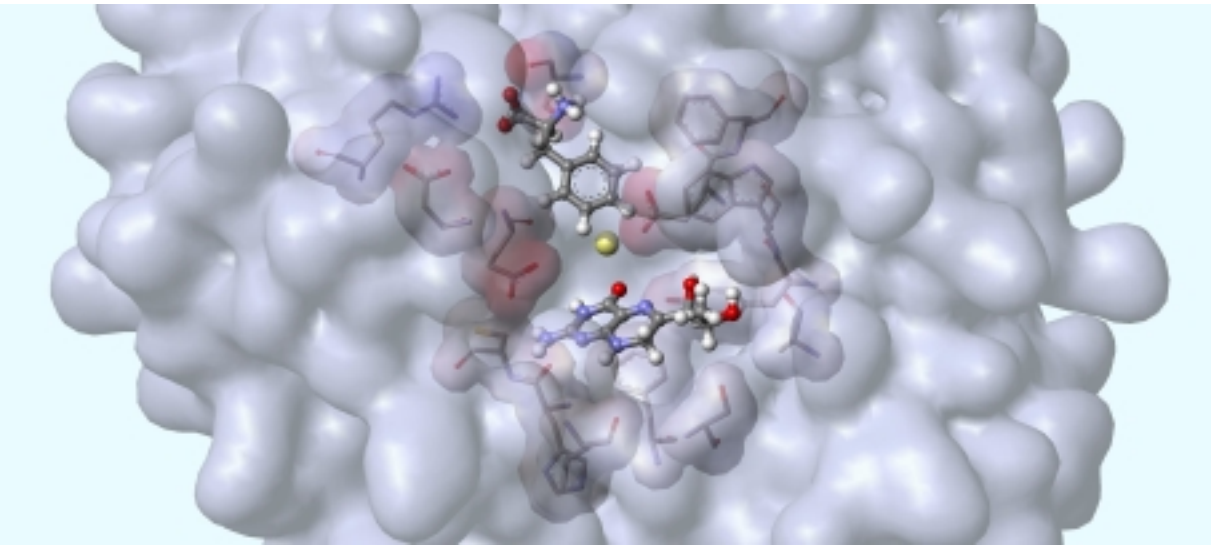
Binding of substrate and pterin cofactor analogue to human phenylalanine hydroxylase studied by multidimensional heteronuclear NMR.



by

Knut Teigen

Submitted in the partial fulfillment of the degree candidatus scientarium
at
The Department of Biochemistry and Molecular Biology, University of Bergen
2000

Binding of substrate and pterin cofactor analogue to human phenylalanine hydroxylase studied by multidimensional heteronuclear NMR.

Contents

Summary

Human phenylalanine hydroxylase (hPAH) is a tetrahydrobiopterin- and non-heme iron-dependent enzyme that hydroxylates L-Phe to L-Tyr using molecular oxygen as additional substrate. A dysfunction of this enzyme leads to phenylketonuria (PKU) which is one of the most prevalent disorder of amino acid metabolism. In this study, the determination of the structure of a ternary complex of recombinant hPAH with L-Phe and the catalytically inactive pterin cofactor analogue 7,8-*erythro*-dihydrobiopterin (BH_2) has been attempted to be solved by nuclear magnetic resonance (NMR) spectroscopy and molecular modeling techniques. This has been done in order to get further insights into the reaction mechanism of hPAH in particular and of the aromatic amino acid hydroxylases in general. Distances of L-Phe and BH_2 protons to the iron and interproton distances in the enzyme bound forms of these ligands were estimated by the paramagnetic probe-T_1 method and by transferred NOESY spectra, respectively. The aromatic protons of L-Phe and H7 of BH_2 are the closest protons to the iron. Transferred NOESY spectra also have shown intermolecular interactions between protons from the substrate and the enzyme, notably from a tryptophan residue, which have helped in the determination of the binding site. The resulting solution conformations generated by distance geometry calculations from the NMR distances of the bound ligands have been docked into the crystal structure of the catalytic domain of hPAH, using the program package DOCK.

In the proposed structure of the ternary complex of hPAH with its substrate and cofactor analogue, L-Phe binds to the enzyme through interactions with Arg270, Ser349 and Trp326. The way of coordination of Glu330 to the iron at the active site seems to determine the amino acid substrate specificity in hPAH and in the homologous enzyme tyrosine hydroxylase (TH). The pterin ring of BH_2 π-stacks with Phe254 and is anchored to Glu286 through interactions with N3 and N2. The ring also establishes specific contacts with His264 and Leu249. Moreover, the O4 atom of BH_2 is at a distance from the iron that is compatible with coordination (2.6 Å$\forall$0.4 Å), a finding that is important for the understanding of the mechanism of the enzyme. The hydroxyls in the side chain at C6 hydrogen-bind with the carbonyl group of Ala322 and the hydroxyl of Ser251, an interaction which seems to have implications for the regulation of the enzyme by substrate and cofactor.

The substrate binding site shows a similar structural arrangement of charged and aromatic residues as other amino acid-binding proteins, while the pterin binding site shows the same motifs for ligand recognition as found in other enzymes in the pathway for the synthesis or regeneration of BH_4. Interestingly, the solution structure of the complex between hPAH and BH_2 differs from the recently reported crystal structure of TH with bound BH_2.

Some frequent mutations causing PKU are located at residues involved in substrate and cofactor binding, i.e. R270K/S, D282N, F331C, S349L, L249H/F, F254I, H264L, C265Y/G, A322T/G. The sites for hydroxylation, C4 in L-Phe and C4a in the pterin are located at a distance of 4.2 and 4.3 Å from the iron, respectively, and at 6.3 Å from each other. These distances are adequate for the intercalation of iron-coordinated molecular oxygen, in agreement with a mechanistic role of the iron both in the binding and activation of dioxygen and in the hydroxylation reaction.

Acknowledgements

I wish to express my warm and sincere feelings of gratitude towards my two supervisors, Prof. Aurora Martínez at Lab D, Dept. of Biochemistry and Molecular Biology and Associate Prof. Nils Åge Frøystein at the Dept. of Chemistry for support and encouragement through every stage of my thesis. Your enthusiasm and expertise has been a great inspiration.

The complete fulfillment of this thesis would not have been possible without the help of Prof. Irwin D. Kuntz at the Molecular Design Institute in San Fransisco, California. Prof. Kuntz and his group are deeply acknowledged for their friendliness and willingness to share their knowledge with a novice in the field of molecular modeling.

Lieutenant-Commander Torhild Frøydis Eid and Lieutenant Atle Hansen Røsseland at Sykestuen, Haakonsvern, Prof. Jan Risberg and Kåre Segadal at nui have been very conciderate in allowing me to work with my thesis during my national service in the navy. Your patience and understanding is sincerely acknowledged.

The students at IBMB are thanked for making PKI a very social and friendly place to work. Bente Berg is thanked for her ever lasting optimism, and for always seeing the brighter side of things. Thanks for all the encouraging chats at the corner office!

My good friend Jan Einar Gravdal at the Dept. of Mathematics is thanked for his expert tutoring when the mathematical concepts of NMR and molecular modeling became too frustrating. The many nights spent discussing docking algorithms over a Jägermeister in Eikedalen is greatly appreciated!

Abbreviations

4a-OH-BH_4	4a-hydroxy-tetrahydropterin
BH_2	7,8-*erythro*-dihydrobiopterin
BH_4	(6R)-L-*erythro*-5,6,7,8-tetrahydrobiopterin
6,7-$DMPH_4$	6,7-dimethyl-tetrahydropterin
FID	free induction decay
HMBC	heteronuclear multiple bond coherence
hPAH	human phenylalanine hydroxylase
HSQC	heteronuclear single quantum coherence
L-Phe	L-phenylalanine
L-Trp	L-tryptophan
L-Tyr	L-tyrosine
MPB	maltose-binding protein
6-MPH_4	6-methyltetrahydropterin
MSB	modified Solomon-Bloembergen
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
PAH	phenylalanine hydroxylase
PKA	cyclic AMP-dependent protein kinase
PKU	phenylketonuria
q-BH_2	quinonoid-7,8-dihydrobiopterin
r.m.s.d.	root mean square deviation
TH	tyrosine hydroxylase
TPH	tryptophane hydroxylase
TRNOESY	transferred nuclear Overhauser effect spectroscopy
wt-hPAH	wild-type human phenylalanine hydroxylase

1 Introduction

1.1 General Introduction

In 1934, Dr. Asbjørn Følling described the first inborn error of metabolism shown to affect the minds of two severely mentally retarded siblings (Følling, 1934). The patients were found to excrete phenylpyruvic acid in their urine. This disease was later named phenylketonuria (PKU), and Følling showed the pattern of an autosomal recessive genetic disease, probably caused by a block in the phenylalanine metabolism. In later years it has been shown that PKU is mostly the result of a deficiency in the enzyme phenylalanine hydroxylase (PAH, phenylalanine 4-monooxygenase, EC 4.14.16.1), arising from mutations in the PAH gene.

PAH catalyzes the hydroxylation of L-phenylalanine (L-Phe) to tyrosine (L-Tyr), using (6R)-L-*erythro*-5,6,7,8-tetrahydrobiopterin (BH_4) and molecular oxygen as additional substrates (Fig 1.1). This is the rate-limiting step in the pathway to catabolize L-Phe mainly in the liver.

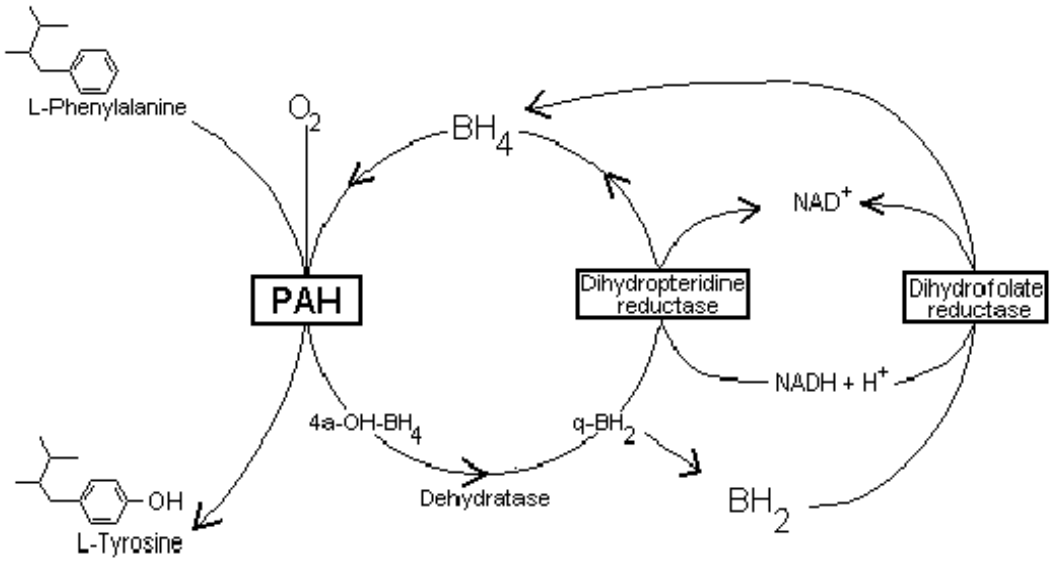


Figure 1.1
The phenylalanine hydroxylase (PAH) system. BH_4, (6R)-L-*erythro*-5,6,7,8-tetrahydrobiopterin; q-BH_2, quinonoid-7,8-dihydrobiopterin; 4a-OH-BH4, 4a-hydroxy-tetrahydrobiopterin.

Together with tyrosine hydroxylase (TH) and tryptophan hydroxylase (TPH), PAH constitutes a superfamily of aromatic amino acid hydroxylases, which catalyze key steps in important metabolic pathways (Fig 1.2) (Martinez *et al*., 1993a; Martinez *et al*., 1998; Mildvan *et al*., 1980).

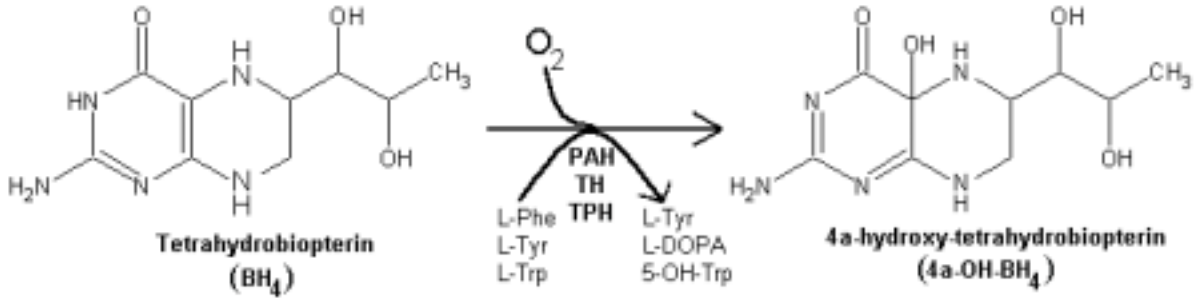


Figure 1.2
The aromatic amino acid hydroxylase family of enzymes. PAH, phenylalanine hydroxylase, TH, tyrosine hydroxylase, TPH, tryptophan hydroxylase.

1.2 Molecular Properties of Phenylalanine Hydroxylase

Rat phenylalanine hydroxylase shows two to three closely spaced bands after SDS gel electrophoresis which range in size from 50 to 52 kD (Kappock & Caradonna, 1996). The molecular origin of the multiple bands has not been established in any case. The amino acid composition predicts a molecular weight of 51.9 kD for rat PAH and 51.8 for human PAH (hPAH). Under nondenaturing conditions of gel exclusion chromatography, active, purified hPAH from liver and the recombinant hPAH have been reported to have a molecular weight corresponding to a dimer and a tetramer (Woolf, 1976). Gel exclusion chromatography gives an apparent molecular weight of about 170 kDa (Kaufman & Fisher, 1970; Nakata & Fujisawa, 1980; Shiman *et al*., 1979) for the tetramer, corresponding to a Stokes radius of 55 Å.

All mammallian PAH have an absolute requirement for iron, and bind about one iron atom per subunit with enzyme activity proportional to the iron content (Knappskog *et al*., 1996). Recombinant PAH has been overexpressed in *Escherichia coli* and insect cells, and the purified recombinant enzyme forms have revealed essentially the same physical and kinetic properties as protein purified from endogenous sources (Citron *et al*., 1992; Martinez *et al*., 1995). PAH and the rest of the aromatic amino acid hydroxylases are organized in a regulatory N-terminal domain, a catalytic domain and a C-terminal tetramerization domain, and they show extensive sequence similarity at the catalytic domains. Therefore, they are believed to have a similar reaction mechanism.

1.3 Regulation of Phenylalanine Hydroxylase

Since phenylalanine is an essential amino acid and also toxic through its metabolites, a tight regulation of PAH activity seems physiologically very important. Several factors which regulate PAH activity have been discovered, the most important being activation by its amino acid substrate and phosphorylation. Inhibition of the enzyme by its natural cofactor BH_4 is also considered to be a regulatory mechanism.

1.3.1 *Activation by Phenylalanine*

Substrate activation is an attractive mechanism for regulating the degradation of an essential amino acid, which is toxic in high levels, since it allows the level of available active enzyme to vary in response to the amount of available substrate. The reaction progress curve of rat liver PAH shows a lag which is found to be temperature dependent. Preincubation of the enzyme with L-Phe completely abolishes this lag. These results imply that activation of PAH by L-Phe is required for PAH activity (Shiman & Gray, 1980). The activation of PAH by L-Phe is thought to result from a cooperative binding of the substrate to the enzyme, either to an allosteric binding site distant to the active site (Parniak & Kaufman, 1981; Shiman *et al*., 1990; Shiman *et al*., 1994) or to the active site near the iron (Martinez et al., 1993a; Martinez *et al*., 1991). The main argument for the presence of an allosteric binding site has been that a suprastoichiometric amount (1.5 mol) of L-Phe seems to bind per PAH subunit. Moreover, the enzyme activated with N-ethylmaleimide, only bound 1 mol L-Phe per subunit, leading to the conclusion that the allosteric site binds 0.5 mol per subunit (Parniak & Kaufman, 1981). Furthermore, it was suggested that this allosteric site lies on the N-terminal domain since the enzyme looses its coopreativity after removal of this domain. However, Martínez *et al*. (Martinez et al., 1993a; Martinez et al., 1991) have shown that the inhibitor L-noradrenaline, which exclusively coordinate to PAH at the active site iron, also binds with positive cooperativity and induces similar conformational changes on rat PAH as the binding of L-Phe does. This suggested that binding to the active site is responsible for the activation.

The activation by L-Phe has been found to result in confomational changes involving the ternary and quaternary structure (Kappock *et al*., 1995; Phillips *et al*., 1984b), changing the dimer ≡ tetramer equilibrium towards the tetrameric form both for rat PAH (Doskeland *et al*., 1982) and recombinant hPAH (Martinez et al., 1995). This conformational change results in an 10 % increase in volume (Kappock & Caradonna, 1996) and is also accompanied by an increase in surface hydrophobicity (Shiman et al., 1979).

1.3.2 *Activation by Phosphorylation*

It has been shown that phosphorylation of rat PAH is dependent on two protein kinases, the cyclic AMP-dependent protein kinase (PKA) (Abita *et al*., 1976) or Ca^{2+}/calmodulin-dependent protein kinase II (Døskeland *et al*., 1984). Phosphorylation takes place at Ser16, which has a surrounding sequence typical of many substrates of PKA (Wretborn *et al*., 1980). Model peptides corresponding to the amino acid sequence between Ser12 and Gly19 had much higher K_m values for PKA than the wild-type enzyme, suggesting that the conformation around Ser16 in PAH has a significant role in recognition by the protein kinase (Wretborn et al., 1980).

The effects of phosphorylation on PAH activity have been intensively studied by Døskeland *et al.* (Doskeland et al., 1982; Doskeland & Flatmark, 1996; Døskeland et al., 1984). A 1.6-1.9 increase in specific activity was reported after phosphorylation of hPAH when the activity was measured without preincubation with L-Phe and with BH_4 as cofactor. If 6-methyltetrahydropterin (6-MPH_4) is used as a cofactor, no appartent effect of phosphorylation is observed (Abita et al., 1976). When the enzyme activity was assayed under the standard conditions, which includes preincubation with L-Phe, no effect of phosphorylation was detected. These results are similar to the effects of phosphorylation of the rat PAH where phosphorylation at

Ser16 resulted in a 4- to 6- fold increase in catalytic activity, only when activity was measured without preincubation of L-Phe. However, the phosphorylated form needed only half the concentration of L-Phe to obtain half-maximal substrate activation but with no significant change in the cooperative binding of L-Phe, both for the human and the rat enzyme (Doskeland & Flatmark, 1996). This suggests that the main effect of the phosphorylation is to make the enzyme more sensitive for substrate activation by L-Phe, indicating that both regulatory mechanisms, activation by L-Phe and phosphorylation, act synergetically. For the rat PAH it has also been found that L-Phe and BH_4 effect the rate of phosphorylation in the sense that L-Phe stimulates and BH_4 inhibits the rate of phosphorylation (Hufton *et al*., 1995).

1.3.3 *Inhibition by Tetrahydrobiopterin*

BH_4 is proposed to be a negative effector that blocks L-Phe activation by forming an inactive BH_4-enzyme complex. It has been suggested that inhibition by BH_4 *in vivo* is important to protect BH_4 from degradation and to control its metabolic availability when the level of L-Phe in the cell is low (Xia *et al*., 1994).

Addition of sufficient BH_4 prior to L-Phe activation will completely block the activating effect of L-Phe (Kappock & Caradonna, 1996). This inhibition of L-Phe activation is specific for BH_4. Thus, 6-MPH_4 is 1000-fold less effective than BH_4 and 50-fold less effective than L-erythro-7,8-dihydrobiopterin (BH_2), an oxidized inactive cofactor analogue to BH_4, as inhibitor in the presence of 5 mM L-Phe (Xia et al., 1994).

BH_4 has been shown to inhibit the *in vitro* rate of phosphorylation at Ser16. The effect is most pronounced for the natural isomer, (6*R*)-BH_4, while (6*S*)-BH_4 is half as inhibitory; neither 6-MPH_4 nor 6,7-dimethyl-tetrahydropterin (6,7-$DMPH_4$) inhibit the enzyme (Phillips *et al*., 1984a).

1.4 Inhibition by Catecholamines

PAH, as well as the homologous enzyme TH, is inhibited by catecholamines, i.e. dopamine, noradrenaline and adrenaline. This inhibition by catecholamines is thought to be of physiological significance for TH, due to its location in catecholaminergic neuroendocrine cells containing catecholamines. The inhibition of PAH by catecholamines is not thought to be of physiological significance.

PAH activity is found to be inhibited by catechols competitively with respect to the BH_4 cofactor and non-competetively with respect to L-Phe. Catecholamines bind to PAH by bidentate coordination to Fe(III) at the active site (Erlandsen *et al*., 1998). The formation of the tight catecholate-Fe(III) complex seems to lower the redox potential and stabilize the ferric state, in agreement with a kinetic competition between the catecholamine inhibitors and the tetrahydropterin cofactors (Kaufman, 1993; Martinez et al., 1991). It has also been postulated a steric hindrance to the binding of the cofactor by the catecholamine as explanation for the competitive type of inhibition (Erlandsen et al., 1998).

1.5 Proposed Reaction Mechanism

Many experiments have been carried out in order to determine the catalytic mechanism of both PAH and TH. Despite of this, few details are known about the exact nature of the substrate hydroxylating species. The recent determination of the

3D-structure of the catalytic domains of both rat TH (Goodwill *et al*., 1997) and hPAH (Erlandsen *et al*., 1997; Fusetti *et al*., 1998) has revealed the structural similarities in both enzymes, including the 2-His-1-carboxylate facial triad (Lange & Que, 1998) anchoring the catalytic mononuclear non-heme iron at the active site. In addition, the crystal structure of the catalytic domains of rat TH with bound BH_2 (Goodwill *et al*., 1998), and of human PAH with bound catechol inhibitors (Erlandsen et al., 1998) have recently been described. In the reported binary complex of TH with BH_2, the pterin binds forming an aromatic π-stacking interaction with Phe300 (Phe254 in human PAH), with a distance from the iron to the pterin 4a carbon of 5.6 Å (Goodwill et al., 1998). No crystal structures of complexes of the enzymes with amino acid substrates have been yet reported and the crystallization of the binary complex of L-Phe and PAH has been unsuccessful so far (Erlandsen et al., 1998). Nevertheless, it has been shown by ^{1}H NMR spectroscopy that the aromatic ring of the amino acid substrate binds to recombinant human TH at the second coordination sphere of the active site iron (Martinez et al., 1993a). A study of the paramagnetic relaxation of solvent protons by the non-heme iron has also indicated that a water molecule is displaced from coordination to the iron following the binding of L-Phe to PAH (Martinez *et al*., 1993c; Olafsdottir & Martinez, 1999). From the observed temperature factors and bond lengths of the three iron-coordinated water molecules in the crystal structure of the catalytic domain of PAH it seems that either H_2O (1), distal to His285, or H_2O (2), distal to His290 (Erlandsen et al., 1997), are the most likely candidates for displacement by the substrate.

Several mechanisms of hydroxylation have been proposed for the aromatic amino acid hydroxylases (Dix & Benkovic, 1988; Kappock & Caradonna, 1996), but the actual catalytic mechanism, including the chemical nature of the hydroxylating intermediate, is still not clear. It has been proven for mammalian PAH that in the catalytic cycle, the iron, which is in the ferric form in the enzymes as isolated, is prereduced to its ferrous form by the pterin cofactor and participates in catalysis (Kappock & Caradonna, 1996; Kaufman, 1993). For TH it has also been shown that the steady-state kinetic mechanism is sequential with the tetrahydropterin cofactor binding first followed by molecular oxygen and then the substrate (Fitzpatrick, 1991). The kinetic mechanism of PAH seems to be sequential as well, with some degree of randomness in the order of substrate addition (Kappock & Caradonna, 1996; Kaufman, 1993). It seems clear for both enzymes that no product or intermediate is released prior to the binding of all substrates. The first observable product of the pterin cofactor in the TH and PAH catalyzed reactions is a 4a-hydroxy-tetrahydropterin, in which the oxygen atom in position 4a is derived from molecular oxygen (Dix & Benkovic, 1988; Haavik & Flatmark, 1987; Kappock & Caradonna, 1996). The other half of the oxygen molecule is found in the hydroxylated product. By analogy with the flavoprotein monooxygenases in which a 4a-peroxyflavin seems to be the hydroxylating intermediate, several authors have deduced that the hydroxylating species in TH and PAH is a 4a-peroxy-tetrahydropterin (Dix & Benkovic, 1988; Kappock & Caradonna, 1996). The ferrous iron at the active site might participate in the formation of this intermediate by prior formation of an iron-oxo or iron-peroxo compound. A highly reactive iron-oxygen intermediate has been postulated to be in fact the hydroxylating intermediate itself in rat PAH (Davis & Kaufman, 1989). None of these reactive species and intermediates have been directly detected or unequivocally proven in enzymatic systems. However, the distances between the iron and the amino acid substrate and pterin cofactors bound to TH, as estimated by NMR spectroscopy, suggest a direct role of the metal ion not

only in the activation of dioxygen, but also in the hydroxylation reaction (Martinez et al., 1993a; Martinez et al., 1998).

1.6 Phenylketonuria

Phenylketonuria (PKU) is an autosomal recessive disease caused by mutations in the PAH gene (Eisensmith & Woo, 1991; Erlandsen & Stevens, 1999). The mutations result in a complete loss or a variable degree of reduced activity of PAH, leading to an accumulation of L-Phe in the blood and excretion of its metabolite phenylpyruvate in the urine. Classical PKU was first described by Følling in two mentally retarded siblings (Følling, 1934). However, the biochemical basis for the defect was not established until Jervis showed that *post mortem* liver tissue from normal individuals could convert L-Phe into tyrosine, whereas liver extracts from PKU patients could not (Jervis, 1953). The discovery of this error of metabolism has lead to the treatment of PKU patients with a low-phenylalanine diet, resulting in improvements in mental development and behavioral performance (Bickel *et al*., 1954; Woolf *et al*., 1955). With the cloning of the hPAH gene (Kwok *et al*., 1985) studies directed to unreveal the genetic basis of PKU were initiated, and since then more than 300 different PKU mutations have been described.

1.7 *Purpose of the Present Study*

The aim of the present study was to determine the structure of a ternary complex of hPAH with L-Phe and the catalytically inactive pterin cofactor analogue 7,8-*erythro*-dihydrobiopterin (BH_2) by using NMR spectroscopy and molecular modeling techniques. This has been done in order to get further insights into the reaction mechanism of hPAH in particular and of the aromatic amino acid hydroxylases in general.

2 Theoretical considerations

2.1 NMR Spectroscopy

In this study, NMR was used to estimate:

(i) the intramolecular proton distances in the enzyme bound ligands by transferred nuclear Overhauser effect spectroscopy, TRNOESY

(ii) the intermolecular distances between some of the ligand protons and protons of the protein by TRNOESY

(iii) the intermolecular distances from the ligand protons to the active site iron by relaxation measurements

2.1.1 *General Theory*

The NMR phenomenon is a consequence of the existence of nuclear spin. Nuclear spin is associated to the spin quantum number (I), and is an intrinsic property of a given nucleus, dependent on the relationship between protons and neutrons of that particular nucleus.

The number of spin states, or eigenvalues (m_I) is given by:

$$2I+1 \qquad (2.1)$$

^{12}C, ^{14}N and ^{16}O are examples of nuclei with I=0, and hence they posess no magnetic spin. ^{1}H and ^{13}C both have spin ½, and therefore have two spin states. The energy difference between the spin states is proportional to the strength of the applied magnetic field (B_0):

$$\Delta E = \gamma h B_0/2\pi \qquad (2.2)$$

where γ is the magnetogyric ratio, an intrinsic property of a given nuclei, h is the Planck constant.

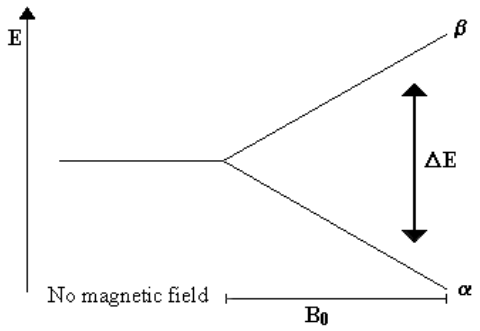


Figure 2.1
The energy dependence on the applied magnetic field.

If a magnetic field is applied to a sample of proton nuclei (e.g. when the nuclei are placed in the NMR spectrometer), their magnetic moments will distribute between the two spin states (α and β), with the energy difference given by 2.2.

The two energy states will be unequally populated, the ratio of the populations being given by the Boltzmann equation:

$$N_\beta/N_\alpha = \exp(-\Delta E/kT) \qquad (2.3)$$

where N_α is the population in the lower energy state, N_β is the population in the upper energy state, k is the Boltzmann constant and T is the absolute temperature.

When a population of nuclei are introduced into a magnetic field (the NMR magnet) their magnetic moments will start to rotate (precess) around the axis of the applied field. The frequency of precessing is dependent on the strength of the applied magnetic field, and is called the Larmor frequency, ν_0.

The distribution of precessing nuclei can be described using vectors.

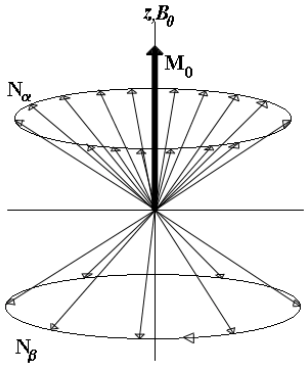


Figure 2.2
The vectors (nuclear spins) precess around the *z*-axis (the applied magnetic field) with frequency ν_0. Since there is a small excess of vectors (spins) aligned with the *z*-axis (B_0), this creates a net magnetization in the sample along the magnetic field.

If we apply a pulse with a carrier frequency corresponding to the energy equal to the difference between the energy levels (ΔE), the Boltzmann distribution of the spins will be perturbed.

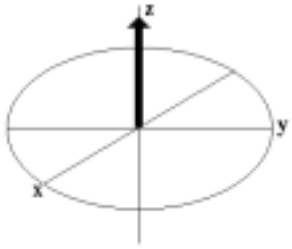


Figure 2.3
Before any pulse is applied (at equlibrium), the net magnetization is oriented along the z-axis (aligned with B_0).

.

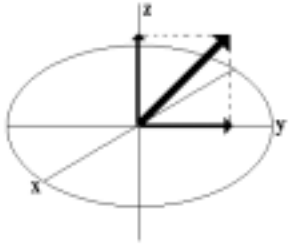


Figure 2.4
Applying a pulse (the B_1 field) along the y-axis will induce a torque which moves the magentization toward the xy-plane (at the same time as it is precessing around B_0). The final position of the magnetization will depend upon the length of time for which the pulse is applied.

The angle, θ (the tip angle or flip angle), through which the magnetization is tipped from the *z*-axis is given by:

$$\theta = \gamma B_1 t_p \qquad (2.4)$$

where t_p is the duration of the pulse (usually a few μs)

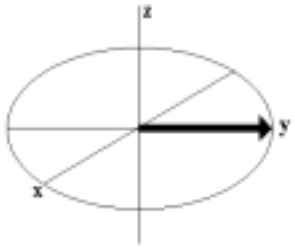


Figure 2.5
If the pulse is exactly long enough to tip the *x*-magnetization from the *z*-axis to the *xy*-plane, this is called the 90 degree pulse, or $\pi/2$ pulse. A $\pi/2$ pulse removes the *z*-axis magnetization and equalizes the populations. If the pulse is applied for twice the time of a $\pi/2$ pulse (a π-pulse), the *z*-axis magnetization and the populations are inverted.

The magnetization will not remain indefinitely along the y-axis after a $\pi/2$ pulse. It will continue to precess around the B_0 at a constant frequency. It will also return back to the z-axis, this is called relaxation (the population difference induced by the B_1 field will return back to equilibrium).

The precession frequency is dependent on the applied magnetic field, but the local chemical environment also influences this precession frequency.

2.1.2 *Fourier Transformation*

If we have a coil wound about the x-axis, the precessing magnetization from the different nuclei will induce an oscillating current which can be detected. This current is called the free induction decay (FID): *free* of the influence of the B_1 field, *induced* in the coil, and *decaying* back to equilibrium. The FID can be transformed to give the frequencies of each of the precessing nuclei giving rise to it. This is done by Fourier transformation.

2.1.3 *The Chemical Shift*

In diamagnetic molecules, the nuclei are shielded from the applied magnetic field (B_0) by the electrons surrounding them. The applied field induces currents in the electrons that give rise to a magnetic field opposing the B_0 field. The effective field (B_{eff}) experienced by the nucleus becomes smaller or larger than B_0. Factors like hydrogen-bonding, aromatic ring currents and anisotropic effects will influence B_{eff} and hence the precession frequency in which they precess around the B_0 field (the z-axis in figure 2.5).

Chemically non-equivalent nuclei are shielded to a different extent, and give rise to different frequencies after Fourier transformation of the FID. In NMR it is not common practise to give the precessing frequency in Hz, but rather the difference in frequency relative to a reference compound. This gives us a unit independent of the

applied magnetic field. This unit is called the chemical shift, δ, measured in parts per million, ppm, given by:

$$\delta = [\nu_{nucleus} - \nu_{reference})/ \nu_{reference}] \cdot 10^6 \quad (2.5)$$

2.1.4 *Scalar Coupling*

Nuclei experiencing the same chemical environment or chemical shift are called equivalent. Those nuclei experiencing different environment or having different chemical shifts are nonequivalent. Nuclei which are close to one another in terms of chemical bond exert an influence on each others effective magnetic field. This effect shows up as splitting of the resonance lines in the NMR spectrum when the nuclei are nonequivalent. If the distance between non-equivalent nuclei is less than or equal to three bond lengths, this effect is observable. This effect is called spin-spin coupling, scalar coupling or simply J coupling.

2.1.5 *Relaxation*

After perturbing a sample from magnetic equilibrium, the net magnetization is not aligned along the applied magnetic field. The mechanisms bringing the magnetization back to equilibrium is time dependent and called relaxation.

The longitudinal ($1/T_1$) relaxation rate of a population of nuclei is the first order rate constant for the equilibration of magnetization along the magnetic field (the *z*-axis). The transverse ($1/T_2$) relaxation rate is the decay of the magnetization in the plane perpendicular to the applied field (the *x,y*-plane). $1/T_1 \leq 1/T_2$ because of dephasing of the magnetization in the *x,y*-plane.

2.1.6 *Relaxation in the Presence of a Paramagetic Probe*

Magnetic nuclei undergo relaxation because they exchange magnetic energy with the environment. Unpaired electrons, with a 657-fold greater magnetic moment than protons, are especially effective in bringing about magnetic relaxation of nuclei.

When a paramagnetic probe is bound at an unique site of an enzyme, as discussed below, the longitudinal relaxation rate ($1/T_1$) of the substrate nuclei may be used to estimate the distances from the probe to these substrate nuclei, based on the modified Solomon-Bloembergen (MSB) theory.

The paramagnetic effect of an unpaired electron (e.g. the unpaired electrons in high-spin (S=5/2) Fe(III) in a metalloenzyme) on the longitudinal relaxation rate ($1/T_{1p}$) of a nearby magnetic nucleus (e.g. a proton nucleus of a substrate which is exchanging into the paramagnetic enzyme complex) depends predominantly on four parameters:

(i) the lifetime of the complex, τ_m,
(ii) the relative stoichiometry of the substrate and paramagnet in the complex, q,
(iii) the correlation time for electron-nuclear dipolar interaction, τ_c,
(iv) the distance from the unpaired electron to the nucleus in the complex, r.

The relations among the parameters can be expressed as follows:

$$\frac{1}{fT_{1p}} = \frac{q}{T_{1\mathrm{M}} + \tau_{\mathrm{M}}} + \frac{1}{fT_{o.s}} \tag{2.6}$$

where f is the ratio of the concentrations of the bound probe and of the total substrate in the solution. $1/T_{\mathrm{o.s.}}$ is the outer-sphere contribution to the relaxation rate, and $1/T_{1\mathrm{M}}$ is the relaxation rate of the substrate nucleus in the ternary enzyme-probe-substrate complex, which is given by:

$$\frac{1}{T_{1\mathrm{M}}} = \left(\frac{\mathrm{C}}{r}\right)^6 \left(f\left(\tau_{\mathrm{c}}\right)\right) \tag{2.7}$$

where C is the product of known constants. Values of C for various probe-nucleus interactions are tabulated elsewhere (Mildvan et al., 1980). The correlation function $f(\tau_c)$ is given by

$$f\left(\tau_{\mathrm{c}}\right) = \frac{3\tau_{\mathrm{c}}}{1+\omega_{\mathrm{I}}{}^2\tau_{\mathrm{c}}{}^2} + \frac{7\tau_{\mathrm{c}}}{1+\omega_{\mathrm{S}}{}^2\tau_{\mathrm{c}}{}^2} \tag{2.8}$$

where ω_I and ω_S are the nuclear and electron resonance frequencies, respectively, at the given magnetic field. Solving (2.6) and (2.8) for the distance *r*, gives us the general equation:

$$r = \mathrm{C}\left[\left(\frac{qfT_{1p}T_{o.s.}}{T_{o.s.} - T_{1p}} - \tau_{\mathrm{M}}\right)\left(\frac{3\tau_{\mathrm{c}}}{1+\omega_{\mathrm{I}}{}^2\tau_{\mathrm{c}}{}^2} + \frac{7\tau_{\mathrm{c}}}{1+\omega_{S}{}^2\tau_{\mathrm{c}}{}^2}\right)\right]^{\frac{1}{6}} \tag{2.9}$$

If the outer sphere contribution ($1/T_{\mathrm{o.s.}}$) can be shown to be small, and the relaxation time ($fT_{1\mathrm{p}}$) is not limited by the lifetime of the complex (τ_M), then equation (2.9) can be simplified as follows:

$$r = C\left[qfT_{1p}\frac{3\tau_c}{1+\omega_I{}^2\tau_c{}^2} + \frac{7\tau_c}{1+\omega_S{}^2\tau_c{}^2}\right] \tag{2.10}$$

Since the electron magnetic moment is 657-fold greater than the nuclear magnetic moment ($\omega_s >> \omega_I$) we can simplify equation (2.10) even more:

$$r = \mathrm{C}\left[qfT_{1p}\left(\frac{3\tau_{\mathrm{c}}}{1+\omega_{\mathrm{I}}{}^2\tau_{\mathrm{c}}{}^2}\right)\right]^{\frac{1}{6}} \tag{2.11}$$

In our study, this equation was used to estimate distances (*r*) from the probe (e.g. Fe(III) in the active site of PAH) to the substrate nuclei (e.g. the protons of L-Phe and BH_2).

The C value in eq. (2.7) for the interaction between high-spin Fe(III) with 5 unpaired electrons (as in PAH) and proton is 812 (Mildvan et al., 1980); ω_I is the precession frequency in radians (e.g. for protons at 600 MHz: $\omega_I = 600 \cdot 10^6 \cdot 2\pi$); q is the binding stoichiometry of the substrate with respect to the paramagnet in the system. For the binding of L-Phe and BH_2 to PAH, q=1.

$1/fT_{1p}$ is a function of the relaxation rate of the ligand (L) protons on the ratio of holoenzyme (holoE) and ligand (L) concentration:

$$\frac{1}{fT_{1p}} = \frac{1/T_1}{[holoE]/[L]} \tag{2.12}$$

Tetrameric wild type PAH has been found to bind 0.5 Fe(III) per subunit, while the truncated dimeric form of PAH (Gly103-Asn428) has been found to bind 0.4 Fe(III) per subunit (Knappskog et al., 1996). The amount of bound holoenzyme forming a hPAH-BH_2-Phe complex, is calculated from the K_D (or K_m/K_i-values) for the binding of the ligand. Both enzyme forms of hPAH used in this study were found to be 98% saturated at the substrate concentrations used (i.e. 5 mM), and the amount of bound holoenzyme (holoE) has been considered to be the same as the added holoenzyme in the calculations.

τ_c is the correlation time, which is defined by:

$$\frac{1}{\tau_c} = \frac{1}{\tau_s} + \frac{1}{\tau_m} + \frac{1}{\tau_r} \tag{2.13}$$

where τ_s is the electron-spin relaxation time, τ_m the lifetime of the substrate-enzyme complex and τ_r is the rotational correlation time.

$1/\tau_c$ is most directly evaluated by measurements of $1/fT_{1p}$ at several magnetic fields. If $1/\tau_r$ and $1/\tau_m$ are found to be small compared to $1/\tau_c$, then $1/\tau_s \cong 1/\tau_c$. To justify the deletion of $1/\tau_m$ in (2.13), the transverse relaxation of the ligand at different enzyme-concentrations can be measured. $1/T_{2p}$ is used to estimate the exchange rate of the substrate into the enzyme. When $1/T_{2p}$ exceeds $1/T_{1p}$ by an order of magnitude or more it may safely be concluded that the rate constant for dissociation of the substrate ($1/\tau_M$) greatly exceeds $1/T_{1p}$, i.e. that the lifetime of the complex (τ_M) contributes little to T_{1p} (Mildvan et al., 1980). $1/\tau_r$ can be estimated from the Stokes' law for the rotation of a rigid sphere in a viscous medium as $\cong 2.7 \cdot 10^{-13} \cdot M_r$. For paramagnetic proteins, $1/\tau_c$ is usually dominated by $1/\tau_s$, and exceeds $1/\tau_r$ by several orders of magnitude.

2.1.7 *Dipolar relaxation*

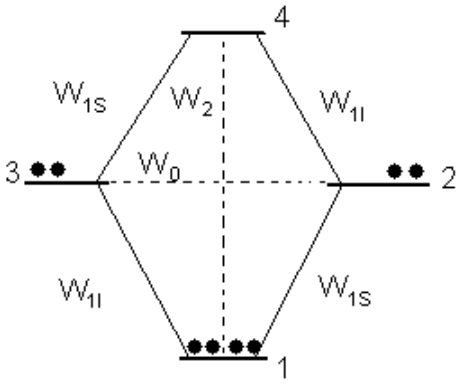


Figure 2.6
Energy levels and equlibrium population distribution for dipole-dipole relaxation.

Figure 2.6 shows the energy level diagram for two protons, I and S, which are relaxing each other but are not J-coupled. The diagram also shows the distribution of spins at equilibrium. After the populations are excited (i.e. moved to a higher energy level in the diagram), they will try to relax back to this equilibrium state. There are several ways the nuclei can relax, each way with a probability W. When the nuclei relax they loose (or gain) energy corresponding to the energy difference between the energy levels. The NMR signal we observe is a result of single quantum processes of relaxation, with probability W_1. Dipole (through space) relaxation occurs when spins exchange energy through double quantum relaxation (W_2) or zero-quantum processes (W_0). These processes are not observed directly, but they can be detected through their effect on single quantum relaxation, and this is the basis of the nuclear Overhauser effect.

If we irradiate the sample at the S nuclei resonance frequency, the population differences for spin S will be equalized, i.e. saturated. The population differences for spin I is not *immediately* influenced by this irradiation. However, the population differences for both W_0 and W_2 are now changed, and the system will try to relax back to equilibrium through these coherence pathways.

Relaxation through W_1 requires magnetic field fluctuations (magnetic noise) near the Larmor precession frequency, ν_0, while W_2 requires fluctuations around $2\nu_0$. Due to the fact that the necessary field fluctuations are produced by protons tumbling at a rate τ_c^{-1}, W_1 and W_2 are most efficient when $\omega_0\tau_c=1$, where ω_0 is the strength of the magnetic field in radians per second ($\omega_0 = \nu_0 \cdot 2\pi$).

If the molecule is very small, it tumbles very fast in solution, and $\omega_0\tau_c>1$. Dipole relaxation is then most efficient through W_2. The population difference will be increased by this relaxation, and give a positive population increase for spin I in Figure 2.6, and hence stronger signal for spin I, i.e. a positive nuclear Overhauser effect (NOE).

On the contrary, when the molecule tumbles very slowly in solution (e.g. a protein), $\omega_0\tau_c<1$, and relaxation through W_0 is more efficient than W_2. The population difference across its transition will be decreased, and give a negative population increase (i.e. a negative NOE) for spin I in Figure 2.6.

2.1.8 *Nuclear Overhauser Effect Spectroscopy (NOESY)*

The NOESY (Nuclear Overhauser Effect Spectroscopy) experiment correlates ^{1}H resonances that are connected via dipolar (through space) coupling, i.e the cross peaks will show which protons cross relax together. The horizontal and vertical axes represent identical proton chemical shift axes; the NOESY spectrum should therefore be completely symmetric with respect to the diagonal.

The NOESY pulse sequence is as follows:

$(\pi/2)$ - t_1 - $(\pi/2)$ - τ_m - $(\pi/2)$ - acquire (t_2)

The first $\pi/2$ pulse creates tranvserse xy magnetization, and the spins now precess during t_1 in the xy-plane. The second $\pi/2$ pulse rotates components of the magnetization along the –z axis. During the subsequent mixing time, τ_m, z-magnetization components exchange under the influence of dipolar relaxation. The third $\pi/2$ pulse regenerates observable magnetization. If this sequence is repeated for a larger value of t_1, the magnetization vectors will dephase further, and a smaller –z component is created, which will pass through zero and become positive for increasing values of t_1. The changes are 'read' by the final $\pi/2$ pulse.

In 2D-NOESY spectra, the diagonal arises from spin vectors that fail to migrate during τ_m. Cross peaks over and under the diagonal are generated from the magnetization transfer between spins, i.e. spins that are dipolar relaxed and experience NOE. The magnitude of the vector represents the population difference between the spins. If this difference is modified by NOE during τ_m, the size of the vectors will be modified before the third $\pi/2$ pulse is applied, and dipolar relaxation can be detected. For large molecules, W_0 dominates over W_2, and the NOESY cross preaks and diagonal have the same phases, i.e all peaks in the NOESY spectrum should be positive. For small molecules W_2 dominates over W_0, and the NOESY cross peaks and diagonal have opposite phases.

The initial build-up rates of cross-peaks in the a NOESY experiment are proportional to the sixth power of the distance between the spins. This gives us a method for quantitative distance determination. Given that you have a known fixed internuclear distance, r_{IM}, r_{IS} can be estimated from the r^{-6} dependence of the cross relaxation rate. The relation between initial cross-relaxation rate and distance is given by:

$$\frac{r_{IM}}{r_{IS}} = \left[\frac{\sigma_{IS}}{\sigma_{IM}}\right]^{1/6} \tag{2.14}$$

where r is the distance between the spins and σ is the initial cross relaxation rate.

2.1.9 *Transferred Nuclear Overhauser Effect*

The TRansferred Nuclear Overhauser Effect (TRNOE) is the extension of the two-dimensional NOE to exchanging systems such as ligand-protein complexes. The intramolecular TRNOE allows the transfer of information concerning cross-relaxation between two nuclei in the bound ligand to the free ligand resonances via chemical

exchange. In the unbound form, the ligand is generally characterized by short correlation times, and thus normally is in the extreme narrowing limit ($\omega_0\tau_c>1$), where the NOEs are positive. When bound to the protein, the ligand is characterized by the long correlation time of the protein, and thus is in the spin diffusion limit ($\omega_0\tau_c<1$), where NOEs are large and negative.

In the presence of chemical exchange of the ligand between its free and bound states, negative NOEs conveying conformational information of the bound ligand are transferred to the free ligand resonances where they are more easily measured because these resonances are much narrower.

The effects of different variables on the TRNOE has been evaluated and reviewed in several publications (Campbell & Sykes, 1991; Campbell & Sykes, 1993). The determination of the effect of these variables are primarily empirical or based on computer simulations. The important parameters to optimize in a TRNOESY experiment is the mixing time, τ_m, and the fraction of bound ligand *vs* total ligand concentration (p_B). Other factors that influence the TRNOE is the free and bound correlation times, τ_c^F and τ_c^B, respectively, that usually can not be optimized.

It has been shown that for fast exchange at equilibrium, the longitudinal magnetization of the spin system decays as a function of the population weighted average of the individual relaxation for the bound and free states (Campbell & Sykes, 1991). Moreover, the intensity of the TRNOE is essentially zero for p_B approaching zero, indicating that the contribution to the TRNOE intensity from the ligand in the free state is negligible. Thus, the observed NOE can be taken to arise solely from the ligand in the bound state. By choosing an appropriate internal standard it is possible to calculate the internal interproton distances in the bound conformation of the ligand. Using these distances as restraints in a distance geometry calculation, it is possible to determine the conformation of the ligand when bound to the receptor.

At short mixing times, the NOE cross-peak intensity is linear on τ_m, but then reaches a maximum and starts to decrease with longer mixing times. It is only the initial buildup rate of NOE that can be correlated directly to distance between the spins. The buildup rate is also dependent on p_B, the fraction of bound ligand. Thus, these parameters have to be optimized in order to use the NOEs for calculating distances. The TRNOE intensity *versus* mixing time (τ_m) for various bound correlation times (τ_c^B) and fraction of bound ligand (p_B) is shown in Figure 2.7.

As seen in this figure, at high p_B values, the TRNOE increase rapidly to reach a maximum and then decays. At lower p_B values, the TRNOE increase linearly over a longer range of τ_M values. Thus, it seems best to use low p_B values, accompanied with reasonable τ_M values in order to obtain optimal experimental signal-to-noise. Especially when the receptor is large (large τ_c^B) and spin diffusion influence the development of TRNOE, it is important to use low p_B values. This may be one of the few examples in biochemistry where less protein gives better results.

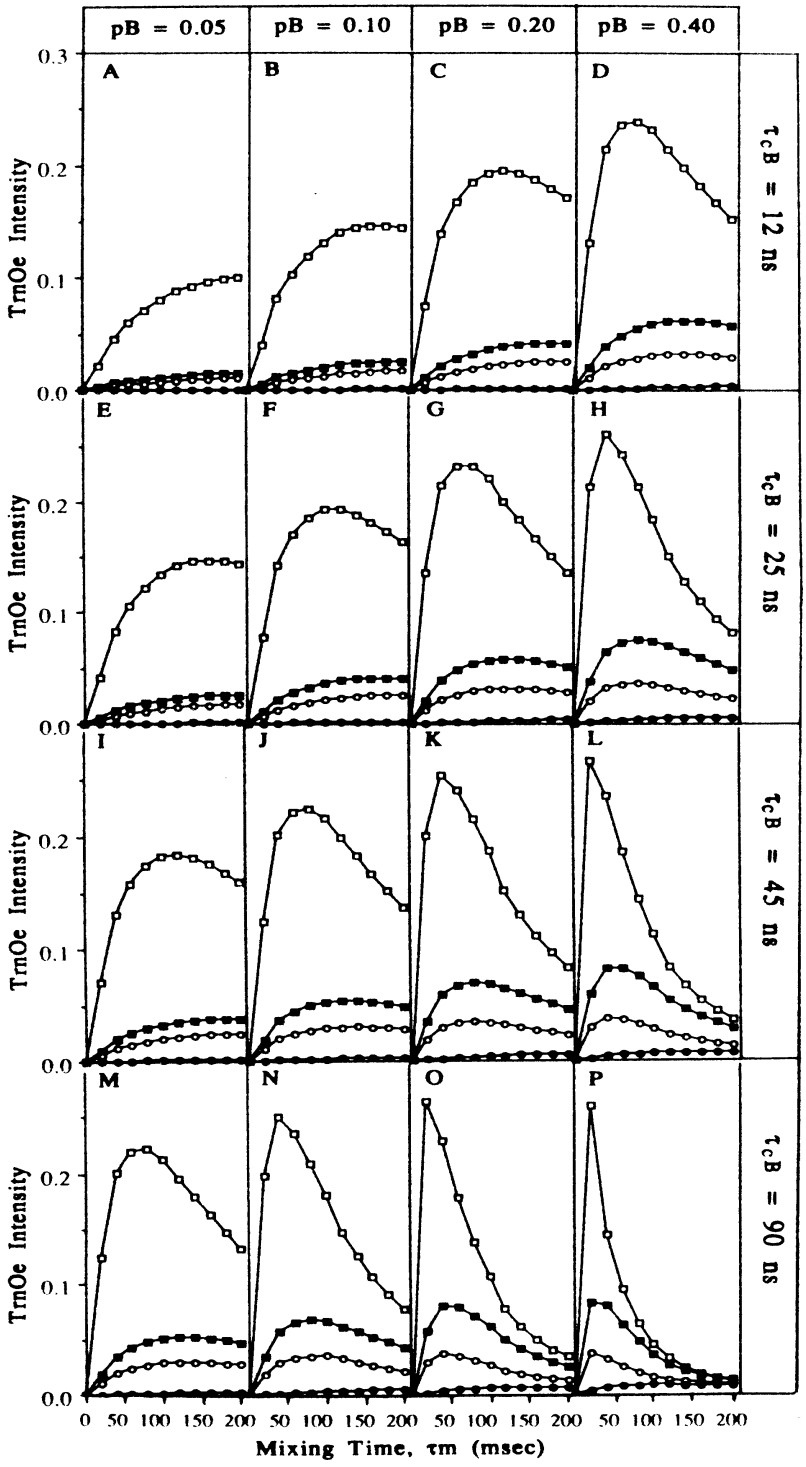


Figure 2.7
TRNOE intensity *versus* mixing time (τ_M) for various bound correlation times (τ_c^B) and fraction of bound ligand (p_B), from (Campbell & Sykes, 1991).

2.1.10 *Assignment of Resonances*

In order to gain any information from the TRNOESY and proton relaxation experiments, the different signals and cross peaks have to be assigned to specific protons of the ligands. This is accomplished by the use of heteronuclear quantum coherence correlation experiments.

2.1.10.1 Heteronuclear Single Quantum Coherence (HSQC)

The HSQC experiment correlate ^{1}H and ^{13}C chemical shifts through one-bond heteronuclear scalar coupling. The cross peaks in a 2D ^{1}H-^{13}C HSQC spectrum show the chemical shifts of ^{1}H and ^{13}C resonances that belong to H- and C-atoms that are directly bonded to each other. The horizontal axis represents the ^{1}H chemical shifts, whereas the vertical axis represents the ^{13}C chemical shifts. The sequence includes so called inverse detection, i.e. ^{1}H is detected rather than ^{13}C. This gives us a considerable increase in the intensity of the ^{1}H-^{13}C-cross peak signals.

2.1.10.2 Heteronuclear Multible Bond Coherence (HMBC)

A related experiment to the HSQC technique is the HMBC experiment, which correlate ^{1}H and ^{13}C chemical shifts through multiple bond heteronuclear scalar coupling. The cross peaks in a 2D ^{1}H-^{13}C HMBC spectrum show the chemical shifts of ^{1}H and ^{13}C resonances that belong to H- and C-atoms that are separated by (usually) two or three chemical bonds.

2.1.11 *Gradients*

Application of constant field gradient G_k along an axis k (x,y or z) alters the precession frequency of the spins as a function of position:

$$\Delta\omega=\gamma G_k k \tag{2.15}$$

If a field gradient is applied for a time t_g following a 90° pulse, there will be a position-dependent dephasing ($\delta\omega t_g$) of the spin alignment responsible for the transverse magnetization over the dimensions of the sample. The bulk NMR signal decays much faster than by T_2 processes alone. However, this dephasing is spatially encoded and can be reversed (refocused) if a gradient with the same amplitude but reversed polarity is applied for the same time t_g.

The description above holds for the effects of a gradient on precessing single-quantum coherence. If two radio frequency pulses are applied to a homonuclear coupled spin system higher order of coherence is generated. Application of a gradient to precessing double-quantum coherence, for example, has twice the effect, and the induced phase shift is $2\delta\omega t_g$.

Gradient selected experiements rely on the fact that another identical gradient applied at a later stage of the pulse sequence can rephase the coherences if their coherence level was changed, for example by a 180° pulse. Thus, one is able to select coherence pathways by combining radio frequency pulses and pulsed field gradients in one pulse sequence. Since the NMR receiver detects now only the desired signals, its gain can be set much higher, reducing the time requirement for an

experiment drastically compared to a non gradient selected experiment. This technique is also excellent for solvent supression.

2.1.12 *Water Supression*

Since the molarity of protons in water is 110 M, for ^{1}H NMR of 1-5 mM samples in 90% water, there is a real problem in detecting the sample resonances in the presence of such a large water resonance. Two very important types of approaches to water (or solvent) suppression are:

(i) pre-irradiation with low-power frequency-selective continous-wave radio frequency,

(ii) selective excitation of the water resonance with a series of either long, weak (soft) pulses or short, strong (hard) pulses separated by delays.

All experiments in this study use selective excitation (or actually inversion) of the water resonance (approach ii) in order to suppress the solvent signal. In the 1D experiments, the water signal was suppressed using WATERGATE (WATER suppression by GrAdient Tailored Excitation) (Piotto *et al*., 1992). Following the initial nonselective pulse, a strong gradient pulse dephases both solvent and solute magnetization. Solute magnetization is unaffected by the selective pulses. The nonselective 180° pulse inverts the coherence order of the solute magnetization; therefore, the second gradient pulse rephases the solute magnetization to form a gradient echo. In contrast, the combination of the two selective 90° pulses and the nonselective 180° pulse leaves the coherence order of the solvent magnetization unchanged; therefore, the second gradient pulse continues to dephase the solvent solvent magnetization, and no gradient echo is formed. In the NOESY experiments the 3-9-19 modification of the Watergate pulse sequence for suppression of the water signal (Sklenar *et al*., 1993) has been used.

2.2 Molecular Modeling

In this study, molecular modeling is used to:

(i) Construct conformers of the ligands compatible with restraints obtained from NMR (by distance geometry calculations)

(ii) Find the most optimal positioning of the ligands within the active site of the enzyme (by docking)

2.2.1 *Distance Geometry*

Methods of calculating the conformations of biological molecules from restraints including the possible values of their interatomic distances and dihedral angles, together with chirality constraints, are important tools in structural biochemistry. These calculations, known as distance geometry calculations, have been used in drug design (Billeter *et al*., 1986; Sheridan *et al*., 1986), in structural interpretation of NMR data (Kuntz *et al*., 1989; Wütrich, 1986) and in prediction of protein structure from homologous sequence alignment (Havel, 1991). Distance geometry calculations do not yield a single spatial structure, but rather an ensemble of different conformations that satisfy the restraints and constraints.

In this study, distance geometry calculations were conducted using the program package DGII under InsightII. The package is a combination of the EMBED algorithm (Havel *et al*., 1983) and simulated annealing.

The DGII algorithm proceeds in three distinct steps:

(i) *Bound Smoothing*

In this step, the tightest possible limits on the values of all the interatomic distances are extrapolated from the incomplete matrix of distance bounds that is usually available (obtained for instance by TRNOESY NMR experiments).

The simplest way to accomplish this is to use the lower and upper distance bounds, l_{ij} and u_{ij}, respectively, together with the triangle inequality to compute limits on the distance d_{ij}, *via* the relation

$$l_{ik}-u_{jk} \leq d_{ij} \leq u_{ik} + u_{jk}$$

This is called triangle bound smoothing (Dress & Havel, 1988). With the DGII package, it is also possible to do tetrangle bound smoothing. This is done by using a system of four-atom relations derived from the so called tetrangle inequality. Much tighter limits on the distances can be accumulated, but at the expense of increased computer time. But as long as the structure studied is not a large protein, this can be achieved in reasonable computer time. The distances obtained are used in the embedding step of the DGII procedure.

(ii) *Embedding*

In this step, a matrix of exact values for the interatomic distances from between their respective lower and upper limits is guessed, and a set of coordinates whose distances are a best-fit to this guess is computed.

The embedding step consists of three parts:

(a) Choosing a matrix of random trial distances which obey the triangle inequality from between their limits by a procedure known as metrization.

With sufficiently complete and precise sets of distance restraints, we could simply choose the distances from between their lower and upper limits with an uniform distribution. However, a procedure known as metrization is advisable in order to obtain the widest possible sampling of conformations. The idea is to use the triangle distances calculated in the preceding bound smoothing. The triangle inequality limits are equal to the extremes that the distances can assume in any metric space consistent with the bounds. If we arbitrarily set any distance to any value between its limits, we then practically recalculate the triangle inequality limits. This is demanding in terms of computer capacity, and another procedure called shortest-paths tree is used. This enables the limits from one atom to all the others to be efficiently recomputed after each change. Thus, the distances from one atom to all the other atoms must be set first, followed by the distances from the

next atom to all the remaining atoms, and so on. This is known as prospective metrization. It is also possible to compute the shortest paths tree from each new atom to all the preceding atoms, fix these distances, and proceed to the next atom, and so on. In this case, the algorithm is called retrospective metrization.

(b) Converting this random distance matrix into a set of random atomic coordinates by a procedure known as embedding.

The basic idea of embedding is to compute conformers of the structure with coordinates whose distances are a close fit to the trial distances.

(c) Improving a weighted least squares fit between the trial distances and the coordinates by a procedure known as majorization.

The fit obtained from embedding probably does not yield the best possible starting coordinates for optimization. By successive transformations of the coordinates, known as Guttmann transformations (Havel, 1991), more chemically reasonable starting coordinates are generated.

(iii) *Optimization*

In this step, an error function that measures the violation of the distance restraints and chirality constraints is minimized by these coordinates until the constraints are fully satisfied.

In order to reduce the violations of the constraints by the embedded coordinates to an acceptable level, further optimization is necessary. This involves minimizing a function which measures the total violation of the constraints by the coordinates, which is called an error function. This error function tries to optimize the coordinates, while enforcing the upper and lower bounds and the chirality constraints present.

2.2.1.2 Optimization by Simulated Annealing

The simulated annealing procedure employed by DGII differs from most other simulated anneling procedures. Traditionally, simulated annealing has been used to eliminate the residual violations in a structure that has already been minimized. Rather than taking a structure that has become trapped in a local minimum during minimization and use dynamics to shake it out of that local minimum, simulated anneling is applied directly to the embedded structures.

The error function is scaled to an energy that is sufficient to cause the system to heat up naturally to the desired temperature (i.e. average kinetic energy per degree of freedom). Whenever the temperature rises too rapidly or higher than the parameter T_{max}, the error function is rescaled to keep the temperature within the bounds. In traditional simulated annealing, T_{max} is 1000 K or more. In the DGII

procedure, T_{max} is typically 200 K and reduced to zero during the annealing procedure.

The coordinates obtained from this annealing procedure are generally close to a minimum, hopefully the global minimum. The ensemble of structures generated has to be carefully studied for violations of experimental data. The final structures accepted generally can be divided into families of conformers. In order to reduce the number of families, optimization of the distance geometry calculations have to be conducted, or further experimental data has to be collected.

2.2.2 *Docking*

The DOCK program explores possible orientations of a molecule within a macromolecular active site by superimposing atoms onto precomputed site points. The docking procedure can be divided into six steps:

(i) *Preparing the target receptor for docking*

Normally, the coordinate-file for the receptor is retrieved in pdb-format. This format identifies the atoms and their position, but it does not contain information about the potential of the atoms. This information has to be included in the coordinate file in order to do use it in DOCK. One way of doing this is to read the file into InsightII and save it as mol2 format. InsightII will then calculate the atom type and potential of each atom and include it in the mol2 file, using one of the force fields incorporated in InsightII (i.e. Amber, CVFF, CFF or ESFF).

(ii) *Generating molecular surface for receptor*

There are several ways this can be done. In this study, the DMS (Dot Molecular Surface) program under UCSF MidasPlus was used to generate the solvent accessible surface of the receptor. Only the surface for a region around the active site needs to be generated.

(iii) *Generating spheres to fill the active site*

The shape of the cavities in the receptor is used to define spheres. The centers of the spheres become potential locations for ligand atoms. For this purpose, the program SPHGEN (DesJarlais *et al*., 1988) under the DOCK suite of programs is used. SPHGEN generates sets of overlapping spheres to describe the shape of the molecular surface. The spheres have varying radii and touch the molecular surface at just two points. Spheres are calculated over the entire surface, producing approximately one sphere per surface point. This very dense representation is then filtered to keep only the largest sphere associated with each receptor surface atom. A file is generated that consists of the coordinates for all the sphere center points. The sphere centers can be manually edited to only include surface points that are within the active site.

(iv) *Calculation and assignment of properties for each site point*

To make the actual docking less time consuming, information about the steric and electrostatic environment at each site point is calculated and saved. This is done

by using the program GRID within the DOCK suite of programs. Four output files are generated by grid which hold the bump grid, contact grid, chemical grid and force field grid. These files are used independently in the docking, so that ligand orientations can be scored rapidly. The content and use of the different grid files are explained under step (vi), scoring (see below).

(v) *Matching*

Sphere centers are matched to the ligand atoms to determine possible orientations for the ligand. Typically on the order of tens of thousands of orientations are generated for each ligand molecule.

The matching is done between the distances among ligand atoms and distances among receptor sphere centers. This is visualized in figure 2.8.

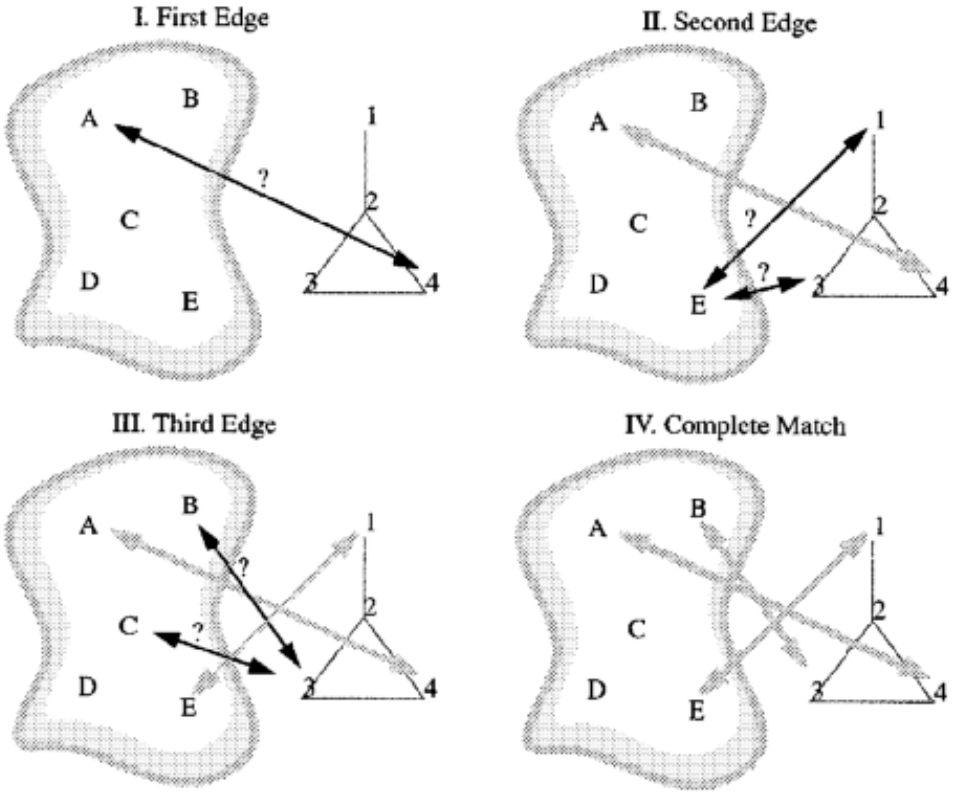


Figure 2.8
The matching procedure. A-E are receptor site points, 1-4 are separate ligand atom centers. I. Out of the 20 possible combinations (5 site points and 4 atoms), an arbitrary site point and atom center is chosen. In this case, atom 4 is superimposed onto site point A. This match would fix three of six orientational degrees of freedom. II. Second site points and atom centers are then considered. Of the 12 possibilities now available, A·4 is considered. Because $r_{A-E} > r_{4-3}$, E·3 is discareded as fit. Then E·1 is considered. Because $r_{A-E} = r_{4-1}$, E·1 is retained as a fit. In three dimensions, this match would superimpose atoms 4 and 1 onto points A and E, respectively. This match would fix two more orientational degrees of freedom. III. Third sets of site points and atom centers are considered. Of the six to be tried (3 site points left x 2 atoms left), C·3 is considered. Through $r_{A-C} = r_{4-3}$, $r_{E-C} < r_{1-3}$, so C·3 must be discarded as a third fit. Then B·3 is considered. Because $r_{A-B} = r_{4-3}$ and $r_{E-B} = r_{1-3}$, B·3 is retained as a third fit. This match fixes the last of six orientational degrees of freedom. IV. The match is large enough to define a unique orientation which superimposes atoms 4,1 and 3 onto site points A, E and B, respectively.

(vi) *Scoring*

Each oriented molecule is then scored for fit. This is done using the precalculated grid files (step 4). The grid files are independent of each other, and the three different scoring grids might result in different scoring for the same oriented molecule. In this study, only the *bump grid* and *energy grid* was used in the docking procedure.

Bump checking

Prior to scoring, each orientation can be processed with the bump filter to reject orientations that are in severe steric overlap with a receptor atom. The bump grid stores an atomic radius which corresponds to the smallest radius of a ligand atom at the grid position which would trigger a bump. During bump checking, for a given orientation, the position of each atom is checked with the bump grid. The penalty of a bump is controlled by the user. It is usual to make a bump give a great negative contribution to the total score of an orientation, so that orientations that overlap with the receptor will not be processed further.

Energy scoring

The energy scoring in DOCK is based on the implementation of force field scoring, consisting of van der Waals and electrostatic components:

$$E = \sum_{i=1}^{lig} \sum_{j=1}^{rec} \left(\frac{A_{ij}}{r_{ij}^{\,a}} - \frac{B_{ij}}{r_{ij}^{\,b}} + 332 \frac{q_i q_j}{D r_{ij}} \right) \qquad (2.15)$$

where each term is a double sum over ligand atoms *i* and receptor atoms *j*, and

E	is the intermolecular interaction energy,
r_{ij}	is the distance between atoms *i* and *j,*
A_{ij} and B_{ij}	are the van der Waals repulsion and attraction parameters, respectively,
a and *b*	are the van der Waals repulsion and attraction exponents, respectively,
D	is the dielectric function,
332	is a factor for converting electrostatic energy in coloumbs/mol to kcal/mol.

The van der Waals component of the scoring function is generalized in the scoring procedure, so that any combination of repulsive and attractive exponents can be used (as long as a>b).

The van der Waals interaction between two atoms is given in equation (2.16).

$$E_{vdw} = \varepsilon\left(\frac{b}{a-b}\right)\left(\frac{2R}{r}\right)^{a} - \varepsilon\left(\frac{a}{a-b}\right)\left(\frac{2R}{r}\right)^{b} \qquad (2.16)$$

where ε is the well depth of the interaction energy (see figure 2.9) and R is the van der Waals radius of atoms.

The most common values for *a* and *b* are 12 and 6, respectively. This gives us the 6-12 Lennard-Jones potential, which is computationally fast to calculate, since the r^{-12} term can be calculated from the square of the r^{-6} term. Also, the r^{-6} term can be calculated from the square of the distance, without having to perform computationally expensive square root calculations.

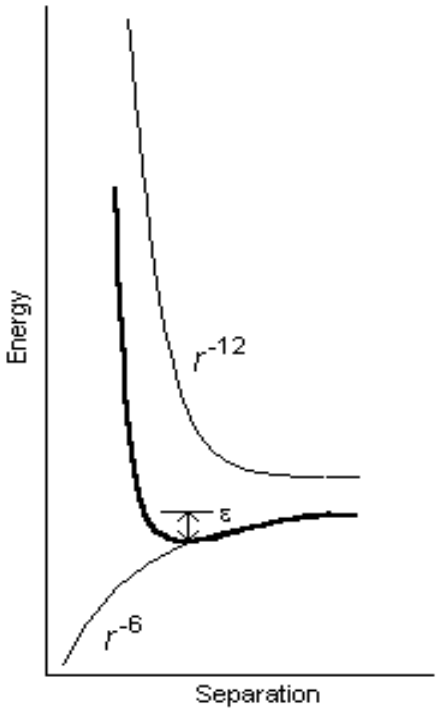


Figure 2.9
The Lennard-Jones potential is constructed from a repulsive (r^{-12}) and an attractive component (r^{-6}).

In order to evaluate the interaction energy quickly, the van der Waals and electrostatic potentials are precomputed for the receptor and stored in a grid file. Three values are stored for every grid point, each a sum over receptor atoms that are within a user-defined cutoff distance of the point:

$$B_{rec} = \sum_{j=1}^{rec} \frac{\sqrt{B_{jj}}}{r_{ij}^{b}} \tag{2.17}$$

$$A_{rec} = \sum_{j=1}^{rec} \frac{\sqrt{A_{jj}}}{r_{ij}^{a}} \tag{2.18}$$

$$Q_{rec} = 332 \sum_{j=1}^{rec} \frac{q_j}{Dr_{ij}} \tag{2.19}$$

Equations (2.17), (2.18) and (2.19) represents the van der Waals attraction parameter, the van der Waals repulsive parameter and the electrostatic parameter, respectively. These values are multiplied by the appropriate ligand values to give the interaction energy. The receptor values are read from the grid file during a DOCK run and used for force field scoring.

The interaction energy is given by:

$$E = \sum_{i=1}^{lig} \left(\sqrt{A_{ii}} A_{rec} - \sqrt{B_{ii}} B_{rec} + q_i Q_{rec} \right) \tag{2.20}$$

Atoms that fall outside the grid, if any, are given interaction energies of zero.

Contact scoring

The contact score is a summation of the heavy atom contacts (every atom except hydrogen) between the ligand and the receptor. A contact is defined as an approach of two atoms within some user defined cutoff distance (default 4.5 Å). If the two atoms approach close enough to bump, then the interaction will be penalized by an amount specified by the user. The contact scoring can be seen as a simplification of the energy scoring. Since the calculations are rather simple, this type of scoring might be the choice if a large database of ligands is to be screened (since energy scoring will be more time consuming). But for this study, contact scoring gave very poor results.

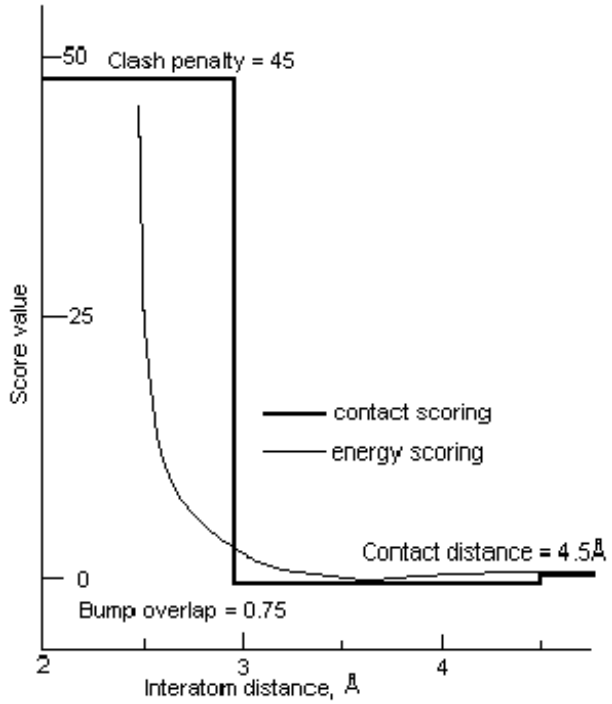


Figure 2.10
Comparison of the contact and energy scoring functions.

Chemical scoring

Chemical scoring is also based on energy scoring, but incorporates empirical concepts of molecular interactions. It includes the van der Waals and electrostatic terms from energy scoring, but the attractive portion of the van der Waals term has been modified. The attractive portion is scaled depending on the chemical nature of the interacting atoms. In order to scale atoms differently, the user has to define chemical labels of specific atoms in a file used by grid. This file arrange different atoms in groups based on atom type and types of adjacent atoms. A group of atoms might be labeled as hydrophobic, polar, donor or some arbitrary name defined by the user. Then, another file has to be made which tell which chemical labels can perform an interaction in matching. This scoring scheme might be the choice if the user wants to force the docking of a ligand to a specific site in the receptor. One of the atoms in the ligand might be given a label that match only with one of the site points in the grid. In this way, a ligand atom might be forced to overlap with a specific site point in the active site.

Anchor Search

The anchor search is based on the separation of the ligand into rigid segments. Each segment contains the largest set of adjacent atoms separated by non-rotatable bonds. Segments are separated by rotatable bonds. By default, the largest segment is chosen to be the anchor in the docking, but this can be controlled by the user. An example is shown in figure 3.3 where phenylalanine is separated into

segments. The basic idea of the anchor search procedure is to first dock the anchor in the active site, either manually or letting DOCK find the optimal orientation using one of the scoring schemes described above. The remaining segments are subsequently re-attached during the conformation search. The segment closest to the anchor is naturally attached first. If a segment is connected to several others, then the largest segment will be attached first. The orientation of the complex can be reoriented or optimized after each addition of a segment, until the whole ligand molecule is built. It is possible to manually orient the anchor in the active site, and not let DOCK reorient it until after the whole ligand is built. In this way, you might have a high degree of control with the final position of the ligand within the active site.

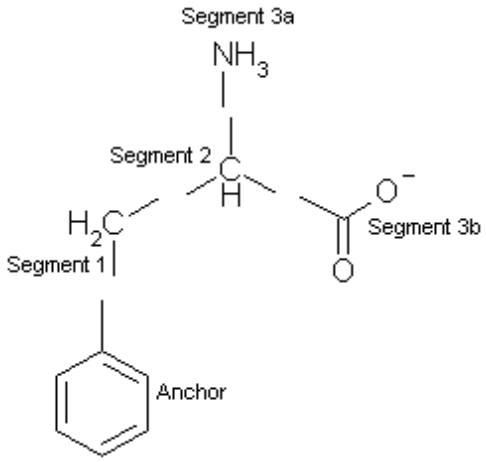


Figure 2.11
Segments in anchor-search procedure for docking.

3 Materials and Methods

3.1 Expression and Purification of hPAH

Expression in *Escherichia coli* (TB1 cells) of wild-type hPAH (wt-hPAH) and the truncated hPAH(Gly103-Gln428), i.e. ΔN102/ΔC24-hPAH, as fusion proteins with maltose-binding protein (MBP), preparation of the cell lysates, purification of the fusion proteins by affinity chromatography on amylose resin and cleavage by the restriction protease factor X_a was performed as described (Knappskog et al., 1996; Martinez et al., 1995). The tetrameric form of wt-hPAH and the dimeric form of hPAH(Gly103-Gln428) were separated from aggregated forms and from MBP and factor X_a by size exclusion chromatography on a HiLoad Superdex (Pharmacia) column (16cm·60cm) (Knappskog et al., 1996). Protein concentration was estimated spectrophotometrically using the absorption coefficients $A_{280\ nm}$ (1 mg × ml^{-1} × cm^{-1}) = 1.0 for wt-hPAH and 0.9 for hPAH(Gly103-Gln428) (Knappskog et al., 1996). The absorption coefficients were calculated from the absorbance at 280 nm, the content of pure protein estimated by amino acid analyses and the amino acid composition of the protein.

3.2 Activity Measurements

The assay of hPAH activity was performed as described (Martinez et al., 1995). Steady-state kinetic studies on the inhibition of hPAH by BH_2 were performed at 1 mM L-Phe (Sigma Chemical Co.), variable concentrations (0-0.5 mM) of (6*R*)-L-*erythro*-5,6,7,8-tetrahydrobiopterin (BH_4, Dr. B. Schircks Laboratories) and in the absence and the presence of variable concentrations (up to 0.2 mM) of L-*erythro*-7,8-dihydrobiopterin (BH_2, Dr. B. Schircks Laboratories). The enzyme samples were preincubated with 1 mM L-Phe at 25 °C for 5 min prior to the 1 min assay in order to study the BH_4-dependent hPAH activity with fully activated enzyme.

3.3 NMR Data Collection and Processing

3.3.1 *General*

Experiments were performed on a Bruker DRX 600 MHz spectrometer equipped with pulse-field-gradient accessories. The probe temperature was 298 K in all experiments. Data processing was done using the Xwinnmr (Bruker) software package.

3.3.2 *Preparation of Samples for NMR*

hPAH samples (about 50 mg ml^{-1}) were initially prepared in 20 mM Na-Hepes, 0.2 M NaCl, pH 7.0. The samples were then subjected to several cycles of dilution/concentration in Amicon Centricon 30 microconcentrators in order to exchange the buffer to 20 mM 10% deuterated potassium phosphate, 0.2 M KCl, pH 7.2. BH_2 and L-Phe at the indicated concentrations were prepared together or in separated samples in the same buffer.

3.3.3 *Assignment of Resonances*

Resonance assignments were made from ^{1}H-^{13}C HSQC and HMBC correlation experiments. The horizontal axis in these experiments represents the ^{1}H chemical shifts, whereas the vertical axis represents the ^{13}C chemical shifts. Carbon and proton atoms that are linked through bonds are identified as cross peaks in the spectrum (see under Theoretical Considerations, 2.1.10), spectra are shown in the Appendix I. The assignment is essential for the interpretation of the NOESY and relaxation data.

3.3.4 *Relaxation Measurements*

The paramagnetic effects of enzyme bound Fe(III) (S=5/2) on the longitudinal relaxation rates ($1/T_{1P}$) of proton resonances of enzyme bound ligands were measured on samples (0.5 ml) containing initially L-Phe and/or BH_2 (5 mM each) which were titrated with enzyme, up to a final concentration of 0.12 mM hPAH subunit. T_1-values were measured by the inversion recovery method with presaturation of the water signal. A 180° pulse is applied to invert the magnetization to the –z axis. The magnetization decays exponentially. After a delay of τ seconds, a 90° pulse is applied to detect the remaining magnetization along the z-axis.

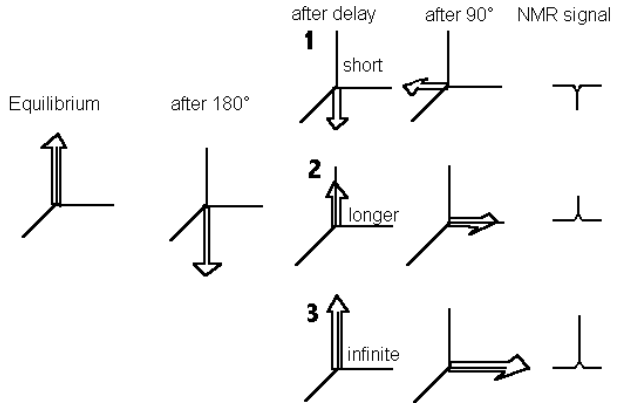
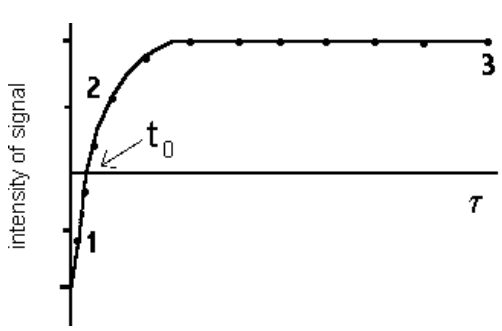


Figure 3.1
Measuring longitudinal relaxation by the inversion recovery method.

The intensity of the signal is plotted *versus* the delay (τ). The τ-value for which the z-magnetization is zero (t_0), may be used to calculate T_1. In practice, the XWinNMR program software uses a three parameter non linear fit for the curve to calculate T_1.



$$T_1 = t_0 / \ln 2$$

Figure 3.2
Intensity of NMR signal *vs* delay time between the 180° and 90° pulse in the inversion recovery experiment.

In the T_1 measurements, 16384 data points were collected for each transient, the spectral width was 4800 Hz, 4 transients were averaged for each τ-value and the recycling delay was 10 s. The observed longitudinal relaxation rate ($1/T_1$) at each enzyme concentration is equal to $1/T_{1P} + 1/T_{1D}$, where $1/T_{1D}$ is the diamagnetic contribution of the protein due to the effects of protein residues-solvent interactions and $1/T_{1P}$ is the paramagnetic contribution to the longitudinal relaxation rate. $1/T_{1D}$ was estimated using apoenzyme forms without iron (Olafsdottir & Martinez, 1999) and was found to be non-significant compared with $1/T_{1P}$. The T_2-values were estimated from the width of these resonances at half-height ($\Delta\nu_{1/2}$), where $1/T_2 = \pi \cdot \Delta\nu_{1/2}$.

The paramagnetic effects of the hPAH-bound Fe(III) on the relaxation of the ligands were analyzed by plotting the increase in relaxation rate as a function of added enzyme in order to calculate the normalized paramagnetic relaxation rates ($1/fT_{1P}$) (Martinez et al., 1993a).

The effective correlation time for the electron-nuclear dipolar interaction (τ_c) was determined to be $(1.2 \pm 0.2) \times 10^{-10}$ s from the frequency dependence of the $1/fT_{1P}$-values at 200, 400 and 600 MHz (Olafsdottir & Martinez, 1999). Estimates of the iron-proton distances were calculated from the $1/fT_{1P}$-values using the modified Solomon-Bloembergen equations (Chapter 2, Theoretical Considerations).

3.3.5 *Nuclear Overhauser Effect Spectroscopy*

NOESY spectra of L-Phe and BH_2 in the absence of enzyme were acquired with a mixing time of 1 s in the phase-sensitive mode using the States-TPPI quadrature detection scheme in ω_1 and utilizing the «3-9-19» modification (Sklenar et al., 1993) of the «Watergate» pulse sequence (Piotto et al., 1992) for suppression of the water signal. The acquisition parameters included a 2 s recycling delay, 4807.6 Hz spectral width, 32 transients per t_1-increment. A 2D raw data set comprising 256 x 2048

States-TPPI time domain data points (t_1,t_2) was obtained. The processing parameters included linear prediction in t_1 with 512 estimated points and a squared sine-bell apodization function shifted by $\pi/2$ in both directions. Matrices of 2048 x 2048 points real data were then obtained. Transferred NOEs from the bound to the free state of both pterin and amino acid substrates were measured in NOESY spectra acquired at different mixing times (20, 35, 50, 75 and 100 ms) in samples containing L-Phe and/or BH_2 (5 mM each) and either wt-hPAH or hPAH(Gly103-Gln428) at a final subunit concentration of 0.12 mM subunit (or 0.5 mM when indicated). The acquisition parameters included a 1.1 s recycling delay, a 6000.6 Hz spectral width and 16 transients per t_1-value. A 2D raw data set of 1024 x 2048 States-TPPI time domain data points (t_1, t_2) was acquired. The processing parameters for TRNOESY spectra included linear prediction in t_1 with 2048 estimated points and the rest of the processing was performed as for the NOESY experiments with the ligands in the absence of enzyme. Intensities of transferred NOE cross peaks were measured by volume integration. Approximate interproton distances were calculated by plotting the increase in cross peak volume as a function of mixing time. The distance of 2.46 $\pm$ 0.02 Å between the H4 and the H3,5 protons of L-Phe as measured by neutron diffraction (Al-Karaghouli & Koetzle, 1975) was used as a reference, thus allowing the use of initial slopes for estimating internuclear distances.

3.4 Distance Geometry Calculations

Graphical visualization, structure generation, and structural analysis of the calculated conformers were performed using the modules NMR Refine, Discover (with the CVFF forcefield) and Analysis within the Insight II 95.0 software package (BIOSYM/MSI). A series of conformers were generated from metal-proton and NOE distance restraints supplemented by chirality constraints by simulated annealing using NMR Refine DGII (see Chapter 2). Degenerated H7 and CH_3 protons of BH_2 were treated considering an average distance from their geometric center. Degeneracy of the chemical shift of the 3 and 5 protons and the 2 and 6 of L-Phe was treated by considering asymmetric and symmetric interactions in the intermolecular NOE-based distances and the Fe(III)-proton distances (Martinez *et al*., 1993b). The symmetric and asymmetric cases were considered separately. Calculations included triangle bond smoothing and embedding in four dimensions using metrized distance matrices. The embedded set of coordinates were refined using simulated annealing optimization followed by a conjugate gradient minimization of the error function to obtain the converged distance geometry based structures (Havel, 1991).

3.5 Docking

The DOCK 4.0.1 suite of programs (University of California, San Francisco) (Ewing & Kuntz, 1996) was used in the final modeling steps in order to fit the conformers of the bound ligands obtained by ^{1}H NMR into the crystal structure of the catalytic domain hPAH(Gly103-Gln428) (Erlandsen et al., 1997). A grid was constructed with a distance of 20 Å around the Fe(III) atom. A grid space of 0.2 Å and a sampling size of 7000 structures were used. The "bound"-structure of BH_2 obtained by NMR and DGII was docked into the enzyme without allowing flexibility of the ligand, i.e. no optimization of the structure was allowed during the procedure. Due to complications with delocalization of charge from the active site iron in hPAH, the iron

was treated as a sphere without charge in the docking procedure. The sampled structures were scored using the energy scoring function (see Chapter 2). Contact and chemical scoring was also tried, but gave poor results.

L-Phe was docked into the active site using the anchor-grow procedure (see Chapter 2). The ring of L-Phe was initially manually docked into the active site of hPAH, positioned in close proximity to Trp326 and the proton at position 4 in L-Phe pointing towards Glu330. Then DOCK was allowed to optimize the position of the ring before the rest of the amino acid was grown from the phenyl-anchor. The observation of interproton cross peaks from the L-Phe ring protons with ppm-values in the protein corresponding to a tryptophan and a glutamate in the TRNOESY spectra (later identified as Trp326 and Glu330) justify the positioning of the ring prior to the anchor-grow procedure.

3.6 Programs Used

3.6.1 *Processing of NMR data*

Data processing was done using the Xwinnmr (Bruker) software package. SigmaPlot (Jandel Scientific) was used to analyze the processed NMR data.

3.6.2 *Distance Geometry Calculations*

Construction of coordinate files was done using the Builder module within the InsightII II 95.0 software package (BIOSYM/MSI). Preparation and optimization of the coordinate files were done using Discover (with the CVFF forcefield) and Analysis within InsightII. The distance geometry calculations were done using DGII, included in the NMR refine module of InsightII.

3.6.3 *Docking*

The most favorable orientations of the ligands in the receptor was done using the DOCK 4.0.1 suite of programs (University of California, San Francisco) (Ewing & Kuntz, 1996). MidasPlus (UCSF) (Ferrin *et al*., 1988) and SYBYL (Tripos Associates, 1991) were used to visualize the docked structures and to measure intra- and inter-molecular distances and angles.

3.6.4 Preparation of Figures

The programs WebLabWiever (MSI), MidasPlus (UCSF) (Ferrin *et al*., 1988), Raster-3D (Merrit & Murphy, 1994) and Molscript (Kraulis, 1991) were used in addition to InsightII (MSI) to visualize the coordinate files of the bound conformers and to prepare figures of structures.

4 Results

The present study consist of mainly three parts.

(i) Kinetic studies
(ii) NMR studies
(iii) Molecular modeling

The kinetic properties of the system were studied by activity measurement assays. The kinetic study gave an estimate for the affinity of the hPAH and its truncated form for the ligands L-Phe and BH_2. This is required to estimate certain NMR parameters (e.g. ratio of ligand and enzyme concentration and mixing time for the NOESY experiements). However, these parameters still had to be optimized in order to get the experiments to work properly. From the NMR experiments, several restraints for the ligand proton-proton distances could be extracted. Distance geometry calculations were used to find conformers compatible with these restraints. The conformers compatible with all the restraints obtained from NMR were docked into the active site of the enzyme.

4.1 Kinetic Studies

Two forms of recombinant hPAH were studied, i.e. the tetrameric wt-hPAH, which requires prior incubation with L-Phe to show full BH_4-dependent activity, and the dimeric hPAH(Gly103-Gln428), which is a truncated form, i.e. ΔN102/ΔC24-hPAH, including the catalytic domain, that does not require activation by L-Phe (Knappskog et al., 1996). Wild-type hPAH shows positive cooperativity for L-Phe with a $S_{0,5}$-value for L-Phe of 154 μM and V_{max} of 1550 nmol Tyr min^{-1} mg^{-1}, while hPAH(Gly103-Gln428) shows hyperbolic saturation curves with a K_m for the substrate of 60 μM and a V_{max} of 1600 nmol Tyr min^{-1} mg^{-1}. The K_m-values for BH_4 were 25 μM for wt-hPAH and 31 μM for hPAH(Gly103-Gln428) (Knappskog et al., 1996). Moreover, the effect of BH_2 on the BH_4-dependent activity of hPAH was studied.

BH_2 has been shown to be a competitive inhibitor of rat PAH *versus* BH_4, with a K_i about 50-60 μM (Bailey & Ayling, 1983; Haavik *et al*., 1986). BH_2 also inhibits recombinant human TH competitively *vs* BH_4 (Martinez et al., 1993a) with a K_i of 70 μM. In order to calculate the K_i-values of BH_2 *vs* BH_4 for recombinant hPAH, activity assays at varying concentrations of inhibitor (BH_2) were conducted (Figure 4.1). K_i-values of 100 μM for wt-hPAH and 120 μM for hPAH(Gly103-Gln428) were measured.

The estimated $S_{0,5}/K_m$-values for L-Phe and K_i-values for BH_2 were used to calculate both the fraction of free (p_F) and bound (p_B) L-Phe and BH_2 in the transferred NOESY experiments and the normalized paramagnetic contributions to the longitudinal relaxation rates ($1/fT_{1P}$) (see below).

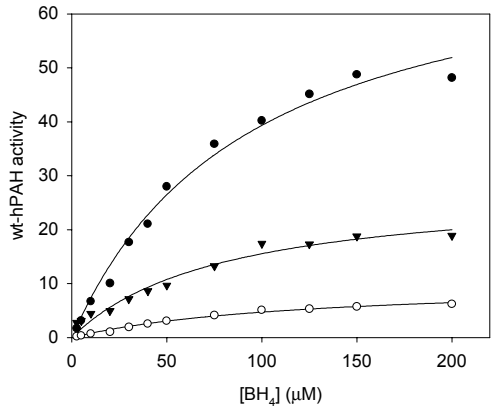


Figure 4.1
Specific activity (nmol tyrosine·min^{-1}·mg wt-hPAH) *versus* BH_4 concentration at three different BH_2 concentrations: 0 (●), 200 μM ( ) and 400 μM (○).

A similar value of p_B for both ligands indicates that the cross-relaxation rate of the H4 - H3,5 protons of L-Phe, separated by 2.46 ± 0.02 Å (Al-Karaghouli & Koetzle, 1975), can be used as a reference to calculate the intramolecular proton distances both in BH_2 and L-Phe (Table 4.2). The distances were calculated using equation 2.14 and the slope of the increase in NOE observed for the H4 - H3,5 protons as a reference.

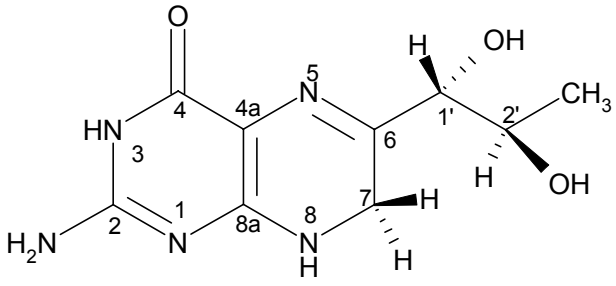
4.2 Assignment of Resonances

Spectra from the HMBC and HSQC experiments with the assignments are shown in Appendix I. The ppm values of the different carbon and proton signals are tabulated in table 4.1.

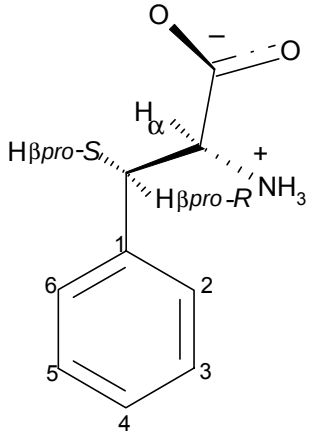
Table 4.1 Assigned resonance frequencies for L-Phe and BH_2 protons.

	Nucleus	δ (ppm)	Nucleus	δ (ppm)
L-Phe	H3,5	7.42	C=O	175
	H4	7.45	C2,4,6	129
	H2,6	7.33	C3,5	128
	Hα	3.96	C1'	78
	Hβ*pro-S*	3.27	C2'	68
	Hβ*pro-R*	3.12	Cα	57
			Cβ	37
BH_2	H7	4.17	C6	154
	H1'	4.11	C1'	77
	H2'	4.01	C2'	69
	CH_3	1.21	C7	42
			CH_3	19
	Carbon atoms 2, 4, 4a and 8a of BH_2 is not assigned. This is because these carbon atoms are not having any non-exchangeable protons connected to them by less than 3 bonds, and therefore cannot be detected in either HSQC nor HMBC.			

The structure of BH_2 and L-Phe with numbering of atoms is shown in Figure 4.2.



BH_2



L-Phe

Figure 4.2
BH_2 and L-Phe with numbering of atoms

The assigned 600-MHz ^{1}H NMR spectrum of a mixture of L-Phe and BH_2 in the presence of hPAH(Gly103-Gln428) is shown in Figure 4.3. Except for the overlapping Hα proton of L-Phe and the H2' of BH_2, the proton signals are well separated and are observed beyond the protein envelope. Moreover, for BH_2 the two H7 protons in the pyrazine-ring reveal a second order splitting pattern, showing that they are not equivalent.

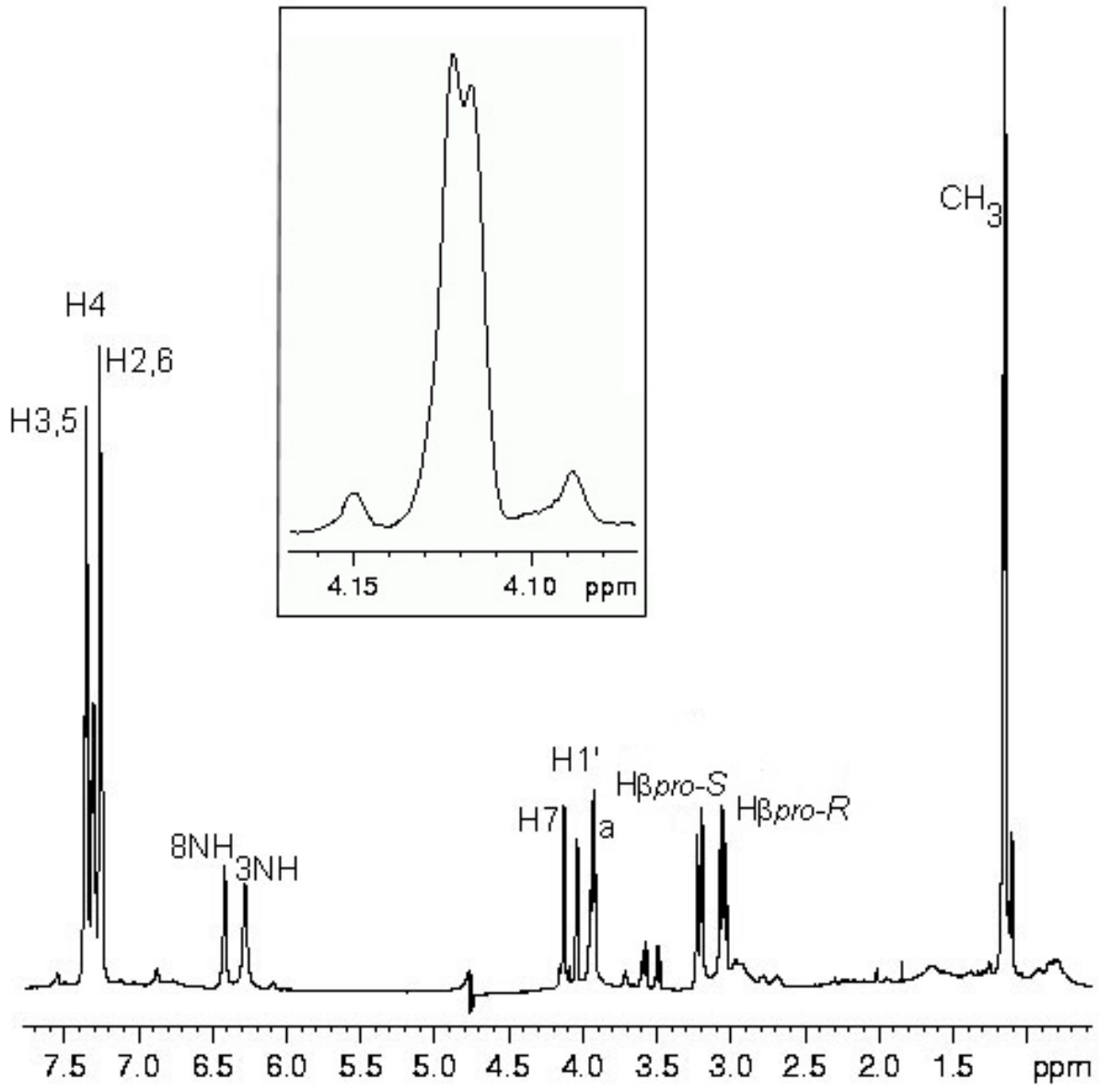


Figure 4.3
Proton NMR spectrum of L-Phe (5 mM) and BH_2 (5 mM) in the presence of 0.12 mM hPAH(Gly103-Gln428) subunit. a corresponds to the overlapping signal of Hα of L-Phe and the H2' of BH_2. Inset, the H7 protons of BH_2.

4.3 NOESY and TRNOESY

The NOESY spectrum of L-Phe and BH_2 free in solution is shown in Figure 4.4. Figure 4.4(b) is an expanded region of the NOESY spectrum, showing strong cross-relaxation between the H7 and H2' protons of BH_2.

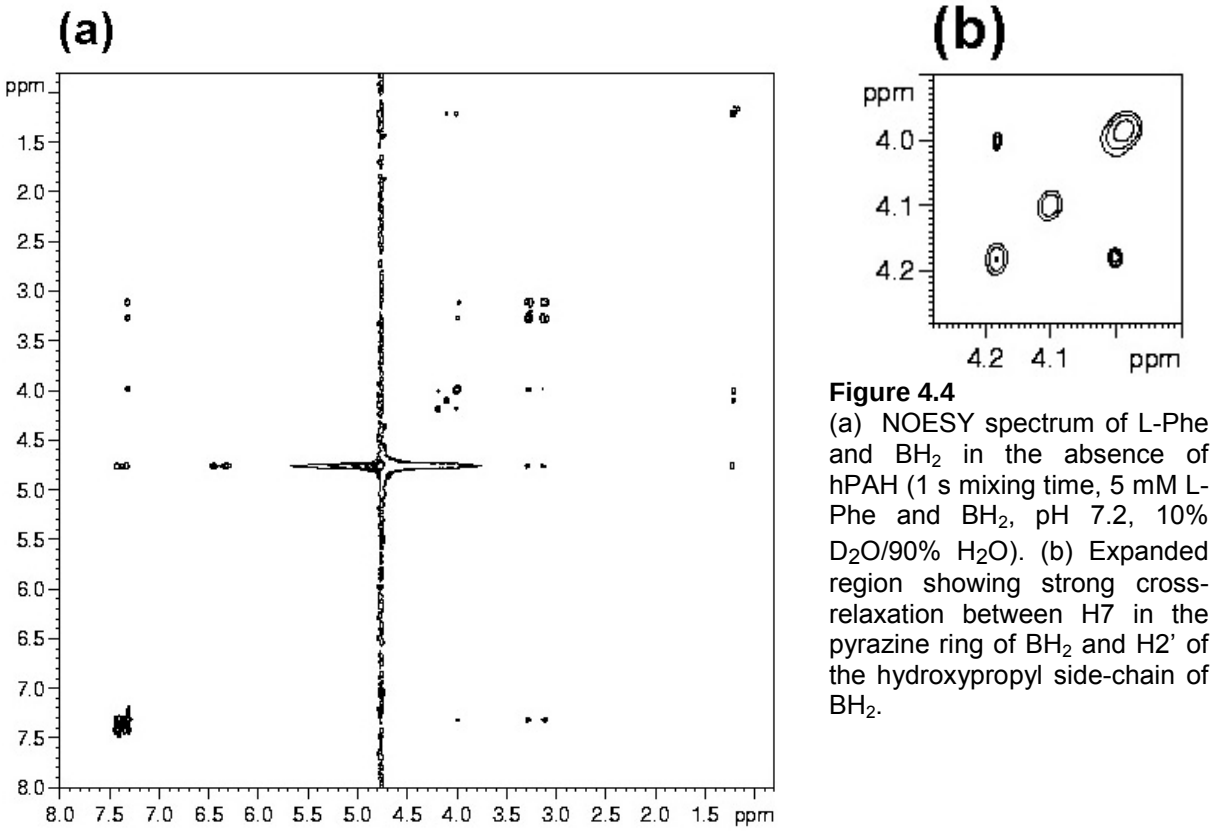


Figure 4.4
(a) NOESY spectrum of L-Phe and BH_2 in the absence of hPAH (1 s mixing time, 5 mM L-Phe and BH_2, pH 7.2, 10% D_2O/90% H_2O). (b) Expanded region showing strong cross-relaxation between H7 in the pyrazine ring of BH_2 and H2' of the hydroxypropyl side-chain of BH_2.

The intramolecular proton distances in the enzyme bound BH_2 and L-Phe were estimated from TRNOESY spectra. Similar spectra were obtained when hPAH(Gly103-Gln428) (Figure 4.5(a)) or wt-hPAH was used as the enzyme source. According to the respective $S_{0,5}$-values for L-Phe and K_i-values for BH_2 (see above), similar fractions of bound BH_2 and L-Phe ($p_B \approx 0.024$) were estimated at the concentrations of enzyme (0.12 mM subunit) and ligands (5 mM) used in the TRNOESY experiments, corresponding to saturated enzyme preparations.

This p_B-value was found to be optimal to obtain cross peak intensities significantly above the noise level, to reduce spin diffusion effects and to obtain linear build-up curves of cross peak intensity *vs* mixing time (Campbell & Sykes, 1993). Transferred NOEs from the bound to the free state were observed for both BH_2 and L-Phe as positive cross peaks in the TRNOESY spectra, reflecting negative NOEs (Figure 4.5(a)). Representative build-up curves of TRNOESY cross peak intensity for L-Phe and BH_2 are shown in Figures 4.6 and 4.7, respectively.

Transferred NOESY spectra showed cross peaks between the aromatic signals of L-Phe and aliphatic protein protons with chemical shift of 2.2, 2.1 and 1.9 ppm (Figure 4.5(c)). Moreover, the H2,6 signal of L-Phe also gave cross peaks with aromatic protein signals at 7.54 and 7.51 ppm (Figure 4.5(d)).

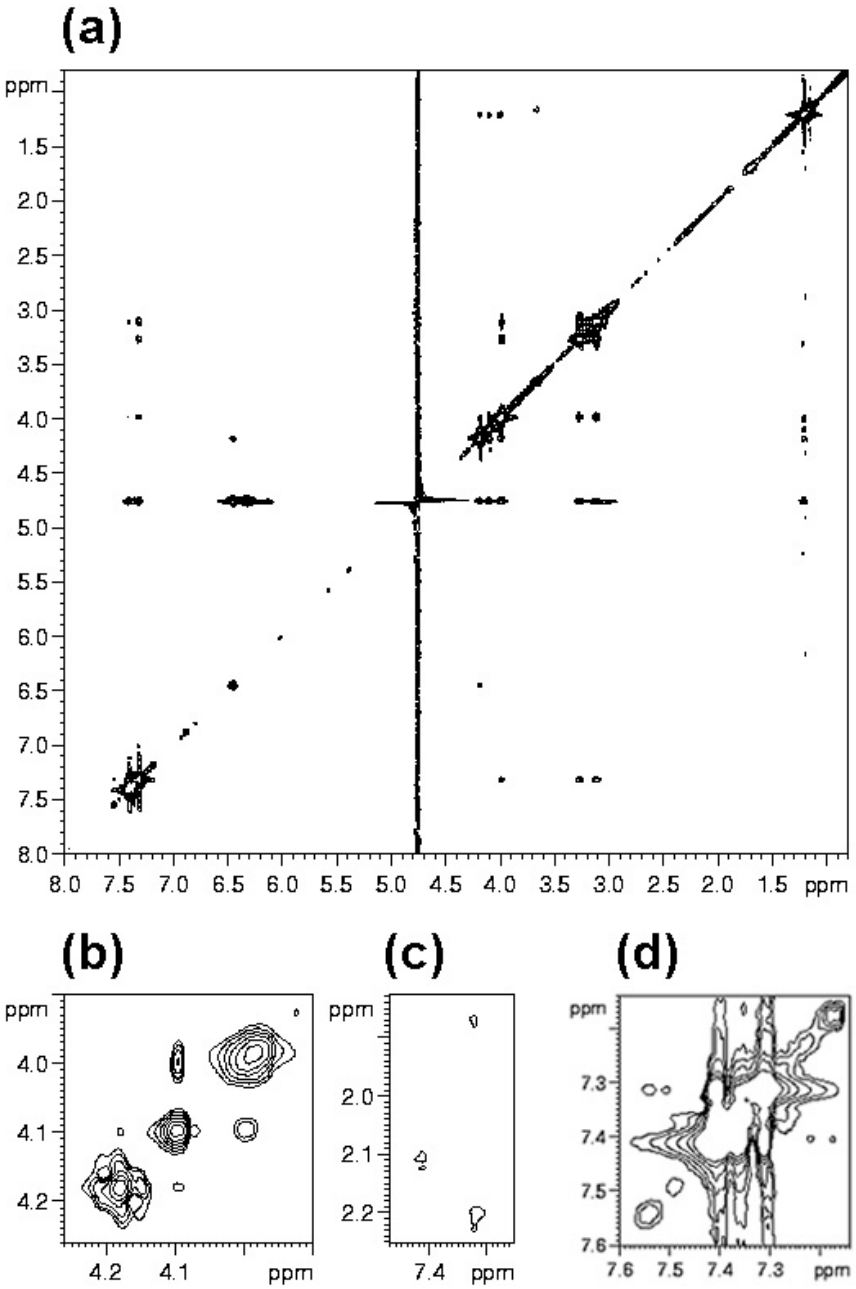


Figure 4.5
(a) Transferred NOESY spectrum (75 ms mixing time) of enzyme-bound L-Phe and BH_2 (5 mM L-Phe and BH_2 and 0.12mM subunit hPAH(Gly103-Gln428), pH7.2, 10% D_2O/90% H_2O).

(b) Cross relaxation between H1' and H2' in the hydroxypropyl side-chain for the enzyme-bound BH_2.

(c) Cross relaxation between the aromatic H3,5 and H2,6 protons of L-Phe and protein signals at 2.2, 2.1 and 1.9 ppm.

(d) Cross relaxation between the aromatic protons of L-Phe and aromatic signals in the enzyme at 7.54, 7.51 and 7.18 ppm.

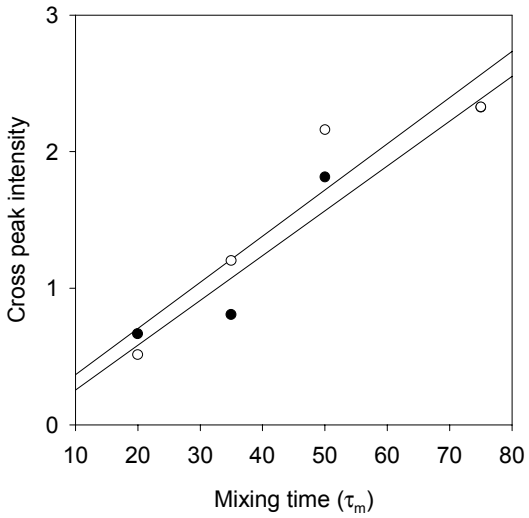


Figure 4.6
Cross peak intensity *vs* mixing time for the interaction between the aromatic Hα proton and HβS proton of L-Phe (5 mM), over (●) and under (○) the TRNOESY diagonal in the presence of 0.12 mM wt-hPAH.

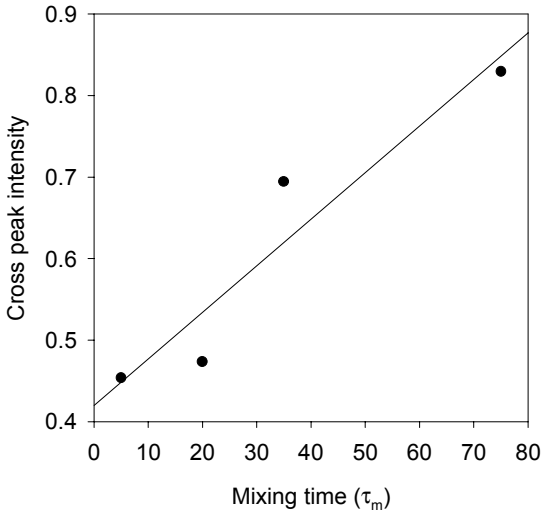


Figure 4.7
Cross peak intensity plotted *versus* mixing time for the interaction between the H7 protons and methyl protons in BH_2 (5 mM) in the presence of 0.12 mM wt-hPAH.

Interproton distances in the bound forms of the ligands were estimated from the slopes of the build-up curves (Figures 4.6 and 4.7). and are shown in Table 4.2. Similar intramolecular proton distances were obtained when spectra were taken in ternary complexes with both L-Phe and BH_2 simultaneously bound to the hPAH enzyme forms (hPAH·L-Phe·BH_2) or when they were taken in binary complexes of the enzyme with each of the ligands (data not shown).

Table 4.2 Interproton distances for L-Phe and BH_2 bound to hPAH(Gly103-Gln428) as estimated by TRNOESY.

Proton pair	(Å)[a]
L-Phe:	
H3,5 – H4	2.4 ± 0.1[b]
H3,5 - H2,6	2.3 ± 0.7
H2,6- Hα	3.9 ± 0.1
H2,6- Hβ*pro-S*	3.7 ± 0.1
H2,6- Hβ*pro-R*	3.4 ± 0.2
Hα - Hβ*pro-S*	3.5 ± 0.6
Hα - Hβ*pro-R*	3.8 ± 0.9
BH2:	
$H7^d$- CH_3 [c]	4.7 ± 0.5
$H7^d$ - H1'	3.9 ± 0.9
$H7^d$ - H2'	3.8 ± 0.6
H1' - CH_3 [c]	4.2 ± 0.2
H2' - CH_3 [c]	2.1 ± 0.1
H1' – CH_3 [c]	4.2 ± 0.2

[a] Distances estimated from TRNOESY spectra. Distances from aliphatic to degenerate pairs of aromatic protons were treated as indicated in Materials and Methods and as described (Martinez et al., 1993a).
[b] Internal standard, based on neutral diffraction study of L-Phe (Al-Karaghouli & Koetzle, 1975).
[c] Average position of these protons.

In addition to the intramolecular cross peaks within the ligands, intermolecular NOE interaction between the aromatic H2,6 and H3,5 protons of L-Phe and protein protons at 3.5 and 2.1 ppm was observed (Figure 4.8).

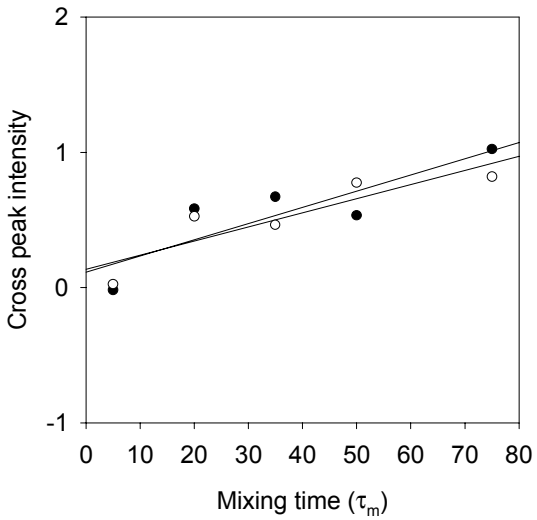


Figure 4.8
Cross peak intensity *vs* mixing time (τ_m) for the interaction between H2,6 and protein signal at 2.2 ppm(●), and H3,5 and protein signal at 2.1 ppm (○).

4.4 Relaxation Measurements

Both rat and recombinant human PAH contain one non-heme high-spin Fe(III) (*S*=5/2) at the active site (Erlandsen et al., 1997; Kappock & Caradonna, 1996; Kaufman, 1993)). Thus, the addition of either hPAH(Gly103-Gln428) (Fig 4.9) or wt-hPAH (Figure 4.10) to solutions of L-Phe and BH_2 forming a catalytically inactive complex (hPAH·L-Phe·BH_2), results in line broadening and increase in the longitudinal relaxation rates ($1/T_1$) of the proton resonances of both ligands. As seen in the figures, no significant differences were found when titrations were made with the full-length or the N-terminal-lacking truncated form.

The largest paramagnetic effect was found on the H7 protons of BH_2 and the H3,5 of L-Phe. Similar slopes for the dependency of the $1/T_{1P}$-values *vs* the concentration of enzyme were obtained from measurements performed in spectra of binary complexes (hPAH·L-Phe or hPAH·BH_2) which also provided information about the paramagnetic effect on the Hα of L-Phe and the H2' of BH_2 (Figure 4.11 and 4.12, respectively). The effect of enzyme concentration on the longitudinal relaxation rates of the two amino protons N3 and N8 were also attempted to be measured (Figure 4.13). However, due to much poorer signal intensities of these protons in the presence of protein, these results were excluded from the ananlysis.

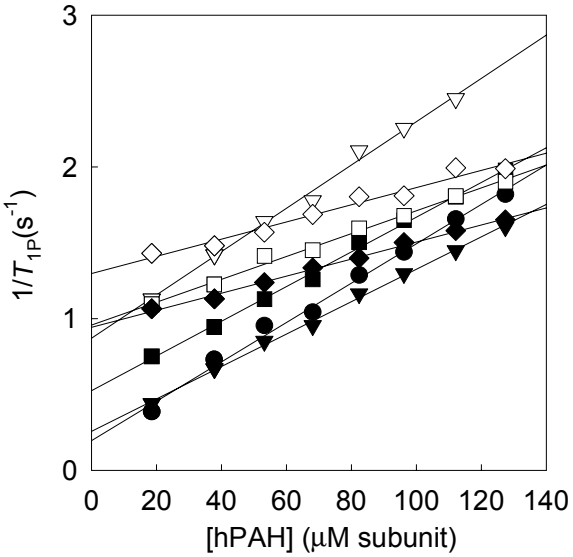


Figure 4.9
Paramagnetic relaxation rates ($1/T_{1p}$) *vs* hPAH(Gly103-Gln428) concentration for the aromatic H3,5 (●) and H2,6 (▼) and the aliphatic Hβ*pro-S* ( ) and Hβ*pro-R* (◆) of L-Phe (5 mM) and for the H7 (∇), H1' (■) and CH_3 (♬) protons of BH_2 (5 mM). Final concentration of enzyme is 0.12 mM.

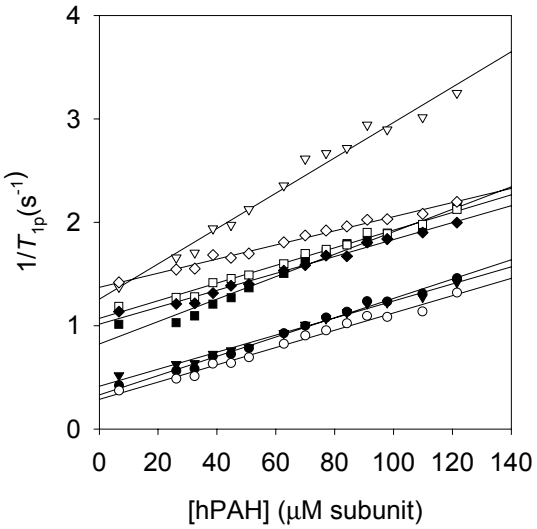


Figure 4.10
Paramagnetic relaxation rates ($1/T_{1p}$) *vs* wt-hPAH concentration for L-Phe and BH_2 (both 5 mM). The labeling is the same as in figure 4.9.

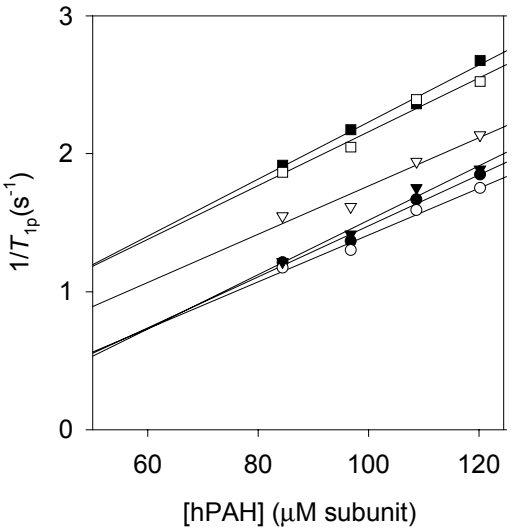


Figure 4.11
Paramagnetic relaxation rates ($1/T_{1p}$) *vs* wt-hPAH concentration for the aromatic H3,5 (●) and H2,6 (▼) and the aliphatic Hβ*pro-S* ( ), Hβ*pro-R* (◆) and Hα (∇) of L-Phe (5 mM).

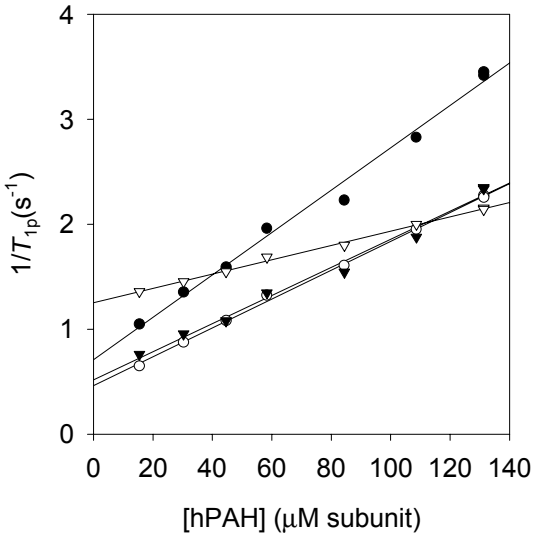


Figure 4.12
Paramagnetic relaxation rates ($1/T_{1p}$) *vs* wt-hPAH concentration for the H7 (●), H1' (▼) H2' (○) and CH_3 (σ) protons of BH_2 (5 mM).

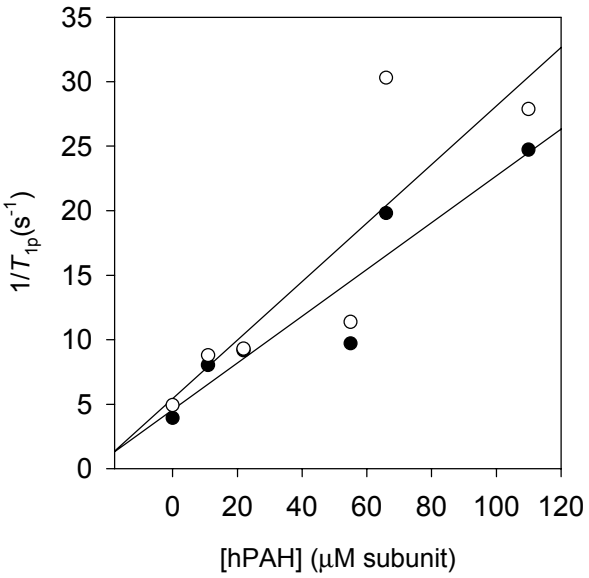


Figure 4.13
Paramagnetic relaxation rates ($1/T_{1p}$) *vs* wt-hPAH concentration for the amine protons at N8 and N3 of BH_2 (5mM).

In order to try to displace L-Phe and BH2 from their enzyme complex and to calculate the outer-sphere contribution to the relaxation rate, several competing ligands were used. Catecholamines have earlier been used to displace L-Phe bound to TH (Martinez et al., 1993a). However, addition of increasing concentrations of noradrenaline (Figure 4.14) (up to 3.5 mM) did not show any effect in the paramagnetic relaxaton rates of the substrate (Figure 4.14).

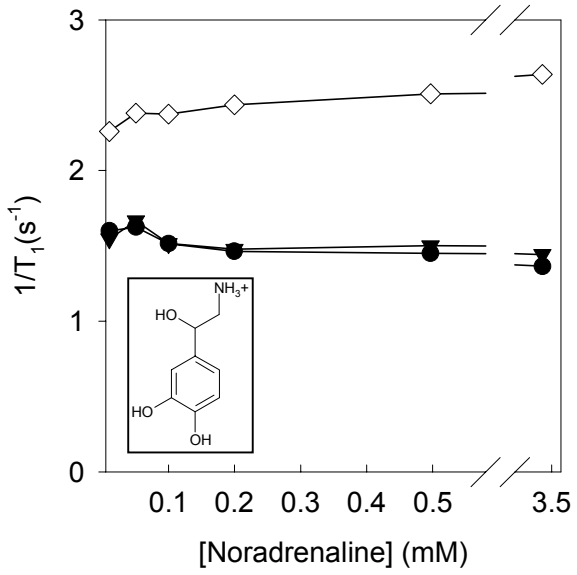


Figure 4.14
The longitudinal relaxation rate of the methyl protons of BH_2 (♬), the aromatic H3,5 (●) and H2,6 (▼) of L-Phe as a function of noradrenaline concentration. The starting solution contained L-Phe and BH_2 (both 5mM) and 0.12 mM hPAH(Gly103-Gln428). Inset, the structure of noradrenaline.

Dopamine was then added to samples of L-Phe and BH_2 in the presence of wt-hPAH in an attempt to displace the ligands. This, however, had only little effect on the paramagnetic contributions to the relaxation of either BH_2 or L-Phe protons (Figure 4.15).

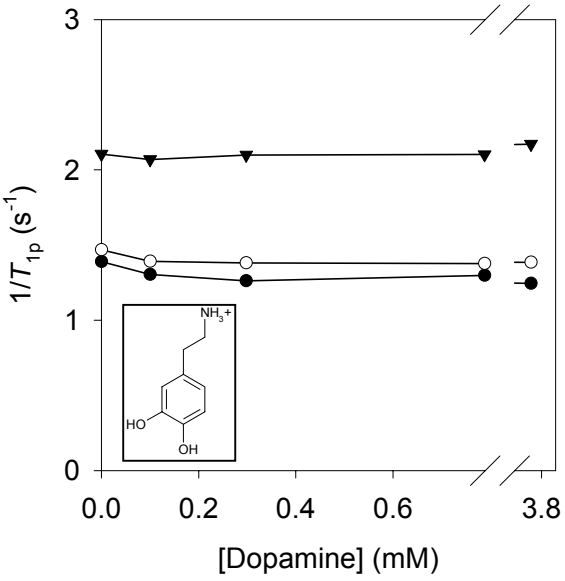


Figure 4.15
Effect of dopamine on the longitudinal relaxation rate of the methyl protons of BH_2 (▼), the aromatic H3,5 (○) and H2,6 (●) of L-Phe (both 5 mM) in the presence of wild-type hPAH (0.12 mM). Inset, the structure of dopamine.

Finally, the substrate analogue L-β-thienylalanine (Kaufman, 1993) was added in the presence of 3.5 mM dopamine. This completely eliminated the paramagnetic contributions to the relaxation of both L-Phe and BH_2 protons, and the relaxation rates returned for the values of the free ligands (Figure 4.16).

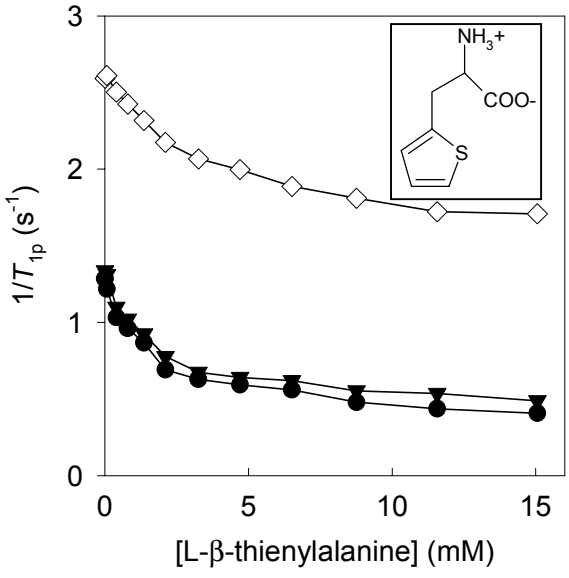


Figure 4.16
Effect of L-β-thienylalanine on the longitudinal relaxation rate of the methyl protons of BH_2 (♬), the aromatic H3,5 (●) and H2,6 (▼) of L-Phe (both 5 mM) in the presence of hPAH (Glu104-Gln428) (0.12 mM) and dopamine (3.5 mM). Inset, the structure of L-β-thienylalanine.

The T_2 values of the methyl protons of BH_2 and Hβ-pro*R* resonances of L-Phe were measured from the width at half-height of the peaks in the ^{1}H NMR spectrum, and the normalized paramagnetic transverse relaxation rate ($1/fT_{2P}$) were calculated from titration with hPAH(Gly103-Gln428) (Figure 4.17). The $1/fT_{2P}$ values were calculated to be 211 ± 17 s^{-1} for the methyl protons of BH_2 and and 669 ± 62 s^{-1} for the Hβ-pro*R* proton of L-Phe, i.e. much larger than their $1/fT_{1p}$ -value (Table 4.3).

The much larger $1/fT_{2P}$-values than $1/fT_{1P}$-values for the methyl protons of BH_2 and Hβ-pro*R* of L-Phe, indicate a fast exchange of the ligands in the complex (Martinez et al., 1993a; Mildvan et al., 1980). Thus, the lifetime of the complex (τ_M) contributes little to the T_{1P}-values and can be excluded in the calculation of distances. The distances from the metal ion to the ligand protons were estimated using the value of the effective dipolar correlation time for the paramagnetic relaxation in human wt-hPAH complexed with L-Phe (τ_c =(1.2 ± 0.2) x 10^{-10} s) (Olafsdottir & Martinez, 1999), which seems to correspond to the longitudinal electron spin relaxation time of Fe(III) in the complex. The distances obtained by using this procedure can be considered only as reasonable approximations (shown in table 4.3 for the measurements with the truncated form of hPAH) due to possible violations of the various assumptions inherent in the modified Solomon-Bloembergen (MSB) approach to nuclear relaxation in paramagnetic systems (see Chapter 2, Theoretical Considerations).

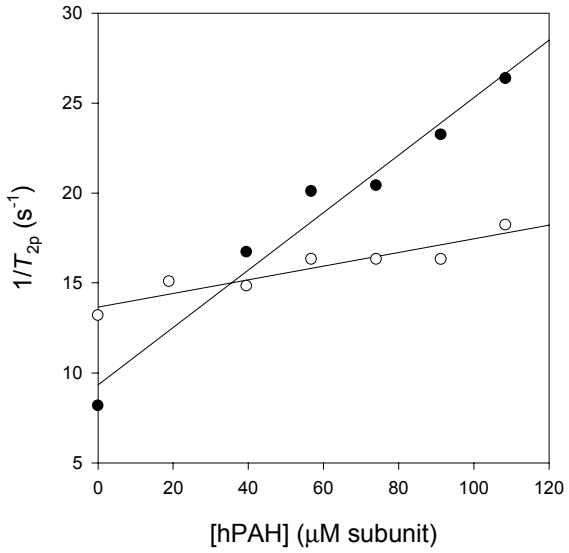


Figure 4.17
The transverse relaxation rate of the methyl protons of BH_2 (●) and Hβ-pro*R* proton of L-Phe (○) as a function of hPAH(Glu103-Gln428) subunit concentration. When corrected for substrate concentration, the slope give the normalized longitudinal relaxation rate ($1/fT_{2P}$).

At the concentrations used in the titrations, both forms of hPAH were calculated to be >98 % saturated with BH_2 and L-Phe for determining the normalized paramagnetic relaxation rates ($1/fT_{1P}$ and $1/fT_{2P}$) (3.2 in Materials and Methods) and the $1/fT_{1P}$-values are summarized in Table 4.3.

Table 4.3 Paramagnetic effects and distances (*r*) from the observable protons of L-Phe and BH_2 simultaneously bound to hPAH(Gly103-Gln428) to the Fe(III) at the active site.

Nucleus	δ (ppm)	$1/fT_{1P}$ (s^{-1})	*r* (Å)[a]
L-Phe:			
H3,5	7.42	127.3	7.4±0.6
H2,6	7.33	105.0	7.7±0.6
Hβ*pro-S*	3.27	73.8	8.1±0.7
Hβ*pro-R*	3.12	55.2	8.5±0.6
BH_2:			
H7[b]	4.20	140.1	7.3±0.5
H1'	4.11	112.2	7.6±0.6
CH_3[b]	1.21	55.8	8.5±0.7

The data are the average of 4 independent titrations.
[a] Estimated using a $\tau_C = (1.2 \pm 0.2) \times 10^{-10}$ s (Olafsdottir & Martinez, 1999). Errors in *r* include contributions from errors in $1/fT_{1P}$ (about 7 %) and in τ_C.
[b] Average position of these protons.

4.5 Distance Geometry

Distance geometry calculations (DGII) for L-Phe and BH_2 were performed using the interproton distance restraints (Table 4.2), the estimated metal-proton distances (Table 4.3) and three chirality constraints, *R* and *S*, for the C1' and C2' chiral centers for BH_2 (L-*erythro*-1',2'-dihydroxypropyl), respectively, and one for the Cα of L-Phe. Thirty sets of coordinates were computed in the embedding step of the calculation and refined as described (Havel, 1991). The resulting conformers after DGII for BH_2 and L-Phe are shown in Figures. 4.18 and 4.19, respectively.

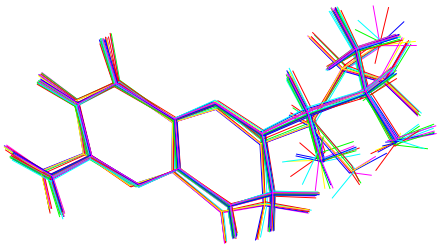


Figure 4.18
Conformers of hPAH-bound BH_2 generated from DGII

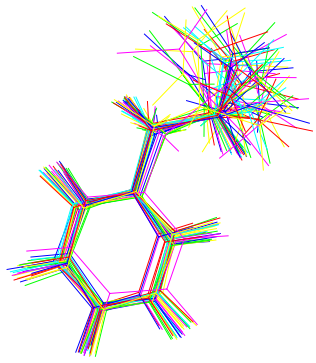


Figure 4.19
Conformers of hPAH-bound L-Phe genereated from DGII

One family of conformers was obtained for bound L-Phe, with pairwise root-mean-squared deviation (r.m.s.d.) values of 0.49 Å for all atoms in the ligands. The refined sets of structures were analyzed for the presence of chiral and NMR distance restraints violations as well as covalent geometry violations. The conformers of bound L-Phe were all accepted and 5 representative structures are shown in Figure 4.20. For bound BH_2 two families were first obtained by DGII, i.e. one in which the torsion angle N5-C6-C1'-C2' was negative and one in which this angle was positive (Figure 4.18). Preliminary molecular docking using one representative structure (the structure with lowest r.m.s.d. relative to the ensemble) from both families indicated that the conformers with negative torsion angle gave only very low-score complexes with hPAH(Gly103-Gln428). A restraint was then included in the distance geometry calculations in order to keep the angle N5-C6-C1'-C2' positive. The nine resulting conformers, after elimination of those with NMR distance and covalent geometry violations, are shown in Figure 4.21. (the r.m.s.d. for all the atoms is 0.22 Å). The intramolecular distances in the "bound" conformers of the ligands were measured after DGII calculations (in the structure with lowest r.m.s.d. relative to the ensemble) and shown in Table 4.4 for comparison with the estimated distances from NMR.

Figure 4.20
Five structures of L-Phe with lowest r.m.s.d. relative to the ensemble. Iron is shown in yellow.

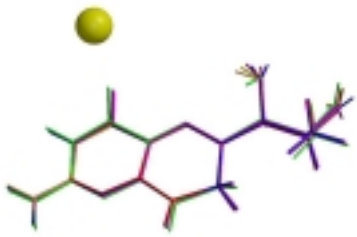


Figure 4.21
Nine conformers of BH_2 generated in DGII after including constraint in the side-chain.

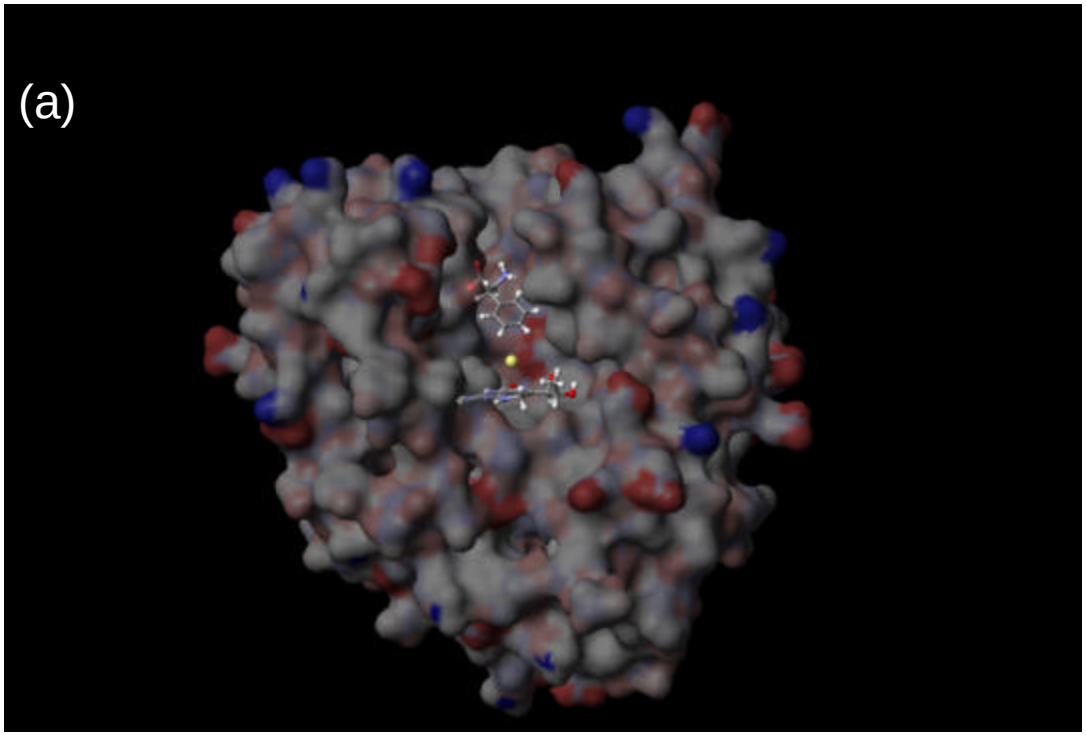
Table 4.4 Interproton distances in L-Phe and BH_2 bound to hPAH(Gly103-Gln428) as estimated by TRNOESY and after distance geometry calculations.

Proton pair	NMR distances (Å)[a]	DGII distances (Å)[b]
L-Phe:		
H3,5 – H4	2.4 ± 0.1	2.4
H3,5 - H2,6	2.3 ± 0.7	2.4
H2,6- Hα	3.9 ± 0.1	4.2
H2,6- Hβ*pro-S*	3.7 ± 0.1	3.7
H2,6- Hβ*pro-R*	3.4 ± 0.2	3.2
Hα - Hβ*pro-S*	3.5 ± 0.6	2.6
Hα - Hβ*pro-R*	3.8 ± 0.9	2.7
BH2:		
H7[c] - CH_3	4.7 ± 0.5	3.9
H7[c] - H1'	3.9 ± 0.9	3.7
H7[c] - H2'	3.8 ± 0.6	3.4
H1' - CH_3 [c]	4.2 ± 0.2	4.1
H2' - CH_3 [c]	2.1 ± 0.1	2.7

[a] Distances estimated from TRNOESY spectra.
[b] Distances measured after distance geometry calculation (in the structure with lowest r.m.s.d. relative to the emsemble).
[c] Average position of these protons.

4.6 *Docking*

As seen from Figure 4.21, one family of BH_2 was generated from DGII. The represantative with lowest r.m.s.d. relative to the ensemble of the nine BH_2 confomers was docked into the crystal structure of hPAH(Gly103-Gln428) using DOCK 4.0 (Ewing & Kuntz, 1996). In agreement with the intermolecular distances estimated by ^{1}H NMR, a GRID-region of 20 Å around the active site iron was defined for the docking procedure. The L-Phe structure with lowest r.m.s.d. relative to the ensemble was used in the anchor-grow docking procedure of the amino acid substrate. Due to the fact that the bond angles are optimized during this procedure, some of the conformers generated violated the distance restraints obtained from NMR. However, the structure with lowest r.m.s.d. relative to the ensemble of docked conformers (from DOCK) practically overlap with the five structures of L-Phe with lowest r.m.s.d. relative to the ensemble of bound conformers (from DGII). The top scoring conformer obtained from DOCK have a r.m.s.d. of 0.22 Å relative to the best scoring conformer with lowest r.m.s.d. to the ensemble of conformers from DGII. For BH_2, the 30 top scoring complexes with respect to molecular mechanics interaction energy (AMBER force field score) practically overlap (r.m.s.d. for all the atoms was 0.53 Å for the hPAH·L-Phe complex and 0.25 Å for the hPAH·BH_2 complex). Representative structures for the hPAH·L-Phe·BH_2 complex are shown in Figure 4.22. The iron-proton distances in the "bound" conformers were measured after docking (in the structure with lowest r.m.s.d. relative to the ensemble) and shown in Table 4.5 for comparison with the estimated distances from the paramagnetic relaxation measurements.



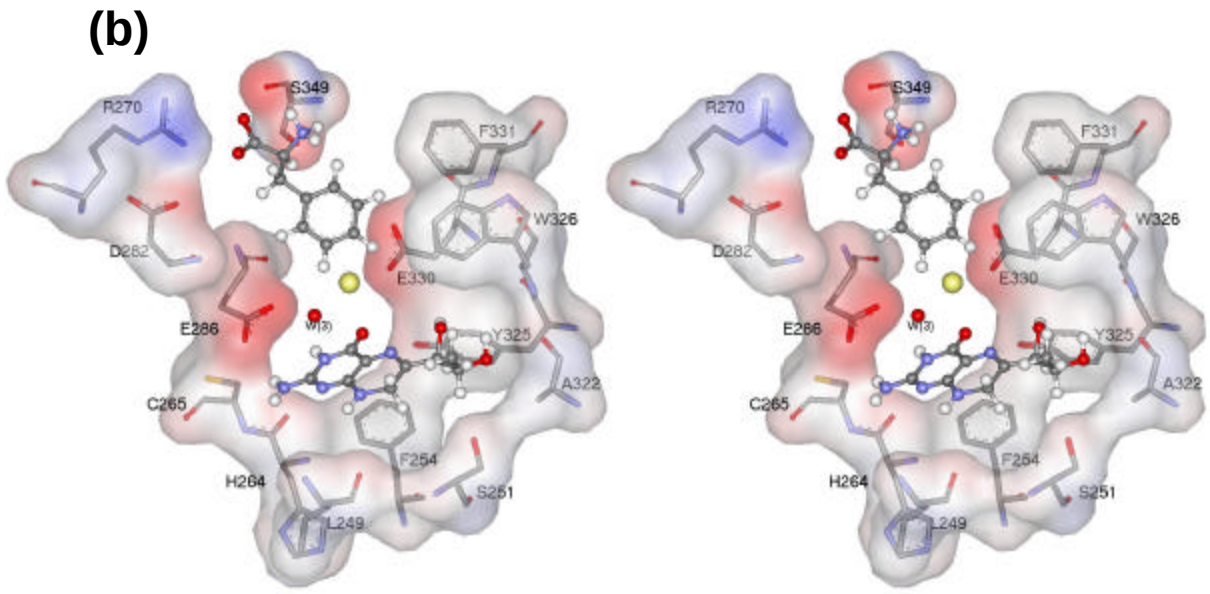


Figure 4.22
Representative distance refined structure of bound conformers of L-Phe and BH_2 (the L-Phe structure with lowest r.m.s.d. relative to the ensemble in Figure 4.20 and the BH_2 conformer with lowest r.m.s.d. relative to the ensemble in Figure 4.21) docked into the crystal structure of PAH(Gly103-Gln428) (PDB accession number 1PAH). (a) Overview of the monomer of the enzyme with the bound ligands. (b) A closer stereo view showing the active site residues involved in the binding of substrate and cofactor analogue. The iron is shown in yellow.

Table 4.5 Distances (*r*) from the observable protons of L-Phe and BH_2 simultaneously bound to hPAH(Gly103-Gln428) to the Fe(III) at the active site

Nucleus	δ (ppm)	*r* (Å)[a]	*r* (Å)[b]
L-Phe:			
H3,5	7.42	7.4±0.6	5.0
H2,6	7.33	7.7±0.6	5.8
Hβ*pro-S*	3.27	8.1±0.7	6.5
Hβ*pro-R*	3.12	8.5±0.6	7.5
BH_2:			
H7[c]	4.20	7.3±0.5	7.3
H1'	4.11	7.6±0.6	6.6
CH_3[c]	1.21	8.5±0.7	8.6

The data are the average of 4 independent titrations.

[a] Estimated from the paramagnetic relaxation measurements.

[b] Distances measured after distance geometry and docking calculation (in the structure with lowest r.m.s.d. relative to the emsemble).

[c] Average position of these protons.

5 Discussion

In this study, NMR is used as a tool to extract information about the conformation of the ligands (i.e. L-Phe and BH_4) while bound to the receptor (i.e. hPAH). Transferred NOESY experiments give intramolecular proton distances of the ligands, and some indications about the residues located at the binding site, while relaxation measurements give intermolecular distances from the ligand protons to the active site iron. These distance restraints, together with intrinsic chirality constraints of the ligands are used to generate possible conformers of the bound ligands. When a plausible conformer is obtained, i.e. one that is compatible with all the given NMR restraints, this conformer is docked into the active site of the enzyme. The aim of the docking procedure is to find the most energetically optimal orientation of the ligands within the structure of the enzyme. After generating several ligand-receptor complexes, these complexes are checked for compatibility with intermolecular NMR restraints. This procedure is depicted in Figure 5.1.

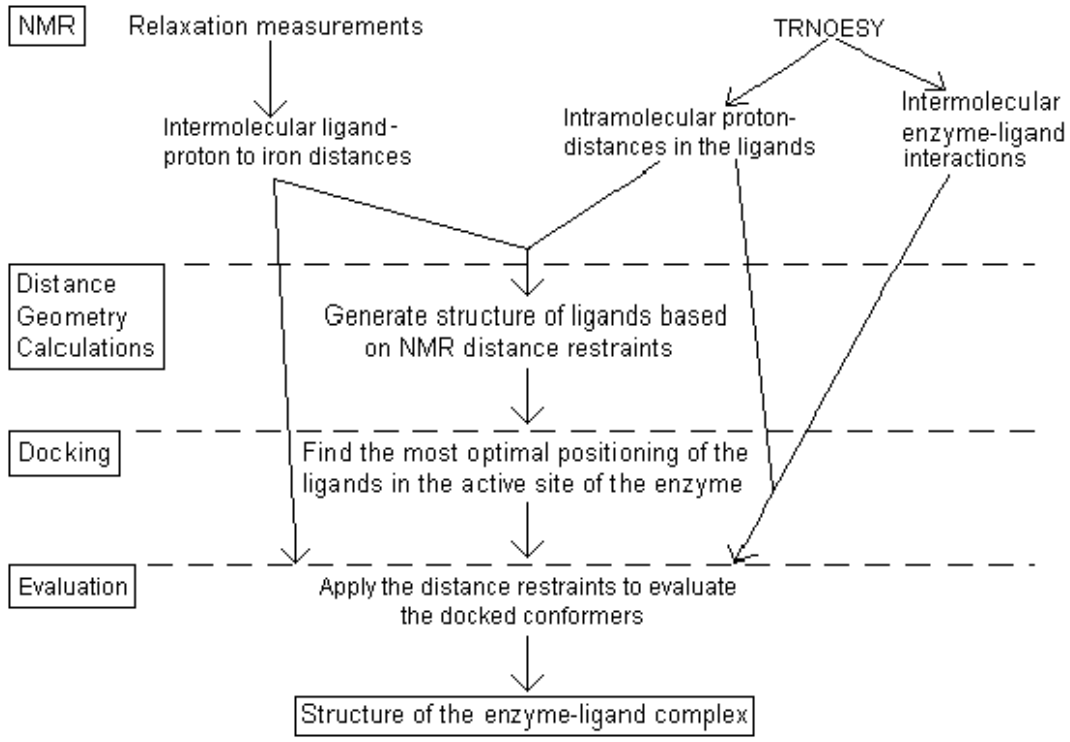


Figure 5.1
Schematic view of the procedure to obtain the structure of the enzyme-ligand complex.

5.1 Results from NMR

The fact that the H7 protons in the pyrazine ring show second order splitting, might indicate that the pyrazine ring is not completely planar or that the hydroxypropyl substituent could affect the chemical shifts of the two H7 protons differently. In the crystal structure of 6-methyl-7,8-dihydropterin both heterocyclic rings are flat (Bieri, 1977). However, after optimization of the BH_2 structure using DGII (EMBED algorithm), the pyrazine ring obtained is not completely planar, i.e. C7 is 0.56 Å above the plane formed by the other 5 atoms in the ring. But is should be noted that these calculations are beyond the scope of the DGII algorithm.

Transferred NOESY spectra of the ternary complexes show strong cross-relaxation between the H1' and H2' in the hydroxypropyl side chain at C6 in BH_2 (Figure 4.4(a)), indicating that the hydroxyls adopt a *cis* conformation when bound to either hPAH(Gly103-Gln428) or wt-hPAH. However, the hydroxyls have been found to be *trans* in BH_4 both by X-ray analysis (Matsuura *et al*., 1985) and theoretical calculations by the Molecular-Orbital Method (Katoh *et al*., 1993). Accordingly, the NOESY spectrum of free BH_2 both in the presence and the absence of L-Phe shows a week cross peak between H1' and H2' and a strong cross peak between the H7 protons in the pyrazine ring and the H2', but no interaction between the H7 and the H1' (Figure 4.3(a)), indicating that the hydroxyls adopt a *trans* configuration in solution as well. Thus, it seems that the binding to the enzyme imposes constrictions to the conformation of the dihydroxypropyl side chain (see below) which may be related to the specific kinetic and regulatory properties manifested by the natural tetrahydrobiopterin cofactor (Kappock & Caradonna, 1996; Kaufman, 1993).

The NOESY cross peaks between the aromatic signals of L-Phe and protein signals at 7.54 and 7.51 ppm (Figure 4.4 (d)) very likely belong to a tryptophan spin system. Since there are only three tryptophans in hPAH, this gives a strong restriction on the positioning of the L-Phe ring. Furthermore, since none of the ring protons of the amino acid substrate are found to be further away from the active site iron than 10 Å (Table 4.5), this indicates that the actual tryptophan residue that gives us the interaction with the amino acid substrate has to be within a distance from the iron of 15 Å (the distance from the iron to the protons of the L-Phe ring plus the maximum distance for which NOEs can be observed (5Å)). This narrows the actual protein residue that gives an interaction with L-Phe down to one, namely Trp326.

Furthermore, weak interproton cross peaks were also found between the H4 and H2,6 ring protons of L-Phe and protein signals at 1.9, 2.1 and 2.2 ppm (Figure 4.4(c)). These could not be unambiguously assigned, since many protein protons have signals at these ppm values. However, since the ring of L-Phe has to be within 5Å from Trp326, and within 10 Å from the active site iron, this restrict the possibilities down to a few residues within the active site. The most probable candidate is the β- and γ-protons of Glu330. But since the signals are so weak, and could not be unambiguosly assigned, they were not given any more attention at this point. However, the bound substrate conformation as obtained by docking, supports the suggestion that these cross peaks arise from interaction with Glu330 (see below).

Cross peaks were also found between protons from BH_2 and from the enzyme, although they could not be unambiguosly assigned.

Catecholamines have been shown to competitively inhibit rat PAH with respect to the pterin cofactor (Bublitz, 1971) and completely displace BH_2 from its binding site near the iron in human TH (Martinez et al., 1998). However, this could not be showed for hPAH. The changes observed are compatible with a rearrangement of the ligands, but displacement was not achieved (Figures 4.14 and 4.15). Thus, competition of the catecholamine for the binding site of cofactor, as recently proposed (Erlandsen et al., 1998), was not observed. However, the addition of the substrate analog L-β-thienylalanine (Kaufman, 1993), in the presence of 3.5 mM dopamine, eliminated the paramagnetic contributions to the relaxation of both L-Phe and BH_2 protons, and the relaxation rates returned for the values of the free ligands (Figure 4.16). This is in agreement with a total displacement of both ligands at this condition, confirming that the outer-sphere contribution to the paramagnetic relaxation rates of both ligands in the complex (PAH·L-Phe·BH_2) was negligible since the enzyme-bound Fe(III) has no effect on free (displaced) L-Phe and BH_2.

5.2 Results from Molecular Modeling

The intramolecular distances in the «bound» conformers of the ligands were measured after the distance geometry calculations (in the structure with lowest r.m.s.d. relative to the emsemble) and shown in Table 4.4 for comparison with the estimated distances from NMR. Except for the interproton distances involving the aliphatic chain of the substrate, there is a good agreement between both data sets for the intramolecular distances (Table 4.4). There is also a good fit between the absolute Fe(III)-proton distances for bound BH_2 estimated from the paramagnetic probe-T_1 method and those obtained after modeling, and between the relative distances for the L-Phe protons (Table 4.5). Significant shortening of the absolute metal-proton distances after modeling was however encountered for the conformation of bound L-Phe, and especially for the aromatic ring (Table 4.3). Although the latter discrepancies could well be due to limitations of the MSB relaxation theory rather than an erroneous model of the structure, the possibility should be considered that the modeling procedure results in conformers with shortened Fe(III)-proton distances in the ring. Thus, it should be mentioned that some of the conclusions inferred from the conformation of the modeled bound L-Phe, and especially those realted to the structure of the putative quaternary reaction complex of the enzyme with L-Phe, O_2 and BH_4 and the proposed catalytic mechanism (see below) are not so adequate for L-Phe bound at the average distances directly obtained from the paramagnetic relaxation measurements (Table 4.5).

In all conformers of BH_2 it was manifested an almost-*cis* conformation of the two hydroxyl groups at C1' and C2' in the side chain at C6, with torsion angles OH-C1'-C2'-OH = -60 $\pm$ 0.2° and C6-C1'-C2'-C3' = -66 $\pm$ 0.2°. As studied by NMR, the two hydroxyl groups in TH-bound BH_2 were also found to be almost *cis* (torsion angle OH-C1'-C2'-OH = -75° to -83°) (Martinez et al., 1998). However, while the H1'-H7 distance was found to be 2.7 $\pm$ 0.2 Å in TH-bound BH_2 (Martinez et al., 1998), it was estimated to be 3.9 $\pm$ 0.9 Å in hPAH-bound BH_2 (Table 4.2), indicating a different conformation of the dihydroxypropyl chain in both enzymes. L-Phe has been found not to coordinate to the metal ion in TH (Martinez et al., 1993a) and, as shown in this work, this substrate neither coordinates to the catalytic iron in hPAH.

The distances from the iron to the N5 in the pyrazine ring and to the 4-oxo in the pyrimidine ring are 4.4 $\pm$ 0.4 Å and 2.6 $\pm$ 0.3 Å, respectively, which is in agreement with a direct coordination to the metal through the 4-oxo group. Pterins are known to coordinate to metals through the N5 and 4-oxo groups in small pteridine-containing metal complexes (Odani *et al*., 1992). Moreover, a coordination of BH_2 to the active-site iron is in agreement with the change in the coordination geometry of the metal ion observed by EPR spectroscopy and has important implications to understand the catalytic mechanism of the enzyme (see below).

5.3 The Substrate Binding Site in PAH

L-Phe binds to human PAH at residues placed at both sides of the wide channel guiding to the active site iron (Figure 4.22(a)). The residues involved in L-Phe binding are shown in Figure 4.22(b). The residue Arg270 from one side of the channel is involved in an ionic interaction with the carboxyl group of the amino acid substrate, while Ser349, from the other side, hydrogen-bonds with the amino group of L-Phe. The phenyl ring of L-Phe binds very close to Phe331, with distances between the H6 proton of the substrate and the H4 and H3 protons of Phe331 of 3.6 and 5.0 Å, respectively. The rings tilt approximately 16° relative to one another. The phenyl ring also binds very close to Trp326 (the distance between C4 of L-Phe and C6($C^{\eta 2}$) of Trp326 is 4.7 Å and these rings tilt approximately 30° relative to one another). Thus, protons H4($H^{\varepsilon 3}$) and H7($H^{\zeta 2}$) of Trp326 seems to be the aromatic protons involved in the intermolecular NOE interaction with the H2,6 protons of L-Phe, while the aliphatic signals with ppm values 1.9, 2.1 and 2.2, that also cross-relax with the L-Phe ring protons (Figure 4.4(c)), seem to correspond to the β- and γ-protons of the neighboring Glu330. It is very likely that on binding of the substrate to the enzyme a better stacking of the side-chain and substrate rings is induced through displacement of Trp326. In addition to Phe331 and Trp226, the ring of Tyr377 is also in the proximity of the binding site of the phenyl ring of L-Phe, creating a cluster of aromatic amino acids. All interacting residues with the substrate are conserved in the aromatic amino acid hydroxylases, except for Trp326, which, interestingly, is a phenylalanine in tryptophan hydroxylase. In agreement with a role of this residue in the binding of the substrate, the W326F mutant form of recombinant hPAH showed decreased affinity for L-Phe, while the catalytic activity and the affinity for BH_4 was not affected (Knappskog & Haavik, 1995). Mutations at Arg270, Ser349, Phe331 and Tyr377 in human PAH are associated with PKU (Bjørgo *et al*., 1998; Erlandsen et al., 1997). Expression analyses of the PKU-mutant forms S349P (Knappskog *et al*., 1995) and R270S (Bjørgo et al., 1998) have shown that the mutations cause structural perturbations resulting in protein instability. In addition to their role in the binding of substrate these residues participate in a network of hydrogen bonds and ionic interaction with other residues important for the structural integrity of the enzyme (Bjørgo et al., 1998; Erlandsen et al., 1997). Results from recent mutagenesis studies support that Arg316 in TH (corresponding to Arg270 in hPAH) is involved in the binding of the substrate (Daubner & Fitzpatrick, 1999). Thus, a conservative substitution at Arg316 in TH, i.e. R316K, results in stable enzyme with 3000-fold increase in K_m. Based on the results from mutagenesis, these authors also indicated a role for Asp328 (Asp282 in hPAH) in the binding of the substrate (Daubner & Fitzpatrick, 1999). However, this residue does not seem to establish contacts with the substrate but it is involved in an ionic interaction with the conserved arginine (270/316) and helps to position this residue in the substrate binding site (Figure 4.22(b)). Mutations at Asp282 in recombinant rat PAH affect protein stability and activity without altering pterin binding ability (Dickson *et al*., 1994). hPAH seems to share the motifs for L-Phe binding with other amino acid binding proteins. Thus, both in the bacterial periplasmic histidine-binding protein (Oh *et al*., 1994) and the hystidyl-tRNA synthetase (Aberg *et al*., 1997) Arg, Ser and Asp or Glu residues, positioned in a similar way as in hPAH, are involved in the binding of L-His through ionic interactions. Moreover, in these histidine-binding proteins, the imidazole ring of the substrate stacks with aromatic residues.

5.4 Implications for Substrate Specificity

hPAH shows a high specificity for L-Phe, although it can hydroxylate other amino acids including L-Trp, but not L-Tyr (Kappock & Caradonna, 1996; Kaufman, 1993). The substrate specificity of the enzyme, as well as of the rest of the aromatic amino acid hydroxylases, has been found to be determined by the catalytic domain, and not by the regulatory domain (Daubner *et al*., 1997). The stereospecificity of hPAH for the L-isomers is explained by the arrangement of the binding site with specific interactions for the carboxyl and the amino groups of the substrate (Figure 4.22(b)). Interestingly, the determinants for the specificity for a *para*-unsubstituted phenyl ring in the side chain of the substrate seem to be at the iron site of the enzyme, since these substituents would come in both steric and ionic conflict with the carboxylic $O^{\varepsilon 1}$ of Glu330 (Figure 5.2(a)).

Accordingly, both L-Tyr and *p*-Cl-L-Phe bind to hPAH with low affinity (K_d about 2 mM) (Kappock & Caradonna, 1996). *para*-substituted amino acids function as inhibitors of the enzyme and when used as substrates they cause an uncoupled reaction, i.e. hydroxylation of the tetrahydropterin cofactor without hydroxylation of the substrate. An examination of the coordination geometry around the iron site in both hPAH and rat TH shows that while the six coordinating ligands (His285, His290, Glu330 and three water molecules) arrange around the iron in hPAH with an octahedral geometry (Erlandsen et al., 1997), the five coordinating ligands in TH (His331, His336, Glu376 and two water molecules) coordinate in a square pyramidal geometry (Goodwill et al., 1997). Glu330 coordinates to the iron in hPAH in a «monodentate» manner through only one of the carboxylic oxygen atoms (O1…Fe, 3.6 Å; O2…Fe, 2.1 Å) (Figure 5.2(a)), while Glu376 seems to coordinate to the iron in TH almost in a «bidentate» mode (O2…Fe, 2.0 Å; O1…Fe, 2.6 Å) (Figure 5.2(b)).

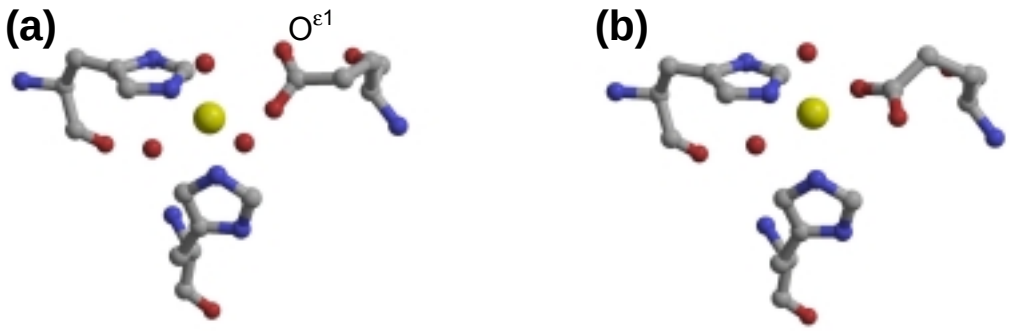


Figure 5.2
A view of the iron-coordinating residues in human PAH and rat TH. The iron atom is shown in yellow and oxygen atoms from water in red. Note the different coordination geometry of Glu330 (in hPAH) and Glu376 (in TH).

Thus, neither $O^{\varepsilon 1}$ nor $O^{\varepsilon 2}$ in Glu376 protrude towards the substrate binding site in TH, if a similar binding mode for the substrate in both TH and hPAH is assumed, allowing the binding and hydroxylation of L-Tyr. The differences in coordination geometry around the iron in both hydroxylases do not seem to be defined by individual changes of residues around the active site, but by the entire structure of the catalytic domain. Two loops outlining the shape of the active site (residues 244-250 and 375-382 in hPAH) come closer to the iron in rat TH than in hPAH, probably affecting the second coordination sphere of the metal and, in turn, the first

coodination (Erlandsen et al., 1997; Goodwill et al., 1997). Moreover, the sulfhydryl group in Cys265 (311 in rat TH) at the second coordination sphere points towards the iron in rat TH but not in hPAH, which may influence the coordination geometry of the metal ion. No structural information is presently available about tryptophan hydroxylase but since this enzyme resembles hPAH with respect to substrate specificity (Kappock & Caradonna, 1996), it is reasonable to expect that both hydroxylases have a similar arrangement of the conserved coordinating residues around the iron.

5.5 The Pterin Binding Site in PAH

As seen from Figure 4.22(b), the side-chain carboxylate of Glu286 hydrogen-bonds the N3 and the amine at C2 of the pterin ring of BH_2. The amine at C2 also forms an additional hydrogen bond with the carbonyl group of His264, while N8 hydrogen bonds with the carbonyl group of Leu249. Moreover, the pterin π-stacks with Phe254, with a medium interatomic distances between the pterin and the phenyl ring of 3.6 Å. Both rings are tilted by approximately 30° relative to one another, although the distance and angle between the rings may be altered in the «bound» structure of the enzyme. Ring stacking of the pterin with aromatic amino acids with similar offset from being parallel is also found in the structures of dihydrofolate reductase complexed with biopterin (McTigue *et al*., 1992) and nitric oxide synthase with tetrahydrobiopterin (Crane *et al*., 1998). The distance between the C4a and the iron is about 4.3 Å and the most significant interaction of the pterin ring with the enzyme is via the 4-oxo atom which is placed at coordinating distances from the Fe(III), at about the same position as the oxygen atom of the coordinating H_2O (1) in the unbound form of the enzyme (Figure 4.22(b)). This water molecule is thus expected to be removed from the iron coordination sphere when the pterin binds. The binding site of the pterin ring is in agreement with the effect of substitutions at positions 2, 5, 6, 7, 8 (Kappock & Caradonna, 1996) and 3 (Martinez et al., 1998) (Almås *et al*., in press) of tetrahydropterin on the binding affinity, the ability of the substituted tetrahydropterins to act as cofactors and the degree of uncoupling in the hydroxylation.

Additional interactions between BH_2 and hPAH(Gly103-Gln428) involve the two hydroxyls of the side chain at C6, with C2' OH hydrogen-bonding with the carbonyl group of Ala322, and both C1' OH and C2’ OH at putative hydrogen-bonding distances of the carbonyl group of Ala322 and the hydroxyl group of Ser251, respectively. Moreover, the residue Tyr325, which is highly conserved in the aromatic amino acid hydroxylases and is hydrogen bonded to H_2O (1) in the unbound form of the enzyme (Erlandsen et al., 1997), also seems to establish van der Waals interactions with several atoms of BH_2 contributing to the observed arrangement of the dihydroxypropyl chain and to the right binding of the pterin in relation to the iron. These interactions of the dihydroxypropyl side chain with enzyme residues seem to be responsible for the change from a *trans* conformation to an almost-*cis* for the H1' and H2' on binding to the enzyme, as observed in the TRNOESY spectra (Figure 4.4(c)). Mutations at Ala322 are associated with a mild-PKU phenotype (Erlandsen et al., 1997)). Expression analysis have shown that the mutants A322T and A322G do not show reduced protein stability and are just kinetic variant forms (Svensson *et al*., 1992), in agreement with a role of Ala322 in the arrangement of the dihydroxypropyl side chain of the pterin.

Since C6 is sp^2 hybridized in BH_2 and sp^3 in BH_4, the orientation of the pterin ring to the metal may be different for BH_2 and BH_4. We have modeled the conformation of (6*R*,1'*R*,2'*S*)-6-(1',2'-dihydroxypropyl)-5,6,7,8-tetrahydropterin (BH_4) when bound to hPAH using the corresponding ring structure (Matsuura et al., 1985; Williams & Storm, 1985), superimposing the O1', O2' and C2' atoms of BH_4 to the corresponding ones in bound BH_2 and giving to the torsion angles [O2'-C2'-C1'-O1'] and [C3'-C2'-C1'-C6] similar values as measured for the hPAH-bound BH_2 (see below and Figure 5.4) For these BH_4 conformers, the distances from C4a to the metal and the interactions between the hydroxypropyl side chain and the enzyme were calculated to be about the same as in the bound BH_2 conformers.

Carboxylic acids show high affinity for the 2-amino pyrimidine moiety (Etter & Adsmond, 1990) and the recognition of the guanidino group of the pyrimidine ring in the pterin by either an Asp or Glu residue, as found in this work for Glu286 in human PAH (Figure 4.22(b)), seems to be a common motif for pterin binding proteins. Thus, in enzymes from the pathway of BH_4 biosynthesis such as GTP cyclohydrolase I (Nar *et al*., 1995), 6-pyruvovoyl tetrahydropterin synthase (Ploom *et al*., 1999) and sepiapterin reductase (Auerbach *et al*., 1997) carboxylate oxygens form hydrogen bonds with the N3 and the amine at C2, although there is no sequence identity among these enzymes. Dihydrofolate reductase also binds biopterin using this carboxylate anchor (McTigue et al., 1992) and, in addition, a water molecule hydrogen bonds with the two hydroxyls in the dihydroxypropyl side chain rendering an almost-*cis* conformation for the hydroxyls. Moreover, the molecular recognition of BH_4 in both inducible (Crane et al., 1998) and constitutive endothelial (Raman *et al*., 1998) nitric oxide synthase is mediated by one of the heme propionate groups, hydrogen bonding with the N3 and the amine at C2 of the pterin.

The homologous residue to Glu286, is Glu332 in rat TH. In the structures of the catalytic domains of both TH (Goodwill et al., 1997) and hPAH (Erlandsen et al., 1997) the side chain of this Glu residue is solvent exposed. However, in the recently reported crystal structure of the binary complex of TH with BH_2 (Goodwill et al., 1998), this residue does not have the same role as in hPAH. In the reported crystal structure, the pterin also binds to TH forming an aromatic π-stacking interaction with Phe300 (Phe254 in human PAH). However, the orientation of the pterin is different in TH, and the distance between the iron and the pterin C4a is 5.6 Å, *vs* 4.3 Å in the PAH·BH_2 complex. In addition, BH_2 forms hydrogen bonds from N8 to the main-chain carbonyl of Leu295 (Leu249 in hPAH), from the 4-oxo to both Tyr371 (Tyr325 in hPAH) and Glu376 (Glu330 in hPAH), and from the C1' OH to the main-chain amides of Leu294 and Leu295. Thus, in this crystal structure the N3 and the amine at C2 of BH_2 do not hydrogen bond with TH residues and the pterin seems to be rotated about 180° with respect to its conformation when bound to hPAH as studied by NMR. The reasons for this different binding mode of the pterin cofactor in TH are not clear. Although the high degree of structural similarity between both enzymes, the crystal structure of the binary complex of TH and BH_2 does not fit straightforward for hPAH. Thus, residues Leu248 and Leu249 (Leu294 and 295 in rat TH), located in a loop with different conformation in both enzymes, are about 1.5 Å longer away from the iron and the stacking Phe residue in hPAH than in TH. Furthermore, and in agreement with an important function of the residue Glu286 in the binding of the pterin cofactor, expression analysis of mutants at this residue in the catalytic domain of rat PAH have shown that the resultant protein is stable but shows a large decrease in affinity for BH_4, i.e. the K_m increased 76-fold for the mutant E286A (Dickson et al., 1994). Interestingly, mutations at the equivalent residue in rat TH (Glu332) also

results in a 10-fold increase in K_m for 6-methyltetrahydropterin (Daubner & Fitzpatrick, 1999), suggesting that the solution structure for the pterin bound to TH may be similar to that found in this work for hPAH, and different to that in the reported crystal structure (Goodwill et al., 1998).

Catecholamines bind to the enzyme by bidentate coordination to the iron (Erlandsen et al., 1998; Kappock & Caradonna, 1996; Martinez et al., 1991). The formation of the tight catecholate-Fe(III) complex seems to lower the redox potential and stabilize the ferric state, in agreement with a kinetic competition between the catecholamine inhibitors and the tetrahydropterin cofactors (Kaufman, 1993; Martinez et al., 1991). It has also been postulated a steric hindrance to the binding of the cofactor by the catecholamine as explanation for the competitive type of inhibition (Erlandsen et al., 1998). By aligning the structure of the PAH·dopamine complex (PDB accession number 5PAH) with that of the (PAH·L-Phe·BH_2) complex determined in this work, we found that there is overlapping of van der Waals ratios of the C5 and H6 atoms of the dopamine with the pterin (Figure 5.3).

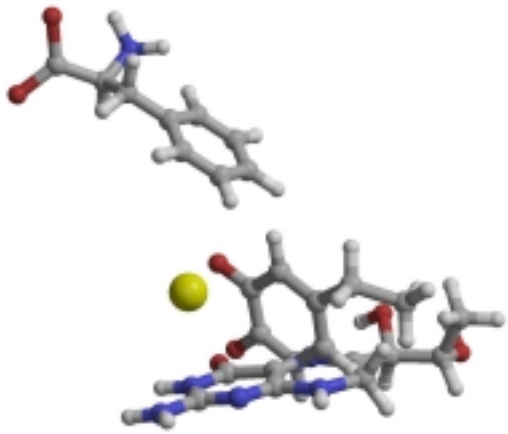


Figure 5.3
A view of the (PAH·L-Phe·BH_2) complex superimposed on the (PAH·dopamine) complex (PDB accession number 5PAH) (Erlandsen et al., 1998).

However, the fact that neither noradrenaline nor dopamine were able to displace BH_2 from its binding site in hPAH (Figures 4.14 and 4.15), indicates a rearrangement of the cofactor (and probably the substrate) when the catecholamine binds, forming a (PAH·L-Phe·BH_2·dopamine) complex.

5.6 The Ternary Complex and Implications for Catalysis

Earlier spectroscopic and binding studies had indicated that pterins bind in the proximity of, but not coordinating to, the non-heme iron in both hPAH (Kemsley *et al*., 1999; Shiman et al., 1994) and TH (Martinez et al., 1998; Meyer-Klaucke *et al*., 1996). It is shown in this study, however, that the pterin cofactor analogue BH_2 binds to the recombinant human PAH at a distance from the iron compatible with coordination *via* the 4-oxo, most probably replacing H_2O (1) (Figure 4.22(b)). Recent EPR spectroscopic studies have shown a change in the coordination geometry of the

Fe(III) following the binding of BH_2 to hPAH which is also compatible with direct coordination and substitution of a soft ligand by the pterin (Schmidt *et al*., 1998). Coordination is also expected for BH_4 and this finding has important mechanistic implications. Thus, the introduction of the electron-rich cofactor in the iron coordination sphere may trim the iron to bind and activate dioxygen, as it has been found in other enzymes that contain a 2-His-1-carboxylate motif coordinating the catalytic iron (Lange & Que, 1998). Although controversy exists about the sequential kinetic mechanism of hPAH regarding the order of substrate addition (Kappock & Caradonna, 1996; Kaufman, 1993), it seems clear that no product or intermediate is released prior to the binding of all substrates. This is also the case in uncoupled reactions occurring with substrate analogues, as L-Tyr or *p*-Cl-L-Phe or with BH_4 analogues with substituents at the 7 position, yielding oxidation of tetrahydropterin and formation of H_2O_2 (Kappock & Caradonna, 1996). This may be explained by the fact that the binding of the substrate/substrate analogue is accompanied by displacement of an additional water molecule (Martinez et al., 1993c; Olafsdottir & Martinez, 1999), probably H_2O (2) (Figure 4.22(b)), originating a free coordination site at the iron. In the presence of the substrate and the tetrahydropterin cofactor, dioxygen can then bind forming an iron-4a-peroxy-tetrahydropterin complex (Figure 5.4).

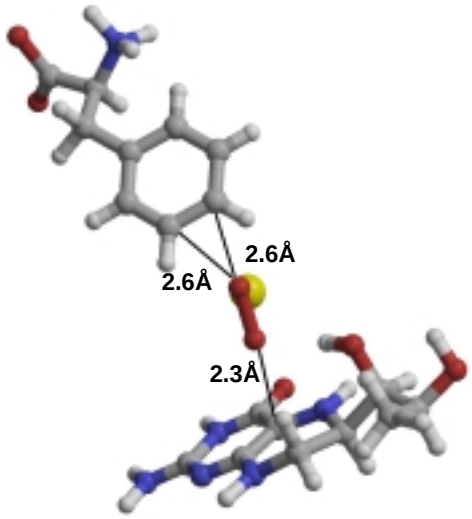


Figure 5.4
A model of the complex of the catalytic domain of human PAH with bound L-Phe, BH_4 and dioxygen molecule (in red) forming an iron-peroxo complex.

^{18}O kinetic isotope effect studies with TH have revealed that the rate-limiting step in the reaction is the reductive activation of molecular oxygen required in the formation of the hydroxylating species (Francisco *et al*., 1998). The reactive iron-peroxo species may then perform an electrophilic attack on the aromatic ring of the substrate, either directly to C4, which is the hydroxylation target and is placed at 4.2 Å from the iron, or to both C3 and C4 forming an arene oxide intermediate (Kappock & Caradonna, 1996; Kaufman, 1993). As a result, one of the dioxygen atoms would become attached to the C4a of the pterin ring and the other to the C4 of the substrate after spontaneous rearrangement of substituents at C3/C4. As previously shown by CD and NMR spectroscopies, oxygen addition to the pterin occurs in a *cis*

configuration with respect to the 6*R* substituents in naturally occurring tetrahydropterins (Dix & Benkovic, 1988), in agreement with the disposition of BH_4 and dioxygen in the modeled structure of the quaternary reaction complex (Figure 5.4).

The distance between the C4 atom of L-Phe and the C4a of BH_2 (or modeled BH_4) in the complex is about 6.3 Å, and is very suited to this proposed mechanism of hydroxylation (Figure 5.4). Accordingly, a peroxy-tetrahydropterin might be formed transiently in the reaction but not as a separated species, explaining why it has never been detected. This proposed mechanism holds for both the productive hydroxylation of the aromatic substrate and the uncoupled oxidation of the tetrahydropterin and also explains the formation of the 4a-hydroxypterin intermediate as the first product of the pterin cofactor (Dix & Benkovic, 1988; Haavik & Flatmark, 1987; Kappock & Caradonna, 1996). While the BH_4 substrate is quasiplanar (with half-semi chair conformation and C6 going out of plane), the product (pterin-4a-carbinolamine) has a pyramidal (sp^3) configuration at position 4a. The resulting bend in the pterin ring may impair the stacking with Phe254 resulting in a decreased affinity for this product and its expulsion from the binding site. Moreover, the product L-Tyr could then be expelled from its binding site due to steric and ionic repulsion with the carboxylic oxygen of Glu330.

5.7 Implications for Regulation of PAH by Substrate and Pterin Cofactor

The interactions of the two hydroxyl groups of the side chain at C6 with the carbonyl group of Ala322 and the hydroxyl group of Ser251 seem to explain the high specificity of hPAH for the natural cofactor BH_4, containing the dihydroxypropyl side chain, as well as its stereospecificity for the 6*R*- isomer (Haavik et al., 1986; Kappock & Caradonna, 1996; Kaufman, 1993). It is interesting to note that although Ser251 and Ala322 are conserved among mammalian PAH they are not conserved among the aromatic amino acid hydroxylases, a fact that may be related to the specific regulatory properties elicited by the cofactors in the three enzymes. No additional interactions seem to occur between the enzyme and BH_2 when the docking was performed using the crystal structure of the enzyme containing the regulatory domain (Kobe *et al*., 1999) (Figure 5.5).

Figure 5.5
The crystal structure of dimeric rat PAH containing the regulatory and catalytic domains (PDB accession number 2PHM) (Kobe et al., 1999) with L-Phe, BH_2 and dopamine (bound as in Figure 5.3). The N-terminal 33 residues are shown in green, and the rest of the regulatory domain in red. The catalytic domain is shown in blue. Residues Ser23, Tyr24 and Ile25 are shown as green sticks, Ser251 and Ala322 as red sticks and the Fe(III) atom in yellow.

Some interesting features were found that may be related to the negative regulation that the natural cofactor BH_4, and not synthetic cofactors as 6-methyltetrahydropterin, exerts in hPAH. Thus, both Ser251 and Ala222 establish polar and non-polar interactions with residues 23-25 at the regulatory N-terminal of the rat enzyme, in the region referred to as the autoregulatory sequence (Kobe et al., 1999) (Figure 5.5) and which is assumed to have an inhibitory function in the non-activated hPAH. This network of specific contacts from the dyhydroxypropyl side-chain towards the N-terminal sequence may lock the enzyme in an inhibited state.

Tetrameric full-length hPAH is activated by preincubation with L-Phe and also shows positive cooperativity for substrate binding (Kappock & Caradonna, 1996; Kaufman, 1993). Many investigators have interpreted this substrate activation as being the result of binding of the substrate to two physically different sites: an allosteric regulatory site and a catalytic site near the iron. We have interpreted the activation of L-Phe as being the result of the homotropic cooperative binding of L-Phe at the active sites (Martinez *et al*., 1990; Martinez et al., 1993c). As shown in this work, the conformation of bound L-Phe, including the distances to the iron, was found to be similar for the tetrameric wt-hPAH and the truncated dimeric hPAH(Gly103-Gln428) (see above), further supporting that L-Phe only binds at one site. Apparently, no additional contact between the enzyme and L-Phe is observed when the conformation of the bound substrate is docked into the crystal structure of the dimeric form of rat PAH containing the regulatory sequence (Figure 5.5).

Moreover, residues from the autoregulatory sequence (or from the rest of the N-terminal domain) do not interact with residues involved in the binding of L-Phe. However, in the crystal structure of this form containing the regulatory domain, there is no structural information about the N-terminal 19 residues, which may have some interactions with the substrate or with residues at the substrate binding site, a binding of L-Phe to hPAH may result in a displacement of the regulatory sequence, resulting in an alleviation of the inhibitory effect exerted by this sequence and in a triggering of positive cooperative conformational changes affecting the other subunits. We have earlier shown that catecholamines bind to the rat enzyme with positive cooperativity inducing similar conformational changes in hPAH as L-Phe does (Martinez et al., 1990; Martinez et al., 1993c). As seen by the displacement experiments shown in Figures 4.14 and 4.15 and in the superimposed structures of the PAH·L-Phe·BH_2 and PAH·dopamine complexes (Figure 4.22(b)) no overlapping exists between the substrate and the catecholamine binding sites. The positive cooperativity manifested by the full-length enzyme for the binding of substrate and inhibitor may thus be the result of similar conformational changes induced upon binding of both ligands. Interestingly, modeling of the catecholamine in the crystal structure of the dimeric rat PAH containing the regulatory sequence shows that the amino group of the catecholamines overlaps the position of Ile25 (Figure 5.5). Thus, binding of the catecholamine may induce the displacement of the autoregulatory sequence initiating cooperative conformational changes. Nevertheless, the complete understanding of the cooperative transmission of conformational changes between subunits must await the determination of the structure of the tetrameric form of the enzyme in the resting and activated forms.

References

Aberg, A., Yaremchuk, A., Tukalo, M., Rasmussen, B. & Cusack, S. (1997). Crystal structure analysis of the activation of histidine by Thermus thermophilus histidyl-tRNA synthetase. *Biochemistry* **36**, 3084-94.

Abita, J. P., Milstien, S., Chang, N. & Kaufman, S. (1976). In vitro activation of rat liver phenylalanine hydroxylase by phosphorylation. *J. Biol. Chem.* **251**, 5310-4.

Al-Karaghouli, A. R. & Koetzle, T. F. (1975). Neutron difraction study of L-phenylalanine hydrochloride. *Acta Crystallogr.* **B31**, 2461-2465.

Almås, B., Teigen, K., Toska, K., Groehn, V., Pfleiderer, W., Martínez, A., Flatmark, T. & Haavik, J. (in press). Binding of tetrahydropterins to tyrosine hydroxylase. A kinetic and molecular docking study.

Auerbach, G., Herrmann, A., Gutlich, M., Fischer, M., Jacob, U., Bacher, A. & Huber, R. (1997). The 1.25 A crystal structure of sepiapterin reductase reveals its binding mode to pterins and brain neurotransmitters. *Embo J.* **16**, 7219-30.

Bailey, S. W. & Ayling, J. E. (1983). 6,6-Dimethylpterins: stable quinoid dihydropterin substrate for dihydropteridine reductase and tetrahydropterin cofactor for phenylalanine hydroxylase. *Biochemistry* **22**, 1790-8.

Bickel, H., Gerrard, J. & Hickmans, E. M. (1954). The influence of phenylalanine intake on chemistry and behavior of a phenylkteonuria child. *Acta Paeditr. Scand.* **43**, 64-77.

Bieri, J. H. (1977). The crystal structure of 6-methyl-7,8-dihydropterine-monohydrochloride-monohydrate. *Helv. Chim. Acta.* **60**, 2303-2308.

Billeter, M., Havel, T. F. & Kuntz, I. D. (1986). A new approach to the problem of docking of two molecules: The ellipsoid algorithm. *Biopolymers* **26**, 777-793.

Bjørgo, E., Knappskog, P. M., Martínez, A., Stevens, R. C. & Flatmark, T. (1998). Partial characterization and three-dimensional-structural localization of eight mutations in exon 7 of the human phenylalanine hydroxylase gene associated with phenylketonuria. *Eur. J. Biochem* **257**, 1-10.

Bublitz, C. (1971). Two mechanisms for the inhibition in vitro of phenylalanine hydroxylase by catecholamines. *Biochem. Pharmacol.* **20**, 2543-53.

Campbell, A. P. & Sykes, B. D. (1991). Theoretical evaluation of the two-dimensional transferred nuclear Overhauser effect. *J. Magn. Res.* **93**, 77-92.

Campbell, A. P. & Sykes, B. D. (1993). The two-dimensional transferred nuclear Overhauser effect: theory and practice. *Annu. Rev. Biophys. Biomol. Struct.* **22**, 99-122.

Citron, B. A., Davis, M. D. & Kaufman, S. (1992). Purification and biochemical characterization of recombinant rat liver phenylalanine hydroxylase produced in Escherichia coli. *Protein Expr.Purif.* **3**, 93-100.

Crane, B. R., Arvai, A. S., Ghosh, D. K., Wu, C., Getzoff, E. D., Stuehr, D. J. & Tainer, J. A. (1998). Structure of nitric oxide synthase oxygenase dimer with pterin and substrate. *Science* **279**, 2121-6.

Daubner, S. C. & Fitzpatrick, P. F. (1999). Site-directed mutants of charged residues in the active site of tyrosine hydroxylase. *Biochemistry* **38**, 4448-54.

Daubner, S. C., Hillas, P. J. & Fitzpatrick, P. F. (1997). Characterization of chimeric pterin-dependent hydroxylases: contributions of the regulatory domains of tyrosine and phenylalanine hydroxylase to substrate specificity. *Biochemistry* **36**, 11574-11582.

Davis, M. D. & Kaufman, S. (1989). Evidence for the formation of the 4a-carbinolamine during the tyrosine- dependent oxidation of tetrahydrobiopterin by rat liver phenylalanine hydroxylase. *J. Biol. Chem.* **264**, 8585-96.

DesJarlais, R. L., Sheridan, R. P., Seibel, G. L., Dixon, J. S., Kuntz, I. D. & Venkataraghavan, R. (1988). Using shape complementary as an initial screen in designing ligands for a receptor binding site of known three-dimensional structure. *J. Med. Chem.* **31**, 722-729.

Dickson, P. W., Jennings, I. G. & Cotton, R. G. (1994). Delineation of the catalytic core of phenylalanine hydroxylase and identification of glutamate 286 as a critical residue for pterin function. *J.Biol.Chem.* **269**, 20369-20375.

Dix, T. A. & Benkovic, S. J. (1988). Mechanism of oxygen activation by pteridine-dependent moooxygenases. *Acc. Chem. Res.* **21**, 101-107.

Døskeland, A., Ljones, T., Skotland, T. & Flatmark, T. (1982). Phenylalanine 4-monooxygenase from bovine and rat liver: some physical and chemical properties. *Neurochem.Res.* **7**, 407-421.

Døskeland, A. P. & Flatmark, T. (1996). Recombinant human phenylalanine hydroxylase is a substrate for the ubiquitin-conjugating enzyme system. *Biochem. J.* **319**, 941-945.

Dress, A. W. & Havel, T. F. (1988). Shortest-path problems and molecular conformation. *Discrete appl. Math.* **19**, 129-144.

Døskeland, A. P., Døskeland, S. O., Ogreid, D. & Flatmark, T. (1984). The effect of ligands of phenylalanine 4-monooxygenase on the cAMP-dependent phosphorylation of the enzyme. *J.Biol.Chem.* **259**, 11242-11248.

Eisensmith, R. C. & Woo, S. L. (1991). Phenylketonuria and the phenylalanine hydroxylase gene. *Mol. Biol. Med.* **8**, 3-18.

Erlandsen, H., Flatmark, T., Stevens, R. C. & Hough, E. (1998). Crystallographic analysis of the human phenylalanine hydroxylase catalytic domain with bound catechol inhibitors at 2.0 A resolution [In Process Citation]. *Biochemistry* **37**, 15638-46.

Erlandsen, H., Fusetti, F., Martínez, A., Hough, E., Flatmark, T. & Stevens, R. C. (1997). Crystal structure of the catalytic domain of human phenylalanine hydroxylase reveals the structural basis for phenylketonuria [letter]. *Nat.Struct.Biol.* **4**, 995-1000.

Erlandsen, H. & Stevens, R. C. (1999). The Structural Basis of Phenylketonuria. *Mol. Genet. Metab.* **68**, 103-125.

Etter, M. C. & Adsmond, D. A. (1990). The use of cocrystallization as a method of stidying hydrogen bond preferences of 2-aminopyrimidine. *J. Chem. Soc. Chem. Commun.* **8**, 589-591.

Ewing, T. J. A. & Kuntz, I. D. (1996). Critical evaluation of search algorithms for automated molecular docking and database screening. *J. Comp. Chem.* **18**, 1175-1189.

Ferrin, T. E., Huang, C. C., Jarvis, L. E. & Langridge, R. (1988). The MIDAS display system. *J. Mol. Graphics* **6**, 13-27.

Fitzpatrick, P. F. (1991). Steady-state kinetic mechanism of rat tyrosine hydroxylase. *Biochemistry* **30**, 3658-62.

Francisco, W. A., Tian, G. C., Fitzpatrick, P. F. & Klinman, J. P. (1998). Oxygen-18 kinetic isotope effect studies of the tyrosine hydroxylase reaction: Evidence of rate limiting oxygen activation. *J. Am. Chem. Soc.* **120**, 4057-4062.

Fusetti, F., Erlandsen, H., Flatmark, T. & Stevens, R. C. (1998). Structure of tetrameric human phenylalanine hydroxylase and its implications for phenylketonuria. *J. Biol. Chem.* **273**, 16962-7.

Følling, A. (1934). Über ausscheidung von phenylbrentztraumbensäure in der harn als stoffwechselanomaile in verbindung mit imbezillität. *Hoppe-Seylers Z. Physiol.Chem.* **227**, 169-176.

Goodwill, K. E., Sabatier, C., Marks, C., Raag, R., Fitzpatrick, P. F. & Stevens, R. C. (1997). Crystal structure of tyrosine hydroxylase at 2.3 A and its implications for inherited neurodegenerative diseases. *Nat.Struct.Biol.* **4**, 578-585.

Goodwill, K. E., Sabatier, C. & Stevens, R. C. (1998). Crystal structure of tyrosine hydroxylase with bound cofactor analogue and iron at 2.3 A resolution: self-hydroxylation of Phe300 and the pterin-binding site. *Biochemistry* **37**, 13437-45.

Havel, T. F. (1991). An evaluation of computational strategies for use in the determination of protein structure from distance constraints obtained by nuclear magnetic resonance. *Prog. Biopys. molec. Biol.* **56**, 43-78.

Havel, T. F., Kuntz, I. D. & Crippen, G. M. (1983). Theory and practice of distance geometry. *Bull. math. Biol.* **45**, 665-720.

Hufton, S. E., Jennings, I. G. & Cotton, R. G. (1995). Structure and function of the aromatic amino acid hydroxylases. *Biochem. J.* **311**, 353-66.

Haavik, J., Døskeland, A. P. & Flatmark, T. (1986). Stereoselective effects in the interactions of pterin cofactors with rat-liver phenylalanine 4-monooxygenase. *Eur. J. Biochem.* **160**, 1-8.

Haavik, J. & Flatmark, T. (1987). Isolation and characterization of tetrahydropterin oxidation products generated in the tyrosine 3-monooxygenase (tyrosine hydroxylase) reaction. *Eur. J. Biochem.* **168**, 21-6.

Jervis, G. A. (1953). Phenylpyruvic oligophrenia deficiency of phenylalanine-oxidizing system. *Proc. Soc. Expt. Biol. Med.* **82**, 514-515.

Kappock, T. J. & Caradonna, J. P. (1996). Pterin-Dependent Amino Acid Hydroxylases. *Chem.Rev.* **96**, 2659-2756.

Kappock, T. J., Harkins, P. C., Friedenberg, S. & Caradonna, J. P. (1995). Spectroscopic and kinetic properties of unphosphorylated rat hepatic phenylalanine hydroxylase expressed in Escherichia coli - Comparison of resting and activated states. *J. Biol. Chem.* **270**, 30532-30544.

Katoh, S., Sueoka, T. & Kurihara, T. (1993). Theoretical Stereostructure of the neutral form of natural tetrahydrobiopterin. *Pteridines* **4**, 27-31.

Kaufman, S. (1993). The phenylalanine hydroxylating system. *Adv Enzymol Relat Areas Mol. Biol.* **67**, 77-264.

Kaufman, S. & Fisher, D. B. (1970). Purification and some physical properties of phenylalanine hydroxylase from rat liver. *J. Biol. Chem.* **245**, 4745-50.

Kemsley, J. N., Mitic, N., Zaleski, K. L., Caradonna, J. P. & Solomon, I. (1999). Circular dichroism and magnetic circular dichroism spectroscopy of the catalytically competent ferrous active site of phenylalanine hydroxylase and its interaction with pterin cofactor. *J. Am. Chem. Soc.* **121**, 1528-1536.

Knappskog, P. M., Eiken, H. G., Martínez, A., Flatmark, T. & Apold, J. (1995). The PKU mutation S349P causes complete loss of catalytic activity in the recombinant phenylalanine hydroxylase enzyme. *Hum. Genet.* **95**, 171-3.

Knappskog, P. M., Flatmark, T., Aarden, J. M., Haavik, J. & Martínez, A. (1996). Structure/function relationships in human phenylalanine hydroxylase. Effect of terminal deletions on the oligomerization, activation and cooperativity of substrate binding to the enzyme. *Eur. J. Biochem.* **242**, 813-21.

Knappskog, P. M. & Haavik, J. (1995). Tryptophan fluorescence of human phenylalanine hydroxylase produced in Escherichia coli. *Biochemistry* **34**, 11790-11799.

Kobe, B., Jennings, I. G., House, C. M., Michell, B. J., Goodwill, K. E., Santarsiero, B. D., Stevens, R. C., Cotton, R. G. & Kemp, B. E. (1999). Structural basis of autoregulation of phenylalanine hydroxylase [see comments]. *Nat. Struct. Biol.* **6**, 442-8.

Kraulis, P. J. (1991). MOLSCRIPT - a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946-950.

Kuntz, I. D., Thomason, J. F. & Oshiro, C. M. (1989). Distance geometry. *Meth. Enzymol.* **177**, 159-204.

Kwok, S. C., Ledley, F. D., DiLella, A. G., Robson, K. J. & Woo, S. L. (1985). Nucleotide sequence of a full-length complementary DNA clone and amino acid sequence of human phenylalanine hydroxylase. *Biochemistry* **24**, 556-561.

Lange, S. & Que, L. J. (1998). Oxygen activating nonheme iron enzymes. *Curr. Opin. Chem. Biol.* **2**, 159-172.

Martínez, A., Abeygunawardana, C., Haavik, J., Flatmark, T. & Mildvan, A. S. (1993a). Conformation and interaction of phenylalanine with the divalent cation at the active site of human recombinant tyrosine hydroxylase as determined by proton NMR. *Biochemistry* **32**, 6381-6390.

Martínez, A., Abeygunawardana, C., Haavik, J., Flatmark, T. & Mildvan, A. S. (1993b). Interaction of substrate and pterin cofactor with the metal of human tyrosine hydroxylase as determined by 1H-NMR. *Adv.Exp.Med.Biol.* **338**, 77-80.

Martínez, A., Andersson, K. K., Haavik, J. & Flatmark, T. (1991). EPR and 1H-NMR spectroscopic studies on the paramagnetic iron at the active site of phenylalanine hydroxylase and its interaction with substrates and inhibitors. *Eur. J. Biochem.* **198**, 675-682.

Martínez, A., Haavik, J. & Flatmark, T. (1990). Cooperative homotropic interaction of L-noradrenaline with the catalytic site of phenylalanine 4-monooxygenase. *Eur.J.Biochem.* **193**, 211-219.

Martínez, A., Knappskog, P. M., Olafsdottir, S., Døskeland, A. P., Eiken, H. G., Svebak, R. M., Bozzini, M., Apold, J. & Flatmark, T. (1995). Expression of recombinant human phenylalanine hydroxylase as fusion protein in Escherichia coli circumvents proteolytic degradation by host cell proteases. Isolation and characterization of the wild-type enzyme. *Biochem. J.* **306**, 589-597.

Martínez, A., Olafsdottir, S. & Flatmark, T. (1993c). The cooperative binding of phenylalanine to phenylalanine 4- monooxygenase studied by 1H-NMR paramagnetic relaxation. Changes in water accessibility to the iron at the active site upon substrate binding. *Eur. J. Biochem.* **211**, 259-266.

Martínez, A., Vageli, O., Pfleiderer, W. & Flatmark, T. (1998). Proton NMR studies on the conformation of the pterin cofactor bound at the active site of recombinant human tyrosine hydroxylase. *Pteridines* **9**, 44-52.

Matsuura, S., Sugimoto, T., Murata, S., Sugawara, Y. & Iwasaki, H. (1985). Stereochemistry of biopterin cofactor and facile methods for the determination of the stereochemistry of a biologically active 5,6,7,8- tetrahydropterin. *J. Biochem.* **98**, 1341-8.

McTigue, M. A., Davies, J. F. d., Kaufman, B. T. & Kraut, J. (1992). Crystal structure of chicken liver dihydrofolate reductase complexed with NADP+ and biopterin. *Biochemistry* **31**, 7264-73.

Merrit, E. A. & Murphy, M. E. P. (1994). RASTER3D version-2.0a program for photorealistic molecular graphics. *Acta Cryst. Section D* **50**, 869-873.

Meyer-Klaucke, W., Winkler, H., Schunemann, V., Trautwein, A. X., Nolting, H. F. & Haavik, J. (1996). Mossbauer, electron-paramagnetic-resonance and X-ray-absorption fine- structure studies of the iron environment in recombinant human tyrosine hydroxylase. *Eur. J. Biochem.* **241**, 432-9.

Mildvan, A. S., Granot, J., Smith, G. M. & Liebman, M. n. (1980). Nuclear magnetic relaxation rates. *Adv. Inorg. Biochem.* **2**, 211-236.

Nakata, H. & Fujisawa, H. (1980). Purification and characterization of phenylalanine 4-monooxygenase from rat liver. *Biochim. Biophys. Acta* **614**, 313-27.

Nar, H., Huber, R., Auerbach, G., Fischer, M., Hosl, C., Ritz, H., Bracher, A., Meining, W., Eberhardt, S. & Bacher, A. (1995). Active site topology and reaction mechanism of GTP cyclohydrolase I. *Proc Natl. Acad. Sci. U S A* **92**, 12120-5.

Odani, A., Masuda, H., Inukai, K. & Yamauchi, O. (1992). Ptedirine-containing ternaty and quaternary complexes as models for metalloenzyme-pterin cofactor-substrate association-structure of ternary copper(II)-2,2'-bipyridine lumazine complex and successful equilibrium study of a quaternary copper(II) system. *J. Am. Chem. Soc.* **114**, 6294-6300.

Oh, B. H., Kang, C. H., De Bondt, H., Kim, S. H., Nikaido, K., Joshi, A. K. & Ames, G. F. (1994). The bacterial periplasmic histidine-binding protein. structure/function analysis of the ligand-binding site and comparison with related proteins. *J. Biol. Chem.* **269**, 4135-43.

Olafsdottir, S. & Martínez, A. (1999). The accessibility of iron at the active site of recombinant human phenylalanine hydroxylase to water as studied by 1H NMR paramagnetic relaxation. Effect of L-Phe and comparison with the rat enzyme. *J. Biol. Chem.* **274**, 6280-6284.

Parniak, M. A. & Kaufman, S. (1981). Rat liver phenylalanine hydroxylase. Activation by sulfhydryl modification. *J. Biol. Chem.* **256**, 6876-6882.

Phillips, R. S., Parniak, M. A. & Kaufman, S. (1984a). The interaction of aromatic amino acids with rat liver phenylalanine hydroxylase. *J. Biol. Chem.* **259**, 271-277.

Phillips, R. S., Parniak, M. A. & Kaufman, S. (1984b). Spectroscopic investigation of ligand interaction with hepatic phenylalanine hydroxylase: evidence for a conformational change associated with activation. *Biochemistry* **23**, 3836-3842.

Piotto, M., Saudek, V. & Sklenar, V. (1992). Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR* **2**, 661-5.

Ploom, T., Hau#mann, C., Hof, P., Steinbacher, S., Bacher, A., Richardson, J. & Huber, R. (1999). Crystal structure of 7,8-dihydroneopterin triphosphate epimerase. *Structure* **7**, 509-516.

Raman, C. S., Li, H., Martasek, P., Kral, V., Masters, B. S. & Poulos, T. L. (1998). Crystal structure of constitutive endothelial nitric oxide synthase: a paradigm for pterin function involving a novel metal center. *Cell* **95**, 939-50.

Schmidt, P. P., Martínez, A., Barra, A. L., Flatmark, T. & Andersson, K. K. (1998). EPR characterisations on the Fe(III) site of full-length and truncated forms of recombinant hyman phenylalanine hydroxylase at various frequencies. *Fourth European Biological Inorganic Chemistry Conference, Sevilla, Spain.*, 180.

Sheridan, R. P., Nilakantan, R. S., Dixon, J. & Venkataraghavan, R. (1986). The ensemble approach to distnace geometry: Application to the nicotinic pharmacophore. *J. Med. Chem.* **29**, 899-906.

Shiman, R. & Gray, D. W. (1980). Substrate activation of phenylalanine hydroxylase. A kinetic characterization. *J. Biol. Chem.* **255**, 4793-800.

Shiman, R., Gray, D. W. & Pater, A. (1979). A simple purification of phenylalanine hydroxylase by substrate-induced hydrophobic chromatography. *J. Biol. Chem.* **254**, 11300-6.

Shiman, R., Jones, S. H. & Gray, D. W. (1990). Mechanism of phenylalanine regulation of phenylalanine hydroxylase. *J. Biol. Chem.* **265**, 11633-11642.

Shiman, R., Xia, T., Hill, M. A. & Gray, D. W. (1994). Regulation of rat liver phenylalanine hydroxylase. II. Substrate binding and the role of activation in the control of enzymatic activity. *J. Biol. Chem.* **269**, 24647-24656.

Sklenar, V., Piotto, M., Leppik, R. & Saudek, V. (1993). Gradient-tailored water suppression for H-1-N-15 HSQC experiments optimized to full sensitivity. *J. Mag. Res.* **102**, 241-245.

Svensson, E., Eisensmith, R. C., Dworniczak, B., von Dobeln, U., Hagenfeldt, L., Horst, J. & Woo, S. L. (1992). Two missense mutations causing mild hyperphenylalaninemia associated with DNA haplotype 12. *Hum. Mutat.* **1**, 129-37.

Williams, T. C. & Storm, C. B. (1985). Tetrahydrobiopterin analogues: solution conformations of 6-methyltetrahydropterin, 7-methyltetrahydropterin, and cis- and trans- 6,7-dimethyltetrahydropterins as determined by proton nuclear magnetic resonance. *Biochemistry* **24**, 458-66.

Woolf, L. I. (1976). The isolation, properties, and assay of phenylalanine hydroxylase from human and rat liver. *Biochem.Med.* **16**, 284-291.

Woolf, L. I., Griffiths, R. & Moncrief, A. (1955). Treatment of phenylketonuria with a diet low in phenylalanine. *Brit. Med. J.* **1**, 57-64.

Wretborn, M., Humble, E., Ragnarsson, U. & Engstrom, L. (1980). Amino acid sequence at the phosphorylated site of rat liver phenylalanine hydroxylase and phosphorylation of a corresponding synthetic peptide. *Biochem. Biophys. Res. Commun.* **93**, 403-8.

Wütrich, K. (1986). NMR of proteins and nucleic acids. *J.Wiley and sons, New York, NY* ISBN 0 471 82893 9.

Xia, T., Gray, D. W. & Shiman, R. (1994). Regulation of rat liver phenylalanine hydroxylase. III. Control of catalysis by (6R)-tetrahydrobiopterin and phenylalanine. *J. Biol. Chem.* **269**, 24657-24665.

Appendix I

Assignment of Resonances

Figure 6.1 and 6.2 show the HMBC and HSQC spectra, respectively of L-Phe and BH_2 (5 mM) in 10% deuterated 20 mM potassium phosphate buffer, pH 7.2.

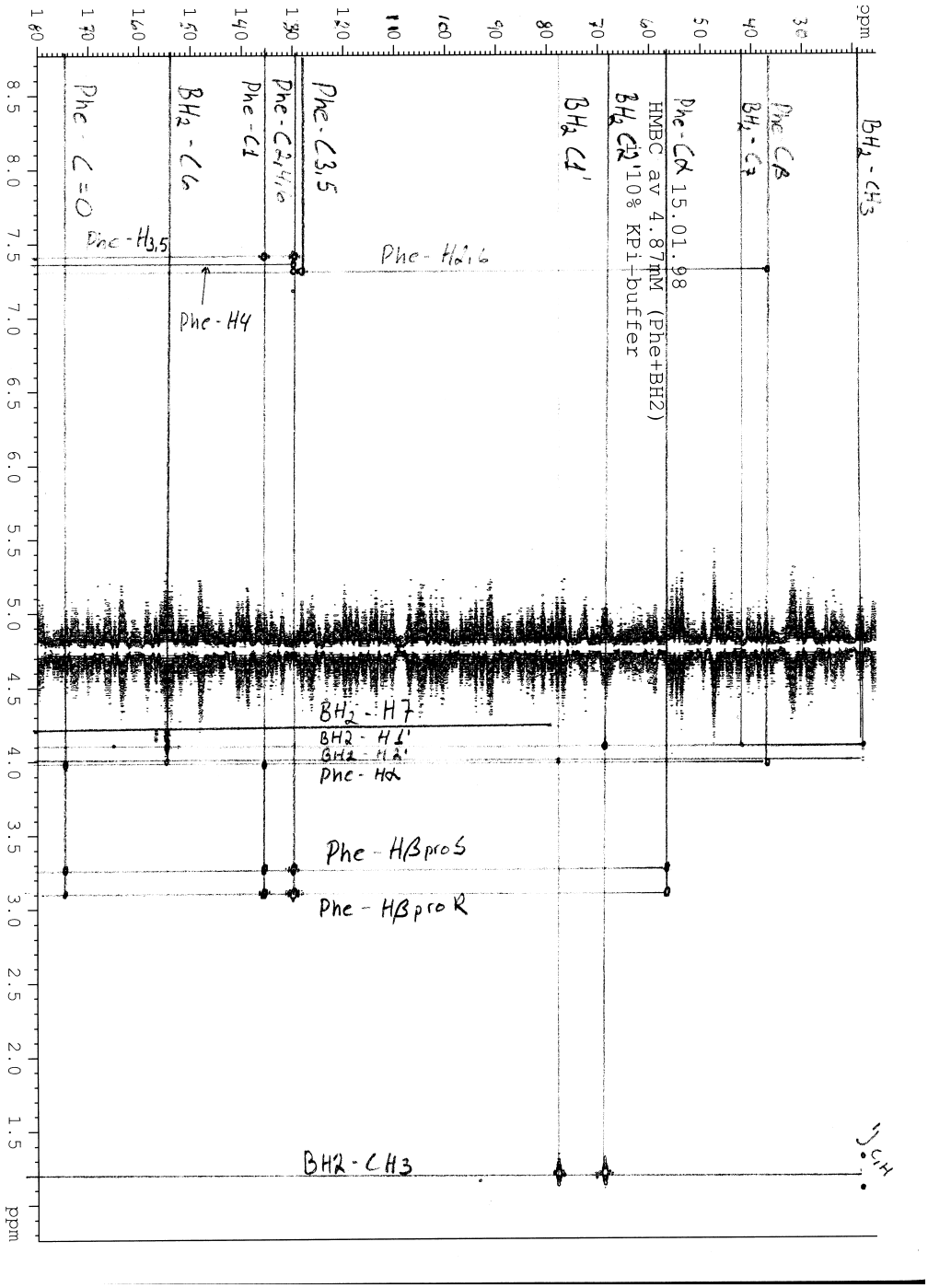


Figure 6.1

HMBC spectrum of L-Phe and BH_2 (5 mM). Cross peaks correlate proton chemical shifts and carbon chemical shifts of atoms separated by two or three bonds.

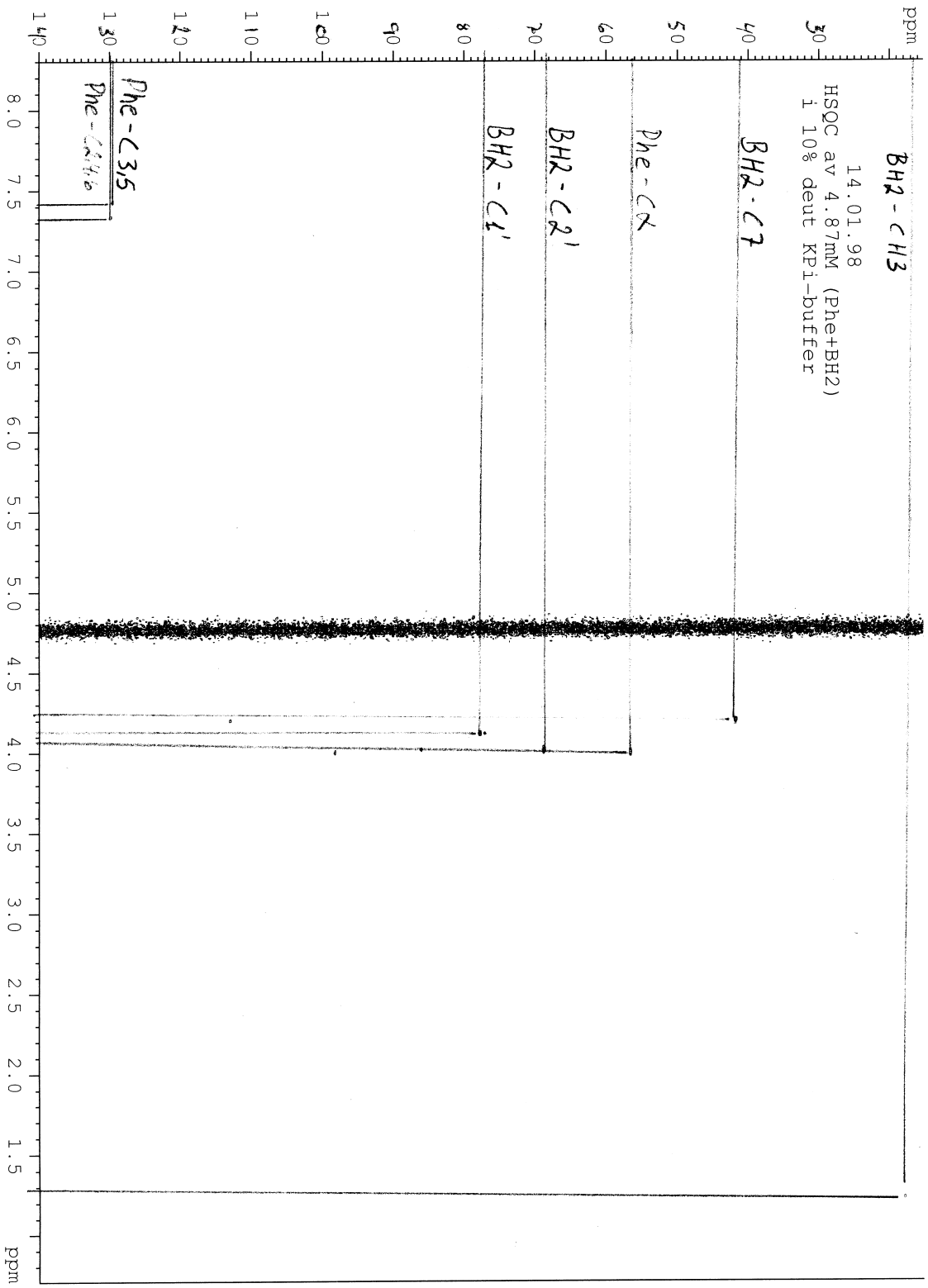


Figure 6.2
HSQC spectrum of L-Phe and BH_2 (5mM). Cross peaks correlate the chemical shifts of protons to the chemical shift of their respective carbon atoms direclty bonded to each other.

Appendix II

Trp326 signal

Figure 6.3 shows the aromatic part of the TRNOESY spectrum of L-Phe and BH_2 (5 mM) in the presence of 0.5 mM hPAH(Gly103-Gln428).

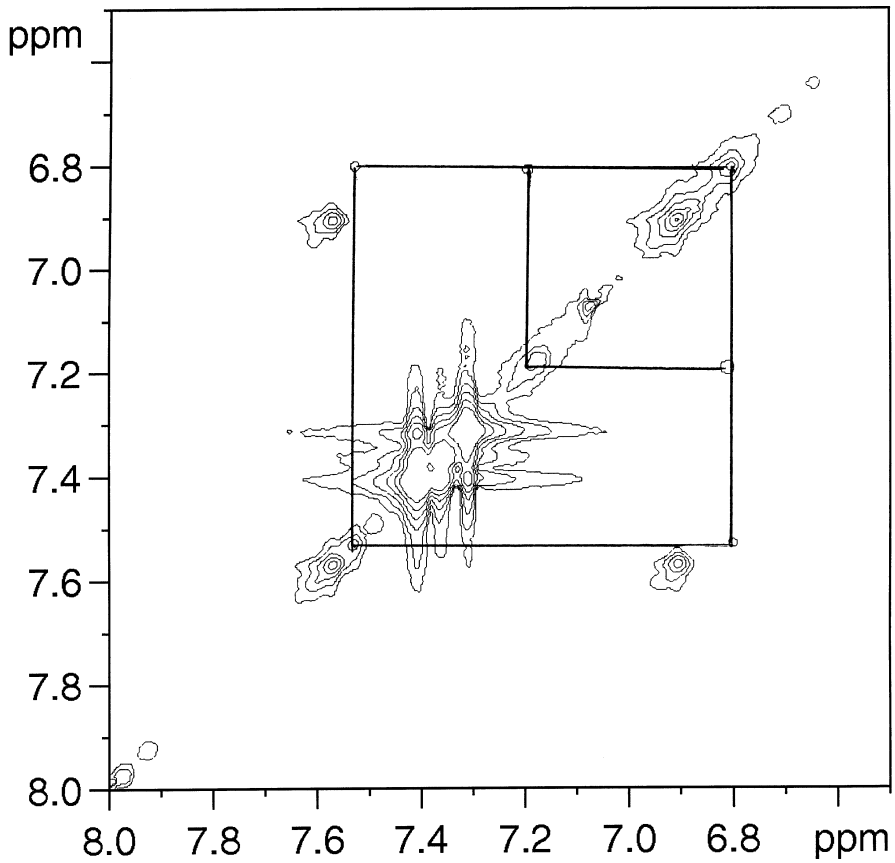


Figure 6.3
The three aromatic signals at 7.54 ppm, 7.21 ppm and 6.80 ppm from the protein showing strong NOE cross peaks. These are believed to arise from Trp326. The three resonances between 7.25 ppm and 7.5 ppm on the diagonal are from the aromatic protons of the substrate.

Appendix III

Grid run

The first part (first page) of the file shown below is a copy of the grid input file. The last part of the file (second page) gives information about the performance of the grid run. Calculating the grid is usually the most time consuming process of the docking procedure. The time required to run grid is strongly dependent on the chosen parameters, in particular the size of the grid box. The Octane (SGI) workstation spent 4 hours on this particular run. When the grid box was reduced to only include the residues around Glu286, the grid calculation took 20 minutes.

```
        UUUUUUU    CCCCCCCC    SSSSSSS/  FFFFFFFFF
      UU/    UU/  CC/    CC/     SS/     FF/    FF/
     UU/         CCCCCCCC/      SS/     FF/    FF/
    UU/  UUUU/  CC/ CC\         SS/     FF/    FF/
   UU/    UU/  CC/   CC\       SS/     FF/    FF/
   UUUUUUU/   CC/     CC\  SSSSSS/  FFFFFFFFF/


University of California at San Francisco, DOCK 4.0.1

____________________Job_Information_____________________
launch_time                    Sun Jun 27 23:32:55 1999
host_name                      octane.bio.uib.no
memory_limit                   255524864
working_directory              /disk2/people/kteigen/dock/PAH_27_6_99/grid
user_name                      kteigen

_________________General_Parameters_____________________
compute_grids                  yes
grid_spacing                   0.2
output_molecule                yes

_________________Scoring_Parameters_____________________
contact_score                  yes
contact_cutoff_distance        4.5
chemical_score                 yes
energy_score                   yes
energy_cutoff_distance         15
atom_model                     united
attractive_exponent            6
repulsive_exponent             12
distance_dielectric            yes
dielectric_factor              4
bump_filter                    yes
bump_overlap                   0.5

_____________________File_Input_____________________
receptor_file                  ../struc/PAH.mol2
box_file                       ../../test/27_1_1999/showbox_out2.pdb
vdw_definition_file            /disk2/software/dock/parameter/vdw.defn
chemical_definition_file       /disk2/software/dock/parameter/chem.defn

_____________________File_Output_____________________
score_grid_prefix              grid
receptor_out_file              PAH_gridout.mol2
```

```
Reading in coordinates of receptor.
CHARGED RESIDUE ARG+123                 :    1.000
CHARGED RESIDUE GLU-127                 :   -1.000
CHARGED RESIDUE ASP-129                 :   -1.000
CHARGED RESIDUE ARG+130                 :    1.000
.
.
.
.
CHARGED RESIDUE ASP-415                 :   -1.000
CHARGED RESIDUE ARG+420                 :    1.000
CHARGED RESIDUE GLU-422                 :   -1.000

Total charge on SUBUNA_UTEN             :   -5.000

Writing out processed receptor.

Reading in grid box information.
Box center of mass                      :   -0.506    1.539  -29.228
Box dimensions                          :   33.303   26.533   26.983
Number of grid points per side [x y z]  :      168      134      136
Total number of grid points             :  3061632

Generating scoring grids.
Percent of protein atoms processed      :        0
Percent of protein atoms processed      :       10
Percent of protein atoms processed      :       20
Percent of protein atoms processed      :       30
Percent of protein atoms processed      :       40
Percent of protein atoms processed      :       50
Percent of protein atoms processed      :       60
Percent of protein atoms processed      :       70
Percent of protein atoms processed      :       80
Percent of protein atoms processed      :       90
Percent of protein atoms processed      :      100
Writing general grid info to grid.bmp
Writing bump grid to grid.bmp
Writing contact grid to grid.cnt
Writing chemical grids to grid.chm
  Writing null chemical grid
  Writing hydrophobic chemical grid
  Writing donor chemical grid
  Writing acceptor chemical grid
  Writing polar chemical grid
Writing energy grids to grid.nrg
  Writing attractive VDW energy grid
  Writing repulsive VDW energy grid
  Writing electrostatic energy grid

Finished calculation.
```

Appendix IV

Dock run

The file shown below is generated when starting a dock run. The first part of the file (three first pages) is a copy of the dock input file. All the parameters are listed, also the parameters not used (e.g. Flexible Ligand Parameters and Chemical Screen Parameters). The last page of the file gives information about the performance of the dock run. The input file for vdw definitions (my_vdw.defn) is edited to include Fe(III) and is also shown in the appendix (V).

```
      UUUUUUUUU     CCCCCCC     SSSSSSS    FF/   FFF/
      UU/    UU/  CC/    CC/  SS/    SS/  FF/ FFF/
     UU/    UU/  CC/    CC/  SS/         FFFFF/
    UU/    UU/  CC/    CC/  SS/          FF/ FF\
   UU/    UU/  CC/    CC/  SS/    SS/  FF/   FF\
 UUUUUUUUU/    CCCCCCC/    SSSSSSS/   FF/     FF\

University of California at San Francisco, DOCK 4.0.1
DOCK 4.0.1 was released on May 17, 1998.

____________________Job_Information____________________
launch_time                    Mon Jun 28 01:26:39 1999
host_name                      octane.bio.uib.no
memory_limit                   255524864
working_directory              /disk2/people/kteigen/dock/PAH_27_6_99/dock
user_name                      kteigen

_________________General_Parameters_________________
flexible_ligand                no
orient_ligand                  yes
score_ligand                   yes
minimize_ligand                no
multiple_ligands               no
chemical_screen                no
parallel_jobs                  no
random_seed                    0

_____________Flexible_Ligand_Parameters_____________
anchor_search                  no
multiple_anchors               no
anchor_size                    0
peripheral_search              no
write_partial_structures       no
torsion_drive                  no
clash_overlap                  0
conformation_cutoff_factor     1
torsion_minimize               no
reminimize_layer_number        0
minimize_anchor                no
reminimize_anchor              no
reminimize_ligand              no
flexible_bond_maximum          <infinity>

_____________Orient_Ligand_Parameters_____________
match_receptor_sites           yes
random_search                  no
ligand_centers                 no
```

```
automated_matching              yes
maximum_orientations            500
write_orientations              yes
rank_orientations               yes
rank_orientation_total          30

____________________Match_Parameters____________________
nodes_minimum                   3
nodes_maximum                   10
distance_tolerance              0.25
distance_minimum                2
check_degeneracy                no
reflect_ligand                  no
critical_points                 no
multiple_points                 no
chemical_match                  no

___________________Scoring_Parameters___________________
intramolecular_score            no
intermolecular_score            yes
gridded_score                   yes
grid_version                    4
grid_points                     0
receptor_atom_grid_spacing      0
bump_filter                     yes
bump_maximum                    3
contact_score                   no
contact_cutoff_distance         0
contact_clash_overlap           0
contact_clash_penalty           0
chemical_score                  no
energy_score                    yes
energy_cutoff_distance          0
distance_dielectric             no
dielectric_factor               0
attractive_exponent             0
repulsive_exponent              0
atom_model                      u
vdw_scale                       1
electrostatic_scale             1
output_atom_scores              no
rmsd_score                      no
contact_maximum                 0
contact_size_penalty            0
chemical_maximum                0
chemical_size_penalty           0
energy_maximum                  0
energy_size_penalty             0
rmsd_maximum                    0
rmsd_size_penalty               0
rmsd_override                   0

_________________Minimization_Parameters_________________
contact_minimize                no
chemical_minimize               no
energy_minimize                 no
rmsd_minimize                   no
initial_translation             0
initial_rotation                0
initial_torsion                 0
maximum_iterations              0
```

```
contact_convergence             0
chemical_convergence            0
energy_convergence              0
rmsd_convergence                0
maximum_cycles                  0
cycle_convergence               0
contact_termination             0
chemical_termination            0
energy_termination              0
rmsd_termination                0

_____________Chemical_Screen_Parameters_____________
construct_screen                no
screen_ligands                  no
pharmacophore_screen            no
similarity_screen               no
fold_keys                       yes
dissimilarity_maximum           0.25
distance_begin                  0
distance_end                    0
distance_interval               0

_________________Parallel_Job_Parameters_____________
parallel_server                 no
server_name                     server
client_total                    0

_____________Multiple_Ligand_Parameters_____________
ligands_maximum                 1
initial_skip                    0
interval_skip                   0
heavy_atoms_minimum             0
heavy_atoms_maximum             <infinity>
rank_ligands                    no
rank_ligand_total               1
restart_interval                0

_____________________File_Input_____________________
ligand_atom_file                BH2.mol2
ligand_center_file              ligand_center.sph
receptor_site_file              ../site/clusters.sph
score_grid_prefix               ../grid/grid
receptor_atom_file              receptor.mol2
vdw_definition_file             ../site/my_vdw.defn
chemical_definition_file        /disk2/software/dock/parameter/chem.defn
chemical_match_file             /disk2/software/dock/parameter/chem_match.tbl
chemical_score_file             /disk2/software/dock/parameter/chem_score.tbl
chemical_screen_file            /disk2/software/dock/parameter/chem_screen.tbl
flex_definition_file            /disk2/software/dock/parameter/flex.defn
flex_drive_file                 /disk2/software/dock/parameter/flex_drive.tbl
quit_file                       docking.quit
dump_file                       docking.dump

____________________File_Output_____________________
ligand_out_file                 docking_out.mol2
ligand_contact_file             docking_cnt.mol2
ligand_chemical_file            docking_chm.mol2
ligand_energy_file              BH2_nrg.mol2
ligand_rmsd_file                docking_rmsd.mol2
info_file                       docking.info
restart_file                    docking.rst
```

```
Processing ****
Reading general grid info from ../grid/grid.bmp
Reading bump grid from ../grid/grid.bmp
Reading energy grids from ../grid/grid.nrg
  VDW grids use a 6-12 Lennard-Jones potential with an all atom model.
  Reading attractive VDW energy grid.
  Reading repulsive VDW energy grid.
  Reading electrostatic energy grid.

_____________________Docking_Results_____________________
Name         : ****
Description  : ****
Orientations tried                                :      17326
Orientations scored                               :       7000

Best intermolecular energy score                  :     -23.94
RMSD of best energy scorer (A)                    :      25.69

Elapsed cpu time (sec)                            :     223.24


_____________________Docking_Performance_____________________

Procedure timings                              time (s)  percent
Read                                               0.06        0
Screen                                             0.00        0
Orientation Search                                 5.88        3
Orientation Score                                217.32       97
Conformation Anchor                                0.02        0
Conformation Peripheral                            0.00        0
Other                                              0.22        0
Total                                            223.50      100

Minimizer usage                        minimum  average  maximum
Calls per molecule                        7000     7000     7000
Score improvement per call               0.032  8.2e+03  1.8e+05
Vertices per call                            6        6        6
Cycles per call                              1        1        1
Iterations per cycle                        18       89      541

Finished processing molecule in 223.498 seconds.
```

Appendix V

Van der Waals Definition File

This is part of the van der Waals definition file. The complete file contains data for all the atoms defined in the AMBER forcefield. Shown here are the definitions for three nitrogen atom types and Fe(III). Iron is not defined in the AMBER force field, and the vdw.defn file had to be edited to include iron.

```
AMBER-based Lennard-Jones parameters for van der Waals energy potential.

 The radius corresponds to half the equilibrium distance (in Ångstroms)
 between atoms.
 The well depth corresponds to the energy (in kcal/mol) at the equilibrium
 distance.
 Lennard-Jones parameters are derived as follows:

 sqrt(A) = sqrt [e * b/(a-b) * (2r)^a]
 sqrt(B) = sqrt [e * a/(a-b) * (2r)^b]

 where:
      r = radius                    a = repulsive exponent
      e = well depth                b = attractive exponent
_____________________________________
name         Iron_III
atom_model   either
radius       1.170
well_depth   0.100
heavy_flag   1
valence            0

definition   Fe
_____________________________________
name         Nitrogen
atom_model   either
radius       1.750
well_depth   0.160
heavy_flag   1
valence            4

definition   N
_____________________________________
name         Nitrogen_quaternary
atom_model   either
radius       1.850
well_depth   0.080
heavy_flag   1
valence            4

definition   N.4
_____________________________________
name         Nitrogen_sp3
atom_model   either
radius       1.850
well_depth   0.120
heavy_flag   1
valence            3

definition   N.3
```

Appendix VI

NMR Refine

This is a manual on how to run distance geometry calculations with DGII under InsightII to generate structures compatible with user defined restraints and constraints.

```
Start out with optimized structure(s).
Define intra- and inter molecular distances togehter with chirality
constraints in restraint-file.  The file is made using jot or xedit.
NB!!  Spaces in your atom name definition must be replaced by an ”_”
      eks.:  BH2ASS:BIO1 3   must be written    BH2ASS:BIO1_3
The file must have the extension .rstrnt

example of a file:

!BIOSYM restraint 1
!
#chiral
BH2ASS:BIO1_3:C9         R
BH2ASS:BIO1_3:C10  S
!
#distance
BH2ASS:BIO1_3:H11* BH2ASS:BIO1_3:H10  2.16  4.16  1.000 1.000 1000.0
BH2ASS:BIO1_3:H11* BH2ASS:BIO1_3:H9   3.16  5.16  1.000 1.000 1000.0
BH2ASS:BIO1_3:H11* BH2ASS:BIO1_3:H7*  3.74  5.74  1.000 1.000 1000.0
BH2ASS:BIO1_3:H7*  BH2ASS:BIO1_3:H10  2.57  4.57  1.000 1.000 1000.0
BH2ASS:BIO1_3:H7*  BH2ASS:BIO1_3:H9   3.04  5.04  1.000 1.000 1000.0
BH2ASS:BIO1_3:H7*  BH2ASS:BIO1_3:H8   2.47  4.47  1.000 1.000 1000.0
BH2ASS:OH61_1:FE6 BH2ASS:BIO1_3:H7*  3.87  4.87  1.000 1.000 1000.0
BH2ASS:OH61_1:FE6  BH2ASS:BIO1_3:H9   4.86  5.86  1.000 1.000 1000.0
BH2ASS:OH61_1:FE6  BH2ASS:BIO1_3:H11* 5.44  6.44  1.000 1.000 1000.0


From the IsightII shell choose the NMR_Refine module.
Read in the molecules of interest, preferably in mol2 format.

To check the molecule, go to the Builder module, choose Modify from the
pulldown menu.  Then choose Hydrogens.  In the pop up menu, click set_pH,
and Capping Mode Charged (unless you have a good reason not to).
If there is a problem with the number of Hydrogens added to the structure,
start by defining single and double bonds.  This is done from the Modify
menu. You can also modify hybridisation and potensials of atoms from the Atom
pull down menu.

From the NMR_DB menu, click on Files.
Select Get, give a name to the project, the name of the molecule on screen
should appear in the NMR_Molecule_Name.  Give the complete localization of
the Restraint_File you have defined for your molecule.
De-select Peaks.
Click on Execute.  Check the output from InsightII that it has accepted all
the restraints given in the restraint file!

You are now ready to run DGII.
From the DGII pull-down, select DGII_Run.
Start with only smooth selected (not Embed or Optimize)
```

Toggle Triangle_Smooth on. Sequential Tetrangle_Strategy on, the values of 0.1 for Convergence_Crit and 1 for Max_Passes are OK.
In the Global_Setup, choose Detail as Output_Level, give a Project_Description. DG_Num_Structures somewhere between 30 and 50. Toggle Incremnt_Files on.Select Execute.
If you get the status 0 as output, your calculation has not run into any problems.

You are now ready to run Embed of the DGII procedure.
Untoggle Smooth, and toggle Embed on in the DGII_Run menu. Select Prospective Metrization and 4 Embed_Dimensions, you can also try Majorize. Deselect Ovrwrt_Emd_Structs. Select Execute.

The last procedure in the DG calculation is to do Optimization.
Toggle Smooth and Embed off, Optimize on.
Deselect Calc_Init_Engy. The energy value to input should be equal to the number of atoms in your specific structure. Select Execute.

The structures generated are written to an arc-file.
'Given_project_name'.arc