J Mol Model (2009) 15:1013–1025
DOI 10.1007/s00894-008-0365-1

ORIGINAL PAPER

Analysis of oligomeric proteins during unfolding by pH and temperature

Pradip Bhattacharya · Tamil Ganeshan · Soumiyadeep Nandi · Alok Srivastava · Prashant Singh · Mohommad Rehan · Reshmi Rashkush · Naidu Subbarao · Andrew Lynn

Received: 25 January 2008 / Accepted: 22 September 2008 / Published online: 11 February 2009

Abstract During thermal transition and variation of pH, structural properties of 35 proteins and their complexes (bound with substrate and co-factor) were analyzed in detail. During pH alteration, these proteins were shown to have substantial differences in conformations. pH conformers were analyzed in detail. Free energy and other energy parameters were also estimated for these proteins at various pH and temperatures. Detailed structural analysis and binding interfaces of various substrates, inhibitors and cofactor of these proteins were also investigated using docking and molecular dynamic simulation.

Keywords Analysis of structure and conformation at various pH · Binding free energy · Docking · Electrostatic charge of protein · Free energy · Molecular dynamic simulation · Oligomeric proteins · Unfolding of proteins · Variation of pH and temperature

Introduction

Protein folding is inherently a heterogeneous process because of the very large number of microscopic pathways that connect the myriads of unfolded conformations to the unique conformation of the native state in the ensemble. Protein folding/unfolding is a highly cooperative process. It has been shown that the folding/unfolding of small globular proteins occurs via a two-state process, whereas the folding/unfolding of larger proteins (>100 amino acids) is complex and often involves the formation of intermediate(s) [1–15]. The most thorough investigations of protein folding and stability have been done with unusually small proteins, which are folded into single domains and display simple two-state unfolding processes. Report of analysis of conformation of proteins during pH alteration by computation has not been found, even though some reports of conformation changes of proteins in the experimental set-up were documented earlier using circular dichroism(CD), UV-spectroscopy, differential scanning calorimetry(DSC), and infra-red spectroscopy [12, 13].

During thermal transition and variation of pH, the structural properties of 35 proteins and their complexes (bound with substrate and co-factor) were analyzed in detail. During pH alteration, these proteins were shown to have substantial differences in conformations. Free energy and other energy parameters were also estimated for these proteins at various pH and temperatures. Detailed structural analysis and binding interfaces of various substrates, inhibitors, and cofactor of these proteins were also investigated using docking. Molecular dynamic simulation of proteins was performed for 1 fentosecond.

Computational methods

Electrostatic charges of protein (alone or complexed with cofactor and substrate) was determined as described before [8, 16, 17]. This procedure automated addition of a limited

Electronic supplementary material The online version of this article (doi:10.1007/s00894-008-0365-1) contains supplementary material, which is available to authorized users.

P. Bhattacharya (✉)
School of Life Sciences, Jawaharlal Nehru University,
New Delhi 110067, India
e-mail: pradip.bhattacharya13@gmail.com

T. Ganeshan · S. Nandi · A. Srivastava · P. Singh · M. Rehan · R. Rashkush · N. Subbarao · A. Lynn
School of Information Technology, Jawaharlal Nehru University,
New Delhi 110067, India

number of missing heavy atoms into biomolecular structures, estimation of titration states, protonation of biomolecules in a manner consistent with favorable hydrogen bonding, and assignment of charge and radius parameters from a variety of force fields, such as charmm22, amber99, and parse force fields. His, Asn and Gln sidechain χ angles were sampled via Monte Carlo for optimum hydrogen-bonding conformation in water–water and water–protein hydrogen bonding network [8, 16]. The calculation of Poisson-Boltzman electrostatic charges [16] was based on the partial equalization of orbital electronegativities (PEOE) procedure. In the PEOE procedure, orbital electronegativities were linked to partial atomic charges (q) by a polynomial expansion [16].

$$q = a + b \cdot q + c \cdot q2 + d \cdot q3 \quad (1)$$

The coefficients a, b, c, and d were optimized using gas phase data on ionization potentials and electron affinities. PEOE algorithm [4] had been optimized to obtain better agreement between theoretical and experimental solvation energies for a set of small molecules including the polar amino acids. The relative free energy difference ($\Delta\Delta G_{AB}$) between the two states (ΔG_A, ΔG_B) was the basis of their titration curves.

$$\Delta\Delta G_{AB} = \Delta G_A - \Delta G_B \quad (2)$$

The relative free energy ($\Delta\Delta G_{AB}$) of conformer A and B for the absolute stability was given by Eq. 3.

$$\Delta\Delta G_{AB}(pH) = \Delta G_A(pH) - \Delta G_B(pH) \quad (3)$$

The following linear combination of empirical terms was used to calculate free energy (kcal mol^{-1}) at ionic strength from 0.01 M to 1.0 M as described before [5–7, 10].

$$\begin{aligned}\Delta G = {} & a.\Delta G_{vdw} + b.\Delta G_{solvH} + c.\Delta G_{solP} + d.\Delta G_{wb} \\ & + e.\Delta G_{hbond} + f.\Delta G_{Gel} + g.\Delta G_{kon} \\ & + h.T\Delta S_{mc} + k.T\Delta S_{sc} + l.\Delta G_{clash}\end{aligned} \quad (4)$$

In this expression (a…l) were relative weights of the different energy terms used for the free energy calculation. The bulk solvent was treated as a desolvation term that was continuously scaled with the burial of an atom and separated into contributions from hydrophobic (ΔG_{solvH}) and polar (ΔG_{solvP}) groups. Those water molecules that had a persistent interaction with groups of the protein, i.e., made more than two hydrogen bonds with the protein, were calculated explicitly in the ΔG_{wb} term. Van der Waals term, ΔG_{vdw}, was taken into account of experimental transfer energies from water to vapor. Hydrogen bonds were calculated on the basis of simple geometric considerations and their energy, ΔG_{hbond}, was inferred from protein engineering double mutant cycles. The electrostatic contribution to the free energy, ΔG_{el}, was calculated from a simple implementation of Coulomb's law, in which the dielectric constant was scaled with the burial of the bond under consideration. For protein complexes, an additional electrostatic contribution was calculated between atoms of different polypeptide chains, ΔG_{kon}. The entropic penalty for fixing the backbone in a given conformation, ΔG_{mc}, was derived from a statistical analysis of the ϕ-ψ distribution of a given amino acid as observed in a set of non-redundant high-resolution crystal structures. The entropy cost of fixing a side chain in a particular conformation, ΔS_{sc}, was obtained by scaling a set entropy parameters to the burial of the side chain. Finally, the ΔG_{clash} term provided a measure of the steric overlaps among atoms in the structure. Binding energy of the substrate, cofactor, and complex of two identical subunits was calculated as the difference between the total energy of the complex and total energy of subunit alone. Binding free energy of the substrate, cofactor, and complex of two identical subunits was calculated as the difference between the free energy of the complex and free energy of subunit alone. The free energy difference ($\Delta\Delta G_{D\text{-}N}$) of the native state ($\Delta G_N$) and the denatured state (ΔG_D) was calculated using Eq. 5 [5–7, 10].

$$\Delta\Delta G_{D-N} = \Delta G_D - \Delta G_N \quad (5)$$

The interaction(or adaptive) binding energy was analyzed using Eq. 6.

$$\Delta G_{binding} = \Delta G_{AB} - (\Delta G_A + \Delta G_B) = RTlnK_d \quad (6)$$

$\Delta\Delta G_{binding}$, ΔG_{AB}, ΔG_A, and ΔG_B were binding free energy, free energy for complex, free energy of A, and free energy of B, respectively. In Foldx program [18–21], the salt concentration was varied from 0.01 M to1 M and temperature was varied from 0 °C to 100 °C. The calculation of hydrogen bonds (in water), Van der Waals radii (using Skhake procedure), polar/nonpolar accessible surface area(ASA), charged ASA (using Shrake procedure) and volume (using Standard Voronoi procedure) was performed as described earlier [17].

The transition rate between free energy minima is controlled by the dynamics of passing through an unstable transition region determined by saddlepoints in the free energy surface. Accordingly, the rate is expected to follow Arrhenius form.

$$k_f = k_0 e^{-\beta \Delta G^+} (\beta = 1/k_B T) \quad (7)$$

$1/k_BT$ is the inverse temperature and ΔG^+ is the free energy difference between the unfolded and transition-

state ensembles. The exponential factor reflects the equilibrium population of the transition-state ensemble relative to unfolded ensemble and the prefactor, k_0, is the time scale associated with the dynamics of crossing the free energy barrier [14, 15, 22]. The folding parameters (k_f, k_u) of proteins in the ensemble were also determined from the following equations.

$$k_f = k_o . e^{-\beta \Delta G}{}_{D-N} \quad (8)$$

$$\Delta G_{N(\text{or } D)} = -RTlnK_d \quad (9)$$

$$K_d = k_u/k_f \quad (10)$$

K_d is the equilibrium dissociation constant, k_f is the folding rate constant, k_u is the unfolding rate constant, k_B is the Boltzman constant, and k_o is the prefactor or time scale associated with the protein molecule to cross the energy barrier [14, 15, 23].

The accession numbers of proteins of the Protein Data Bank (PDB) were mentioned in Table 1. The accession numbers of proteins of the Protein Model Data Bank were from PM0074680 to PM0074692, from PM0074716 to PM0074718, from PM0074741 to PM0074812, from PM0074861 to PM0074929 and from PM0074994 to PM0075006. All of these proteins and their complexes determined by docking were deposited in the data banks (PDB and PMDB) from this center. Modeling and docking were performed according to the published procedures [26–29]. Structural properties of proteins were determined from several programs (Supplementary Table 1) as described before [8, 16, 17, 20–22, 30–33]. Besides, Gromacs, VMD, Deepview, Molsoft (ICM), Rastop, Rasmol were also used to study the 3D models of both protein and their complexes. In the PDBsum program, proteins of PMDB [34] were loaded to generate various structural files for examination.

Molecular dynamic simulation (Gromacs 3.3.2 version) [35, 36] was performed with the steepest descent method. Steepest descents never converged to the machine precision Fmax < 1. A more popular FF99 force field and particle-mesh Ewald (PME) method were used to calculate the electrostatic interactions. One fentosecond was the limit for time scale simulation in a 32 processor with a 32 bit machine. This simulation was performed for 2I95(335 steps), 2I3K (54 steps), 2I98 (50 steps), 2I97 (97 steps), 2I3J (52 steps), 2I3N (20 steps), 2IB2 (144 steps), 2IB3 (64 steps), 2IB4 (39 steps), 2E0E (23 steps), 2E0F (26 steps), 2I8X (31 steps), 2I8H (28 steps), 2IG1 (78 steps), 2I8Y (85 steps), 2I8Z (19 steps), 2I90 (140 steps), and CD150-MVH complex PM0074783 (499 steps).

Results and discussion

Energy calculation of protein during thermal transition

The subunits of proteins of PDB and PMDB were listed in Table 1. The free energy (-ΔG) of unfolding and folding of proteins was determined as described before [2, 29] from 273°K to 373°K (Fig. 1) at ionic strength 0.05 M-1 M. The patterns of change in free energy (-ΔG) with respect to unfolding were similar in nature for all individual proteins (Fig. 1), when the temperature was raised from 273°K to 373°K at ionic strength 0.05 M. In all cases, free energy decreased with increasing temperature. As a model system, the patterns of energy parameters (such as conformational stability, change in free energy, total energy, electrostatic energy, Van der Waals' energy, mainchain entropy, sidechain entropy, etc.) of 2DXO protein during thermal denaturation were illustrated in Supplementary Table 2. The quantitative differences of each energy parameter were also different for different proteins during thermal transition (data not shown). During thermal transition of proteins, the energy differences were found to be in the range of 1–25 kcal mol^{-1}. Van der Waals' energy, solvation energy(polar), Van der Waals' clash, cis-bond energy, torsional clash, backbone clash and backbone hydrogen bond energy almost remained constant. All of these calculations of thermal denaturation were performed at pH 7 and 0.5 M ionic strength. No variation of torsion angles(ϕ, ψ, ω and chi1–5), bond angles, bond distances, distances among atom pairs and distance matrix were observed during thermal denaturation(Procheck and PDBsum analysis). Since unfolding of proteins during thermal denaturation involved the breaking of hydrogen bonds and salt bridges, the sidechain hydrogen bond energy was found to be lower at higher temperature and higher at lower temperature (Supplementary Table 2). However, solvation hydrophobic energy increased during increase of temperature (Supplementary Table 2). Thermal denaturation can be explained that the only flexible regions of proteins where hydrogen bond energies were lower at high temperature had more rotational degrees of freedom. During denaturation, these flexible regions of proteins where sidechain hydrogen bond energies were comparatively lower at higher temperature had apparently more rotational degrees of freedom that imposed more contsrain in the rigidity of the overall structure. Most of the residues of α-helix and β-strands retained the covalent bonds as such (apparently rigid cluster). From one to few residues of some proteins were shown to be flexible [22]. The thermal stability of homodimer of protein was discussed before [37].

Table 1 List of all models submitted in PDB and PMDB

Organism	Protein sequence	Protein name	PDB or PMDB ID	References
Gallus gallus (red jungle fowl)	XP_415327	LAAO	2I8X(PDB)	This lab
Mus musculus (house mouse)	AAH17599	LAAO	2I8Y(PDB)	This lab
Scomber japonicus(Eukaryota;Metazoa)	CAC00499	LAAO	2I8Z(PDB)	This lab
Crotalus atrox (western diamondback rattlesnake)	AAD45200	LAAO	2I90(PDB)	This lab
Homo sapiens	CAI54292	interleukin 4 induced protein1 (@LAAO)	2I8H(PDB)	This lab
Mus musculus(house mouse)	NP_034345	interleukin 4 induced protein1 (@LAAO)	2I8W(PDB)	This lab
Mus musculus(house mouse)	CAA46268	TK	2I8J(PDB)	This lab
Drosophila melanogaster (fruit fly)	BAA04489	TK	2I8I(PDB)	This lab
Xenopus laevis (African clawed frog)	BAB87808	TK	2I8M(PDB)	This lab
Sycon raphanus(marine sponge)	CAC14731	TK	2I8K(PDB)	This lab
Rubrobacter xylanophilus DSM 9941(bacteria)	YP_643312	DAAO	2DZH(PDB	This lab
Mus musculus(house mouse)	BAA01063	DAAO	2DXO(PDB)	This lab
Mycobacterium tuberculosis CDC1551(bacteria)	NP_336413	DAAO	2I3M(PDB)	This lab
Streptomyces coelicolor A3(2)(bacteria)	NP_630813	DAAO	2I95(PDB)	This lab
Oryctolagus cuniculus (rabbit)	P22942	DAAO	2I3K(PDB)	This lab
Rattus norvegicus (Norway rat)	NP_446078	DAAO	2I98(PDB)	This lab
Mycobacterium leprae(bacteria)	CAC30966	DAAO	2I97(PDB)	This lab
Homo sapiens(human)	NP_001908	DAAO	2I3J(PDB)	This lab
Rhodosporidium toruloides	CAA96323	DAAO	2DZG(PDB)	This lab
Danio rerio (zebrafish)	NP_999897	DAAO	2I3N(PDB)	This lab
Cricetulus griseus(Chinese hamster)	Q9Z302	DAAO	2IB2(PDB)	This lab
Cavia porcellus(domestic guinea pig)	CAA07616	DAAO	2IB3(PDB)	This lab
Nectria haematococca(fungi)	BAA00692	DAAO	2IB4(PDB)	This lab
Photobacterium sp. SKA34(bacteria)	ZP_01162949	DAAO	2EOE(PDB)	This lab
Nocardioides sp. JS614(bacteria	ZP_00659314	DAAO	2EOF(PDB)	This lab
Xanthomonas campestris(ATCC33913)(bacteria)	AAM42948	DAAO	PM0074690(PMDB)	This lab
Candida albicans SC5314(fungi)	XP_7116070	DAAO	PM0074682(PMDB)	This lab
Erythrobacter sp. NAP1(bacteria)	EAQ29060	DAAO	PM0074692(PMDB	This lab
Aspergillus fumigatus Af293(bacteria)	EAL91482	DAAO	PM0074685(PMDB)	This lab
Xanthomonas axonopodis(bacteria)	NP_642864	DAAO	PM0074684/86(PMDB)	This lab
Candida albicans SC5314(fungi)	XP_721563	DAAO	PM0074683(PMDB	This lab
Candidatus Pelagibacter ubique HTCC1062(bacteria)	YP_265971	DAAO	PM0074681(PMDB)	This lab
Xanthomonas oryzae KACC10331(bacteria)	YP_201779	DAAO	PM0074680(PMDB)	This lab
Roseovarius sp. 217(bacteria)	ZP_01037079	DAAO	PM0074774/691(PMDB)	This lab
Drosophila melanogaster	AAF25683	NOS	2I86(PDB)	This lab
Bombyx mori	NP_001036963	NOS	2I7L(PDB)	This lab
Candida boidinii	BAB12222	DAAO	PM0074687/88(PMDB	This lab
Anthrobacter protophormiae	AAP70489	DAAO	PM0074689(PMDB)	This lab
Synechococchoccus sp	ABB36398	DAAO	PM0074677(PMDB)	This lab
Homo sapiens	Q13291	CD150	2DZF(PDB)	This lab
Homo sapiens and measles virus		CD150-measles virus H protein complex	PM0074783(PMDB	This lab
Measles virus	AAC29443	Nucleoprotein	2IG1(PDB)	This lab
Measles virus	AAD29100	RNA polymerase	2DZ6(PDB)	This lab
Measles virus	AAF85679	Matrix protein	2IHH(PDB)	This lab

Effect of salt concentration on unfolding of proteins

When the ionic strength was increased from 0.01 to 0.5 M at 300°K, the conformational stability, molecular total energy, backbone H-bond energy, Van der Waals' energy, electrostatic energy, solvation(polar) energy, Van der Waals' clash, cis-bond energy and m-loop entropy remained unchanged, whereas side chain H-bond energy and backbone clash increased (data not shown). During unfolding from 273°K to 373°K at 0.05 M ionic strength (Fig. 1) of these proteins, torsion angles (ϕ, ψ, ω and Chi1–5) of each residues of a protein remained unchanged (data not shown).

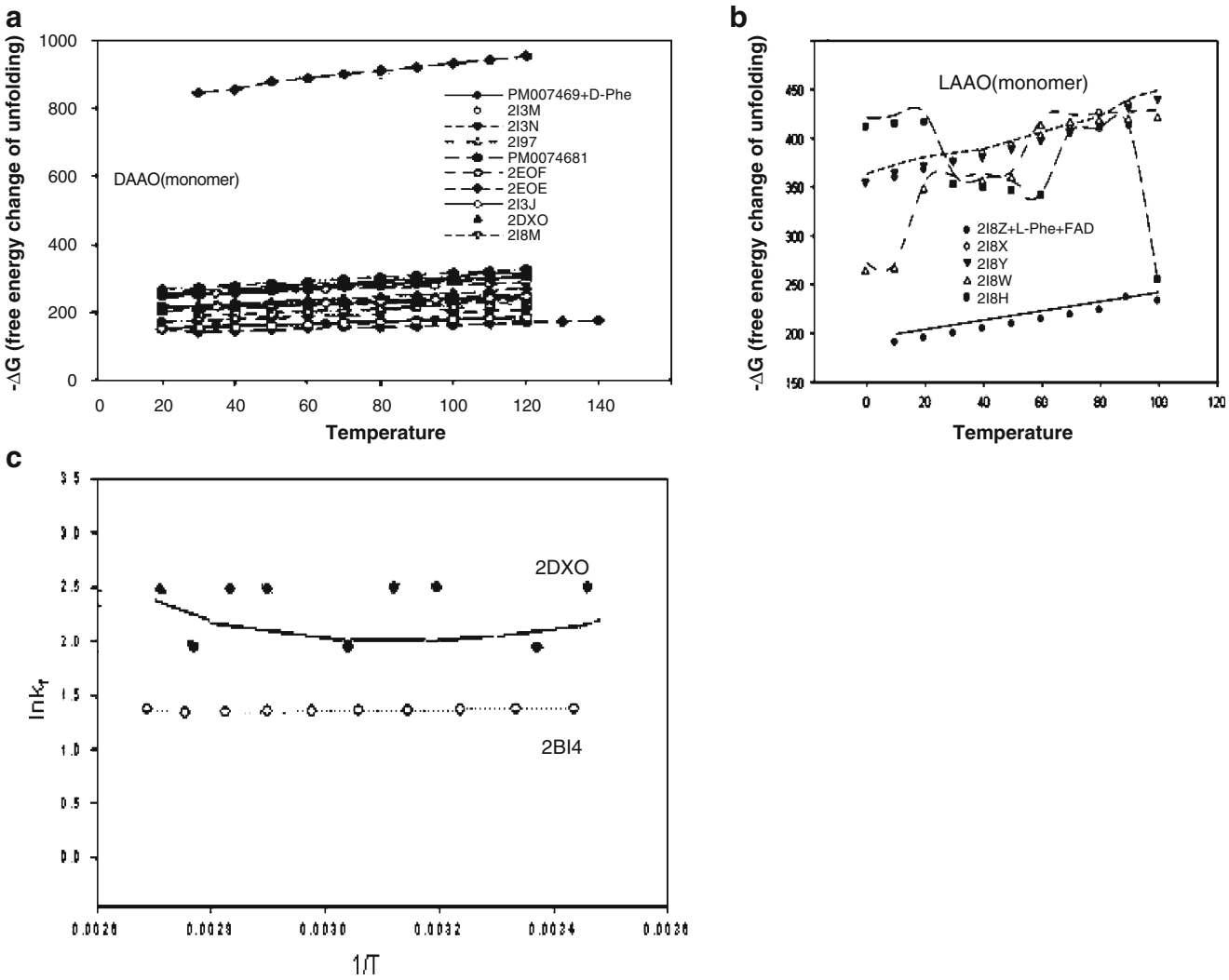

Fig. 1 Free energy vs temperature plot of proteins of DAAO(**a**) and LAAO(**b**) as well as $\ln k_f$ vs 1/T plot of DAAO(**c**)

Effect of pH on energy parameters, electrostatic charge, and conformation of proteins

The free energy and binding free energy of substrate, cofactor, and protein subunit (Tables 2, 3, 4, 5, Supplementary Table 3, and Supplementary Table 4) were calculated at ionic strength 0.05M and different pH (pH1-pH13). The free energy of protein subunit was shown to decrease after binding of substrate and cofactor (Table 3). Again, the free energy of protein was shown to decrease first and then increase with increasing pH, but the optimum pH values are different for each protein. The change of total electrostatic charge of some proteins with respect to pH was presented in Table 4. Although pK_a of most of the amino acids of proteins during pH alteration(pH1-pH13) remained unchanged, pK_a values of few amino acids (Table 5) of few proteins were found to be changed during pH variation. Proteins generally acquired more positive charges at pH1. At pH13, proteins generally obtained more negative charges. Proteins generally acquired lower (either positive or negative) charges near the vicinity of pH 7. Binding of substrate and cofactor into a single subunit was usually found to decrease the total electrostatic charge of the complex protein molecule. For example, total charge (+48.0e) of *C. boidini* DAAO (PM0074687) was reduced to -19.0e after binding with FAD and D-Trp.

The substantial differences in conformation and geometry of pH conformers (pH1, pH7, and pH13) were shown in Fig. 2 after superimposition of these pH conformers. The differential coloring (two colors) was performed to show differences in pH conformers (Fig. 2). For example, DAAO {*Roseovarius sp.* 217}(PM0074774) complexed with FAD and D-Trp exhibited change in conformation among few amino acids (blue color) (Glu60,Gln231, Asp270, Arg281, Glu321, Phe133, Vall64, Arg174, Pro167, Pro325, Pro138, Gln231, Lys84, Val81, Ser302, Asn319, Arg281, Gln36, Leu87 and Asp270) at pH1 and pH7, whereas human DAAO (2I3J) exhibited change in conformation among

Table 2 Folding parameters of some proteins and their complexes

Protein (PDB/PMDB)	lnkf (sec^{-1})	-ΔGD-N (kcal $mole^{-1}$)	k0 (sec^{-1})
2I3J	-6.412	58.65	0.0019
2DXO	-8.98	57.45	0.0593
2I3N	-11.69	56.55	0.0000791
2I98	-6.25	10.92	0.0019
2IB2	16.19	40.32	0.0019
PM0074682	7.35	40.8	-
PM0074684	15.2	31.05	-
PM0074687	-1.87	51.9	0.19
PM0074761	-6.38	94.8	0.0012
PM0074762	18.9	60.75	-
PM0074763	-10.8	75.6	0.000025
PM0074765	-23.6	57.95	-
PM0074766	18.9	62.19	-
PM0074767	11.6	85.2	-
PM0074768	-10.8	69.26	0.025
PM0074965	28.8	34.8	-

Table 4 Total electrostatic charge of protein subunit at three different pHs

Electrostatic charge(culoumb)			
Protein	pH 1.0	pH 7.0	pH 13.0
2I86	56.0 e	5.0 e	-46.0e
2I90	74.0 e	-9.0e	-55.0e
2IHG	47.0e	8.0e	-29.0e
2DZF	40.0e	7.0e	-37.0e
2I3J	45.0e	-4.0e	-31.0e
2I3N	43.0e	0.0e	-31.0e
2DXO	43.0e	2.0e	-31.0e
2DZ6	130.0e	nd	-105.0e
2E0E	50.0e	6.0e	-36.0e
2E0F	39.0e	5.0e	-3.0e
2I3K	43.0e	-1.0e	-32.0e
2I3M	35.0e	-4.0e	-20.0e
2I7L	52.0e	-1.0e	-52.0e

nd, not determined

many amino acids (blue color) (Fig. 2). The helices and strands were also different for three pH conformers (2I3J) in Procheck analysis (Fig. 2). Figures pH conformers of other proteins were not shown. The programs mentioned in the method section [5–7, 11, 18, 21–26, 34] were used in this study for analysis of pH conformers. The different conformations of the same molecule at different pH can be attributed to breakage of non-covalent bonds, viz. hydrogen bonds and saltbridges, ionic-ionic interactions and hydrophobic-hydrophilic interactions [22, 36–43]. The

Table 3 Free energy of protein complex as a function of pH

ΔG(kcal $mole^{-1}$)															
	G. gallus LAAO			*M. leprae* DAAO			*M. tuberculosis* DAAO			*C. porcellus* DAAO			*C. griseus* DAAO		
pH	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
0.00	-0.39	-0.44	-1.12	-0.00	-0.01	-0.26	-0.02	-0.02	-0.17	-0.08	-0.10	-0.41	-0.06	-0.07	-0.29
1.00	-5.74	-6.48	-15.48	-0.16	-0.39	-4.38	-0.62	-0.62	-2.92	-1.79	-2.01	-7.79	-1.49	-1.61	-5.59
2.00	-15.75	-17.37	-39.28	-1.27	-2.35	-12.16	-3.1	-3.43	-11.08	-5.94	-6.67	-21.55	-4.70	-5.33	-15.79
3.00	-33.55	-35.95	-79.69	-5.31	-7.11	-23.61	-8.48	-9.43	-26.74	-15.59	-16.76	-45.40	-11.11	-12.63	-33.34
4.00	-55.34	-58.23	-126.44	-11.96	-13.49	-35.07	-3.88	-16.22	-43.87	-30.21	-31.33	-76.72	-20.63	-22.69	-55.55
5.00	-68.95	-72.46	-153.71	-16.50	-17.48	-39.04	-16.17	-9.33	-51.61	-41.39	-43.04	-100.62	-28.20	-30.49	-71.15
6.00	-76.81	-80.41	-166.88	-19.90	-20.84	-41.70	-17.89	-21.48	-55.63	-48.63	-50.47	-115.41	-34.1	-36.45	-82.14
7.00	-80.88	-84.23	-170.90	-22.45	-23.39	-44.85	-19.78	-23.02	-57.79	-52.38	-54.02	-122.86	-36.82	-39.12	-88.14
8.00	-81.72	-84.31	-167.80	-24.00	-24.84	-47.29	-21.35	-24.13	-59.09	-53.78	-55.08	-125.78	-37.67	-39.79	-90.65
9.00	-82.36	-83.67	-164.65	-24.48	-25.24	-48.27	-22.01	-25.17	-61.88	-54.69	-55.72	-128.58	-38.1	-40.12	-92.99
10.00	-80.26	-80.22	-158.98	-21.77	-22.50	-44.02	-9.96	-24.52	-60.97	-52.37	-53.30	-125.36	-35.68	-37.36	-89.52
11.00	-71.79	-70.56	-140.56	-18.43	-19.40	-40.73	-18.96	-24.85	-60.74	-46.88	-47.17	-113.70	-30.5	-31.12	-78.46
12.00	-62.30	-59.62	-117.91	-19.10	-20.27	-44.85	-22.19	-29.16	-68.72	-45.37	-44.64	-108.58	-28.65	-27.60	-71.93
13.00	-52.03	-49.23	-95.88	-19.88	-21.21	-45.31	-25.49	-33.51	-76.03	-45.35	-44.35	-106.90	-28.21	-26.71	-69.08

For each protein, column 1 denoted single subunit; column 2 denoted single subunit bound with FAD and substrate, and column 3 denoted complex of two identical subunits bound with FAD and substrate amino acid. *G. gallus* LAAO: subunit(2I8X, column1), subunit complexed with FAD and L-Trp(PM0074797, column 2) and two subunits complexed with FAD and L-Tyr(PM0074887, column 3); *M. leprae* DAAO:subunit (2I97, column 1), subunit complexed with FAD and D-His(PM0074742, column 2) and two subunits complexed with FAD and D-His(PM0074880, column 3); *M. tuberculosis* DAAO: subunit (2I3M, column 1), subunit complexed with FAD and D-Trp(PM0074760, column 2) and two subunits complexed with FAD and D-Trp(PM0074885, column 3); *C. porcellus* DAAO: subunit (2IB3, column 1), subunit complexed with FAD and D-His(PM0074749, column 2) and two subunits complexed with FAD and D-His(PM0074884, column 3); *C. griseus* DAAO subunit (2IB2, column 1), subunit complexed with FAD and D-His(PM0074773, column 2) and two subunits complexed with FAD and D-His (PM0074882, column 3).

Table 5 pK_a of some amino acid residues that change significantly during alteration of pH

Protein	Amino acid residues
2I8X	Arg(435, 432, 278, 226, 208), Lys(502, 272), Tyr(439, 438, 295, 267, 202), His(402), Glu(462), Asp(506, 482, 479, 460, 459, 431, 421, 322, 315, 305, 304, 291, 286)
2I8Z	Asn(512)
2I86	Trp(629), Asp(443)
2I90	Glu(503)
2I97	Asp(95), His(197, 224, 295, 299), Cys(230, 310), Tyr (140, 221), Arg(55,86, 227, 254, 275)
2IB2	Asp(72), Glu(21, 84, 120, 139, 153, 266), His(77, 79, 211), Cys(262, 263), Tyr(94), Lys(337, 210, 233, 270), Arg(114, 119)
2IB3	Asp(37, 73), Glu(64, 85, 121, 150, 235, 267, 294), His (78, 80, 134, 212, 217, 307, 311, 319, 338), Cys(181, 263, 264, 322), Tyr(55, 95, 128, 144, 228, 279, 309, 314), Lys(142, 211, 225), Arg(38, 115, 120, 271, 283, 286, 290)
PM0074893	Asp(39,47, 150, 189, 201, 263, 475, 476), Glu(51, 168, 188, 215, 221, 226, 378, 473, 499), His(91, 129, 133, 185, 217, 347, 381), Arg(130, 216), Lys55, Tyr(50, 267, 485, 493)
PM0074680	Asp(404,405), Glu(225, 338), His(341, 347, 486), Tyr (237,238), Lys(285,324), Arg(337,380)

electrostatic energy at each pH depends not only on the net number of total charges of the protein, but also on the location of such charges. The variation of electrostatic energy with pH also depends on the order in which such charged groups are ionized or deionized with change in pH. Also, the dielectric constant of the medium, its ionic strength and the surface shape of the protein affect the value of electrostatic energy in such media [43]. Using Deepview program (Supplementary Table 1), one can analyze structural details of electrostatic surface energy and molecular surface of pH conformers. Three different conformers of pH1, pH7, and pH13 had apparently identical ψ and Chi1–5 angles. In general, the differences in the three pH conformers(pH1, pH7, and pH13) were found in Ramachandran plot, bond lengths(CA-CB, N-CA, CA-C), bond angles (CA-C-N, C-N-CA, CA-C-O), RMS distances of planarity, beta turns (deletion/addition, chi i to i + 3 distances), hairpins (strand1, strand2, number of residues, hairpin class), helix interactions(helix1, helix2, distance geometry, omega angle, number of interacting residues), helix geometry (helix number, residues per turn, pitch), helices(3–10, α), β-α-β units, β-sheets(parallel, antiparallel, mixed, toplogy, sequences), residue distance matrix, ω angle, ϕ(i + 1) angle, significant pK_a changes (Table 9) of some amino acids, energy parameters (viz. free energy, sidechain entropy, mainchain entropy, Van der Waals' energy, polar desolvation energy, water bridge energy, hydrophobic desolvation energy, electrostatic energy, total energy and electrostatic charge), B-factor analysis, topology, dipole moment (also quadrapole moment, dipole vectors and mass moment vectors), number and pair of residues forming salt bridges (2–7 Å), free energy of folding and unfolding, pI and residue-residue distance matrix, average pK_a values of amino acid residues (nature and position of some amino acid residues were also different), total electrostatic charges (coulomb), Cα trace analysis (backbone analysis), electrostatic charged surface area, and other properties. The tools (Supplementary Table 1) that are available in the website can be used to analyze a desired property or structure of a molecule. However, the free energy difference ($\Delta\Delta G$) among three pH conformers was in the range of 0.5–5 kcal $mole^{-1}$ (Suppementary Table 4). This difference of energy values (0.5–5 kcal $mole^{-1}$) was almost the same for all energy parameters. The molecular co-ordinates of proteins of pH conformers obtained from PDB2PQR program [4, 8, 16] were analyzed using several programs (Supplementary Table 1). Salt bridges were identified between the negatively charged groups of aspartate, glutamate, or the carboxy-terminus of the protein, and the positively charged groups of histidine, lysine, arginine, or the amino-terminus. This calculation of salt bridges was limited to intra-subunit salt bridges connection network (data not shown). This variation may arise from different mode of polarization associated with hydrogen bonding (…OH-OH…) of sidechain(possibly N-glycosyl residues) by solvent(water) H^+ ions [43, 44]. The negative logarithm of association constant (pK_a) of individual residue amino acid of a protein was calculated as described before [8, 16]. The pK_a of one or more amino acid residues of some single subunit models were found to be altered significantly during variation of pH from 1 to 13 (Table 5). The reason is the different mode of protonation and deprotonation states of some amino acids of some proteins (Table 5) at different pH. Intermolecular over intramolecular hydrogen bonding with specific hydrophobicity and charge profiles that correlated to solubility and assembly of some proteins in the context of lowering pH and vaporization were explained with recent domain mapping studies (silk proteins), 2D Raman spectroscopy, NMR, and DLS studies [43, 44].

Prediction of binding interfaces of substrate, cofactor, inhibitor and ligand

The predicted binding surfaces (Supplementary Table 5 and Table 6) of L-amino acid (LAA), D-amino acid(DAA), deoxynucleotide triphosphate (dNTP), flavin-adenine dinucleotide (FAD), inhibitors (benzoate and anthranilate) and complex of two identical subunits were determined using

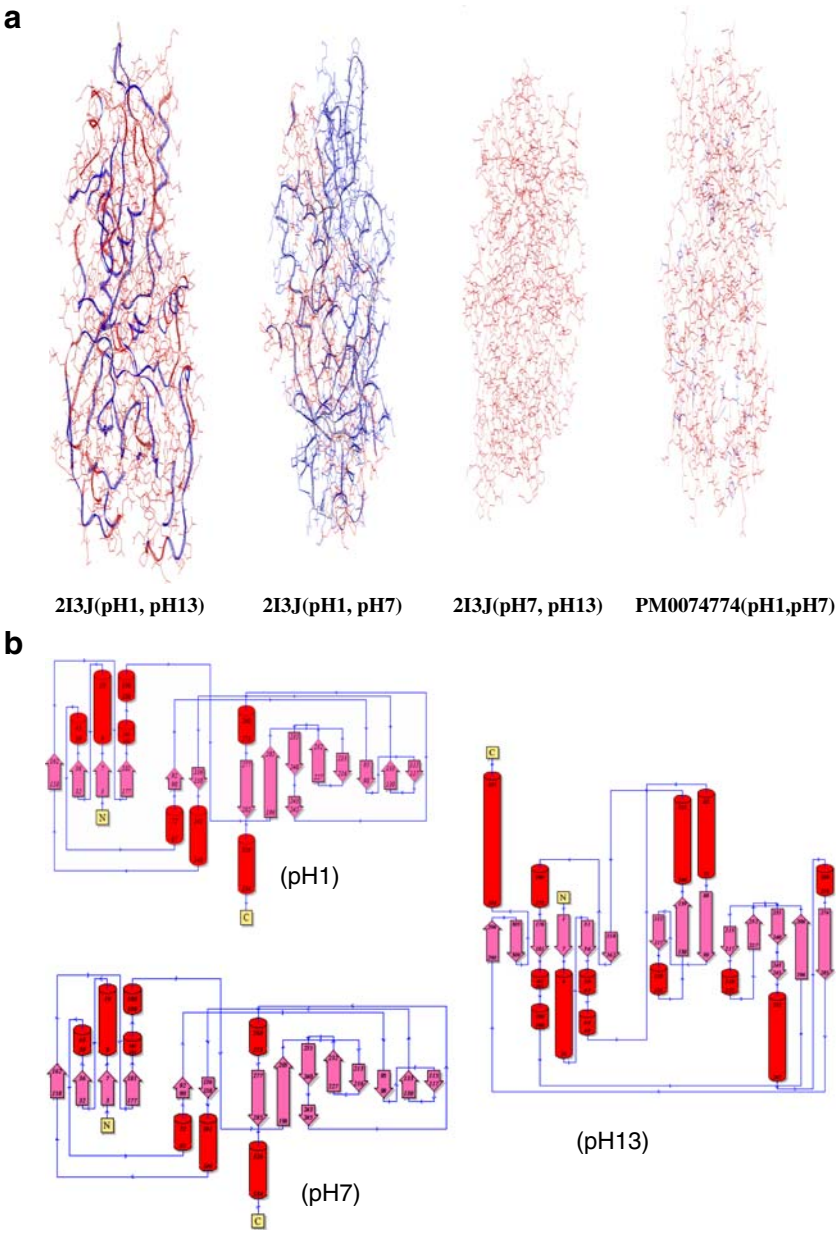

Fig. 2 Superimposition of pH conformers of proteins at three pH (viz. 1, 7 and 13) and their topolgy. Superimposition of pH conformers of human DAAO (2I3J) and *Roseovarius sp.* 217 DAAO(PM0074774) complexed with FAD and D-Trp (**a**), topology of 2I3J(**b**) and Ramachandran plot of 2I3J(**c**). 3D structures were drawn in wireframe

docking [28, 29], ICM (Molsoft) and Ligplot program of PDBsum. The predicted binding surfaces were found to be little different from the predicted binding sites estimated by PROSITE program (in PDBsum). PROSITE program is based mainly on UniProtKB (Swiss-Prot or TrEMBL) ID and protein sequence database, whereas the docking programs [32, 33] used in this study utilized algorithms to define a cleft in a protein molecule and attached a molecule preferably at that point. Some examples of protein binding interfaces and their geometry that were predicted from the computational analysis were shown in Fig. 3. The free energy difference(-$\Delta\Delta G$) and binding energy for LAA, DAA and FAD were discussed in Table 3. Despite some similarity of primary sequences of each category of proteins (NCBI sequences), 3D structures of these individual protein were unique and distinct (figure not shown), because of bonding network (hydrogen bonding, salt bridges, ionic interactions) that was unique and different for each protein molecule.

Measles virus hemagglutinin protein (MVH) (www.pepscan.nl/downloads/measlesH.pdb) was complexed with human CD150 (2DZF) (Fig. 3). *In vitro* mutagenesis analysis [45, 46] indicated that the amino acid binding residues were N-terminal V-domain(58–67) of human CD150 spanning Gly27-Leu135 region, whereas the amino acid binding residues of MVH were Ala429-Gly438, Tyr481, Asp505, Asp507, Leu522, Val525, Ser526, Ala527, Asp530, Asp533, His536, Tyr524, Thr531, and

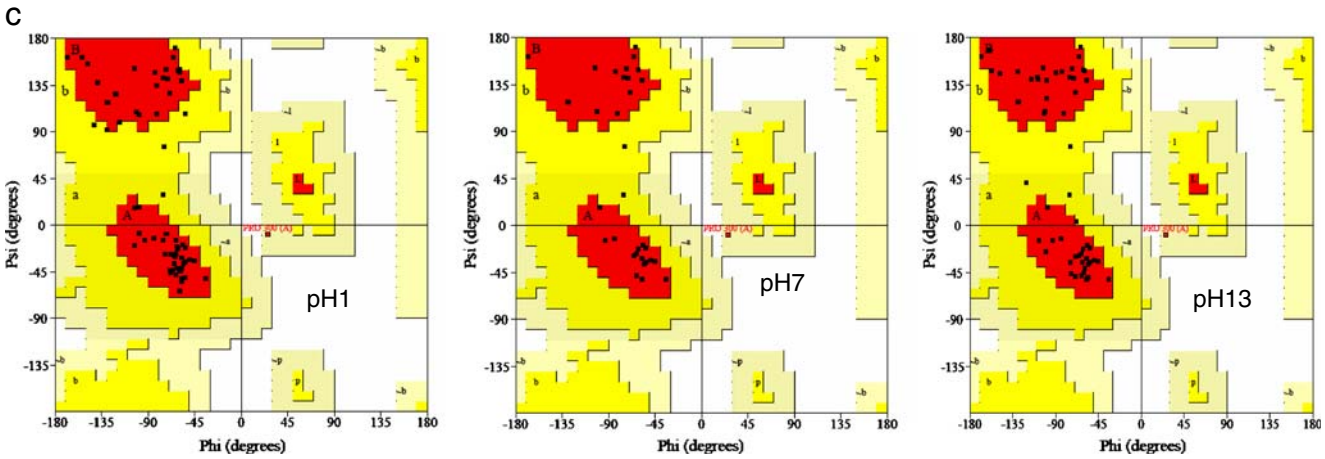

Fig. 2 (continued)

Ser546-Arg549, respectively. Gly27-Leu135 amino acids of CD150 were used for 10-mer phage display library to identify two peptides SGFDPLITHA and SDWDPLFTHK that were highly homologous with amino acid residues Ser429-Gly438(SGFGPLITHG) of MVH. These peptides specifically inhibited MV binding to SLAM [46]. In Fig. 3, the predicted binding residues of human CD150 were shown to be Leu36, Ile35, Leu36, Pro46, Asn72, Arg95, Tyr96, and Phe101, whereas the predicted binding residues of MVH were found to be Ser435, Ser439, Tyr471, Gln523, and Leu577. Other amino acid residues in the binding interface of the complex were also shown in Fig. 3. This result of docking can be interpreted such that the residual amino acids of the mutagenesis experiments [45, 46] could influence the binding amino acid residues (contact residues) of human CD150 and MVH during oligomeric interaction of these two protein partners by hydrogen bonding network, and may not participate directly in their association. The distance geometry, bond angles, bond distances, bond lengths, etc, were analyzed with Procheck and WhatIf analysis. In Fig. 3, only single subunit-subunit interaction of MVH and human CD150 are shown. ΔG of folding of CD150 (receptor), MVH (ligand) and receptor-ligand complex was found to be 316.08 kcal mole^{-1}, 329.40 kcal mole^{-1}, and 800.10 kcal mole^{-1}, respectively. Binding free energy ($\Delta\Delta G$) of receptor-ligand complex was 154.62 kcal mole^{-1}. These values were obtained by computing at ambient temperature (300°K). The hydrogen bond energy (65.70 kcal mole^{-1}), Van der Waals energy (-28.86 kcal mole^{-1}), polar desolvation energy (54.26 kcal mole^{-1}), side chain entropy (5.87 kcal mole^{-1}), main chain entropy (-7.52 kcal mole^{-1}), water bridge energy (-2.77 kcal mole^{-1}), hydrophobic energy (-51.46 kcal mole^{-1}), electrostatic energy (-4.27 kcal mole^{-1}) and total energy (154.62 kcal mole^{-1}) of this receptor-ligand complex were also determined.

Molecular dynamic simulation

The RMSD of each amino acid residue (Fig. 3 and Supplementary Fig. 1) was found to be 0.07–0.1 (Z-score 7.9–8.1) and 0.018–0.232 (z-score 7.9–8.1) for MD simulations (1 fentosecond at 300°K) after superimposition (profit program analysis) of the original PDB structure and MD simulated structure. Potential energy and total energy of protein subunit remained unchanged during this simulation. The deviations of MD simulated structures were determined using the similar procedures that were explained for pH conformers. The nature of these structural differences was also similar with the nature of the structural differences of pH conformers discussed in the previous section (figures not shown).

Conclusion

These mammalian enzymes (cellular, not mitochondrial) such as DAAO, LAAO and RTK are usually oligomer (homo or hetero multimer) [49 and references therein]. The only exception is *E.coli* DAAO which is a heterodimer [48]. The stoichiometry of mammalian enzyme subunit (DAAO, LAAO, and IL4-induced Protein1), FAD and substrate (or inhibitor) is usually 1:1:1 [48 and references therein]. DAAO has industrial importance [48 and references therein]. IL4-induced Protein 1 (human and mouse) is more or less equivalent to LAAO in function [48–55]. RTK genes were found in vertebrates, yeast, multicellular prokaryotes (e.g., *M. xanthus*), *Arabidopsis thaliana*, early metazoans and *Caenorhabditis elegans* [56–58]. Various RTK receptors have a tendency to phosphorylate both cis and trans on either Ser/Thr or Tyr. ATP binding residues for all RTKs lie in the C-terminal domain. Several other organisms have orthologs or mammalian paralogs. Until

Table 6 Predicted binding free energy (ΔΔG) and binding energy of substrate and cofactor

Organism	Enzyme	Cofactor and substrates	Binding free energy($\Delta\Delta G$) (kcal mole^{-1})	Binding energy (kcal mole^{-1})
H. sapiens	DAAO	FAD, D-Tyr	-3.513	-9.15
Photobacterium sp.	DAAO	FAD, D-Val	- 2.125	-8.0
M. musculus	DAAO	FAD, D-Trp	-5.137	-9.450
C. griseus	DAAO	FAD, D-His	-5.887	-3.337
M. tuberculosis	DAAO	FAD, D-Trp	- 5.562	-5.125
C. pelagibacter	DAAO	FAD, D-Asp	-2.162	-3.812
C. albicans	DAAO	FAD, D-Cys	-4.925	-9.213
C. albicans	DAAO	FAD, D-His	-4.212	-4.2125
C. atrox	LAAO	FAD, Anthranilate	-3.435	-8.386
C. atrox	LAAO	FAD	-22.34	-30.65
S. japonicus	LAAO	FAD, Anthranilate	-4.675	-18.424
G. gallus	LAAO	FAD, Anthranilate	-2.55	-8.875
G. gallus	LAAO	FAD, L-Trp	-2.551	-
X. campestris	DAAO	FAD, D-Phe	-0.64	-3.825
A. protomorphiae	DAAO	FAD, D-Tyr	-4.324	-8.72
A. fumigatus	DAAO	FAD, D-Met	-0.866	-5.037
Erythrobacter sp.	DAAO	FAD, D-Tyr	-10.874	-6.477
D. rerio	DAAO	FAD, D-Pro	-4.025	-3.844
D. rerio	DAAO	FAD	-1.9125	-6.925
D. rerio	DAAO	D-Pro	-3.163	-12.925
O. cuniculus	DAAO	FAD, D-Trp	-3.862	-5.611
R. norvegicus	DAAO	FAD, Benzoate	-1.625	-0.887
C. porcellus	DAAO	FAD, D-His	-3.863	-4.399
Nocardioides sp.	DAAO	FAD, D-Arg	-6.261	-8.23
S. raphanus	TK	ATP	-1.6	-8.6025

now, there is no report for 3D structures of nucleoprotein, matrix protein and RNA polymerase of measles virus. The complete crystal srtructure of human CD150(SLAM) is also unavailable. The other PDB entries (communicated from this center) of domains of these proteins were 2IG4 (PDB), 2IG5 (PDB), and 2IFL (PDB), 2DZF (PDB) and 2IGI (PDB). Each category of proteins had extensive similarity in amino acid sequences. Besides, all proteins exhibited the variable pattern of unfolding and folding during denaturation caused by pH and temperature. Whereas the X-ray structure is the appropriate model for the native protein in solution, there are insufficient experimental data to determine the structures of unfolded proteins or folding intermediates, in most cases. The available data suggest that the unfolded state of a protein is characterized by an ensemble of rather different conformations [22, 38–43]. Although an ensemble of configurations should be used for the native state, as well as the unfolded state, an average over a sufficient number of conformers is likely to be more important [22, 38–43]. The folding core is stabilized by a network of particularly dense or strong non-covalent interactions, which tend to resist unfolding or denaturation [42]. Flexible regions in proteins were defined by analyzing the constraints on flexibility formed by the covalent and non-covalent bond network [38–43]. Covalent bonds, salt bridges, hydrogen bonds, and hydrophobic interactions were included in the protein representation. Groups of atoms coupled to each other via rigid bonds form a rigid cluster. One or more independent rigid clusters with intervening flexible regions may occur in a protein structure. As a protein is gradually denatured, the covalent bonds remain intact, whereas hydrogen bonds begin to break. The flexibility in the protein will increase as the number of hydrogen bond energies in the protein decrease. Because hydrophobic interactions actually become somewhat stronger with moderate temperature increases [38–43], these interactions are maintained throughout the simulation. Application of the charge-assignment and pK_a-calculation procedure to protein-ligand complexes provides clear structural interpretations of experimentally observed changes of protonation states of functional groups upon complex formation. This information is essential for the interpretation of thermodynamic data of protein-ligand complex formation and provides the basis for the reliable factorization of the free energy of binding in enthalpic and entropic contributions [4]. The desolvation effects and intra-protein interactions, which cause variations in pK_a values of protein ionizable groups, are empirically related to the positions and chemical nature of the groups proximate to the pK_a sites. Unusual pK_a values at buried

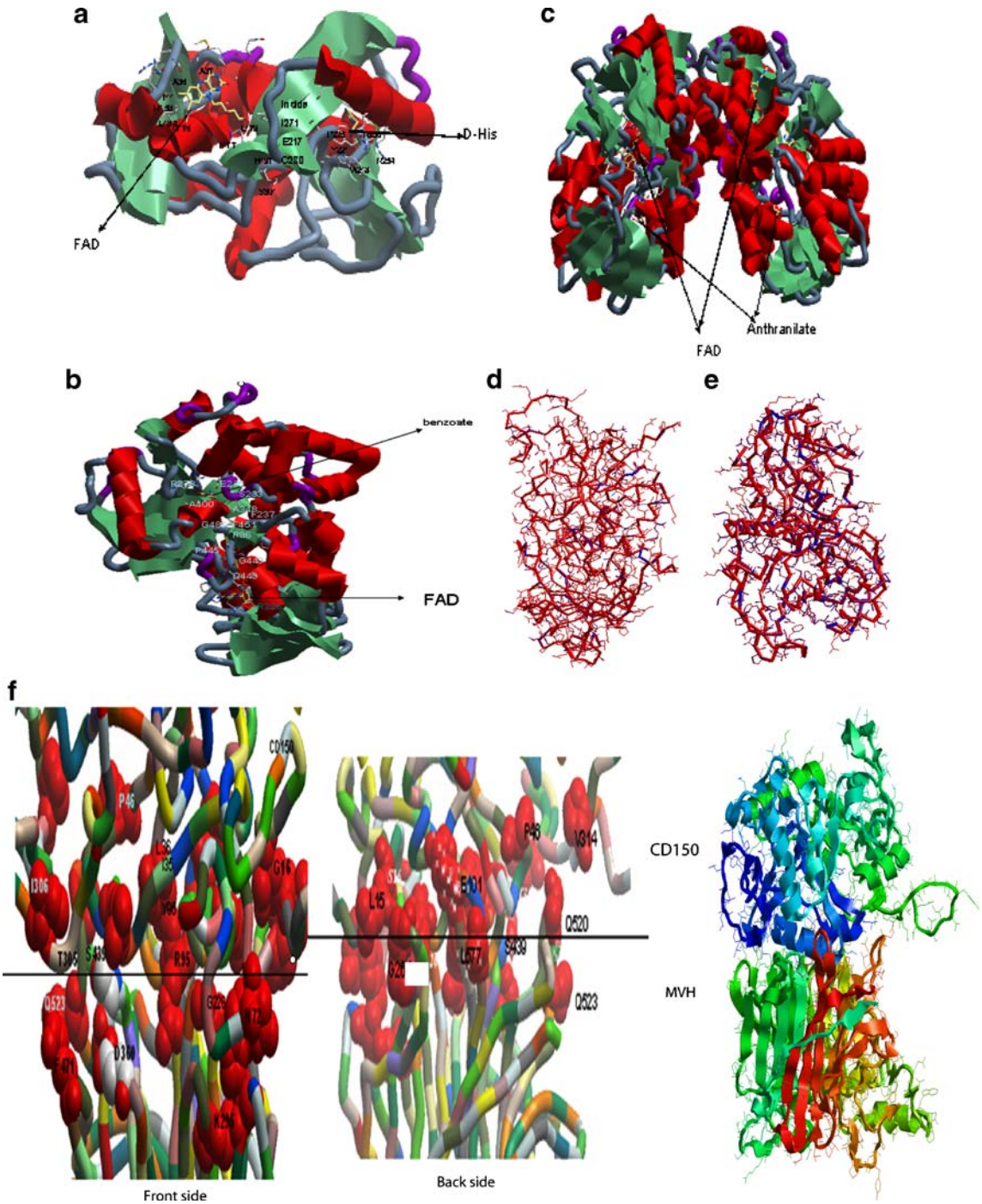

Fig. 3 Docking and MD simulation of proteins. *M. leprae* DAAO subunit bound with FAD and D-His (**a**), mouse IL4-induced Protein1 subunit bound with FAD and benzoate(**b**), *G. gallus* LAAO(complex of two identical subunits) bound with FAD and L-Trp(**c**), superimposition of PDB proteins and their MD simulated structures (2DZG(**d**) and 2I3M(**e**)) and binding complex of human CD150 and measles virus hemagglutinin(MVH)(**f**). 3D structures of (**d**) and (**e**) were drawn in wireframe

active sites, which are among the most interesting protein pK_a values, are predicted very well with the empirical method [9]. The measurement, theoretical concept, significance of algorithm and computational limitation of the detailed structural parameters (including pK_a) of proteins with respect to change in pH and temperature were discussed before [22, 38–43]. Many studies (including X-ray or NMR solution structure) had shown the characteristics of protein–protein interfaces in an effort to search for the factors that contributed to the affinity and specificity of protein–protein interactions [59–63]. These analyses implied that the two surfaces of a protein–protein interface

usually depended on high degrees of geometric and chemical complementarities, electrostatic forces [64–66], residue composition and inter-residue contacts [67, 68], potentials of mean force for inter-residue interactions held for both intra-molecular and inter-molecular interactions [69], hydrophobic forces [70, 71], hydrophilic effect [72], salt bridges, Van der Waals' forces, hydrogen bonds, and the existence of "hot-spot" residues for complex formation [73] among several types and subtypes of the protein–protein interface. Some studies divided the protein–protein interfaces into several subtypes and analyzed the characteristics of each subtype [67–75]. In an experimental set-up, pH conformation was analyzed with circular dichroism (CD), ultra-violet spectroscopy, differential scanning calorimetry, and infrared spectroscopy [12, 13].

Acknowledgements The authors are indebted to Dr. Debikumar Lobiyal and Dr. Vikas Rai of School of Computer Science for discussion. The authors greatly acknowledge Vijyan G, Sarbashis Das, Rahul Sharma, Rakesh Pandey,Anurag Bagaria, Amit Kumar, Mr. Shashi Thakur, Mahinder Kumar and Sailendra Singh Bisht for their help and comments during this work.

Conflict of interest among authors None.

Request data Analysis of data set (15 GB) will be available after request from the corresponding author.

References

1. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Proc Natl Acad Sci USA 98:10037–10041. doi:10.1073/pnas.181342398
2. Chen R, Li L, Weng Z (2003) Proteins 52:80–87. doi:10.1002/prot.10389
3. Comeau SR, Camacho CJ (2005) J Struct Biol 150:233–244. doi:10.1016/j.jsb.2005.03.006
4. Czodrowski P, Dramburg I, Sotriffer CA, Klebe G (2006) Proteins 65:424–437. doi:10.1002/prot.21110
5. Dobson CM, Karplus M (1999) Curr Opin Struct Biol 9:92–101. doi:10.1016/S0959–440X(99)80012–8
6. Englander SW (2000) Annu Rev Biophys Biomol Struct 29:213–238. doi:10.1146/annurev.biophys.29.1.213
7. Englander J, Borg J, Stricher F, Nys R, Rousseau F, Serrano L (2005) Nucleic Acids Res 33:W382–W388. doi:10.1093/nar/gki387
8. Li H, Robertson AD, Jensen JH (2005) Proteins 61:704–721. doi:10.1002/prot.20660
9. Morris AL, MacArthur MW, Hutchinson EG, Thornton JM (1992) Protein Struct Funct Gen 12:345–364. doi:10.1002/prot.340120407
10. Samuel D, Kumar TK, Balamurugan K, Lin WY, Chin DH, Yu C (2001) J Biol Chem 276:4134–4141. doi:10.1074/jbc.M005921200
11. Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F, Serrano L (2005) Nucleic Acids Res 33:W382–W388. doi:10.1093/nar/gki387
12. Shimada H, Caughey WS (1982) J Biol Chem 257:11893–11900
13. Welfle K, Misselwitz R, Hausdorf G, Höhne W, Welfle H (1999) Biochim Biophys Acta 1431:120–131
14. Qi X, Portman JJ (2007) Proc Natl Acad Sci USA 104:10841–10846. doi:10.1073/pnas.0609321104
15. Merlo C, Dill KA, Weikl TR (2005) Proc Natl Acad Sci USA 102:10171–10175. doi:10.1073/pnas.0504171102
16. Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA (2004) Nucleic Acids Res 32:W665–W667. doi:10.1093/nar/gkh381
17. Willard L, Ranjan A, Zhang H, Monzavi H, Boyko RF, Sykes BD et al (2003) Nucleic Acids Res 31:3316–3319. doi:10.1093/nar/gkg565
18. Hooft RWWW, Vriend G, Sander C, Abola EE (1996) Nature 381:272–272. doi:10.1038/381272a0
19. Hutchinson EG, Thornton JM (1996) Protein Sci 5:212–220
20. Kabsch W, Sander C (1983) Biopolymers 22:2577–2637. doi:10.1002/bip.360221211
21. Luthy R, Bowie JU, Eisenberg D (1992) Nature 356:83–85. doi:10.1038/356083a0
22. Melo F, Feytmans E (1998) J Mol Biol 277:1141–1152. doi:10.1006/jmbi.1998.1665
23. Chang I, Cieplak M, Banavar JR, Maritan A (2004) Protein Sci 13:2446–2457. doi:10.1110/ps.04713804
24. Gromiha M (2005) J Chem Inf Model 45:494–501. doi:10.1021/ci049757q
25. Brown MPS, Grundy WN, Lin D, Cristianini N, Sugnet CW, Furey TS, Ares M Jr., Haussler D (2000) Proc Natl Acad Sci USA 97:262–267. doi:10.1073/pnas.97.1.262
26. Lambert C, Leonard N, De Bolle X, Depiereux E (2002) Bioinformatics 18:1250–1256. doi:10.1093/bioinformatics/18.9.1250
27. Mintseris J, Weng Z (2003) Proteins 53:629–639. doi:10.1002/prot.10432
28. Li L, Chen R, Weng Z (2003) Proteins 53:693–707. doi:10.1002/prot.10460
29. Wiehe K, Pierce B, Mintseris J, Tong W, Anderson R, Chen R et al. (2005) Proteins 60:207–221. doi:10.1002/prot.20559
30. Maiti R, van Domselaar GH, Zhang H, Wishart DS (2004) Nucleic Acids Res 32:W590–W594. doi:10.1093/nar/gkh477
31. Shindyalov IN, Bourne PE (1998) Protein Eng 11:739–747. doi:10.1093/protein/11.9.739
32. Halligan BD, Victor Ruootti V, Weihong Jin W, Scott Laffoon S, Simon N, Twigger SN et al (2004) Nucleic Acids Res 32:W638–W644. doi:10.1093/nar/gkh356
33. Laskowski RA (2007) Bioinformatics 23:1824–1827. doi:10.1093/bioinformatics/btm085
34. Castrignanò T, De Meo PD, Domenico Cozzetto D, Talamo IG, Tramontano A (2006) Nucleic Acids Res 34:D306–D309. doi:10.1093/nar/gkj105
35. Berendsen HJC, van der Spoel D, van Drunen R (1995) Comput Phys Commun 91:43–56. doi:10.1016/0010–4655(95)00042-E
36. Lindahl E, Hess B, van der Spoel D (2001) J Mol Model 7:306–317
37. Pollegioni L, Iamettl S, Fessas D, Caldinelli L, Piubelli L, Barbiroli A et al (2003) Protein Sci 12:1018–1029. doi:10.1110/ps.0234603
38. Antosiewicz J, McCammon JA, Gilson M (1994) J Mol Biol 238:415–436 . doi:10.1006/jmbi.1994.1301
39. Dahiyat BI, Gordon DB, Mayo SL (1997) Protein Sci 6:1333–1337
40. Dong X, Tsai C-J, Nussinov R (1997) Protein Eng 10:999–1012. doi:10.1093/protein/10.9.999
41. Fersht AR, Itzhaki LS, ElMasry NF, Matthews JM, Otzen DE (1994) Proc Natl Acad Sci USA 91:10426–10429. doi:10.1073/pnas.91.22.10426
42. Jacobs DJ, Rader AJ, Kuhn LA, Thorpe MF (2001) Proteins Struct Func Genet 44:150–165. doi:10.1002/prot.1081

43. Rader AJ, Hespenheide BM, Kuhn LA, Thorpe MF (2002) Proc Natl Acad Sci USA 99:3540–3545. doi:10.1073/pnas.062492699
44. Olivera-Nappa A, Lagomarsino G, Andrews BA, Asenjo JA (2004) J Chromatogr B Analyt Technol Biomed Life Sci 807:81–86. doi:10.1016/j.jchromb.2004.03.033
45. Massé N, Ainouze M, Néel B, Wild TF, Buckland R, Langedijk JPM (2004) J Virol 78:9051–9063. doi:10.1128/JVI.78.17.9051–9063.2004
46. Santiago CE, Bjorling E, Stehle T, Cassaanovas JM (2002) J Biol Chem 277:32294–32301. doi:10.1074/jbc.M202973200
47. Garcı'a-Alles LF, B Erni B (2002) Eur J Biochem 269:3226–3236. doi:10.1046/j.1432–1033.2002.02995.x
48. Sumathi K, Ananthalakshmi P, Roshan MNA, Sekar K (2005) Nucleic Acids Res 34:W128–W132. doi:10.1093/nar/gkl036
49. Pollegioni L, Piubelli L, Sacchi S, Pilone MS, Molla G (2007) Cell Mol Life Sci 64:1373–1394. doi:10.1007/s00018–007–6558–4
50. Chavan SS, Tian W, Hsueh K, Jawaheer D, Gregersen PK, Chu CC (2002) Biochim Biophys Acta 1576:70–76
51. Copie-Bergman C, Boulland M-L, Dehoulle C, Moller P, Farcet J-P, Dyer MJS et al (2003) Blood 101:2756–2761. doi:10.1182/blood-2002–07–2215
52. Mason JM, Naidu MD, Barcia M, Porti D, Chavan SS, Chu CC (2004) J Immunol 173:4561–4567
53. Pawelek PD, Cheah J, Coulombe R, Macheroux P, Ghisla S, Ghisla A (2000) EMBO J 19:4204–4209. doi:10.1093/emboj/19.16.4204
54. Wiemann S, Kolb-Kokocinski A, Poustka A (2005) BMC Biol 3:1–12. doi:10.1186/1741–7007–3–16
55. Geyer A, Fitzpatrick BTB, Pawelek PD, Kitzing K, Vrielink A, Ghisla S et al (2001) Eur J Biochem 268:4044–4053. doi:10.1046/j.1432–1327.2001.02321.x
56. Fukamachi H, Kawakami Y, Takeit M, Shizakat T, Ishizaka K, Kawakami T (1992) Proc Natl Acad Sci USA 89:9524–9528. doi:10.1073/pnas.89.20.9524
57. Grassot J, Gouy M, Perrie`re G, Mouchiroud G (2006) Mol Biol Evol 23:1232–1241. doi:10.1093/molbev/msk007
58. Murali R, Brennan PJ, Kieber T, Emmons TK, Greene MI (1996) Proc Natl Acad Sci USA 93:6252–6257. doi:10.1073/pnas.93.13.6252
59. Chothia C, Janin J (1975) Nature 256:705–708. doi:10.1038/256705a0
60. Wodak SJ, Janin J (2002) Adv Protein Chem 61:9–73
61. Deremble C, Lavery R (2005) Curr Opin Struct Biol 15:171–175. doi:10.1016/j.sbi.2005.01.018
62. Ponstingl H, Kabir T, Gorse D, Thornton JM (2005) Prog Biophys Mol Biol 89:9–35. doi:10.1016/j.pbiomolbio.2004.07.010
63. Reichmann D, Rahat O, Cohen M, Neuvirth H, Schreiber G (2007) Curr Opin Struct Biol 17:67–76. doi:10.1016/j.sbi.2007.01.004
64. Sheinerman FB, Norel R, Honig B (2000) Curr Opin Struct Biol 10:153–159. doi:10.1016/S0959–440X(00)00065–8
65. Heifetz A, Katchalski-Katzir E, Eisenstein M (2002) Protein Sci 11:571–587. doi:10.1110/ps.26002
66. Vizcarra CL, Mayo SL (2005) Curr Opin Chem Biol 9:622–626
67. Glaser F, Steinberg DM, Vakser IA, Ben-Tal N (2001) Proteins 43:89–102. doi:10.1002/1097–0134(20010501)43:2<89::AID-PROT1021>3.0.CO;2-H
68. Ofran Y, Rost B (2003) J Mol Biol 325:377–387. doi:10.1016/S0022–2836(02)01223–8
69. Keskin O, Bahar I, Badretinov AY, Ptitsyn OB, Jernigan RL (1998) Protein Sci 7:2578–2586
70. Young L, Jernigan RL, Covell DG (1994) Protein Sci 3:717–729
71. Berchanski A, Shapira B, Eisenstein M (2004) Proteins 56:130–142. doi:10.1002/prot.20145
72. Ben-Naima A (2006) J Chem Phys 125:24901. doi:10.1063/1.2205860
73. Keskin O, Mab B, Nussinov R (2005) J Mol Biol 345:1281–1294. doi:10.1016/j.jmb.2004.10.077
74. Jones S, Thornton JM (1996) Proc Natl Acad Sci USA 93:13–20. doi:10.1073/pnas.93.1.13
75. Block JP, Hu¨lermeier E, Sanschagrin P, Sotriffer CA, Klebe G (2006) Proteins 65:607–622. doi:10.1002/prot.21104