

# Molecular modeling and dynamics simulation of a histidine-tagged cytochrome $b_5$

Ying-Wu Lin · Tian-Lei Ying · Li-Fu Liao

Received: 12 October 2009 / Accepted: 25 June 2010 / Published online: 11 July 2010  
© Springer-Verlag 2010

**Abstract** Although an affinity tag such as six consecutive histidines, (His)<sub>6</sub>-tag, has been widely used to obtain high quantity of recombinant proteins, little is known about its influences on heme proteins for lack of structural information. When (His)<sub>6</sub>-tag was introduced to the N-terminus of a small heme protein, cytochrome  $b_5$ , experimental results showed the resultant protein, (His)<sub>6</sub>-cyt  $b_5$ , has similar property and function to that of isolated cyt  $b_5$ . To provide structural information for this observation, we herein performed a structural prediction of (His)<sub>6</sub>-cyt  $b_5$  by molecular modeling in combination with molecular dynamics simulation. The predicted structure, as assessed by a series of criteria with good quality, reveals that the (His)<sub>6</sub>-tag adopts a helical conformation and packs against the hydrophobic core 2 of cyt  $b_5$  through salt bridges, hydrogen bonding and hydrophobic interactions. The heme group, with the axial His ligands slightly rotated, was found to have similar conformation as in isolated cyt  $b_5$ , which indicates that the N-terminal (His)<sub>6</sub>-tag does not alter the heme active site, resulting in similar dynamics properties for core 1. This study

provides valuable information of interactions between (His)<sub>6</sub>-tag and the rest of the protein, aiding in rational design and application of functional His-tagged proteins.

**Keywords** Affinity tag · Heme proteins · Hydrogen bonding · Hydrophobic interactions

## Introduction

Histidine tag such as six consecutive histidines, (His)<sub>6</sub>-tag, has been popularly used for protein purification over the last two decades [1]. It was generally introduced into the N- or C-terminus of a protein of interest to facilitate the purification on a nickel affinity column. Although the tag may be removed by a protease, the cleavage often takes many hours and may considerably impair the target protein by side reactions [2]. Thus, small (His)<sub>6</sub>-tag is often left on the protein, which however raises a question about its effects on the protein structure and function. Recently, by surveying the crystal structures with and without His-tags within the PDB bank, Carson *et al.* showed that His-tags generally have no significant effects on the structure of the native protein [3]. However, one recent study showed by NMR that an N-terminal G (H)<sub>6</sub>LE-tag interacts transiently with the rest of the protein Rob11\_mouse in solution [4].

Comparatively, in solution little structural information is available from literature about the effects of His-tag on a protein. This is due to the fact that in most cases the (His)<sub>6</sub>-tag is too mobile to provide NOR constraints in NMR structure determination. On the other hands, as a complement to experiments, molecular dynamics (MD) simulation offers a valuable approach to provide an atomic level as well as a time-dependent view of protein structure

Y.-W. Lin (✉) · L.-F. Liao  
School of Chemistry and Chemical Engineering,  
University of South China,  
Hengyang 421001, People's Republic of China  
e-mail: linlinying@hotmail.com

Y.-W. Lin  
Department of Chemistry,  
University of Illinois at Urbana-Champaign,  
Urbana, IL 61801, USA

T.-L. Ying  
Chemical Biology Lab, Department of Chemistry,  
Fudan University,  
Shanghai 200433, People's Republic of China

in solution [5]. Therefore, it is possible to investigate the effects of His-tag on the structure of a given protein by means of MD simulation. As shown recently by Freydank *et al.* from a relative short MD simulation (100 ps, 300 K), the (His)<sub>6</sub>-tag attached at the C-terminus of tropinone reductase can interfere with the enzyme active site by steric or electrostatic interactions [6].

In a previous study, in order to simplify the purification procedure of cytochrome *b*<sub>5</sub> (cyt *b*<sub>5</sub>), a small heme protein with heme bound non-covalently by two axial histidine ligands (His39 or His63), we expressed it as a glutathione S-transferase (GST) fusion protein, GST-cyt *b*<sub>5</sub>, and purified it on a one-step affinity column chromatography with glutathione-agarose gel [7]. Alternatively, we cloned the cyt *b*<sub>5</sub> gene into an expression vector of pQE-30 that attached a (His)<sub>6</sub>-tag to the N-terminus of cyt *b*<sub>5</sub>, (His)<sub>6</sub>-cyt *b*<sub>5</sub>, which was then purified by a one-step nickel affinity column [8]. Also, other researchers used this method for cyt *b*<sub>5</sub> purification [9, 10]. Both (His)<sub>6</sub>-cyt *b*<sub>5</sub> and GST-cyt *b*<sub>5</sub> expression systems have their advantages in comparison to conventional system such as pUC19 or pET11, particularly when they aim to obtain cyt *b*<sub>5</sub> mutants with altered axial ligands, e.g., H39S or H39C [11], since these mutants were matured as apoproteins with no heme group incorporated due to the loss of one of the two axial His ligands. The one-step purification process thus overcomes the drawbacks of the colorless apocyt *b*<sub>5</sub> that has low stability in the absence of heme group.

Although we have shown that the GST protein fused to the N-terminus of cyt *b*<sub>5</sub> has no significant effects on the structure and function of cyt *b*<sub>5</sub> [7], we know little about the structural information of (His)<sub>6</sub>-tag, since it is comparatively small and more flexible. Previous study [9] as well as our preliminary investigations [8] showed that it has slight to no impact on cyt *b*<sub>5</sub>. However, we are short of the structural information of (His)<sub>6</sub>-cyt *b*<sub>5</sub> to support the experimental observations, even though isolated cyt *b*<sub>5</sub> was well-studied [12, 13]. Actually, to the best of our knowledge, in the case of heme proteins few efforts have been attempted to investigate the effects of His-tag on the protein structure and function to date. However, this issue is important for heme proteins since the His-tag might have a chance to interact with the heme group or the substrate directly or indirectly, and thus exerts influences on protein functions. In a very recent study, by enzyme assay and docking simulation Yamazaki *et al.* showed that substrate midazolam could not easily access the heme center of histidine-tagged cytochrome P450 3A5, with high interaction energy compared with isolated P450 3A5, resulting in no detectable activities of midazolam hydroxylation [14]. Meanwhile, the detail interactions of His-tag with the polypeptide chain of P450 3A5 were not well defined. With these in mind, we herein performed

molecular modeling in combination with MD simulations for (His)<sub>6</sub>-cyt *b*<sub>5</sub>, aimed at providing structural information at atomic level about the influence of His-tag on the structure and function of heme proteins in solution. Additionally, modeling of (His)<sub>6</sub>-cyt *b*<sub>5</sub> will guide future studies such as self-assembly of (His)<sub>6</sub>-cyt *b*<sub>5</sub> based on the affinity tag, providing deeper insights into native cyt *b*<sub>5</sub> system.

## Materials and methods

### System setup

As expressed in pQE-30, (His)<sub>6</sub>-cyt *b*<sub>5</sub> has a peptide of MRGSHHHHHHG attached at its N-terminus [8]. According to this sequence, the initial structure of (His)<sub>6</sub>-cyt *b*<sub>5</sub> was constructed by combination of the conformation of cyt *b*<sub>5</sub> obtained from NMR structure (PDB entry 1hko [12], model 1, a fragment of Ser5-Ser97 corresponding to lipase proteolysis with 93 residues was selected) and the His-tag generated by Swiss-PdbViewer 4.0.1. [15], a program used previously for construction of a His-tagged cytochrome *c* oxidase model [16]. It should be noted that the His-tag was constructed as a short  $\alpha$ -helix by this program. As a result, the constructed (His)<sub>6</sub>-cyt *b*<sub>5</sub> contains 104 amino acids (Met1-Ser104). The psfgen of program NAMD 2.7 b1 (Nanoscale Molecular Dynamics) [17] was used to add hydrogen atoms and assign charges to the protein, which is set up according to pH 7. Note that according to experimental observations in physiological pH range (~7.4), the  $\delta$ -nitrogen of histidine was protonated for (His)<sub>6</sub>-cyt *b*<sub>5</sub>. The protein was then solvated in a cubic box (about 85 Å × 65 Å × 60 Å) of TIP3 water with periodic boundary conditions, which extended 15 Å away from any given protein atom. The (His)<sub>6</sub>-cyt *b*<sub>5</sub> water system was then neutralized by adding a total of 28 counter ions (17 Na<sup>+</sup> and 11 Cl<sup>-</sup>) by using the autoionize plug-in of program VMD 1.8.7 (Visual Molecular Dynamics) [18], resulting in the physiological ionic strength of 0.15 M.

### Minimization and MD simulation protocols

To search the conformation of His-tag in (His)<sub>6</sub>-cyt *b*<sub>5</sub>, a minimization protocol was performed as described previously [19]. Shortly, the system was minimized 1,000 steps, followed by a MD simulation with the temperature gradually drop from 550 to 300 K (increment: -50 K, with equilibration for 200 ps at each temperature and the step size is 1 fs), and then were reminimized for 50,000 steps. The non-bonded interaction cutoff distance was set to be 9 Å. In the first cycle of minimization, residues

Val15-Ser104 of (His)<sub>6</sub>-cyt *b*<sub>5</sub> were fixed, while the N-terminal residues of Met1-Ala14 containing the His-tag, as well as the counter ions and water molecules, were allowed to move. In the second cycle of minimization, only residues from His26 (at the end of  $\alpha$ 1) to Ile86 (at the end of  $\alpha$ 5) were fixed in (His)<sub>6</sub>-cyt *b*<sub>5</sub>, while both N-terminal residues including His-tag,  $\alpha$ 1 and  $\beta$ 1 (Met1-Lys25) and C-terminal residues including  $\beta$ 2 and  $\alpha$ 6 (Ile87-Ser104) were allowed to move. For further model refinement, the resulting (His)<sub>6</sub>-cyt *b*<sub>5</sub> system was subjected to two MD simulations with temperature hold at 300 K for 2 ns. A control MD simulation on isolated cyt *b*<sub>5</sub> (Ser12-Ser104) without the (His)<sub>6</sub>-tag was also performed under identical conditions. Coordinates were saved every 1,000 steps (corresponding to 2 ps) to produce trajectory files containing 1,000 configurations for each 2-ns of simulation. The simulations were carried out with the program NAMD 2.7 b1. Visualization and data analysis were done with VMD 1.8.7.

#### Criteria to assess the modeling structure

A series of criteria were used to assess the modeling structure of (His)<sub>6</sub>-cyt *b*<sub>5</sub>: (i) The modeling structure was evaluated by ProSA (Protein Structure Analysis) as accessible at <https://prosa.services.came.sbg.ac.at>. It is an established tool that has a large user base and is frequently employed in the refinement and validation of experimental protein structures and in structure prediction and modeling [20]; (ii) The modeling structure was overlapped with the initial NMR structure of cyt *b*<sub>5</sub> as well as the MD simulation result of cyt *b*<sub>5</sub>, and Ramachandran plot analysis was further performed for these structures by using VMD 1.8.7; (iii) The dynamic property of (His)<sub>6</sub>-cyt *b*<sub>5</sub> was compared to that of isolated cyt *b*<sub>5</sub> in this study and previous MD simulations of cyt *b*<sub>5</sub> [21–23] and apocyt *b*<sub>5</sub> [24–27]; (iv) The (His)<sub>6</sub>-tag conformation and its intramolecular interactions with (His)<sub>6</sub>-cyt *b*<sub>5</sub> were compared to several crystal structures from PDB bank with (His)<sub>6</sub>-tag being well determined; (v) The structure of (His)<sub>6</sub>-cyt *b*<sub>5</sub> including both core 1 (heme-binding domain) and core 2 was compared to the initial NMR structure as well as the crystal structure of cyt *b*<sub>5</sub> (PDB entry 1cyo [12]); (vi) And the effects of (His)<sub>6</sub>-tag on the property of cyt *b*<sub>5</sub> was evaluated by comparison to the experimental results of our recent study [8].

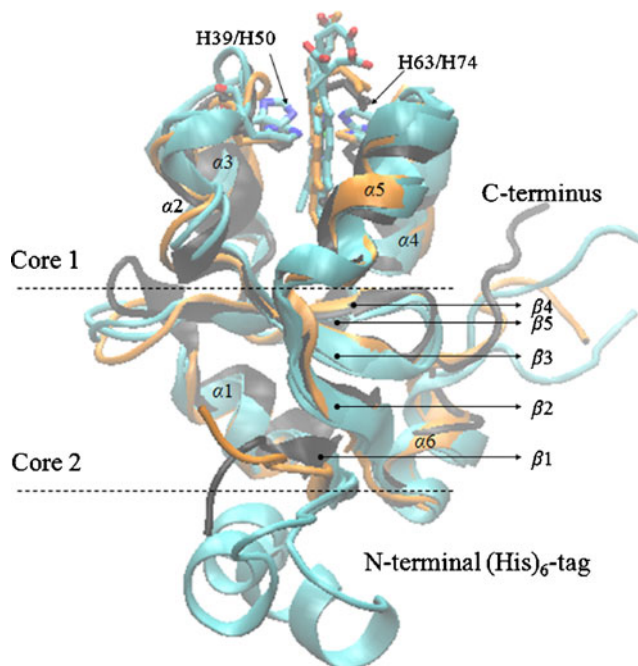
## Results and discussion

### Modeling structure of (His)<sub>6</sub>-cyt *b*<sub>5</sub>

With two cycles of energy minimization, followed by two separated MD simulations, the refined two modeling

structures of (His)<sub>6</sub>-cyt *b*<sub>5</sub> are shown in Fig. 1. The (His)<sub>6</sub>-tag keeps most of the helix structure in both cases and packs against core 2 by interacting with  $\alpha$ 1,  $\alpha$ 6 and  $\beta$ 1 (see next section for details). Although (His)<sub>6</sub>-tag often adopts an extending conformation as shown in some protein structures with His-tag being well determined [3], it does show a well-folded helical structure in some cases such as in PDB 1w3o [28] and 1v30 [29], depending on the local environments. The rest of protein (His)<sub>6</sub>-cyt *b*<sub>5</sub> overlaps well with the initial NMR structure of cyt *b*<sub>5</sub> as well as the control MD simulation structure of cyt *b*<sub>5</sub>, which indicates that the (His)<sub>6</sub>-tag introduced at N-terminus of cyt *b*<sub>5</sub> does not disturb the overall structure in a significant manner. It is interesting to observe that the loop connecting  $\alpha$ 1 and  $\beta$ 1 adopted an open conformation in present simulations. Previous studies also showed that an open-close cleft transition occurred in this region [21–23].

ProSA analysis was performed to check the quality of the modeling structures from minimization and MD simulation. As shown in Table 1, the combined energy z-score decreases from  $-6.28$  of the initial (His)<sub>6</sub>-cyt *b*<sub>5</sub> structure to about  $-6.6$  after model refinement by MD simulation, suggesting that a comparatively stable state was achieved. Control experiment of cyt *b*<sub>5</sub> also shows a slight decrease compared to the initial NMR structure.



**Fig. 1** The overlapping view of two modeling structures of (His)<sub>6</sub>-cyt *b*<sub>5</sub> (cyan) with the initial NMR structure of cyt *b*<sub>5</sub> (PDB entry 1hko, model 1, black) and that after MD simulation (orange). The two distinct regions, core 1 and core 2 (separated by dashed line), the heme group with two axial ligands (His39 or His50 and His63 or His74), the elements of secondary structure, helix  $\alpha$ 1–6, five  $\beta$  strands ( $\beta$ 1–5) (indicated by arrows), as well as the N-terminal (His)<sub>6</sub>-tag and C-terminus are highlighted

**Table 1** Calculated z-score values for the initial and modeling structure of cyt  $b_5$  and (His) $_6$ -cyt  $b_5$  by program ProSA

Structure	z-score value
Cyt $b_5$ (NMR, PDB entry 1hko, model 1) <sup>a</sup>	-7.01
Cyt $b_5$ (MD simulation)	-7.18
(His) $_6$ -cyt $b_5$ (Initial structure) <sup>b</sup>	-6.28
(His) $_6$ -cyt $b_5$ (Minimization)	-6.44
(His) $_6$ -cyt $b_5$ (MD simulation-1)	-6.58
(His) $_6$ -cyt $b_5$ (MD simulation-2)	-6.63

<sup>a</sup>) 93 amino acids (Ser5-Ser97)

<sup>b</sup>) 104 amino acids (Met1-Ser104)

Additionally, Ramachandran plots show that over 90% residues are in favored region and allowed region for two models of (His) $_6$ -cyt  $b_5$ , reflecting the good stereo chemical features. These observations further elucidate the acceptability of the predicted structures of (His) $_6$ -cyt  $b_5$ .

Furthermore, the total energy of (His) $_6$ -cyt  $b_5$  in MD simulations as a function of time step was plotted in Fig. 2a, with that of cyt  $b_5$  shown for comparison. The total energy fluctuates around  $-65600 \text{ kcal mol}^{-1}$  for cyt  $b_5$  and around  $-45600 \text{ kcal mol}^{-1}$  for (His) $_6$ -cyt  $b_5$  in both MD simulations during the last 1-ns, suggesting an energy favorable state was achieved. An analysis of the backbone  $C_\alpha$  root mean square deviations (RMSD) was further performed for both core 1 and core 2 in (His) $_6$ -cyt  $b_5$ , and in cyt  $b_5$  as well, with the C-terminal six residues (Thr99-Ser104) excluded (Fig. 2b–d). It can be observed that RMSD fluctuates during the last 0.5-ns MD simulation, indicating an achievement of stable conformation equilibrium. The RMSD value of core 1 in both (His) $_6$ -cyt  $b_5$  and cyt  $b_5$  is found to be around 1.5 Å, which suggests that the introduction of a (His) $_6$ -tag at the N-terminus of cyt  $b_5$  leaves the heme-binding core 1 largely intact. Meanwhile, a slightly lower RMSD value is observed for core 2 in (His) $_6$ -cyt  $b_5$  (around 0.75 Å) compared to that in isolated cyt  $b_5$  (around 1.0 Å), indicating that (His) $_6$ -tag stabilizes some residues in core 2 to some extent by direct interactions. The different RMSD value between core 1 and core 2 also agrees well with previous MD study of cyt  $b_5$  [21]. Moreover, in the absence of heme group, apocyt  $b_5$  exhibits the RMSD value of core 1 and core 2 around 3.0 and 1.5 Å, respectively [25].

During the last 0.5-ns MD simulation, the  $C_\alpha$  root mean square fluctuations (RMSF) over time of each residue presents more detailed motion of (His) $_6$ -cyt  $b_5$ . As shown in Fig. 3, for heme-binding core 1, (His) $_6$ -cyt  $b_5$  displays very similar RMSF to that of cyt  $b_5$ , lower than that of apocyt  $b_5$  in previous MD simulations due to the destabilization by heme removal from the heme-binding core 1 [25]. In case of hydrophobic core 2, (His) $_6$ -cyt  $b_5$  is found

to have slightly lower RMSF value for some residues compared to cyt  $b_5$ , such as Glu21 and Glu22 in  $\alpha 1$ , and residues 90–94 in  $\beta 1$  and  $\alpha 6$ , due to the strong interactions with (His) $_6$ -tag, as analyzed in next section. As a result, comparatively low RMSF was observed for the (His) $_6$ -tag, even though it locates in the N-terminus.

#### Interactions of (His) $_6$ -tag and cyt $b_5$

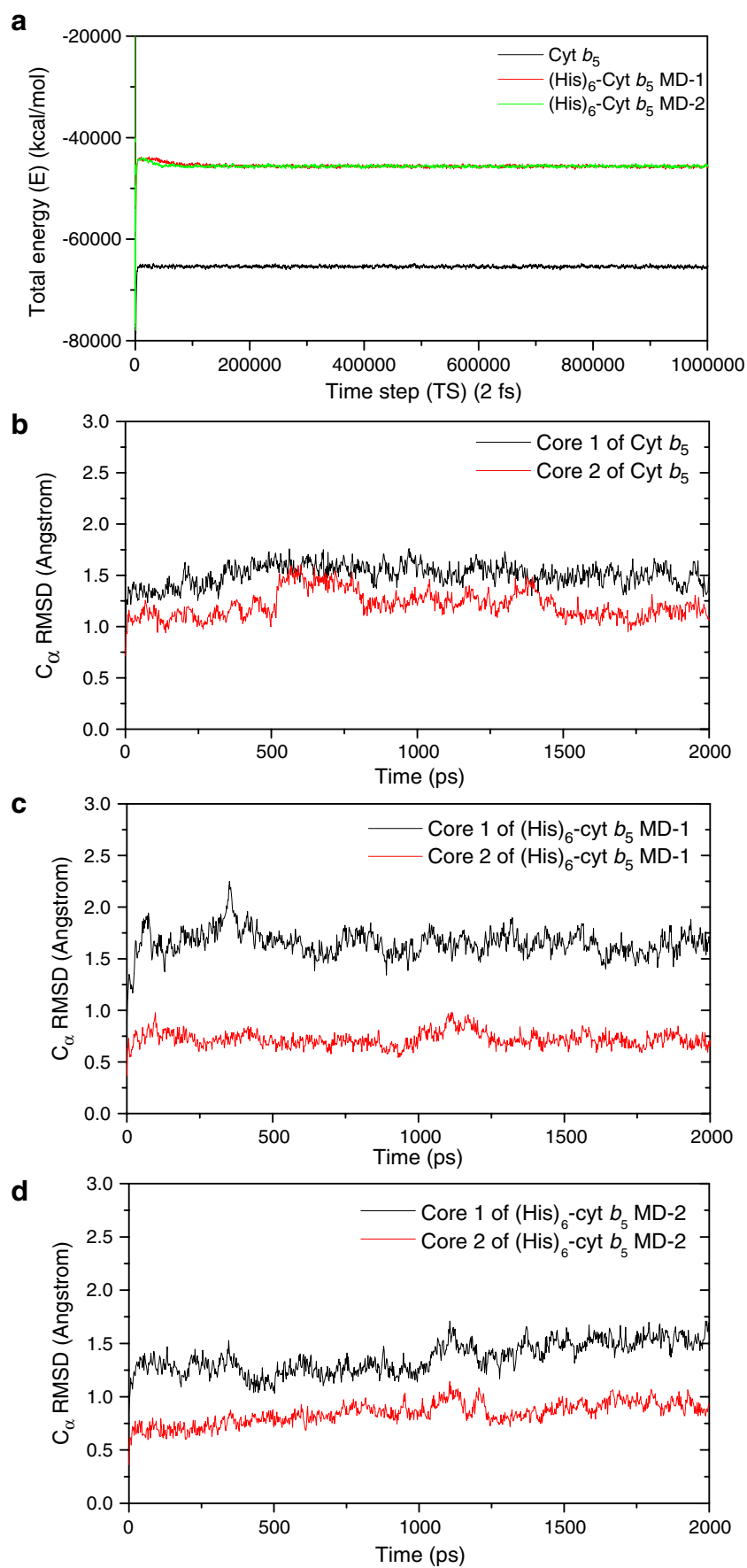
By inspecting the local region of (His) $_6$ -tag in the modeling structure of (His) $_6$ -cyt  $b_5$ , it is found that the (His) $_6$ -tag interacts with the neighboring  $\alpha 1$ ,  $\alpha 6$  and  $\beta 1$  through salt bridges, hydrogen bonding as well as hydrophobic interactions. As shown in Fig. 4, in both structures Arg2 forms salt bridges and hydrogen bonding simultaneously with Glu21 and Asp94 that are located in  $\alpha 1$  and  $\alpha 6$ , respectively. Glu22, another negative charged residue located in  $\alpha 1$ , interacts with the backbone N atoms, and with the neighboring Lys25 in one structure. This kind of interaction was also observed in the crystal structure of *E. Coli* YfhJ protein in which a (His) $_6$ -tag is presented (PDB entry 1uj8) [30].

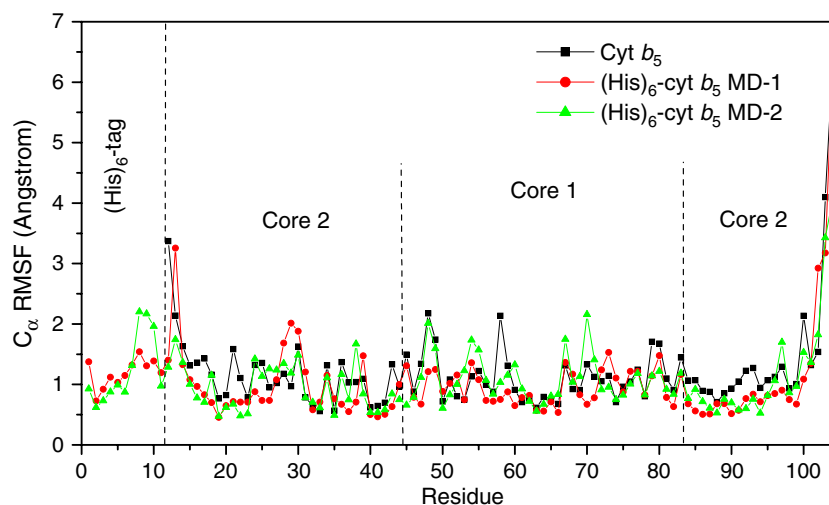
At the same time, hydrophobic interactions are formed among His7, Val15 and Tyr18 between His-tag and core 2 of cyt  $b_5$ , where His7 also forms a hydrogen bond with Gly11 in one structure (Fig. 4b). Note that Tyr18 located in  $\beta 1$  plays a crucial role in stabilizing the hydrophobic core 2 of cyt  $b_5$  and also the entire protein structure, as illustrated recently by high temperature MD simulations [27]. The packing behavior of (His) $_6$ -tag in (His) $_6$ -cyt  $b_5$  will thus extend the hydrophobic interactions from His26-Trp33-Ile87 to Tyr18-Val15-His7, resulting in stabilization of both the His-tag and the hydrophobic core 2 of cyt  $b_5$ , as reflected in Figs. 2 and 3. Three histidine residues, His5, His8 and His9, are found to locate on the same side of (His) $_6$ -tag toward the solution, which likely coordinate to the Ni(II) ions during protein purification.

#### Effects of (His) $_6$ -tag on cyt $b_5$

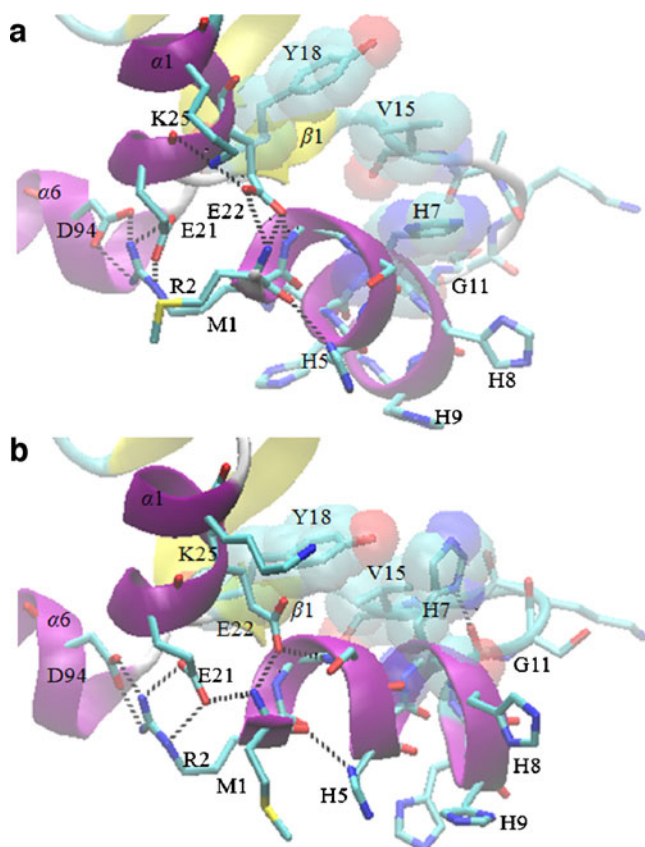
From the spatial alignment of protein polypeptide chain displayed in Fig. 1, it can be found that the (His) $_6$ -tag, when attached at the N-terminus, causes little conformational disturbance to the hydrophobic core 2 of cyt  $b_5$ , the latter is responsible for stabilizing the entire protein by pre-organization of the second heme-binding core 1 [12]. Actually, what one might be interested in is the influence of (His) $_6$ -tag on the heme-binding region, especially on the heme group in (His) $_6$ -cyt  $b_5$  as well as its two axial His ligands, His50 and His70. To probe such effects, we overlapped the heme group in modeling structure of (His) $_6$ -cyt  $b_5$  with the initial NMR structure as well as the crystal structure of cyt  $b_5$ , as shown in Fig. 5, including the

**Fig. 2** (a) Plot of the total energy versus the time step in MD simulations of *cyt b<sub>5</sub>* and (His)<sub>6</sub>-*cyt b<sub>5</sub>*; (b) RMSD as a function of time for Core 1 and Core 2 in *cyt b<sub>5</sub>* from the initial NMR structure and in (His)<sub>6</sub>-*cyt b<sub>5</sub>* from the minimization structure (c and d)





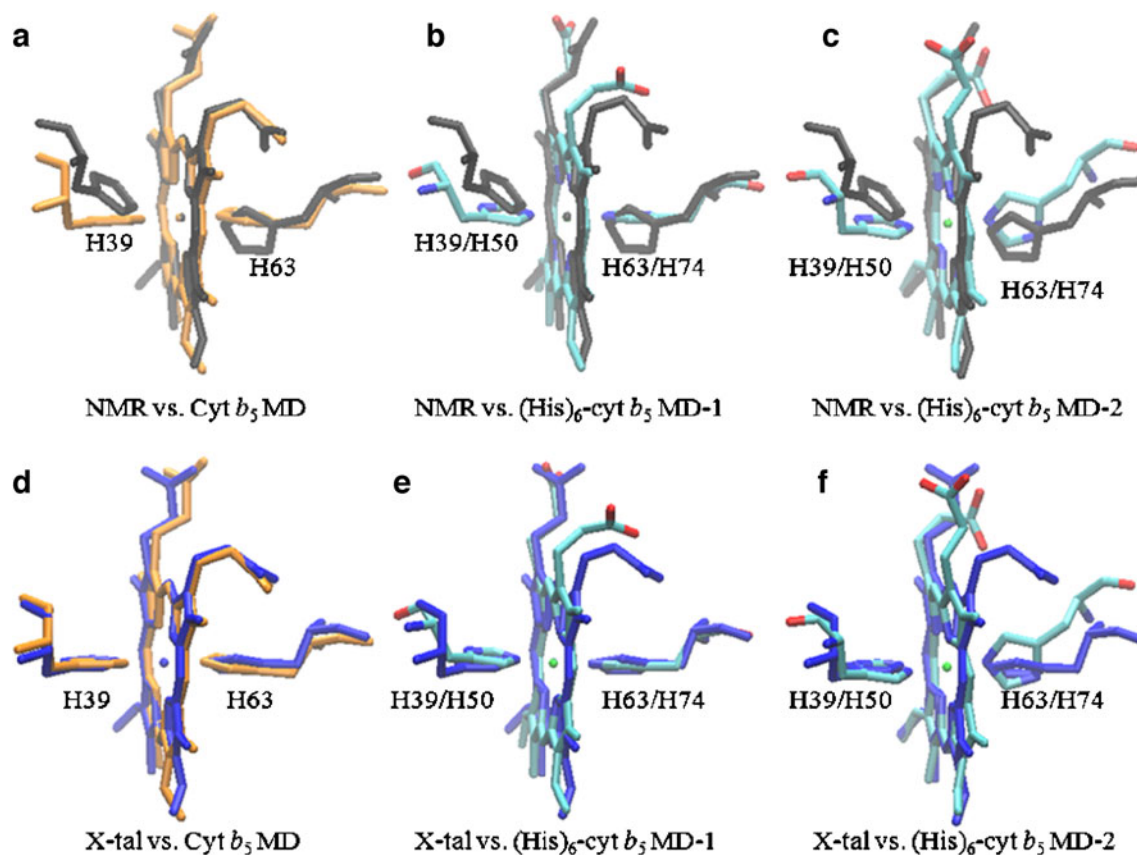
**Fig. 3** Residue  $C_{\alpha}$  RMSF over time during the last 0.5-ns simulation. The dashed lines separate the  $(\text{His})_6$ -tag (residues 1-11), core 1 (residues 44-84) and core 2 (residues 12-43 and 85-104)



**Fig. 4** Interactions between  $(\text{His})_6$ -tag and the rest of protein  $(\text{His})_6$ -cyt  $b_5$  MD-1 (**a**) and  $(\text{His})_6$ -cyt  $b_5$  MD-2 (**b**). The hydrogen bonding networks are shown by dotted lines. The hydrophobic interactions between Y18, V15 and H7 are shown as VMD representation. All other residues discussed in text as well as  $\alpha 1$ ,  $\alpha 6$  and  $\beta 1$  are also highlighted

control simulation of isolated cyt  $b_5$ . When compared to the initial NMR structure (Fig. 5a–c), the conformation of two axial ligands is found to be similar in one  $(\text{His})_6$ -cyt  $b_5$  structure (Fig. 5c), while in the other (Fig. 5b), the two histidine imidazole rings are rotated to be nearly parallel to each other, this also happened to the simulation of cyt  $b_5$  (Fig. 5a). However, as illustrated in Fig. 5d, e, the resultant conformation overlaps well with the crystal structure of cyt  $b_5$  (PDB entry 1cyo). This is very interesting since although the MD simulation was performed in water solution, similar to that in NMR determination, the resultant structure is closer to the crystal structure, indicating a lower energy state can be achieved by MD simulation. The rotation of the axial histidines was also observed in previous MD simulations of other heme proteins [31–33], as well as in experimental studies of cyt  $b_5$  and neuroglobin [34–36], which was believed to be associated with the protein function by modulating the electronic properties of the heme group.

On the other hand, comparison of the coordination distance between axial histidine ligands and heme iron shows that in both  $(\text{His})_6$ -cyt  $b_5$  structures the distance is slightly shorter than that determined by NMR or X-ray technique, while similar to the control simulation results of cyt  $b_5$  (Table 2). This suggests that the slight change of coordination in  $(\text{His})_6$ -cyt  $b_5$  is not due to the effects of  $(\text{His})_6$ -tag, but due to the MD simulation itself. The location and conformation of  $(\text{His})_6$ -tag in  $(\text{His})_6$ -cyt  $b_5$  determines that it can not interact directly with the heme group. This is different from His-tagged cytochrome P450 3A5, where the histidine tag is close to the heme active site and interacts with the substrate directly [14]. Therefore, in case of core 1 in  $(\text{His})_6$ -cyt  $b_5$ , the influence belongs to a distal weak interaction exerted through the hydrophobic core 2. In addition, core 2 exhibits a similar



**Fig. 5** Comparison of heme active sites between *cyt b<sub>5</sub>* MD result (orange),  $(\text{His})_6\text{-cyt } b_5$  (cyan) and *cyt b<sub>5</sub>* in NMR (black) (a, b and e) or crystal structure (blue) (d, e and f)

conformation in  $(\text{His})_6\text{-cyt } b_5$  compared to that in *cyt b<sub>5</sub>* (Fig. 1). These observations interpret the experimental observations that both  $(\text{His})_6\text{-cyt } b_5$  and *cyt b<sub>5</sub>* have the same UV-vis spectrum (413 nm in oxidized state, and 423, 527 and 556 nm in reduced state), as well as the similar chemical stability toward guanidine-induced unfolding (3.12 M of  $(\text{His})_6\text{-cyt } b_5$  vs. 3.18 M of *cyt b<sub>5</sub>*) [8]. At the same time, the visible circular dichroism (CD) spectrum, a good conformational probe for heme proteins [37], identified a slight conformational change occurred around the heme group in  $(\text{His})_6\text{-cyt } b_5$  with respect to *cyt b<sub>5</sub>*. This is also reflected in current simulations, as the rotation of heme axial ligands is observed in both simulations of  $(\text{His})_6\text{-cyt } b_5$  (Fig. 5).

## Conclusions

In this study, the structure of a  $(\text{His})_6\text{-cyt } b_5$  was predicted by computer simulations. Due to the special folding module of *cyt b<sub>5</sub>*, the  $(\text{His})_6$ -tag introduced at the N-terminus does not interfere with the protein active site. The interactions between  $(\text{His})_6$ -tag and the rest of *cyt b<sub>5</sub>* revealed in the study, both in a local position and at a long distance, provide deep insights into the experimental observations. Molecular modeling combined with MD simulations offers a convenient approach to produce a more precise image of the  $(\text{His})_6$ -tag conformation, making it predictable for the consequences of introducing such an affinity tag. This can thus aid in rational design of

**Table 2** Comparison of the distance between heme axial ligands and heme iron for *cyt b<sub>5</sub>* and the modeling  $(\text{His})_6\text{-cyt } b_5$  proteins (Å)

Protein	H39/H50-Fe	H63/H74-Fe
<i>Cyt b<sub>5</sub></i> (NMR, PDB entry 1hko, model 1)	2.22	1.92
<i>Cyt b<sub>5</sub></i> (MD simulation)	1.74	1.74
<i>Cyt b<sub>5</sub></i> (X-tal structure, PDB entry 1cyo)	2.07	2.00
$(\text{His})_6\text{-cyt } b_5$ (MD simulation-1)	1.72	1.74
$(\text{His})_6\text{-cyt } b_5$ (MD simulation-2)	1.74	1.73

a His-tag protein to obtain high quantity of functional protein, and therefore its applications, especially for heme protein with alterations at heme axial His ligands that otherwise is difficult to obtain.

**Acknowledgments** This work was supported by the initial foundation for Ph. D. introduced in University of South China (No. 506XJQ06001). NAMD and VMD were developed by the Theoretical Biophysics Group in the Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana-Champaign, USA. We also thank Fengyun Ni at Rice University, USA, for discussion of energy minimization.

## References

1. Smith MC, Furman TC, Ingolia TD, Pidgeon C (1988) *J Biol Chem* 263:7211–7215
2. Jenny RJ, Mann KG, Lundblad RL (2003) *Protein Expr Purif* 31:1–11
3. Carson D, Johnson DH, McDonald H, Brouillette C, DeLucas LJ (2007) *Acta Crystallogr D* 63:295–301
4. Song J, Markey JL (2007) *Protein Pept Lett* 14:265–268
5. Klepeis JL, Lindorff-Larsen K, Dror RO, Shaw DE (2009) *Curr Opin Struct Biol* 19:120–127
6. Freydank AC, Brandt W, Drager B (2008) *Proteins* 72:173–183
7. Lin YW, Zhao DX, Wang ZH, Huang ZX (2006) *Protein Expr Purif* 45:352–358
8. Lin YW (2005) Doctoral Dissertation, Fudan University, Shanghai, China, 86–106
9. Jung Y, Kwak J, Lee Y (2001) *Appl Microbiol Biotechnol* 55:187–191
10. Sobrado P, Goren MA, James D, Amundson CK, Fox BG (2008) *Protein Expr Purif* 58:229–241
11. Wang WH, Lu JX, Yao P, Xie Y, Huang ZX (2003) *Protein Eng* 6:1047–1054
12. Durlley RC, Mathews FS (1996) *Acta Crystallogr D Biol Crystallogr* 52:65–76
13. Muskett FW, Kelly GP, Whitford D (1996) *J Mol Biol* 258:172–189
14. Emoto C, Murayama N, Wakiya S, Yamazaki H (2009) *Drug Metab Lett* 3:207–211
15. Guex N, Peitsch MC (1997) *Electrophoresis* 18:2714–2723
16. Friedrich MG, Giess F, Naumann R, Knoll W, Ataka K, Heberle J, Hrabakova J, Murgida DH, Hildebrandt P (2004) *Chem Commun* 21:2376–2377
17. Kalé L, Skeel R, Bhandarkar M, Brunner R, Gursoy A, Krawetz N, Phillips J, Shinozaki A, Varadarajan K, Schulten K (1999) *J Comput Phys* 151:283–312
18. Humphrey W, Dalke A, Schulten K (1996) *J Mol Graph* 14:33–38
19. Ni FY, Cai B, Ding ZC, Zheng F, Zhang MJ, Wu HM, Sun HZ, Huang ZX (2007) *Proteins* 68:255–266
20. Wiederstein M, Sippl MJ (2007) *Nucleic Acids Res* 35:W407–W410
21. Storch EM, Daggett V (1995) *Biochemistry* 34:9682–9693
22. Lee KH, Kuczera K (2003) *Biopolymers* 69:260–269
23. Cheng Q, Benson DR, Rivera M, Kuczera K (2006) *Biopolymers* 83:297–312
24. Storch EM, Daggett V (1996) *Biochemistry* 35:11596–11604
25. Lin YW, Wang ZH, Ni FY, Huang ZX (2008) *Protein J* 27:197–203
26. Lin YW, Ying TL, Liao LF (2009) *Chin Chem Lett* 20:631–634
27. Lin YW, Nie CM, Liao LF (2009) *J Mol Struct Theochem* 910:154–162
28. Leiros HKS, Kozielski-Stuhrmann S, Kapp U, Terradot L, Leonard GA, Mcsweeney SM (2004) *J Biol Chem* 279:55840–55849
29. Tajika Y, Sakai N, Tamura T, Yao M, Watanabe N, Tanaka I (2004) *Proteins* 57:862–865
30. Shimomura Y, Takahashi Y, Kakuta Y, Fukuyama K (2005) *Proteins* 60:566–569
31. Galstyan AS, Zarić SD, Knapp E (2005) *J Biol Inorg Chem* 10:343–354
32. Medaković V, Zarić SD (2003) *Inorg Chim Acta* 349:1–5
33. Zarić SD, Popović DM, Knapp E (2001) *Biochemistry* 40:7914–7928
34. Zhang Q, Cao C, Wang ZQ, Wang YH, Wu HM, Huang ZX (2004) *Protein Sci* 13:2161–2169
35. Sarma S, DiGate RJ, Goodin DB, Miller CJ, Guiles RD (1997) *Biochemistry* 36:5658–5668
36. Walker FA (2006) *J Biol Inorg Chem* 11:391–397
37. Myer YP (1970) *Biochim Biophys Acta* 214:94–106