Theoretical Evidence to *Cis* Folding in PhoR Histidine Kinase of *Corynebacterium Pseudotuberculosis*

Gleiciane L. Moraes, Guelber Cardoso, Cláudio N. Alves & Jerônimo Lameira

Introduction

Two-component systems (TCS) are signaling pathways that respond to changes environmental by modifying cellular behaviors. The first component of these systems is a dimeric sensor histidine kinase (HK); the second is its cognate target, termed the response regulator (RR)1. In bacterias the main signal-transduction mechanism is via TCS systems, in which the signaling process starts by the autophosphorylation of a highly conserved histidine residue in the sensor HK. In response to a signal generated in the N-terminus sensor domain of HK, occur transfer γ-phosphate from ATP bound to the catalytic ATP-binding domain (CA) to conserved histidine residue on the dimerization/histidine phosphorylation (DHp) domain in the C-terminus kinase core (DHp_CA)[1]. Depending on the architecture of the DHp domain, the histidine residue is phosphorylated by CA domain of same chain (cis-autophosphorylation) or of the other chain (trans- autophosphorylation)[2] (Fig.1). This being suggested that autophosphorylation mechanism is determined by the loop at the base of the DHp domain four-helix bundle[2,3,4]. Surprisingly, Ashenberg and colleagues (2013) found that the mode of autophosphorylation be highly conserved in orthologous sets of histidine kinases, despite the sequence of the DHp domain loop is highly variable, and that orthologs of PhoR with diverse loops autophosphorylated in cis and orthologs of EnvZ with diverse loops autophosphorylated in trans. Chimeras constructed in which was replaced the loop of EnvZ (trans) by loop of PhoR (cis), change autophosphorylation mechanism from in cis to in trans, indicate that the DHp loop is a functionally important determinant of autophosphorylation mechanism in histidine kinases, supplanting the notion that these loops are simple linkers between helices[2,4]. Here, we obtained structure 3D by comparative molecular modeling of the phosphate sensor PhoR *Corynebacterium pseudotuberculosis* in folding cis and trans, and carried out a molecular dynamics simulation to finding theoretical evidences of folding in HK.

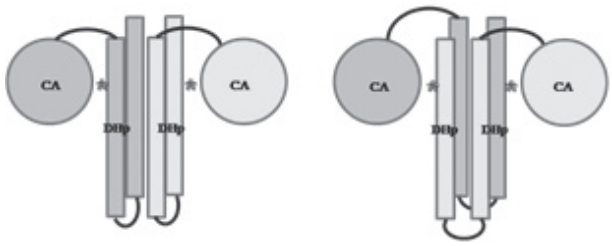


Figure 1. Folding C-terminus kinase core in HK. a) Folding to cis-autophosphorylation and b) Folding to trans-autophosphorylation.

Methods

PROTEIN STRUCTURE MODELING

The amino acid primary sequences enzyme of M. tuberculosis FRC41 were obtained from the NCBI (http://blast.ncbi.nlm.nih.gov/) under accession number YP_003784140.1. The structures of the enzymes were generated by of SWISS-MODEL Server5 and Modeller 9v12[6] software. Modeling of the folding to cis-autophosphorylation was obtained using osmolarity sensor protein EnvZ chimera *E. coli* (EnvZchim) histidine kinase (PDB code: 4KP4)4; and folding to trans-autophosphorylation was obtained using osmolarity sensor protein EnvZ native E. coli ($EnvZ^{wt}$) histidine kinase (PDB code: 4CTI)7. ATP and Mg^{2+} were transfer as a rigid body to generated model using the BLK residue type in Modeller 9v12[6]. The structures of each protein modelled were validated using Ramachandran plot,[8] QMEAN score9 (available in SWISS-MODEL Server5) and ProQ[10].

MOLECULAR DINAMICS SIMULATION

The ionization state of ionizable residues was assessed from PROPKA3[11] calculations, except to catalytic residue histidine (His289), which must be protonated in triphosphate hydrolysis[12].

We carried out molecular dynamics simulations using AMBER 12 software package[13]. The system was immersed in an octahedral box of TIP3P14 water molecules and sodium ions were added to neutralize the system. The parm99SBildn force field parameters in the Amber12[13] package were used for the protein and ions, the Carlson corrected parameters[15] were used for the ATP molecule and the Sticht parameters[16] for the cofactor ion Mg2+. The minimization has been made using multistep protocol, first the adjustment of hydrogens, following water molecules and counterions were refined, and finally the minimization of the whole system. Thermalization of system was performed by heating in 10 steps, initially one step of 20ps from 0 to 100 K at constant volume, 8 steps of 1ns increasing the temperature from 100 to 300 K, and an additional step of 5ns was performed in order to equilibrate the system density at constant pressure (1 bar). The bonds containing hydrogen atoms were kept at the equilibrium distance using the Shake algorithm[17]. Finally, a 50ns trajectory were run in the NTP formalism, the temperature were kept constant Berendsen thermostat[18] using a time step of 2fs, periodic boundary conditions and Ewald sums[19] (grid spacing of 1 Å) for long-range electrostatic interactions.

Results and Discussion

PROTEIN STRUCTURE MODELING

Homology modeling or comparative modeling allows the construction of the tertiary structure of a protein based on the primary structure similarity[20]. The best structures obtained by comparative molecular modeling of the phosphate sensor PhoR C. pseudotuberculosis in folding cis (Fig.2a) and trans (Fig.2b) are presented. The better structure in folding trans was obtained in Modeller 9v12[6] software. The folding trans structure constructed in SWISS-MODEL Server5 had intertwining of DHP loop, because this structural error the model was rejected to continue the work (structure not shown). The better structure in folding cis was obtained in SWISS-MODEL Server[5].

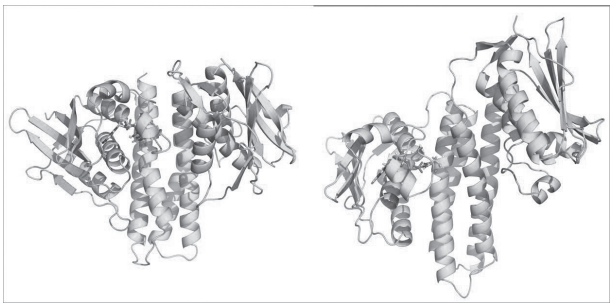


Figure 2. Folding C-terminus kinase core in HK of the C. pseudotuberculosis. a) folding cis and b) folding trans.

Table 1 shows a summary of validation results of build 3D structures.

Table 1. Validation proteins modeled.

Folding	Ramachandran plot	LGscore	QMEAN	RMSD
cis	91.9%	5.021	0.614	1.519
trans	94.4%	4.068	0.683	8.681

Based on the parameters shown in Tab.1, both built models show good values Ramachandran plot8 (most favoured regions), QMEAN[9] and LGscore[10]. However, *trans* folding showed high RMSD to the template (8.681). Furthermore, graphical plots of ANOLEA[21] mean force potential that represents the energy for each amino acid of the protein chain indicates that the packing quality of the 3D structure is better in cis folding (Supporting information). In other modeling performed using HK CpxA of E. coli[22] (32% de identity and 49% de similarity), a folding trans, showed 90,8% of the residues are in the most favoured regions of the Ramachandran plot8, QMEAN9 of 0.453, LGscore[10] of 3.76, RMSD of 7.46 Å and very bad packing quality (structure not shown). It is important to note that, despite the low identity between the osmolarity sensor protein $EnvZ^{chim}$ (26% identity and 46% similarity) of E. coli and the phosphate sensor PhoR of C. pseudotuberculosis, the models obtained from this template were very better including the quality of packaging and RMSD. The superposition of the between target and templates (Fig.3) indicates that all secondary structures in the target protein were conserved in both models, however the model constructed using $EnvZ^{chim}$, trans folding, it does not shows good structural overlap. On the other hand, o model constructed using $EnvZ^{wt}$, cis folding, and the model shows good structural overlap and the generated 3D structure is good according to the parameters analysed.

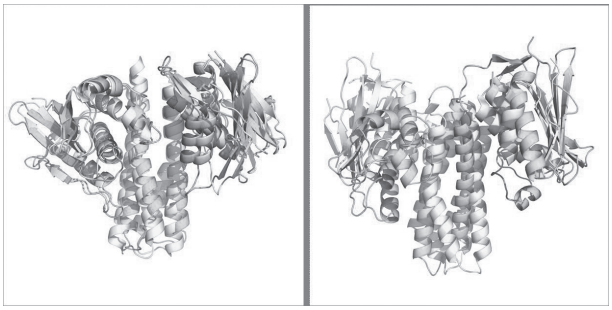


Figure 3. Superposition of the calculated PhoR of C. pseudotuberculosis (green) and a) EnvZchim (PDB code: 4KP44), b) EnvZwt (PDB code: 4CTI7) of E. coli (yellow).

The catalytic core (DHp_Ca) HK of C. pseudotuberculosis N-terminus (DHp) is formed by two long helices (α1 and α2), where α1 contains the phosphorylatable conserved His289 residue, and C-terminus (CA) is formed by an α/β sandwich fold similar to the ATPase domains of the ATPase superfamily[23]. This latter domain binds ATP in a pocket covered by a flexible and variable loop named ATP lid[24].

MOLECULAR DINAMICS SIMULATION

To explore the implications of folding in the stability of catalytic core PhoR complex (protein+ATP+Mg^{2+}), we performed 50 ns of MD simulations at 300 K for structures obtained in the previous step (Fig.4b). Moreover, we carried out molecular dynamics simulations in $EnvZ^{chim}$ (cis) and $EnvZ^{wt}$ (trans) complex (protein+ATP+Mg2+) (Fig.4a)

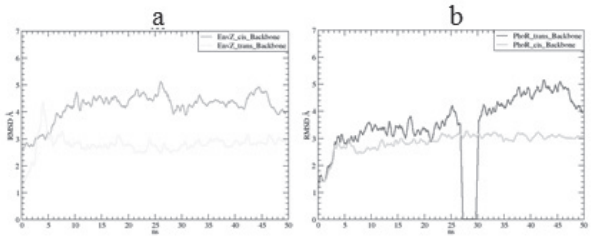


Figure 4. Average RMSD in performed trajectory of 50 ns of MD simulations. a) EnvZchim and EnvZwt of E. coli and b) PhoR of C. pseudotuberculosis folding trans and cis.

After the first 5 ns MD, it is observed that the RMSD of EnvZwt stabilizes, as well as the RMSD folding cis PhoR. It is important to remind that orthologs of PhoR with diverse loops autophosphorylated in cis and orthologs of EnvZ with diverse loops autophosphorylated in trans[2]. Interestingly osmolarity sensor protein EnvZ native structure of E. coli and the more probable structure folding of phosphate sensor PhoR of C. pseudotuberculosis are which present the more stable RMSD during MD simulation trajectory. Even after 50 ns MD simulation the RMSD of chimera EnvZ cis and PhoR trans folding are less stable. In crystallographic studies by Marina et. al. (2014)[4] there

was obtained the EnvZ chimera shown that it is possible to change the autophosphorylation mechanism of cis to trans in histidine kinase by changing the connection handle. However, our studies show that during the MD simulation time produced, this folding is less stable. Regarding the PhoR protein, it is important to note that DHp loop connection was not replaced. In our study, the sequence of the native PhoR protein was modeled in folds cis and trans and has been theoretically evaluated the stability of structure in each folding. We also monitored the interactions of the active site of C. pseudotuberculosis PhoR (Fig 5). It is expected that a Mg^{2+} ion bridges the three nucleotide phosphates, an invariant Asn and two water molecules complete the octahedral coordination sphere of the Mg^{2+} ion in the structures[4, 25].

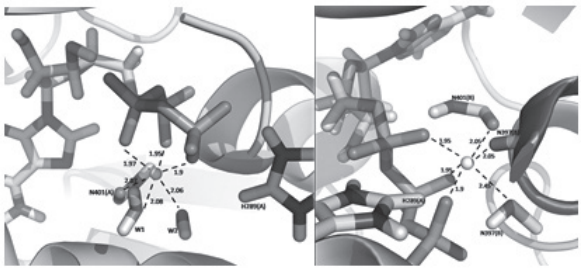


Figure 5. Octahedral coordination sphere of the Mg2+ ion in the C. pseudotuberculosis. a) folding cis and b) folding trans.

During the MD simulations of the different PhoR system was observed the formation of an octahedral coordination sphere. The octahedral coordination sphere observed in cis folding PhoR showed Mg^{2+} ion bridges with three ATP phosphates, Asn401 carbonyl and two water molecules (W1 and W2) (Fig.5a), as described for other HK[4, 25]. On the other hand, the octahedral coordination sphere observed in trans folding PhoR showed Mg^{2+} ion bridges with three ATP phosphates of chain A, Asn397 and Asn401 carbonyl and Thr400 amino acid residues (Fig.5b). The latter octahedral coordination sphere is not according to the procedure described for other HK.

Conclusions

The results obtained by comparative modeling and MD simulations indicate the fact that the probable folding to phosphate sensor PhoR of C. pseudoturberculosis (based in orthologs of PhoR) is the better structure obtained by comparative modeling and show the more stable RMSD during DM simulations. Therefore, the HK folding can be predicted theoretically.

Acknowledgments

The authors are grateful for the support given from the UFPA and CAPES.

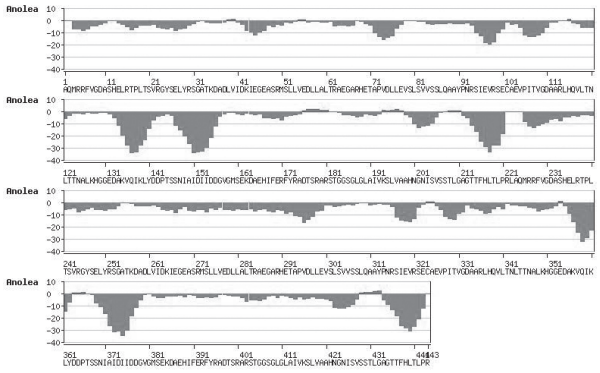
References

1. R. Gao, A. M. Stock, Annu. Rev. Microbiol. 63, 133, (**2009**).
2. O. Ashenberg, A. E. Keating, M. T. Laub, J. Mol. Biol. 425, 1198, (**2013**).
3. P. Casino, V. Rubio, A. Marina, Cell. 139, 325, (**2009**).
4. P. Casino, L. Miguel-Romero, A. Marina, Nat. Commun. 5, 3258, (**2014**).
5. K. Arnold, L. Bordoli, J. Kopp, T. Schwede, Bioinformatics. 22, 195, (**2006**).
6. N. Eswar, B. Webb, M. A. Marti-Renom, M. S. Madhusudhan, D. Eramian, M-Y. Shen, U. Pieper, A. Sali. Curr. Protoc. Protein Sci / editorial board, J. E. Coligan [et al] Chapter 2, Unit 2.9, (**2007**).
7. H. U. Ferris, M. Coles, A. N. Lupas, M. D. Hartmann, J. Struct. Biol. 186, 376, (**2014**).
8. R. A. Laskowski, M. W. Macarthur, D. S. Moss, J. M. Thornton. J. Appl. Cryst. 26, 283, (**1993**).
9. B. Wallner, A. Elofsson, Prot. Sci. 12, 1073, (**2003**).
10. P. Benkert, S. C. E. Tosatto, D. Schomburg, proteins struct. funct. bioinf, 71, 261, (**2008**).
11. M. H. M. Olsson, C. R. Sondergaard , M. Rostkowski, J. H. Jensen, J. Chem. Theory Comput. 7, 525, (**2011**).
12. G. Wallin, S. C. Kamerlin, J. Aqvist, Nat. Commun. 4, 1733, (**2013**).
13. D.A. Case, T.A. Darden, T.E. Cheatham III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A.W. Götz, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, R.M. Wolf, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, R. Salomon-Ferrer, C. Sagui,

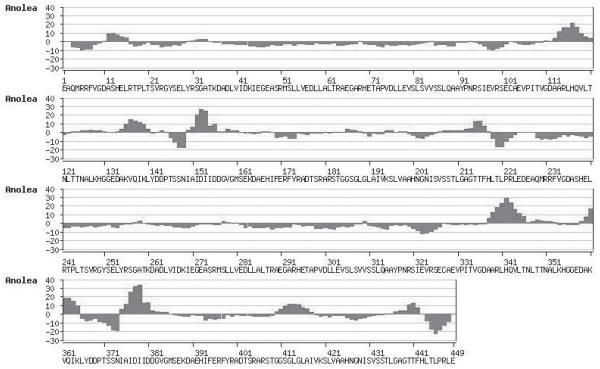
V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P.A. Kollman, AMBER 12, University of California, San Francisco, (**2012**).

14. W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, J. Chem. Phys. 79, 926, (**1983**).
15. K.L. Meagher, L.T. Redman, H.A. Carlson, J. Comput. Chem. 24, 1016, (**2003**).
16. O. Allnér, L. Nilsson, A. Villa, J. Chem. Theory Comput. 8, 1493, (**2012**).
17. J.P. Ryckaert, G. Ciccotti, H.J.C. Berendsen, J. Comput. Phys. 23, 327, (**1977**).
18. H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, J. Chem. Phys. 81, 3684, (**1984**).
19. D. M. York, T. A. Darden, L. G. Pedersen, J. Chem. Phys. 99, 8345, (**1993**).
20. H. D. Höltje, W. Sippl, D. Rognan, G. Folkers, Weinheim: Wiley-VCH, (**2008**).
21. F. Melo, E. Feytmans, J. Mol. Biol. 277, 1141, (**1998**).
22. A. E. MECHALY, N. SASSOON, J. BETTON, P. M. ALZARI, Plos Bio. 12, 1001776, (**2014**).
23. R. Dutta, M. Inouye, Trends. Biochem. Sci. 25, 24, (**2000**).
24. A. Marina, C. Mott, A. Auyzenberg, W. A. Hendrickson, C. D. Waldburger, J. Biol. Chem. 276, 41182, (**2001**).
25. A. M. Bilwes, C. M. Quezada, L. R. Croal1, B. R. Crane, M. I. Simon, Nat. Struct. Biol. 8, 353, (**2001**).

Supporting information



S1. Plots of ANOLEA1 mean force potential. a) Folding cis PhoR of C. pseudotuberculosis and b) Folding trans PhoR of C. pseudotuberculosis



1 F. Melo, E. Feytmans, J. Mol. Biol. 277, 1141, (1998).

Gleiciane L. Moraes[a], Guelber Cardoso[b], Cláudio Nahum Alves[a] & Jerônimo Lameira[b*]

[a] Laboratório de Planejamento de Fármacos, Instituto de Ciências Exatas e Naturais, Universidade Federal do Pará, CEP 66075-110 Belém, PA, Brazil.

[b] Laboratório de Planejamento de Fármacos, Instituto de Ciências Exatas e Naturais, Universidade Federal do Pará, CEP 66075-110 Belém, PA, Brazil; Instituto de Ciências Biológicas, Universidade Federal do Pará, CEP 66075-110 Belém, PA, Brazil

*E-mail: lameira@ufpa.br