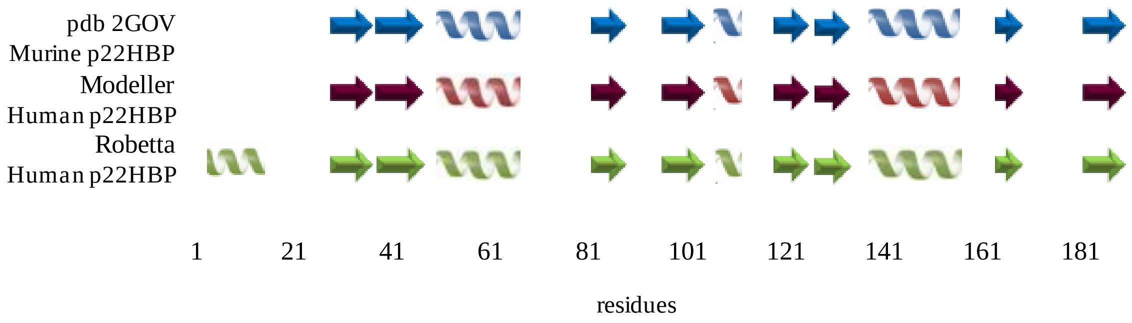


Figure 5.23.Modeller and Robetta secondary structure alignment for human p22HBP aligned with murine p22HBP (pdb 2GOV).

Diffusion tensor

Table 5.4 summarizes the total correlation time τ_m obtained from the diffusion tensor calculations. For human p22HBP, estimates of the rotational diffusion tensor resulted in a significantly small correlation time: τ_m=12.655±0.071 ns. Theta and Phi are the polar angles for the symmetry axis of the diffusion tensor in the coordinate frame of the PDB file. This estimation results in 145° for Theta (φ) angle and 126° for Phi (θ) angle. Axial isotropy diffusion tensor ratio (Dratio) of 0.881 indicates that the molecule behaves as a spherical rotor.

Diffusion parameters	**Fit value**	**Sim value**	**Sim error**
τ_m (ns)	12.655	12.681	0.074
Dratio	0.881	0.872	0.033
Theta (°)	145.515	144.24	13.689
Phi (°)	126.75	128.407	18.061

Table 5.3. Estimation of total correlation time τ_m of p2HBP. Fit value is the value of a parameter obtained by optimization of the input value, Sim value is the mean value of a parameter obtained from Monte Carlo simulations and Sim error the respective error.

The rotational correlation time of a protein is defined as the time that the molecule rotates through an angle of one radian, and is dependent on the size, shape, and dynamics of the molecule, as well as the bulk physical characteristics of the solvent. Thus, it is directly related to the volume and molecular weight of the protein. In Table 5.4 some correlation times for different proteins with different sizes are listed. The τ_m values increase as the protein size increases. In comparison, the value for human p22HBP is in agreement with its size.

Protein	Number of residues	Temperature (K)	τ_m (ns)
Human Ubiquitin [100]	76	303	4.10
Major Cold-Shock Protein (CspA) [101]	70	303	4.88
Rat microsomal cytochrome b5 [102]	98	298	5.00 ± 0.70
Calcium-loaded parvalbumin [103]	109	305	7.60
Photoactive Yellow protein [104]	121	-	6.40 ± 0.60
GMH4CO [105]	147	293	10.3
CDK inhibitor p19INK4d [106]	166	300	13.6 ± 1.10

Table 5.4. Correlation times examples for different proteins with different sizes.

Model Free results

Three parameters were obtained from a MF analysis of the ^{15}N relaxation measurements: S^2, R_{ex} and $\tau_{e.}$ and assigned to 5 different models according to modelfree formalism. The models that best fit the results are models 1 and 2, with the exception of residue 28 which was assigned to model 4. This residue is the unique in that has an R_{ex} term which describes residues affected by motions occurring on the μs-ms time-scale and accounts for chemical exchange processes that contribute to the decay of transverse magnetization. The glutamic acid present at residue 28 is located in a loop region and could experience some flexibility (as shown in Figure 5.24).

In appendix 9.9 a complete table with fast model free results for human p22HBP is shown. It should be noted that models 1-4 have S^2_f=1.0 and S^2_s=S^2.

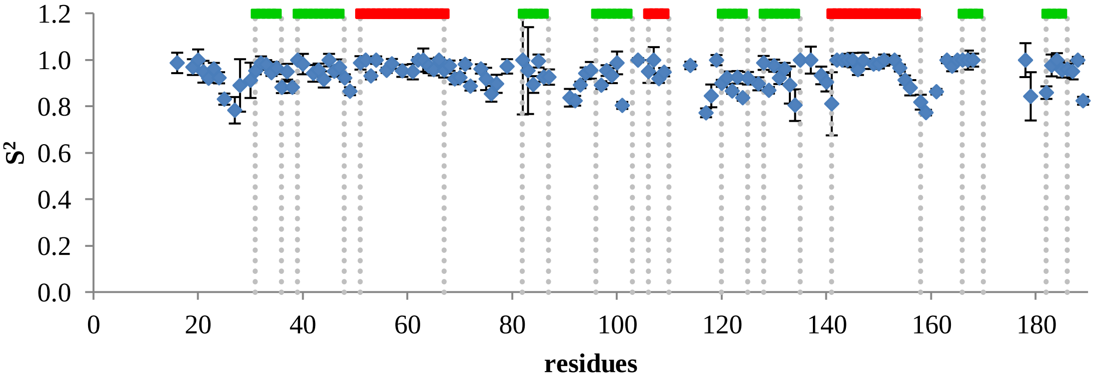


Figure 5.24. Generalized order parameter S^2 of human p22HBP at 303 K using data measured at 500 and 700 MHz.

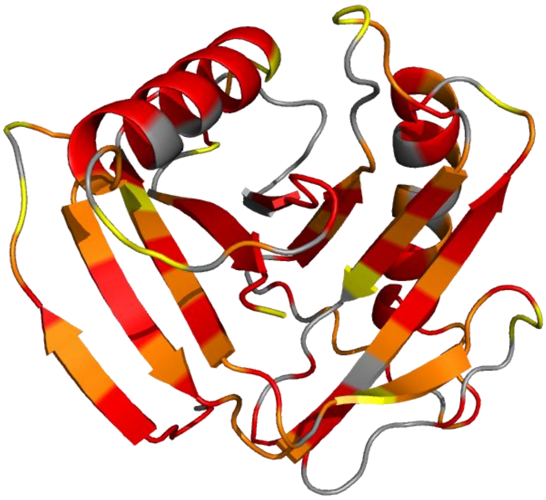


Figure 5.25. Human p22HBP structure according to S^2 values. ($S^2 > 0.95$, color red; $0.95 \geq S^2 > 0.85$, color orange; $0.85 \geq S^2 > 0.75$ color yellow; not assigned colored grey.

S^2 ranges between zero for isotropic internal motions to unity for completely restricted motion in a molecular reference frame and accounts for the degree of spatial restriction for a backbone amide ^{15}N-^{1}H bond vector on the ps-ns time-scale. Internal motions in the secondary structure elements of human p22HBP are highly restricted, with an average squared order parameter of 0.943±0.031 giving a picture of a largely rigid protein which is consistent with HetNOE values previously described. If free diffusion within a cone is assumed as a motional model, this value of S^2 corresponds to a semi-cone angle of approximately 15°.

Regions with higher internal mobility are found in the loops, especially between helix α and β sheet (residues 159-163); these residues were all assigned to model 2. As molecular modelling studies [107], indicated a change in mobility for a flexible region flanked by residues 171-180 when Hemin or PPIX binds this region should warrant particular attention. Unfortunately only residues Y178 and G179 could be assigned in this loop and Y178 shows S^2 value of 0.988 while G179 did not retrieve any value for S^2.

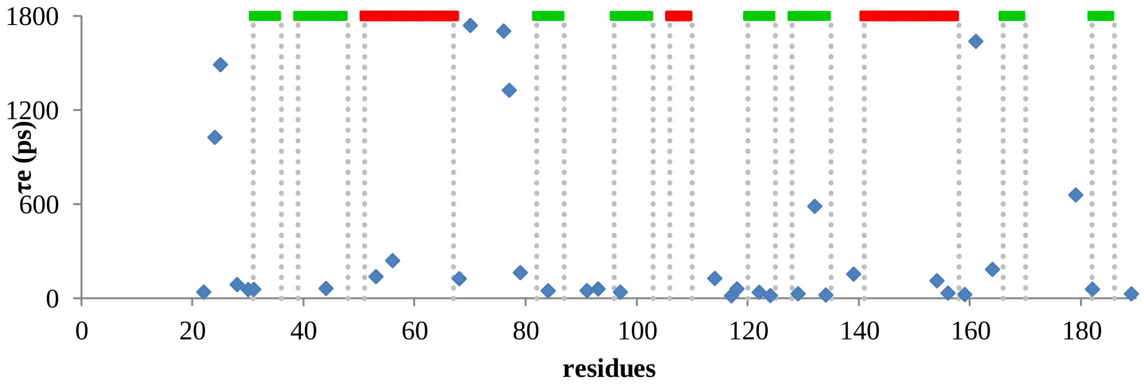


Figure 5.26. Effective correlation time (τ_e) of human p22HBP at 303 K using data measured at 500 and 700 MHz.

Considering the effective correlation time, τ_e, secondary structures have small τ_e values as expected. An exception is observed at residues 24, 25, 33 70, 77, 92 and 160, 161. These residues are located in loops and therefore should show some internal mobility.

5.3.2 PPIX-human p22HBP

^{15}N relaxation parameters

As for human p22HBP, hetNOE values were calculated for PPIX-human p22HBP (Figure 5.27 and Figure 5.28). As hetNOE values suggest, the protein remains well ordered and an averaged NOE value of 0.773 was obtained. Once again NOE values of N-terminus were not available due to the absence of peaks observed in this region. Residues 50, 60, 81, 88, 105, 112, 113, 115, 173, 174 and 177 are Prolines with no NH signal and thus no NOE value. K61, F84, G136, E141 and G163 have high uncertainties associated with their NOE values and those were not taken into account.

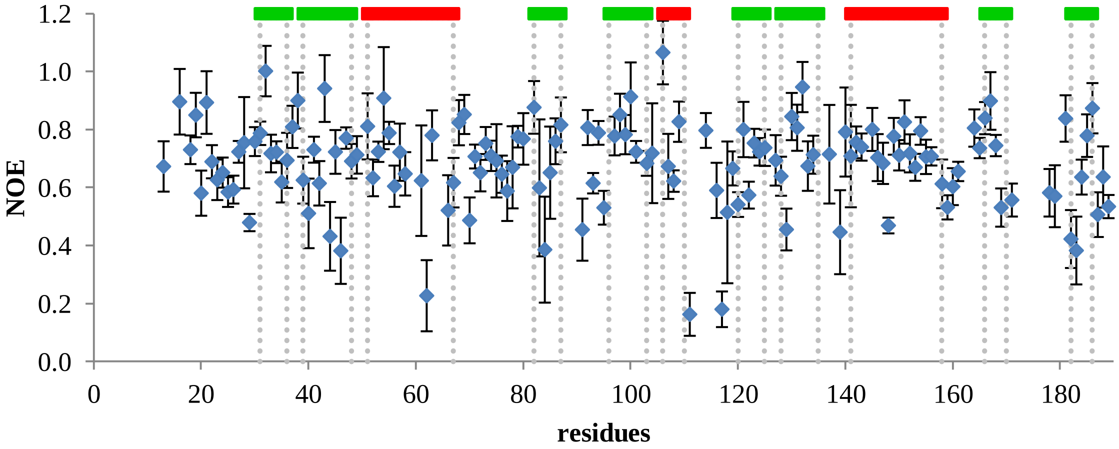


Figure 5.27. Heteronuclear ^{15}N-^{1}H NOE values plotted as a function of human p22HBP sequence, in the presence of PPIX (p22HBP:PPIX, 1:1.4), at 500 MHz.

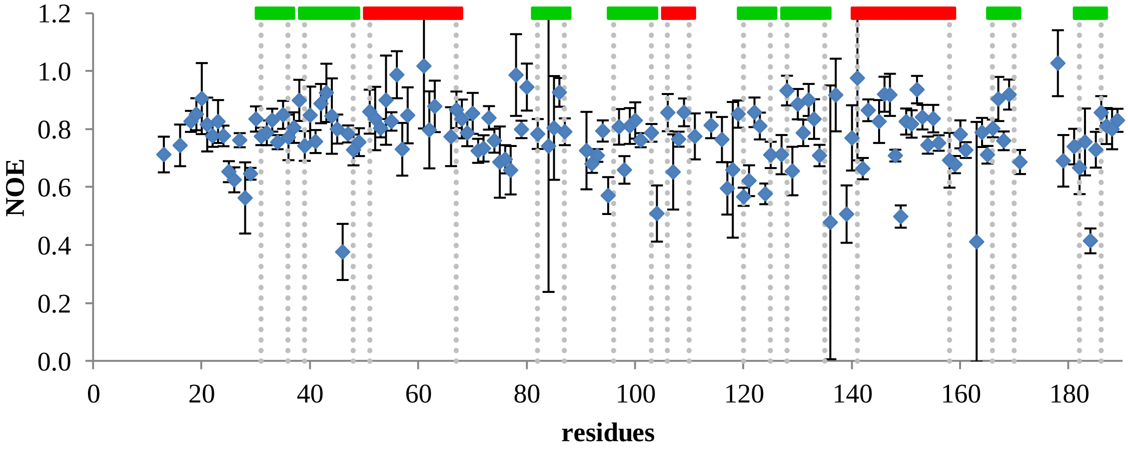


Figure 5.28. Heteronuclear ^{15}N-^{1}H NOE values plotted as a function of human p22HBP sequence, in the presence of PPIX, (p22HBP:PPIX, 1:1.4), at 700 MHz.

Analyzing Figure 5.27 and Figure 5.28, it is possible to conclude that these results have much more dispersion than those obtained for human p22HBP in the free form. Low signal to noise ratios were obtained for this experiments due to the presence of PPIX and consequent low concentration of p22HBP.

Considering secondary structures (Table 5.5), the main differences between human and murine p22HBP when titrated with PPIX were observed in the α1 helix and in the β5 and β8 sheets, although the same behaviour was expected for both proteins. The low signal to noise ratio obtained in human p22HBP spectra limit the correct analysis of the results. These experiments should be performed with high concentration of protein in order to increase the signal to noise ratio.

NOE		β1	β2	α1	β3	β4	β5	β6	α2	β7	β8
PPIX p22HBP	human	0.795	0.761	0.856	0.810	0.774	0.691	0.815	0.802	0.847	0.685
	murine	0.787	0.848	0.644	0.804	0.891	0.800	0.839	0.830	0.833	0.824

Table 5.5. ^{15}N-^{1}H NOE averaged values for human and murine p22HBP secondary structures in the presence of PPIX at 500 MHz.

On average the R_2/R_1 ratio is found to be 15.128±0.532. Figure 5.29 shows that the bound form of human p22HBP undergoes isotropic tumbling in solution due to the fact that no large variations are seen for the R_2/R_1 ratio throughout the structure.

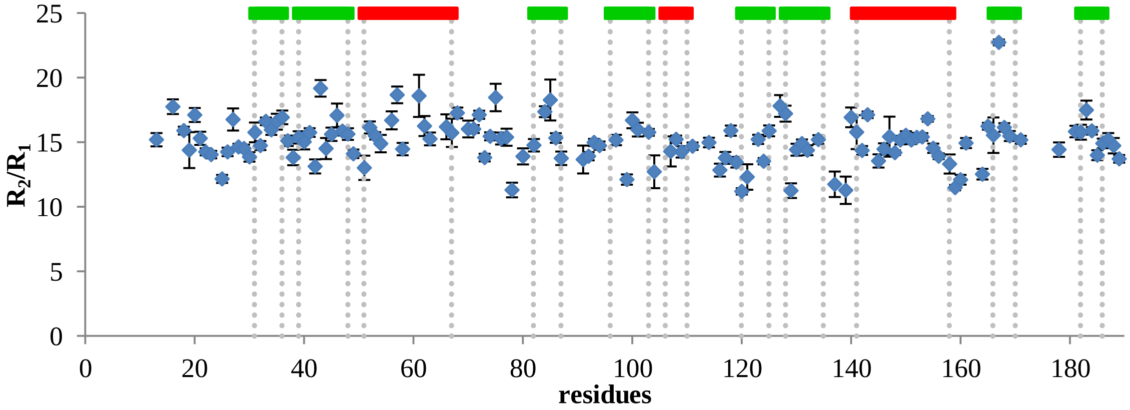


Figure 5.29. R_2/R_1 plot as a function of human p22HBP sequence, in the presence of PPIX, (p22HBP:PPIX, 1:1.4) at 700 MHz, 303 K.

Diffusion tensor

Diffusion tensor estimation was also performed, and a correlation time of 10.435 ns were determined. This value is lower than for free human p22HBP and it is a consequence of the presence of Tween. This limitation results in a significant reduction of theta angle which was determined as 2.5°.

Diffusion parameters	Fit value	Sim value	Sim error
τ_m (ns)	10.435	10.513	0.099
Dratio	1.092	1.111	0.081
Theta (°)	2.5	6.253	34.34
Phi (°)	119.372	115.515	51.183

Table 5.6. Estimation of total correlation time τ_m of p22HBP in the presence of PPIX. Fit value is the value of a parameter obtained by optimization of the input value, Sim value is the mean value of a parameter obtained from Monte Carlo simulations and Sim error the respective error.

Model Free results

In appendix 9.9, a complete table with fast model free results for human p22HBP-PPIX is shown. No large structural changes are apparent, as the S^2 values are similar to those seen for free p22HBP. An averaged S^2 of 0.897 ± 0.054 was obtained for the bound form. It should be noted that for model free analysis of human p22HBP with PPIX only relaxation parameters at 700 MHz were obtained.

Figure 5.30. PPIX-human p22HBP structure according to S^2 values. ($S^2 > 0.95$, color red; $0.95 \geq S^2 > 0.85$, color orange; $0.85 \geq S^2 > 0.75$ color yellow; not assigned colored grey.

Considering the effective correlation time, τ_e, higher values were observed between secondary structures mainly in residues K26, V46, G92, S93, I104, Q134, A139, A142, A148, I165, C168 and I184.

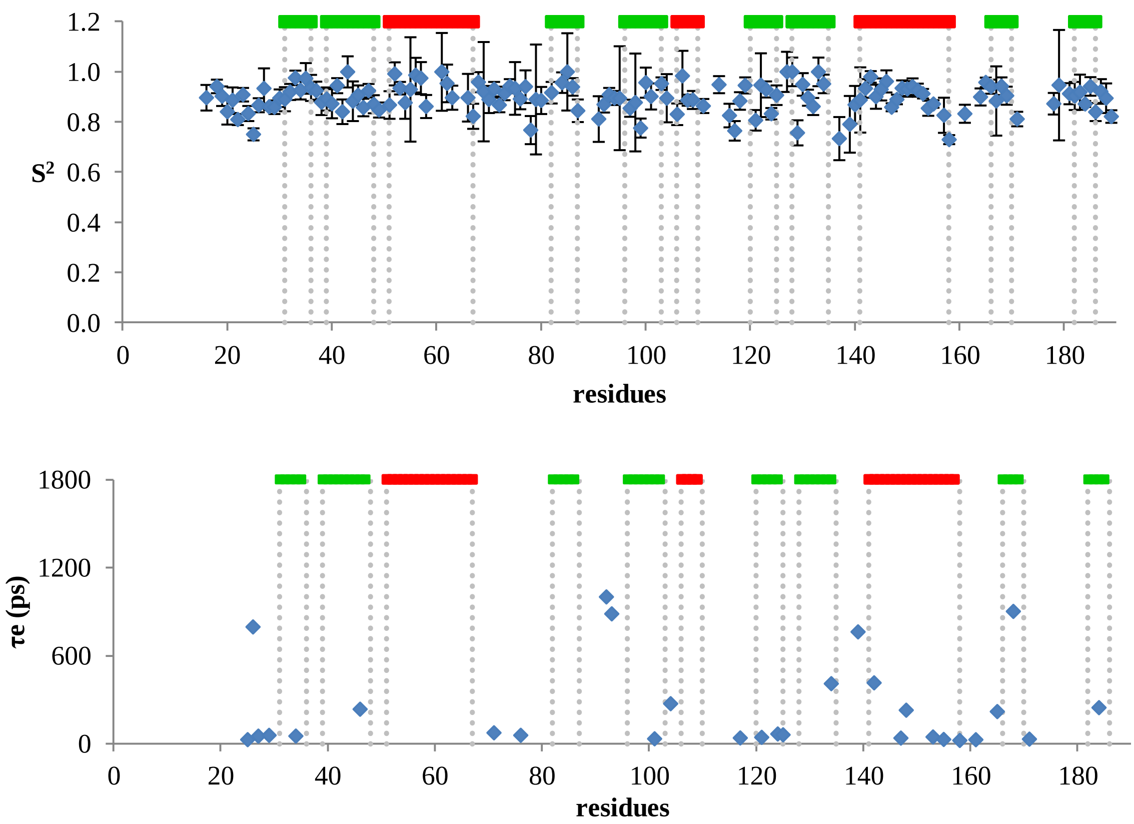


Figure 5.31. Model-free relaxation analysis of human p22HBP with PPIX (1:1.4) at 303 K using data measured at 700 MHz. Generalized order parameter S^2 (top) and effective correlation time (τ_e) (bottom).

Considering the global dynamic behaviour of p22HBP, it is possible to conclude that is a protein well structured, taking in account the NOE values determined for both murine and human proteins. Human p22HBP tumbles isotropically, as confirmed by R_2/R_1 ratios, and a correlation time of 12.655 ns was obtained. The similarity between the human and murine structures and hetNOE data allow us to conclude that murine p22HBP also tumbles isotropically in solution. When bound to PPIX, the diffusion tensor remains isotropic but a lower correlation time of 10.435 ns was obtained. This may be due to the presence of Tween (not present when relaxation experiments were carried out for free human p22HBP). As a surfactant, Tween limits the protein motion in solution and the theta angle was reduced from 145° to 2.5°. In order to confirm this influence from Tween dynamic relaxation studies should be carried out using human p22HBP with Tween and in absence of PPIX. High S^2 values of residues 178 and 179 observed in the absence and in the presence of PPIX, contradict previous molecular modelling studies of this region that identified some mobility as a consequence of binding or dissociation of tetrapyrrole compounds.

6 Chimeric Heme Binding Protein

6.1 Introduction

Despite the structural studies previously reported [11], [21], [40], [108][22], p22HBP function remains unknown. Towards the definition of the functional role(s) of p22HBP, a non functional version of the native protein to be used for protein knockdown by gene silencing studies was designed. Protein knockdown is achieved by siRNA experiments where mRNA of target protein is repressed and inhibits post-trascriptional gene expression [109].

Previous fluorescence quenching and chemical shift mapping studies (chapter 3 and 4) have shown that point mutations in arginine-56 (R56), lysine-64 (K64) and lysine-177 (K177) of murine p22HBP did not significantly affect the hemin-protein interactions. Based on these results, it was decided to build a new construct in which the hydrophobic patch of the heme-binding pocket in p22HBP, mainly located in α1-helix, would be replaced with the homologous hSOUL α1-helix. Human Soul protein (hSOUL), a 23 kDa protein, belongs to the SOUL/HBP heme-binding family of proteins. Of relevance; the hSOUL α1-helix does not contribute to heme binding. This new construct would be for a recombinant chimeric protein, since it would result from the fusion of structural elements from two different proteins. Recent studies involving titration of hSOUL with hemin, shown neither binding nor specific binding of heme to this protein [19]. These experimental data were obtained by chemical shift mapping (Figure 6.1) and UV-visible spectroscopy. In Figure 6.1 it is possible to see that there are no chemical shifts changes in human SOUL residues upon hemin binding suggesting an absence of interaction between them [23].

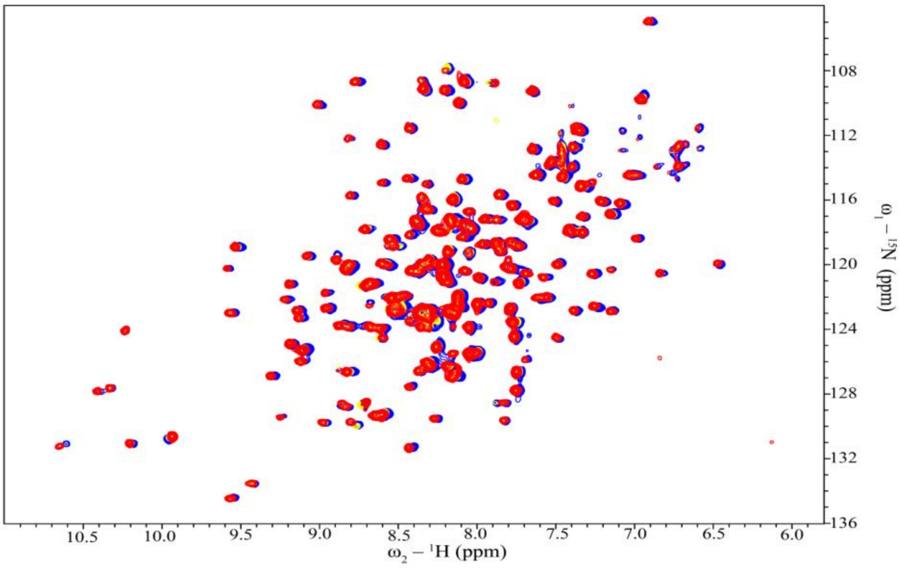


Figure 6.1. ^{1}H, ^{15}N HSQC spectra of hemin hSOUL. ^{15}N-hSOUL: hemin at molar ratio of 1:5 (red), 1:1 (yellow), and ^{15}N-hSOUL alone (blue).

In order to construct a prokaryotic expression vector for a chimeric protein from human p22HBP and human SOUL, an experimental plan was designed to replace the α1-helix involved in binding in human p22HBP with the corresponding helix present in human SOUL. First the phHBP1-28a plasmid used to overexpress human p22HBP was searched to identify any unique restriction sites flanking the human p22HBP α1-helix sequence. While there was a unique restriction enzyme (PstI) site upstream of the human p22HBP α1-helix, no restriction enzyme sites were present downstream of the sequence of interest. Thus, a unique, “downstream” restriction site had to be designed. Although the design was planned in order to introduce as few changes as possible, the final sequence, using the selected restriction enzyme, EcoRI, would lead to the replacement of T14 with phenylalanine. The hSOUL-encoding sequence flanked by PstI and EcoRI sites was achieved by PCR of overlapping oligonucleotides, such that the PCR product after the first round of PCR served as DNA template for subsequent PCR rounds.

The construction the chimeric p22HBP expression vector, in which the human p22HBP α1-helix-encoding sequence was replaced with the hSOUL α1-helix-encoding sequence, included three main stages:

-Stage I: To introduce a unique restriction enzyme site for subcloning of the hbp-soul-hbp subsequence into the pHBP1-28a expression vector (Figure 6.2).

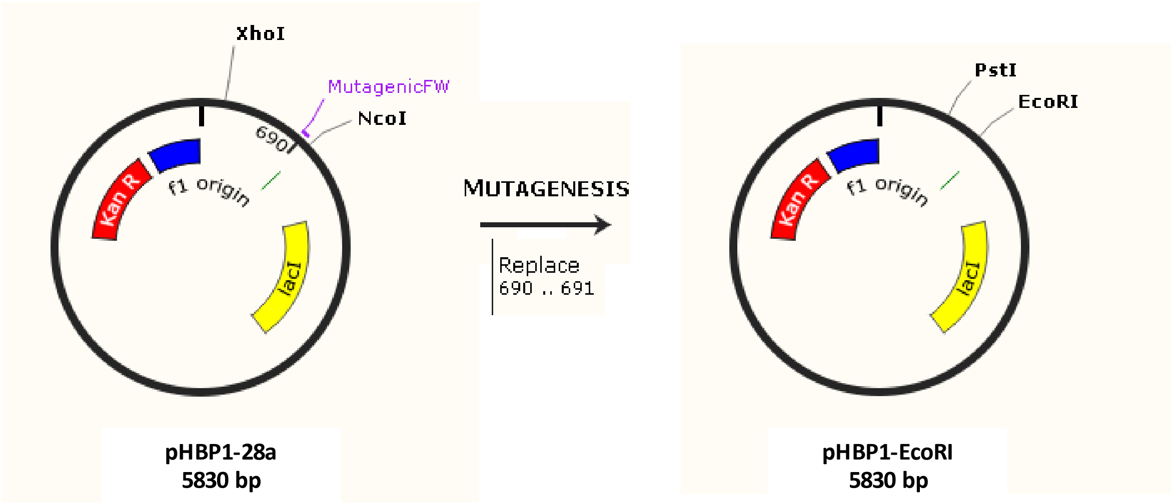


Figure 6.2. Schematic representation of unique RE site introduction. MutagenicFW is the mutagenic primer which contains the desired sequence for RE construction. 690..691 is the location of the RE site in pHBP1-28a.

-Stage II: To design overlapping oligonucleotides that expand the region flanked by the Pst-I and Eco-RI sites (hbp-soul-hbp subsequence) (Figure 6.3).

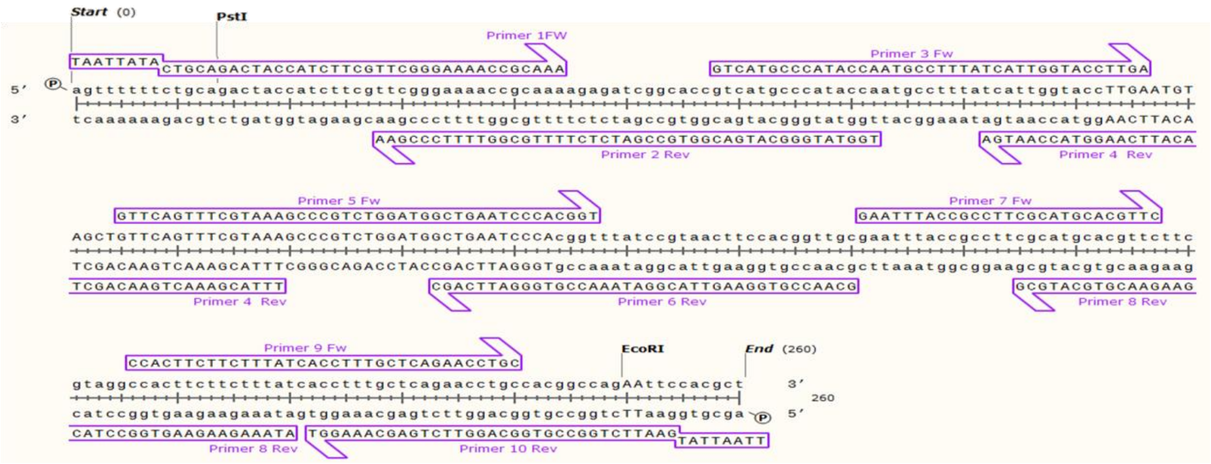


Figure 6.3. Overlapping nucleotides that expand the region flanked by Pst-I and Eco-RI sites.

-Stage 3: To digest the PCR product with PstI and EcoRI and subclone into the pHBP1-28a expression vector.

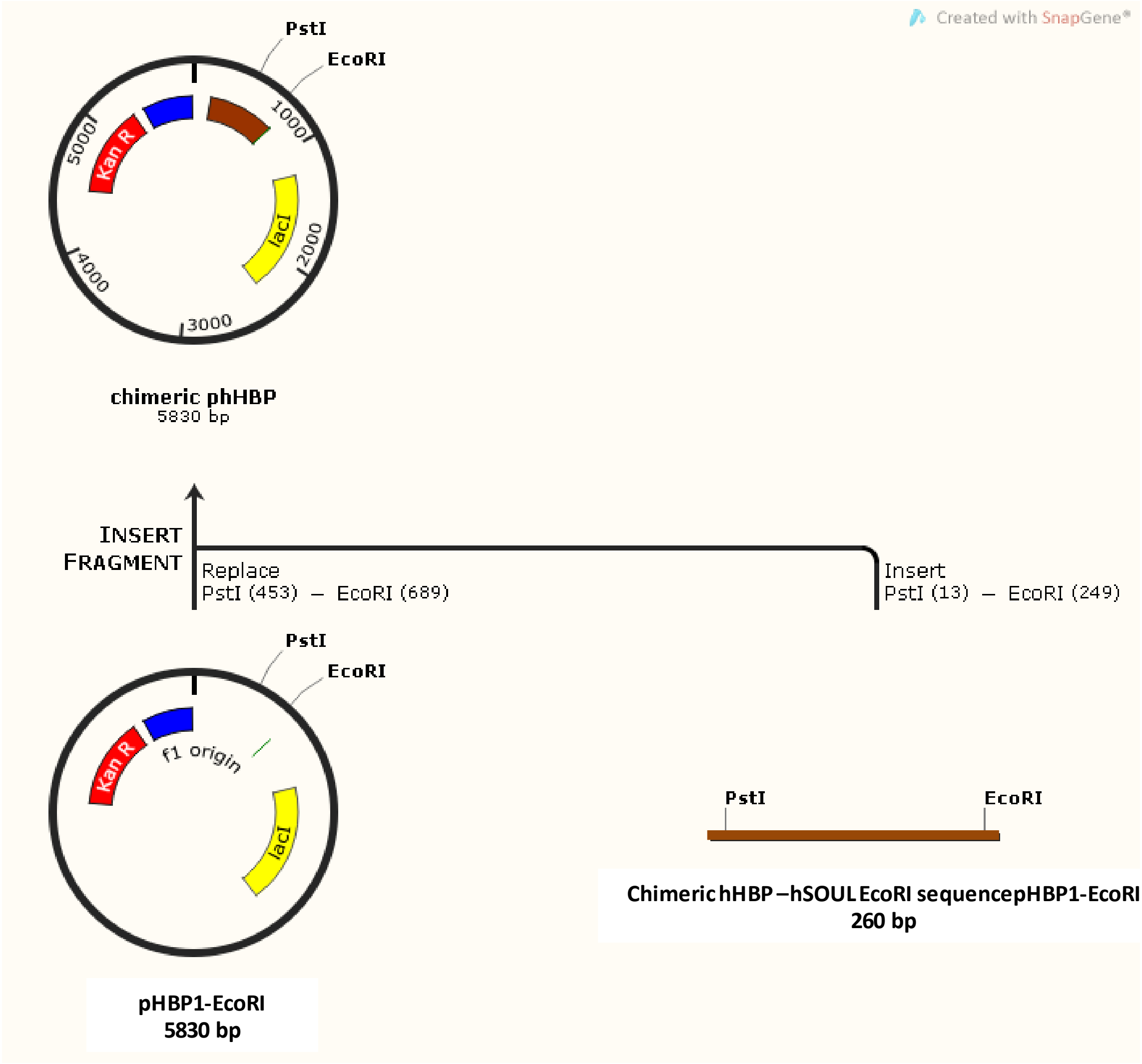


Figure 6.4. Schematic representation of subcloning into pHBP1-28a vector.

6.2 Material and methods

6.2.1 Introduction of a unique restriction enzyme site for subcloning of the "partial HBP+SOUL"-encoding sequence into pHBP1-28a.

Chimeric hHBP was designed based on the sequence of pHBP1-28a, an expression vector for human p22HBP (see chapter 2). After a careful analysis of the pHBP1-28a sequence with pDRAW32 and SnapGene software, with the aim of finding two unique restriction sites, flanking the human p22HBP α1 helix-encoding sequence, PstI and EcoRI were chosen as "the subcloning restriction enzymes". Since the pHBP1-28a vector has a unique PstI restriction site at nucleotide 451 but has no EcoRI site downstream of the p22HBP α1 helix-encoding sequence, a new EcoRI restriction site was created by replacing two base pairs at nucleotide 327 (Figure 6.5 and Figure 6.6).

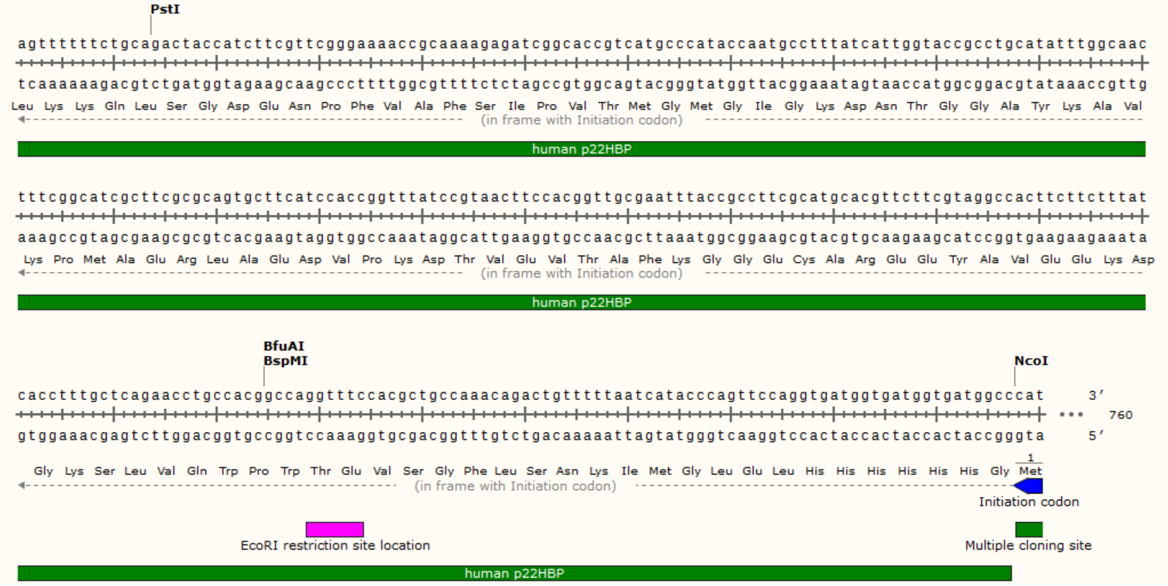


Figure 6.5 pHBP1-28a nucleotide sequence. (atg)Initiation codon. Location of EcoRI restriction site.

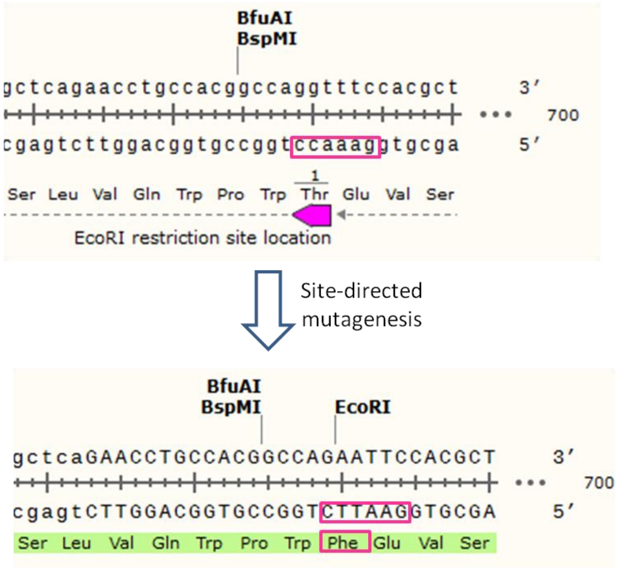


Figure 6.6. Schematic representation of site directed mutagenesis near nucleotide 327 with the replacement of Thr by Phe. Generated mutagenic primer sequence is highlighted with capitals.

The EcoRI site was introduced by simply mutating the sequence “gaaacc” on the 5’-3’ strand, to “gaaTTc”. Unfortunately, the codon for TTC is for a phenylalanine (F) and not Threonine (T). Given the sequence constraints for this expression vector and the need for a unique restriction enzyme this strategy had to be used to carry out subcloning of the hSOUL α1-helix.

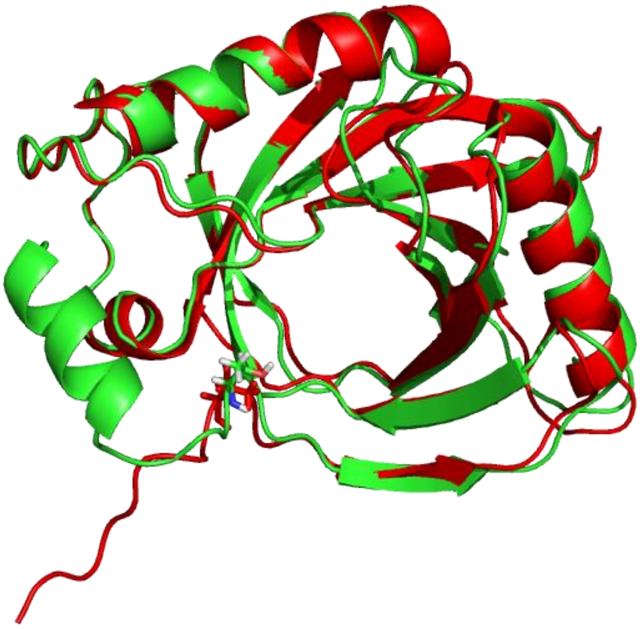


Figure 6.7.Threonine 15 location in human p22HBP structure (green-robetta modeler) and in murine p22HBP (red pdb 2GOV).

However, as the N-terminus in murine p22HBP is unstructured, the replacement of threonine by phenylalanine (an hydrophobic side chain in place of a polar uncharged side chain) should not produce any significant structural changes. To introduce the EcoRI restriction site mutagenic primers were designed using PrimerX software. The input parameters for the design of primers using pHBP1-28a plasmid sequence as a template are described in Table 6.1.

Protocol	Site-Directed Mutagenesis
GC content	40 to 60%
5' flanking region	11 to 21 bp
Terminates in G or C	Yes
Melting temperature	75 to 85°C
Length	25 to 45 bp
3' flanking region	11 to 21 bp
Mutation site at center	yes

Table 6.1: Primer X input parameters for primer design.

The Primer X program indicated 5 possible pairs of primers and the pair shown in table 6.2 were chosen.

primers	sequence
MutagenicFw:	5' GAACCTGCCACGGCCAGAATTCCACGCTGCCAAAC 3'
MutagenicRv:	5' GTTTGGCAGCGTGGAATTCTGGCCGTGGCAGGTTC 3'

Table 6.2: Mutagenic forward and reverse primers for EcoRI restriction site design.

A site-directed mutagenesis protocol provided with a kit from Nzytech (www.nzytech.com) was followed. 10 µL of DNA template [5ng/µL] (pHBP1-28a), 5 µL reaction buffer (10X) (Nzytech exclusive receipt), 1 µL dNTP mix, 1.2 µL Mutagenic FW [10 µM], 1.2 µL MutagenicRv [10 µM], were mixed in a PCR tube, and the volume was adjusted to 49 µL with ddH$_2$O. 1 µL of NZYDNAChange polymerase was added to the reaction and mixed gently. The PCR was performed using the following parameters: initial denaturation at 95 ºC for 2 minutes; 18 cycles at 95 ºC for 1 minute, 60 ºC 1 minute and 68 ºC for 9 minutes; followed by a final extension at 68 ºC for 15 minutes. The samples were then incubated with DpnI at 37 ºC for 1 hour. This restriction enzyme is specific for *dam*-methylated DNA fragments and allows the original template plasmid to be cleaved, while leaving the plasmid generated by PCR, and thus containing the EcoRI restriction site, intact. In addition, DpnI-treated PCRs were purified with NzyGelpure from Nzytech, to eliminate undesired DNA fragments resulting from PCR and DpnI digestion and to eliminate excess primer, ddNTPs and enzyme.

E. coli DH5α competent bacterial cells were used for transformation. Competent cells are *E. coli* cells that are especially treated to transform efficiently. There are two types of competent cells: chemically competent and electro-competent. Chemically competent cells are treated with a buffer that contains $CaCl_2$ and other salts that disrupt the cell membrane creating "holes" [110]. Heat shocking these cells opens the pores of cell membranes allowing an exogenous plasmid to pass into the cell [111]. Electro-competent cells are placed in an electroporation device that delivers a pulse of electricity to disrupt the cell membrane allowing a plasmid to enter the cell. In this work chemically competent cells, prepared in our laboratory, were used.

To prepare competent cells, 3 mL of LB medium was inoculated with 1 fresh colony of DH5α cells and incubated overnight at 37 ºC, with shaking. This bacterial culture was used to inoculate 250 mL of LB medium in an Erlenmeyer flask and incubated for 2-3 hours at 37 ºC with shaking until the culture reached an OD_{600nm} of 0.4. Then, the bacterial culture was transferred to a sterile centrifuge bottle, incubated on ice for 15 minutes and the cells were collected by centrifugation at 3000 rpm for 15 minutes. The supernatant was discarded, while the bacterial pellet was resuspended in 20 mL of sterile and ice-cold 0.1M $CaCl_2$ and incubated for 30 minutes on ice and transferred to 50 mL falcon tube. Based on the volume, 20% glycerol was added to resuspend the cells, which then were incubated on ice for 2 hours. Subsequently, the prepared competent cells were aliquoted (in 200 µL aliquots), quickly frozen in liquid nitrogen and stored at -80 ºC.

Competent *E. coli* DH5α cells were then transformed with plasmids obtained after *DpnI* treatment. A positive control with pHBP1-28a wild type plasmid was performed to check the transformation efficiency. The transformation was carried out by heat shock: 10 µL of each sample (plasmid with desired mutation) were mixed in 100 µL of DH5α competent cells. After 30 minutes at 4 ºC, the cells were heated to 42 ºC and held for 40 seconds and then incubated at 4ºC for 2 minutes. Transformed cells were resuspended in 900 µL of LB media, containing 20 % (m/v) of sterile glucose, and incubated at 37 ºC, 200 rpm for 1 hour. The cells were collected by centrifugation (5000 rpm for 1 minute). 900 µL of supernatant were removed and cells were then spread on LB plates with kanamycin at 50 mg/mL. Inverted plates were incubated for 12 hours, at 37 ºC and the presence of colonies indicated successful transformation. One colony was picked from LB plate, and used to inoculate LB media with Kanamycin 50 mg/mL for growth at 37ºC, overnight. The pHBP1-EcoRI plasmid was extracted from cells and purified with NzyMidiprep kit from Nzytech. To determine the concentration and purification level of the pHBP1-EcoRI plasmid sample, its absorbance was measured at 260 nm using a Nanodrop ND-1000 spectrometer. Purified pHBP1-EcoRI was finally sequenced using the T7 Fwd (5’ TAA TAC GAC TCA CTA TAG G3’ and T7 Rev (5’ GCT AGT TAT TGC TCA GCG G 3’) primers.

6.2.2 The design of overlapping oligonucleotides to expand the region flanked by the Pst I and Eco RI sites

The new pHBP1-EcoRI plasmid with unique PstI and EcoRI restriction sites now allows excision and subcloning of an insert containing the a1-helix of hSOUL. To construct this new insert, overlapping oligonucleotides covering the hSOUL α1 helix-encoding sequence flanked by the PstI and EcoRI sites were designed, such that this region could be expanded by annealing the oligonucleotides, extending them by DNA polymerization and amplifying the generated DNA fragment by sequential PCR [112]

```
agtttttt**ctgcag**actaccatcttcgttcgggaaaaccgcaaaagagatcggcaccgtcatgcccataccaatgcctttatcattg

TAATTATACTGCAGACTACCATCTTCGTTCGGGAAAACCGCAAA → primer 1fw
    ← oligo 2rev 3’AAGCCCTTTTGGCGTTTTCTCTAGCCGTGGCAGTACGGGTATGGT-5’
                                          3fw 5’GTCATGCCCATACCAATGCCTTTATCATTG
                                                    ← oligo 4rev 3’AGTAAC


gtaccTTGAATGTAGCTGTTCAGTTTCGTAAAGCCCGTCTGGATGGCTGAATCCCAcggtttatccgtaacttccacggttgcgaat

GTACCTTGA → oligo 3fw
CATGGAACTTACATCGACAAGTCAAAGCATTT-5’4rev
          5’GTTCAGTTTCGTAAAGCCCGTCTGGATGGCTGAATCCCACGGT-3’ → oligo 5fw
                 ← oligo 6rev   3’-CGACTTAGGGTGCCAAATAGGCATTGAAGGTGCCAACG-5’
                                                               5’- GAAT


ttaccgccttcgcatgcacgttcttcgtaggccacttcttctttatcacctttgctcagaacctgccacggccagAAttccacgct

TTACCGCCTTCGCATGCACGTTC-3’ → oligo 7fw
oligo 8rv3’-GCGTACGTGCAAGAAGCATCCGGTGAAGAAGAAATA-5’
←
                   5’-CCACTTCTTCTTTATCACCTTTGCTCAGAACCTGC-3’ → oligo 9fw
                ← oligo 10rev 3’   ACGAGTCTTGGACGGTGCCGGTCTTAAGTATTAATT
```

Figure 6.8. Chimeric hHBP encoding sequence flanked by PstI and EcoRI restriction sites. The α1-helix hSOUL-encoding sequence is highlighted in yellow. The original pHBP1-encoding sequence is highlighted in grey.

InFigure 6.8 the primers used are shown and in Figure 6.9 the strategy used to anneal and extend the insert region is outlined.

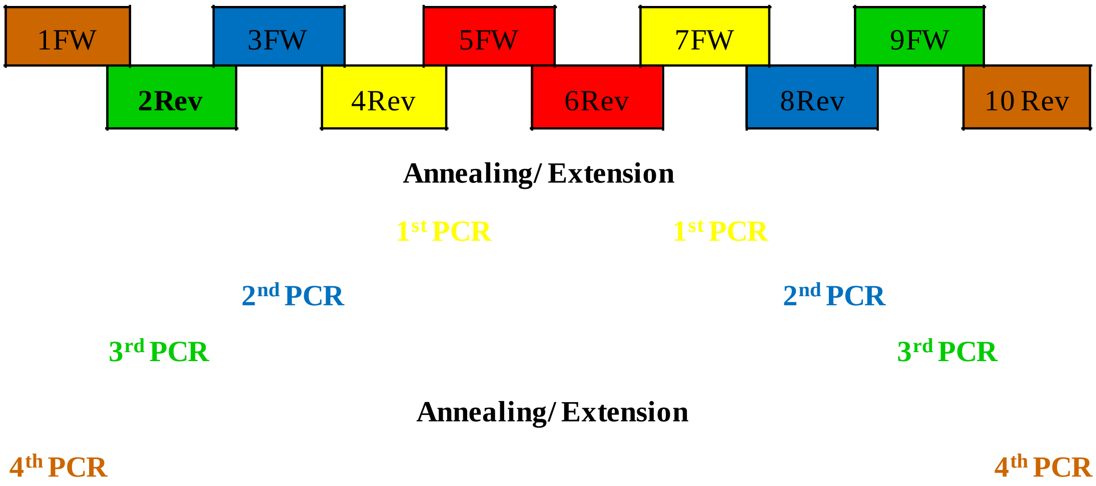


Figure 6.9. Sequential PCR strategy used to produce the hSOUL containing insert flanked by PstI and EcoRI. 1st PCR: primers 7FW and 4Rev; 2nd PCR: primers 3FW and 8Rev; 3rd PCR: primers 9FW and 2Rev; 4th PCR: primers 1FW and 10Rev.

Annealing and Extension

In a sterile PCR tube, 2 µL of each primer, 5FW [100 µM] and 6Rev [100 µM], were gently mixed. An initial denaturation step at 94ºC for 1 minute was applied followed by annealing at 60ºC for 2 minutes. This reaction was ended by slowly cooling down to room temperature at a rate of 5ºC for 5 minutes [112][113]. The hybrid DNA fragments were then extended with 2.5 Units of Klenow fragment of E.coli DNA polymerase I (Nzytech), in the presence of Klenow buffer reaction (2.5 µL of 10x reaction buffer Nzytech) and dNTPs (Nzytech) 1 µL of [2mM]. The reaction mixture was left at room temperature for 20 minutes. After this incubation, 5 µL EDTA [60 µM] was added to the reaction and heated at 75 ºC for 10 minutes, in order to end the extension reaction.[113]

1st PCR

The double-stranded DNA resulting from annealing and extension was used as a template and the primers 7Fw and 4Rev were used in this 1st PCR run. The DNA product obtained in this first PCR step was expected to have 116 bp. The reaction was prepared in a PCR tube, with 2 µL (sample A) and 4 µL (sample B) of previously generated DNA, 1 µL of each primer 7Fw [10 µM] and 4Rev [10 µM], 5 µL of Taq polymerase reaction buffer

[10x] Nzytech, 10 µL of ddNTPs [2 mM], 2 µL $MgCl_2$ [50mM], 1 µL Taq polymerase [5U] and the remaining volume with ddH_2O in a total volume of 50 µL. A total of 30, 1 minute cycles at 94ºC followed by 60 ºC for 1 minute, and 72 ºC for 4 minutes was followed by an extension of 10 min at 72 ºC. An initial denaturation step of was also carried out by running the PCR at 94 ºC for 5 minutes. Negative controls of sample A and B were prepared, with the same reactants but were left at 4ºC without PCR cycling. The amplified DNA was quantified using agarose gel electrophoresis. Most agarose gels are made between 0.7% and 2%. A 0.7% gel will show good separation (resolution) of large DNA fragments (5–10 kb) and a 2% gel will show good resolution for small fragments (0.2–1 kb). To prepare a 2% agarose gel, 2 g of agarose (Nzytech) was mixed with 100 mL of TBE (Tris/Borate/EDTA) buffer in a 250 mL conical flask (see appendix 9.1 for TBE receipt). To melt the agarose in the buffer, the flask was heated in a microwave until the agarose was completely molten. After the gel mixture cooled down to 60ºC, 2.5 µL of GreenSafe from Nzytech, a DNA stain, was added to the gel. The gel was slowly poured into an agarose gel casting tray and any bubbles were pushed away to the side using a disposable tip, followed by inserting a comb in the gel to produce the DNA-loading wells. The gels were left at room temperature until it solidification. Afterwards, the gel was submerged in the agarose gel casting tray with TBE buffer used as running buffer. For each run a total of 50 µL of each reaction were mixed with 10 µL NZYloading buffer from Nzytech and applied to the gel. Electrophoresis was performed at 100 V for 40 minutes.

2nd PCR

The reaction product from the 1st PCR was further amplified with another round of PCR using the overlapping primers 5'(3FW) and 3' (8REV) to generate a PCR product [112]. The PCR conditions were the same as for the 1st PCR with the exception of the DNA template concentration: 10 µL of sample B [12 ng/ µL] were used (reaction C) and 5 µL of sample B (reaction D). A negative control was also performed with the same reactants as reaction D and left at 4ºC. Once more, the amplified DNA was analyzed on a 2% agarose gel, and purified with NzyGelPure from Nzytech.

3rd PCR

To proceed with the 3rd PCR, 20 µL of DNA generated in 2nd PCR (sample E) and 10 µL (sample F) were mixed in different PCR tubes with the same reactants as for previous PCRs in a total reaction volume of 50 µL. In this case, the forward and reverse primers were 9FW and 2REV respectively. The DNA fragment generated in this step should have 213 bp.

Annealing and extension reactions

5 µL of generated DNA [10ng/µL] (sample C) were mixed with 2 µL of each primer 2Rev [10 ng/µL] and 9 FW [14 ng/µL]. Annealing and extension reactions were carried out as described above at the beginning of this strategy.

4th and final PCR

The last PCR was planned using 1Fw and 10Rev as forward and reverse primers, respectively, and 20 µL (sample G) or 10 µL (sample H) of extension reaction as DNA template. 1 µL of each primer [10 µM], 5 µL of Taq polymerase reaction buffer [10x], 10 µL of ddNTPs [2 mM], 2 µL $MgCl_2$ [50mM], 1 µL Taq Polymerase [5U] and ddH_2O up to a final reaction volume of 50 □L were mixed in a PCR tube. The PCR conditions were the same as described above. The generated fragment was expected to have 260 bp. A 2% agarose gel was used to analyse the DNA produced in this reaction.

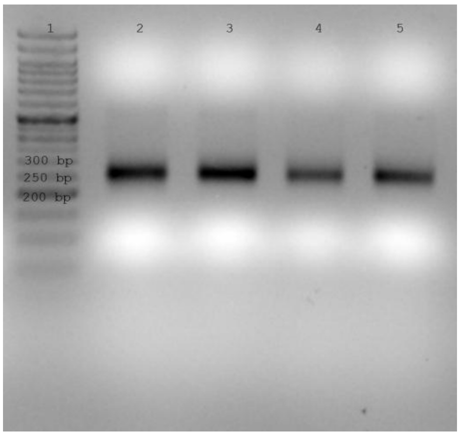


Figure 6.10. 2% agarose gel of annealing/extension followed by PCR reactions. Lane 1: Nzyladder VI; Lanes 2 and 3: 30 µL sample G; Lanes 4 and 5: 30 µL sample H..

Analyzing Figure 6.10, it can be seen that the PCR product has the expected fragment size (*i.e.*, 260 bp) for the DNA fragment encoding the hSOUL □..-helix flanked by the PstI and EcoRI sites. These bands were extracted from the agarose gel and purified using NzyGelPure from Nzytech. The purified DNA was sent for sequencing at Stabvida.

6.2.3 PstI/ EcoRI Double digestion

Once the presence of the sequence of the engineered hSOUL1 α1-helix and of the EcoRI site was confirmed in the mutated pHBP1-28a plasmid (pHBP1-EcoRI), the hSOUL1 α1-helix DNA fragment and pHBP1-EcoRI were digested with PstI and EcoRI. The purified products of these digestions, corresponding to the DNA fragments used in the ligation reaction to yield the expression vector for the chimeric protein, were named “insert” and “vector”, respectively. 4 γg of vector and 1 γg of the insert were independently mixed, *i.e.*, in different tubes, with 1 µL EcoRI (10 U/µL), 1 µL PstI (10 U/µL) and 5 µL of buffer O+ (Roche). These samples were incubated at 37 ºC for 2 hours and then the reaction products were analyzed in a 3 % agarose gel. 5 µL of the doubly digested products were loaded with 1 µL Tp 6x loading dye (Thermo Scientific) and 1 µL SyBr Gold (Thermo Scientific). The doubly digested vector was treated with 1 µL of calf intestinal alkaline phosphatase (Fermentas) [200 U] and incubated at 37 ºC for 1 hour in order to remove phosphates from the 5’ and 3’ends and ensure that only the insert would bind thus avoiding phHBP-EcoRI recircularization.

6.2.4 Vector purification after digestion

1 volume of biophenol:chloroform:isoamyl alcohol 25:24:1 was added to 50 µL of digested plasmid, mixed gently and centrifuged at 13000 rpm for 2 minutes. The aqueous phase (top) was collected and the organic phase discarded. 50 µL H_2O was added to the aqueous phase, mixed gently and centrifuged at 13000 rpm for 2 minutes. Once more the aqueous phase was collected to a total volume of 100 µL. Then 10 µL of 3M sodium acetate was added to increase the ionic strength along with 220 µL of absolute ethanol. This mixture was kept at -20ºC for 1 hour, centrifuged for 30 minutes at 16000 x g and the supernatant discarded. A final wash was performed with 100 µL of cold 70% ethanol and centrifuged for 15 minutes at 16000 g. The supernatant was removed and the tube left at 40 ºC to dry the DNA. The purified DNA pellet was resuspended in 10 µL of ultrapure H_2O.

6.2.5 Purification of digested DNA insert

The digested insert was purified with a Geneclean TurboKit. 210 µL of turbo salt solution (GENECLEAN) was added to 42 µL of insert, mixed and transferred to a Geneclean Turbo cartridge. The salt solution was removed by centrifuging at 14000 g for 5 seconds, with the DNA remaining bound to the cartridge. The cartridge resin was washed with 500 µL of "Geneclean turbo wash", and centrifuged at 14000 g for 5 seconds. Cartridge was dried with an extra centrifugation step of 4 minutes at 14000 g and placed in a new tube. The DNA was eluted from the cartridge resin by adding 30 µL of the Geneclean Turbo elution solution to the cartridge, leaving it for 5 minutes at room temperature and collecting the DNA upon centrifugation at 14000 g for 1 minute.

6.2.6 Ligase reactions and transformation of competent bacterial cells

Ligation reactions should be performed using a 1:3-10 molar ratio of vector: insert. Equation 6.1 is useful to calculate optimal amounts of insert DNA.

$$ng\ of\ insert = \frac{ng\ of\ vector\ x\ kb\ size\ of\ insert}{kb\ size\ of\ vector}\ molar\ ratio\,\frac{insert}{vector}$$

Equation 6.1

Ligation of the purified plasmid vector and insert was performed in order to obtain the new chimeric pHBP plasmid. Thus 1.5 µL of purified vector, 10 µL of purified insert, 1.2 µL

T4 DNA ligase buffer [10X] (Fermentas) and 1 µL T4 DNA Ligase (Fermentas) were mixed and incubated overnight at 22ºC. A negative control was performed using 10 µL of H_2O instead of purified insert. In a PCR tube, 6 µL of ligation product were gently mixed with 75 µL competent DH5α cells and incubated for 30 minutes at 4 ºC, then heat shocked for 40 seconds at 42 ºC and finally for 2 minutes at 4 ºC to transform the cells. The transformed cells were transferred to an Eppendorf tube, mixed with 950 µL of SOC media (see appendix 9.1 for SOC receipt) and incubated at 37 ºC for 1 hour at 800 rpm to avoid cell precipitation. The cells were then collected by centrifugation at 6000 rpm for 3 minutes. 750 µL of supernatant was discarded and remaining volume was resuspended and spread on LB plates with kanamycin at 50mg/mL. These plates were incubated inverted, overnight at 37 ºC. 4 colonies were randomly picked and independently used to inoculate 100 mL of TB (Terrific Broth) media (see appendix 9.1 for TB receipt) containing 50 □g/mL kanamycin. The cells harbouring the putative construct for the chimeric protein were grown at 37ºC, 150 rpm, overnight. Plasmid DNAs from the four overnight bacterial cultures were purified by anion exchange chromatography using a Macherey-Nagel Nucleobond Ax kit for quick purification of nucleic acids. Purified plasmids were sent for sequencing at Stabvida (www.stabvida.com) with T7FW and T7Rev as sequencing primers. Sequencing results confirmed the successful of chimeric human p22HBP construction.

6.2.7 XhoI/ NcoI subcloning of chimeric hHBP

In an attempt to prevent any possible additional mutations introducing during PCR, in the chimeric phHBP plasmid, XhoI and NcoI restriction enzymes were used for subcloning a region containing chimeric insert into wild type phHBP1-28a (see Figure 6.4) obtained by extraction and purification from DH5α cultures [114]. 5 µl of phHBP1-28a [1400 ng/µL] were mixed with 1 µL XhoI [10 U/µL] (Nzytech), 1 µL NcoI [10 U/µL] (Nzytech), 2 µL Nzytech buffer U [10x] and H_2O to a total volume of 20 µL. The same reaction was performed with 8 µL of chimeric phHBP [560 ng/µL]. These reactions were incubated for 12 hours at 37 ºC followed by analysis of the resulting DNA fragments by agarose gel (1%) electrophoresis. The digested phHBP1-28a was treated with alkaline phosphatase: 1 µL of alkaline phosphatase (Nzytech) [10 U/µL], 3.5 µL phosphatase buffer [10x] (Nzytech) were added to 30 µL of the doubly digested plasmid and incubated at 37ºC for 1 hour. Next, plasmid and insert were ligated with T4 DNA ligase, as described in section 6.2.6. Competent *E. coli* DH5α cells were transformed with the plasmid obtained from the ligation reaction. These reactions were performed under similar conditions to those previously described in section 6.2.6.

6.2.8 Chimeric hHBP overexpression and purification

Protocols for chimeric hHBP overexpression and purification were the same as for human p22HBP, previously described in section 2.4. Bacterial growth was performed in 2 L LB media with 2 mL Kanamycin 50 mg/mL at 37ºC and 180 rpm for 5 hours. Cells were harvested by centrifugation at 8000 rpm for 5 minutes and resuspended in M9 medium enriched with glucose and ammonium chloride. An adaptation to M9 medium was performed for 2 hours at 30ºC and 150 rpm. Induction of protein production was carried out by the addition of IPTG [0.5 mM], followed by overnight growth at 30 ºC. Protein purification was carried out as described in chapter 2, using affinity chromatography and SDS-PAGE gels for the analysis of protein purity.

6.2.9 Chimeric hHBP-W51V

After a detailed analysis of the protocols and also of the putative chimeric hHBP structure (robetta model) a possible structural problem with a tryptophan present at residue 51 could be prohibiting the correct folding of the chimeric protein. Figure 6.23 shows the position of W51 which is in close proximity to F84. It was therefore decided to construct a new chimeric mutant containing W51V to remove this putative steric clash.

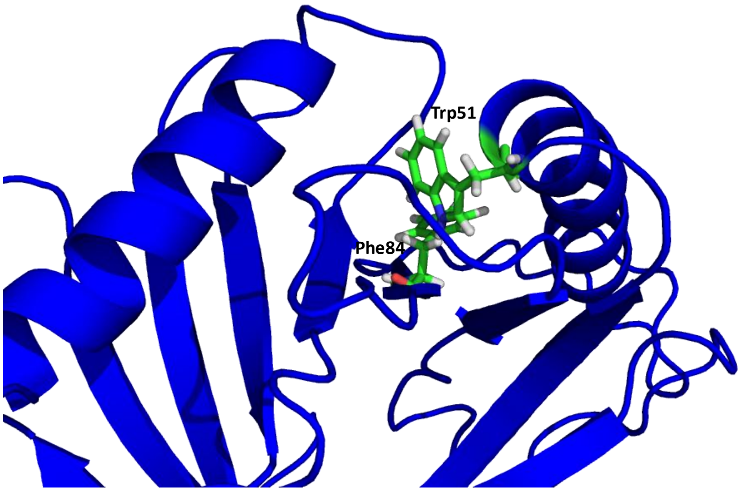


Figure 6.11: Representative structure of chimeric hHBP.

A new strategy was designed in order to replace tryptophan by a valine in position 51 of chimeric hHBP. The sequence CCGTTG at nucleotide 5250 was replaced by CCGGTG.

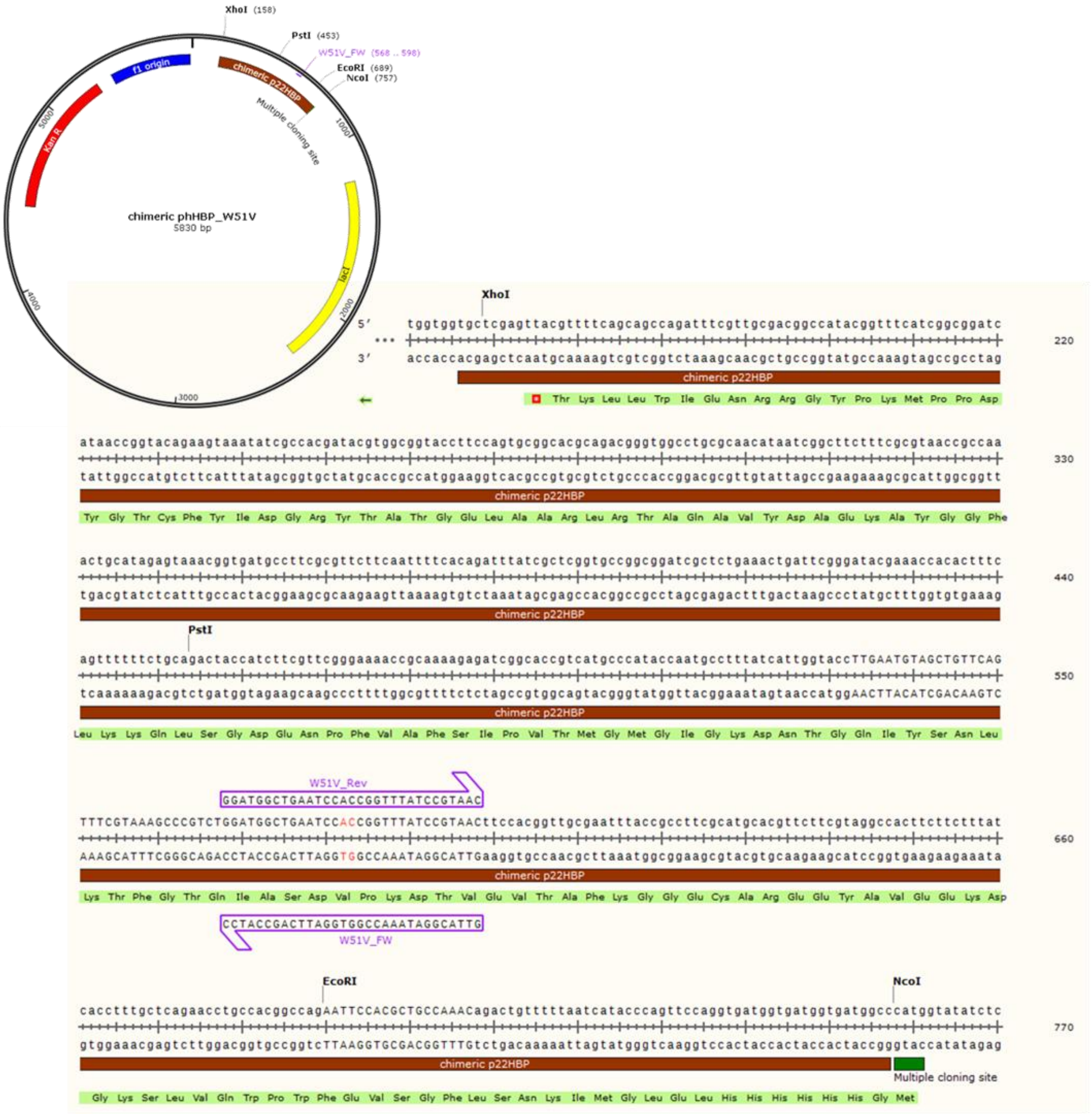


Figure 6.12. Chimeric pHBP1 plasmid and respective Forward (W51V_FW) and Reverse (W51V_Rev) primer annealing to replace tryptophan 51 with a valine.

To introduce this point mutation, Primer X software was used with the same parameters as described previously for EcoRI restriction site construction (section 6.2), and one pair of mutagenic primers was chosen:

primers	**sequence**
W51V-FW:	5' GTTACGGATAAACCGGTGGATTCAGCCATCC 3'
W51V-Rev::	5' GGATGGCTGAATCCACCGGTTTATCCGTAAC 3'

Table 6.3: Mutagenic Forward and Reverse primers for W51V site directed mutagenesis.

PCR was performed with 2 µL of chimeric phHBP [50 ng/µL], 1 µL of each primer W51V_Fw [10 µM] and W51V_Rev [10 µM], 5 µL of NZYDNAchange reaction buffer [10x], 10 µL of ddNTPs [2 mM], 2 µL $MgCl_2$ [50mM], 1 µL Taq Polymerase [5U] and ddH_2O up to a final volume of 50 □L. A total of 18 cycles of 95 ºC for 1 minute, 60 ºC for 1 minute, and 68 ºC for 9 minutes was followed by an extension of 15 min at 68 ºC. An initial denaturation step was also included by running the PCR at 95ºC for 2 minutes. To eliminate the chimeric phHBP plasmid without the W51V mutation, the PCR product was treated with DpnI. 100 µL of competent *E. coli* DH5α cells were transformed with 10 µL of this plasmid, by heat-shocking the bacterial cells for 40 seconds at 42º C preceded by an incubation of 30 minutes at 4 ºC. Cells were added to SOC media and incubated at 37ºC with shaking at 800 rpm for 1 hour. The transformed bacterial cells were harvested by centrifugation at 6000 rpm for 3 minutes and plated onto LB agar containing 50 □g/mL kanamycin. Overnight incubation at 37 ºC was performed in order to get new chimeric phHBP_W51V constructs. One colony was chosen and used to inoculate 100 mL LB of with 50 □g/mL kanamycin. Overnight cultures were purified using NZYMiniprep (Nzytech) and purified DNA was treated with NcoI and XhoI for subsequent subcloning. 1 µg of chimeric phHBP was mixed with 1 µL XhoI [10 U/µL] (Nzytech), 1 µL NcoI [10 U/µL] (Nzytech), 2 µL Nzytech buffer U [10x] and H_2O to a total volume of 20 µL. The same sort of reactions were performed with 8 µL of chimeric phHBP [560 ng/µL]. These reactions were incubated for 12 hours at 37 ºC followed by analysis of double digested products using a 1% agarose gel.

After double digestion, the region flanked by NcoI and XhoI in chimeric phHBP_w51v was extracted from the gel, purified and used to perform a ligase reaction. A 1:10, vector:insert molar ratio, and 1 µL of phHBP [50 ng/ µL] previously digested with NcoI and XhoI was mixed with 7 µL of digested insert [8 ng/ µL] (regarding Equation 6.1), 2 µL T4 DNA ligase buffer [10X] (Nzytech) and 1 µL T4 DNA Ligase (Nzytech). This reaction was incubated overnight at 22ºC, and DH5α competent cells were transformed with the generated plasmids according to previously described for heat shock transformation. Colonies were used to inoculate LB media enriched with kanamycin 50 mg/mL and overexpression was induce in M9 media as previously described. Purification was also carried out using Ni-NTA columns and fractions analyzed by SDS-Page.

6.3 Results and Discussion

6.3.1 Introduction of a unique restriction enzyme site for subcloning of the "partial HBP+SOUL"-encoding sequence into pHBP1-28a.

In order to evaluate whether the unique EcoRI restriction site had been introduced, LB plates with kanamycin were spread with transformed cells. Colonies were found on the LB plates indicating success and the purified plasmid resulting from overnight growth of randomly picked colonies were purified and analyzed by a Nanodrop ND-1000. The DNA concentration can be calculated from its absorbance at 260 nm where an absorbance of 1 (1 cm path length) is equivalent to 50 μg DNA/mL. The plasmid purity can also be evaluated by UV spectroscopy. A ratio of A_{260}/A_{280} between 1.80–1.90 and A_{260}/A_{230} around 2.0 indicates "pure plasmid DNA". An A_{260}/A_{230} ratio above 2.0 is a sign for too much RNA in the DNA preparation, an A_{260}/A_{280} ratio below 1.8 indicates protein contamination. The ratio A_{260}/A_{280} of 1.86 and A_{260}/A_{230} of 1.67 indicated successful plasmid purification, and gives a concentration of 122 ng/µL. Finally, sequencing results indicated that an EcoRI restriction site (GAATTC) had been created in the pHBP1-28a vector (sequence results not shown).

6.3.2 To design overlapping oligonucleotides that expand the region flanked by the Pst I and Eco RI sites

1st PCR

The size of the PCR product was expected to be 116 bp, and the DNA bands on the agarose gel were around 100 bp (Figure 6.13). Bands on lanes 4 and 5, which correspond to a larger sample volume, were extracted and the PCR product was purified.

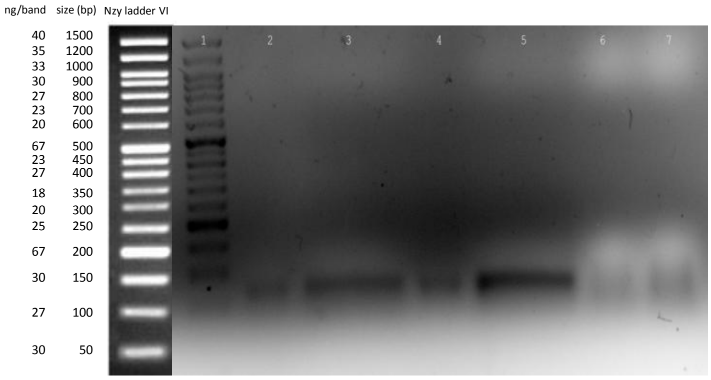


Figure 6.13. 2% agarose gel of 1st PCR reaction. Lane 1: 5 µL Nzyladder VI; Lane 2: 5 µL sample A; Lane 3: 55 µL sample A; Lane 4: 5 µL sample B; Lane 5: 55 µL sample B; Lane 6 and 7: A and B negative controls.

To estimate the amount of DNA present in each band, the amount of DNA in each fragment as visualized as a band from the molecular weight standard Ladder as reference was used. Similarity intensity for bands with the same DNA volume corresponds to approximately the same amount of DNA. Analysing the band in lane 4, Figure 6.13, and comparing its intensity with the ladder, it is possible to estimate that band corresponds to 30 ng. Considering the total volume of sample eluted (60 µL), 360 ng of DNA had been obtained. This total amount of DNA was extracted from the gel, purified (NZYGelPure Nzytech) and eluted from the resin with 30 µL of ddH2O, which resulted in a concentration of 12 ng/µL.

2nd PCR

The expected size for the 2nd PCR product was 163 bp. The results illustrated in Figure 6.17 (lanes 2- 5) are in agreement and the estimated concentration of the purified 2nd PCR product was 12 ng/ µL.

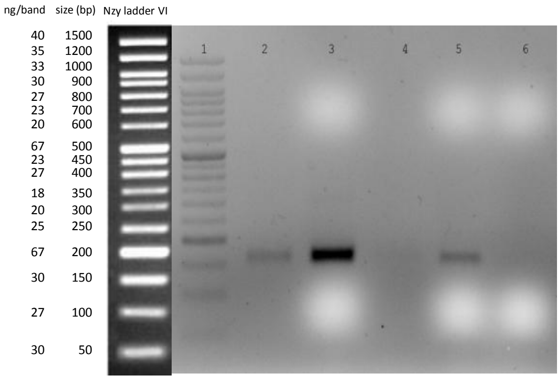


Figure 6.14. 2% agarose gel of 2nd PCR reaction. Lane 1: Nzyladder VI; Lane 2: 5 µL sample C; Lane 3: 55 µL sample C; Lane 4: 5 µL sample D; Lane 5: 55 µL sample D; Lane 6: negative control.

3^{rd} PCR

Analysis of the 3rd PCR products on a 2% agarose gel indicated that the expected 213 bp fragment was not generated (Fig. 6.17). Instead, DNA fragments of different sizes (i.e. 120 bp and 160 bp) were produced during the 3rd PCR, possibly suggesting non-specific binding of the primers to the DNA template. To overcome this problem, an annealing and extension reaction using 2nd PCR product as template and oligonucleotides 2Rev and 9Fw was performed. The results from this approach revealed that a single 260 bp DNA fragment was generated (Figure 6.16).

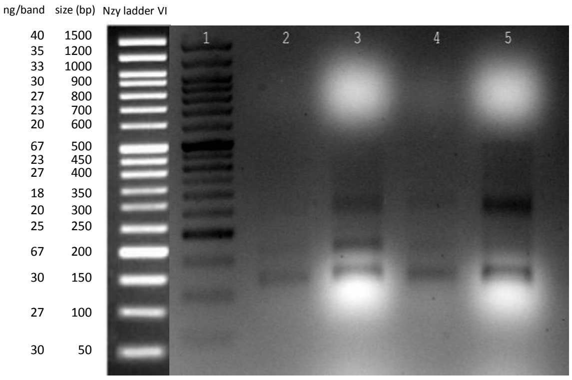


Figure 6.15. 2% agarose gel of 3rd PCR reaction. Lane 1: Nzyladder VI; Lane 2: 5 µL sample E; Lane 3: 55 µL sample E; Lane 4: 5 µL sample F; Lane 5: 55 µL sample F.

Final PCR

Analysis of Figure 6.16, PCR product bands have the expected fragment size for the region flanked by the PstI and EcoRI.

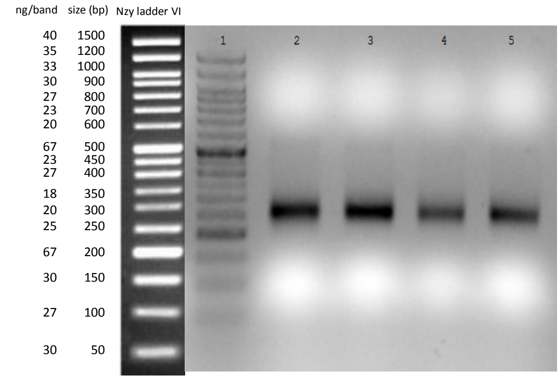


Figure 6.16. 2% agarose gel of annealing/extension followed by PCR reactions. Lane 1: Nzyladder VI; Lanes 2 and 3: 30 µL sample G; Lanes 4 and 5: 30 µL sample H.

6.3.3 Insert and pHBP1-EcoRI double digestion and purification

Double digestion was analyzed using a 3% agarose gel as shown in Figure 6.17. At the top of lane 2, it is possible to visualize a band that corresponds to double digested pHBP1-EcoRI and a smaller one (around 260 bp) that corresponds to the fragment excised from pHBP1-EcoRI. The bands detected in lanes 3 and 4 are the double digested inserts with approximately 260 bp.

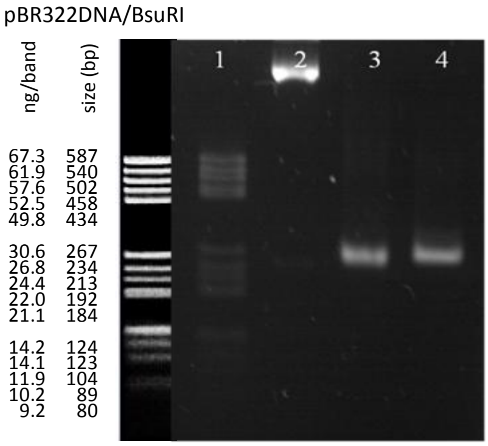


Figure 6.17. 3% agarose gel of PstI and EcoRI double digested products. Lane 1:pBR322DNA/BsuRI. Lane 2 phHBP-EcoRI double digested; Lane 3 and 4: Double digested insert.

Pure double digested pHBP1-EcoRI gave an Abs_{260} of 0.089 which corresponds to 222 ng/µL with a ratio $OD_{260/280}$ of 1.816 and $OD_{260/230}$ of 2.225. Digested DNA insert absorbance measurement results in OD_{260} of 0.012, which corresponds to 30 ng/µL of DNA, with a ratio $OD_{260/280}$ of 1.714 and $OD_{260/230}$ of 0.706.

6.3.4 Ligase reactions and transformation of competent bacterial cells

The negative control did not present any colonies which meant that the double digestion protocol and phosphatase treatment was successful and the positive experiments gave more than 20 colonies. 4 constructs were purified from 4 colonies and sequenced. The sequences were correct except that a point mutation (A39E) was detected (sequencing results not shown).

6.3.5 XhoI NcoI subcloning of chimeric hHBP

For chimeric hHBP subcloning an agarose gel with double digest pHBP1 wild type plasmid and chimeric phHBP-EcoRI plasmid was run on the purified products. Figure 6.18

show that the digestion was successful and it is possible to identify the digested insert highlighted in red with an approximate size of between 600 and 800 bp, in agreement with the expected size of the insert. The XhoI restriction site is located at nucleotide 158 and the NcoI restriction site is located at nucleotide 757 which gives a 599 bp fragment.

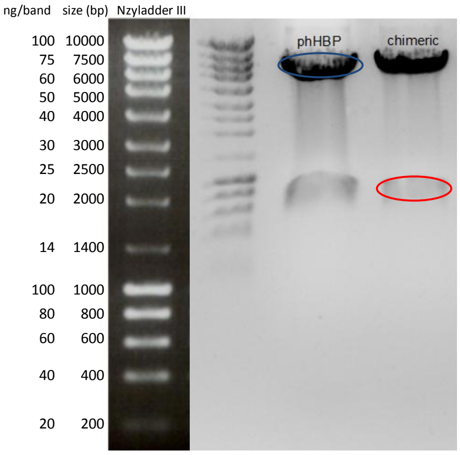


Figure 6.18. 1% agarose gel NcoI and XhoI double digested phHBP1 (highlighted in blue) and double digested chimeric phHBP insert (highlighted in red).

6.3.6 Chimeric hHBP overexpression and purification

Figure 6.19 shows a SDS-Page gel run after overexpression and purification of the chimeric protein. Only a very weak band can be seen in the 75mM imidazole fraction indicating problems with the process. An optimization of the IPTG concentration was attempted and the results of a SDS PAGE gel are shown in Figure 6.20. None of the three different IPTG concentrations tested (0.25 mM, 0.5 mM and 1 mM) resulted in an improvement.

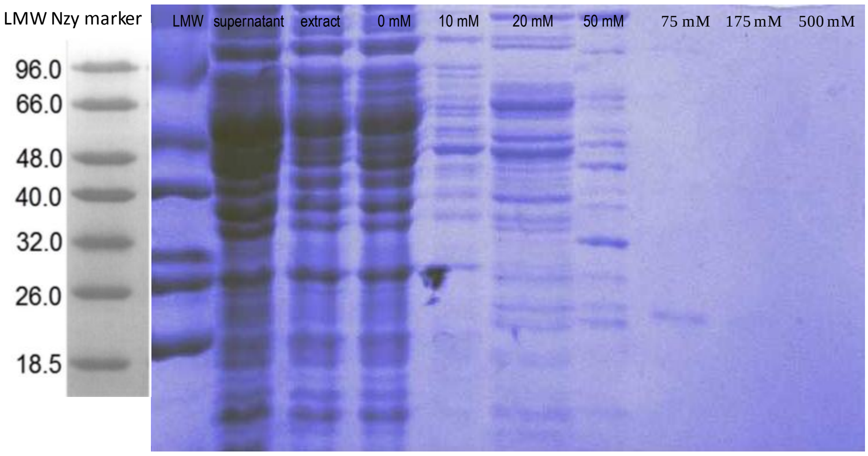


Figure 6.19. Chimeric hHBP purification analysis by SDS Page of different fractions of Ni-NTA purification column.

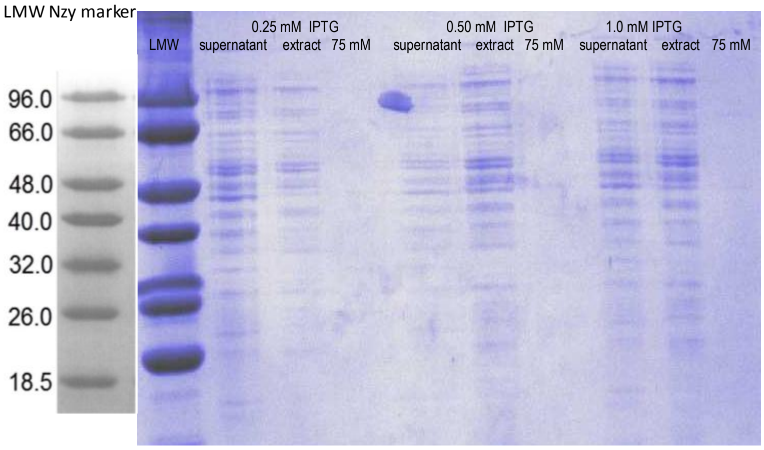


Figure 6.20. Chimeric hHBP purification analysis by SDS Page of supernatant, extract after Ni-NTA elution and 75 mM imidazole fractions for different IPTG concentrations.

6.3.7 Chimeric hHBP-W51V

Transformation of *E.coli* cells with the new plasmid was successful and more than 10 colonies were obtained on LB plates with kanamycin. As described in above, a single colony was isolated and grown overnight. The purified plasmid obtained from this growth was double digested. After double digestion, the region flanked by NcoI and XhoI in chimeric phHBP-W51V was extracted from the gel, purified and used to perform a ligase reaction. As shown in

Figure 6.21, lane 2, a fragment with approximately 600 bp was obtained.

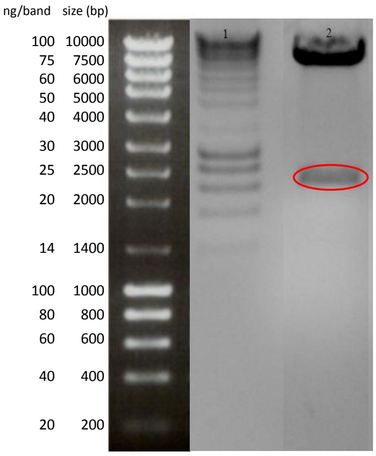


Figure 6.21. 1% agarose gel NcoI and XhoI double digested chimeric phHBP_w51v. Lane 1: Nzyladder III; Lane 2. chimeric phHBP-W51V digested plasmid.

Overexpression and purification of chimeric p22HBP-W51V was then performed but, as illustrated in Figure 6.22, no band corresponding to p22HBP appears in the 75 mM imidazole fraction.

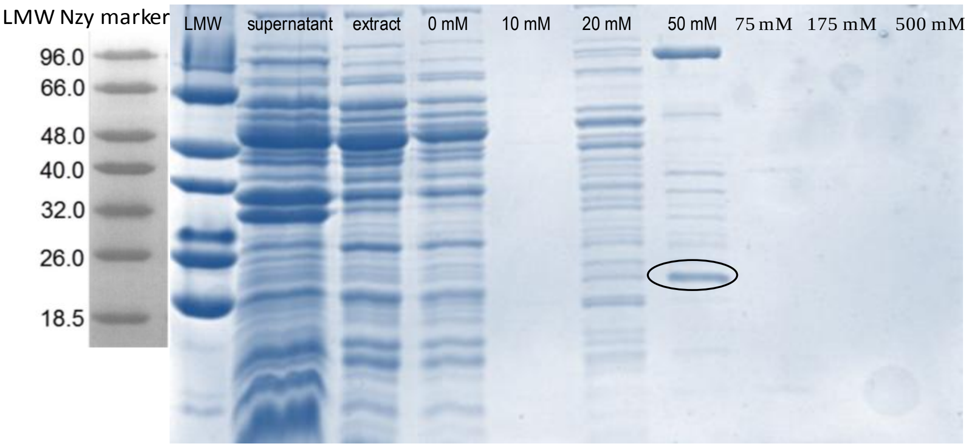


Figure 6.22. Chimeric hHBP_w51v purification analysis by SDS Page of different fractions of Ni-NTA purification column.

However, a band with the expected size for chimeric p22HBP appeared in the 50 mM imidazole fraction. In an attempt to separate this band from the others with higher sizes, the fraction was concentrated in a 30 kDa centricon, at 7000 g but no absorbance at 280 nm was seen indicating that little or no protein was present in the sample.

For chimeric hHBP design and construction a sequential procedures were taken in account to achieve the desired protein. PCR reactions were performed with success and enough amplified DNA was used in each reaction once visible bands were obtained in agarose gel analysis extracted and purified. Double digestion procedures were also carried out with success as the digested fragments obtained had the expected size and the subsequent subcloning retrieve the well structures plasmids. This plasmids were inserted in competent cells whose shown kanamycin resistance which was indicative of viable plasmids. Although these steps were obtained with success it was not possible to overexpress chimeric hHBP. Despite subcloning chimeric hHBP insert into new generated pHBP1-EcoRI plasmid, some deletion or insertion of a base pair could occurred which consequently change initiation codons (signal to start protein translation) and consequently prevent chimeric phHBP overexpression. In order to confirm base pair deletion and/or insertion, chimeric phHBP plasmid sequence should be verified by sequencing all plasmid. This sequencing order was expensive, considering financial resources available, and was not carried out.

7 Conclusions

This work presented in this thesis is based on interactions studies of p22HBP with PPIX/Hemin using NMR and FQ. The overall aim was to probe, in more detail than has been carried out before, the dynamics and molecular interactions involved in tetrapyrrole binding to p22HBP and identify key residues involved in their interaction. Previous structural studies resulted in the structure of murine p22HBP solution by NMR in 2006 by Dias *et al* and more recently (2011) Ambrosi *et al* [24] solved the 3D structure of human SOUL by X-ray crystallography. Molecular modelling studies have also been carried out in order to characterize the molecular recognition process that takes place upon the binding of p22HBP with intermediates of heme synthesis. These studies have confirmed that a hydrophobic patch on the surface of p22HBP is responsible for tetrapyrrole binding. However some residues, namely R56, K64 and K177 may be involved in electrostatic interactions with key residues at the edge of the p22HBP binding pocket. These results were the basis of our work and key residues were the starting point and murine p22HBP variants were constructed to analyze how changes in hydrophobicity and polarity would influence the strength of PPIX/Hemin interactions. Molecular biological techniques were performed in order to obtain p22HBP variants and the resulted constructs were overexpressed and analyzed by fluorescence quenching and NMR and when titrated with tetrapyrrole compounds. As Micaelo and its collaborators [107] identified a mobile region in region flanked by β8 and β9 sheets (residues 171-180), near the protein-binding site and changed with binding and release of the tetrapyrrole rings, dynamics studies were carried out in order to confirm this flexibility upon tetrapyrrole binding. With the aim of supporting functional studies involving siRNA knockdown, a chimeric human p22HBP was also constructed. Chimeric proteins have found wide application for the study of protein folding and protein structure stability [31]

Two main mutations were performed: the introduction of nonpolar side chains by replacing R56, K64 and K177 by alanine or the introduction of negatively charged side chain by replacing R56, K64 and K177 by Glutamic acid which has a carboxylate group on its side chain. Overexpression and purification using affinity chromatography were performed for these variants as well for murine and human version of p22HBP. Fluorescence quenching studies were performed in order to evaluate the PPIX interactions with the new variants, and dissociation constants were determined. Although dissociation constants for p22HBP variants and human p22HBP were in the same order of magnitude (nanomolar range) as previously reported, the main differences were observed in variants with negatively charged side chain of Glutamic acid in key residues. These results demonstrated that residues R56, K64 and K177 are not crucial by themselves or all together (as demonstrated with triple mutant) for hemin interaction with p22HBP and modifications in polarity and hydrophobicity do not prevent neither favor Hemin interactions.

Chemical shift mapping by NMR was also performed for p22HBP variants and human p22HBP and residues involved in the heme-binding were identified. These results shown that the binding pocket identified for murine p22HBP remains in respective variants and human p22HBP. The main differences in chemical shifts were observed near the binding site. K64 and R56 variants shown mainly differences in residues 61-66 and K177 variants shown mainly differences in regions flanked by 172 to 182. Triple variant shown differences in residues ranging from 54 to 66 and 171 to 182. These results suggest that human p22HBP as well as p22HBP variants have similar structures to murine p22HBP. Backbone resonances assignments were performed for human p22HBP and 92% of residues located in secondary structures were assigned while only 61% of residues located in unstructured regions were assigned. This difference is related to flexibility inherent to unstructured regions. Out of 168 spin systems found in ^{1}H-^{15}N HSQC, 150 were assigned mainly in structured regions for the protein. 18 peaks were unassigned and considering 12 Prolines that do not show NH signals, 9 spin systems were missing in ^{1}H-^{15}N HSQC corresponding to N-terminal.

Dynamic studies of human p22HBP were also carried out by NMR, with hetNOE as well as longitudinal and transverse relaxation rates being measured. Model Free analysis was applied to relaxation parameters and the diffusion tensor was determined. Human p22HBP tumbles isotropically in solution, with a correlation time of 12.655 ±0.071 ns. When bound to PPIX this correlation time decreased to 10.435±0.099, a possible consequence of Tween present in solution. Tween works as a surfactant and reduces the surface tension of tetrapyrrole compounds and also reduced the surface tension of p22HBP in solution. As a consequence the p22HBP, despite being linked to PPIX with an increasing of the molecular weight of the complex, tumbles more rapidly in solution. Averaged generalized order parameter of 0.943 suggested a restricted protein with small variations in flexible regions located between secondary structures especially between α2 helix and β7 sheet (residues 159-163). Dynamic behaviour of loop flanked by residues 171 to 180, suggest an inflexible part of protein, either in the presence of PPIX, in contrast with molecular modelling studies previously reported. Experimental conditions of relaxation measurements performed with human p22HBP with PPIX should be optimized in order to obtain more consistent results.

The chimeric p22HBP was the last work of this thesis and we designed it but unfortunately we were not able to overexpress it and consequently purified it. Before overexpression and purification protocols optimization, it is crucial to confirm the chimeric p22HBP encoding sequence is in frame in chimeric phHBP1 plasmid. Thus, a complete sequencing analysis of chimeric p22HB1 plasmid should be carried out. Chimeric p22HBP will be an important support in functional studies to confirm if p22HBP play a central role in heme transport as initially studied by Jean Marck Moulis. Knock down studies shown that an absence of p22HBP in K562- human erythroleukemia cells, strongly up regulated heme oxygenase as a consequence of heme accumulation. And if we have p22HBP in this cells but with the binding pocket replaced by a non-binding homologous sequence? Would heme accumulate in the cells? Would be heme oxygenase upregulated due the presence of this non-binding version? This is the big challenge that should be carried out in the future in order to specifically define p22HBP function.

8 References

[1] S. Taketani, “Aquisition, mobilization and utilization of cellular iron and heme: endless findings and growing evidence of tight regulation.,” *Tohoku J. Exp. Med.*, vol. 205, no. 4, pp. 297–318, Apr. 2005.

[2] R. S. Ajioka, J. D. Phillips, and J. P. Kushner, “Biosynthesis of heme in mammals,” *Biochim. Biophys. Acta-Molecular Cell Res.*, vol. 1763, no. 7, pp. 723–736, 2006.

[3] I. J. Schultz, C. Chen, B. H. Paw, and I. Hamza, “Iron and porphyrin trafficking in heme biogenesis.,” *J. Biol. Chem.*, vol. 285, no. 35, pp. 26753–9, Aug. 2010.

[4] C. J. Reedy and B. R. Gibney, “Heme protein assemblies,” *Chem. Rev.*, vol. 104, no. 2, pp. 617–649, 2004.

[5] N. Wijayanti, N. Katz, and S. Immenschuh, “Biology of heme in health and disease,” *Curr. Med. Chem.*, vol. 11, no. 8, pp. 981–986, 2004.

[6] G. C. Ferreira, “Ferrochelatase,” *Int. J. Biochem. Cell Biol.*, vol. 31, no. 10, pp. 995–1000, 1999.

[7] G. C. Ferreira, “Heme-Biosynthesis - Biochemistry, Molecular-Biology, and Relationship to Disease,” *J. Bioenerg. Biomembr.*, vol. 27, no. 2, pp. 147–150, 1995.

[8] G. C. Ferreira and J. Gong, “5-Aminolevulinate Synthase and the First Step of Heme-Biosynthesis,” *J. Bioenerg. Biomembr.*, vol. 27, no. 2, pp. 151–159, 1995.

[9] B. J. Goodfellow, J. S. Dias, G. C. Ferreira, P. Henklein, V. Wray, and A. L. Macedo, “The solution structure and heme binding of the presequence of murine 5-aminolevulinate synthase,” *Febs Lett.*, vol. 505, no. 2, pp. 325–331, 2001.

[10] J. S. Dias, “NMR studies of murine ALAS and murine p22HBP:two proteins involved in heme metabolism,” Universidade Nova de Lisboa, Lisboa, 2007.

[11] S. Iwahara, H. Satoh, D. X. Song, J. Webb, A. L. Burlingame, Y. Nagae, and U. Mullereberhard, “Purification, Characterization, and Cloning of a Heme-Binding Protein (23 Kda) in Rat-Liver Cytosol,” *Biochemistry*, vol. 34, no. 41, pp. 13398–13406, 1995.

[12] S. Taketani, Y. Adachi, H. Kohno, S. Ikehara, R. Tokunaga, and T. Ishii, “Molecular characterization of a newly identified heme-binding protein induced during differentiation of urine erythroleukemia cells,” *J. Biol. Chem.*, vol. 273, no. 47, pp. 31388–31394, 1998.

[13] L. J. Smith, A. Kahraman, and J. M. Thornton, “Heme proteins-Diversity in structural characteristics, function, and folding,” *Proteins-Structure Funct. Bioinforma.*, vol. 78, no. 10, pp. 2349–2368, 2010.

[14] B. J. Blackmon, T. A. Dailey, L. C. Xiao, and H. A. Dailey, “Characterization of a human and mouse tetrapyrrole-binding protein,” *Arch. Biochem. Biophys.*, vol. 407, no. 2, pp. 196–201, 2002.

[15] J. S. Dias, A. L. Macedo, G. C. Ferreira, F. C. Peterson, B. F. Volkman, and B. J. Goodfellow, “The first structure from the SOUL/HBP family of heme-binding proteins, murine P22HBP. (vol 281, pg 31553, 2006),” *J. Biol. Chem.*, vol. 281, no. 50, p. 38966, 2006.

[16] E. Sato, I. Sagami, T. Uchida, A. Sato, T. Kitagawa, J. Igarashi, and T. Shimizu, “SOUL in mouse eyes is a new hexameric heme-binding protein with characteristic optical absorption, resonance Raman spectral, and heme-binding properties,” *Biochemistry*, vol. 43, no. 44, pp. 14189–14198, 2004.

[17] J. T. Lathrop and M. P. Timko, “Regulation by heme of mitochondrial protein transport through a conserved amino acid motif.,” *Science*, vol. 259, no. 5094, pp. 522–5, Jan. 1993.

[18] W. S. Shin, H. Yamashita, and M. Hirose, “Multiple effects of haemin binding on protease susceptibility of bovine serum albumin and a novel isolation procedure for its large fragment.,” *Biochem. J.*, vol. 304 (Pt 1, pp. 81–6, Nov. 1994.

[19] F. Freire, “Integrated Study by NMR and X-Ray Cristallography on the Analysis of the Molecular Interactions in Heme-Binding Proteins,” Universidade Nova de Lisboa, 2012.

[20] A. Szigeti, S. Bellyei, B. Gasz, A. Boronkai, E. Hocsak, O. Minik, Z. Bognar, G. Varbiro, B. Sumegi, and F. Gallyas, “Induction of necrotic cell death and mitochondrial permeabilization by heme binding protein 2/SOUL.,” *FEBS Lett.*, vol. 580, no. 27, pp. 6447–54, Nov. 2006.

[21] M. J. Zylka and S. M. Reppert, “Discovery of a putative heme-binding protein family (SOUL/HBP) by two-tissue suppression subtractive hybridization and database searches,” *Mol. Brain Res.*, vol. 74, no. 1–2, pp. 175–181, 1999.

[22] N. M. Micaelo, A. L. Macedo, B. J. Goodfellow, and V. Félix, “Tetrapyrrole binding affinity of the murine and human p22HBP heme-binding proteins.,” *J. Mol. Graph. Model.*, vol. 29, no. 3, pp. 396–405, Nov. 2010.

[23] F. Freire, M. J. Romao, A. L. Macedo, S. S. Aveiro, B. J. Goodfellow, and A. L. Carvalho, “Preliminary structural characterization of human SOUL, a haem-binding protein,” *Acta Crystallogr. Sect. F-Structural Biol. Cryst. Commun.*, vol. 65, pp. 723–726, 2009.

[24] E. Ambrosi, S. Capaldi, M. Bovi, G. Saccomani, M. Perduca, and H. L. Monaco, “Structural changes in the BH3 domain of SOUL protein upon interaction with the anti-apoptotic protein Bcl-xL.,” *Biochem. J.*, vol. 438, no. 2, pp. 291–301, Sep. 2011.

[25] M. Babusiak, P. Man, R. Sutak, J. Petrak, and D. Vyoral, “Identification of heme binding protein complexes in murine erythroleukemic cells: Study by a novel two-dimensional native separation - liquid chromatography and electrophoresis,” *Proteomics*, vol. 5, no. 2, pp. 340–350, 2005.

[26] I. Migeotte, E. Riboldi, J.-D. Franssen, F. Grégoire, C. Loison, V. Wittamer, M. Detheux, P. Robberecht, S. Costagliola, G. Vassart, S. Sozzani, M. Parmentier, and D. Communi, “Identification and characterization of an endogenous chemotactic ligand specific for FPRL2.,” *J. Exp. Med.*, vol. 201, no. 1, pp. 83–93, Jan. 2005.

[27] T. Devosse, R. Dutoit, I. Migeotte, P. De Nadai, V. Imbault, D. Communi, I. Salmon, and M. Parmentier, “Processing of HEBP1 by Cathepsin D Gives Rise to F2L, the Agonist of Formyl Peptide Receptor 3,” *J. Immunol.*, vol. 187, no. 3, pp. 1475–1485, 2011.

[28] K. Tiemann and J. J. Rossi, “RNAi-based therapeutics-current status, challenges and prospects.,” *EMBO Mol. Med.*, vol. 1, no. 3, pp. 142–51, Jun. 2009.

[29] D. Lochmatter and P.-E. Mullis, “RNA interference in mammalian cell systems.,” *Horm. Res. pædiatrics*, vol. 75, no. 1, pp. 63–9, Jan. 2011.

[30] R. W. Carthew and E. J. Sontheimer, “Origins and Mechanisms of miRNAs and siRNAs.,” *Cell*, vol. 136, no. 4, pp. 642–55, Feb. 2009.

[31] R. Grandori, K. Struck, K. Giovanielli, and J. Carey, “A three-step PCR protocol for construction of chimeric proteins,” *Protein Eng.*, vol. 10, no. 9, pp. 1099–1100, 1997.

[32] G. Sezonov, D. Joseleau-Petit, and R. D’Ari, “Escherichia coli physiology in Luria-Bertani broth.,” *J. Bacteriol.*, vol. 189, no. 23, pp. 8746–9, Dec. 2007.

[33] C. B. Harley and R. P. Reynolds, “Analysis of E. coli promoter sequences.,” *Nucleic Acids Res.*, vol. 15, no. 5, pp. 2343–61, Mar. 1987.

[34] S. C. Makrides, “Strategies for achieving high-level expression of genes in Escherichia coli.,” *Microbiol. Rev.*, vol. 60, no. 3, pp. 512–38, Sep. 1996.

[35] M. Rai and H. Padh, “Expression systems for production of heterologous proteins,” vol. 80, no. 9, 2001.

[36] Novagen, “Novagen pET system manual,” 2003.

[37] J. A. Asenjo and B. A. Andrews, “Protein purification using chromatography: selection of type, modelling and optimization of operating conditions.,” *J. Mol. Recognit.*, vol. 22, no. 2, pp. 65–76, 2009.

[38] T. Dingermann, D. Steinhilber, and G. Folkers, Eds., *Molecular Biology in Medicinal Chemistry*. Wiley, 2004, p. 413.

[39] J. Arnau, C. Lauritzen, G. E. Petersen, and J. Pedersen, “Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins.,” *Protein Expr. Purif.*, vol. 48, no. 1, pp. 1–13, Jul. 2006.

[40] J. S. Dias, A. L. Macedo, G. C. Ferreira, F. C. Peterson, B. F. Volkman, and B. J. Goodfellow, “The first structure from the SOUL/HBP family of heme-binding proteins, murine P22HBP.,” *J. Biol. Chem.*, vol. 281, no. 42, pp. 31553–61, Oct. 2006.

[41] J. R. Lakowicz, Ed., *Principles of Fluorescence Spectroscopy*. Boston, MA: Springer US, 2006.

[42] L. A. Munishkina and A. L. Fink, “Fluorescence as a method to reveal structures and membrane-interactions of amyloidogenic proteins,” *Biochim. Biophys. Acta-Biomembranes*, vol. 1768, no. 8, pp. 1862–1885, 2007.

[43] S. Doose, H. Neuweiler, and M. Sauer, “Fluorescence Quenching by Photoinduced Electron Transfer: A Reporter for Conformational Dynamics of Macromolecules,” *Chemphyschem*, vol. 10, no. 9–10, pp. 1389–1398, 2009.

[44] J. W. Taraskaz and W. N. Zagotta, “Fluorescence Applications in Molecular Neurobiology,” *Neuron*, vol. 66, no. 2, pp. 170–189, 2010.

[45] M. Moller and A. Denicola, “Protein tryptophan accessibility studied by fluorescence quenching,” *Biochem. Mol. Biol. Educ.*, vol. 30, no. 3, pp. 175–178, 2002.

[46] B. Valeur and M. N. Berberan-Santos, *Molecular Fluorescence: Principles and Applications*, 2nd ed. Wiley-VCH, 2012.

[47] E. C. Hulme and M. A. Trevethick, “Ligand binding assays at equilibrium: validation and interpretation.,” *Br. J. Pharmacol.*, vol. 161, no. 6, pp. 1219–37, Nov. 2010.

[48] S. Swillens, “Interpretation of binding curves obtained with high receptor concentrations: practical aid for computer analysis,” *Mol. Pharmacol.*, vol. 47, no. 6, pp. 1197–1203, 1995.

[49] S. E. Harding and B. Z. Chowdhry, “Protein-Ligand Interactions: structure and spectroscopy.” 2001.

[50] X. L. Zhang, A. Sanger, R. Hemmig, and W. Jahnke, “Ranking of High-Affinity Ligands by NMR Spectroscopy,” *Angew. Chemie-International Ed.*, vol. 48, no. 36, pp. 6691–6694, 2009.

[51] C. J. Markin and L. Spyracopoulos, “Increased precision for analysis of protein-ligand dissociation constants determined from chemical shift titrations.,” *J. Biomol. NMR*, vol. 53, no. 2, pp. 125–38, Jun. 2012.

[52] L. P. C. Delgado, “Estudos estruturais da versão humana da proteína p22HBP Structural protein studies of human Leonildo Pedro Cabral Delgado Estudos estruturais da versão humana da proteína p22HBP Structural protein studies of human,” Universidade de Aveiro, 2011.

[53] Q. Teng, *Structural Biology: Pratical NMR Applications*. New York: Springer Science+Business Media, Inc., 2005.

[54] M. Billeter, A. D. Kline, W. Braun, R. Huber, and K. Wuthrich, “Comparison of the High-Resolution Structures of the Alpha-Amylase Inhibitor Tendamistat Determined by Nuclear Magnetic-Resonance in Solution and by X-Ray-Diffraction in Single-Crystals,” *J. Mol. Biol.*, vol. 206, no. 4, pp. 677–687, 1989.

[55] W. Jahnke, “Spin labels as a tool to identify and characterize protein-ligand interactions by NMR spectroscopy,” *Chembiochem*, vol. 3, no. 2–3, pp. 167–173, 2002.

[56] T. Maurer, “NMR studies of Protein-Ligand Interactions,” in *Protein-Ligand Interactions: Methods and Applications*, G. U. Nienhaus, Ed. Totowa, New Jersey: Humana Press Inc., 2005.

[57] R. A. de Graaf, *In Vivo NMR Spectroscopy: principles and tecniques*, 2nd ed. Wiley, 2007.

[58] V. M. S. Gil and C. F. G. C. Geraldes, *Ressonância Magnética Nuclear*. Coimbra: Fundação Calouste Gulbenkian, 2002.

[59] K. Wüthrich, *NMR of proteins and nucleic acids*. New YorK: John Wiley & Sons, 1986.

[60] A. L. Skinner and J. S. Laurence, “High-Field Solution NMR Spectroscopy as a Tool for Assessing Protein Interactions with Small Molecule Ligands,” *J. Pharm. Sci.*, vol. 97, no. 11, pp. 4670–4695, 2008.

[61] N. L. of Medicine, “Statistics Calculated for Selected Chemical Shifts from Atoms in the 20 Common Amino Acids.” Biological Magnetic Resonance Data Bank , 2011.

[62] M. Pellecchia, I. Bertini, D. Cowburn, C. Dalvit, E. Giralt, W. Jahnke, T. L. James, S. W. Homans, H. Kessler, C. Luchinat, B. Meyer, H. Oschkinat, J. Peng, H. Schwalbe, and G. Siegal, “Perspectives on NMR in drug discovery: a technique comes of age,” *Nat. Rev. Drug Discov.*, vol. 7, no. 9, pp. 738–745, 2008.

[63] J. Clarkson and I. D. Campbell, “Studies of protein-ligand interactions by NMR,” *Biochem. Soc. Trans.*, vol. 31, pp. 1006–1009, 2003.

[64] M. Sattler, J. Schleucher, and C. Griesinger, “Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed

field gradients," *Prog. Nucl. Magn. Reson. Spectrosc.*, vol. 34, no. 2, pp. 93–158, 1999.

[65] J. Cavanagh, *Protein NMR spectroscopy: principles and practice*, Second . Academic Press, 2007.

[66] K. Pervushin, N. Krishna, and L. Berliner, "Transverse Relaxation Optimized Spectroscopy: Protein NMR for the Millennium," vol. 20, Springer US, 2002, pp. 3–34.

[67] C. Fernandez and G. Wider, "TROSY in NMR studies of the structure and function of large biological macromolecules," *Curr. Opin. Struct. Biol.*, vol. 13, no. 5, pp. 570–580, 2003.

[68] K. Pervushin, "Impact of Transverse Relaxation Optimized Spectroscopy (TROSY) on NMR as a technique in structural biology," *Q. Rev. Biophys.*, vol. 33, no. 2, pp. 161–197, 2000.

[69] K. Pervushin, R. Riek, G. Wider, and K. Wuthrich, "Attenuated T-2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 94, no. 23, pp. 12366–12371, 1997.

[70] K. Wüthrich, "Protein-Structure Determination in Solution by NMR-Spectroscopy," *J. Biol. Chem.*, vol. 265, no. 36, pp. 22059–22062, 1990.

[71] M. Schubert, "Development and application of novel triple-resonance experiments for the assignment of protein NMR spectra," University of Berlin, Berlin, 2002.

[72] K. H. Gardner and L. E. Kay, "The use of H-2, C-13, N-15 multidimensional NMR to study the structure and dynamics of proteins," *Annu. Rev. Biophys. Biomol. Struct.*, vol. 27, pp. 357–406, 1998.

[73] J. S. Dias, A. L. Macedo, G. C. Ferreira, N. Jeanty, S. Taketani, B. J. Goodfellow, F. C. Peterson, and B. F. Volkman, "H-1, N-15 and C-13 resonance assignments of the heme-binding protein murine p22HBP," *J. Biomol. Nmr*, vol. 32, no. 4, p. 338, 2005.

[74] L. E. Kay, "NMR studies of protein structure and dynamics," *J. Magn. Reson.*, vol. 173, no. 2, pp. 193–207, 2005.

[75] B. Meyer and T. Peters, "NMR Spectroscopy techniques for screening and identifying ligand binding to protein receptors," *Angew. Chemie-International Ed.*, vol. 42, no. 8, pp. 864–890, 2003.

[76] S. B. Shuker, P. J. Hajduk, R. P. Meadows, and S. W. Fesik, "Discovering high-affinity ligands for proteins: SAR by NMR," *Science (80-.).*, vol. 274, no. 5292, pp. 1531–1534, 1996.

[77] B. J. Goodfellow, “NMR Structural Studies of Desulforedoxin,” Aveiro, University of Aveiro, 1997.

[78] J. Marley, M. Lu, and C. Bracken, “A method for efficient isotopic labeling of recombinant proteins,” *J. Biomol. Nmr*, vol. 20, no. 1, pp. 71–75, 2001.

[79] J. Sambrook and D. W. Russel, *Molecular Cloning: A Laboratory Manual*, Third. New York: Cold Spring Harbor Laboratory Press , 2001.

[80] J. Fiaux, E. B. Bertelsen, A. L. Horwich, and K. Wuthrich, “Uniform and residue-specific N-15-labeling of proteins on a highly deuterated background,” *J. Biomol. Nmr*, vol. 29, no. 3, pp. 289–297, 2004.

[81] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, and A. Bax, “NMRPipe: a multidimensional spectral processing system based on UNIX pipes.,” *J. Biomol. NMR*, vol. 6, no. 3, pp. 277–93, Nov. 1995.

[82] R. Keller, *The Computer Aided Resonance Assignment Tutorial*. Goldau: Cantina Verlag, 2004.

[83] Y. Shen, F. Delaglio, G. Cornilescu, and A. Bax, “TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts.,” *J. Biomol. NMR*, vol. 44, no. 4, pp. 213–23, Aug. 2009.

[84] D. S. Wishart, B. D. Sykes, and F. M. Richards, “The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy,” *Biochemistry*, vol. 31, no. 6, pp. 1647–1651, Feb. 1992.

[85] J. S. Fraser, M. W. Clarkson, S. C. Degnan, R. Erion, D. Kern, and T. Alber, “Hidden alternative structures of proline isomerase essential for catalysis.,” *Nature*, vol. 462, no. 7273, pp. 669–73, Dec. 2009.

[86] D. Fushman, “BioNMR in drug research,” O. Zerbe, Ed. Wiley-VCH, 2006.

[87] K. Hung Sze and P. Man Lai, “Probing Protein Dynamics by Nuclear Magnetic Resonance,” *Protein Pept. Lett.*, vol. 18, no. 4, pp. 373–379, Apr. 2011.

[88] A. Mittermaier and L. E. Kay, “New tools provide new insights in NMR studies of protein dynamics.,” *Science*, vol. 312, no. 5771, pp. 224–8, Apr. 2006.

[89] M. Heller, “Protein Dynamics of ADT and KdpBN by NMR Spectroscopy,” Technische Universität München, 2004.

[90] D. Neuhaus and M. P. Williamson, *The Nuclear Overhauser Effect in Structural and Conformational Analysis, 2nd Edition*, 2nd Editio. Wiley-VCH, 2000, p. 656.

[91] A. Viegas, “Molecular Determinants of Ligand Specificity in Carbohydrate-Binding Modules: an NMR and X-ray crystallography integrated study,” Universidade Nova de Lisboa, 2012.

[92] T. D. W. Claridge, *High-resolution NMR Techniques in Organic Chemistry*. Newnes, 2009, p. 383.

[93] S. Grzesiek, “From biomolecular structure to functional understanding : new NMR developments narrow the gap,” pp. 585–595, 2009.

[94] V. Kanelis, J. D. Forman-Kay, and L. E. Kay, “Multidimensional NMR methods for protein structure determination.,” *IUBMB Life*, vol. 52, no. 6, pp. 291–302, Dec. 2001.

[95] N. Bloembergen, E. Purcell, and R. Pound, “Relaxation Effects in Nuclear Magnetic Resonance Absorption,” *Phys. Rev.*, vol. 73, no. 7, pp. 679–712, Apr. 1948.

[96] R. Ishima and D. a Torchia, “Protein dynamics from NMR.,” *Nat. Struct. Biol.*, vol. 7, no. 9, pp. 740–3, Sep. 2000.

[97] V. A. Jarymowycz and M. J. Stone, “Fast time scale dynamics of protein backbones: NMR relaxation methods, applications, and functional consequences.,” *Chem. Rev.*, vol. 106, no. 5, pp. 1624–71, May 2006.

[98] D. E. Kim, D. Chivian, and D. Baker, “Protein structure prediction and analysis using the Robetta server.,” *Nucleic Acids Res.*, vol. 32, no. Web Server issue, pp. W526–31, Jul. 2004.

[99] N. Tjandra, H. Kuboniwa, H. Ren, and a Bax, “Rotational dynamics of calcium-free calmodulin studied by 15N-NMR relaxation measurements.,” *Eur. J. Biochem.*, vol. 230, no. 3, pp. 1014–24, Jun. 1995.

[100] A. J. Wand, J. L. Urbauer, R. P. McEvoy, and R. J. Bieber, “Internal dynamics of human ubiquitin revealed by 13C-relaxation studies of randomly fractionally labeled protein.,” *Biochemistry*, vol. 35, no. 19, pp. 6116–25, May 1996.

[101] S. R. Site, W. Feng, R. Tejero, D. E. Zimmerman, M. Inouye, and G. T. Montelione, “Solution NMR Structure and Backbone Dynamics of the Major Cold-Shock Protein (CspA) from Escherichia coli : Evidence for Conformational Dynamics in the,” vol. 2960, no. 98, pp. 10881–10896, 1998.

[102] L. Banci, I. Bertini, C. Cavazza, I. C. Felli, and D. Koulougliotis, “Probing the backbone dynamics of oxidized and reduced rat microsomal cytochrome b5 via 15N rotating frame NMR relaxation measurements: biological implications.,” *Biochemistry*, vol. 37, no. 35, pp. 12320–30, Sep. 1998.

[103] C. Baldellon, J. R. Alattia, M. P. Strub, T. Pauls, M. W. Berchtold, A. Cavé, and A. Padilla, “15N NMR relaxation studies of calcium-loaded parvalbumin show tight

dynamics compared to those of other EF-hand proteins.," *Biochemistry*, vol. 37, no. 28, pp. 9964–75, Jul. 1998.

[104] R. Gnanasekaran and D. M. Leitner, "Dielectric response and vibrational energy relaxation in photoactive yellow protein: A molecular dynamics simulation study," *Chem. Phys. Lett.*, vol. 516, no. 1–3, pp. 102–105, Nov. 2011.

[105] B. F. Volkman, S. L. Alam, J. D. Satterlee, and J. L. Markley, "Solution structure and backbone dynamics of component IV Glycera dibranchiata monomeric hemoglobin-CO.," *Biochemistry*, vol. 37, no. 31, pp. 10906–19, Aug. 1998.

[106] C. Renner, R. Baumgartner, A. A. Noegel, and T. A. Holak, "Backbone dynamics of the CDK inhibitor p19(INK4d) studied by 15N NMR relaxation experiments at two field strengths.," *J. Mol. Biol.*, vol. 283, no. 1, pp. 221–9, Jan. 1998.

[107] N. M. Micaelo, A. L. Macedo, B. J. Goodfellow, and V. Felix, "Tetrapyrrole binding affinity of the murine and human p22HBP heme-binding proteins," *J. Mol. Graph. Model.*, vol. 29, no. 3, pp. 396–405, 2010.

[108] S. Immenschuh, E. Baumgart-Vogt, M. Tan, S. Iwahara, G. Ramadori, and H. D. Fahimi, "Differential cellular and subcellular localization of heme-binding protein 23/peroxiredoxin I and heme oxygenase-1 in rat liver," *J. Histochem. Cytochem.*, vol. 51, no. 12, pp. 1621–1631, 2003.

[109] S. Mocellin and M. Provenzano, "RNA interference: learning gene knock-down from cell physiology.," *J. Transl. Med.*, vol. 2, no. 1, p. 39, Jan. 2004.

[110] D. Hanahan, "Studies on transformation of Escherichia coli with plasmids," *J. Mol. Biol.*, vol. 166, no. 4, pp. 557–580, Jun. 1983.

[111] D. Hanahan, J. Jessee, and F. R. Bloom, "Plasmid Transformation of Escherichia-Coli and Other Bacteria," *Methods Enzymol.*, vol. 204, pp. 63–113, 1991.

[112] T. Lendrihas, G. A. Hunter, and G. C. Ferreira, "Targeting the Active Site Gate to Yield Hyperactive Variants of 5-Aminolevulinate Synthase," *J. Biol. Chem.*, vol. 285, no. 18, pp. 13704–13711, 2010.

[113] J. Gong and G. C. Ferreira, "Aminolevulinate Synthase - Functionally Important Residues at a Glycine Loop, a Putative Pyridoxal-Phosphate Cofactor-Binding Site," *Biochemistry*, vol. 34, no. 5, pp. 1678–1685, 1995.

[114] K. Struhl, "Subcloning of DNA fragments.," *Curr. Protoc. Mol. Biol.*, vol. Chapter 3, p. Unit3.16, May 2001.

9 Appendix

9.1 Buffers and media receipts

LB (1L)	
Yeast extract	5 g
Bactotriptone	10 g
NaCl	5 g
up to 1000 mL ddH_2O	

M solution	
MOPS	42 g
Tricine	4 g
NaCl	14.6 g
KOH	8 g
NH_4Cl	2.55 g
up to 1000 mL ddH_2O pH should be 7.3 to 7.4	

T solution (100 mL)	
HCl conc.	8 mL
$CaCl_2$	18.4 mg
H_3BO_3	64 mg
$MnCl_2.4H_2O$	40 mg
$CoCl_2.6H_2O$	18 mg
$CuCl_2.2H_2O$	4 mg
$ZnCl_2$	340 mg
$Na_2MoO_4.2H_2O$	605 mg
up to 100 mL ddH_2O	

MOPS Media	
H_2O	745 mL
M solution	200 mL

O solution (50 mL)
Dissolve 0.1 g $FeCl_2.4H_2O$ in 10 mL concentrated HCl.
Add 10 mL H_2O.
Add 1 mL of T solution.
Add 2.68 g $MgCl_2.6H_2O$.
Up to 50 mL ddH_2O. Filter sterilize. Store at room temperature.

P solution
KH_2PO_4 1.0 M

S solution
K_2SO_4 0.276 M

M9 Salts (10×) 1L	
$Na_2HPO_4.7H_2O$	128 g
KH_2PO_4	30 g
NH_4Cl	10 g
NaCl	5 g

TBE	
Tris base	108 g
Boric acid	55 g
EDTA	9.3 g
up to 1000 mL ddH_2O	

SOC media (1 L)	
Tryptone	20 g
Yeast extract	5 g
NaCl 5M	2 mL
KCl 1M	2.5 mL
$MgCl_2$ 1M	10 mL
$MgSO_4$ 1M	10 mL
Glucose 1M	20 mL
up to 1000 mL ddH_2O	

TB broth 1L	
Tryptone	12 g
Yeast extract	24 g
Glycerol	4 mL
Up to 900 mL ddH_2O. Autoclave	
Add 100 mL sterile solution of:	
KH_2PO_4 0.17 M/ K_2HPO_4 0.72 M	

9.2 Human p22HBP chemical shifts

Residue	δCO(ppm)	δCα(ppm)	δCβ(ppm)	δHα(ppm)	δHN(ppm)	δN(ppm)
13 V	175.340	62.076	32.151	3.993	8.077	121.857
14 E						
15 T						
16 W					10.832	128.588
17 P						
18 W	174.480	55.131		5.380	7.545	118.793
19 Q	175.517	53.740	31.180	4.843	8.173	117.202
20 V	176.080	63.366		4.691	9.234	126.680
21 L	177.653	55.903	41.661	4.298	9.193	130.084
22 S	171.212	58.484	64.356	4.481	7.737	110.873
23 K	174.984	55.485	34.856	4.718	8.299	118.320
24 G	170.732	44.552		4.173	7.578	108.426
25 D	175.166	53.560	43.770	5.252	7.967	117.185
26 K	175.691	56.081	33.605	4.426	8.758	122.469
27 E		57.981	26.487		9.155	124.910
28 E	175.737	57.682	27.347	4.086	8.693	116.298
29 V	173.292	62.435	31.493	4.089	7.886	121.541
30 A	176.902	50.702	20.519	4.988	8.231	130.169
31 Y	174.274	55.483	42.240	5.709	8.354	118.433
32 E	173.932	54.159	33.057	4.868	9.016	117.819
33 E	175.691	54.059	30.789	5.078	8.895	127.121
34 R	173.794	53.644	33.996	5.052	9.099	125.126
35 A	177.474	50.767	17.414	4.845	8.828	125.357
36 C	175.486	57.549	26.252	5.052	9.123	124.892
37 E	178.869	58.526	29.546		8.788	124.332
38 G	172.126	44.243	8.819	4.663		111.467
39 G	171.966	43.928			8.965	111.644
40 K	176.469	55.721	33.683	5.036	9.804	121.756
41 F	174.640	56.304	39.784		9.358	120.003
42 A	176.738	50.309	19.369	5.493	9.559	122.937
43 T	173.640	57.329	71.059	5.319	9.488	113.898
44 V	173.977	61.272	36.973	4.300	6.941	115.918
45 E	175.029	54.990	31.493	5.146	9.074	129.693
46 V	175.257	61.486	34.387	4.413	8.765	126.019
47 T	173.269	61.022	70.758	4.717	8.673	119.506
48 D	174.649	56.670	39.080	4.355	8.206	115.025

Residue	δCO(ppm)	δCα(ppm)	δCβ(ppm)	δHα(ppm)	δHN(ppm)	δN(ppm)
49 K	172.731	53.524	33.917	4.856	7.731	117.997
50 P						
51 V	176.573	67.720	31.806		8.520	121.807
52 D	179.325	58.934	40.175		9.189	116.791
53 E	178.983	58.458		4.032	7.154	118.656
54 A	179.737	55.195	18.509	4.693	8.867	123.857
55 L	178.983	58.258	41.504	4.987	8.565	119.022
56 R	177.254	58.679	29.459	3.983	7.012	116.966
57 E	178.210	57.483	29.772	4.210	7.207	116.189
58 A	179.131	55.056	20.386		8.442	119.790
59 M	176.644	59.078	28.208	4.655	8.231	114.884
60 P						
61 K	180.257	59.908	31.727	3.956	6.808	113.371
62 V	176.194	66.137	30.241	3.435	7.250	117.873
63 A	174.696	54.990	17.023		7.599	121.379
64 K	179.061	59.553	30.945		8.253	117.782
65 Y	176.738	60.835	37.046		7.498	122.011
66 A	177.653	53.863	17.257		7.890	120.633
67 G	174.549	44.665	4.491		4.186	103.009
68 G		44.545		4.342	6.818	105.189
69 T	173.634	61.219	65.752	4.449	8.634	122.300
70 N	175.869	52.094	41.817	5.145	8.310	121.163
71 D	176.788	57.053	39.705		9.338	120.102
72 K	176.526	54.937	33.213	4.490	7.513	116.821
73 G	175.025	46.435	7.855		4.698	111.808
74 I		59.633	40.800	4.209	7.698	116.926
75 G	173.634	44.884			8.534	114.416
76 M	175.943	55.728	29.459	4.100	8.547	123.103
77 G		44.494		4.318	7.619	106.764
78 M	174.151	53.915	32.822	4.989	7.869	112.017
79 T		61.239	69.119	4.982	8.554	119.033
80 V	180.674	61.742		5.002	6.295	119.576
81 P	175.851					
82 I	178.046	56.626	29.772	4.378	8.994	119.891
83 S	174.914	60.517	62.780		8.420	107.708
84 F	172.537	55.295	39.926	5.593	7.196	117.558
85 A	175.509	51.013	20.933	4.440	10.003	126.611
86 V	173.634	59.077	32.979	4.438	7.912	111.733

Residue		δCO(ppm)	δCα(ppm)	δCβ(ppm)	δHα(ppm)	δHN(ppm)	δN(ppm)
87	F	173.749	55.311	39.940	4.960	9.038	116.551
88	P						
89	N						
90	E		59.758				
91	D	177.019	53.361	39.549	4.490	7.775	114.780
92	G	173.749	44.984		4.682	7.958	109.378
93	S	172.913	58.641	64.031	4.265	8.076	115.043
94	L	178.629	54.391	40.916	4.733	8.742	118.468
95	Q	177.681	55.954	29.694	4.009	8.385	119.774
96	K	175.509	57.767	27.816			
97	K	173.269	56.500	34.621	4.757	7.820	121.435
98	L	175.564	55.242	46.119	5.297	8.623	123.399
99	K	175.303	56.015	35.560	5.456	8.314	124.175
100	V	173.886	63.399	32.822	3.238	9.354	130.376
101	W		57.629	29.928	5.051	9.117	125.001
102	F						
103	R	174.755	56.689				
104	I		58.487			7.202	129.842
105	P						
106	N	178.122	57.212			10.101	123.445
107	Q	175.851	58.211	27.425	4.681	9.582	118.947
108	F	174.462	57.556		4.579	8.125	115.273
109	Q	177.040	59.560	27.973		7.178	118.985
110	S	175.189	54.201				
111	D	180.354	63.611	39.681	4.056	7.682	126.128
112	P						
113	P						
114	A	176.263	50.336	17.570		8.910	128.780
115	P						
116	S	173.612	59.648	62.388	4.689	7.469	119.177
117	D	176.400	52.907		6.931	8.268	123.235
118	K	177.268	57.383		4.159	8.358	122.984
119	S		60.332	63.253	4.284	8.885	115.481
120	V	174.227	61.029	30.476	4.438	7.906	124.251
121	K	175.212	53.787	33.683	4.520	8.672	124.807
122	I	176.961	57.697		5.329	8.261	120.480
123	E	174.183	55.124	30.632	4.877	9.324	126.869
124	E	175.966	54.912	31.571	5.495	8.757	121.694

Residue		**δCO(ppm)**	**δCα(ppm)**	**δCβ(ppm)**	**δHα(ppm)**	**δHN(ppm)**	**δN(ppm)**
125	R	174.343	56.002	33.839	4.564	9.191	123.672
126	E						
127	G	173.177	46.516		4.704	8.516	106.776
128	I	174.206	59.465	41.192	4.674	7.632	114.058
129	T		61.179		4.993	8.288	120.553
130	V	174.191	56.786	34.826	5.279	8.629	117.425
131	Y	174.395	57.012	39.705	4.991	8.901	118.693
132	S		55.898	64.973	6.091	9.331	116.185
133	M	173.875	54.701	37.672	4.998	8.676	127.106
134	Q	174.983	54.484	31.180	3.785	8.963	127.005
135	F	174.255	56.671	40.253	4.998	9.161	120.694
136	G		43.971			8.732	110.011
137	G		44.287			7.557	104.305
138	Y		57.448	36.655		8.534	120.711
139	A	176.080	52.264	18.821		8.292	130.235
140	K	177.395	54.123	34.543	4.844	9.420	123.561
141	E	177.102	62.036	28.912		10.626	125.349
142	A	181.039	55.072	17.883		8.565	115.593
143	D	177.517	57.127	41.035	4.426	7.801	116.677
144	Y	178.523	61.961			7.993	118.725
145	V	179.014	66.179	31.966	3.833	8.146	118.070
146	A	181.618	55.361	17.805		8.043	123.485
147	Q	178.320	57.712	25.785	1.386	8.716	117.340
148	A	179.323	55.787	16.397		7.880	121.598
149	T	176.181	66.723	68.774		7.817	114.437
150	R	178.662	59.335	29.068		7.333	121.781
151	L	177.747	57.846	39.393	4.401	7.738	120.881
152	R	179.554	60.808	29.615		8.396	116.866
153	A	179.965	54.766	16.944	4.088	7.718	119.836
154	A	178.169	54.086	17.573	4.036	7.469	120.180
155	L	177.223	54.510	41.270	3.812	7.255	114.708
156	E	177.680	58.280	28.520		7.293	121.686
157	G	174.341	45.539	19.760			
158	T	173.118	63.092	70.210	4.373	8.203	114.193
159	A	175.517	51.383	22.353	4.637	7.844	124.136
160	T	173.175	60.142	70.210	4.299	8.322	111.733
161	Y	170.884	52.452	40.331	4.783	7.190	117.448
162	R	175.937	58.413				

Residue	δCO(ppm)	δCα(ppm)	δCβ(ppm)	δHα(ppm)	δHN(ppm)	δN(ppm)
163 G		45.799			8.410	107.661
164 D	174.480	53.594	39.549	4.662	9.499	115.203
165 I	172.772	59.429	30.319	4.689	7.250	118.134
166 Y	171.463	56.854	38.610	4.306	7.956	119.220
167 F		55.076	41.504	5.737	9.097	113.521
168 C	173.520	55.475		5.057	8.622	119.150
169 T		58.442	70.145		8.992	111.380
170 G	173.986	45.228				
171 Y	176.961	55.565	37.750	4.211	8.016	120.125
172 D						
173 P						
174 P						
175 M						
176 K						
177 P						
178 Y	175.166	55.688	39.784		7.128	115.737
179 G	174.797	44.367			7.729	106.878
180 R						
181 R		55.923	33.619		9.404	121.520
182 N	175.851	50.421	42.147	5.416	8.247	121.436
183 E	176.309	54.574	31.969	5.808	8.222	115.981
184 I	174.456	58.199	37.672	4.847	8.393	118.392
185 W	176.034	53.099	31.962	5.787	9.513	125.191
186 L	177.131	53.189	41.797	5.005	8.964	122.102
187 L	175.166	55.156	42.287	4.855	10.040	125.885
188 K	175.966	58.014	33.333	3.717	8.137	127.477
189 T	178.594	62.775	69.975	4.118	7.684	123.332

9.3 Chemical shift index for human p22HBP

Residue		CA (exp)	HA (exp	CA (ref)	HA (ref)	$\Delta\delta_{CA}$	$\Delta\delta_{HA}$
7	N			53.60	4.75		
8	S			58.30	4.50		
9	L			55.70	4.17		
10	F			57.90	4.66		
11	G			45.00	3.97		
12	S			58.30	4.50		
13	V	62.08	3.99	63.00	3.95	-0.92	0.04
14	E			56.70	4.29		
15	T			63.10	4.35		
16	W			57.80	4.70		
17	P			62.90	4.44		
18	W	55.13	5.38	57.80	4.70	-2.67	0.68
19	Q	53.74	4.84	56.20	4.37	-2.46	0.47
20	V	63.37	4.69	63.00	3.95	0.37	0.74
21	L	55.90	4.30	55.70	4.17	0.20	0.13
22	S	58.48	4.48	58.30	4.50	0.18	-0.02
23	K	55.49	4.72	56.70	4.36	-1.22	0.36
24	G	44.55	4.17	45.00	3.97	-0.45	0.20
25	D	53.56	5.25	54.10	4.76	-0.54	0.49
26	K	56.08	4.43	56.70	4.36	-0.62	0.07
27	E	57.98		56.70	4.29	1.28	
28	E	57.68	4.09	56.70	4.29	0.98	-0.20
29	V	62.44	4.09	63.00	3.95	-0.56	0.14
30	A	50.70	4.99	52.50	4.35	-1.80	0.64
31	Y	55.48	5.71	58.60	4.60	-3.12	1.11
32	E	54.16	4.87	56.70	4.29	-2.54	0.58
33	E	54.06	5.08	56.70	4.29	-2.64	0.79
34	R	53.64	5.05	56.30	4.38	-2.66	0.67
35	A	50.77	4.85	52.50	4.35	-1.73	0.50
36	C	57.55	5.05	58.80	4.65	-1.25	0.40
37	E	58.53		56.70	4.29	1.83	
38	G	44.24	4.66	45.00	3.97	-0.76	0.69
39	G	43.93		45.00	3.97	-1.07	
40	K	55.72	5.04	56.70	4.36	-0.98	0.68
41	F	56.30		57.90	4.66	-1.60	
42	A	50.31	5.49	52.50	4.35	-2.19	1.14

Residue		CA (exp)	HA (exp	CA (ref)	HA (ref)	$\Delta\delta_{CA}$	$\Delta\delta_{HA}$
43	T	57.33	5.32	63.10	4.35	-5.77	0.97
44	V	61.27	4.30	63.00	3.95	-1.73	0.35
45	E	54.99	5.15	56.70	4.29	-1.71	0.86
46	V	61.49	4.41	63.00	3.95	-1.51	0.46
47	T	61.02	4.72	63.10	4.35	-2.08	0.37
48	D	56.67	4.36	54.10	4.76	2.57	-0.40
49	K	53.52	4.86	56.70	4.36	-3.18	0.50
50	P			62.90	4.44		
51	V	67.72		63.00	3.95	4.72	
52	D	58.93		54.10	4.76	4.83	
53	E	58.46	4.03	56.70	4.29	1.76	-0.26
54	A	55.20	4.69	52.50	4.35	2.70	0.34
55	L	58.26	4.99	55.70	4.17	2.56	0.82
56	R	58.68	3.98	56.30	4.38	2.38	-0.40
57	E	57.48	4.21	56.70	4.29	0.78	-0.08
58	A	55.06		52.50	4.35	2.56	
59	M	59.08	4.66	56.60	4.52	2.48	0.14
60	P			62.90	4.44		
61	K	59.91	3.96	56.70	4.36	3.21	-0.40
62	V	66.14	3.44	63.00	3.95	3.14	-0.52
63	A	54.99		52.50	4.35	2.49	
64	K	59.55		56.70	4.36	2.85	
65	Y	60.84		58.60	4.60	2.24	
66	A	53.86		52.50	4.35	1.36	
67	G	44.67		45.00	3.97	-0.34	
68	G	44.55	4.34	45.00	3.97	-0.45	0.37
69	T	61.22	4.45	63.10	4.35	-1.88	0.10
70	N	52.09	5.15	53.60	4.75	-1.51	0.40
71	D	57.05		54.10	4.76	2.95	
72	K	54.94	4.49	56.70	4.36	-1.76	0.13
73	G	46.44		45.00	3.97	1.44	
74	I	59.63	4.21	62.60	3.95	-2.97	0.26
75	G	44.88		45.00	3.97	-0.12	
76	M	55.73	4.10	56.60	4.52	-0.87	-0.42
77	G	44.49	4.32	45.00	3.97	-0.51	0.35
78	M	53.92	4.99	56.60	4.52	-2.69	0.47
79	T	61.24	4.98	63.10	4.35	-1.86	0.63
80	V	61.74	5.00	63.00	3.95	-1.26	1.05

Residue		CA (exp)	HA (exp	CA (ref)	HA (ref)	$\Delta\delta_{CA}$	$\Delta\delta_{HA}$
81	P			62.90	4.44		
82	I	56.63	4.38	62.60	3.95	-5.97	0.43
83	S	60.52		58.30	4.50	2.22	
84	F	55.30	5.59	57.90	4.66	-2.61	0.93
85	A	51.01	4.44	52.50	4.35	-1.49	0.09
86	V	59.08	4.44	63.00	3.95	-3.92	0.49
87	F	55.31	4.96	57.90	4.66	-2.59	0.30
88	P			62.90	4.44		
89	N			53.60	4.75		
90	E	59.76		56.70	4.29	3.06	
91	D	53.36	4.49	54.10	4.76	-0.74	-0.27
92	G	44.98	4.68	45.00	3.97	-0.02	0.71
93	S	58.64	4.27	58.30	4.50	0.34	-0.24
94	L	54.39	4.73	55.70	4.17	-1.31	0.56
95	Q	55.95	4.01	56.20	4.37	-0.25	-0.36
96	K	57.77		56.70	4.36	1.07	
97	K	56.50	4.76	56.70	4.36	-0.20	0.40
98	L	55.24	5.30	55.70	4.17	-0.46	1.13
99	K	56.02	5.46	56.70	4.36	-0.69	1.10
100	V	63.40	3.24	63.00	3.95	0.40	-0.71
101	W	57.63	5.05	57.80	4.70	-0.17	0.35
102	F			57.90	4.66		
103	R	56.69		56.30	4.38	0.39	-4.38
104	I	58.49		62.60	3.95	-4.11	
105	P			62.90	4.44		
106	N	57.21		53.60	4.75	3.61	
107	Q	58.21	4.68	56.20	4.37	2.01	0.31
108	F	57.56	4.58	57.90	4.66	-0.34	-0.08
109	Q	59.56		56.20	4.37	3.36	
110	S	54.20		58.30	4.50	4.10	4.50
111	D	63.61	4.06	54.10	4.76	-9.51	0.70
112	P			62.90	4.44		
113	P			62.90	4.44		
114	A	50.34		52.50	4.35	-2.16	
115	P			62.90	4.44		
116	S	59.65	4.69	58.30	4.50	1.35	0.19
117	D	52.91	6.93	54.10	4.76	-1.19	2.17
118	K	57.38	4.16	56.70	4.36	0.68	-0.20

Residue		**CA (exp)**	**HA (exp**	**CA (ref)**	**HA (ref)**	$\Delta\delta_{CA}$	$\Delta\delta_{HA}$
119	S	60.33	4.28	58.30	4.50	2.03	-0.22
120	V	61.03	4.44	63.00	3.95	-1.97	0.49
121	K	53.79	4.52	56.70	4.36	-2.91	0.16
122	I	57.70	5.33	62.60	3.95	-4.90	1.38
123	E	55.12	4.88	56.70	4.29	-1.58	0.59
124	E	54.91	5.50	56.70	4.29	-1.79	1.21
125	R	56.00	4.56	56.30	4.38	-0.30	0.18
126	E			56.70	4.29		
127	G	46.52	4.70	45.00	3.97	1.52	0.73
128	I	59.47	4.67	62.60	3.95	-3.14	0.72
129	T	61.18	4.99	63.10	4.35	-1.92	0.64
130	V	56.79	5.28	63.00	3.95	-6.21	1.33
131	Y	57.01	4.99	58.60	4.60	-1.59	0.39
132	S	55.90	6.09	58.30	4.50	-2.40	1.59
133	M	54.70	5.00	56.60	4.52	-1.90	0.48
134	Q	54.48	3.79	56.20	4.37	-1.72	-0.59
135	F	56.67	5.00	57.90	4.66	-1.23	0.34
136	G	43.97		45.00	3.97	-1.03	
137	G	44.29		45.00	3.97	-0.71	
138	Y	57.45		58.60	4.60	-1.15	
139	A	52.26		52.50	4.35	-0.24	
140	K	54.12	4.84	56.70	4.36	-2.58	0.48
141	E	62.04		56.70	4.29	5.34	
142	A	55.07		52.50	4.35	2.57	
143	D	57.13	4.43	54.10	4.76	3.03	-0.33
144	Y	61.96		58.60	4.60	3.36	
145	V	66.18	3.83	63.00	3.95	3.18	-0.12
146	A	55.36		52.50	4.35	2.86	
147	Q	57.71	1.39	56.20	4.37	1.51	
148	A	55.79		52.50	4.35	3.29	
149	T	66.72		63.10	4.35	3.62	
150	R	59.34		56.30	4.38	3.04	
151	L	57.85	4.40	55.70	4.17	2.15	0.23
152	R	60.81		56.30	4.38	4.51	
153	A	54.77	4.09	52.50	4.35	2.27	-0.26
154	A	54.09	4.04	52.50	4.35	1.59	-0.31
155	L	54.51	3.81	55.70	4.17	-1.19	-0.36
156	E	58.28		56.70	4.29	1.58	

Residue		CA (exp)	HA (exp	CA (ref)	HA (ref)	$\Delta\delta_{CA}$	$\Delta\delta_{HA}$
157	G	45.54		45.00	3.97	0.54	-3.97
158	T	63.09	4.37	63.10	4.35	-0.01	0.02
159	A	51.38	4.64	52.50	4.35	-1.12	0.29
160	T	60.14	4.30	63.10	4.35	-2.96	-0.05
161	Y	52.45	4.78	58.60	4.60	-6.15	0.18
162	R	58.41		56.30	4.38	2.11	
163	G	45.80		45.00	3.97	0.80	
164	D	53.59	4.66	54.10	4.76	-0.51	-0.10
165	I	59.43	4.69	62.60	3.95	-3.17	0.74
166	Y	56.85	4.31	58.60	4.60	-1.75	-0.29
167	F	55.08	5.74	57.90	4.66	-2.82	1.08
168	C	55.48	5.06	58.80	4.65	-3.33	0.41
169	T	58.44		63.10	4.35	-4.66	
170	G	45.23		45.00	3.97	0.23	-3.97
171	Y	55.57	4.21	58.60	4.60	-3.04	-0.39
172	D			54.10	4.76		
173	P			62.90	4.44		
174	P			62.90	4.44		
175	M			56.60	4.52		
176	K			56.70	4.36		
177	P			62.90	4.44		
178	Y	55.69		58.60	4.60	-2.91	
179	G	44.37		45.00	3.97	-0.63	
180	R			56.30	4.38		
181	R	55.92		56.30	4.38	-0.38	
182	N	50.42	5.42	54.10	4.76	-3.68	0.66
183	E	54.57	5.81	56.70	4.29	-2.13	1.52
184	I	58.20	4.85	62.60	3.95	-4.40	0.90
185	W	53.10	5.79	57.80	4.70	-4.70	1.09
186	L	53.19	5.01	55.70	4.17	-2.51	0.84
187	L	55.16	4.86	55.70	4.17	-0.54	0.69
188	K	58.01	3.72	56.70	4.36	1.31	-0.64
189	T	62.78	4.12	63.10	4.35	-0.33	-0.23

9.4 Talos prediction for human p22HBP

residue	S^2	secondary structure	prediction
13 V	0.00	X	None
14 E	0.00	X	None
15 T	0.00	X	None
16 W	0.00	X	None
17 P	0.00	X	None
18 W	0.32	Loop	Warn
19 Q	0.59	Sheet	Good
20 V	0.92	Sheet	Good
21 L	0.97	Sheet	Good
22 S	0.81	Sheet	Good
23 K	0.73	Sheet	Good
24 G	0.76	Sheet	Warn
25 D	0.69	Sheet	Good
26 K	0.56	Loop	Good
27 E	0.54	Loop	Warn
28 E	0.46	Loop	Warn
29 V	0.30	Loop	Warn
30 A	0.80	Sheet	Good
31 Y	0.98	Sheet	Good
32 E	1.00	Sheet	Good
33 E	0.97	Sheet	Good
34 R	0.91	Sheet	Good
35 A	0.64	Sheet	Good
36 C	0.63	Loop	Good
37 E	0.90	Loop	Warn
38 G	0.90	Loop	Warn
39 G	0.53	Loop	Warn
40 K	0.40	Sheet	Good
41 F	0.70	Sheet	Good
42 A	0.88	Sheet	Good
43 T	0.97	Sheet	Good
44 V	0.96	Sheet	Good
45 E	0.98	Sheet	Good
46 V	0.89	Sheet	Good
47 T	0.24	Sheet	Warn
48 D	0.70	Loop	Good
49 K	0.73	Loop	Good
50 P	0.59	Loop	Good
51 V	0.88	Helix	Good
52 D	0.99	Helix	Good
53 E	1.00	Helix	Good
54 A	1.00	Helix	Good
55 L	0.99	Helix	Good
56 R	0.97	Helix	Good
57 E	0.95	Helix	Good
58 A	0.94	Helix	Good
59 M	0.94	Helix	Good
60 P	0.96	Helix	Good
61 K	0.95	Helix	Good
62 V	0.98	Helix	Good
63 A	0.97	Helix	Good
64 K	0.98	Helix	Good
65 Y	0.93	Helix	Good
66 A	0.50	Helix	Good
67 G	0.64	Loop	Good
68 G	0.86	Loop	Warn
69 T	0.79	Loop	Good
70 N	0.88	Loop	Warn
71 D	0.94	Loop	Good
72 K	0.84	Loop	Good
73 G	0.59	Loop	Good
74 I	0.40	Sheet	Good
75 G	0.43	Sheet	Warn
76 M	0.67	Loop	Warn
77 G	0.84	Loop	Warn
78 M	0.66	Loop	Good
79 T	0.39	Loop	Good
80 V	0.66	Loop	Warn
81 P	0.00	X	None
82 I	0.00	X	None

residue	**S^2**	**secondary structure**	**prediction**
83 S	0.65	Loop	Good
84 F	0.63	Sheet	Warn
85 A	0.92	Sheet	Good
86 V	0.71	Sheet	Good
87 F	0.40	Sheet	Warn
88 P	0.00	X	None
89 N	0.00	X	None
90 E	0.00	X	None
91 D	0.91	Loop	Good
92 G	0.95	Loop	Good
93 S	0.91	Loop	Warn
94 L	0.79	Loop	Warn
95 Q	0.78	Loop	Good
96 K	0.45	Loop	Warn
97 K	0.76	Sheet	Good
98 L	0.86	Sheet	Good
99 K	0.89	Sheet	Good
100 V	0.75	Sheet	Good
101 W	0.73	Sheet	Warn
102 F	0.00	X	None
103 R	0.00	X	None
104 I	0.00	X	None
105 P	0.00	X	None
106 N	0.62	Helix	Good
107 Q	0.86	Helix	Good
108 F	0.64	Helix	Good
109 Q	0.46	Helix	Good
110 S	0.00	X	None
111 D	0.00	X	None
112 P	0.00	X	None
113 P	0.00	X	None
114 A	0.00	X	None
115 P	0.98	Loop	Good
116 S	0.99	Loop	Warn
117 D	0.94	Loop	Good
118 K	0.78	Loop	Good
119 S	0.61	Loop	Good

residue	**S^2**	**secondary structure**	**prediction**
120 V	0.43	Loop	Warn
121 K	0.71	Sheet	Good
122 I	0.88	Sheet	Good
123 E	0.82	Sheet	Good
124 E	0.57	Sheet	Good
125 R	0.83	Sheet	Good
126 E	0.15	Sheet	Good
127 G	0.46	Loop	Warn
128 I	0.60	Sheet	Good
129 T	0.89	Sheet	Good
130 V	0.94	Sheet	Good
131 Y	0.91	Sheet	Good
132 S	0.98	Sheet	Good
133 M	1.00	Sheet	Good
134 Q	1.00	Sheet	Good
135 F	0.84	Sheet	Good
136 G	0.15	Sheet	Warn
137 G	0.41	Loop	Warn
138 Y	0.38	Sheet	Good
139 A	0.18	Loop	Good
140 K	0.69	Loop	Good
141 E	0.71	Helix	Good
142 A	0.98	Helix	Good
143 D	0.96	Helix	Good
144 Y	1.00	Helix	Good
145 V	1.00	Helix	Good
146 A	1.00	Helix	Good
147 Q	1.00	Helix	Good
148 A	1.00	Helix	Good
149 T	1.00	Helix	Good
150 R	1.00	Helix	Good
151 L	1.00	Helix	Good
152 R	1.00	Helix	Good
153 A	0.99	Helix	Good
154 A	0.91	Helix	Good
155 L	0.49	Helix	Good
156 E	0.07	Loop	Warn

residue		**S^2**	**secondary structure**	**prediction**
157	X	0.00	X	None
158	T	0.21	Loop	Warn
159	A	0.57	Sheet	Good
160	T	0.78	Sheet	Good
161	Y	0.70	Sheet	Good
162	R	0.00	X	None
163	G	0.00	X	None
164	D	0.87	Loop	Good
165	I	0.94	Loop	Warn
166	Y	0.25	Sheet	Warn
167	F	0.00	X	None
168	C	0.00	X	None
169	T	0.00	X	None
170	G	0.00	X	None
171	Y	0.00	X	None
172	D	0.00	X	None
173	P	0.00	X	None
174	P	0.00	X	None
175	M	0.00	X	None
176	K	0.00	X	None
177	P	0.00	X	None
178	Y	0.00	X	None
179	G	0.00	X	None
180	R	0.00	X	None
181	R	0.00	X	None
182	N	0.00	X	None
183	E	0.29	Loop	Warn
184	I	0.73	Sheet	Good
185	W	0.95	Sheet	Good
186	L	0.90	Sheet	Good
187	L	0.52	Sheet	Good
188	K	0.38	Loop	Good
189	T	0.03	Sheet	Good
190	V	0.00	X	None

9.5 Longitudinal and transverse relaxation rates, ^{1}H-^{15}N steady state NOE values and R_2/R_1 ratios for human p22HBP at 500 MHz

residue	hetNOE	R_1 (1/s)	R_2 (1/s)	R_2/R_1
16 W	0.721 ± 0.065	1.299 ± 0.191	17.299 ± 1.111	13.317 ± 0.651
17 P				
18 W	0.695 ± 0.037	1.336 ± 0.189	16.825 ± 0.659	12.598 ± 0.424
19 Q	0.683 ± 0.049	1.410 ± 0.186	16.300 ± 0.847	11.564 ± 0.517
20 V	0.662 ± 0.072	1.342 ± 0.277	15.225 ± 1.591	11.341 ± 0.934
21 L	0.864 ± 0.070	1.474 ± 0.195	16.400 ± 1.277	11.124 ± 0.736
22 S	0.859 ± 0.041	1.288 ± 0.198	15.473 ± 0.489	12.013 ± 0.343
23 K	0.769 ± 0.039	1.343 ± 0.252	15.730 ± 0.637	11.716 ± 0.444
24 G	0.676 ± 0.033	1.359 ± 0.184	14.511 ± 0.480	10.681 ± 0.332
25 D	0.574 ± 0.033	1.300 ± 0.305	14.167 ± 0.525	10.895 ± 0.415
26 K	0.662 ± 0.032	1.347 ± 0.195	14.216 ± 0.374	10.553 ± 0.284
27 E	0.731 ± 0.097	1.115 ± 0.274	14.064 ± 1.336	12.609 ± 0.805
28 E	0.658 ± 0.051	1.452 ± 0.335	17.754 ± 1.119	12.231 ± 0.727
29 V	0.690 ± 0.017	1.364 ± 0.156	15.551 ± 0.240	11.398 ± 0.198
30 A	0.729 ± 0.035	1.375 ± 0.210	16.371 ± 0.543	11.907 ± 0.377
31 Y	0.805 ± 0.027	1.337 ± 0.175	16.351 ± 0.373	12.231 ± 0.274
32 E	0.757 ± 0.050	1.322 ± 0.154	16.088 ± 0.843	12.171 ± 0.498
33 E	0.720 ± 0.031	1.361 ± 0.153	15.192 ± 0.485	11.165 ± 0.319
34 R	0.806 ± 0.023	1.363 ± 0.176	15.339 ± 0.354	11.255 ± 0.265
35 A	0.778 ± 0.047	1.351 ± 0.181	15.054 ± 0.668	11.140 ± 0.424
36 C	0.699 ± 0.041	1.308 ± 0.196	14.791 ± 0.617	11.306 ± 0.407
37 E	0.661 ± 0.048	1.342 ± 0.239	16.423 ± 0.874	12.238 ± 0.556
38 G	0.908 ± 0.053	1.441 ± 0.264	14.687 ± 0.662	10.196 ± 0.463
39 G	0.828 ± 0.046	1.328 ± 0.201	18.520 ± 0.962	13.942 ± 0.582
40 K		1.392 ± 0.270	16.860 ± 1.153	12.109 ± 0.712
41 F	0.850 ± 0.055	1.514 ± 0.155	16.435 ± 0.643	10.858 ± 0.399
42 A	0.834 ± 0.062	1.355 ± 0.167	16.751 ± 1.029	12.364 ± 0.598
43 T	0.816 ± 0.049	1.344 ± 0.151	17.418 ± 0.846	12.958 ± 0.499
44 V	0.698 ± 0.054	1.197 ± 0.161	16.820 ± 1.038	14.050 ± 0.599
45 E	0.820 ± 0.044	1.379 ± 0.213	18.444 ± 0.749	13.375 ± 0.481
46 V	0.776 ± 0.044	1.284 ± 0.092	15.485 ± 0.623	12.058 ± 0.357
47 T	0.814 ± 0.059	1.370 ± 0.319	16.404 ± 0.876	11.972 ± 0.597
48 D	0.760 ± 0.029	1.330 ± 0.143	16.089 ± 0.439	12.098 ± 0.291
49 K	0.755 ± 0.033	1.390 ± 0.170	14.934 ± 0.458	10.746 ± 0.314

residue	**hetNOE**	**R_1 (1/s)**	**R_2 (1/s)**	**R_2/R_1**
50 P				
51 V	0.790 ± 0.046	1.294 ± 0.161	17.527 ± 0.737	13.550 ± 0.449
52 D	0.753 ± 0.037	1.317 ± 0.117	18.133 ± 0.630	13.771 ± 0.374
53 E	0.684 ± 0.026	1.273 ± 0.122	16.879 ± 0.412	13.262 ± 0.267
54 A	0.772 ± 0.033	1.326 ± 0.111	17.417 ± 0.541	13.138 ± 0.326
55 L	0.697 ± 0.023	1.339 ± 0.161	18.645 ± 0.431	13.923 ± 0.296
56 R	0.682 ± 0.047	1.301 ± 0.114	18.077 ± 0.866	13.890 ± 0.490
57 E	0.758 ± 0.036	1.294 ± 0.121	17.433 ± 0.602	13.473 ± 0.361
58 A	0.813 ± 0.028	1.350 ± 0.179	17.317 ± 0.459	12.823 ± 0.319
59 M	0.709 ± 0.035	1.354 ± 0.158	16.307 ± 0.614	12.042 ± 0.386
60 P				
61 K	0.842 ± 0.061	1.233 ± 0.127	15.824 ± 0.907	12.836 ± 0.517
62 V	0.789 ± 0.025	1.408 ± 0.153	17.139 ± 0.350	12.169 ± 0.252
63 A	0.814 ± 0.052	1.345 ± 0.222	19.223 ± 1.148	14.290 ± 0.685
64 K	0.830 ± 0.051	1.290 ± 0.159	16.876 ± 0.811	13.082 ± 0.485
65 Y	0.823 ± 0.042	1.404 ± 0.190	17.730 ± 0.815	12.625 ± 0.503
66 A	0.719 ± 0.036	1.411 ± 0.173	16.042 ± 0.675	11.370 ± 0.424
67 G	0.734 ± 0.039	1.378 ± 0.144	18.103 ± 0.898	13.141 ± 0.521
68 G	0.762 ± 0.038	1.244 ± 0.180	18.032 ± 0.656	14.496 ± 0.418
69 T	0.820 ± 0.038	1.344 ± 0.177	16.461 ± 0.570	12.249 ± 0.373
70 N	0.720 ± 0.022	1.298 ± 0.189	16.425 ± 0.349	12.651 ± 0.269
71 D	0.905 ± 0.033	1.369 ± 0.107	17.216 ± 0.440	12.580 ± 0.273
72 K	0.712 ± 0.036	1.201 ± 0.157	15.245 ± 0.591	12.696 ± 0.374
73 G	0.687 ± 0.024	1.333 ± 0.155	16.710 ± 0.489	12.532 ± 0.322
74 I	0.728 ± 0.031	1.341 ± 0.156	16.566 ± 0.521	12.355 ± 0.339
75 G	0.776 ± 0.083	1.163 ± 0.331	15.564 ± 1.347	13.378 ± 0.839
76 M	0.686 ± 0.053	1.407 ± 0.165	14.794 ± 0.684	10.517 ± 0.425
77 G	0.621 ± 0.046	1.445 ± 0.209	14.638 ± 0.818	10.127 ± 0.513
78 M	0.975 ± 0.035	1.396 ± 0.146	15.835 ± 0.371	11.342 ± 0.258
79 T	0.796 ± 0.051	1.374 ± 0.138	18.576 ± 1.160	13.521 ± 0.649
80 V	0.815 ± 0.126	1.357 ± 0.297	27.749 ± 0.000	20.443 ± 0.148
81 P				
82 I	0.456 ± 0.159	1.283 ± 0.563	24.662 ± 10.201	19.224 ± 5.382
83 S	0.655 ± 0.193	1.976 ± 1.002	14.745 ± 5.845	7.462 ± 3.423
84 F	0.696 ± 0.016	1.294 ± 0.222	15.449 ± 0.238	11.938 ± 0.230
85 A	0.762 ± 0.059	1.347 ± 0.224	15.967 ± 0.999	11.858 ± 0.612
86 V	0.816 ± 0.051	1.448 ± 0.142	15.332 ± 0.670	10.588 ± 0.406
87 F	0.735 ± 0.057	1.299 ± 0.220	16.302 ± 0.943	12.554 ± 0.581

residue	**hetNOE**	**R_1 (1/s)**	**R_2 (1/s)**	**R_2/R_1**
88 P				
89 N				
90 E				
91 D	0.697 ± 0.063	1.327 ± 0.346	14.801 ± 1.089	11.151 ± 0.718
92 G	0.680 ± 0.032	1.436 ± 0.291	14.437 ± 0.424	10.054 ± 0.358
93 S	0.724 ± 0.025	1.301 ± 0.240	15.632 ± 0.358	12.015 ± 0.299
94 L	0.713 ± 0.043	1.378 ± 0.156	15.357 ± 0.639	11.144 ± 0.398
95 Q	0.754 ± 0.049	1.345 ± 0.194	16.232 ± 0.932	12.066 ± 0.563
96 K				
97 K	0.762 ± 0.032	1.318 ± 0.215	15.816 ± 0.462	12.001 ± 0.338
98 L	0.797 ± 0.041	1.319 ± 0.179	16.042 ± 0.551	12.158 ± 0.365
99 K	0.860 ± 0.064	1.304 ± 0.100	15.540 ± 0.890	11.921 ± 0.495
100 V	0.777 ± 0.067	1.378 ± 0.190	17.732 ± 1.378	12.866 ± 0.784
101 W	0.897 ± 0.037	1.353 ± 0.183	13.769 ± 0.409	10.173 ± 0.296
102 F				
103 R				
104 I	0.801 ± 0.086	1.478 ± 0.171	21.094 ± 0.737	14.269 ± 0.454
105 P				
106 N	0.882 ± 0.084	1.294 ± 0.200	16.811 ± 1.297	12.997 ± 0.748
107 Q	0.814 ± 0.085	1.671 ± 0.571	18.433 ± 1.873	11.029 ± 1.222
108 F	0.778 ± 0.032	1.333 ± 0.156	16.201 ± 0.493	12.155 ± 0.325
109 Q	0.893 ± 0.034	1.382 ± 0.146	16.528 ± 0.432	11.956 ± 0.289
110 S				
111 D				
112 P				
113 P				
114 A	0.835 ± 0.036	1.385 ± 0.128	17.600 ± 0.476	12.712 ± 0.302
115 P				
116 S	0.598 0.044	1.360 0.242	15.092 1.161	11.098 0.702
117 D	0.704 ± 0.036	1.365 ± 0.232	13.540 ± 0.469	9.921 ± 0.351
118 K	0.680 ± 0.074	1.228 ± 0.275	14.000 ± 1.296	11.397 ± 0.785
119 S	0.765 ± 0.040	1.427 ± 0.287	16.871 ± 0.651	11.823 ± 0.469
120 V	0.755 ± 0.032	1.290 ± 0.178	15.557 ± 0.399	12.057 ± 0.289
121 K	0.759 ± 0.044	1.371 ± 0.113	15.418 ± 0.725	11.244 ± 0.419
122 I	0.687 ± 0.019	1.296 ± 0.184	15.293 ± 0.297	11.799 ± 0.241
123 E	0.816 ± 0.045	1.386 ± 0.167	15.010 ± 0.674	10.828 ± 0.420
124 E	0.735 ± 0.029	1.345 ± 0.151	14.928 ± 0.360	11.097 ± 0.255
125 R	0.782 ± 0.048	1.290 ± 0.188	15.717 ± 0.679	12.188 ± 0.434

residue	**hetNOE**	**R_1 (1/s)**	**R_2 (1/s)**	**R_2/R_1**
126 E				
127 G	0.693 ± 0.043	1.527 ± 0.334	14.345 ± 0.679	9.397 ± 0.506
128 I	0.797 ± 0.047	1.300 ± 0.218	16.921 ± 0.792	13.017 ± 0.505
129 T	0.715 ± 0.023	1.320 ± 0.190	15.048 ± 0.350	11.399 ± 0.270
130 V	0.800 ± 0.043	1.445 ± 0.159	17.247 ± 0.679	11.935 ± 0.419
131 Y	0.795 ± 0.050	1.444 ± 0.147	16.301 ± 0.781	11.288 ± 0.464
132 S	0.704 ± 0.041	1.410 ± 0.178	15.651 ± 0.729	11.099 ± 0.453
133 M	0.827 ± 0.070	1.392 ± 0.149	15.589 ± 1.067	11.199 ± 0.608
134 Q	0.722 ± 0.035	1.418 ± 0.210	15.766 ± 0.573	11.118 ± 0.392
135 F	0.857 ± 0.074	1.247 ± 0.259	19.783 ± 0.966	15.866 ± 0.613
136 G				
137 G	0.588 ± 0.093	1.168 ± 0.402	20.464 ± 1.325	17.514 ± 0.864
138 Y				
139 A	0.698 ± 0.054	1.406 ± 0.350	17.009 ± 1.042	12.095 ± 0.696
140 K	0.794 ± 0.078	1.192 ± 0.165	15.484 ± 1.160	12.987 ± 0.662
141 E	0.736 ± 0.089	1.300 ± 0.277	16.270 ± 1.633	12.511 ± 0.955
142 A	0.738 ± 0.030	1.403 ± 0.127	17.206 ± 0.474	12.266 ± 0.300
143 D	0.784 ± 0.031	1.333 ± 0.156	17.380 ± 0.535	13.041 ± 0.345
144 Y	0.774 ± 0.047	1.414 ± 0.133	16.951 ± 0.812	11.985 ± 0.472
145 V	0.755 ± 0.048	1.374 ± 0.134	16.936 ± 0.821	12.328 ± 0.477
146 A	0.767 ± 0.040	1.367 ± 0.132	16.125 ± 0.665	11.799 ± 0.398
147 Q	0.836 ± 0.057	1.444 ± 0.192	17.063 ± 0.883	11.820 ± 0.537
148 A	0.676 ± 0.017	1.362 ± 0.151	15.315 ± 0.230	11.243 ± 0.190
149 T	0.739 ± 0.035	1.404 ± 0.212	16.742 ± 0.611	11.924 ± 0.412
150 R	0.956 ± 0.043	1.448 ± 0.156	17.288 ± 0.550	11.940 ± 0.353
151 L	0.829 ± 0.040	1.359 ± 0.188	17.215 ± 0.660	12.667 ± 0.424
152 R	0.653 ± 0.035	1.426 ± 0.193	16.261 ± 0.586	11.399 ± 0.390
153 A	0.784 ± 0.029	1.363 ± 0.163	18.097 ± 0.501	13.277 ± 0.332
154 A	0.772 ± 0.028	1.309 ± 0.130	17.142 ± 0.454	13.093 ± 0.292
155 L	0.764 ± 0.037	1.410 ± 0.155	15.166 ± 0.564	10.755 ± 0.360
156 E	0.755 ± 0.019	1.392 ± 0.130	15.384 ± 0.266	11.053 ± 0.198
157 G				
158 T	0.704 ± 0.061	1.398 ± 0.349	14.054 ± 0.887	10.052 ± 0.618
159 A	0.670 ± 0.026	1.336 ± 0.255	12.562 ± 0.288	9.401 ± 0.271
160 T	0.610 ± 0.036	1.285 ± 0.316	12.552 ± 0.413	9.766 ± 0.364
161 Y	0.668 ± 0.015	1.294 ± 0.224	15.434 ± 0.231	11.932 ± 0.228
162 R				
163 G	0.804 ± 0.109	1.423 ± 0.314	21.061 ± 0.975	14.796 ± 0.645

residue	hetNOE	R_1 (1/s)	R_2 (1/s)	R_2/R_1
164 D	0.781 ± 0.046	1.569 ± 0.303	16.977 ± 0.742	10.817 ± 0.523
165 I	0.769 ± 0.018	1.382 ± 0.156	17.232 ± 0.297	12.473 ± 0.226
166 Y	0.825 ± 0.034	1.393 ± 0.177	17.056 ± 0.581	12.247 ± 0.379
167 F	0.840 ± 0.063	1.363 ± 0.265	17.875 ± 1.172	13.118 ± 0.719
168 C	0.771 ± 0.045	1.368 ± 0.168	17.169 ± 0.736	12.547 ± 0.452
169 T	1.028 ± 0.060	1.343 ± 0.178	18.595 ± 0.912	13.843 ± 0.545
170 G				
171 Y				
172 D				
173 P				
174 P				
175 M				
176 K				
177 P				
178 Y	0.644 ± 0.078	1.312 ± 0.376	18.369 ± 1.987	14.002 ± 1.181
179 G	0.394 ± 0.084	1.430 ± 0.466	14.326 ± 2.179	10.021 ± 1.323
180 R				
181 R				
182 N	0.653 ± 0.037	1.417 ± 0.252	14.926 ± 0.633	10.535 ± 0.443
183 E	0.770 ± 0.063	1.296 ± 0.172	16.773 ± 1.261	12.938 ± 0.717
184 I	0.757 ± 0.049	1.417 ± 0.197	16.767 ± 0.732	11.828 ± 0.465
185 W	0.773 ± 0.051	1.340 ± 0.172	15.353 ± 0.829	11.456 ± 0.501
186 L	0.852 ± 0.071	1.379 ± 0.337	17.389 ± 1.127	12.607 ± 0.732
187 L	0.848 ± 0.062	1.348 ± 0.172	14.826 ± 0.935	10.995 ± 0.554
188 K	0.829 ± 0.083	1.278 ± 0.212	16.704 ± 1.314	13.070 ± 0.763
189 T	0.696 ± 0.034	1.332 ± 0.151	14.114 ± 0.473	10.593 ± 0.312

9.6 Longitudinal and transverse relaxation rates, ^{1}H-^{15}N steady state NOE values and R_2/R_1 ratios for human p22HBP at 700 MHz

residue	hetNOE	R_1 (1/s)	R_2(1/s)	R_2/R_1
16 W	0.799 ± 0.065	0.858 ± 0.291	19.307 ± 1.362	22.495 ± 0.826
17 P				
18 W	1.014 ± 0.039	0.966 ± 0.241	19.276 ± 0.570	19.945 ± 0.406
19 Q	0.796 ± 0.056	0.898 ± 0.233	18.536 ± 0.936	20.649 ± 0.585
20 V	0.907 ± 0.124	0.915 ± 0.341	20.517 ± 1.053	22.419 ± 0.697
21 L	0.660 ± 0.069	0.957 ± 0.259	16.665 ± 1.238	17.417 ± 0.748
22 S	0.719 ± 0.030	0.916 ± 0.188	17.318 ± 0.464	18.914 ± 0.326
23 K	0.826 ± 0.052	0.864 ± 0.232	18.301 ± 0.674	21.174 ± 0.453
24 G	0.733 ± 0.027	0.952 ± 0.273	18.245 ± 0.504	19.167 ± 0.388
25 D	0.738 ± 0.025	0.945 ± 0.357	16.635 ± 0.431	17.609 ± 0.394
26 K	0.519 ± 0.024	0.969 ± 0.234	14.390 ± 0.347	14.848 ± 0.291
27 E	0.368 ± 0.157	0.793 ± 0.438	14.248 ± 1.828	17.969 ± 1.133
28 E	0.587 ± 0.046	1.014 ± 0.331	23.555 ± 0.000	23.235 ± 0.166
29 V	0.338 ± 0.011	0.969 ± 0.193	16.662 ± 0.206	17.194 ± 0.200
30 A	0.705 ± 0.039	0.945 ± 0.211	16.585 ± 0.543	17.552 ± 0.377
31 Y	0.752 ± 0.024	0.906 ± 0.167	17.869 ± 0.394	19.715 ± 0.281
32 E	0.757 ± 0.036	0.945 ± 0.221	18.611 ± 0.721	19.697 ± 0.471
33 E	0.820 ± 0.026	0.923 ± 0.181	19.544 ± 0.486	21.175 ± 0.334
34 R	0.819 ± 0.020	0.940 ± 0.219	17.618 ± 0.362	18.733 ± 0.291
35 A	0.834 ± 0.038	0.928 ± 0.221	18.822 ± 0.657	20.292 ± 0.439
36 C	0.797 ± 0.049	0.924 ± 0.254	17.167 ± 0.718	18.570 ± 0.486
37 E	0.820 ± 0.047	0.978 ± 0.293	17.758 ± 0.929	18.156 ± 0.611
38 G	0.901 ± 0.034	0.967 ± 0.278	16.373 ± 0.585	16.929 ± 0.432
39 G	0.803 ± 0.038	0.945 ± 0.286	23.188 ± 0.000	24.545 ± 0.143
40 K	0.810 ± 0.058	0.834 ± 0.191	19.001 ± 1.204	22.792 ± 0.697
41 F	0.705 ± 0.028	0.949 ± 0.167	17.555 ± 0.633	18.496 ± 0.400
42 A	0.761 ± 0.049	0.905 ± 0.161	17.820 ± 0.983	19.701 ± 0.572
43 T	0.930 ± 0.048	0.945 ± 0.159	18.031 ± 0.726	19.089 ± 0.442
44 V	0.704 ± 0.044	0.879 ± 0.151	17.342 ± 0.954	19.735 ± 0.552
45 E	0.850 ± 0.035	0.929 ± 0.157	19.140 ± 0.709	20.595 ± 0.433
46 V	0.766 ± 0.037	0.908 ± 0.149	19.191 ± 0.669	21.136 ± 0.409
47 T	0.968 ± 0.063	0.917 ± 0.362	19.038 ± 0.930	20.763 ± 0.646
48 D	0.792 ± 0.034	0.916 ± 0.225	17.198 ± 0.412	18.782 ± 0.319
49 K	0.770 ± 0.031	0.962 ± 0.182	16.632 ± 0.424	17.282 ± 0.303

residue	**hetNOE**	**R_1 (1/s)**	**R_2(1/s)**	**R_2/R_1**
50 P				
51 V	0.839 ± 0.046	0.881 ± 0.158	18.974 ± 0.821	21.537 ± 0.490
52 D	0.784 ± 0.045	0.973 ± 0.167	20.820 ± 0.065	21.388 ± 0.116
53 E	0.611 ± 0.025	0.916 ± 0.126	17.850 ± 0.432	19.497 ± 0.279
54 A	0.840 ± 0.028	0.942 ± 0.144	19.862 ± 0.386	21.092 ± 0.265
55 L	0.819 ± 0.024	0.951 ± 0.190	20.342 ± 0.066	21.401 ± 0.128
56 R	0.672 ± 0.036	0.906 ± 0.153	±	±
57 E	0.797 ± 0.038	0.885 ± 0.141	19.072 ± 0.550	21.561 ± 0.345
58 A	0.746 ± 0.023	0.929 ± 0.182	18.464 ± 0.398	19.869 ± 0.290
59 M	0.997 ± 0.068	0.939 ± 0.169	17.885 ± 0.769	19.050 ± 0.469
60 P				
61 K	0.746 ± 0.046	0.909 ± 0.147	19.387 ± 0.974	21.316 ± 0.560
62 V	0.929 ± 0.026	0.972 ± 0.206	19.671 ± 0.343	20.247 ± 0.274
63 A	0.857 ± 0.072	0.888 ± 0.195	17.864 ± 1.664	20.128 ± 0.930
64 K	0.829 ± 0.059	0.957 ± 0.159	18.961 ± 0.915	19.812 ± 0.537
65 Y	0.834 ± 0.035	0.885 ± 0.200	18.063 ± 0.611	20.403 ± 0.406
66 A	0.789 ± 0.030	0.968 ± 0.159	20.243 ± 0.231	20.917 ± 0.195
67 G	0.807 ± 0.042	0.910 ± 0.137	16.625 ± 0.896	18.275 ± 0.517
68 G	0.773 ± 0.030	0.847 ± 0.178	18.447 ± 0.584	21.791 ± 0.381
69 T	0.886 ± 0.039	0.893 ± 0.238	17.154 ± 0.657	19.218 ± 0.448
70 N	0.806 ± 0.023	0.832 ± 0.237	17.989 ± 0.382	21.613 ± 0.310
71 D	0.792 ± 0.026	0.920 ± 0.203	18.722 ± 0.498	20.347 ± 0.351
72 K	0.819 ± 0.033	0.837 ± 0.159	17.498 ± 0.454	20.899 ± 0.306
73 G	0.565 ± 0.026	0.926 ± 0.187	18.335 ± 0.500	19.794 ± 0.344
74 I	0.879 ± 0.039	0.955 ± 0.165	18.838 ± 0.526	19.725 ± 0.346
75 G	0.865 ± 0.072	0.854 ± 0.287	18.036 ± 1.231	21.118 ± 0.759
76 M	0.744 ± 0.045	0.958 ± 0.244	16.573 ± 0.802	17.307 ± 0.523
77 G	0.820 ± 0.059	1.057 ± 0.224	19.354 ± 0.966	18.310 ± 0.595
78 M	0.786 ± 0.019	0.971 ± 0.160	17.719 ± 0.358	18.244 ± 0.259
79 T	0.718 ± 0.044	0.966 ± 0.131	18.230 ± 1.333	18.876 ± 0.732
80 V	0.781 ± 0.094	0.911 ± 0.235	30.417 ± 0.000	33.395 ± 0.117
81 P				
82 I		0.980 ± 1.490	27.084 ± 6.514	27.625 ± 4.002
83 S	0.867 ± 0.225	0.961 ± 0.451	17.777 ± 5.178	18.495 ± 2.815
84 F	0.713 ± 0.017	0.912 ± 0.319	17.274 ± 0.253	18.931 ± 0.286
85 A	0.849 ± 0.054	0.901 ± 0.194	20.039 ± 0.646	22.237 ± 0.420
86 V	0.857 ± 0.026	0.941 ± 0.195	17.477 ± 0.465	18.569 ± 0.330
87 F	0.845 ± 0.044	0.856 ± 0.247	17.028 ± 0.819	19.890 ± 0.533

residue	**hetNOE**	**R_1 (1/s)**	**R_2(1/s)**	**R_2/R_1**
88 P				
89 N				
90 E				
91 D	0.582 ± 0.047	0.866 ± 0.346	16.210 ± 0.962	18.728 ± 0.654
92 G	0.726 ± 0.023	1.030 ± 0.333	16.423 ± 0.371	15.950 ± 0.352
93 S	0.647 ± 0.018	0.924 ± 0.328	17.364 ± 0.369	18.784 ± 0.349
94 L	0.820 ± 0.045	0.944 ± 0.230	17.661 ± 0.642	18.705 ± 0.436
95 Q	0.732 ± 0.056	0.920 ± 0.104	16.980 ± 1.027	18.466 ± 0.566
96 K				
97 K	0.701 ± 0.028	0.896 ± 0.235	17.018 ± 0.426	18.997 ± 0.331
98 L	0.798 ± 0.038	0.928 ± 0.146	18.148 ± 0.625	19.554 ± 0.386
99 K	0.811 ± 0.055	0.869 ± 0.159	18.637 ± 0.992	21.452 ± 0.576
100 V	0.862 ± 0.062	0.975 ± 0.238	17.877 ± 1.366	18.345 ± 0.802
101 W	0.789 ± 0.034	0.924 ± 0.236	15.775 ± 0.385	17.067 ± 0.310
102 F				
103 R				
104 I	0.888 ± 0.096	0.949 ± 0.241	22.434 ± 0.175	23.635 ± 0.208
105 P				
106 N	0.781 ± 0.066	1.011 ± 0.345	18.193 ± 1.414	18.002 ± 0.879
107 Q	0.811 ± 0.081	0.959 ± 0.369	19.797 ± 1.307	20.649 ± 0.838
108 F	0.804 ± 0.031	0.913 ± 0.259	17.710 ± 0.489	19.405 ± 0.374
109 Q	0.841 ± 0.035	0.945 ± 0.176	17.798 ± 0.507	18.839 ± 0.341
110 S				
111 D				
112 P				
113 P				
114 A	0.751 ± 0.024	0.969 ± 0.179	18.864 ± 0.427	19.460 ± 0.303
115 P				
116 S	0.884 0.046	0.921 0.308	17.271 0.742	18.747 0.525
117 D	0.737 ± 0.037	0.930 ± 0.309	14.187 ± 0.500	15.249 ± 0.405
118 K	0.561 ± 0.054	0.898 ± 0.304	16.073 ± 1.294	17.898 ± 0.799
119 S	0.779 ± 0.027	1.016 ± 0.381	19.505 ± 0.531	19.192 ± 0.456
120 V	0.801 ± 0.021	0.926 ± 0.219	17.285 ± 0.379	18.675 ± 0.299
121 K	0.833 ± 0.046	0.926 ± 0.194	18.272 ± 0.785	19.734 ± 0.490
122 I	0.709 ± 0.023	0.891 ± 0.181	16.632 ± 0.318	18.666 ± 0.250
123 E	0.798 ± 0.044	0.902 ± 0.169	18.969 ± 0.807	21.021 ± 0.488
124 E	0.756 ± 0.023	0.931 ± 0.193	15.973 ± 0.343	17.155 ± 0.268
125 R	0.763 ± 0.046	0.889 ± 0.233	18.047 ± 0.766	20.292 ± 0.500

residue	hetNOE	R_1 (1/s)	R_2(1/s)	R_2/R_1
126 E				
127 G	0.841 ± 0.039	0.972 ± 0.363	18.060 ± 0.758	18.576 ± 0.560
128 I	0.900 ± 0.045	0.966 ± 0.288	19.398 ± 0.817	20.084 ± 0.552
129 T	0.743 ± 0.026	0.955 ± 0.219	16.875 ± 0.395	17.668 ± 0.307
130 V	0.817 ± 0.037	0.938 ± 0.192	17.260 ± 0.674	18.406 ± 0.433
131 Y	0.811 ± 0.038	0.929 ± 0.187	16.772 ± 0.631	18.056 ± 0.409
132 S	0.752 ± 0.040	0.935 ± 0.219	17.921 ± 0.674	19.166 ± 0.447
133 M	0.771 ± 0.054	1.007 ± 0.211	21.931 ± 0.049	21.772 ± 0.130
134 Q	0.711 ± 0.028	0.947 ± 0.198	21.211 ± 0.000	22.386 ± 0.099
135 F	0.761 ± 0.059	0.910 ± 0.297	22.615 ± 0.000	24.845 ± 0.148
136 G				
137 G	0.849 ± 0.110	1.021 ± 0.618	16.114 ± 1.797	15.788 ± 1.208
138 Y				
139 A	0.572 ± 0.061	1.013 ± 0.369	18.179 ± 1.484	17.944 ± 0.927
140 K	0.893 ± 0.069	0.853 ± 0.179	17.326 ± 1.053	20.304 ± 0.616
141 E	0.814 ± 0.093	0.991 ± 0.299	24.734 ± 0.000	24.951 ± 0.149
142 A	0.893 ± 0.030	0.972 ± 0.167	19.502 ± 0.493	20.074 ± 0.330
143 D	0.818 ± 0.028	0.936 ± 0.183	19.768 ± 0.418	21.129 ± 0.301
144 Y	0.784 ± 0.035	0.910 ± 0.165	18.927 ± 0.687	20.789 ± 0.426
145 V	0.806 ± 0.051	0.999 ± 0.195	19.032 ± 0.824	19.054 ± 0.510
146 A	0.801 ± 0.034	0.977 ± 0.168	18.471 ± 0.564	18.899 ± 0.366
147 Q	0.764 ± 0.064	0.955 ± 0.248	18.853 ± 0.988	19.734 ± 0.618
148 A	0.838 ± 0.016	0.944 ± 0.206	16.946 ± 0.238	17.945 ± 0.222
149 T	0.809 ± 0.028	0.896 ± 0.191	18.540 ± 0.489	20.699 ± 0.340
150 R	0.839 ± 0.027	0.938 ± 0.166	18.250 ± 0.519	19.452 ± 0.343
151 L	0.773 ± 0.040	0.909 ± 0.173	18.551 ± 0.594	20.407 ± 0.383
152 R	0.888 ± 0.033	0.972 ± 0.158	18.543 ± 0.579	19.072 ± 0.369
153 A	0.784 ± 0.030	0.962 ± 0.145	19.651 ± 0.443	20.432 ± 0.294
154 A	0.739 ± 0.022	0.892 ± 0.162	18.413 ± 0.368	20.634 ± 0.265
155 L	0.800 ± 0.036	0.935 ± 0.165	17.084 ± 0.490	18.274 ± 0.328
156 E	0.713 ± 0.018	0.962 ± 0.189	16.678 ± 0.257	17.331 ± 0.223
157 G				
158 T	0.853 ± 0.063	0.999 ± 0.370	15.208 ± 0.764	15.225 ± 0.567
159 A	0.670 ± 0.022	0.976 ± 0.321	14.836 ± 0.298	15.205 ± 0.310
160 T	0.711 ± 0.031	0.945 ± 0.372	14.190 ± 0.400	15.020 ± 0.386
161 Y	0.735 ± 0.016	0.912 ± 0.320	17.185 ± 0.231	18.839 ± 0.276
162 R				
163 G	0.835 ± 0.088	0.889 ± 0.335	22.928 ± 0.255	25.795 ± 0.295

residue	**hetNOE**	**R_1 (1/s)**	**R_2(1/s)**	**R_2/R_1**
164 D	0.721 ± 0.042	1.032 ± 0.314	18.119 ± 0.725	17.553 ± 0.520
165 I	0.809 ± 0.019	0.964 ± 0.210	19.609 ± 0.320	20.341 ± 0.265
166 Y	0.779 ± 0.024	0.964 ± 0.200	18.891 ± 0.446	19.598 ± 0.323
167 F	0.862 ± 0.053	0.882 ± 0.219	19.303 ± 0.990	21.891 ± 0.604
168 C	1.009 ± 0.060	0.872 ± 0.228	19.666 ± 0.724	22.549 ± 0.476
169 T	0.786 ± 0.037	0.943 ± 0.269	23.335 ± 0.000	24.745 ± 0.135
170 G				
171 Y				
172 D				
173 P				
174 P				
175 M				
176 K				
177 P				
178 Y	0.651 ± 0.118	0.982 ± 0.379	19.112 ± 2.011	19.462 ± 1.195
179 G		1.120 ± 0.517	18.586 ± 3.057	16.595 ± 1.787
180 R				
181 R				
182 N	0.607 ± 0.045	0.945 ± 0.296	16.357 ± 0.803	17.311 ± 0.549
183 E	0.888 ± 0.081	0.830 ± 0.269	18.683 ± 1.336	22.522 ± 0.803
184 I	0.848 ± 0.042	0.880 ± 0.208	18.931 ± 0.812	21.501 ± 0.510
185 W	0.843 ± 0.049	0.976 ± 0.244	18.531 ± 0.833	18.993 ± 0.538
186 L	0.793 ± 0.050	0.928 ± 0.121	17.356 ± 0.850	18.701 ± 0.486
187 L	0.770 ± 0.044	0.989 ± 0.274	18.161 ± 0.854	18.361 ± 0.564
188 K	0.778 ± 0.064	0.924 ± 0.284	21.076 ± 0.278	22.802 ± 0.281
189 T	0.689 ± 0.031	0.910 ± 0.235	16.031 ± 0.454	17.608 ± 0.344

9.7 ^{1}H-^{15}N steady state NOE values for PPIX- human p22HBP at 500 MHz

residue		hetNOE		
13	V	0.673	±	0.087
16	W	0.896	±	0.113
17	P			
18	W	0.731	±	0.049
19	Q	0.851	±	0.077
20	V	0.581	±	0.078
21	L	0.894	±	0.108
22	S	0.690	±	0.057
23	K	0.627	±	0.070
24	G	0.652	±	0.052
25	D	0.585	±	0.051
26	K	0.594	±	0.048
27	E	0.724	±	0.038
28	E	0.755	±	0.157
29	V	0.480	±	0.030
30	A	0.759	±	0.050
31	Y	0.788	±	0.040
32	E	1.002	±	0.087
33	E	0.718	±	0.065
34	R	0.723	±	0.038
35	A	0.620	±	0.070
36	C	0.694	±	0.095
37	E	0.810	±	0.073
38	G	0.901	±	0.096
39	G	0.626	±	0.081
40	K	0.512		0.120
41	F	0.731	±	0.045
42	A	0.615	±	0.077
43	T	0.942	±	0.115
44	V	0.433	±	0.118
45	E	0.724	±	0.075
46	V	0.383	±	0.114
47	T	0.771	±	0.037
48	D	0.692	±	0.060
49	K	0.713	±	0.065
50	P			

residue		**hetNOE**		
51	V	0.812	±	0.114
52	D	0.634	±	0.063
53	E	0.724	±	0.035
54	A	0.909	±	0.176
55	L	0.789	±	0.038
56	R	0.605	±	0.071
57	E	0.723	±	0.099
58	A	0.648	±	0.075
59	M		±	
60	P			
61	K	0.624	±	0.191
62	V	0.228	±	0.123
63	A	0.781	±	0.086
64	K		±	
65	Y		±	
66	A	0.522	±	0.121
67	G	0.618	±	0.085
68	G	0.824	±	0.078
69	T	0.853	±	0.067
70	N	0.488	±	0.079
71	D	0.708	±	0.041
72	K	0.652	±	0.064
73	G	0.753	±	0.057
74	I	0.710	±	0.052
75	G	0.693	±	0.126
76	M	0.647	±	0.064
77	G	0.589	±	0.103
78	M	0.671	±	0.142
79	T	0.776	±	0.037
80	V	0.768	±	0.089
81	P			
82	I	0.877	±	0.091
83	S	0.600	±	0.236
84	F	0.387	±	0.183
85	A	0.652	±	0.159
86	V	0.760	±	0.079
87	F	0.817	±	0.094
88	P			

residue		**hetNOE**		
89	N			
90	E			
91	D	0.456	±	0.107
92	G	0.808	±	0.060
93	S	0.616	±	0.035
94	L	0.789	±	0.041
95	Q	0.531	±	0.058
96	K			
97	K	0.778	±	0.067
98	L	0.852	±	0.072
99	K	0.783	±	0.068
100	V	0.914	±	0.118
101	W	0.724	±	0.038
102	F			
103	R	0.684		0.043
104	I	0.719	±	0.172
105	P			
106	N	1.066	±	0.110
107	Q	0.674	±	0.112
108	F	0.623	±	0.037
109	Q	0.828	±	0.069
110	S			
111	D	0.164		0.074
112	P			
113	P			
114	A	0.797	±	0.060
115	P			
116	S	0.591		0.095
117	D	0.181	±	0.061
118	K	0.515	±	0.244
119	S	0.667	±	0.059
120	V	0.542	±	0.043
121	K	0.801	±	0.095
122	I	0.575	±	0.046
123	E	0.754	±	0.049
124	E	0.724	±	0.047
125	R	0.738	±	0.063
126	E			

residue		**hetNOE**		
127	G	0.694	±	0.087
128	I	0.640	±	0.067
129	T	0.456	±	0.072
130	V	0.845	±	0.081
131	Y	0.807	±	0.080
132	S	0.947	±	0.087
133	M	0.675	±	0.085
134	Q	0.714	±	0.066
135	F		±	
136	G			
137	G	0.716	±	0.170
138	Y			
139	A	0.447	±	0.145
140	K	0.792	±	0.153
141	E	0.709	±	0.176
142	A	0.757	±	0.055
143	D	0.740	±	0.047
144	Y		±	
145	V	0.801	±	0.074
146	A	0.704	±	0.081
147	Q	0.684	±	0.071
148	A	0.470	±	0.027
149	T	0.777	±	0.063
150	R	0.712	±	0.052
151	L	0.827	±	0.075
152	R	0.719	±	0.066
153	A	0.670	±	0.046
154	A	0.796	±	0.048
155	L	0.707	±	0.059
156	E	0.708	±	0.032
157	G			
158	T	0.613	±	0.084
159	A	0.532	±	0.042
160	T	0.604	±	0.064
161	Y	0.656	±	0.033
162	R			
163	G		±	
164	D	0.806	±	0.064

residue		**hetNOE**		
165	I	0.737	±	0.035
166	Y	0.840	±	0.055
167	F	0.899	±	0.099
168	C	0.746	±	0.037
169	T	0.532	±	0.066
170	G			
171	Y	0.558		0.057
172	D			
173	P			
174	P			
175	M			
176	K			
177	P			
178	Y	0.583	±	0.082
179	G	0.571	±	0.107
180	R			
181	R	0.839		0.080
182	N	0.423	±	0.100
183	E	0.384	±	0.117
184	I	0.637	±	0.060
185	W	0.780	±	0.073
186	L	0.874	±	0.087
187	L	0.507	±	0.077
188	K	0.637	±	0.105
189	T	0.535	±	0.040

9.8 Longitudinal and transverse relaxation rates, ^{1}H-^{15}N steady state NOE values and R_2/R_1 ratios for PPIX- human p22HBPat 700 MHz

residue	hetNOE	R_1 (1/s)	R_2(1/s)	R_2/R_1
13 V	0.713 ± 0.062	0.946 ± 0.150	14.385 ± 0.880	15.203 ± 0.515
14 E				
15 T				
16 W	0.744 ± 0.072	0.861 ± 0.085	15.297 ± 1.052	17.761 ± 0.568
17 P				
18 W	0.827 ± 0.036	0.957 ± 0.165	15.226 ± 0.444	15.909 ± 0.304
19 Q	0.852 ± 0.055	1.004 ± 2.160	14.463 ± 0.621	14.399 ± 1.390
20 V	0.906 ± 0.122	0.785 ± 0.255	13.437 ± 0.829	17.120 ± 0.542
21 L	0.816 ± 0.093	0.925 ± 0.129	14.164 ± 0.891	15.313 ± 0.510
22 S	0.777 ± 0.038	0.924 ± 0.182	13.184 ± 0.364	14.265 ± 0.273
23 K	0.827 ± 0.074	1.047 ± 0.132	14.713 ± 0.452	14.052 ± 0.292
24 G	0.777 ± 0.036	±	13.564 ± 0.486	± 0.243
25 D	0.654 ± 0.036	0.996 ± 0.226	12.125 ± 0.397	12.168 ± 0.312
26 K	0.626 ± 0.043	1.021 ± 0.284	14.563 ± 0.372	14.258 ± 0.328
27 E	0.762 ± 0.024	0.931 ± 0.161	15.610 ± 1.552	16.770 ± 0.857
28 E	0.563 ± 0.123	0.935 ± 0.278	13.697 ± 0.210	14.646 ± 0.244
29 V	0.646 ± 0.020	0.963 ± 0.141	13.977 ± 0.458	14.521 ± 0.299
30 A	0.835 ± 0.043	1.021 ± 0.076	14.165 ± 0.654	13.870 ± 0.365
31 Y	0.775 ± 0.030	0.921 ± 0.742	14.519 ± 0.787	15.771 ± 0.765
32 E	0.788 ± 0.043	1.003 ± 0.174	14.798 ± 0.522	14.750 ± 0.348
33 E	0.831 ± 0.040	0.960 ± 0.092	15.967 ± 0.495	16.626 ± 0.294
34 R	0.755 ± 0.024	0.930 ± 0.165	14.848 ± 0.610	15.967 ± 0.387
35 A	0.850 ± 0.048	0.944 ± 0.192	15.665 ± 1.035	16.599 ± 0.614
36 C	0.772 ± 0.078	0.888 ± 0.207	15.041 ± 0.852	16.934 ± 0.529
37 E	0.806 ± 0.053	0.973 ± 0.230	14.709 ± 0.575	15.119 ± 0.403
38 G	0.899 ± 0.071	1.008 ± 0.371	13.938 ± 0.827	13.827 ± 0.599
39 G	0.742 ± 0.051	0.942 ± 0.161	14.489 ± 0.789	15.382 ± 0.475
40 K	0.848 ± 0.099	0.928 ± 0.166	13.957 ± 1.004	15.037 ± 0.585
41 F	0.758 ± 0.040	0.955 ± 0.186	15.045 ± 0.481	15.760 ± 0.333
42 A	0.888 ± 0.068	1.017 ± 0.282	13.363 ± 0.804	13.135 ± 0.543
43 T	0.925 ± 0.100	0.870 ± 0.236	16.696 ± 1.048	19.183 ± 0.642
44 V	0.845 ± 0.130	0.988 ± 0.279	14.339 ± 1.353	14.520 ± 0.816
45 E	0.801 ± 0.050	0.949 ± 0.393	14.825 ± 0.638	15.625 ± 0.516
46 V	0.377 ± 0.097	0.898 ± 0.330	15.347 ± 1.505	17.089 ± 0.917

residue	hetNOE	R_1 (1/s)	R_2(1/s)	R_2/R_1
47 T	0.784 ± 0.029	0.963 ± 0.225	15.288 ± 0.458	15.870 ± 0.341
48 D	0.729 ± 0.053	0.912 ± 0.306	14.269 ± 0.594	15.642 ± 0.450
49 K	0.755 ± 0.048	0.973 ± 0.164	13.730 ± 0.496	14.112 ± 0.330
50 P				
51 V	0.860 ± 0.076	1.086 ± 0.986	14.143 ± 0.895	13.027 ± 0.940
52 D	0.837 ± 0.109	0.990 ± 0.170	15.996 ± 0.763	16.150 ± 0.466
53 E	0.804 ± 0.032	0.969 ± 0.159	15.026 ± 0.409	15.505 ± 0.284
54 A	0.900 ± 0.154	0.960 ± 0.256	14.300 ± 1.091	14.900 ± 0.674
55 L	0.827 ± 0.032	0.968 ± 0.217	±	± 0.108
56 R	0.988 ± 0.081	0.966 ± 0.171	16.138 ± 1.225	16.703 ± 0.698
57 E	0.731 ± 0.091	0.871 ± 0.168	16.276 ± 1.129	18.676 ± 0.648
58 A	0.848 ± 0.097	0.971 ± 0.168	14.063 ± 0.789	14.477 ± 0.479
59 M				
60 P				
61 K	1.018 ± 0.215	0.951 ± 0.308	17.688 ± 2.958	18.601 ± 1.633
62 V	0.798 ± 0.133	0.959 ± 0.160	15.596 ± 1.391	16.263 ± 0.775
63 A	0.879 ± 0.089	0.940 ± 0.200	14.343 ± 0.793	15.265 ± 0.497
64 K				
65 Y				
66 A	0.774 ± 0.102	0.899 ± 0.311	14.563 ± 1.621	16.202 ± 0.966
67 G	0.867 ± 0.063	0.834 ± 1.401	13.122 ± 0.791	15.739 ± 1.096
68 G	0.835 ± 0.067	0.897 ± 0.169	15.496 ± 0.665	17.277 ± 0.417
69 T	0.787 ± 0.046	0.944 ± 0.203	±	± 0.101
70 N	0.855 ± 0.071	0.897 ± 0.414	14.375 ± 0.885	16.033 ± 0.650
71 D	0.725 ± 0.041	0.952 ± 0.109	15.258 ± 0.500	16.031 ± 0.305
72 K	0.734 ± 0.045	0.836 ± 0.148	14.326 ± 0.479	17.136 ± 0.314
73 G	0.839 ± 0.041	1.051 ± 0.150	14.530 ± 0.446	13.819 ± 0.298
74 I	0.760 ± 0.041	0.971 ± 0.167	15.015 ± 0.453	15.464 ± 0.310
75 G	0.687 ± 0.123	0.832 ± 0.324	15.379 ± 1.803	18.473 ± 1.064
76 M	0.696 ± 0.049	0.937 ± 0.262	14.325 ± 0.660	15.295 ± 0.461
77 G	0.659 ± 0.084	0.986 ± 0.219	15.187 ± 1.105	15.396 ± 0.662
78 M	0.987 ± 0.141	1.050 ± 0.169	11.869 ± 0.967	11.307 ± 0.568
79 T	0.799 ± 0.030	0.949 ± 0.234	±	± 0.117
80 V	0.945 ± 0.081	1.018 ± 0.374	14.164 ± 0.873	13.911 ± 0.623
81 P				
82 I	0.784 ± 0.051	0.985 ± 0.265	14.546 ± 0.700	14.774 ± 0.482
83 S				
84 F	0.742 ± 0.503	0.919 ± 0.096	15.972 ± 0.747	17.372 ± 0.421

residue	**hetNOE**	**R_1 (1/s)**	**R_2(1/s)**	**R_2/R_1**
85 A	0.804 ± 0.179	0.961 ± 0.380	17.573 ± 2.793	18.279 ± 1.586
86 V	0.927 ± 0.050	0.993 ± 0.165	15.229 ± 0.577	15.336 ± 0.371
87 F	0.791 ± 0.046	0.999 ± 0.293	13.757 ± 0.749	13.765 ± 0.521
88 P				
89 N				
90 E				
91 D	0.726 ± 0.133	0.969 ± 0.650	13.252 ± 1.510	13.669 ± 1.080
92 G	0.681 ± 0.031	1.017 ± 0.239	14.155 ± 0.398	13.924 ± 0.318
93 S	0.709 ± 0.022	0.976 ± 0.213	14.630 ± 0.345	14.992 ± 0.279
94 L	0.794 ± 0.037	0.969 ± 0.180	14.293 ± 0.460	14.744 ± 0.320
95 Q	0.571 ± 0.064	0.945 ± 0.218		
96 K				
97 K	0.808 ± 0.061	0.922 ± 0.246	13.990 ± 0.560	15.179 ± 0.403
98 L	0.660 ± 0.048	0.906 ± 0.201		
99 K	0.811 ± 0.062	1.032 ± 0.139	12.510 ± 0.655	12.117 ± 0.397
100 V	0.830 ± 0.063	0.941 ± 0.175	15.729 ± 1.053	16.707 ± 0.614
101 W	0.762 ± 0.024	0.928 ± 0.181	14.836 ± 0.887	15.986 ± 0.534
102 F				
103 R	0.787 ± 0.031	0.973 ± 0.109	15.354 ± 0.421	15.776 ± 0.265
104 I	0.509 ± 0.097	1.205 ± 0.587	15.333 ± 1.955	12.720 ± 1.271
105 P				
106 N	0.857 ± 0.064	±	13.523 ± 0.696	
107 Q	0.653 ± 0.130	1.102 ± 0.709	15.747 ± 1.621	14.295 ± 1.165
108 F	0.766 ± 0.026	0.958 ± 0.205	14.592 ± 0.367	15.237 ± 0.286
109 Q	0.858 ± 0.048	0.999 ± 0.295	14.246 ± 0.583	14.263 ± 0.439
110 S				
111 D	0.775 ± 0.079	0.938 ± 0.070	13.783 ± 0.500	14.700 ± 0.285
112 P				
113 P				
114 A	0.814 ± 0.044	1.032 ± 0.169	15.469 ± 0.573	14.986 ± 0.371
115 P				
116 S	0.764 ± 0.078	1.020 ± 0.224	13.105 ± 0.775	12.849 ± 0.500
117 D	0.596 ± 0.090	0.894 ± 0.277	12.333 ± 0.638	13.795 ± 0.457
118 K	0.660 ± 0.234	0.904 ± 0.203	14.385 ± 0.584	15.906 ± 0.394
119 S	0.852 ± 0.047	1.132 ± 0.297	15.243 ± 0.526	13.460 ± 0.411
120 V	0.567 ± 0.031	1.003 ± 0.186	11.240 ± 0.308	11.202 ± 0.247
121 K	0.623 ± 0.053	1.079 ± 1.287	13.283 ± 0.682	12.315 ± 0.985
122 I	0.858 ± 0.051	0.973 ± 0.130	±	± 0.065

residue	**hetNOE**	**R_1 (1/s)**	**R_2(1/s)**	**R_2/R_1**
123 E	0.811 ± 0.046	0.978 ± 0.189	14.892 ± 0.523	15.230 ± 0.356
124 E	0.577 ± 0.036	0.985 ± 0.092	13.337 ± 0.394	13.536 ± 0.243
125 R	0.711 ± 0.045	0.915 ± 0.204	14.532 ± 0.639	15.889 ± 0.421
126 E				
127 G	0.712 ± 0.068	1.043 ± 0.373	18.577 ± 1.310	17.818 ± 0.841
128 I	0.933 ± 0.051	0.966 ± 0.296	16.631 ± 0.940	17.224 ± 0.618
129 T	0.656 ± 0.084	1.059 ± 0.338	11.923 ± 0.798	11.256 ± 0.568
130 V	0.886 ± 0.052	1.042 ± 0.172	15.041 ± 0.751	14.431 ± 0.461
131 Y	0.787 ± 0.045	0.957 ± 0.098	14.266 ± 0.540	14.903 ± 0.319
132 S	0.901 ± 0.055	0.953 ± 0.220	13.714 ± 0.548	14.387 ± 0.384
133 M	0.835 ± 0.068	±	16.619 ± 0.922	
134 Q	0.709 ± 0.037	1.029 ± 0.163	15.655 ± 0.524	15.212 ± 0.343
135 F				
136 G	0.479 ± 0.472			
137 G	0.918 ± 0.125	1.016 ± 0.547	11.941 ± 1.426	11.750 ± 0.987
138 Y				
139 A	0.507 ± 0.099	1.165 ± 0.262	13.144 ± 1.856	11.286 ± 1.059
140 K	0.770 ± 0.112	0.844 ± 0.226	14.287 ± 1.303	16.936 ± 0.765
141 E	0.976 ± 0.284	0.909 ± 0.317	14.353 ± 2.308	15.790 ± 1.313
142 A	0.664 ± 0.037	1.041 ± 0.171	14.965 ± 0.512	14.370 ± 0.341
143 D	0.865 ± 0.038	0.936 ± 0.119	16.030 ± 0.390	17.135 ± 0.254
144 Y				
145 V	0.827 ± 0.074	1.056 ± 0.235	14.321 ± 0.787	13.559 ± 0.511
146 A	0.920 ± 0.060	1.023 ± 0.107	14.812 ± 0.778	14.478 ± 0.442
147 Q	0.918 ± 0.073	1.027 ± 2.376	15.853 ± 0.724	15.442 ± 1.550
148 A	0.709 ± 0.020	0.973 ± 0.108	13.823 ± 0.255	14.201 ± 0.181
149 T	0.499 ± 0.038	0.959 ± 0.163	14.559 ± 0.499	15.176 ± 0.331
150 R	0.826 ± 0.045	0.971 ± 0.111	15.091 ± 0.524	15.543 ± 0.317
151 L	0.818 ± 0.047	0.991 ± 0.075	15.043 ± 0.558	15.179 ± 0.317
152 R	0.937 ± 0.047	0.981 ± 0.172	15.092 ± 0.536	15.386 ± 0.354
153 A	0.842 ± 0.042	0.971 ± 0.155	14.968 ± 0.428	15.421 ± 0.291
154 A	0.745 ± 0.030	0.895 ± 0.113	15.054 ± 0.324	16.818 ± 0.219
155 L	0.836 ± 0.047	0.950 ± 0.228	13.810 ± 0.489	14.544 ± 0.359
156 E	0.752 ± 0.027	0.996 ± 0.204	13.865 ± 0.244	13.921 ± 0.224
157 G				
158 T	0.693 ± 0.094	0.985 ± 0.342	13.117 ± 1.140	13.321 ± 0.741
159 A	0.678 ± 0.030	1.004 ± 0.118	11.535 ± 0.299	11.491 ± 0.208
160 T	0.782 ± 0.048	0.916 ± 0.308	11.084 ± 0.447	12.101 ± 0.377

residue	hetNOE	R_1 (1/s)	R_2(1/s)	R_2/R_1
161 Y	0.728 ± 0.027	0.888 ± 0.184	13.268 ± 0.594	14.944 ± 0.389
162 R				
163 G	0.412 ± 0.413	±	±	±
164 D	0.789 ± 0.050	1.144 ± 0.268	14.337 ± 0.554	12.534 ± 0.411
165 I	0.712 ± 0.030	0.947 ± 0.150	15.384 ± 0.382	16.247 ± 0.266
166 Y	0.803 ± 0.031	0.968 ± 2.304	15.051 ± 0.447	15.552 ± 1.375
167 F	0.904 ± 0.076	0.926 ± 0.144	21.055 ± 0.310	22.742 ± 0.227
168 C	0.760 ± 0.032	0.941 ± 0.206	15.204 ± 0.452	16.156 ± 0.329
169 T	0.919 ± 0.052	0.963 ± 0.154	14.915 ± 0.619	15.490 ± 0.386
170 G				
171 Y	0.687 ± 0.041	0.881 ± 0.133	13.382 ± 0.483	15.196 ± 0.308
172 D				
173 P				
174 P				
175 M				
176 K				
177 P				
178 Y	1.028 ± 0.114	0.981 ± 0.415	14.157 ± 0.701	14.435 ± 0.558
179 G	0.691 ± 0.089	0.989 ± 0.230		
180 R				
181 R	0.740 ± 0.061	0.939 ± 0.201	14.865 ± 0.694	15.838 ± 0.447
182 N	0.667 ± 0.091	0.941 ± 0.129	14.853 ± 1.028	15.786 ± 0.578
183 E	0.756 ± 0.115	0.865 ± 0.285	15.144 ± 1.178	17.501 ± 0.731
184 I	0.415 ± 0.043	0.929 ± 0.176	14.764 ± 0.402	15.890 ± 0.289
185 W	0.729 ± 0.062	1.083 ± 0.122	15.172 ± 0.641	14.004 ± 0.381
186 L	0.857 ± 0.057	0.903 ± 0.124	13.477 ± 0.616	14.918 ± 0.370
187 L	0.811 ± 0.062	0.966 ± 0.221	14.691 ± 0.798	15.207 ± 0.510
188 K	0.800 ± 0.069	0.969 ± 0.143	14.286 ± 1.050	14.739 ± 0.597
189 T	0.831 ± 0.040	0.972 ± 0.172	13.336 ± 0.418	13.721 ± 0.295

9.9 Relaxation parameters for human p22HBP

Residue	Model	S^2	τ_e (ps)	R_{ex}
16 W	1	0.968 ± 0.047		
17 P				
18 W				
19 Q	1	0.990 ± 0.027		
20 V	1	0.978 ± 0.057		
21 L	1	0.963 ± 0.051		
22 S	2	0.904 ± 0.031	3.120E+01 ± 1.605E+01	
23 K	1	0.937 ± 0.053		
24 G	2	0.935 ± 0.021	8.850E+02 ± 4.005E+02	
25 D	2	0.851 ± 0.022	1.370E+03 ± 2.095E+02	
26 K	2	0.786 ± 0.026	4.430E+01 ± 8.098E+00	
27 E	1	0.820 ± 0.060		
28 E	4	0.898 ± 0.096	9.630E+01 ± 1.730E+02	3.123 ± 1.819
29 V				
30 A	2	0.915 ± 0.025	5.610E+01 ± 2.594E+01	
31 Y	2	0.957 ± 0.016	6.020E+01 ± 3.634E+01	
32 E	1	0.966 ± 0.034		
33 E	2	0.970 ± 0.020	9.600E+02 ± 1.031E+03	
34 R	1	0.935 ± 0.021		
35 A	1	0.952 ± 0.036		
36 C	1	0.886 ± 0.037		
37 E	1	0.956 ± 0.032		
38 G	1	0.876 ± 0.020		
39 G	1	1.000 ± 0.043		
40 K	1	0.984 ± 0.031		
41 F	2	0.943 ± 0.026	7.850E+01 ± 5.812E+01	
42 A	1	1.000 ± 0.017		
43 T	1	0.985 ± 0.025		
44 V	2	0.883 ± 0.050	4.230E+01 ± 2.429E+01	
45 E	1	1.000 ± 0.021		
46 V	1	0.950 ± 0.029		
47 T	1	0.969 ± 0.046		
48 D	2	0.905 ± 0.014	2.160E+01 ± 1.020E+01	
49 K	2	0.877 ± 0.024	2.070E+01 ± 8.472E+00	
50 P				
51 V	1	0.962 ± 0.035		

Residue	Model	S^2	τ_e (ps)	R_{ex}
52 D	1	1.000 ± 0.019		
53 E	2	0.924 ± 0.018	1.180E+02 ± 5.806E+01	
54 A	1	0.995 ± 0.028		
55 L				
56 R	2	0.958 ± 0.023	2.560E+02 ± 3.201E+02	
57 E	1	0.963 ± 0.025		
58 A	2	0.943 ± 0.018	4.620E+01 ± 2.206E+01	
59 M	1	0.947 ± 0.039		
60 P				
61 K	1	0.930 ± 0.037		
62 V	1	1.000 ± 0.012		
63 A	1	0.986 ± 0.044		
64 K	1	0.980 ± 0.036		
65 Y	1	0.959 ± 0.027		
66 A	1	0.992 ± 0.022		
67 G	1	0.892 ± 0.028		
68 G	1	0.947 ± 0.031		
69 T	1	0.900 ± 0.031		
70 N	2	0.965 ± 0.018	9.140E+02 ± 6.697E+02	
71 D	1	0.953 ± 0.020		
72 K	1	0.886 ± 0.025		
73 G	2	0.923 ± 0.018	1.350E+02 ± 6.236E+01	
74 I	1	0.978 ± 0.016		
75 G	1	0.904 ± 0.058		
76 M	1	0.912 ± 0.040		
77 G	2	0.887 ± 0.045	1.430E+03 ± 5.924E+02	
78 M				
79 T	1	1.000 ± 0.028		
80 V	1	1.000 ± 0.057		
81 P				
82				
83	1	1.000 ± 0.140		
84 F	2	0.904 ± 0.019	5.380E+01 ± 1.337E+01	
85 A	1	0.969 ± 0.029		
86 V	1	0.932 ± 0.029		
87 F	1	0.918 ± 0.041		
88 P				
89 N				

Residue	**Model**	**S^2**	**τ_e (ps)**	**R_{ex}**
90 E				
91 D	2	0.859 ± 0.046	5.960E+01 ± 2.908E+01	
92 G	2	0.849 ± 0.029	1.680E+03 ± 2.907E+02	
93 S	2	0.898 ± 0.021	6.440E+01 ± 1.805E+01	
94 L	1	0.933 ± 0.029		
95 Q	1	0.956 ± 0.039		
96 K				
97 K	2	0.917 ± 0.030	5.400E+01 ± 2.668E+01	
98 L	1	0.953 ± 0.021		
99 K	1	0.938 ± 0.046		
100 V	1	0.986 ± 0.038		
101 W	1	0.948 ± 0.046		
102 F				
103 R				
104 I				
105 P				
106 N	1	0.961 ± 0.050		
107 Q	1	1.000 ± 0.048		
108 F	1	0.922 ± 0.028		
109 Q	1	0.974 ± 0.025		
110 S				
111 D				
112 P				
113 P				
114 A	2	0.969 ± 0.021	8.680E+01 ± 1.131E+02	
115 P				
116 S				
117 D	2	0.786 ± 0.021	1.750E+01 ± 4.966E+00	
118 K				
119 S	1	0.999 ± 0.025		
120 V	1	0.889 ± 0.032		
121 K	1	0.933 ± 0.037		
122 I	2	0.868 ± 0.020	3.930E+01 ± 8.795E+00	
123 E	1	0.973 ± 0.026		
124 E	2	0.866 ± 0.022	2.300E+01 ± 6.668E+00	
125 R	1	0.925 ± 0.028		
126 E				
127 G	1	0.938 ± 0.020		

Residue	Model	S^2	τ_e (ps)	R_{ex}
128 I	1	0.982 ± 0.037		
129 T	2	0.892 ± 0.029	3.700E+01 ± 1.331E+01	
130 V	1	0.942 ± 0.028		
131 Y	1	0.925 ± 0.040		
132 S	2	0.950 ± 0.019	9.160E+01 ± 5.849E+01	
133 M	1	0.980 ± 0.053		
134 Q	2	0.956 ± 0.044	1.350E+02 ± 2.438E+02	
135 F	1	1.000 ± 0.034		
136 G				
137 G	1	0.858 ± 0.138		
138 Y				
139 A	2	0.921 ± 0.049	1.190E+02 ± 1.427E+02	
140 K	1	0.885 ± 0.052		
141 E	1	1.000 ± 0.065		
142 A	1	1.000 ± 0.014		
143 D	1	1.000 ± 0.016		
144 Y	1	0.995 ± 0.023		
145 V	1	1.000 ± 0.024		
146 A	1	0.955 ± 0.025		
147 Q	1	0.988 ± 0.032		
148 A				
149 T	1	0.993 ± 0.018		
150 R	1	0.980 ± 0.021		
151 L	1	0.976 ± 0.026		
152 R				
153 A	1	1.000 ± 0.014		
154 A	2	0.971 ± 0.010	1.420E+02 ± 1.291E+02	
155 L	1	0.907 ± 0.031		
156 E	2	0.894 ± 0.015	3.850E+01 ± 8.294E+00	
157 G				
158 T	1	0.817 ± 0.045		
159 A	2	0.773 ± 0.013	2.400E+01 ± 3.241E+00	
160 T	2	0.703 ± 0.026	2.000E+03 ± 2.274E+02	
161 Y	2	0.895 ± 0.024	1.340E+03 ± 2.528E+02	
162 R				
163 G				
164 D	1	0.983 ± 0.025		
165 I	1	1.000 ± 0.013		

Residue	**Model**	**S^2**	**τ_e (ps)**	**R_{ex}**
166 Y	1	1.000 ± 0.021		
167 F	1	1.000 ± 0.036		
168 C	1	1.000 ± 0.021		
169 T	1	1.000 ± 0.037		
170 G				
171 Y				
172 D				
173 P				
174 P				
175 M				
176 K				
177 P				
178 Y	1	0.988 ± 0.094		
179 G				
180 R				
181 R				
182 N	2	0.858 ± 0.033	5.780E+01 ± 1.918E+01	
183 E	1	0.966 ± 0.057		
184 I	1	0.999 ± 0.022		
185 W	1	0.936 ± 0.034		
186 L	1	0.951 ± 0.043		
187 L	1	0.953 ± 0.044		
188 K	1	1.000 ± 0.032		
189 T	2	0.826 ± 0.025	2.880E+01 ± 7.139E+00	

9.10 Relaxation parameters for PPIX-human p22HBP

Residue	Model	S^2	τ_e (ps)	R_{ex}
16 W	1	0.897 ± 0.051		
17 P				
18 W	1	0.942 0.027		
19 Q	1	0.903 ± 0.039		
20 V	1	0.841 ± 0.051		
21 L	1	0.887 ± 0.051		
22 S	1	0.810 ± 0.022		
23 K	1	0.909 ± 0.027		
24 G	1	0.834 ± 0.030		
25 D	2	0.751 ± 0.024	25.531 ± 6.176	
26 K	2	0.867 0.028	794.085 ± 221.502	
27 E	2	0.934 ± 0.080	49.844 ± 72.001	
28 E	1	0.861 0.013		
29 V	2	0.859 0.027	54.996 ± 14.988	
30 A	1	0.895 ± 0.035		
31 Y	1	0.891 ± 0.048		
32 E	1	0.920 ± 0.032		
33 E	1	0.976 ± 0.029		
34 R	2	0.927 ± 0.037	49.895 ± 33.217	
35 A	1	0.974 ± 0.061		
36 C	1	0.936 ± 0.052		
37 E	1	0.925 ± 0.036		
38 G	1	0.879 ± 0.051		
39 G	1	0.892 0.046		
40 K	1	0.873 ± 0.058		
41 F	1	0.945 ± 0.030		
42 A	1	0.841 ± 0.049		
43 T	1	1.000 ± 0.062		
44 V	1	0.883 ± 0.079		
45 E	1	0.907 ± 0.039		
46 V	2	0.858 ± 0.035	233.151 ± 938.772	
47 T	1	0.925 ± 0.027		
48 D	1	0.871 ± 0.036		
49 K	1	0.848 ± 0.030		
50 P				
51 V	1	0.868 ± 0.055		

Residue	Model	S^2	τ_e (ps)	R_{ex}
52 D	1	0.992 ± 0.046		
53 E	1	0.935 ± 0.025		
54 A	1	0.877 ± 0.064		
55 L	1	0.930 0.208		
56 R	1	0.987 ± 0.069		
57 E	1	0.975 ± 0.064		
58 A	1	0.862 ± 0.046		
59 M		±		
60 P				
61 K	1	1.000 ± 0.155		
62 V	1	0.954 ± 0.075		
63 A	1	0.897 ± 0.048		
64 K		±		
65 Y		±		
66 A	1	0.897 ± 0.095		
67 G	1	0.823 ± 0.050		
68 G	1	0.959 ± 0.040		
69 T	1	0.921 ± 0.198		
70 N	1	0.890 ± 0.054		
71 D	2	0.931 ± 0.029	72.341 ± 48.317	
72 K	1	0.868 ± 0.029		
73 G	1	0.914 ± 0.027		
74 I	1	0.943 ± 0.028		
75 G	1	0.934 ± 0.105		
76 M	2	0.892 ± 0.041	55.244 ± 32.702	
77 G	1	0.941 ± 0.065		
78 M	1	0.768 0.056		
79 T	1	0.890 ± 0.219		
80 V	1	0.886 0.054		
81 P				
82	1	0.917 ± 0.043		
83		±		
84 F	1	0.957 ± 0.041		
85 A	1	1.000 ± 0.154		
86 V	1	0.940 ± 0.035		
87 F	1	0.845 ± 0.045		
88 P				
89 N				

Residue	Model	S^2	τ_e (ps)	R_{ex}
90 E				
91 D	1	0.812 ± 0.091		
92 G	2	0.870 ± 0.032	998.645 ± 258.747	
93 S	2	0.909 ± 0.027	883.211 ± 279.091	
94 L	1	0.896 ± 0.028		
95 Q	1	0.895 ± 0.207		
96 K				
97 K	1	0.855 ± 0.034		
98 L	1	0.878 ± 0.195		
99 K	1	0.776 ± 0.038		
100 V	1	0.957 ± 0.060		
101 W	2	0.902 ± 0.052	31.416 ± 21.210	
102 F				
103 R	1	0.953 0.025		
104 I	2	0.895 ± 0.096	270.587 ± 2692.936	
105 P				
106 N	1	0.831 ± 0.043		
107 Q	1	0.984 ± 0.100		
108 F	1	0.887 ± 0.022		
109 Q	1	0.888 ± 0.036		
110 S				
111 D	1	0.864 0.028		
112 P				
113 P				
114 A	1	0.949 ± 0.034		
115 P				
116 S	1	0.826 0.047		
117 D	2	0.765 ± 0.039	37.145 ± 16.954	
118 K	1	0.884 ± 0.035		
119 S	1	0.946 ± 0.032		
120 V	2	±		
121 K	2	0.807 ± 0.041	41.346 ± 15.948	
122 I	1	0.947 ± 0.127		
123 E	1	0.930 ± 0.032		
124 E	2	0.833 ± 0.023	64.396 ± 15.789	
125 R	2	0.907 ± 0.039	58.598 ± 37.546	
126 E				
127 G	1	1.000 ± 0.080		

Residue	**Model**	**S^2**	**τ_e (ps)**	**R_{ex}**
128 I	1	1.000 ± 0.057		
129 T	1	0.757 ± 0.050		
130 V	1	0.950 ± 0.045		
131 Y	1	0.897 ± 0.032		
132 S	1	0.862 ± 0.034		
133 M	1	1.000 ± 0.057		
134 Q	2	0.952 ± 0.037	407.220 ± 719.917	
135 F		±		
136 G	1			
137 G	1	0.734 ± 0.086		
138 Y				
139 A	2	0.791 ± 0.113	760.412 ± 387.456	
140 K	1	0.869 ± 0.075		
141 E	1	0.888 ± 0.130		
142 A	2	0.934 ± 0.037	412.880 ± 512.764	
143 D	1	0.980 ± 0.023		
144 Y	1	0.901 ± 0.048		
145 V	1	0.930 ± 0.044		
146 A	1	0.962 ± 0.044		
147 Q	2	0.859 ± 0.016	36.349 ± 7.589	
148 A	2	0.890 0.019	226.337 ± 428.608	
149 T	1	0.935 ± 0.031		
150 R	1	0.932 ± 0.031		
151 L	1	0.941 ± 0.033		
152 R	1	0.927 0.026		
153 A	2	0.913 ± 0.019	43.926 ± 18.530	
154 A	1	0.855 ± 0.030		
155 L	2	0.872 ± 0.015	27.107 ± 9.325	
156 E		±		
157 G	1	0.827 0.070		
158 T	2	0.730 ± 0.018	20.180 ± 4.268	
159 A		±		
160 T	1	±		
161 Y	2	0.833 ± 0.036	25.162 ± 9.157	
162 R				
163 G	1	±		
164 D	1	0.900 ± 0.034		
165 I	2	0.958 ± 0.019	216.922 ± 634.736	

Residue	**Model**	**S^2**	**τ_e (ps)**	**R_{ex}**
166 Y	1	0.941 ± 0.028		
167 F	3	0.884 ± 0.138		6.629 ± 1.862
168 C	2	0.943 ± 0.035	899.877 ± 619.926	
169 T	1	0.905 0.036		
170 G				
171 Y	2	0.812 0.029	29.272 ± 9.847	
172 D				
173 P				
174 P				
175 M				
176 K				
177 P				
178 Y	1	0.873 ± 0.043		
179 G	1	0.947 ± 0.220		
180 R				
181 R	1	0.913 0.042		
182 N	1	0.904 ± 0.056		
183 E	1	0.919 ± 0.070		
184 I	2	0.872 ± 0.012	243.649 ± 578.343	
185 W	1	0.942 ± 0.037		
186 L	1	0.840 ± 0.036		
187 L	1	0.922 ± 0.049		
188 K	1	0.895 ± 0.059		
189 T	1	0.822 ± 0.025		