Journal of Molecular Graphics and Modelling 29 (2010) 396–405

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Tetrapyrrole binding affinity of the murine and human p22HBP heme-binding proteins

Nuno M. Micaelo[a,1], Anjos L. Macedo[b], Brian J. Goodfellow[c], Vítor Félix[a,*]

[a] Departamento de Química, CICECO and Secção Autónoma de Ciências da Saúde, Universidade de Aveiro, 3810-193 Aveiro, Portugal
[b] REQUIMTE, Departamento de Quìmica, FCT-UNL, 2829-516 Caparica, Portugal
[c] Departamento de Quìmica, CICECO, Universidade de Aveiro, 3810-193 Aveiro, Portugal

ARTICLE INFO

Article history:
Received 9 March 2010
Received in revised form 1 July 2010
Accepted 26 July 2010
Available online 6 August 2010

Keywords:
Heme-binding protein
Tetrapyrrole
Docking
Molecular dynamics
Binding constant

ABSTRACT

We present the first systematic molecular modeling study of the binding properties of murine (mHBP) and human (hHBP) p22HBP protein (heme-binding protein) with four tetrapyrrole ring systems belonging to the heme biosynthetic pathway: iron protoporphyrin IX (HEMIN), protoporphyrin IX (PPIX), coproporphyrin III (CPIII), coproporphyrin I (CPI). The relative binding affinities predicted by our computational study were found to be similar to those observed experimentally, providing a first rational structural analysis of the molecular recognition mechanism, by p22HBP, toward a number of different tetrapyrrole ligands. To probe the structure of these p22HBP protein complexes, docking, molecular dynamics and MM–PBSA methodologies supported by experimental NMR ring current shift data have been employed. The tetrapyrroles studied were found to bind murine p22HBP with the following binding affinity order: HEMIN > PPIX > CPIII > CPI, which ranged from −22.2 to −6.1 kcal/mol. In general, the protein–tetrapyrrole complexes are stabilized by non-bonded interactions between the tetrapyrrole propionate groups and basic residues of the protein, and by the preferential solvation of the complex compared to the unbound components.

1. Introduction

The biosynthesis of heme takes place mainly in the mitochondria of erythroid cells (~85%) and in hepatocytes, although heme can be biosynthesized in virtually all tissues. The role of heme in many biological processes is well known and most heme is used in heme-containing proteins. However, due to the poor solubility and reactivity of heme in aqueous solution under physiological conditions, the levels of free heme in cells must be tightly controlled. In order to regulate this uncommitted heme pool, it has been hypothesized that intracellular heme-binding proteins must exist to act as chaperones or as a buffer during induced heme synthesis.

In 1998, p22HBP, a 22 kDa protein, was first purified from mouse liver cell extracts and characterized as a cytosolic heme-binding protein by Taketani et al. [1]. Subsequently, Blackmon et al. [2] determined that other tetrapyrroles, in addition to hemin, can bind to p22HBP although its functional role in the cell remains unknown. Recent studies have suggested possible roles for p22HBP in cell death and infection *via* the discovery that an acetylated N-terminal fragment of p22HBP (residues 1–21) has potent chemoattractant activity [3]. Furthermore, a proteomic study, of hemoglobin biosynthesis demonstrated that p22HBP is a component in one of four identified multi-protein complexes in murine erythroleukemia cells induced to undergo differentiation [4].

SOUL and p22HBP are part of an evolutionarily conserved heme-binding protein family. The SOUL protein is expressed in retina and pineal gland in the domestic chicken and solely in the retina in the murine form [5]. Murine SOUL (mSOUL) has 27% sequence identity to murine p22HBP (mHBP) and also binds heme. A recent study of murine SOUL indicated that this protein is a dimer in its apo form and becomes hexameric when bound to heme. The heme is thought to bind *via* coordination of the Fe^{III} heme to a histidine side chain [6]. Recently SOUL has been identified as a novel member of the BH3 domain-only protein family that cannot induce cell death alone but can facilitate both outer and inner mitochondrial membrane permeabilization and predominantly necrotic cell death in oxidative stress [7].

In 2006, our group and collaborators determined the first structure of a protein from the SOUL/HBP family, mHBP, using NMR spectroscopy [8,9]. The mHBP structure was found to be a novel fold in eukaryotes composed of a 9-stranded twisted beta-barrel flanked by two alpha helices. Dissociation constants for the mHBP–HEMIN and mHBP–PPIX complexes were determined by fluorescence quenching to be in the low nanomolar range. Chemi-

* Corresponding author.
E-mail address: vitor.felix@ua.pt (V. Félix).
[1] Current address: Centro de Quìmica, Universidade do Minho, Campus de Gualtar, 4710-057 Braga, Portugal.

1093-3263/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2010.07.008

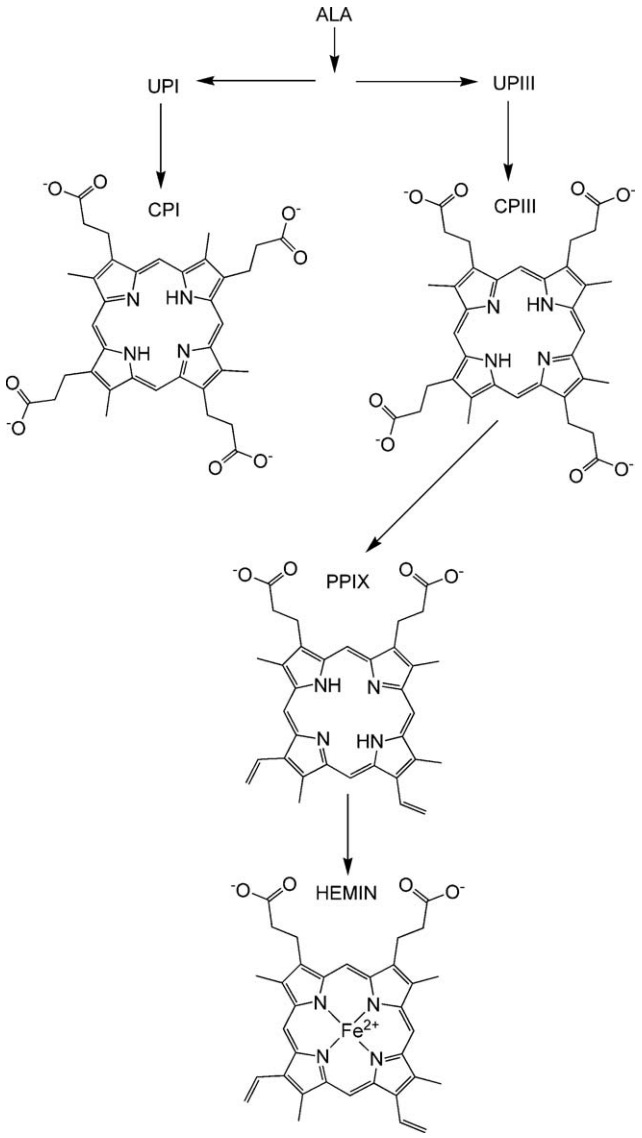

Fig. 1. Precursors and intermediate products of the heme biosynthetic pathway. Iron protoporphyrin IX (HEMIN), protoporphyrin IX (PPIX), coproporphyrin III (CPIII), coproporphyrin I (CPI), uroporphyrin III (UPIII), uroporphyrin I (UPI) and δ-aminolevulinate (ALA).

cal shift perturbations arising from the addition of HEMIN and PPIX were mapped to the mHBP structure allowing the binding site to be determined. From sequence alignments and homology modeling, it was concluded that mHBP and SOUL should share a conserved tertiary fold, although they most probably bind heme at different sites within this fold.

As no structure exists for the bound form of mHBP, molecular modeling has been used to extend our knowledge of the tetrapyrrole binding properties of this protein and also of the human homologue (hHBP). The structure of hHBP has not been determined, although the high sequence identity (86%) to mHBP suggests the structure should be very similar. The physiological role of m/hHBP is not yet clear although the protein has been localized in the cytoplasm. Possible binding modes to a number of tetrapyrroles belonging to the heme biosynthetic pathway (Fig. 1) including iron protoporphyrin IX (HEMIN), protoporphyrin IX (PPIX), coproporphyrin III (CPIII), coproporphyrin I (CPI), have been studied and the experimental binding affinities towards HBP have been determined.

This work aims to provide insights into modes and affinity of, tetrapyrrole binding to this class of proteins using molecular modeling methodologies. This will be accomplished by docking the tetrapyrroles to mHBP and hHBP followed by unconstrained molecular mechanics/dynamics (MM/MD) simulations in water. The binding affinity and the protein-tetrapyrrole molecular interactions have been analyzed and compared with available experimental data, especially NMR ring current shift (RCS) differences [8] in the case of PPIX, and published dissociation constants [1,2,6,8]. The latter can be compared with those found from the binding free energies of the mHBP and hHBP estimated by molecular mechanics–Poisson Bolzmann surface area (MM–PBSA) calculations. The role of a mobile loop near the protein-binding site in binding and release of the tetrapyrrole rings is also discussed. This modeling study will provide insights into the, currently unknown, mechanism of binding and will attempt to identify key residues involved in molecular recognition and stabilization of the complexes between tetrapyrroles and HBP.

2. Computational methods

2.1. Tetrapyrrole force field parameters

Four tetrapyrrole ring systems belonging to the heme biosynthetic pathway; HEMIN, PPIX, CPIII, CPI (Supplementary Material Fig. S1) have been parameterized [10]. The HEME type B ring found in the X-ray single crystal structure (1.25 Å resolution) of human hemoglobin (PDB 2DN2 [11]) was used as a structural base from which the tetrapyrrole derivatives were constructed using the molecular visualization program PyMOL [12]. Geometry optimization of the all-atom model of each tetrapyrrole was performed at the B3LYP level with GAUSSIAN03, using the 6-31G(d) basis set for organic atoms (C, H, N and O) and the 6-31G(3df) set for the iron(II) in HEMIN [13]. Partial atomic charges (Supplementary Material, Table S1) were obtained using the RESP [14] fitting method, based on electrostatic potentials calculated at the Hartree Fock level with the 6-31G(d) basis set in GAUSSIAN03 [13]. A two-step RESP fitting procedure was used in order to build a united atom model in agreement with the GROMOS 53A6 [15] force field. Initially, for the propionate groups, the charges of the abstracted hydrogens (a process used in GROMOS to collapse an hydrogen to its attached aliphatic carbon) were given zero partial charge and the charges of the oxygen atoms were equalized. Bonded and non-bonded parameters for the tetrapyrroles, excluding partial atomic charges, were derived from the parameterized model of HEMIN found in the GROMOS FF [13]. Lennard–Jones parameters for the iron atom in HEMIN were taken from the universal force field [16].

2.2. Protein set-up

The structure of the apo form of mHBP has recently been solved by NMR and is deposited in the PDB with code 2GOV [8,9]. Among the 20 NMR conformers deposited, the ninth listed was chosen since it has the lowest-energy in the GROMOS force field. A theoretical structural model, using the mHBP NMR structure as a template, of the hHBP protein (NCBI Entrez Protein accession number: AAD32098) was obtained using MODELLER [17]. The human and murine proteins share 86% sequence similarity. 20 models were generated using an initial alignment between the murine and human protein sequences carried out with MODELLER. The model with the lowest objective function [17] was chosen and its quality was evaluated by Procheck based on the stereochemistry of the generated model [18]. A high quality model of the hHBP protein was obtained with no residues in disallowed regions of the Ramachandran plot. The protonation states of the acidic and basic residues

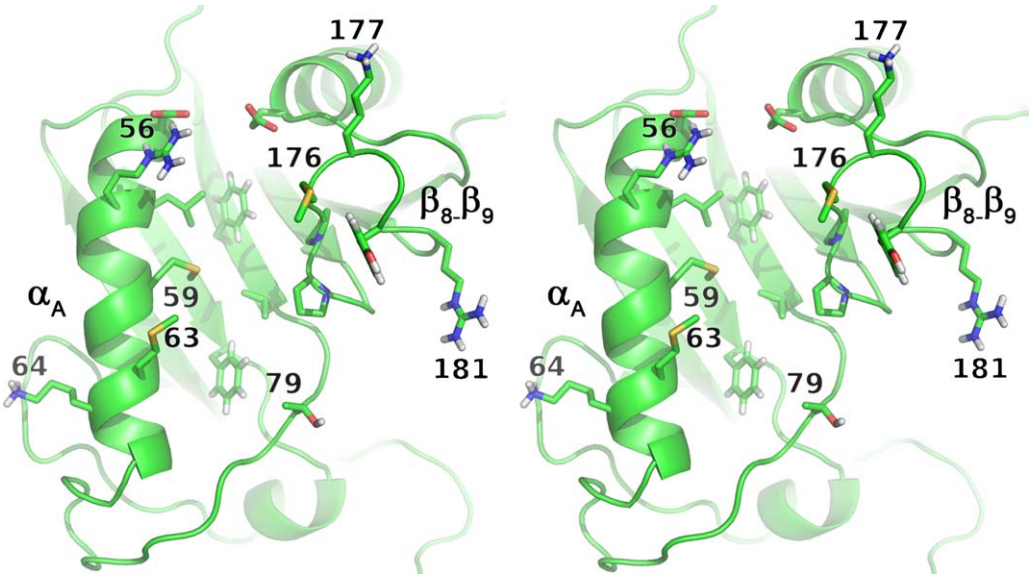

Fig. 2. Cross-eye stereo view of the apo mHBP protein-binding site. Selected side chain residues of the binding pocket are rendered in sticks: arginine 56, methionine 59 and 63, lysine 64, threonine 79, methionine 176, lysine 177 and arginine 181.

were set to their standard states found in aqueous solution at pH 8. All acidic residues of the protein were fully deprotonated, and all basic residues were fully protonated. The protonation state of the sole histidine (only present in the murine form) was considered to be neutral with the proton at the Nδ atom. The N-terminus was protonated and the C-terminus deprotonated for both proteins.

2.3. Tetrapyrrole docking

The docking study of the different tetrapyrroles with mHBP and hHBP was carried out with AutoDock 4 [19]. Even though experimental binding data is only available for HEMIN, PPIX, and CPIII binding to mHBP, experiments with CPI were also performed. As experimental binding data is only available for hHBP with HEMIN and PPIX and the structure of this protein was not determined experimentally, docking experiments were only carried out for hHBP with these two tetrapyrroles. The tetrapyrrole and protein structures, as well as their atomic charges were identical to those described above. All possible rotatable bonds were set as active for the tetrapyrroles, resulting in 8 rotatable bonds in HEMIN and PPIX, and 12 in CPIII and CPI. Three residues potentially involved in tetrapyrrole binding: arginine 56, methionine 176 (position 175 in human sequence) and lysine 177 (position 176 in human sequence) were located near the experimentally determined protein-binding pocket [8] (Fig. 2). As these residues are at the protein surface they will experience higher conformational freedom. Also, the side chain conformations of these residues found in the NMR structures are for the apo form and, as such, their positions may not reflect the orientations found in the bound form. Arginine 56 and lysine 177/176 are positive residues that can interact with the negative propionate groups of the tetrapyrroles and were selected to be flexible in order to explore electrostatic interactions with the tetrapyrrole ligands during the docking process. Methionine 176/175 was also set as flexible due to its possible role in the coordination of the iron in HEMIN. All possible rotatable bonds were set as active with exception of the C–N amide bond of arginine 56. The grid for probe–target energy calculations was placed with its centre at the binding pocket previously identified by NMR [8]. The grid was built along the cleft bordered by the α_A helix and the β_8–β_9 loop (Fig. 2) and contained $60 \times 48 \times 46$ grid points with 0.375 Å spacing. Here it has been assumed that other binding sites do not exist, a reasonable assumption considering the NMR chemical shift mapping results [8]. A more comprehensive docking search would be possible but it is outside of the scope of this current work. For each ligand 150 runs using a Lamarckian genetic algorithm with 150 individuals in each population were carried out. For each run, the maximum number of generations was set to 27×10^3 and the maximum number of energy evaluations to 25×10^5. The resulting docking solutions were clustered using AutoDock with a rtol cut-off of 1 Å. The lowest-energy protein–tetrapyrrole complex from the lowest-energy cluster was selected for subsequent molecular mechanics (MM) and molecular dynamics (MD) studies.

2.4. Simulation set-up

MM/MD simulations were performed with the GROMACS 3.3.2 package [20,21] using the GROMOS 53A6 force field [15]. Each lowest-energy protein–tetrapyrrole complex obtained from the docking calculation, the apo form of the protein and the tetrapyrroles were hydrated in a pre-equilibrated truncated octahedral box of SPC waters at 300 K. The system size was chosen according to the minimum image convention taking into account a cut-off of 14 Å. A system with zero net charge was obtained by replacement of solvent molecules by equivalent number of Na^+ atoms. Protein and tetrapyrrole bond lengths were constrained with LINCS [22] and with SETTLE for water [23]. Non-bonded interactions were calculated using a twin-range method with short and long range cut-offs of 8 Å and 14 Å, respectively. Neighbor searching was carried out up to 14 Å and updated every five steps. A time step of integration of 2 fs was used. A reaction field correction for electrostatic interactions was applied using a dielectric constant of 54 [24]. The initial systems were energy minimized for 5000 steps using the steepest descent method with all protein heavy atoms harmonically restrained using a force constant of 10^6 kJ/mol nm^2. Subsequently, 5000 steps of energy minimization with positional restrains applied to the Cα protein atoms were carried out followed by 5000 steps of unrestrained full system energy minimization. The different protein–tetrapyrrole complexes were initialized in the canonical ensemble (NVT) for 100 ps with all protein heavy atoms harmonically restrained using a force constant of 10^6 kJ/mol nm^2, followed by 100 ps with positional restrains applied to the protein Cα atoms. The simulation was continued for

10 ns in the isothermal–isobaric ensemble (NPT) to ensure the full equilibration of all system properties. Pressure control was implemented using the Berendsen barostat [25] with a reference pressure of 1 atm, 0.5 ps relaxation time, and isothermal compressibility of 4.5×10^{-5} bar^{-1}. Temperature control was set using the Berendsen thermostat [25] at 300 K. The protein, tetrapyrrole and ions were grouped in the same heat bath and solvent molecules in separate heat bath with temperature coupling constants of 0.01 ps in the initialization phase and with 0.1 ps in the equilibration phase. Five replica simulations of 10 ns length were carried out using different initial velocities taken from a Maxwell–Boltzmann distribution at 300 K leading to a total simulation time of 50 ns for each system. The conformations were recorded at a frequency of 1/ps, leading to a trajectory files containing 10,000 frames. Properties were calculated from the last 5 ns of each simulation run.

2.5. Analysis of MD simulations

Several detailed analyses were carried out for the trajectories obtained from the last 5 ns of the equilibrated simulations. Tetrapyrrole binding free energy was estimated from five independent MD simulations from the complexes, free protein and tetrapyrrole using 5000 conformations of each system (configurations from the last 5 ns of a 10 ns simulation run), by the MM–PBSA method according to Gilson and Zhou [26] MM–PBSA ΔG_{bind} was estimated according to:

$$\Delta G_{\text{bind}} = \Delta E_{\text{MM}} + \Delta G_{\text{solv}} + \Delta G_{\text{np}} - T\Delta S \quad (1)$$

The "MM" subscript corresponds to the internal energy of the protein or ligand (bond, angle, dihedral, the electrostatic and the van der Waals interactions), the "solv" subscript refers to the free energy of polar solvation, and the "np" subscript the free energy of nonpolar solvation. In summary, ΔE_{MM} refers to the molecular mechanics internal energy, ΔG_{sol} the polar solvation free energy, ΔG_{np} nonpolar solvation free energy, ΔS is the entropic contribution for the binding free energy and T is the temperature. Each term in Eq. (1) is given by:

$$\Delta E_{\text{MM}} = E_{\text{MM}}^{\text{complex}} - (E_{\text{MM}}^{\text{protein}} + E_{\text{MM}}^{\text{lig}})$$

$$\Delta G_{\text{solv}} = G_{\text{solv}}^{\text{complex}} - (G_{\text{solv}}^{\text{protein}} + G_{\text{solv}}^{\text{lig}})$$

$$\Delta G_{\text{np}} = G_{\text{np}}^{\text{complex}} - (G_{\text{np}}^{\text{protein}} + G_{\text{np}}^{\text{lig}})$$

$$\Delta S = S^{\text{complex}} - (S^{\text{protein}} + S^{\text{lig}})$$

The electrostatic potential and solvation free energy (ΔG_{solv}) calculations were carried out using the Poisson–Boltzmann method as implemented in the MEAD package version 2.2.0 [27]. The atomic charges for the protein were taken from the GROMOS96 [15] force field and those for the tetrapyrroles as described above. Atomic radii equal to half of R_{min} were used [28–30]. This distance corresponds to the minimum Lennard–Jones energy computed using a pair of identical atoms types taken from the GROMOS96 [13] force field (radii of the tetrapyrroles and proteins are given as Supplementary Material in Tables S2 and S3). A grid of 81 points with 1 Å spacing and a focusing grid of 81 points with 0.5 Å spacing were used to enclose the whole system. A dielectric constant of 80 and 4 was used for the solvent and the solute interior respectively [28]. An ionic strength of 100 mM was used. The nonpolar solvation energy was determined as detailed elsewhere [31] from the solvent accessible surface calculated with a probe radius of 1.4 Å and constant of 5 cal/mol Å^2. The entropy was calculated using the quasiharmonic approximation [32] from the full trajectories. Cluster analysis of the MM/MD trajectories was carried out using the single-linkage method with a root mean square deviation (RMSD) cut-off of 1 Å. The selection of the representative conformer for the PPIX mHBP complex was carried out by comparison with experimental NMR chemical shifts. The largest component to the chemical shifts of the protein protons on complexation with PPIX is from the ring current shifts (RCS) of the aromatic PPIX. The RMSD between the experimental and theoretical RCS were calculated as follows:

$$\text{RMSD} = \sqrt{[(\text{RCS}_{\text{complex}}^{\text{experimental}} - \text{RCS}_{\text{protein}}^{\text{experimental}}) - (\text{RCS}_{\text{complex}}^{\text{theoretical}} - \text{RCS}_{\text{protein}}^{\text{theoretical}})]^2} \quad (2)$$

where the theoretical RCS of each protein and complex conformation were calculated with SHIFTS-4.1 [33]. The conformer with the lowest RMSD was chosen as being the representative complex. Due to the absence of experimental RCS for the remaining protein–tetrapyrrole complexes, the representative conformation of each protein complex with HEMIN, CPIII and CPI was that closest to the average structure of the MD trajectories.

3. Results and discussion

3.1. Tetrapyrrole–mHBP/hHBP docking

The lowest-energy protein–tetrapyrrole complexes obtained from the docking experiments reveal that the tetrapyrrole derivatives are clearly assembled into the protein-binding pocket. HEMIN and PPIX were found to have identical binding orientations for the lowest-energy conformation, indicating that stabilization of the propionate side chains is mainly achieved by electrostatic interactions with lysine 177 (position 176 in human) and arginine 56 located at the edges of the protein-binding pocket. The disposition of these residues in the binding pocket is outlined in Fig. 2. This binding mode for HEMIN and PPIX is also observed in the hHBP complexes. The binding pocket of the human homologue is almost identical to the murine protein. The sequentially (and structurally) conserved lysine and arginine residues seem to play an identical role in stabilizing the tetrapyrrole propionate side chains in both protein complexes.

CPIII and CPI lowest-energy binding modes with mHBP are comparable to those found for the HEMIN and PPIX complexes. No experimental binding data is available for CPI, and consequently these theoretical findings are the first pieces evidence for a possible molecular association between this tetrapyrrole and mHBP.

3.2. Tetrapyrrole–mHBP/hHBP MM/MD and binding affinity

The protein–tetrapyrrole complexes obtained from docking experiments were subjected to 50 ns of MM/MD simulation, allowing the system to fully reorganize in order to sample the representative protein–tetrapyrrole complex conformational space. In general, all MD simulations show that HEMIN, PPIX, CPIII, CPI are located in the mHBP protein-binding pocket for the entire duration of the MD simulation.

The free energies of binding (ΔG_{bind}) calculated using MM–PBSA [26] for the different tetrapyrroles with mHBP were found to be (from highest to lowest affinity): −22.2, −15.3, −9.8 and −6.1 kcal/mol for HEMIN, PPIX, CPIII, CPI, respectively(Table 1). A straight comparison in absolute terms between experimental and theoretical values is not possible, however we can compare relative binding affinities for the different tetrapyrroles. Here our theoretical values are consistent with published experimental data (Table 1): HEMIN is reported experimentally to bind more strongly to mHBP [1,2,8] and hHBP [2] and, we find it has the highest free energy of binding −14.6 kcal/mol. It should be noted that, in our modeling study the iron atom in HEMIN is not coordinated by any residue. This is consistent with experimental EPR data [2].

Table 1
MM–PBSA contributions to the binding free energy (ΔG_{bind}) of the protein–tetrapyrrole complexes with mHBP and hHBP compared with experimental data.

Ligand	Theoretical ΔG_{bind}							Experimental ΔG_{bind}			
	ΔE_{MM}	ΔG_{np}	ΔG_{solv}	ΔG_{tot}	$-T\Delta S$	ΔG_{bind}		Ref. [7]	Ref. [8]	Ref. [1]	Ref. [2]
HEMIN[a]	18.4	−1.1	−39.3	−22.0	−0.2	−22.2	(2.82)	−14.6	−11.8	−9.6	−8.2
PPIX[a]	3.7	−1.8	−17.0	−15.1	−0.2	−15.3	(2.88)		−12.9	−9.7	−6.7
CPIII[a]	1.9	−1.0	−12.5	−11.6	1.9	−9.8	(2.40)			−9.0	
CPI[a]	1.5	−0.8	−8.4	−7.7	1.6	−6.1	(2.50)				
HEMIN[b]	4.2	−0.9	−20.2	−16.9	−0.6	−17.5	(2.93)				−6.4
PPIX[b]	−4.6	−1.0	−11.5	−17.1	−0.5	−17.6	(2.70)				−6.2

ΔE_{MM}, molecular mechanics internal energy; ΔG_{sol}, polar solvation free energy; ΔG_{np} nonpolar solvation free energy; $-T\Delta S$, entropy; $\Delta G_{tot} = \Delta G_{solv} + \Delta G_{np} + \Delta E_{MM}$; $\Delta G_{bind} = \Delta G_{tot} - T\Delta S$. ΔG_{bind} standard error is given in parentheses and was calculated by error propagation from each term of Eq. (1). See Section 2 for the details of each contribution calculation. Experimental K_d were converted to ΔG_{bind} through the $\Delta G_{bind} = -RT \times \ln(1/K_d)$.

[a] Tetrapyrrole binding to mHBP.
[b] Tetrapyrrole binding to hHBP. Energy units are in kcal/mol.

The experimental binding data shown in Table 1 also indicates that PPIX has very similar binding to HEMIM again suggesting that there is no iron coordination. The discrepancies seen between different experimental studies of tetrapyrrole binding (Table 1) are most probably due to the different experimental conditions and techniques used. For example, Blackmon et al. [2] used high micromolar concentrations of p22HBP protein whereas Dias et al. [8] used low nanomolar protein concentrations. Our MM–PBSA binding free energies are overestimated relative to the experimental values, which is not unexpected when this methodology is applied

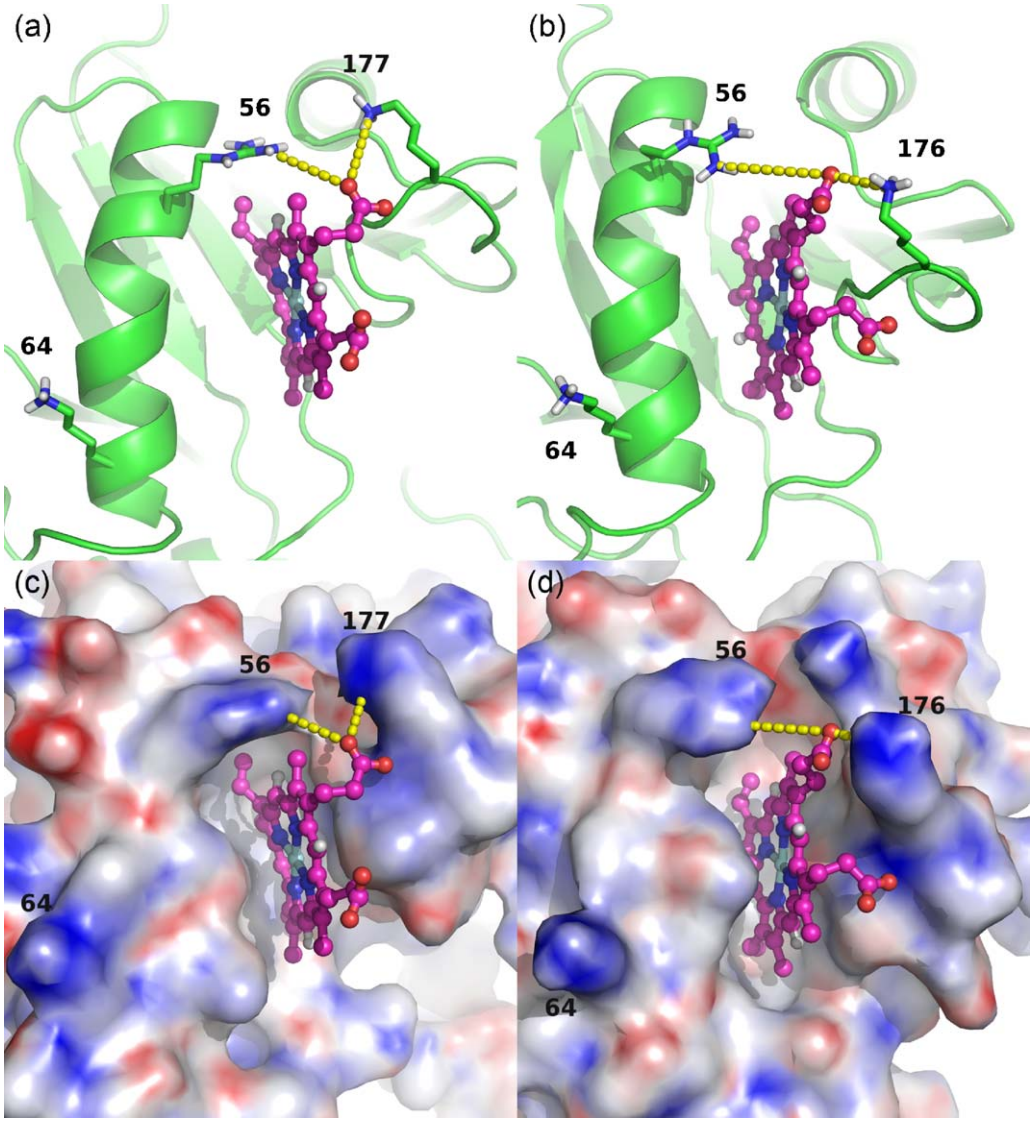

Fig. 3. Representative structure of the HEMIN mHBP, (a) and (c), and HEMIN hHBP, (b) and (d) complexes. The binding site of each complex is shown with HEMIN rendered in ball and stick. The protein is rendered in cartoon in (a) and (b). Key side chain residues are rendered in sticks (see Fig. 2 for residue names). The corresponding electrostatic potential mapped on the protein molecular surface from $-5k_bT$ (red) to $+5\ k_bT$ (blue) is also shown for each complex in (c) and (d).

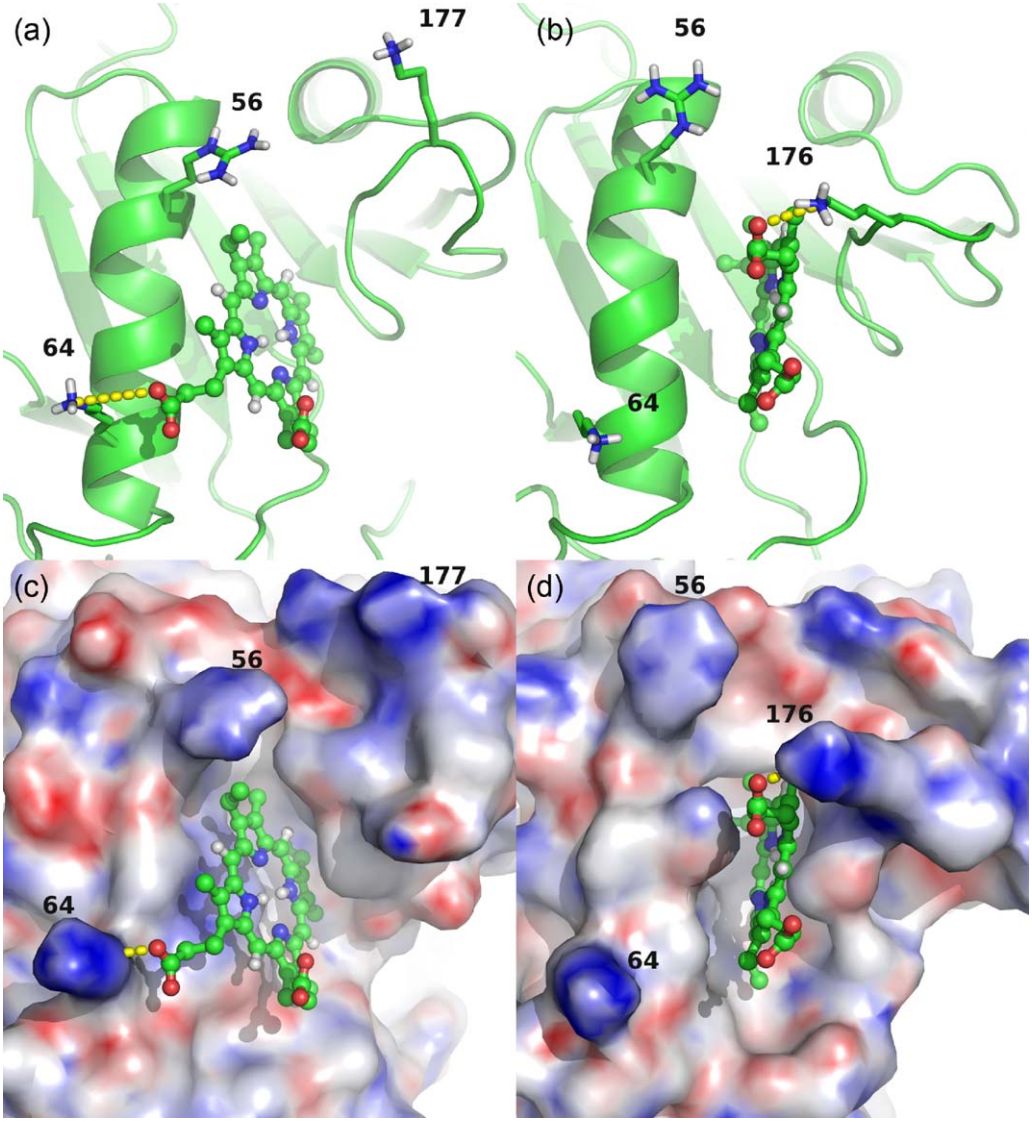

Fig. 4. Representative structure of the PPIX mHBP, (a) and (c), and PPIX hHBP (b) and (d) complexes. The binding site of each complex is shown with PPIX rendered in ball and stick. The protein is rendered in cartoon in (a) and (b). Key side chain residues are rendered in sticks (see Fig. 2 for residue names). The corresponding electrostatic potential mapped on the protein molecular surface from $-5k_bT$ (red) to $+5\ k_bT$ (blue) is also shown for each complex in (c) and (d). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

[26]. A better agreement between experiment and theory could be in principle obtained by choosing other starting structures as well as by the use of extended sampling time rather than the five replicas of 10 ns (total simulation time 50 ns) used here. Furthermore, our molecular dynamics simulations do not take into account the changes in protonation of the tetrapyrroles upon binding and only allow limited non-planar deformations. However, the methodology and the level of approximation followed are sufficiently accurate to obtain relative binding affinities for the proteins with different tetrapyrrole ligands. Indeed, for CPIII and CPI, the data from Taketani et al. [1] suggest that CPIII has lower affinity than HEMIN and PPIX, which is in agreement with our calculations. We also find that CPI binds the weakest to m22HBP, a result that needs to be confirmed experimentally. For hHBP, the theoretical binding affinity of HEMIN and PPIX towards this protein is −17.5 and −17.6 kcal/mol, respectively, which is again an overestimate relative to the currently available experimental data [2]. Overall, however, our theoretical calculations agree with the experimental data and predict similar binding affinities of HEMIN and PPIX towards mHBP.

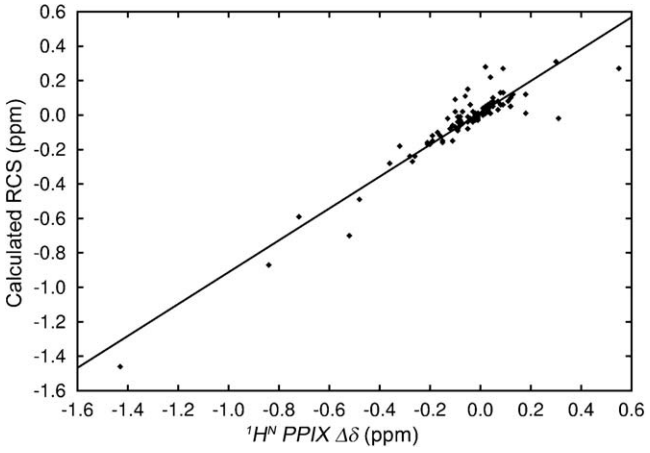

Fig. 5. Calculated versus experimental RCS differences for the PPIX mHBP protein complex. Best fit is given by: $y = 0.92x + 0.013$, $R^2 = 0.88$.

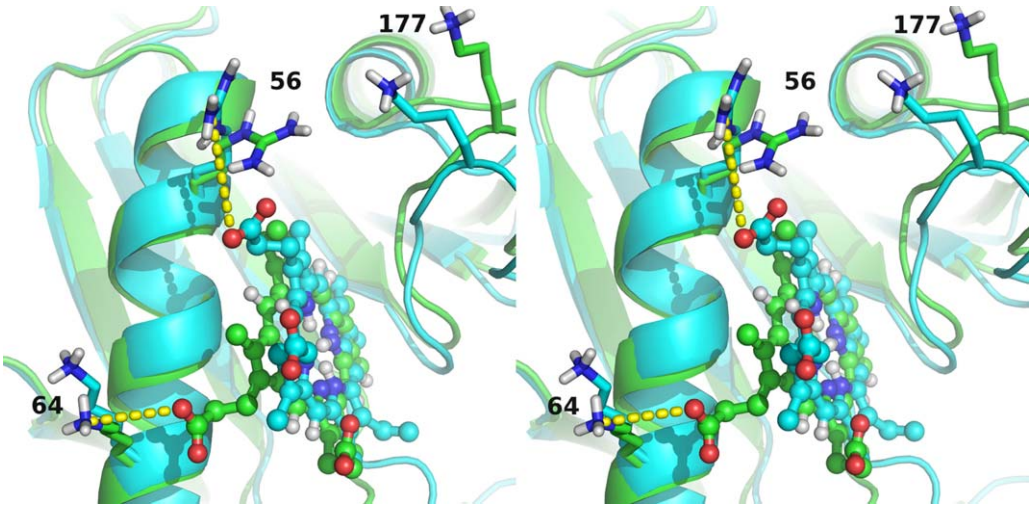

Fig. 6. Cross-eye stereo view of the mHBP protein-binding site with the two representative binding conformations for PPIX. The two conformations are colored in blue and green. The overall protein is rendered in cartoon, and key side chain residues are rendered in sticks (see Fig. 2 for residue names). PPIX is rendered in ball and stick. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

The analysis of the MM–PBSA calculations indicates that the favorable solvation free energy (ΔG_{solv} and ΔG_{np}) of the protein–tetrapyrrole complex relative to the free protein and ligand is the driving force that promotes and stabilizes the complexes (Table 1). In the paper describing the structure determination of mHBP by NMR [8] the interaction between PPIX and the protein was discussed in terms of an hydrophobic pocket. The results presented here confirm that the interaction is essentially hydrophobic in that, the poor free energy of solvation for free PPIX and for the free protein (with a hydrophobic patch) is outweighed in the complex where the solvation free energy is more negative due to PPIX being positioned on top of the hydrophobic patch. The same interpretation can be applied to the HEMIN interaction with the murine and human protein: the binding seems to be driven by the preferential solvation of the complex relative to the free protein and tetrapyrrole. Furthermore, it is also evident that this balance is also favorable, but to a lesser degree, in the case of CPIII and CPI binding to mHBP.

3.3. *HEMIN–mHBP/hHBP complex*

The formation of the HEMIN mHBP/hHBP complex is important for its putative role as a cellular buffer [1,2,8]. We have confirmed that HEMIN is strongly bound to mHBP and propose for the first time a model of the complex of HEMIN with hHBP. The representative murine and human HEMIN protein complexes are presented in Fig. 3 and they show several key interactions. The complexes shown are those closest to the average structure from each MD trajectory, that is, the structure that yields the lowest root mean square deviation variance when all other conformations are fitted against it.

The HEMIN tetrapyrrole ring is clearly buried inside the protein, filling the hydrophobic protein-binding pocket, and is flanked by the α_A helix and the β_8–β_9 loop. Iron coordination by mHBP has been discussed in the literature and it has been shown by EPR that mHBP HEMIN binding does not involve iron coordination. The only possible residues in the vicinity of the binding pocket that could be involved in HEME iron coordination are methionines 59, 63 and 176. Methionine coordination has been seen in other heme systems such as the bacterioferritins of *Azotobacter vinelandii* [34], *Desulfovibrio desulfuricans* [35] and *Escherichia coli*. [36] Our results show that methionine 176, located in the β_8–β_9 loop is not found near the HEMIN iron and thus does not seem to be involved in coordination. On the other hand, the side chain of methionine 59 and 63 are found in the opposite site of HEMIN but not in a conformation that could indicate a clear HEMIN iron coordination. Furthermore, these methionines are not conserved among p22HBP homologs, human lacks the methionine in position 63, xenopus and zebrafish lacks both methionines at 59 and 63 [8]. Instead, the binding of HEMIN seems to be stabilized by lysine 177 (position 176 in human) and arginine 56 *via* a strong electrostatic interactions with one of the propionate groups in both the murine and human complexes (Fig. 3a and b). This can be seen *via* an electrostatic potential mapped on the molecular surface of the murine and human protein complexes (Fig. 3c and d). It can be seen that HEMIN is buried inside the hydrophobic binding pocket and the exposed propionates are stabilized by strong electrostatic interactions with neighboring positively charged regions. These strong interactions add to the hydrophobic interactions described above and provide a molecular explanation for the observed strong affinity of mHBP for HEMIN.

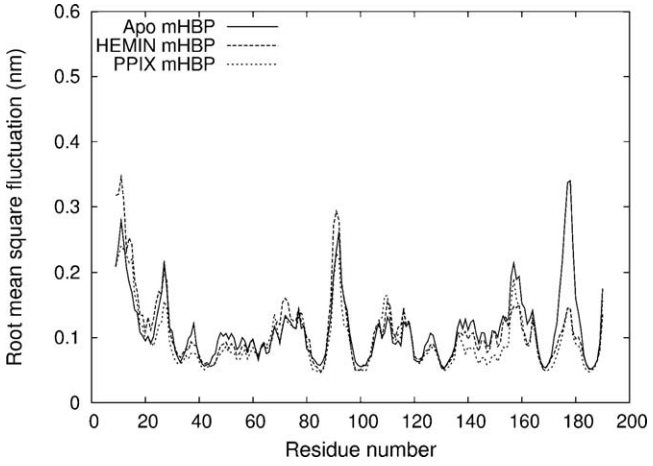

Fig. 7. Overlay of the mHBP protein Cα root mean square fluctuation (averaged over the five replicate simulations for each system) for the apo protein (solid line), HEMIN mHBP complex (long dashed line) and the PPIX mHBP complex (short dashed line).

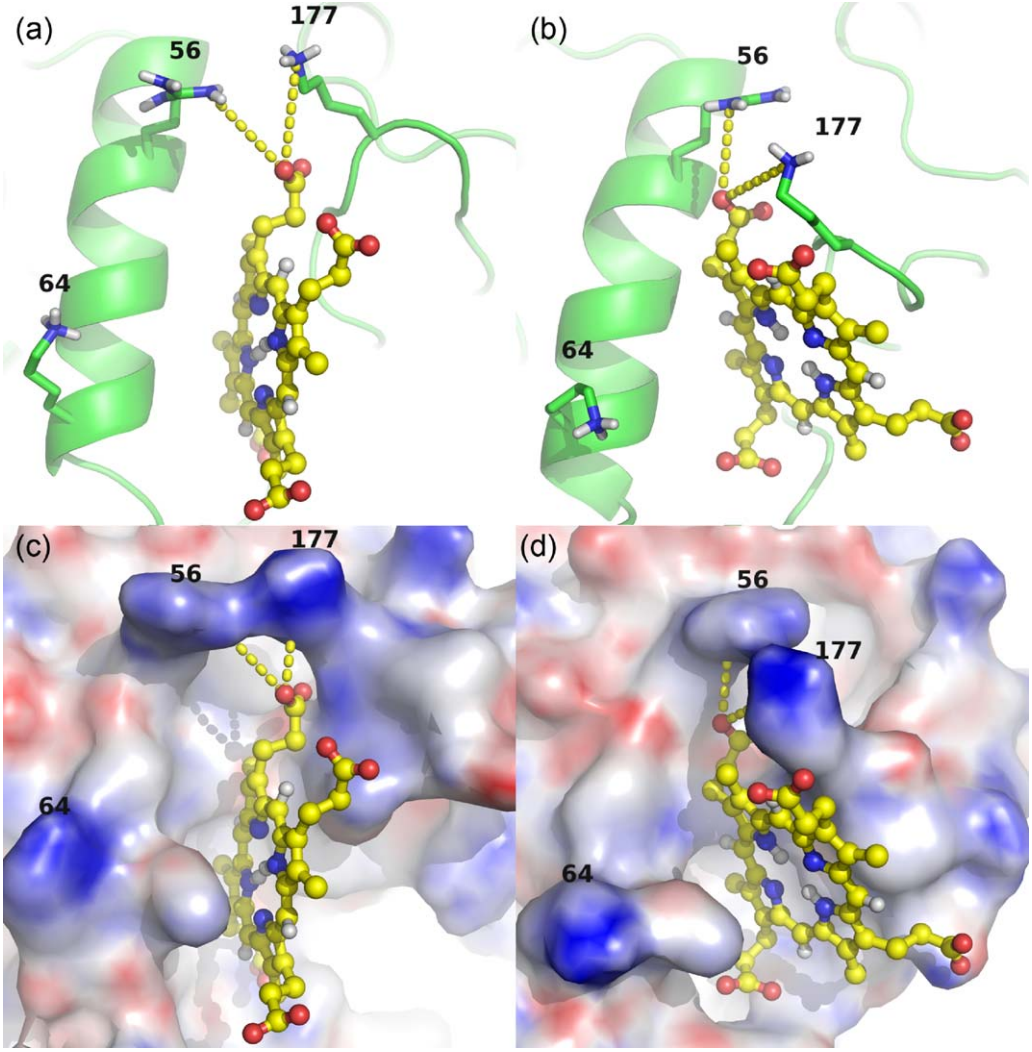

Fig. 8. Representative structure of the CPIII mHBP, (a) and (c), and CPI mHBP (b) and (d) complexes. The binding site of each complex is shown with CPIII and CPI rendered in ball and stick. The protein is rendered in cartoon in (a) and (b). Key side chain residues are rendered in sticks (see Fig. 2 for residue names). The corresponding electrostatic potential mapped on the protein molecular surface from $-5k_bT$ (red) to $+5\ k_bT$ (blue) is also shown for each complex in (c) and (d). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

3.4. PPIX–mHBP/hHBP complex

The representative structures of the murine and human PPIX complexes are shown in Fig. 4. The hHBP PPIX complex is that found to be closest to the average structure (Fig. 4b). The PPIX hHBP complex is very similar to the HEMIN hHBP complex (Fig. 3b). In both cases one of the propionate groups is stabilized by lysine 176. Due to the availability of experimental chemical shift data for the mHBP–PPIX complex (Fig. 4a) this model was determined using the available ring current shift (RCS) data. The complex shown in Fig. 4a is that which has the best agreement between the theoretical and experimental RCS differences given by Eq. (2) (see Section 2). The fitting between the calculated and the experimental shift differences is plotted in Fig. 5. In this complex one of the propionate groups is stabilized by an electrostatic interaction with lysine 64 of the α_A helix. Furthermore as for HEMIN, the preferential solvation of the complex relative to the free protein and tetrapyrrole seems to contribute to complex stabilization (Table 1).

It should be mentioned that the available NMR results [8] suggest that PPIX can occupy the mHBP binding pocket in two slightly different orientations. In order to clarify this, a cluster analysis of the PPIX–mHBP binding pocket (residues at least 5 Å from PPIX) using all conformations of the MM/MD trajectories was carried out. Two main clusters with similar number of elements corresponding to two PPIX orientations were in fact discernible. The representative complex of each cluster was selected by comparison with experimental RCS differences. The representative complexes of these two clusters are depicted in Fig. 6 showing two different binding orientations in the binding pocket, which have similar theoretical RCS differences. In the first orientation, PPIX interacts preferentially *via* a propionate group with lysine 64 as shown previously in Fig. 4, while in the second this group is involved in an electrostatic interaction with arginine 56. These two binding orientations give very similar calculated chemical shifts but correspond in fact to two different binding orientations. This supports the NMR suggestion of two populations of protein PPIX complexes and indicates the possible role of the charged residues located at the edge of the binding pocket.

Our MD simulations also indicate that the mHBP β_8–β_9 loop (residues 171–180), near the binding pocket of the protein, has a distinct mobility change upon tetrapyrrole binding (Fig. 7). This loop exhibits a significant reduction in root mean square fluctuation when HEMIN or PPIX is bound to the protein and may indicate that the β_8–β_9 loop is involved in the binding of these ligands. Experimental data are required, however, for confirmation of this fact.

3.5. CPIII/I–mHBP complexes

The complexes of CPIII and CPI with mHBP are shown in Fig. 7. To the best of our knowledge, experimental binding data is only available for CPPIII with mHBP (Table 1) [1]. The mHBP binding affinity was estimated to be −9.8 kcal/mol for CPIII and −6.1 kcal/mol for CPI, which is lower than that seen for HEMIN and PPIX complexes with mHBP. The binding estimate for CPIII is in good agreement with the experimental value reported by Taketani et al. [1]. CPIII and CPI have four propionate side chains attached to each pyrrole ring, thus, are more negatively charged than PPIX or HEMIN (Fig. 1). This may be why there is an observed reduction in binding affinity. There appears to be a steric and energetic penalty caused by the placement of the charged side chains of these ligands in the hydrophobic binding pocket. This hydrophobicity is clearly seen in Fig. 7c and d where the protein electrostatic properties are mapped on the molecular surface. The CPIII ligand is exposed to the solvent with binding supported by electrostatic interactions between the propionates and lysine 177 and arginine 56. The interaction of CPI with the protein pocket is similar to that seen for CPIII with the propionates interacting with lysine 177 and arginine 56 (Fig. 8).

4. Conclusions

The binding properties of murine and human p22HBP towards the main tetrapyrrole metabolic intermediates of heme biosynthesis have been characterized using a molecular modeling approach. The protein-binding pocket recognizes the tetrapyrrole derivatives *via* multiple and cooperative non-bonded interactions, which results in relative binding affinities to mHBP with the order: HEMIN > PPIX > CPIII > CPI. Similar relative affinities for HEMIN and PPIX towards hHBP were also obtained.

The mHBP and hHBP protein-binding pocket is mainly composed of nonpolar residues that create a solvent exposed hydrophobic binding region that is highly conserved between these two protein sequences. This structural feature allows the entry of the tetrapyrrole systems leaving their propionates exposed towards the solvent and neighboring basic residues of the protein. The binding mechanism seems to be driven by compensation of the poor free energy of solvation for the free tetrapyrroles and the apo protein (with its hydrophobic binding pocket) that occurs on binding. In mHBP and hHBP the architecture of the binding site enables the exposed propionates to be stabilized by conserved positively charged residues located at the edge of the binding site: arginine 56, and lysine 64, lysine 177 (176 in human).

In conclusion these theoretical studies have allowed a deeper understanding of the molecular recognition process that occurs between tetrapyrroles and HBP and have identified aminoacids that may be important targets for future functional mutagenesis studies.

Acknowledgements

The authors acknowledge the Fundação para a Ciência e Tecnologia (FCT) for financial support under project PTDC/QUI/64203/2006 with co-participation European Community funds from the FEDER, QREN and COMPET. Nuno M. Micaelo also thanks the FCT for the grant SFRH/BPD/35003/2007.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2010.07.008.

References

[1] S. Taketani, Y. Adachi, H. Kohno, S. Ikehara, R. Tokunaga, T. Ishii, Molecular characterization of a newly identified heme-binding protein induced during differentiation of urine erythroleukemia cells, J. Biol. Chem. 273 (1998) 31388–31394.
[2] B.J. Blackmon, T.A. Dailey, L.C. Xiao, H.A. Dailey, Characterization of a human and mouse tetrapyrrole-binding protein, Arch. Biochem. Biophys. 407 (2002) 196–201.
[3] I. Migeotte, E. Riboldi, J.D. Franssen, F. Gregoire, U. Loison, V. Wittamer, M. Detheux, P. Robberecht, S. Costagliola, G. Vassart, S. Sozzani, M. Parmentier, D. Communi, Identification and characterization of an endogenous chemotactic ligand specific for FPRL2, J. Exp. Med. 201 (2005) 83–93.
[4] M. Babusiak, P. Man, R. Sutak, J. Petrak, D. Vyoral, Identification of heme binding protein complexes in murine erythroleukemic cells: study by a novel two-dimensional native separation – liquid chromatography and electrophoresis, Proteomics 5 (2) (2005) 340–350.
[5] M.J. Zylka, S.M. Reppert, Discovery of a putative heme-binding protein family (SOUL/HBP) by two-tissue suppression subtractive hybridization and database searches, Brain Res. Mol. Brain Res. 74 (1–2) (1999) 175–181.
[6] E. Sato, I. Sagami, T. Uchida, A. Sato, T. Kitagawa, J. Igarashi, T. Shimizu, SOUL in mouse eyes is a new hexameric heme-binding protein with characteristic optical absorption, resonance Raman spectral, and heme-binding properties, Biochemistry 43 (2004) 14189–14198.
[7] A Szigeti, E. Hocsak, E. Rapolti, B. Racz, A. Boronkai, E. Pozsgai, B. Debreceni, Z. Bognar, S. Bellyei, B. Sumegi, F. Gallyas, A. Szigeti Jr., Facilitation of mitochondrial outer and inner membrane permeabilization and cell death in oxidative stress by a novel Bcl-2 homology 3 domain protein, J. Biol. Chem. 285 (3) (2010) 2140–2151.
[8] J.S. Dias, A.L. Macedo, G.C. Ferreira, F.C. Peterson, B.F. Volkman, B.J. Goodfellow, The first structure from the SOUL/HBP family of heme-binding proteins, murine P22HBP, J. Biol. Chem. 281 (2006) 31553–31561.
[9] J.S. Dias, A.L. Macedo, G.C. Ferreira, N. Jeanty, S. Taketani, B.J. Goodfellow, F.C. Peterson, B.F. Volkman, H-1, N-15 and C-13 resonance assignments of the heme-binding protein murine p22HBP, J. Biomol. NMR 32 (4) (2005) 338–1338.
[10] S.W. Ryter, R.M. Tyrrell, The heme synthesis and degradation pathways: role in oxidant sensitivity – heme oxygenase has both pro- and antioxidant properties, Free Radic. Biol. Med. 28 (2000) 289–309.
[11] S.Y. Park, T. Yokoyama, N. Shibayama, Y. Shiro, J.R.H. Tame, 1.25 angstrom resolution crystal structures of human haemoglobin in the oxy, deoxy and carbonmonoxy forms, J. Mol. Biol. 360 (2006) 690–701.
[12] W.L. DeLano, in: D. Scientific (Ed.), The PyMOL Molecular Graphics System, Palo Alto, CA, USA, 2002.
[13] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, J.A. Montgomery, T. Vreven, K.N. Kudin, J.C. Burant, J.M. Millam, S.S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G.A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J.E. Knox, H.P. Hratchian, J.B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, P.Y. Ayala, K. Morokuma, G.A. Voth, P. Salvador, J.J. Dannenberg, V.G. Zakrzewski, S. Dapprich, A.D. Daniels, M.C. Strain, O. Farkas, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J.V. Ortiz, Q. Cui, A.G. Baboul, S. Clifford, J. Cioslowski, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, C. Gonzalez, J.A. Pople, Gaussian 03, Revision C. 02, Gaussian, Inc., Wallingford, CT, 2004.
[14] C.I. Bayly, P. Cieplak, W.D. Cornell, P.A. Kollman, A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges – the RESP model, J. Phys. Chem. 97 (1993) 10269–10280.
[15] C. Oostenbrink, A. Villa, A.E. Mark, W.F. Van Gunsteren, A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6, J. Comput. Chem. 25 (2004) 1656–1676.
[16] A.K. Rappe, C.J. Casewit, K.S. Colwell, W.A. Goddard, W.M. Skiff, Uff, a full periodic-table force-field for molecular mechanics and molecular dynamics simulations, J. Am. Chem. Soc. 114 (1992) 10024–10035.
[17] A. Sali, T.L. Blundell, Comparative protein modeling by satisfaction of spatial restraints, J. Mol. Biol. 234 (1993) 779–815.
[18] R.A. Laskowski, M.W. Macarthur, D.S. Moss, J.M. Thornton, PROCHECK – a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291.
[19] G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, A.J. Olson, AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility, J. Comput. Chem. 30 (16) (2009) 2785–2791.
[20] H.J.C. Berendsen, D. van der Spoel, R. van Drunen, GROMACS: a message-passing parallel molecular-dynamics implementation, Comput. Phys. Commun. 91 (1995) 43–56.
[21] E. Lindahl, B. Hess, D. van der Spoel, GROMACS 3.0: a package for molecular simulation and trajectory analysis, J. Mol. Model. 7 (2001) 306–317.
[22] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS: a linear constraint solver for molecular simulations, J. Comput. Chem. 18 (1997) 1463–1472.
[23] S. Miyamoto, P.A. Kollman, SETTLE: an analytical version of the shake and rattle algorithm for rigid water models, J. Comput. Chem. 13 (1992) 952–962.
[24] P.E. Smith, W.F. van Gunsteren, Consistent dielectric properties of the simple point charge and extended simple point charge water models at 277 and 300 K, J. Chem. Phys. 100 (1994) 3169–3174.

[25] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. Dinola, J.R. Haak, Molecular-dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684–3690.
[26] M.K. Gilson, H.X. Zhou, Calculation of protein–ligand binding affinities, Annu. Rev. Biophys. Biomol. Struct. 36 (2007) 21–42.
[27] D. Bashford, An object-oriented programming suite for electrostatic effects in biological molecules, in: Y. Ishikawa, R.R. Oldehoeft, J.V.W. Reynders, M. Tholburn (Eds.), Scientific Computing in Object-oriented Parallel Environments, ISCOPE97, Springer, Berlin, 1997, pp. 233–240.
[28] E. Demchuk, R.C. Wade, Improving the continuum dielectric approach to calculating pK(a)s of ionizable groups in proteins, J. Phys. Chem. 100 (1996) 17373–17387.
[29] V. Mohan, M.E. Davis, J.A. Mccammon, B.M. Pettitt, Continuum model-calculations of solvation free-energies-accurate evaluation of electrostatic contributions, J. Phys. Chem. 96 (1992) 6428–6431.
[30] V.H. Teixeira, C.A. Cunha, M. Machuqueiro, A.S.F. Oliveira, B.L. Victor, C.M. Soares, A.M. Baptista, On the use of different dielectric constants for computing individual and pairwise terms in Poisson–Boltzmann studies of protein ionization equilibrium, J. Phys. Chem. B 109 (2005) 14691–14706.
[31] A.H. Elcock, D. Sept, J.A. McCammon, Computer simulation of protein-protein interactions, J. Phys. Chem. B 105 (2001) 1504–1518.
[32] H. Schafer, A.E. Mark, W.F. van Gunsteren, Absolute entropies from molecular dynamics simulation trajectories, J. Chem. Phys. 113 (2000) 7809–7817.
[33] K. Osapay, D.A. Case, A new analysis of proton chemical-shifts in proteins, J. Am. Chem. Soc. 113 (1991) 9436–9444.
[34] H.L. Liu, H.N. Zhou, W.M. Xing, H.F. Zhao, S.X. Li, J.F. Huang, R.C. Bi, 2.6 angstrom resolution crystal structure of the bacterioferritin from *Azotobacter vinelandii*, FEBS Lett. 573 (2004) 93–98.
[35] S. Macedo, C.V. Romao, E. Mitchell, P.M. Matias, M.Y. Liu, A.V. Xavier, J. LeGall, M. Teixeira, P. Lindley, M.A. Carrondo, The nature of the di-iron site in the bacterioferritin from *Desulfovibrio desulfuricans*, Nat. Struct. Biol. 10 (2003) 285–290.
[36] F. Frolow, A.J. Kalb, J. Yariv, Structure of a unique twofold symmetrical heme-binding site, Nat. Struct. Biol. 1 (1994) 453–460.