Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Crystal structure of wild-type and mutant human Ap₄A hydrolase

Honghua Ge a,b,*, Xiaofang Chen a,†, Weili Yang b, Liwen Niu a,b, Maikun Teng b,†

a Institute of Health Sciences and Modern Experiment Technology Center, Anhui University, Hefei 230601, People’s Republic of China
b School of Life Sciences, University of Science and Technology of China, Hefei 230026, People’s Republic of China

Abstract

Ap₄A hydrolase (asymmetrical diadenosine tetraphosphate hydrolase, EC 3.6.1.17), an enzyme involved in a number of biological processes, is characterized as cleaving the polyphosphate chain at the fourth phosphate from the bound adenosine moiety. This paper presents the crystal structure of wild-type and E58A mutant human Ap₄A hydrolase. Similar to the canonical Nudix fold, human Ap₄A hydrolase shows the common x/β-x/sandwich architecture. Interestingly, two sulfate ions and one diphosphate coordinated with some conserved residues were observed in the active cleft, which affords a better understanding of a possible mode of substrate binding.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

The dinucleotide diadenosine tetraphosphate (Ap₄A) can be found in all living cells from Archaea to humans. It is produced enzymatically as a side product of protein synthesis catalyzed predominantly by aminoacyl-tRNA synthetases [1]. Ap₄A has been proposed to be an intracellular “alarmone” both in prokaryotes [2] and in eukaryotes [3]. Intracellularly it affects DNA repair [4], RNA processing, cell division [5], heat shock and oxidative stress [6], apoptosis [8,9] and transcriptional regulation [1]. In bacteria, Ap₄A levels may also play a role in invasion [10,11]. In addition, as an extracellular signaling molecule, Ap₄A may play an important role as a neurotransmitter in the cardiovascular system [12,13]. Ap₄A concentration appeared to increase after exposure of cells to various forms of metabolic stress such as heat, oxidative, nutritional, and DNA damage [7]. In addition, high intracellular levels of Ap₄A have been associated with reduced replication times [14]. Thus the regulation of Ap₄A levels must be tightly controlled.

Ap₄A hydrolase is responsible for metabolizing the “alarmone” nucleotide Ap₄A and is therefore involved in all the above biological processes. Recent evidence further implicates that the SARS-CoV 7a protein interacts with human Ap₄A-hydrolase and may participate in common pathways leading to cell cycle arrest and apoptosis [15]. Ap₄A hydrolase belongs to Nudix (nucleoside diphosphate linked to other moieties, X, and contains the conserved XGₓEXGᵤ) family of enzymes which catalyze the hydrolysis of nucleoside diphosphates linked to other moieties, X, and contains the conserved Nudix sequence GXₓEXGᵤ (where U represents a bulky aliphatic residue usually Ile, Leu or Val, and X represents any residue) [16]. It is characterized as cleaving the polyphosphate chain at the fourth phosphate from the bound adenosine moiety. According to phylogenetic analysis [17], Ap₄A hydrolases were classified into two distinct groups, animal-archaeal type and plant-bacterial type enzymes. The latter contain loop insertions that undergo a large translation on substrate binding and exhibit substantial dynamic changes [18]. Previous studies of Ap₄A structures from both groups revealed that the enzyme has the x/β-x-sandwich architecture of a Nudix fold [18–20]. However, these studies suggest that important differences exist in the binding sites between animal-archaeal type and plant-bacterial type enzymes. According to the analysis of the NMR structure of human Ap₄A hydrolase (E58A), a proposal suggests recognition trigger is not required as the side chains of some key residues are on average predisposed with the correct orientation to unite key interactions for binding [19].

In this paper, we present the crystal structure of wild-type human Ap₄A hydrolase and its E58A mutant with/without DPO. The first structure was determined by molecular replacement at 2.7 Å resolution using the Caenorhabditis elegans Ap₄A Hydrolase structure as a search model. The mutant structures were all determined by molecular replacement at 2.1 Å resolution using the wild-type structure as a search model. Interestingly, two sulfate ions (in wild-type protein) and one diphosphate (in E58A mutant) coordinated with some conserved residues were observed in the active site, which may have important implications in the substrate binding mode in this class of enzymes.

2. Materials and methods

2.1. Cloning, expression and purification

The DNA coding sequence for human Ap₄A hydrolase (NCBI entry code P50583) was amplified by PCR from human brain cDNA li-
library (BD biosciences) using the following oligonucleotide primers containing artificial Ndel and XhoI sites (in bold): 5′-CTGTGCATATG GCGTGGAGAGCATGGCTGTG-3′ and 5′-GACCTCGAGG GCCCTATG-GAGCAAAAGAAC-3′. The PCR product was digested with Ndel and XhoI and was ligated into the Ndel and XhoI sites of the bacterial expression vector pET22b (incorporating a C-terminal hexahistidine tag). The cloning junctions were confirmed by DNA sequencing.

The recombinant plasmid was transformed into Escherichia coli strain BL21 (DE3). Cells were grown at 310 K in Luria–Bertani medium containing 100 µg/mL ampicillin and were harvested and sonicated in 20 mM Tris–HCl buffer pH 8.0 containing 150 mM NaCl. The lysate was clarified by centrifugation at 15,000 g for 30 min. The soluble fraction was applied to nickel-chelating resin (Amersham Biosciences) preequilibrated with equilibration buffer [20 mM Tris pH 8.0, 20 mM imidazole, 150 mM NaCl]. The resin was washed with wash buffer [20 mM Tris–HCl pH 8.0, 50 mM imidazole, 150 mM NaCl] and the target protein was eluted with elution buffer [20 mM Tris–HCl pH 8.0, 200 mM imidazole, 150 mM NaCl]. The eluate was exchanged to low-salt non-imidazole buffer (20 mM Tris pH 8.0, 20 mM NaCl) and was further purified by chromatography on a MonoQ anion-exchange column (GE Healthcare). The target protein was eluted using a linear gradient of 20–800 mM NaCl in 20 mM Tris pH 8.0, and then was applied to a gel filtration column (16/60 Superdex 200, GE Healthcare) in 20 mM Tris–HCl buffer pH 8.0 containing 150 mM NaCl. The flow-through fraction, which contained the target protein, was further purified by crystallization buffer[20 mM Tris–HCl pH 8.0, 20 mM NaCl, 5 mM DTT] and concentrated for crystallization assays to 10 mg/mL by centrifugal ultrafiltration (Millipore). The protein concentration was determined by Bradford method (BioRad Protein Assay), using bovine serum albumin as standard. The presence and purity of the recombinant protein was then analyzed on SDS–PAGE (better than 95% purity) and was judged to be suitable for crystallization.

The E58A mutant was constructed by replacing the GAG codon for Gln58 with GCC. The construct was transformed into E. coli strain BL21 (DE3). The expression and purification protocols are the same as for the wild-type protein described earlier, except using the equilibration buffer instead of the wash buffer to wash the resin due to the lower affinity between the mutant protein and nickel-chelating resin.

2.2. Crystallization and data collection

The wild-type enzyme was crystallized in substrate-free form. Initial crystallization trials were set up with Crystal Screen I and Crystal Screen II reagent kits (Hampton Research) at 283 K by using the hanging-drop vapor-diffusion method. The best crystals were obtained by mixing 2.0 L of protein solution and an equal volume of reservoir solution containing 0.1 M Sodium citrate (pH 5.6), 1.2 M Lithium sulfate, 0.2 M ammonium sulfate and incubated at 277 K.

The crystals were harvested using cryoloops and immersed briefly in a cryoprotectant solution consisting of reservoir solution with the glycerol concentration raised to 15%. The crystals were subsequently flash-frozen and stored in liquid nitrogen and transferred to beamline 3W1A of the BSRF (Beijing Synchrotron Radiation Facility) for X-ray diffraction analysis and data collection.

After screening for diffraction quality, a complete data set to 2.7 Å resolution was collected using a single crystal maintained at 100 K at a wavelength of 1.0000 Å, using 1° oscillations and an exposure time of 8 s per image. The diffraction data were processed with the HKL2000 program [21].

The wild-type human Ap4A hydrolase crystal belongs to space group P43 with unit cell parameters a = b = 72.49, c = 133.49 Å. Assuming the presence of four 17 kDa molecules in the asymmetric unit, a Matthews coefficient of 2.58 Å³ Da⁻¹ and a solvent content of 52.33% were calculated [22].

The E58A mutant crystals were grown by mixing the protein solution with a reservoir solution containing 0.1 M Tris (pH 8.4), 2.0 M Ammonium phosphate monobasic, 5 mM magnesium chloride and incubated at 285 K. Data were collected from crystals soaked in mother liquor with 15% glycerol added prior to flash-cooling in a liquid nitrogen stream at 100 K. A diffraction dataset of the E58A mutant crystal to 2.1 Å was collected at beamline BL17U of the SSRF (Shanghai Synchrotron Radiation Facility) at a wavelength of 1.00584 Å using 1° oscillations with a crystal-to-detector distance of 250 mm and an exposure time of 1 s per image. The diffraction data were processed with iMoSFLM [23] and scaled with SCALA from the CCP4 program suite [24].

E58A-DPO crystals were obtained by soaking E58A mutant crystals with 10 mM AP4A and 5 mM magnesium chloride. Data were collected as detailed above for the E58A mutant crystal except at a wavelength of 0.97915 Å. The diffraction data were processed with the HKL2000 program [21].

Both mutant crystals belong to space group P43212 with two molecules in the asymmetric unit and with unit cell dimensions: E58A, a = b = 72.19, and c = 133.51 Å; E58A-DPO, a = b = 72.37 Å, and c = 133.38 Å, with Vm = 2.5 Å³ Da⁻¹ solvent content of approximately 51% (w/w) [22]. All crystal parameters and data collection statistics are summarized in Table 1.

2.3. Structure determination and refinement

Structure determination of the wild-type enzyme by molecular replacement was implemented with Phaser [25]. The C. elegans Ap4A Hydrolase structure (PDB accession code 1KT9 [26]), which had 48% identity to the target structure, was used as the search model. The program Arp-Warp was then used to build from the molecular replacement model in an automated fashion. The model was completed with iterative rounds of manual building in COOT [27] and refinement in REFMAC [28]. The final refined model contains four human Ap4A hydrolase molecules in the asymmetric unit and was refined to an R factor (Rfree) of 21.7% (28.4%). The final model contains 432 residues, 148 water molecules, 18 sulfate ions and one glycerol. The three N-terminal residues and the last 8 C-terminal amino acids (belong to the recombinant 6–His tag) were not observed in the electron-density map.

Both mutant structures were solved by molecular replacement (MR) using the refined wild-type enzyme structure as a search model. The final E58A model has an R factor of 18.9% and an Rfree of 23.5%, while the final E58A-DPO model was refined to an R factor (Rfree) of 19.1% (23.2%). Some regions (including residues 19–21 on chain B, several N- and C-terminal residues) were not modeled because of the poor electron density.

All final crystallographic models were evaluated using MolProbity [29] with all parameters within the expected value range at the resolutions. The refinement statistics are summarized in Table 1. The coordinates and structure factors have been deposited in the Protein Data Bank under the accession code 3U53, 4ICK and 4IJK.

2.4. Sequence analysis and structural presentation

Amino-acid sequences were aligned using MultAlin [30] and the structure-based sequence-alignment figure was generated using ESPript [31]. All illustrations were prepared with PyMOL [32].
Ap4A hydrolase also exists as a monomeric form in solution as evidenced by size-exclusion chromatography (data not shown). The E58A mutant crystallized with two monomers which interacted to form a higher order oligomeric assemblies in the crystal lattice, which was consistent with the purification of the enzyme as a monomer [18–20], the potential substrate binding pocket of human Ap4A hydrolase is a large cleft constructed by residues from β1, β3, β4, β5, β6, α1, α3, and several loops (L2, L3, L7). Two conserved aromatic residues (Tyr82, Phe128) lie at the top of the active site cleft. They play a major role in substrate binding as described previously [19,26].

The side chain of Tyr82 is well defined and its orientation would be correct for π-stacking interactions of the substrate adenine ring. However, benzene ring of Phe128 is roughly perpendicular to the phenol ring of Tyr82 based on the density for the side chain of Phe128 (although the density in monomer A and B is not as well defined as that in monomer C and D). In the E58A-DPO structure from the mutant protein crystal soaked with AP4A, which was consistent with the purification of the enzyme as a monomer [18–20], the potential substrate binding pocket of human Ap4A hydrolase is a large cleft constructed by residues from β1, β3, β4, β5, β6, α1, α3, and several loops (L2, L3, L7). Two conserved aromatic residues (Tyr82, Phe128) lie at the top of the active site cleft. They play a major role in substrate binding as described previously [19,26].

Table 1
Data collection and refinement statistics.

|                      | Wild-type | E58A     | E58A-DPO |
|----------------------|-----------|----------|----------|
| Wavelength (Å)       | 1.0000    | 1.00584  | 0.97915  |
| Space group          | P4₁       | P4₁2₁2   | P4₁2₁2   |
| Molecules in ASU     | 4         | 2        | 2        |
| Cell parameters a/b/c (Å) | 72.49/72.49/133.49 | 72.19/72.19/133.51 | 72.37/72.37/133.38 |
| Resolution range(out shell) (Å) | 50–2.7 (2.775–2.705) | 20–2.1 (2.154–2.10) | 30–2.1 (2.157–2.103) |
| No. of unique reflections | 17881    | 21385    | 20214    |
| Corresponding % solvent | 52.33    | 50.71    | 50.9     |
| Rmerge* (out shell) (%) | 9.83 (37.56) | 12.7 (42.9) | 8.8 (44.3) |
| I/Ω(t)               | 8.5 (2.5) | 15.7 (6.3) | 26 (3.8) |
| Redundancy           | 7.7       | 13.7     | 13.4     |
| Completeness (out shell) (%) | 100 (100) | 100 (100) | 99.9 (100) |

The side chain of Tyr82 is well defined and its orientation would be correct for π-stacking interactions of the substrate adenine ring. However, benzene ring of Phe128 is roughly perpendicular to the phenol ring of Tyr82 based on the density for the side chain of Phe128 (although the density in monomer A and B is not as well defined as that in monomer C and D). In the E58A-DPO structure from the mutant protein crystal soaked with AP4A, which was consistent with the purification of the enzyme as a monomer [18–20], the potential substrate binding pocket of human Ap4A hydrolase is a large cleft constructed by residues from β1, β3, β4, β5, β6, α1, α3, and several loops (L2, L3, L7). Two conserved aromatic residues (Tyr82, Phe128) lie at the top of the active site cleft. They play a major role in substrate binding as described previously [19,26].

3.3. Substrate binding cleft

Based on the analysis of other Ap4A hydrolases of known structure [18–20], the potential substrate binding pocket of human Ap4A hydrolase is a large cleft constructed by residues from β1, β3, β4, β5, β6, α1, α3, and several loops (L2, L3, L7). Two conserved aromatic residues (Tyr82, Phe128) lie at the top of the active site cleft. They play a major role in substrate binding as described previously [19,26].
Examination of the electron density map for the E58A-DPO structure revealed a density feature in the active cleft of Chain B. Although there was no interpretable density for the entire AP4A molecule, the density for two phosphate groups are clearly identifiable.

3.4. The sulfate- and diphosphate- binding sites in active cleft

In the crystal structure of the wild-type human Ap4A hydrolase, each monomer has one sulfate ion near the top of the substrate-binding pocket. The sulfate forms direct interactions with the protein via hydrogen bonds to His37 NE2, Lys42 NZ and Tyr82 OH (Fig. 2B). In addition, it has one oxygen atom at a distance of about 3.9 Å from the positively charged residue Lys89 NZ. In the E58A-DPO structure, one phosphate group of DPO was observed in the relevant position in Chain B, and forms the similar binding mode with the protein, while the other phosphate group, which may be the \( \beta \)-phosphate of the substrate, makes hydrogen binds to Lys42 NZ. (Fig. 2C).

According to the analysis of related Ap4A hydrolase-substrate complexes [18–20], the analogous positions for these residues are conserved and coordinate the \( \alpha \)-phosphate of the substrate. Hence, His37, Lys42, Tyr82 and Lys89 probably facilitate \( \alpha \)-phosphate group binding of the real substrate. The observation of the Lys42 also suggests it could stabilize the \( \alpha \)-phosphate as well as the \( \beta \)-phosphate of the substrate.

One of the monomers (chain B) in the wild-type enzyme has another sulfate ion near the Nudix sequence motif in the substrate-binding pocket, where its oxygen atoms are at hydrogen-bonding distance from the side chain ND1 and the main-chain amide of His44, OE2 of Glu110, OE1 and OE2 of Glu58. In monomer A, the equivalent position is most likely a sulfate although a water molecule was modeled in the density. In contrast, there is no relevant density in this region in monomers C and D, and the residues His44 and Glu110 show a different conformation compared to that in monomers A and B. In addition, inspection of the active site shows there is sufficient flexibility in the location of \( \beta \)- and \( \gamma \)-phosphates between the two sulfate ions sites. Thus, this sulfate position is at a distance of about 3.8–5 Å from some other Nudix signature sequence residues Arg57, Cln61 and Glu62. Probably due to the lack of magnesium ions, this sulfate...
position is closer to the Nudix sequence motif than the proposed δ-phosphate binding site in the crystal structure of C. elegans counterpart (Fig. 2D). This observation of the sulfate would suggest these residues (His44, Arg57, Glu58, Glu61, Glu62 and Glu110) are sufficient to stabilize the substrate and orientate the δ-phosphate by hydrogen bond or via metal coordination for nucleophilic attack.

Acknowledgments

We thank the staff at SSRF beamline BL17U for assistance with synchrotron data collection and Dr. Weiguo Li for plasmids. This work was supported by grants from the Chinese Ministry of Science and Technology (2012CB917200), the National Natural Science Foundation of China (31270770, 31130018 and 10979039), Anhui Provincial Natural Science Foundation (Grant No. 11040606M66) and the Natural Science Foundation of the Department of Education of Anhui Province (Grant No. KJ2011A015).

References

[1] I. Carmi-Levy, N. Yannay-Cohen, G. Kay, E. Razin, H. Nechushtan, Diadenosine tetraphosphate hydrolase is part of the transcriptional regulation network in immunologically activated mast cells, Mol. Cell Biol. 28 (2008) 5777–5784.
[2] P.C. Lee, B.R. Bocner, B.N. Ames, AppppA, heat-shock stress, and cell oxidation, Proc. Natl. Acad. Sci. USA 80 (1983) 7496–7500.
[3] A. Varshavsky, Diadenosine 5',5'-O-(P1), P4-tetraphosphate: a pleiotropically acting alarmone?, Cell 34 (1983) 711–712.
[4] A.G. McLennan, Dinucleoside polyphosphates-friend or foe?, Pharmacol. Ther. 87 (2000) 103–115.
[5] A. Nishimura, The timing of cell division: Ap4A as a signal, Trends Biochem. Sci. 23 (1998) 157–159.
[6] M.D. Baxi, J.K. Vishwanatha, Diadenosine polyphosphates: their biological and pharmacological significance, J. Pharmaco. Toxicol. Methods 33 (1995) 121–128.
[7] LL. Kisselev, J. Justesen, A.D. Wolfson, L.Y. Frolova, Diadenosine oligophosphates (Ap(n)A), a novel class of signalling molecules?, FEBS Lett 427 (1998) 157–163.
[8] A. Vartanian, I. Prudovsky, H. Suzuki, I. Dal Pra, L. Kisselev, Opposite effects of cell differentiation and apoptosis on Ap3A/Ap4A ratio in human cell cultures, FEBS Lett. 415 (1997) 160–162.
[9] A.G. McLennan, The MutT motif family of nucleotide phosphohydrolases in man and human pathogens (review), Int. J. Mol. Med. 4 (1999) 79–89.
[10] J.L. Badger, C.A. Wass, K.S. Kim, Identification of Escherichia coli K1 genes contributing to human brain microvascular endothelial cell invasion by differential fluorescence induction, Mol. Microbiol. 36 (2000) 174–182.
[11] T.M. Ismail, A.C. Hart, A.G. McLennan, Regulation of dinucleoside polyphosphate pools by the YgdP and Apaf hydrolyases is essential for the ability of Salmonella enterica serovar typhimurium to invade cultured mammalian cells, J. Biol. Chem. 278 (2003) 32602–32607.
[12] J. Pintor, M. Diaz-Hernandez, J. Gualian, R. Gomez-Villafuertes, F. Hernando, M.T. Miras-Portugal, Diadenosine polyphosphate receptors. from rat and guinea-pig brain to human nervous system, Pharmacol. Ther. 87 (2000) 103–115.
[13] R.M. Stavrou, Diadenosine polyphosphates: postulated mechanisms mediating the cardiac effects, Curr. Med. Chem. Cardiovasc. Hematol. Agents 1 (2003) 151–169.
[14] E. Rapaort, P.C. Zamecnik, Presence of diadenosine 5',5'-O-(P1), P4-tetraphosphate (Ap4A) in mamalian cells in levels varying widely with proliferative activity of the tissue: a possible positive "pleiotypic activator", Proc. Natl. Acad. Sci. USA 73 (1976) 3984–3988.
[15] N. Vasilenko, I. Moslynskyy, A. Zakhartchouk, SARS coronavirus protein 7a interacts with human Ap4A-hydrolase, Virol. J. 7 (2010) 31.
[16] W. Xu, C.A. Dunn, C.R. Jones, G. D’Souza, M.J. Bessman, The 26 Nudix hydrolases of Caenorhabditis elegans, Biochim. Biophys. Acta 1550 (2001) 27–36.
[17] J. Fletcher, J.D. Swarbrick, D. Maksel, K.R. Gayler, P.R. Gooley, The structure of Ap4A hydrolase complexed with ATP-MgF(x) reveals the basis of substrate binding, Structure 10 (2002) 205–213.
[18] J.D. Swarbrick, S. Buyya, D. Gunawardana, K.R. Gayler, A.G. McLennan, P.R. Gooley, Structure and substrate-binding mechanism of human Ap4A hydrolase, J. Biol. Chem. 280 (2005) 8471–8481.
J. Jeyakanthan, S.P. Kanaujia, Y. Nishida, N. Nakagawa, S. Praveen, A. Shinkai, S. Kuramitsu, S. Yokoyama, K. Sekar, Free and ATP-bound structures of Ap4A hydrolase from Aquifex aeolicus V5, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 116–124.

Z. Otwinowski, W. Minor, Processing X-ray diffraction data collected in oscillation mode, Methods Enzymol. 276 (1997) 307–326.

B.W. Matthews, Solvent content of protein crystals, J. Mol. Biol. 33 (1968) 491–497.

T.G. Battye, L. Kontogiannis, O. Johnson, H.R. Powell, A.G. Leslie, IMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM, Acta Crystallogr. D Biol. Crystallogr. 67 (2011) 271–281.

M.D. Winn, C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R.M. Keegan, E.B. Krissinel, A.G. Leslie, S.J. McNicholas, G.N. Murshudov, N.S. Pannu, A.E. Pannier, H.R. Powell, R.J. Read, A. Vagin, K.S. Wilson, Overview of the CCP4 suite and current developments, Acta Crystallogr. D Biol. Crystallogr. 67 (2011) 235–242.

A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R.J. Read, Phaser crystallographic software, J. Appl. Crystallogr. 40 (2007) 658–674.

S. Bailey, S.E. Sedelnikova, G.M. Blackburn, H.M. Abdelghany, P.J. Baker, A.G. McLennan, J.B. Rafferty, The crystal structure of diadenosine tetraphosphate hydrolase from Caenorhabditis elegans in free and binary complex forms, Structure 10 (2002) 589–600.

P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics, Acta Crystallogr. D Biol. Crystallogr. 60 (2004) 2126–2132.

G.N. Murshudov, A.A. Vagin, E.J. Dodson, Refinement of macromolecular structures by the maximum-likelihood method, Acta Crystallogr. D Biol. Crystallogr. 53 (1997) 240–255.

S.C. Lovell, I.W. Davis, W.B. Arendall 3rd, P.I. de Bakker, J.M. Word, M.G. Prisant, J.S. Richardson, D.C. Richardson, Structure validation by Calpha geometry: phi, psi and Cbeta deviation, Proteins 50 (2003) 437–450.

F. Corbet, Multiple sequence alignment with hierarchical clustering, Nucleic Acids Res. 16 (1988) 10881–10890.

P. Gouet, E. Courcelle, D.J. Stuart, F. Metoz, ESPript: analysis of multiple sequence alignments in postscript, Bioinformatics 15 (1999) 305–308.

W.L. DeLano, The PyMOL Molecular Graphics System, DeLano Scientific, Palo Alto, CA, USA, 2002.

C.S. Bond, Topdraw: a sketchpad for protein structure topology cartoons, Bioinformatics 19 (2003) 311–312.