Journal of Biomolecular NMR (2006) 36:199–204
DOI 10.1007/s10858-006-9073-2
© Springer 2006

NMR structure note

NMR solution structure of the acylphosphatase from *Escherichia coli*

Katiuscia Pagano[a], Matteo Ramazzotti[b], Paolo Viglino[a], Gennaro Esposito[a], Donatella Degl'Innocenti[b], Niccolò Taddei[b] & Alessandra Corazza[a,*]
[a]*Department of Biomedical Sciences and Technologies, University of Udine, Udine, Italy;* [b]*Department of Biochemical Sciences, University of Florence, Florence, Italy*

Received 11 May 2006; Accepted 17 July 2006

Key words: acylphosphatases, NMR structure

Abstract

The solution structure of *Escherichia coli* acylphosphatase (*E. coli* AcP), a small enzyme catalyzing the hydrolysis of acylphosphates, was determined by ^{1}H and ^{15}N NMR and restrained modelling calculation. In analogy with the other members of AcP family, *E. coli* AcP shows an α/β sandwich domain composed of four antiparallel and one parallel β-strand, assembled in a five-stranded β-sheet facing two antiparallel α-helices. The pairwise RMSD values calculated for the backbone atoms of *E. coli* and *Sulfolobus solfataricus* AcP, Bovine common type AcP and Horse muscle AcP are 2.18, 5.31 and 5.12 Å, respectively. No significant differences are present in the active site region and the catalytic residue side chains are consistently positioned in the structures.

Biological context

Acylphosphatase (AcP), one of the smallest enzyme known, is widely expressed in different organisms, spanning from insects to higher vertebrates, including man (Stefani et al., 1997). It catalyzes the hydrolysis of acylphosphates, both synthetic and of physiological relevance, such as benzoylphosphate, aspartylphosphate, and 1,3-bisphosphoglycerate (Stefani et al., 1997). Although it has been proposed that AcP might play a role in the regulation of calcium transport across membranes or in the control of the glycolytic flux, the exact function of the enzyme remains to be established.

In human and vertebrate tissues, AcP is expressed as two different isozymes named common-type (ctAcP) and muscle (mAcP) forms (Stefani et al., 1997). These share over 50% sequence identity and a common three-dimensional structure consisting of a globular α/β fold formed by a five stranded β-sheet, composed by four antiparallel and one parallel β-strand, packed against two antiparallel α-helices (Pastore at al., 1992; Thunnissen et al., 1997). The sequence and structural conservation is maintained to a large extent in all protein variants isolated so far from different sources, including the *Drosophila melanogaster* AcP (Dro2 AcP) purified and characterized more recently (Degl'Innocenti et al., 2003). Moreover, it has been established that AcP Open Reading Frames are widely present in the prokaryotic world. This has allowed the structures of some putative AcPs from archea such as *Pyrococcus horikoshii* (Ph AcP) and *Sulfolobus solfataricus* (Sso AcP) to be determined (Cheung et al., 2005; Corazza et al., 2006).

An AcP from *Escherichia coli* has been isolated and characterized very recently (manuscript submitted for publication). Despite remarkable

*To whom correspondence should be addressed. E-mail: acorazza@mail.dstb.uniud.it

sequence conservation, this variant shows at least two distinctive differences compared to the AcPs from eukarya: a higher conformational stability and a dramatically reduced catalytic activity; in addition, the protein contains two cysteine residues forming a disulfide bond in the native state, an unprecedented feature in the whole AcP family (manuscript submitted for publication).

In the last 10 years, AcP family members have been proven to be very useful models for studies on protein folding (Bemporad et al., 2004 and references therein) and amyloid aggregation (Chiti et al., 2003 and references therein). mAcP has been one of the first proteins not associated to disease to be induced to aggregate *in vitro* forming amyloid fibrils similar to those found in human pathology. A later study on mAcP based on a systematic mutational approach, enabled Chiti and coworkers to define an equation for the prediction of the effects of point mutations on the aggregation propensity of a protein of choice (Chiti et al., 2003). Intriguingly, further investigations on aggregation have shown that AcPs from different sources behave in a rather different manner (Plakoutsi et al., 2004), suggesting that comparative studies of proteins belonging to the same family can provide important clues for the elucidation of the molecular mechanism underlying protein aggregation. From this point of view, the three-dimensional structure of AcP from *E. coli* represents a fundamental prerequisite for any study aimed at defining the folding process and the aggregation tendency of a member of the AcP family that shows distinctive features.

Methods and results

E. coli AcP, cloned in pET28a vector (Novagen), has been expressed in *E. coli* BL21(DE3) strain as HIS-fused recombinant protein. The protein has been purified from bacterial lysates with Ni-NTA affinity matrix (Qiagen) and the HIS-tag has been removed by O/N treatment with thrombin (manuscript submitted for publication).

The amino acidic sequence of *E. coli* AcP is GSH-MSKVC IIAWVY GRVQG VGFRY TTQYE AKRLG LTGYA KNLDD GSVEV VACGE EGQVE KLMQW LKSGG PRSAR VERVL SEPHH PSGEL TDFRI R, in which the N-terminal Gly-Ser-His tripeptide results from gene cloning.

All NMR spectra were acquired at 310 K with a Bruker Avance-500 spectrometer, working at 11.74 T, on two 0.4 mM *E. coli* AcP samples (unlabelled and ^{15}N labelled, respectively) dissolved in H_2O/D_2O (95/5) containing 50 mM sodium phosphate buffer, 30 mM deuterated sodium acetate and 50 mM NaCl, at pH 4.95 where the enzyme has its activity maximum (manuscript submitted for publication). For structure determination, a set of 2D ^1H TOCSY, ^1H DQF-COSY, ^1H NOESY and ^1H-^{15}N HSQC, and a 3D ^1H-^{15}N NOESY-HSQC was used. The ^1H-^{15}N HSQC spectrum is shown in Figure 1. 2D ^1H NOESY and 3D ^1H-^{15}N-NOESY-HSQC spectra were recorded with a mixing time of 150 ms. Data were processed with XWINNMR (Bruker Biospin) and analyzed with Felix (Accelrys, San Diego, CA). All spectra were referenced on Ala36 $(H^{\beta})_3$ (1.473 ppm) or alternatively on Ala36 H^N (7.869 ppm) in the case of ^{15}N edited experiments. These two resonances,

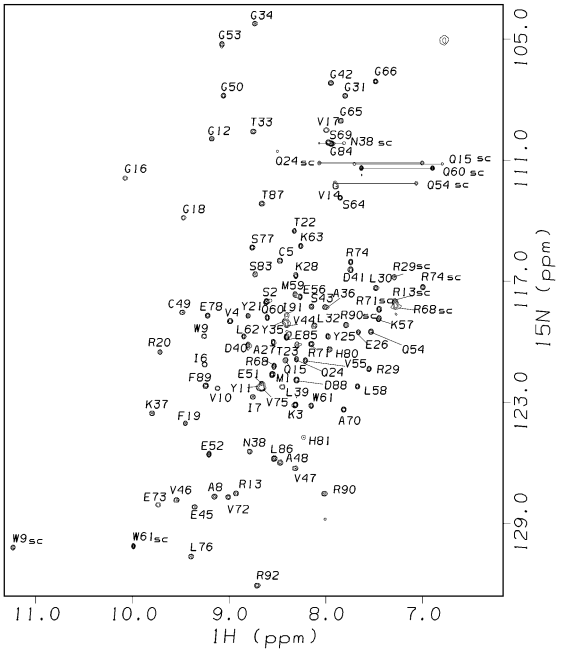

Figure 1. The ^1H-^{15}N HSQC spectrum of *E. coli* AcP. The spectrum was acquired with water suppression obtained with a flip-back pulse and t_1 quadrature detection using Echo/Antiecho-TPPI gradient selection. The acquisition was performed over a spectral width of 7002.801 and 1700.000 Hz in t_2 and t_1 respectively, with matrix size of 2048 and 256 points in t_2 and t_1, respectively, and 48 scans/t_1 FID. The assignments of the backbone amide groups are labelled. The "sc" labels indicate sidechain peaks from tryptophans, asparagines, glutamines or arginines residues. The latter ones are folded downfield by 33.55 ppm.

in turn, had been previously referenced using the dioxane peak as internal reference. ^{15}N chemical shifts were referenced indirectly to Ala36 H^N using the absolute frequency ratio.

Internuclear distances were quantified from NOESY spectra using cross-peaks from Gly66 ($H^{\alpha 2}/H^{\alpha 3}$) and Cys49/Ala48 ($H^N/H^\alpha$) as calibrant (the last one was used to calibrate the 3D ^1H-^{15}N NOESY-HSQC based on the 2D ^1H NOESY). An uncertainty of +/− 20% was allowed to obtain upper and lower limits and 970 upper distance bounds were gained for the final structure calculation. 59 ϕ torsion angles were extracted from 2D ^1H TOCSY by evaluating $J_{HN\text{-}H\alpha}$ from H^N-H^α cross-peak line width at half-height ($\Delta\nu_{1/2}$) (Wang et al., 1997). An uncertainty of 30° was assigned to every ϕ angle. From the analysis of the relative intensities of ^1H DQF-COSY, ^1H TOCSY and ^1H NOESY H^α-$H^{\beta 2/\beta 3}$ and H^N-$H^{\beta 2/\beta 3}$ cross peaks 8 atoms or group pairs were stereospecifically assigned.

The chemical shift assignments and $^3J_{HN\text{-}H\alpha}$ values were deposited in the BioMagResBank (accession code BMRB-7094).

All available experimental information was used as input for restrained dihedral angle molecular dynamics simulations, performed by using the program CYANA 2.1 (Güntert et al., 1997). The 20 best structures of the CYANA ensemble, in terms of target function, were submitted to restrained minimization using DISCOVER (MSI, San Diego CA) with the AMBER force field (Weiner et al., 1984). The analysis of violations and structural parameters of the 20 conformers obtained from CYANA runs are summarized in Table 1. The final structures were validated using AQUA and PROCHECK-NMR (Laskowski et al. 1996) and the main indicators of structure quality are reported in Table 1: the conformational family of *E. coli* AcP in solution matches the accepted criteria for structure validation. The refined ensemble structures have an RMSD from their mean structure, calculated on all residues (1–92), of 0.66 Å and 1.17 Å for backbone and non-hydrogen atoms, respectively. These values indicate a very limited spread within the conformer family (Figure 2b).

The solution structure coordinates were deposited in the Protein Data Bank with the accession code 2GV1.

Table 1. Structure quality and validation parameters for *E. coli* AcP[a].

CYANA Quality Parameter		
Average target function (Å^2)	0.97 ± 0.20 [0.55–1.24]	
Average number of violation:		
Distance upper limit >0.2 Å	1.69 ± 1.31 [0–5]	
Distance lower limit >0.2 Å	1.50 ± 0.50 [0–2]	
Van der Walls >0.2 Å	3.57 ± 1.68 [0–5]	
Dihedral angles >5 deg	0	
Average RMSD to mean[b]*:*		
Backbone (Å)	0.62 ± 0.12	
Heavy atoms (Å)	1.13 ± 0.18	
PROCHECK Statistics[c]	*This work*	*Consensus standard*[d]
Total residue/selected residues	92/92	74 ± 39/62 ± 36
Number of conformers	20	20 ± 15
Restraints per residue[e]	10.5	11.3 ± 4.5
NOE rms violation (Å)	0.047	0.061 ± 0.043
Average of bad steric contacts/100 residues	0.3 ± 0.5	4.2 ± 10.0
Average overall G-factor	−0.12 ± 0.03	−0.4 ± 0.3
Ramachandran quality:		
Residues in most-favored regions	78.8%	73.6 ± 15.5%
Residues in additional allowed regions	20.3%	n.a.
Residues in generously allowed regions	0.5%	n.a.
Residues in disallowed regions	0.5%	n.a.

[a]The experimental set of conformational restraints for torsion angle molecular dynamics comprised 970 inter-atomic distances (427 intra-residue [i–j = 0], 291 sequential [|i–j| = 1], 78 medium range [|i–j| < 5], 174 inter-residue long range [|i–j| ≥ 5]), 59 dihedral angle restraints, 8 stereospecific pair assignments and one disulfide bridge. Minimum and maximum values of the target function, violation numbers, and extents are given in brackets. n.a. not applicable. n.d. not determined. [b]The RMSD values reported in table and in the text are referred to the unrefined and refined ensemble, respectively. [c]The PROCHECK validation parameters were calculated on the refined structure ensemble. [d]From Doreleijers et al. (1998). [e]To avoid the redundancy of considering upper and lower bounds, only a single distance bound per internuclear separation was included in the count. Thus, 970 NOEs and 59 dihedral angle restraints were included in the count.

All structures were visualized and analyzed within the INSIGHTII framework (Dayringer et al., 1986), MOLMOL (Koradi et al., 1996). The secondary structures of the NMR structural family were evaluated according to the Kabsch and Sander classification (Kabsch et al., 1983). Similarly to other structurally determined members of the AcP family (Corazza et al., 2006 and references therein), *E. coli* AcP displays an α/β sandwich domain composed of four antiparallel and one parallel β-strand, assembled in a five-stranded β-sheet facing two antiparallel α-helices. As reported in Figure 2a, the sheet arrangement has the following topology: S4(Arg71-His80), S1(Cys5-Arg13), S3(Val44-Ala48), S2(Gly34-Asn38), S5(Arg90-Ile91). Helices are composed as follows: H1(Thr23-Leu30), H2(Glu52-Ser64). The helices, running parallel to the sheet, protect one side of the latter while the other side is exposed to the solvent. An extended interaction surface between the sheet and the helices provides the hydrophobic core of the enzyme.

By analogy with other AcP structures, the six loops connecting the secondary structure elements have been labelled L1 through L6, starting from the N-terminus. A disulfide bridge, between Cys5 and Cys94, links the N-terminal region of the protein and the L4 loop.

The structural based sequence alignment of *E. coli* AcP, Sso AcP, Ph AcP, Dro2 AcP, Horse mAcP, Bovine ctAcP, and Human mAcP and ctAcP is reported in Figure 3c. As for the other AcPs, whose structure was previously solved, in *E. coli* AcP residues 14–20 match the generally AcP conserved sequence Val-Gln-Gly-Val-X-X-Arg (corresponding to residues 24–30 in Sso AcP, to residues 21–27 in Dro2 AcP, and to residues 17–23 in Horse mAcP, Bovine ctAcP, and Human mAcP and ctAcP) and define a cradle-like conformation of the polypeptide backbone (Figure 3a). In this region, the nitrogen atoms of the main chain point toward the centre of the cradle, to which the phosphate moiety of the substrate is expected to bind, on the basis of homologous AcP structures and of site-directed mutagenesis studies (Taddei et al. 1997). The conserved residues Arg20 and Asn38 are recognized as the main phosphate binding residue and as the residue involved in the orientation and stabilization of catalytic water molecule, respectively (Stefani et al., 1997).

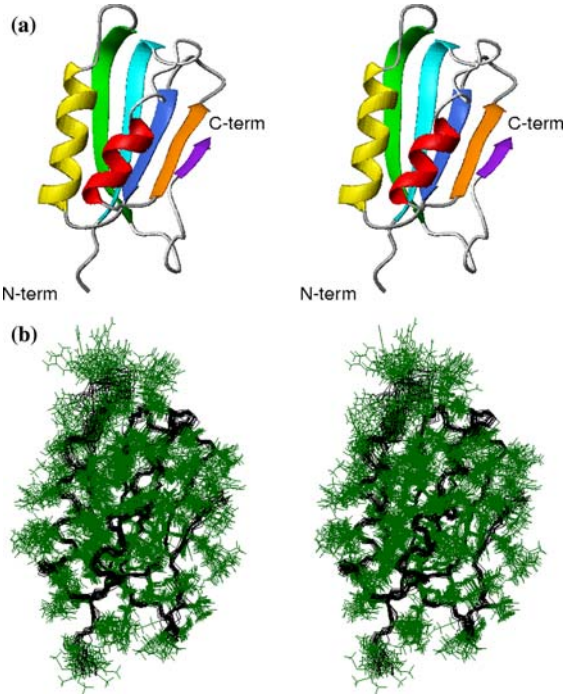

Figure 2. (a) Stereo view of the ribbon diagram of *E. coli* AcP showing the βαββαββ fold. The secondary structure elements are drawn according to the following colour code: Strand 1, Helix1, Strand 2, Strand 3, Helix2, Strand 4, Strand 5. (b) Stereo view of the best-fit superposition of the 20 conformers of *E. coli* AcP after DISCOVER refinement of the NMR-restrained ensemble obtained from CYANA. Backbone atoms are black coloured and the side chain atoms are green coloured.

E. coli Acp surface, depicted in Figure 3a, is colour coded according to the electrostatic potential, calculated using UHBD software package (University of Houston Brownian Dynamics) that solves the Poisson-Boltzmann equation with a finite difference algorithm (Madura et al. 1995). The calculation was performed at ionic strength of 50 mM, 298 K and the protonation state of the protein corresponds to pH 5.3. Results were visualized using the programs GRASP (Nicholls 1993) and DINO (Visualizing Structural Biology, 2002 http://www.dino3d.org).

Discussion and conclusions

The superposition of *E. coli* and hyperthermophilic Sso AcP mean structures with the two mesophilic Horse mAcP and Bovine ctAcP highlights an overall conformational similarity

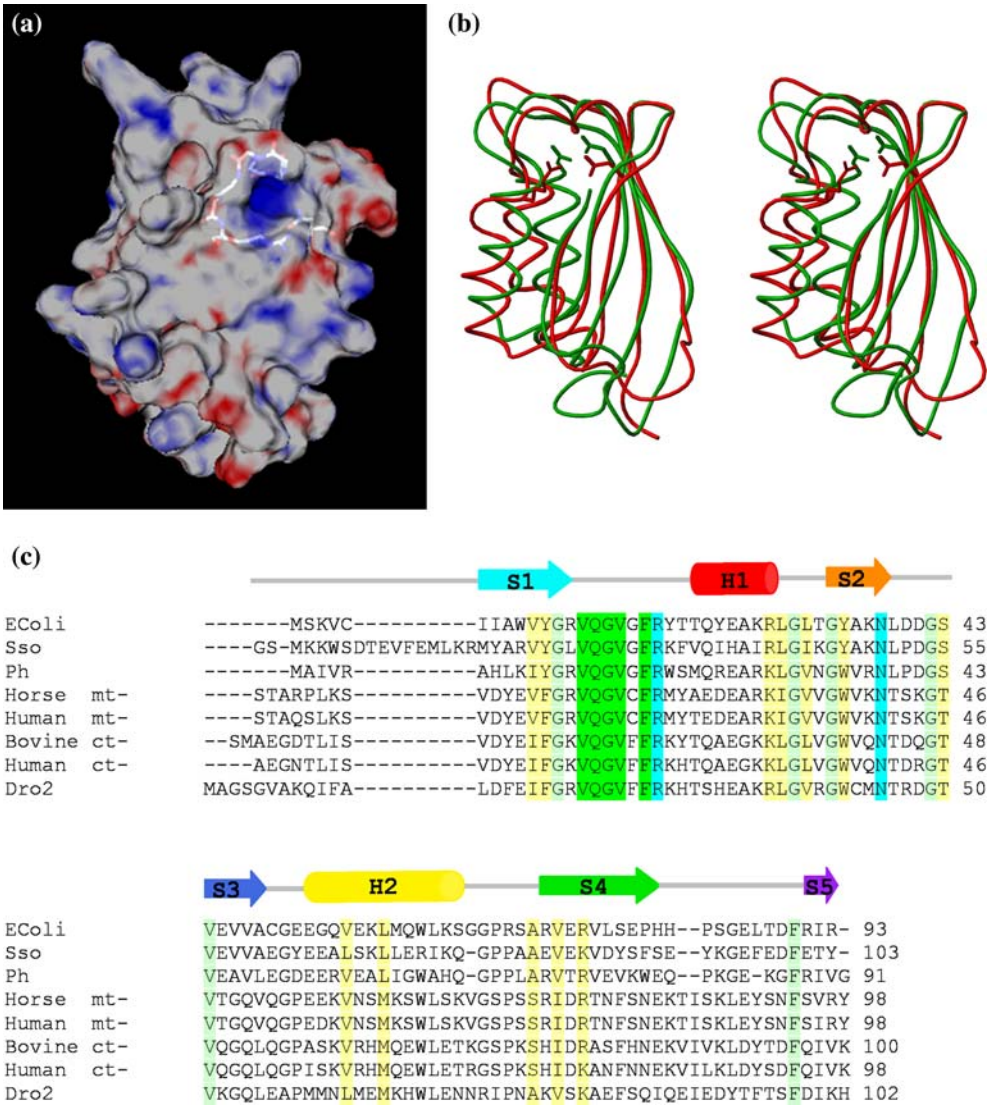

Figure 3. (a) View of the electrostatic charge distribution at the *E. coli* AcP surface with the 14–20 polypeptide cradle depicted in transparency. The active site cavity is characterized by an average electrostatic potential of 9.7 ± 2.4 kcal mol^{-1} q^{-1} calculated with UHBD over the 20 structures of the ensemble (the elementary charge q is set to 1.602×10^{-19}C). (b) Stereo view of the overlaid C^{α} traces for *E. coli* AcP (red) and Bovine ctAcP (green): the side chains of the catalytic residues are showed. (c) Structure-based sequence alignment of selected AcP's from different sources (*E. coli Escherichia coli*; Sso *Sulfolobus solfataricus*: Ph *Pyrococcus horikoshii*; Horse mt- horse muscle type; Human mt- human muscle type; Bovine ct- bovine common type; Human ct- human common type; Dro2 *Drosophila melanogaster*. The residues are highlighted according to the following colour code: conserved residues; conservatively replaced residues; catalytic residues; active site cradle. The secondary structure elements of *E. coli* AcP are reported on the top of the figure.

among *E. coli* and other AcPs from different phyla. The pairwise RMSD values calculated for the backbone atoms of *E. coli* and Sso AcP, Bovine ctAcP and Horse mAcP are 2.18, 5.31 and 5.12 Å, respectively. A stereo view of the backbone superposition between the most distant pairs of structures, i.e. *E. coli* and Bovine ctAcP, is reported in Figure 3b. No significant differences are present in the active site region and the catalytic residue side chains are consistently positioned in the structures, as it is evidenced in Figure 3b. The main variation is the lack in *E. coli* AcP of the first H1 helix turn, that include the catalytic residue Arg20. Although the geom-

etry of the residues involved in this missing helix turn does not fulfill all the requested criteria, according to Kabsh & Sander secondary structure definition, most of characteristics helix-diagnostic NMR connectivities are observed for the same residues. The active site cavity, highlighted by the backbone residues of the segment 14–20 in transparency in Figure 3a, is characterized by a net positive charge that can help to drive the acylphosphate substrate into the catalytic pocket.

The loop L5, connecting helix H2 and strand S4, is highly variable among the four considered molecules: in particular, in *E. coli* AcP, L5 presents a high conformational mobility (local rmsd = 1.16 Å) probably due to the presence of an additional glycine residue. The other loops are instead very tightly packed within the family as it is illustrated by their RMSD values (L1 0.30 Å, L2 0.09 Å, L3 0.09 Å, L4 0.08 Å, L6 0.17 Å).

Peculiarly, *E. coli* AcP S4 β-strand includes a proline residue (Pro79) generally considered the residue with the least β-sheet propensity, together with glycine (Smith et al., 1994). The structural alignment of known AcPs does not evidence however the conservation of this proline position among the phyla.

Only 4 salt bridges on the molecular surface of *E. coli* can be recognized: the first and the second one involve the side chains of Glu26 and Arg29, belonging to H1 helix, and Glu56 and Lys63 belonging to H2 helix. The third one stabilizes the S4 strand (Glu73 and Arg74 side chains) and the last one links the catalytic residue Arg20 to the C-terminal carboxylate of Arg92. Thus, three of the four salt bridges stabilize secondary structure elements whereas the fourth one, linking the C-terminal region, in analogy with Sso AcP, contributes to the three-dimensional topology stabilization (Taddei et al. 1996).

Several previously studied acylphosphatases include a single cystein residue in the primary sequence, e.g. in muscle type and in Dro2 AcPs. Instead, *E. coli* AcP sequence contains two cysteins residues (Cys5 and Cys49) that, at least under non-reducing conditions *in vitro*, arrange to form a disulfide bridge that binds the N-terminal S1 strand and the L4 loop. This latter region is stabilized instead in the case of Sso and mAcP, by the presence of a salt bridge (Corazza et al. 2006).

Acknowledgments

This work was supported by grants from the Italian Ministry of Education PRIN2004 050405, FIRB RBNE01S29H, FIRB RBAU0115B47, PRIN 2005 027330.

References

Bemporad, F., Capanni, C., Calamai, M., Tutino, M.L., Stefani, M. and Chiti, F. (2004) *Biochemistry*, **43**, 9116–9126.
Cheung, Y.Y., Lam, S.Y., Chu, W.K., Allen, M.D., Bycroft, M. and Wong, K.B. (2005) *Biochemistry*, **44**, 4601–4611.
Chiti, F., Stefani, M., Taddei, N., Ramponi, G. and Dobson, C.M. (2003) *Nature*, **424**, 805–808.
Corazza, A., Rosano, C., Pagano, K., Alverdi, V., Esposito, G., Capanni, C., Bemporad, F., Plakoutsi, G., Stefani, M., Chiti, F., Zuccotti, S., Bolognesi, M. and Viglino, P. (2006) *Proteins*, **62**, 64–79.
Dayringer, H.E., Tramontano, A., Sprang, S.R. and Fletterick, R. (1986) *J. Mol. Graphics*, **6**, 82–88.
Degl'Innocenti, D., Ramazzotti, M., Marocchini, R., Chiti, F., Raugei, G. and Ramponi, G. (2003) *FEBS Lett.*, **535**, 171–174.
Doreleijers, J.F., Rullmann, J.A.C. and Kaptein, R. (1998) *J. Mol. Biol.*, **281**, 149–164.
Güntert, P., Mumenthaler, C. and Wüthrich, K. (1997) *J. Mol. Biol.*, **273**, 283–298.
Kabsch, W. and Sander, C. (1983) *Biopolymers*, **22**, 2577–2637.
Koradi, R., Billeter, M. and Wüthrich, K. (1996) *J. Mol. Graphics*, **14**, 51–55.
Laskowski, R.A., Rullmann, J.A.C., MacArthur, M.W., Kaptein, R. and Thornton, J.M. (1996) *J. Biomol. NMR*, **8**, 477–486.
Madura, J.D., Briggs, J.M., Wade, R.C., Davis, M.E., Luty, B.A., Ilin, A., Antosiewicz, J., Gilson, M.K., Bagheri, B., Scott, L.R. and McCammon, J.A. (1995) *Comp. Phys. Comm.*, **91**, 57–95.
Nicholls, A. (1993) *GRASP: Graphical Representation and Analysis of Surface Properties*, Columbia University, New York.
Pastore, A., Saudek, V., Ramponi, G. and Williams, R.J. (1992) *J. Mol. Biol.*, **224**, 427–440.
Plakoutsi, G., Taddei, N., Stefani, M. and Chiti, F. (2004) *J. Biol. Chem.*, **279**, 14111–14119.
Smith, C.K., Withka, J.M. and Regan, L. (1994) *Biochemistry*, **33**, 5510–5517.
Stefani, M., Taddei, N. and Ramponi, G. (1997) *Cell. Mol. Life Sci*, **53**, 141–151.
Taddei, N., Magherini, F., Chiti, F., Bucciantini, M., Raugei, G., Stefani, M. and Ramponi, G. (1996) *FEBS Lett.*, **384**, 172–176.
Taddei, N., Chiti, F., Magherini, F., Stefani, M., Thunnissen, M.M., Nordlund, P. and Ramponi, G. (1997) *Biochemistry*, **36**, 7217–7224.
Thunnissen, M.M., Taddei, N., Liguri, G., Ramponi, G. and Nordlund, P. (1997) *Structure*, **5**, 69–79.
Wang, Y., Nip, A.M. and Wishart, D.S. (1997) *J. Biomol. NMR*, **10**, 373–382.
Weiner, S.J., Kollman, P.A., Case, D.A., Singh, U.C., Ghio, C., Alagona, G., Profeta, S. Jr. and Weiner, P. (1984) *J. Am. Chem. Soc.*, **106**, 765–784.