Crowding, molecular volume and plasticity: An assessment involving crystallography, NMR and simulations

M Selvaraj[1], Rais Ahmad[2], Umesh Varshney[2] and M Vijayan[1,*]

[1]*Molecular Biophysics Unit and* [2]*Department of Microbiology and Cell Biology, Indian Institute of Science, Bangalore 560012*

**Corresponding author (Fax, +91-80-2360-0535, +91-80-2360-0683; Email, mv@mbu.iisc.ernet.in)*

The discrepancy between the X-ray and NMR structures of *Mycobacterium tuberculosis* peptidyl-tRNA hydrolase in relation to the functionally important plasticity of the molecule led to molecular dynamics simulations. The X-ray and the NMR studies along with the simulations indicated an inverse correlation between crowding and molecular volume. A detailed comparison of proteins for which X-ray and the NMR structures appears to confirm this correlation. In consonance with the reported results of the investigations in cellular compartments and aqueous solution, the comparison indicates that the crowding results in compaction of the molecule as well as change in its shape, which could specifically involve regions of the molecule important in function. Crowding could thus influence the action of proteins through modulation of the functionally important plasticity of the molecule.

[Selvaraj M, Ahmad R, Varshney U and Vijayan M 2012 Crowding, molecular volume and plasticity: An assessment involving crystallography, NMR and simulations. *J. Biosci.* **37** 1–11] **DOI** 10.1007/s12038-012-9276-5

1. Introduction

The effect of macromolecular crowding on the shape, folding and action of proteins has received considerable attention in recent years (Ellis and Minton 2003; Zhou *et al*. 2008; Gershenson and Gierasch 2011; Pernilla 2011). This issue is of substantial biological significance as close to half the volume of a typical cell is occupied by biomolecules, and therefore macromolecules like proteins function in a crowded environment in biological systems (Ellis and Minton 2003). This situation is not factored into most of the solution and computational studies on the structure and function of proteins, although there have been a couple of attempts to do so (Ai *et al*. 2006; Roque *et al*. 2007; Homouz *et al*. 2009). Perhaps the only extensive studies which are carried out in an overcrowded environment involve those pursued in crystals where typically the protein molecules occupy 50% of the volume. Admittedly, in crystals a protein molecule is usually surrounded by molecules of the same kind while the environment is much more heterogeneous inside the cell. Furthermore, the mobility of molecules in crystals is much more limited than in the cell. However, the level of overcrowding is comparable in the two cases. It has also been demonstrated that crystal contacts are different in nature from interactions at interfaces involved in specific assemblies Bahadur *et al.* (2003)). Therefore, the situation in protein crystals provides a handle, imperfect though it might be, for approaching the problem. The differences in protein structures derived using crystallography and NMR are particularly interesting in this context. Such differences observed in eubacterial peptidyl-tRNA hydrolase are the genesis of the work reported here.

Peptidyl-tRNA hydrolase (Pth), which catalyses the hydrolysis of stalled peptidyl-tRNA during protein synthesis, is an essential enzyme in eubacteria. Premature stalling of translation, caused by a variety of events, leads to the dropping off of peptidyl-tRNA from the ribosome. Accumulation of peptidyl-tRNA is toxic to the cell. Pth cleaves the ester bond between tRNA and the peptide, thus preventing this toxicity and also releasing tRNA for further use. The enzyme from *Mycobacterium tuberculosis* (*Mt*Pth), is a 191-residue monomeric protein. We deter-

Keywords. Molecular crowding; molecular plasticity; molecular shape; peptidyl-tRNA hydrolase; protein function

Supplementary materials pertaining to this article are available on the *Journal of Biosciences* Website at *http://www.ias.ac.in/jbiosci/dec2012/supp/Selvaraj.pdf*

Published online: 10 November 2012

mined, to start with, the structures of three crystal forms of this protein (Selvaraj *et al*. 2006; 2007) as part of a larger programme on the structural biology of mycobacterial proteins (Vijayan 2005; Krishna *et al*. 2007; Kaushal *et al*. 2008; Roy *et al*. 2008; Prabu *et al*. 2009; Chetnani *et al*. 2010). The X-ray structure of the homologue from *E.coli* (*Ec*Pth) was then already available (Schmitt *et al*. 1997). The C-terminal stretch of *Mt*Pth was disordered in the crystal structures. On the other hand, in the crystal structure of *Ec*Pth, the C-terminus of the molecule was bound to the peptide binding site of a neighbouring molecule. Thus, *Mt*Pth provided the structure of the apoenzyme while *Ec*Pth presented a picture of the structure of the peptide bound enzyme.

The Pth molecule, with an α/β fold involving a twisted β sheet flanked by helices, has two binding regions separated by a gate (figure 1). One of them, the peptide binding region, involves a lid containing a helix and a loop (Gly136 to Val150 in *Mt*Pth). This region is separated from the tRNA binding region by a gate made up of Asp98 and Gly113. In the *Mt*Pth structure without any ligand, the lid has an open conformation and the gate is closed. In the crystal structure of *Ec*Pth, where a peptide stretch is bound to the peptide binding region, the lid is closed on the bound peptide. Concurrently, the gate opens such that the peptide and the tRNA binding regions are now contiguous permitting the binding of peptidyl-tRNA to the molecule prior to hydrolysis. Thus, the *Mt*Pth and *Ec*Pth structures appeared to provide a plausible picture of functionally important plasticity in terms of the correlated motion of the lid and the gate (Selvaraj *et al*. 2007). This picture derived further support from the crystal structures of *Mycobacterium smegmatis* Pth (*Ms*Pth, Pdb code: 3p2j) and *Francisella tularensis* Pth (*Ft*Pth, Pdb code: 3nea), which became available subsequently (Clarke *et al*. 2011). The molecule is unliganded in the *Ms*Pth structures and it has an open lid and a closed gate. The crystal structure of *Ft*Pth is similar to that of *Ec*Pth with the C-terminal peptide bound to a neighbouring molecule. As in the *Ec*Pth structure, the *Ft*Pth molecule has a closed lid and an open gate.

In the meantime, the solution NMR structure of *Mt*Pth became available. The overall structure of the molecule in solution is understandably similar to that in the crystals, although there are differences in detail. The C-terminal stretch of the molecule is disordered to different extents in the crystal structures while it forms a helix in solution. More strikingly, unlike in the crystal structures, the gate is open in the solution NMR structure although the lid has an open conformation in the crystals and solution (Pulavarti *et al*. 2009). To further address this anomaly, more crystal forms of *Mt*Pth, grown under different conditions, were X-ray analysed (Selvaraj *et al*. 2012). As in the case of the *Mt*Pth crystals analysed earlier, the molecule in them had an open lid and a closed gate. Thus, all the crystal structures of Pth from different sources analysed so far give indication for a functionally important inverse correlation between the locations of the lid and the gate. Such a correlation does not exist in the solution NMR structure (Selvaraj *et al*. 2012). As a further effort, molecular dynamics (MD) simulations were undertaken on *Mt*Pth and *Ec*Pth. The results of these simulations, reported here, along with the available X-ray and NMR results appear to indicate a relation among crowding, macromolecular volume and plasticity. Subsequently, proteins for which X-ray as well as NMR results were available were carefully examined. Earlier studies involving such results have led to concerns about packing in NMR structures (Ratnaparkhi *et al*. 1998; Fu *et al*. 2011; Schwieters and Clore 2008). The fairly comprehensive examination

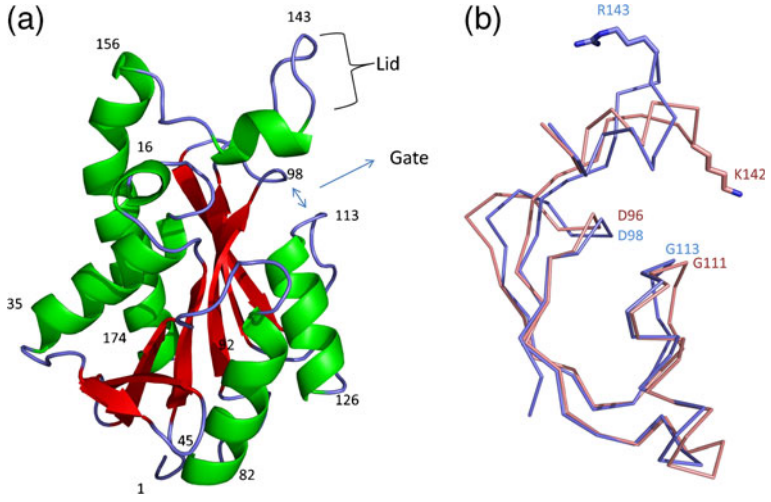

Figure 1. (**a**) Overall architecture of *Mt*Pth, with lid and gate highlighted. (**b**) Close up view of the active site of *Ec*Pth (pink) and *Mt*Pth (blue).

presented here focuses specifically on macromolecular volume and the space available around the protein molecule. This examination appears to lend additional support to the hypothesis on the relation mentioned above.

2. Methods

Molecular dynamics (MD) simulations were performed using the package GROMACSv3.3.1 running on parallel processor with OPLS-AA/L force field (Jorgensen *et al*. 1996; Spoel *et al*. 2005). Initially, the crystallographic water molecules in the structures were removed from the protein model. The conformation of lids in *Mt*Pth and *Ec*Pth was interchanged using COOT. A cubic box around the protein molecule was generated using the *editconf* module of GROMACS with the criterion that the minimum distance between the solute and the edge of the box was at least 7.5 Å. Following this, the protein models were solvated with TIP4P water model using program *genbox* available in the GROMACS suite. Sodium or chloride ions were added to neutralize the overall charge of the system whenever necessary. Energy minimizations were carried out using the conjugate gradient and steepest descent methods with a frequency of latter at 1 in 1000. A maximum force of 1 $kJmol^{-1}nm^{-1}$ was chosen as convergence criterion for minimization. Energy minimizations were followed by solvent equilibration by position restrained dynamics of 10 ps where positions of protein atoms were restrained and solvent was allowed to move. Simulations utilized the NPT ensembles with Parrinello-Rahman isotropic pressure coupling (τ_p=0.5) to 1 bar and Nose-Hoover temperature coupling (τ_t=0.1 ps) to 300 K. Long-range electrostatic interactions were computed using the Particle Mesh Ewald (PME) method with a cut-off of 12 Å (Darden *et al*. 1993). A cut-off of 15 Å was used to compute long-range van der Waals interactions. Simulations were performed with full periodic boundary conditions (PBC). Bonds were constrained with the LINCS algorithm (Berk Hess et al. 1997). Simulations were carried out with a dielectric constant of unity.

Structure superpositions were performed using ALIGN (Cohen 1997). Principal axes of the ellipsoidal representation of the protein molecules were calculated using CHIMERA (Pettersen *et al*. 2004). The ellipsoidal volume of the molecules was calculated using the principal axes.

3. Results and discussion

3.1 *MD simulations*

To obtain an additional independent insight into the structure and plasticity of *Mt*Pth, MD simulations were carried out on the molecule. A molecule obtained from form I was used as the starting model (Selvaraj *et al*. 2007). The C-terminal stretch involving residues from 180 to 191 was not defined in this form. The structure of the stretch in the crystal structure of *Ec*Pth was used to complete the model. Simulations were also performed with *Ec*Pth as the starting model. In order to explore the effect of starting models on the outcome of simulations, similar calculations were carried out with *Mt*Pth having a closed lid conformation as in *Ec*Pth and *Ec*Pth with an open lid conformation as in *Mt*Pth, as starting models. Each simulation was run for 60 ns. The trajectory of r.m.s. deviations in them are given in the supplementary figure 1.

The root mean square fluctuations of Cα positions along the polypeptide chain obtained from MD simulations are given in figure 2. The pattern of the fluctuation is nearly the same in the four simulations. The nature of the flexibility of the molecule indicated by the simulations is nearly the same as that shown by the X-ray and the NMR results (Selvaraj *et al*. 2012). The C-terminal stretch is the most flexible region of the molecule. This flexibility has been suggested to be of biological significance (Selvaraj *et al*. 2012). The lid region (residues 136 to 150), which has an open conformation in the crystal structures in the absence of a bound ligand and closed conformation when bound to a peptide, also exhibits substantial flexibility. One of the gate residues, Gly113, which again moves in crystal structures, is in a flexible region. Another flexible stretch is the 50–60 loop. The biological significance of the flexibility of this loop is unclear. In the three-dimensional structure, this loop is flanked by the highly flexible C-terminal stretch and the moderately flexible N-terminal region.

3.2 *Movement of the lid and the gate*

The movement of the lid is best described in terms of the distance between the Cα of Val92, which is an invariant residue in the rigid core of the molecule, and the Cα of Arg143 (Lys142/143 in *Ec*Pth and *Ft*Pth) at the tip of the lid (Selvaraj *et al*. 2012). The distance between the Cα atoms of Asp98 and Gly113 can be obviously used to describe the movement of the gate. The lid is open and the 92Cα–143Cα distance in the relevant mycobacterial structures varies between 28.0 Å and 29.4 Å. The gate in these structures is closed with a 98 Cα–113Cα distance varying between 6.1 Å and 7.4 Å. The lid is closed in *Ec*Pth and *Ft*Pth with a lid distance of 23.6 Å and 23.4 Å, respectively. The distance between 98 Cα and 113 Cα in the open gate in the two structures is 9.5 Å and 10.0 Å, respectively. The distances are thus inversely correlated in the crystal structures (figure 3a). The lid distances of the 40 models obtained from NMR range between 25.3 Å to 31.7 Å with an average value of 28.5 Å, a value appropriate for an open conformation (Selvaraj *et al*. 2012). The gate is always open. The distance between 98Cα and 113Cα varies between 8.6 Å and 17.7 Å with an average value of 12.6 Å. The correlation between the distances is absent or is very weak at best (figure 3b).

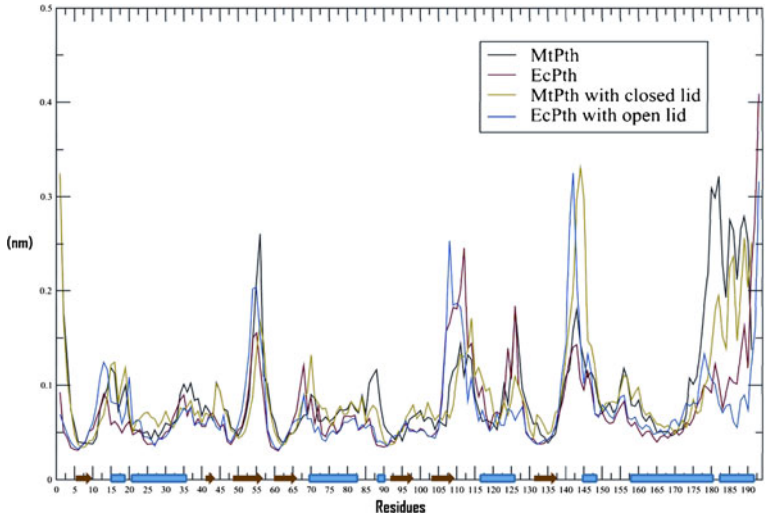

Figure 2. Root mean square fluctuations of simulated structures with the secondary structural elements of the molecule indicated at the bottom of the figure.

The simulations indicate two possible gate distances, one around 10.5 Å and the other around 14 Å (figure 4a), the latter being the prominent one except in *Ec*Pth. Even in *Ec*Pth, the distance eventually stabilizes at around 14 Å (figure 4c). On the contrary, the lid distance appears to critically depend on the starting model. In three out of the four cases, the distance hovers around the value in the starting model (figure 4b). The distance exhibits a wide scatter in the simulation of *Mt*Pth with a closed lid. Even in this case, the distance eventually stabilizes at values close to that in the starting model (figure 4d). Irrespective of the population distribution and the time evolution of the gate and the lid distances, there is no appreciable correlation between them in any of the four simulations (figure 5).

3.3 *Volume of the molecule and the available space around it*

The apparent discrepancy between the X-ray and the NMR results on lid and gate distances in *Mt*Pth was a cause for concern. MD simulations, like X-ray and NMR results, indicated an open conformation for the lid, but yielded gate distances even larger than those obtained through NMR. The crystal structures of Pths showed an inverse correlation between lid distance and gate distance, while NMR studies and MD simulations of *Mt*Pth indicated hardly any correlation between the two. There was no obvious explanation for this intriguing situation. A careful visual examination appeared to suggest that the Pth molecule(s) derived from NMR studies was slightly larger than that seen in crystal structures. To further explore this possibility, ellipsoidal volumes of the molecule in the relevant crystal structures, the NMR structures, and in the course of MD simulations were calculated, along with the principal axes of the ellipsoids (table 1).

It is clear from table 1 that the ellipsoidal volume of the molecule obtained from NMR studies is higher than that obtained from crystal structures. The MD simulations yield an intermediate value. The change in volume appears to be related to the volume of the solution around the molecule (table 1). The unit cell of the crystal containing four molecules has a volume of 1.66×10^5 Å^3 with the solvent content

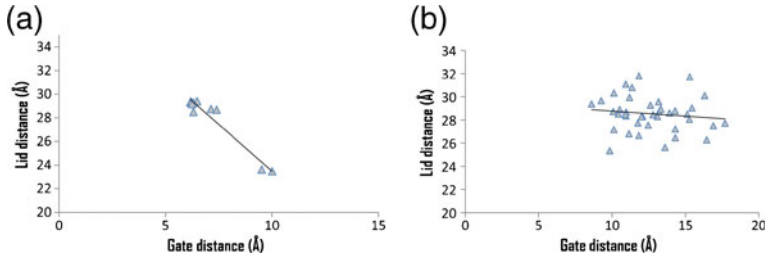

Figure 3. Correlation between gate and lid distances in (**a**) available crystal structures and (**b**) NMR structures. Reproduced with permission from the International Union of Crystallography from Selvaraj *et al.* (2012) (*http://journals.iucr.org/*).

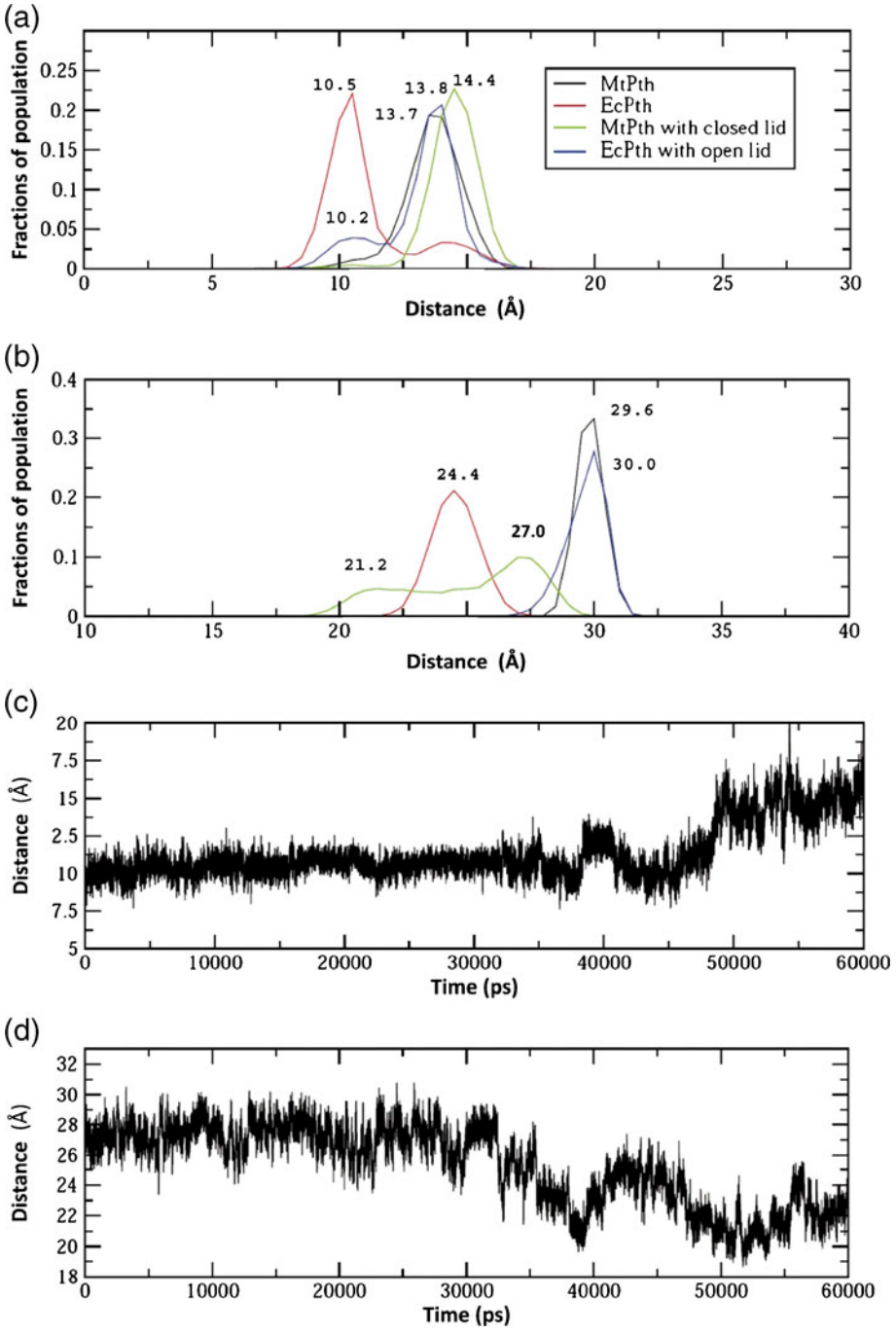

Figure 4. Population distributions of (**a**) the gate and (**b**) the lid distances in MD simulations. The time evolution of (**c**) the gate distance in the simulation of *Ec*Pth and (**d**) the lid distance in the *Mt*Pth with a closed lid.

occupying about 40% of the volume as estimated by the Matthews method (Matthews 1968). Thus, the volume of the solvent around the molecule in the crystal is less than the volume of the molecule. The volume of water surrounding the one molecule in the box in which MD simulations were carried out (3×10^5 Å^3) is more than an order of magnitude higher. At a concentration of 0.8 mM used in the experiment, the volume of the solution per molecule in the medium in which the NMR experiments were performed (2×10^6 Å^3) works out to be still higher by close to one more order of magnitude (Bal *et al*. 2006). Thus there appears to be a correlation, though not linear, between the ellipsoidal volume of the molecule and that of the surrounding medium.

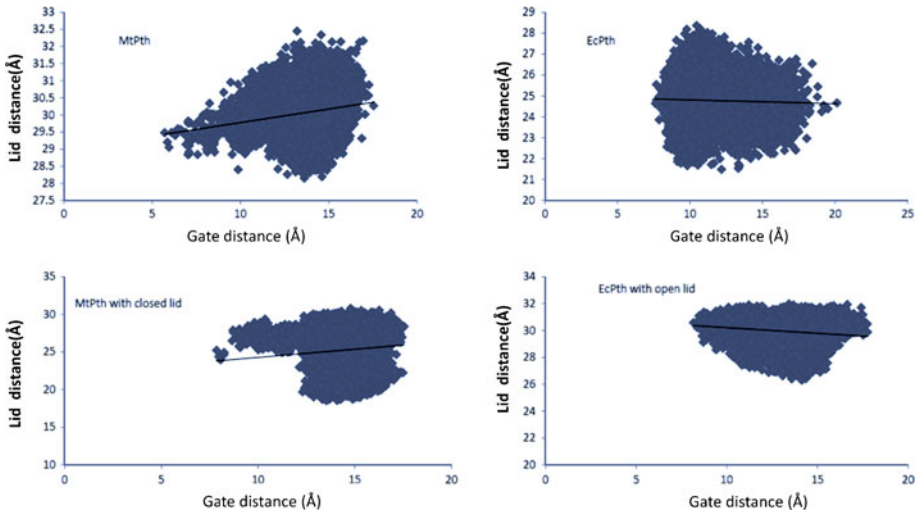

Figure 5. Correlation between gate and lid distances in simulations.

Protein structures can often be described as made up of a core region involving the basic fold elaborated through the addition of loops and other secondary structural elements. It was important to inquire if the correlation mentioned above holds even when the core region alone is considered. According to the SCOP classification (Murzin *et al*. 1995), *Mt*Pth has a phosphorylase/hydrolase-like fold involving a twisted β-sheet (residues 6–9, 41–42, 49–55, 59–63, 90–96, 105–108 and 131–136) flanked by two helices on either side (22–35 and 156–159, and 69–83 and 116–125) (figure 1a). Among these, 9 N-terminal residues of the long 156–179 helix juts out of the bulk of the molecule. The basic fold minus these nine residues can be considered as constituting the core of the molecule made up of close to 80 residues. The ellipsoidal volumes of this core in the X-ray, MD and NMR structures were also compared (table 1). The correlation exhibited by the volumes of the whole molecule is seen in the volumes of the core region as well, adding credence to the hypothesis regarding the relation between the ellipsoidal volume of the molecule and the volume of the surrounding medium. In the three situations considered, only in the crystals molecules and the surrounding medium have comparable volumes and each molecule is in contact with surrounding molecules. Does such crowding in general lead to a reduction in molecular volume as in the case of Pth? Is such crowding required for the correlated movement necessary for enzyme action, to manifest itself ?

3.4 *A comparative survey of X-ray and NMR structures*

In order to address the first question posed above, the structures of proteins determined by X-ray as well as NMR were examined. 1271 such structures are available in the PDB (Berman *et al*. 2002) as on May 31st, 2012. This data base covers a wide range of folds involving α, β, and α/β proteins and also different families with a variety of functions. Among the X-ray and NMR structures in the database, 533 are not directly comparable as the structure determinations involve different segments of the protein molecule. Ellipsoidal volumes of the X-ray and NMR structures were compared in all

Table 1. Ellipsoidal volume of the protein molecule along with the principal axes and the volume of the solvent around the molecule. The C-terminal stretch 180-191 has been removed from all the calculations as it is partly or wholly disordered in crystal structures (See text for details)

S. No.	Structure	Principal axes of ellipsoid (Å)			Ellipsoidal volume of the molecule (Å^3)	Volume of solvent around the molecule (Å^3)	Ellipsoidal volume of the core of the molecule (Å^3)
		a	b	c			
1	Form I *Mt*Pth	24.472	17.629	15.772	28,513	18,893	16,719
2.	*Mt*Pth MD structure (at 5000 ps)	25.356	17.912	15.959	30,373	3,00,675	17,432
3.	*Mt*Pth NMR structure (Model 1)	24.974	19.267	16.267	32,799	20,42,578	19,364

Table 2. Ellipsodial volume of the molecule and volume of solvent around the molecule in directly comparable X-ray and NMR structures (See text for details)

S.No.	Protein	PDB code		Ellipsoidal volume of the molecule ($Å^3$)		Unit cell^3 volume ($Å^3$)	Volume of solvent around the molecule ($Å^3$)	
		X-ray	NMR	X-ray	NMR		X-ray	NMR
1	NEDD4 binding protein from *Homo sapiens*	3FAU	2D9I	12687	14210	139205	24053	1646092
2	Basic fibroblast growth factor from *Homo sapiens*	1BAS	1BLA	19291	22980	33012	17481	1637322
3	Replication protein A from *Homo sapiens*	2B29	1EWI	20039	23916	183091	16091	1636386
4	Interleukin enhancer-binding factor 3 from *Homo sapiens*	3P1X	2 L33	11256	14452	154204	22293	2357408
5	APOA-1 from *Homo sapiens*	1AV1	1GW3	11868	18084	2177494	143830	313976
6	Tyrosine protein phosphatase type 4 from *Homo sapiens*	2VPH	2CS5	16465	18751	185529	21409	903639
7	Spliceosomal U5 snRNP specific 15 kDa protein from *Homo sapiens*	1QGV	1PQN	21219	24556	144565	19006	1635746
8	Peptidyl-tRNA hydrolase 2 from *Archaeoglobus fulgidus*	3ERJ	1RZW	20264	23576	216857	23613	1636726
9	CzrA protein from *Staphylococcus aureus*	1R1U	2KJB	22309	31595	333807	12593	1628707
10	Alpha-amylase from *Elusine coracana*	1B1U	1BIP	19224	25386	104153	55172	804765
11	Neuronal bungarotoxin from *Bungarus multicinctus*	1KBA	2NBT	10958	14488	220588	21089	193049
12	Interleukin-2 from *Homo sapiens*	3INK	1IRL	21991	25314	71045	36025	528120
13	Homeobox protein engrailed from *Homo sapiens*	1ENH	1ZTR	10139	12989	204124	9836	2754181
14	Arc represson from phage P22	1BAZ	1ARQ	15371	17034	228611	32607	621543
15	Regulatory protein E2 from Bovine papillomavirus type 1	1JJH	1DBD	13679	16093	1101126	62755	774527
16	Interleukin-4 from *Homo sapiens*	1HIJ	1BBN	23513	25870	393583	32175	804281
17	Excinuclease ABC subunit from *Escherchia coli*	1QOJ	1E52	6183	7749	331332	46182	1652553
18	Spectrin alpha chain from *Gallus gallus*	1BK2	2LJ3	9201	11706	70008	10602	1648596
19	Prolifin1A from *Acanthamoeba sp.*	1ACF	2PRF	18751	21662	108528	27132	1255493
20	Enzyme –I from *Escherchia coli*	1ZYM	1EZA	50615	56113	611899	86220	1050755
21	Apoliphophorin-III from *Locusta migratoria*	1AEP	1LS4	24989	30428	613983	32548	1629874
22	CAMP- dependent kinase type II from *Mus musculus*	2IZY	1R2A	9671	12363	1370428	57982	817788
23	Cardiotoxin V from *Geobacillus stearothermophilus*	1KXI	1CVO	9754	10868	277108	19235	404207
24	Beta-elicitin cryptogein from *Phytoptora cryptogea*	1BEO	1BEG	15068	17656	291813	25088	535778
25	Fimbrial protein from *Pseudomonas aeruginosa*	1QVE	1HPW	22976	28044	49568	20018	3292560
26	P85-alpha from *Bos taurus*	1QAD	1BFI	18301	20491	114300	15336	1639811
27	YfjZ from *E.coli*	2EA9	2JN7	16734	24794	82993	14969	1484571
28	Mesencephalic astrocyte-derived neurotrophic factor from *Homo sapiens*	2 W51	2KVD	34240	38408	282671	32141	1621894
29	HNRNP K from *Homo sapiens*	1ZZK	1KHM	14974	17213	66226	6591	1643089
30	Peptidyl-tRNA hydrolase from *Mycobacterium tuberculosis*	2Z2I	2JRC	28513	32799	166064	18893	2042578
31	ORF c02003 protein from *Sulfolobus solfataricus*	2Q00	2JPU	19720	24689	1367409	53363	1484676
32	Q251Q8_DESH from *Desulfitobacterium hafniense*	3IPF	2KS0	10226	12477	557575	28911	1647825
33	Plectasin from *Sulfolobus solfataricus*	3E7R	1ZFU	4959	5679	12707	9066	1654623
34	Sso10a from *Sulfolobus solfataricus*	1R7J	1XSX	20908	23793	126886	20134	667999
35	Riboflavin kinase from *Methanocaldococcus jannaschii*	2VBS	2P3M	24446	31833	646412	62956	2522477
36	Replication protein A from *Methanococcus maripaludis*	3E0E	2K5V	14717	16210	77972	9119	1390825
37	Lipoprotein from *Neisseria meningitidis*	3KVD	2KC0	46282	51917	426388	40825	4098838
38	PSPTO_3016 protein from *Pseudomonas syringae*	3H9X	2KFP	21548	23895	145762	79246	1485470
39	UPF0352 protein from *Shewanella oneidensis*	2QTI	2JUW	22185	24819	159942	13283	1635483
40	Putative Diflavin flavoprotein A3 from *Nostac sp*.	3FNI	2KLB	25254	29223	139988	15224	1377812
41	Putative chaperone Q8ZP25 from *Salmonealla typhimurium*	2ES7	2GZP	19414	23726	267354	75804	Not available

Table 2 (continued)

S.No.	Protein	PDB code		Ellipsoidal volume of the molecule (Å^3)		Unit cell^3 volume (Å)	Volume of solvent around the molecule (Å^3)	
		X-ray	NMR	X-ray	NMR		X-ray	NMR
42	Lin2157 protein from *Listeria innocua*	3I1E	2KJK	13736	15332	125747	17901	1391703
43	Conserved hypothetical protein PA2412 from *Pseudomonas aeroginosa*	2PST	2GPF	10190	13857	86594	14475	2753313
44	Conserved lipoprotein from *Ureaplasma parvum*	3JVC	2KRT	19558	22802	833565	57268	1821978
45	Nudix hydrolase DR0079 from *Deinococcus radiodurans*	2O5F	1Q27	28837	33020	674034	40819	797131
46	DR_A0006 from *Deinococcus radiodurans*	3GGN	2KCZ	19740	22137	134396	31713	1822643
47	Coactosin like protein from *Homo sapiens*	1T2L	1TMW	22533	26598	265180	32102	1080270
48	Acylphophatase 1 from *Homo sapiens*	2W4C	2K7K	15040	17906	84128	8247	2749264

the remaining 738. The ellipsoidal volume of the X-ray and the corresponding NMR structures differed by 10% or more in 206 proteins. A 10% difference was taken as highly significant and these 206 structures were carefully examined. In 158 cases, the X-ray and the NMR were not again directly comparable as there were differences in bound ligand, length of defined polypeptide chain, domain swapping, etc. In the remaining 48, which were comparable in almost every aspect, the NMR structure was larger than the X-ray structure in volume by 10% or more (table 2). Thus, it appeared that, in general, the ellipsoidal volumes of NMR structures are larger than those of the corresponding X-ray structures.

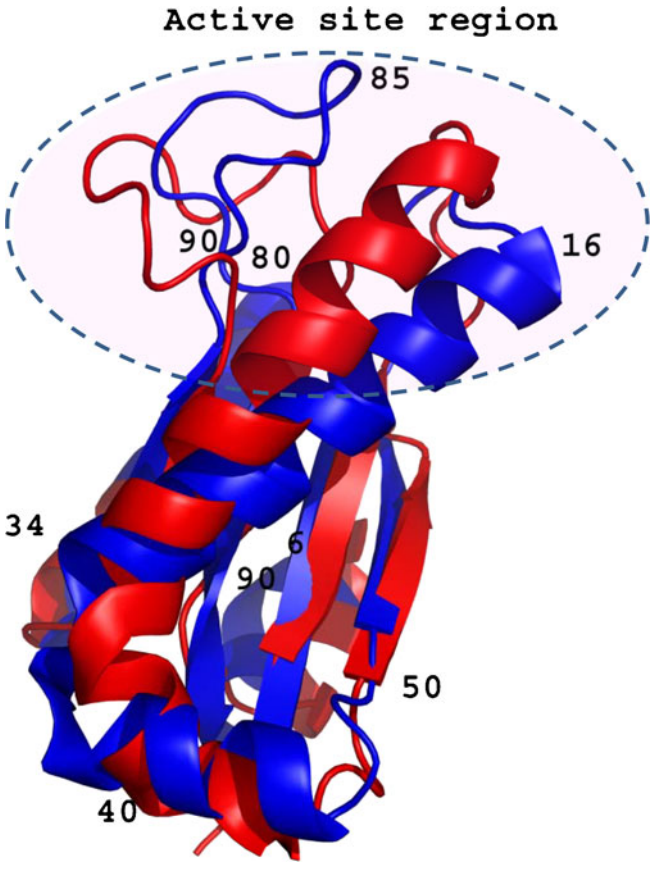

Figure 6. X-ray structure of *Af*Pth2 (blue) superposed over the NMR structure of *Af*Pth2 (red).

It was important to examine if the larger volume of NMR structures is caused by environmental effects. The composition of the crystallization medium and the NMR solvent vary so widely that it is unlikely to lead to a systematic effect. The situation is somewhat different in relation to temperature. Most X-ray structures have been determined at low temperature whereas most NMR structures have been elucidated at room temperature. There are, however, three structures [X-ray (NMR) structure Pdb codes: 2w4c (2k7k), 1prq (2prf) and 1a43 (2lf4)], which have been determined crystallographically and by NMR at the same or nearly the same temperature (Worthylake *et al*. 1999; Bennet *et al*. 2003; Lam *et al*. 2011) . In all the three instances, the ellipsoidal volume of the NMR structure is more than 10% larger than that of the X-ray structures as in the case of structures determined at different temperatures using X-ray and NMR. Furthermore, there are instances where the same protein crystals in the same space group grown under the same or similar conditions have been studied at different temperatures. For example, the ribosome recycling factor from *M. tuberculosis* has almost the same ellipsoidal volume in its crystal at 100 K (1wqg) and at room temperature (1wqf) (Saikrishnan *et al*. 2005). In the case of *M. smegmatis* RecA crystals, the volume at low temperature (2zrn) is slightly higher than at room temperature (2zr7) (Prabu *et al*. 2008). It would thus appear that difference in temperature has only a minor effect on the volume. However, the solvent volume available for each molecule in the crystal is roughly the same as the molecular volume, whereas it is one or two orders of magnitude higher in solutions used for NMR experiments. Thus, the lower molecular volume found in crystal structures, in comparison with that obtained from

NMR studies in solution, is most probably on account of macromolecular crowding.

Among the other proteins for which X-ray and NMR structures are available, there appears none to be directly comparable to eubacterial Pth with concerted motion of two regions of the molecule. The closest to eubacterial Pth among them is an archeal Pth. Archaeal and eukaryotic Pths, often referred to as Pth2, have similar structures which are very different from the structures of eubacterial Pth. The X-ray structures of dimeric Pth2 from *Archaeglobus flucidus* (3erj), *Sulfolobus solfataricus* (1xty), *Pyrococcus horikoshii* (1wn2) and humans (1q7s) are available (Pereda *et al*. 2003; Fromant *et al*. 2005; Shimuzu *et al*. 2008). These proteins which share 40% to 52% sequence identity have very similar structures. The NMR structure of *Af*Pth2 is also available (Powers *et al*. 2005). A superposition of X-ray and NMR structures of *Af*Pth2 is shown in figure 6. Although the rest of the structure superposes well, the active site region involving a loop and one end of the longest helix in the structure exhibit substantial differences. The loop has an open conformation in the NMR structure while it closes on the main body of the molecule in the X-ray structure. The end of the helix that forms part of the active site moves towards the rest of the molecule in the X-ray structure from its position in the NMR structure. Thus, as in the case of *Mt*Pth, *Af*Pth2 has a more voluminous open structure in solution than in crystals (table 2). A similar difference between NMR and X-ray structures involving a functionally important helix or loop has been reported in human hDim1 (Pdb codes 1qgv and 1pqn) (Reuter *et al*. 1999; Zhang *et al*. 2003) and a CTP-dependent archaeal riboflavin kinase as well (Pdb codes 2vbs and 2p3m) (Ammelburg *et al*. 2007).

4. Conclusion

The results presented here indicate that the molecular volume of proteins is dependent on the volume of the surrounding medium. The volume tends to be larger in solutions used in NMR experiments than in crystals where crowding exists. The recent observation that phosphoglycerate kinase has a more compact structure in cellular compartment than in aqueous solution, is of interest in this context (Dhar *et al*. 2011). It has been shown that the protein molecule could change its shape in a crowded environment (Homouz *et al*. 2008). The comparative study of NMR and X-ray structures indicate that the change in shape could specifically involve regions of the molecule important in activity, as indeed has been suggested by an earlier MD and theoretical investigation as well (Dong *et al*. 2010). The study also suggests that the functionally important plasticity of protein molecules could be modulated by crowding. This is in consonance with the observations of the effect of crowding on enzyme kinetics (Norris and Malys 2011). The effect of crowding on the structure, folding and action of proteins is complex and is characterized by subtlety and it needs to be pursued more extensively. However, the work reported here, in conjunction with results available in the literature, suggest that crowding leads to compaction of protein structure and influences the functionally important plasticity of the molecule.

Acknowledgements

Financial support from the Department of Biotechnology (DBT is acknowledged. Part of the computations was performed at the Interactive Graphics Facility supported by the DBT. MV is a DAE Homi Bhabha Professor.

References

Ai X, Zhou Z, Bai Y and Choy WY 2006 ^{15}N NMR spin relaxation dispersion study of the molecular crowding effects on protein folding under native conditions. *J. Am. Chem. Soc*. **126** 3916–3917

Ammelburg M, Hartmann MD , Djuranovic S, Alva V, Koretke KK, *et al*. 2007 A CTP-dependent archaeal riboflavin kinase forms a bridge in the evolution of cradle-loop barrels. *Structure* **15** 1557–1590

Bahadur RP, Chakrabarthi P, Rodier F and Janin J 2003 A dissection of specific and non-specific protein-protein interfaces. *J. Mol. Biol*. **336** 943–955

Bal NC, Agrawal H, Meher AK, Pulavarti SV, Jain A, Kelley G, *et al*. 2006 NMR assignment of peptidyl-tRNA hydrolase from *Mycobacterium tuberculosis* H37Rv. *J. Biomol. NMR* **36** Suppl 1:53

Bennet EM, Li C, Allan PW and Ealick SE 2003 Structural basis for substrate specificity of *Escherichia coli* purine nucleoside phosphorylase. *J. Biol. Chem*. **278** 47110–47118

Berk Hess HB, Berendsen HJC and Fraaije JGEM 1997 LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem*. **18** 1463–1472

Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, *et al*. 2002 The Protein Data Bank. *Acta Crystallog. Sect. D* **58** 899–90

Chetnani B, Kumar P, Surolia A and Vijayan M 2010 M. tuberculosis pantothenate kinase: Dual substrate specificity and unusual changes in ligand locations. J. Mol. Biol. **400** 171–185

Clarke TE, Romanov V, Lam R, Gothe SA, Peddi SR , Razumova EB, *et al*. 2011 Structure of *Francisella tularensis* peptidyl-tRNA hydrolase. *Acta Crystallog. Sect. F* **67** 913–915

Cohen GE 1997 ALIGN: a program to superimpose protein coordinates, accounting for insertions and deletions. *J. Appl. Cryst*. **30** 1160–1161

Darden T, York D and Pedersen L 1993 Particle mesh Ewald: a N.log(N) method for Ewald sums in large systems. *J. Chem. Phy*. **98** 10089–10092

Dhar A, Girdhar K, Singh D, Gelman H, Ebbinghaus S and Gruebele M 2011 Protein stability and folding kinetics in the nucleus and endoplasmic reticulum of eukaryotic cells. *Biophys. J.* **101** 421–430

Dong H, Qin S and Zhou HX 2010 Effect of macromolecular crowding on protein conformational changes. *PLoS Comput. Biol.* **6** e1000833

Ellis RJ and Minton AP 2003 Join the crowd. *Nature* **425** 27–28

Fromant M, Schmitt E, Mechulam Y, Lazennec C, Plateau P and Blanquet S 2005 Crystal structure at 1.8 Å resolution and identification of active site residues of Sulfolobus solfataricus peptidyl-tRNA hydrolase. Biochemistry **44** 4294–4301

Fu R, Wang X, Li C, Miranda AS, Pielak GJ and Tian F 2011 In situ structural characterization of a recombinant protein in native Escherichia coli membrane with solid state magic anlg spinning NMR. *J. Am. Chem. Soc*. **133** 12370–12373

Gershenson A and Gierasch LM 2011 Protein folding in the cell: challenges and progress. *Curr. Opin. Struct. Biol*. **21** 32–41

Homouz D, Perham M, Samiotakis A, Cheung MS and Pernilla WS 2008 Crowded, cell-like environment induces shape changes in aspherical protein. *Proc. Natl. Acad. Sci. USA* **105** 11754–11759

Homouz D, Sanabria H, Waxham MN and Cheung MS 2009 Modulation of calmodulin plasticity by the effect of macromolecular crowding. *J. Mol. Biol*. **392** 933–943

Jorgensen WL, Maxwell DS and Tirado-Rives J 1996 Development and testing of the OPLS all-atom force field on conformation energetic and properties of organic liquids. *J. Am. Chem. Soc*. **118** 11225–11236

Kaushal PS, Talawar RK, Krishna PD, Varshney U and Vijayan M 2008 Unique features of the structure and interactions of mycobacterial uracil-DNA glycosylase: Structure of a complex of the Mycobacterium tuberculosis enzyme in comparison with those from other sources. Acta Crystallogr. Sect. D: Biol. Crystallogr. **64** 551–560

Krishna R, Prabu JR, Manjunath GP, Datta S, Chandra NR, Muniyappa K and Vijayan M 2007 Snapshots of RecA protein involving movement of the C-domain and different conformations of the DNA-binding loops: crystallographic and comparative analysis of 11 structures of Mycobacterium smegmatis RecA. J. Mol. Biol. **367** 1130–1144

Lam SY, Yeung RC, Yu TH, Sze KH and Wong K 2011 A rigidifying salt-bridge favours the activity of thermophilic enzyme at high temperatures at the expense of low-temperature activity. *PLoS Biol*. **9** e1001027

Matthews BW 1968 Solvent content of protein crystals. *J. Mol. Biol*. **33** 491–497

Murzin AG, Brenner SE, Hubbard T, Chothia C 1995 SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* **247** 536–540

Norris MGS and Malys N 2011 What is the true enzyme kinetics in the biological system? An investigation of macromolecular crowding effect upon enzyme kinetics of glucose-6-phosphate dehydrogenase. *Biochem. Biophys. Res. Commun*. **405** 388–392

Pereda JM, Wass WF, Jan Y, Ruoslahti E, Schimmel P, and Pascual J 2003 Crystal structure of a human peptidyl-tRNA hydrolase reveals a new fold and suggests basis for a bifunctional activity. *J. Biol. Chem*. **279** 8111–8115

Pernilla WS 2011 Protein folding inside the cell. *Biophys. J*. **101** 265–266

Pettersen EF , Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, *et al*. 2004 UCSF Chimera- a visualization system for exploratory research and analysis. *J. Comput. Chem*. **25** 1605–1612

Powers R, Acton TB, Huang YJ, Liu J, Ma L, Rost B, *et al*. 2005 The solution structure of the *Archaeglobus flugidis* protein AF2095 Northeast structural genomics consortium target GR4. *Protein Sci.* **14** 2849–2861

Prabu JR, Thamotharan S, Khanduja JS, Chandra NR, Muniyappa K and Vijayan M 2009 Crystallographic and modelling studies on Mycobacterium tuberculosis RuvA: Additional role of RuvB-binding domain and inter species variability. Biochim. Biophys. Acta **1794** 1001–1009

Prabu JR, Manjunath GP, Chandra NR, Muniyappa K and Vijayan M 2008 Functionally important movements in RecA molecules and filaments: studies involving mutation and environmental changes. Acta Crystallogr. Sect. D **64** 1146–1157

Pulavarti SV, Jain A, Pathak PP, Mahmood A and Arora A 2009 Solution structure and dynamics of peptidyl-tRNA hydrolase from *Mycobacterium tuberculosis* H37Rv. *J. Mol. Biol.* **378** 165–177

Ratnaparkhi, GS, Ramachandran S, Udgaonkar, B and Varadarajan, R 1998 Discrepancies between the NMR and X-ray structures of uncomplexed Barstar: Analysis suggests that the packing densities of protein structures determined by NMR are unreliable. *Biochemistry* **37** 6958–6966

Reuter K, Nottrot S, Fabrizio P, Luhrmann R and Fincer 1999 Identification, characterization and crystal structure analysis of the human spliceosomal U5 snRNP-specific 15kD protein. *J. Mol. Biol*. **294** 515–525

Roque A, Ponte I, and Suau P 2007 Macromolecular crowding induces a molten globule state in the C-terminal domain of histone H1. *Biophys. J*. **93** 2170–2177

Roy S, Saraswathi R, Chatterji D and Vijayan M 2008 Structural studies on the second Mycobacterium smegmatis Dps: Invariant and variable features of structure, assembly and function. J. Mol. Biol. **375** 948–995

Saikrishnan K, Kalapala SK, Varshney U and Vijayan M 2005 X-ray structural studies of *Mycobacteriam tuberculosis* RRF and a comparative study of RRFs of known structure. Molecular plasticity and biological implications. *J. Mol. Biol*. **345** 29–38

Schmitt E, Mechulam Y, Fromant M, Plateau P and Blanquet S 1997 Crystal structure at 1.2Å resolution and active site mapping of *Escherchia coli* peptidyl-tRNA hydrolase. *EMBO J*. **16** 4760–4769

Schwieters CD and Clore MG 2008 A pseudopotential for improving the packing of ellipsoidal protein structures determined from NMR data. *J. Phys. Chem B* **112** 6070–6073

Selvaraj M, Ahmad R, Varshney U and Vijayan M 2012 Structures of new crystal forms of *Mycobacterium tuberculosis* peptidyl-tRNA hydrolase and functionally important plasticity of the molecule. *Acta Crystallog. Sect. F* **68** 124–128

Selvaraj M, Roy S, Sangeetha R, Varshney U and Vijayan M 2007 Structural Plasticity and Enzyme Action: Crystal Structures of

Mycobacterium tuberculosis Peptidyl-tRNA Hydrolase. *J. Mol. Biol*. **372** 186–193

Selvaraj M, Singh NS, Roy S, Sangeetha R, Varshney U and Vijayan M 2006 Cloning, expression, purification, crystallization and preliminary X-ray analysis of peptidyl-tRNA hydrolase from *Mycobaterium tuberculosis. Acta. Crystallog. Sect. F*. **62** 913–915

Shimuzu K, Kuroishi C, Sugahara M and Kunishima N 2008 Structure of peptidyl-tRNA hydrolase 2 from Pyrococcus horikoshii OT3: insight into the functional role of its dimeric state. Acta Crystallogr. Sect D **64** 444–453

Spoel VD, Lindahl E, Hess B, Groenhof G, Mark AE and Berendsen HJ 2005 Gromacs: Fast, flexible, and free. *J. Comput. Chem*. **26** 1701–1718

Vijayan M. Structural biology of mycobacterial proteins: The Bangalore effort 2005 *Tuberculosis* **85** 357–366

Worthylake DK, Wang H, Yoo S, Sundquist WI and Hill CP 1999 Structures of HIV-1 capsid protein dimerizationdomain at 2.6 Å resolution. *Acta Crystallog. Sect. D* **55** 85–92

Zhang, YZ, Cheng H, Gould, KL, Golemis EA and Roder H 2003 Structure, stability and function of hDim1 investigated by NMR, circular dichroism and mutational analysis. *Biochemistry* **42** 9609–9618

Zhou HX, Rivas G and Minton AP 2008 Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Annu. Rev. Biophys.* **37** 375–397

MS received 16 July 2012; accepted 07 October 2012

Corresponding editor: DURGADAS P KASBEKAR