Computational Biology and Chemistry 49 (2014) 1–6

Contents lists available at ScienceDirect

Computational Biology and Chemistry

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

Research Article

Collective variable driven molecular dynamics to improve protein–protein docking scoring

Diego Masone[a,*], Solène Grosdidier[b]

[a] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Instituto de Ciencias Básicas (ICB), Universidad Nacional de Cuyo (UNCUYO), Mendoza, Argentina
[b] Integrative Biomedical Informatics Group, Research Programme on Biomedical Informatics (GRIB), Hospital del Mar Research Institute (IMIM), DCEXS, Universitat Pompeu Fabra, Barcelona, Spain

ARTICLE INFO

Article history:
Received 22 October 2013
Received in revised form
28 December 2013
Accepted 28 December 2013

Keywords:
Protein–protein
Docking
Scoring
Collective variable
Molecular dynamics

ABSTRACT

In biophysics, the structural prediction of protein–protein complexes starting from the unbound form of the two interacting monomers is a major difficulty. Although current computational docking protocols are able to generate near-native solutions in a reasonable time, the problem of identifying near-native conformations from a pool of solutions remains very challenging. In this study, we use molecular dynamics simulations driven by a collective reaction coordinate to optimize full hydrogen bond networks in a set of protein–protein docking solutions. The collective coordinate biases the system to maximize the formation of hydrogen bonds at the protein–protein interface as well as all over the structure. The reaction coordinate is therefore a measure for docking poses affinity and hence is used as scoring function to identify near-native conformations.

1. Introduction

As most biological processes involve macromolecular complexes, identifying and characterizing protein interactions as well as the network they create, is critical for understanding the molecular mechanisms within the cell. Protein–protein interactions are fundamental in most cellular processes as for example DNA replication or signal transduction. Knowledge of the structure and properties of protein–protein complexes is essential to understand how proteins function within the cell, in order to identify new targets for therapeutic applications and develop new approaches for drug discovery. Hence, solving the structure of protein–protein complexes might provide the basis for understanding how a biological signal is transmitted or how a biological function is performed (Smith and Sternberg, 2002).

Many docking algorithms rely on the divide-and-conquer strategy (Luo et al., 2010): first, an initial sampling of the configurational space of the interacting proteins is performed by an efficient algorithm specialized on generating docking candidates (or solutions), typically rigid-body or coarse-grained, based and optimized through Fast Fourier Transform (FFT). Second, a scoring step is performed to rank the candidates according to scoring functions that can offer different levels of complexity (Halperin et al., 2002). However, the arduousness of sampling and scoring is not equal, while modern supercomputers allow for very good sampling of the configurational space between two proteins, there are no efficient and accurate methods for refinement and scoring yet. Consequently, even though near-native poses could be generated, it is still extremely difficult to distinguish them from a pool of solutions, making docking protocols to produce significant amounts of false positives (Gabb et al., 1997; Chen et al., 2003). Therefore, efficient computational docking highly depends on the accuracy of the energy functions used to evaluate the strength of binding candidates.

Unfortunately, because of bad protein modeling (missing ions, heteroatoms or unrealistic inter-protein contacts generated by rigid-body docking) scoring is currently a serious issue. The energy function used for scoring analyzes the conformation of both proteins in complex molecule and outputs a value representing a total energy. This number is meaningless alone, but when used in a relative way to compare the evolution of the total energy along the conformational changes that the proteins suffered, it gives an idea of how stable the different docking conformations are. Then, computational studies of protein associations from an energetic point of view, are also important to comprehend their essential principles and thus to improve protein interaction modeling. Energetic landscapes represent at different smoothness the natural behavior of protein–protein interactions. Their knowledge

* Corresponding author at: Padre Jorge Contreras 1300, Parque General San Martín, Mendoza 5500, Argentina. Tel.: +54 261 4236003.
E-mail addresses: dmasone@fing.uncu.edu.ar, diegomasone@yahoo.com.ar (D. Masone).

1476-9271/$ – see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.compbiolchem.2013.12.003

allow to predict stable configurational and conformational spaces, such as the binding funnel. Unfortunately for mathematical simplifications, protein binding energy landscapes are extremely rugged surfaces (Camacho and Vajda, 2001; Hagler et al., 1974a; Hagler and Lifson, 1974b). Since interacting forces upon protein binding act at different length-scales, the spatial frequency spectrum shows rapid variations and though generates many wells of local minima. For this reason, many times search algorithms are confined to a small portion of the sampling space (Phillips et al., 2001), since characterizing every minima of the potential energy landscape of a geometry is a problem with tens to thousands of degrees of freedom (Crescenzi et al., 1998; Atkins and Hart, 1999; Calland, 2003). However, the energy funnel shape is generally observed in the vicinities of the bound form of protein–protein complexes, once optimal interface residue conformations have been reached (Camacho and Vajda, 2001). In general, to test the performance of force-fields in docking applications, the native complex and the lowest energy near-native solution generated by a docking approach are compared energetically (Dixon, 1997; Totrov and Abagyan, 1997). However, this procedure is artificial and may lead to incorrect conclusions (Verkhivker et al., 2000). In a complete docking protocol the possibility to correctly identify near-native conformations as the lowest energy ones stands on the ability of the sampling process to generate them (Diller and Verlinde, 1999). The gradually narrowing of the energetic landscape has a physical meaning in protein–protein docking: there are large amounts of unfavorable high energy poses while only a small part of the configurational space is energetically stable.

Among macromolecular biological interactions hydrogen bond networks are specially interesting due to their key role in protein 3D structure, and as a consequence, in molecular recognition specificity and protein function (Morozov et al., 2004). Thus the appropriate description of their energetics is of great interest in the fields of protein–protein docking and protein folding. In computer simulations the challenging problem emerging from using crystallographic protein structure from the PDB (www.pdb.org) is to position every hydrogen atom. Indeed, the optimization of full hydrogen bond networks requires force-fields to properly describe every possible hydrogen configurations and their interaction energies (Masone et al., 2013). But as observed in previous studies, molecular mechanics force fields show poor accuracy in describing hydrogen bond physics (Fabiola et al., 2002; Hu et al., 2003; Lii and Allinger, 1994, 1998; Morozov et al., 2004; Masone et al., 2012).

Molecular dynamics is the tool by excellence to exhaustively explore the protein potential energy landscape while simulating its flexibility. Fully flexible relaxation tends to increase the amount of recovered native contacts among sets of docking poses (Król et al., 2007a, 2007b). However, it is impossible to assure that extensive molecular dynamics simulations will result in good conformers suitable for docking. Moreover, determining the most important motions for binding purposes and then performing docking experiments may result in equivalent solutions as flexible docking (Cozzini et al., 2008). As pointed out by Alonso and collaborators (Alonso et al., 2006) molecular dynamics have shown to accurately, although expensively, refine a few selected candidates from a previous fast docking stage used to sample large configurational spaces. The full atomistic description in long, and though useful, time scales are still beyond classical molecular dynamics simulations for most biological systems due to the small femtosecond time steps needed for energy conservation. Moreover, crystallographic monomers of a protein–protein complex cannot provide enough information on how interface rearrangements will occur. Hence, in these cases classical molecular dynamics can achieve only a limited phase-space exploration (Tiwary and van de Walle, 2013).

Remarkably, collective variable driven molecular dynamics have shown to adequately reproduce complex conformational changes in biomolecules by accelerating rare events (Fiorin et al., 2013). When biasing the system with a previously chosen collective reaction coordinate, molecular dynamics simulations may surpass intrinsic limitations of the physical model and a more efficient statistical sampling can be performed. However, it is usually difficult to select the proper collective variable that adequately describes the macroscopic phenomena (Fiorin et al., 2013; Laio and Parrinello, 2002; Laio and Gervasio, 2008; Kumar et al., 1996).

The purpose of this work is to propose a collective coordinate to optimize hydrogen bond networks in protein–protein systems by driving molecular dynamics simulations. The collective coordinate is then a measure of the hydrogen bond formation in each docking solution and though is used as a scoring function.

Table 1
Selected complexes from Vakser et al. decoy. Complexi: pdb code of co-crystallized structure. Rec.ii: pdb code of unbound receptor structure. RMSDiii: C_alpha rmsd of unbound receptor and co-crystallized structure [Å]. Lig.iv: pdb code of unbound ligand structure. RMSDv: C_alpha rmsd of unbound ligand and co-crystallized structure [Å]. RMSDvi: The ligand RMSD of the best near-native solution [Å]. HITSvii: The number of near-native solution kept in each decoy set. H-bondviii: The ranking position of the first near-native conformation identified.

Complexi	Rec.ii	RMSDiii	Lig.iv	RMSDv	RMSDvi	Hitsvii	H-bondviii
1bvn	1hx0	0.63	1ok0	0.42	2.24	10	1
1tmq	1jae	0.77	1b1u	1.42	2.07	10	19
1ugh	1akz	0.61	1ugi	2.60	2.86	10	9
1xd3	1uch	2.45	1yj1	2.73	3.64	10	5
3sic	1sup	0.34	3ssi	0.78	3.54	10	2

2. Materials and methods

We used the DockGround (Liu et al., 2008) set of protein–protein solutions generated by Vakser and collaborators, freely available on-line (http://dockground.bioinformatics.ku.edu/) that provides 100 non-native and at least one near-native (ligand RMSD < 5 Å) solution generated by GRAMM-X (Tovchigrechko and Vakser, 2006) docking scan per complex for a total of 61 complexes. To select our complexes we chose three (1xd3, 1ugh, 3sic) that were not included in a previous study in hydrogen bond network optimizations (Masone et al., 2012) using the software PELE (Borrelli et al., 2005), but still containing 10 near-native solutions in the DockGround decoy. Other two complexes that did were studied before (1bvn, 1tmq) were also selected for validation purposes (see Table 1). In all of them the condition that the ligand RMSD to the crystal reference is 5 Å or less for at least one of the poses in the decoy was fulfilled. Previous studies in protein–protein (Masone et al., 2012) and protein–ligand interactions (Borrelli et al., 2010) show that refinement techniques can only return near-native top scores if at least one of the poses in the decoy is close enough to the crystallographic reference. Fig. 1 shows the configurational space explored by the docking scan for the 1bvn protein–protein system.

We performed molecular dynamics in GROMACS 4.5.5 (Hess et al., 2008) patched with PLUMED 1.3 (Bonomi et al., 2009) to use a collective reaction coordinate in order to drive the system into the formation of hydrogen bonds. The PLUMED code provides a variety of different collective variables to performed free-energy calculations. The collective variable we used, available in PLUMED 1.3, counts the number of intra-molecular hydrogen bonds between a group donors and acceptors and it is defined as follows:

$$s = \sum_{ij} \frac{1-(d_{ij}/r_0)^n}{1-(d_{ij}/r_0)^m} \tag{1}$$

where i counts over the group of donors and j over the group of acceptors. For d_{ij} distance calculations all donor–acceptor pairs were included and the user defined values where set to $r_0 = 2.5$, $n = 6$ and $m = 12$. As a general rule, the two monomers of a

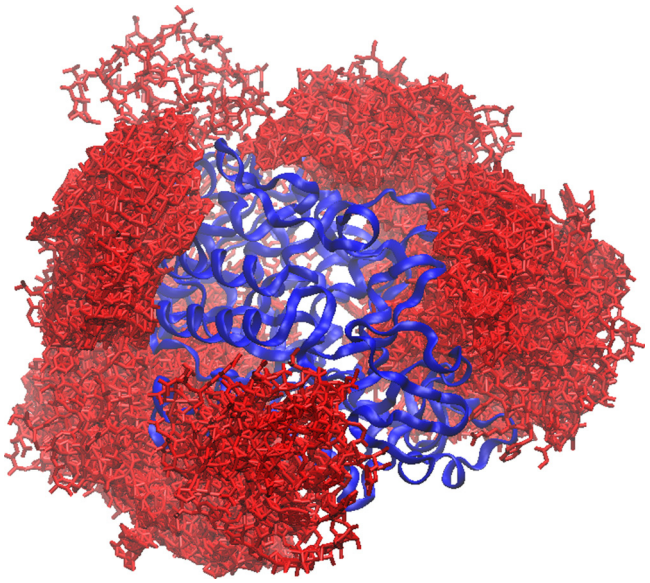

Fig. 1. The alpha-amylase is shown as blue ribbon while the 100 docking positions of its microbial inhibitor are displayed in red sticks all around its surface (PDB code: 1bvn).

protein–protein docking solution were defined as donor/acceptor groups so that all hydrogen bonds (not only the ones located at the protein–protein interface) are included in the collective variable.

All protein–protein docking poses were solvated in explicit SPC water, minimized and equilibrated. All bond-lengths were constrained using the sixth-order LINCS (Hess et al., 1997) algorithm, for an integration time step of 2fs. Simulations were performed in the NVT ensemble using Berendsens' temperature coupling scheme (Berendsen et al., 1984). Molecular dynamics runs including the collective variable for hydrogen bond optimizations were performed for the 100 poses for each protein–protein complex. Simulation time varied according to the reaction coordinate convergence, but was in all cases around 1 ns (an overnight simulation time in a current eight-core processor workstation). Protein figures were generated with VMD (Humphrey et al., 1996).

3. Results and discussion

The major limitation of collective variable based dynamics is the small amount of collective coordinates that can be used to avoid prohibitive computational times. On top of this, the *a priori* intuitive choice of the reaction coordinate may introduce unrealistic biases on how events of interest occur (Abrams and Vanden-Eijnden, 2010), thus choosing a correct set of collective variables remains an unsolved problem. In this work, we propose and describe a collective coordinate based on hydrogen bond optimizations to score protein–protein docking poses. By using molecular dynamics driven by this collective variable, we are able to optimize full hydrogen bond networks in energetic terms and to identify near-native poses. We tested the method in a group of decoys of unbound structures from a standard benchmark containing at least one near native solution. After the molecular dynamics based H-bond optimization we could score near-native solutions among the first poses, Table 1 lists the studied protein–protein complexes and its last column indicates the H-bond reaction coordinate performance in terms of near-native ranking positions.

Additionally, Table 2 shows the ranking positions of the first 10 near-native solutions for all complexes. The method directly allows for additional restraints that might be included in more complex reaction coordinates in order to improve biasing and to sample more efficiently, (*i.e.*, low-frequency normal modes). The method can be as well applied to protein–ligand systems.

Table 2
Ranking position of the first 10 near-native solutions (S1 to S10) identified for each complex studied.

Complex	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
1bvn	1	2	3	4	5	8	12	15	16	43
1tmq	19	44	54	57	60	65	79	101	106	107
1ugh	9	23	26	34	43	70	73	85	91	95
1xd3	5	13	19	22	30	38	44	76	83	84
3sic	2	11	18	28	30	34	38	61	65	103

Table 3
Top 1 ranking for different protein–protein docking + scoring methods.

Complex	ZDOCK 3.0.2	pyDOCK	Our method
1bvn	3	159	1
1tmq	9	1	19
1ugh	–	–	9
1xd3	7	12	5
3sic	–	–	2

Table 3 addresses a comparison between ZDOCK + ZRANK (Hwang et al., 2010a) and FTDOCK (Gabb et al., 1997) + pyDOCK (Cheng et al., 2007), two docking + scoring methods that have shown excellent results in the CAPRI competition, a blind test of protein docking algorithms that predicts the complex structure from the crystal structures of two interacting proteins. Complexes 3sic and 1ugh are not included in benchmark 4.0 (Hwang et al., 2010b) and hence were not studied by most docking and scoring protocols.

Due to the conformational space complexity in protein–protein interfaces, collective variable driven molecular dynamics offers a unique alternative to accelerate molecular dynamics simulations and to sample rare events occurring at long timescales in the all-atom space with explicit solvent. The application of this technique to score docking solutions is a novel approach that relies on the definition of proper collective coordinates. For the 1bvn complex the near-native conformation is ranked as the highest value of the H-bond collective variable, though being identified as top1 (Fig. 2, top left corner). According to the definition of the collective reaction coordinate (see Eq. (1)) most favorable poses get highest ranks.

Although exceptional, this may indicate the ability of the collective coordinate to distinguish between conformations as an initial

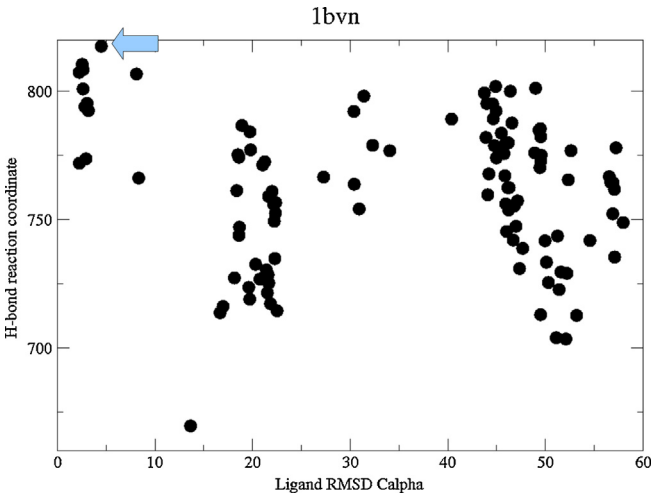

Fig. 2. Reaction coordinate scoring for the 100 docking poses for 1bvn. Near-native conformations are located in the top left corner with the highest values of the reaction coordinate. The arrow indicates TOP1 solution.

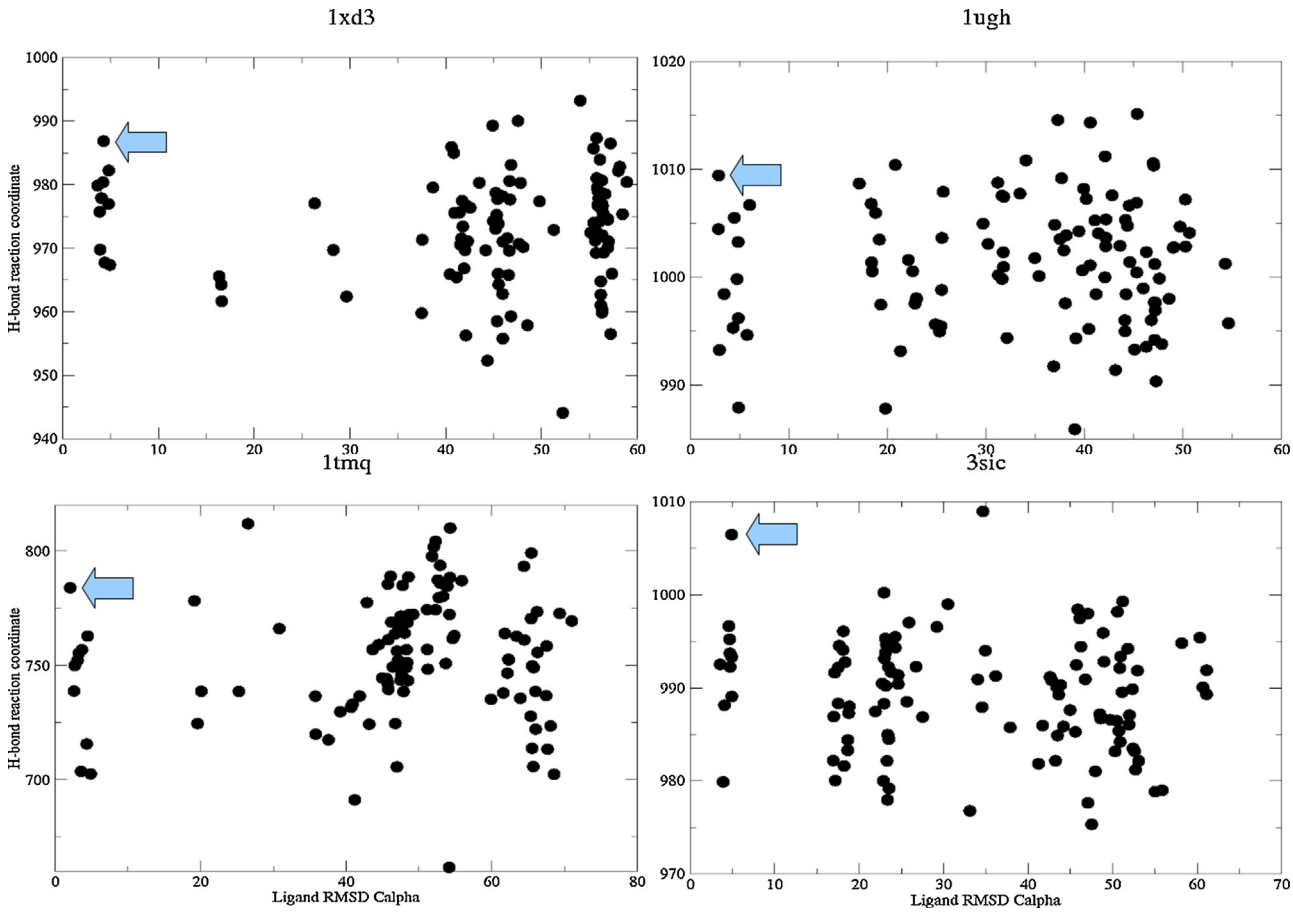

Fig. 3. Reaction coordinate scoring for 1tmq, 1ugh, 1xd3 and 3sic complexes. The arrow indicates in each case the best solution identified.

scoring step to select best candidates and to get rid of incorrect ones, showing a remarkable correlation between the interaction energy (score) and the RMSD to the reference crystal. As demonstrated in a previous work (Masone et al., 2012) the optimization cannot involve only the protein–protein interface region, but needs to include the entire system as the optimization of the interface alone would maximize the interaction energy and not the total energy, introducing false positives. Fig. 3 depicts the other four protein–protein complexes where near native conformations are included between the first ranked solutions.

For all the complexes studied there were no significant changes in the reaction coordinate after 0.8 ns (see Fig. 4). However, it is interesting to analyze how each one of the complexes behaves differently under the collective variable restrain in terms of the observed oscillations. In particular, complexes with a higher steady state value of the reaction coordinate (1ugh, 1xd3 and 3sic) show an underdumped oscillation, while the ones with lower steady state value of the reaction coordinate (1bvn and 1tmq) describe an overdumped oscillation. The reason for this is that complexes 1ugh, 1xd3 and 3sic are capable to form more H-bonds and then to reach their steady state reaction coordinate value in less time (rise times are considerably shorter). On the contrary, complexes 1bvn and 1tmq require more sampling and though higher conformational changes to reach a stationary sate with comparatively less amounts of H-bonds present in the structures.

As pointed out recently (Lindert et al., 2013) the Rosetta Protein Structure Refinement Protocol can distinguish native-like from non-native-like conformations in many cases but it is limited by conformational sampling for larger proteins. This highlights the complexity of protein structural refinement during molecular dynamics runs, which has a practical value separate from protein–protein docking and scoring. In this sense, Fig. 5 shows the protein–protein interface in one of ours 1tmq's docking solutions before and after simulations, where histidine residue HIS14 rotates during interface rearrangements.

We have presented here a scoring protocol to identify near-native poses in protein–protein interactions. The method involves no hierarchical steps but molecular dynamics simulations driven by a properly chosen collective variable. Our results show that

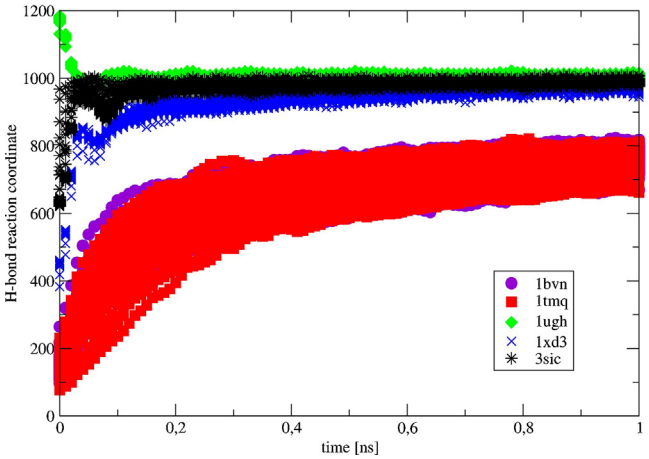

Fig. 4. Time evolution for H-bond optimization reaction coordinate over all complexes studied.

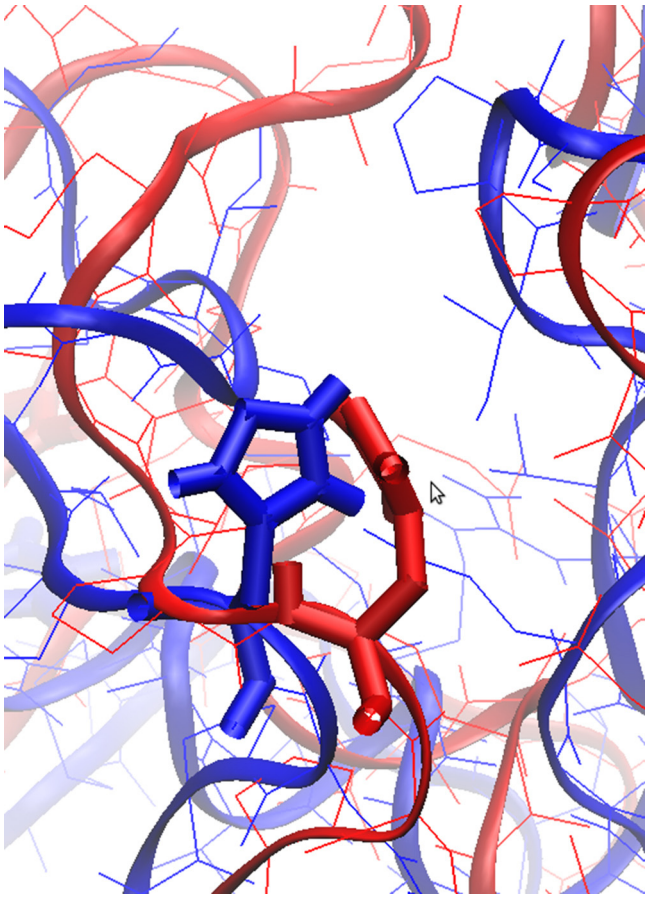

Fig. 5. Interface rearrangements in 1tmq. Comparison between initial configuration (blue) and final (red). Rotating histidine residue HIS14 is highlighted in bond representation in both cases. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

hydrogen bonds can be used as a reliable collective variable for scoring purposes in protein–protein docking. Moreover, hydrogen bonds are an interesting characteristic to distinguish near-native poses although more general collective reaction coordinates could result in even better scoring. For very large protein–protein systems a variable resolution procedure could be designed to alternate with coarse-grained schemes to reduce computational expenses (Samiotakis et al., 2010).

Acknowledgments

This work was supported by the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). The authors would like to thank the Instituto de Ciencias Básicas (ICB) – Universidad Nacional de Cuyo (UNCuyo) and Universidad Nacional de Córdoba (UNC) for supercomputing time.

References

Abrams, C.F., Vanden-Eijnden, E., 2010. Large-scale conformational sampling of proteins using temperature-accelerated molecular dynamics. Proceedings of the National Academy of Sciences of the United States of America 107 (11), 4961–4966.

Alonso, H., Bliznyuk, A.A., Gready, J.E., 2006. Combining docking and molecular dynamic simulations in drug design. Medicinal Research Reviews 26 (5), 531–568.

Atkins, J., Hart, W.E., 1999. On the intractability of the protein folding with finite alphabet of amino-acids. Journal of Computational Biology 5 (3), 423–466.

Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A., Haak, J.R., 1984. Molecular dynamics with coupling to an external bath. Journal of Chemical Physics 81, 3684–3690.

Bonomi, M., Branduardi, D., Bussi, G., Camilloni, C., Provasi, D., Raiteri, P., Donadio, D., Marinelli, F., Pietrucci, F., Broglia, R.A., Parrinello, M., 2009. PLUMED: a portable plugin for free energy calculations with molecular dynamics. Computer Physics Communications 180, 1961.

Borrelli, K.W., Cossins, B., Guallar, V., 2010. Exploring hierarchical refinement techniques for induced fit docking with protein and ligand flexibility. Journal of Computational Chemistry 31, 1224–1235.

Borrelli, K.W., Vitalis, A., Alcantara, R., Guallar, V., 2005. Protein energy landscape exploration. A novel Monte Carlo technique, implications on camphor ligand binding. Journal of Chemical Theory and Computation 6, 1304–1311.

Calland, P.Y., 2003. On the structural complexity of a protein. Protein Engineering 16 (2), 76–86.

Camacho, C., Vajda, S., 2001. Protein docking along smooth association pathways. Proceedings of the National Academy of Sciences of the United States of America 98, 10636–10641.

Chen, R., Mintseris, J., Janin, J., Weng, Z., 2003. A protein–protein docking benchmark. Proteins 52, 88–91.

Crescenzi, P., Goldman, D., Papadimitriou, C.H., Piccolboni, A., Yannakakis, M., 1998. On the complexity of protein folding. Journal of Computational Biology 5 (3), 423–466.

Diller, D.J., Verlinde, C.L., 1999. A critical evaluation of several global optimization algorithms for the purpose of molecular docking. Journal of Computational Chemistry 20, 1740–1751.

Cheng, T.M., Blundell, T.L., Fernandez-Recio, J., 2007. pyDock: electrostatics and desolvation for effective scoring of rigid-body protein–protein docking. Proteins 68, 503–515.

Cozzini, P., Kellogg, G.E., Spyrakis, F., Abraham, D.J., Costantino, G., Emerson, A., Fanelli, F., Gohlke, H., Kuhn, L.A., Morris, G.M., Orozco, M., Pertinhez, T.A., Rizzi, M., Sotriffer, C.A., 2008. Target flexibility: an emerging consideration in drug discovery and design. Journal of Medicinal Chemistry 51 (20), 6237–6255.

Dixon, J.S., 1997. Evaluation of the CASP2 docking section. Proteins 29 (1), 198–204.

Fabiola, F., Bertram, R., Korostelev, A., Chapman, M.S., 2002. An improved hydrogen bond potential: Impact on medium resolution protein structures. Protein Science 11 (6), 1415–1423.

Fiorin, G., Klein, M.L., Hénin, J., 2013. Using collective variables to drive molecular dynamics simulations. Molecular Physics, 1–18.

Gabb, H.A., Jackson, R.M., Sternberg, M.J.E., 1997. Modelling protein docking using shape complementarity, electrostatics and biochemical information. Journal of Molecular Biology 272, 106–120.

Hagler, A.T., Huler, E., Lifson, S., 1974a. Energy functions for peptides and proteins. i. Derivation of a consistent including the hydrogen bond from amide crystals. Journal of the American Chemical Society 96 (17), 5319–5327.

Hagler, A.T., Lifson, S., 1974b. Energy functions for peptides and proteins. ii. The amine hydrogen bond and calculation of amide crystal properties. Journal of the American Chemical Society 96 (17), 5327–5335.

Halperin, I., Ma, B., Wolfson, H., Nussinov, R., 2002. Principles of docking: an overview of search algorithms and a guide to scoring functions. Proteins 47, 409–443.

Hess, B., Bekker, H., Berendsen, H.J.C., Fraaije, J.G.E.M.J., 1997. LINCS: a linear constraint solver for molecular simulations. Computers and Chemistry 18 (12), 1463–1472.

Hess, B., Kutzner, C., van der Spoel, D., Lindahl, E., 2008. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. Journal of Chemical Theory and Computation 4, 435–447.

Hu, H., Elstner, M., Hermans, J., 2003. Comparison of a QM/MM force field and molecular mechanics force fields in simulations of alanine and glycine "dipeptides" (Ace-Ala-Nme and Ace-Gly-Nme) in water in relation to the problem of modeling the unfolded peptide backbone in solution. Proteins 50, 451–463.

Humphrey, W., Dalke, A., Schulten, K., 1996. VMD – visual molecular dynamics. Journal of Molecular Graphics 14 (1), 33–38.

Hwang, H., Vreven, T., Pierce, B., Hung, J., Weng, Z., 2010a. Performance of ZDOCK and ZRANK in CAPRI round 13–19. Proteins 78 (15), 3104–3110.

Hwang, H., Vreven, T., Janin, J., Weng, Z., 2010b. Protein–protein docking benchmark version 4.0. Proteins 78 (15), 3111–3114.

Król, M., Chaleil, R.A., Tournier, A.L., Bates, P.A., 2007a. Implicit flexibility in protein docking: cross-docking and local refinement. Proteins 69 (4), 750–757.

Król, M., Tournier, A.L., Bates, P.A., 2007b. Flexible relaxation of rigid-body docking solutions. Proteins 68 (1), 159–169.

Kumar, S., Payne, P.W., Vasquez, M., 1996. Method for free-energy calculations using iterative techniques. Journal of Computational Chemistry 17, 1269–1275.

Laio, A., Gervasio, F.L., 2008. Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. Reports on Progress in Physics 71, 126601.

Laio, A., Parrinello, M., 2002. Escaping free-energy minima. Proceedings of the National Academy of Sciences of the United States of America 99 (20), 12562–12566.

Lii, J.H., Allinger, N.L., 1994. Directional hydrogen bonding in the MM3 force field. I. Journal of Physical Organic Chemistry 7, 591–609.

Lii, J.H., Allinger, N.L., 1998. Directional hydrogen bonding in the MM3 force field. II. Journal of Computational Chemistry 19, 1001–1016.

Lindert, S., Meiler, J., McCammon, J.A., 2013. Iterative molecular dynamics—Rosetta protein structure refinement protocol to improve model quality. Journal of Chemical Theory and Computation 9 (8), 3843–3847.

Liu, S., Gao, Y., Vakser, I.A., 2008. Dockground protein–protein docking decoy set. Bioinformatics 24 (22), 2634–2635.

Luo, W., Pei, J., Zhu, Y., 2010. A fast protein–ligand docking algorithm based on hydrogen bond matching and surface shape complementarity. Journal of Molecular Modeling 16 (5), 903–913.

Masone, D., Ciocco Aloia, F., Del Pópolo, M.G., 2013. H-bond refinement for electron transfer membrane-bound protein–protein complexes: cytochrome c oxidase and cytochrome c552. Computational Biology and Chemistry 47, 31–36.

Masone, D., Cabeza de Vaca, I., Pons, C., Fernández Recio, J., Guallar, V., 2012. H-bond network optimization in protein–protein complexes: are all-atom force field scores enough? Proteins 80 (3), 818–824.

Morozov, A.V., Kortemme, T., Tsemekhman, K., Baker, D., 2004. Close agreement between the orientation dependence of hydrogen bonds observed in protein structures and quantum mechanical calculations. Proceedings of the National Academy of Sciences of the United States of America 101 (18), 6946–6951.

Phillips, A.T., Rosen, J.B., Dill, K.A., 2001. Convex global underestimation for molecular structure prediction. From Local to Global Optimization: Nonconvex Optimization and Its Applications 53, 1–18.

Samiotakis, A., Homouz, D., Cheung, M.S., 2010. Multiscale investigation of chemical interference in proteins. Journal of Chemical Physics 132, 175101.

Smith, G.R., Sternberg, M.J.E., 2002. Prediction of protein–protein interactions by docking methods. Current Opinion in Structural Biology 12, 28–35.

Tiwary, P., van de Walle, A., 2013. Accelerated molecular dynamics through stochastic iterations and collective variable based basin identification. Physical Review B 87, 094304.

Totrov, M., Abagyan, R., 1997. Flexible protein–ligand docking by global energy optimization in internal coordinates. Proteins 1, 215–220.

Tovchigrechko, A., Vakser, I.A., 2006. GRAMM-X public web server for protein–protein docking. Nucleic Acids Research 34, 310–314.

Verkhivker, G.M., Bouzida, D., Gehlhaar, D.K., Rejto, P.A., Arthurs, S., Colson, A.B., Freer, S.T., Larson, V., Luty, B.A., Marrone, T., Rose, P.W., 2000. Deciphering common failures in molecular docking of ligand–protein complexes. Journal of Computer-Aided Materials Design 14, 731–751.