

This article was originally published in a journal published by Elsevier, and the attached copy is provided by Elsevier for the author's benefit and for the benefit of the author's institution, for non-commercial research and educational use including without limitation use in instruction at your institution, sending it to specific colleagues that you know, and providing a copy to your institution's administrator.

All other uses, reproduction and distribution, including without limitation commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution's website or repository, are prohibited. For exceptions, permission may be sought for such use through Elsevier's permissions site at:

<http://www.elsevier.com/locate/permissionusematerial>

Structural Analysis of *N*-acetylglucosamine-6-phosphate Deacetylase Apoenzyme from *Escherichia coli*

Frederico M. Ferreira¹, Guillermo Mendoza-Hernandez²
Maria Castañeda-Bueno², Ricardo Aparicio³, Hannes Fischer¹
Mario L. Calcagno² and Glaucius Oliva^{1*}

¹Instituto de Física de São Carlos, Universidade de São Paulo, C.P. 369, 13560-970 São Carlos, SP, Brazil

²Departamento de Bioquímica Facultad de Medicina Universidad Nacional Autónoma de México AP. Postal 70-159 Cd. Universitaria, 04510 México, DF México

³Instituto de Química Universidade Estadual de Campinas, C.P. 6154, 13084-862 Campinas, SP, Brazil

We report the crystal structure of the apoenzyme of *N*-acetylglucosamine-6-phosphate (GlcNAc6P) deacetylase from *Escherichia coli* (*Ec*NAGPase) and the spectrometric evidence of the presence of Zn^{2+} in the native protein. The GlcNAc6P deacetylase is an enzyme of the amino sugar catabolic pathway that catalyzes the conversion of the GlcNAc6P into glucosamine 6-phosphate (GlcN6P). The crystal structure was phased by the single isomorphous replacement with anomalous scattering (SIRAS) method using low-resolution (2.9 Å) iodine anomalous scattering and it was refined against a native dataset up to 2.0 Å resolution. The structure is similar to two other NAGPases whose structures are known from *Thermotoga maritima* (*Tm*NAGPase) and *Bacillus subtilis* (*Bs*NAGPase); however, it shows a phosphate ion bound at the metal-binding site. Compared to these previous structures, the apoenzyme shows extensive conformational changes in two loops adjacent to the active site. The *E. coli* enzyme is a tetramer and its dimer–dimer interface was analyzed. The tetrameric structure was confirmed in solution by small-angle X-ray scattering data. Although no metal ions were detected in the present structure, experiments of photon-induced X-ray emission (PIXE) spectra and of inductively coupled plasma emission spectroscopy (ICP-AES) with enzyme that was neither exposed to chelating agents nor metal ions during purification, revealed the presence of 1.4 atoms of Zn per polypeptide chain. Enzyme inactivation by metal-sequestering agents and subsequent reactivation by the addition of several divalent cations, demonstrate the role of metal ions in *Ec*NAGPase structure and catalysis.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: *N*-acetylglucosamine-6-phosphate deacetylase; amino sugar catabolism; amidohydrolases; PIXE; ICP-AES

*Corresponding author

Introduction

The bacterium *Escherichia coli* displays a great versatility in the utilization of different

carbon sources and it can use a great variety of carbohydrates and related compounds as sources of energy. When amino sugars, glucosamine (GlcN) or *N*-acetylglucosamine (GlcNAc), are present in the environment, they are taken up by the cells and used for cell wall and lipid A synthesis as well as a carbon source. Under these conditions, the synthesis of the enzyme glucosamine synthase (GlmS) necessary for the biosynthesis of amino sugars, is repressed while the enzymes necessary for the transport and metabolism of amino sugars are induced.^{1,2} GlcNAc is amongst the best carbon sources known for *E. coli*, producing similar growth rates as glucose. In *E. coli* both GlcN and GlcNAc are phosphotransferase

Abbreviations used: GlcNAc6P, *N*-acetylglucosamine 6-phosphate; NAGPase, *N*-acetylglucosamine-6-phosphate deacetylase; SIRAS, single isomorphous replacement with anomalous scattering; SAXS, small-angle X-ray scattering; PIXE, photon-induced X-ray emission; ICP-AES, inductively coupled plasma emission spectroscopy; DRM, dummy residues model.

E-mail address of the corresponding author: oliva@ifsc.usp.br

system (PTS) sugars, so their transport into the cell is coupled to their phosphorylation. The *N*-acetylglucosamine 6-phosphate (GlcNAc6P) resulting from the transport of GlcNAc is hydrolyzed to GlcN6P and acetate by the enzyme GlcNAc6P deacetylase (EC 3.5.1.25) encoded by the gene *nagA*. GlcN6P, either produced from GlcNAc6P or from the transport of GlcN, is then deaminated and isomerized by the enzyme GlcN6P deaminase (EC 3.5.99.6, encoded by *nagB*), which produces free ammonia and fructose 6-phosphate.³ Utilization of GlcNAc has also been reported for other pathogenic bacteria, such as *Klebsiella pneumoniae* and *Vibrio cholerae* and it was demonstrated that it induces cellular morphogenesis in *Candida albicans* by disruption of the *N*-acetylglucosamine catabolic pathway.⁴ In the absence of an exogenous supply of amino sugars, GlcN6P can be synthesized from Fru6P and glutamine by the enzyme glutamine-fructose-6-phosphate aminotransferase (GlmS, E.C. 2.6.1.16), encoded by one of the genes in the biosynthetic operon, *glmUS*. In the presence of an exogenous supply of amino sugars the expression of the biosynthetic operon is repressed while the expression of the genes necessary for GlcNAc utilization as a carbon source are induced. The genes for the transport (*nagE*) and catabolism (*nagBA*) of GlcNAc are organized in two divergent operons *nagE-nagBACD*. The inducing signal turning on expression of the *nagE-nagBA* genes is GlcNAc6P, which prevents the transcriptional repressor NagC from binding to its operators.¹ GlcNAc6P that is the direct product of GlcNAc transport, has an additional regulatory role; it is the allosteric activator of GlcN6P deaminase (NagB).^{5,6} *Ec*NAGPase is not allosteric, but it is a key component of the coordinated control of amino sugar metabolism, because of the double regulator role of its substrate. The operation *in vivo* of both the transcriptional and allosteric control of amino sugar utilization in *E. coli* has been recently demonstrated.⁷ Although *Ec*NAGPase is primarily a catabolic enzyme, recently it has been shown to have an important role in the amino sugar salvage pathway of peptidoglycan recycling in *E. coli*.⁸ At each generation, *E. coli* breaks down about 60% of its cell wall peptidoglycan, most of which is efficiently recycled in a sequence of reactions that release GlcNAc and 1,6-*anhydro-N*-acetylmuramic acid in the cytoplasm.⁹ The GlcNAc recycling pathway involves the phosphorylation of these intermediates by two specific kinases, NagK and AnmK,^{9,10} and its conversion into GlcN6P by NAGPase (NagA), which is subsequently isomerized to GlcN1P (GlmM) and then *N*-acetylated with the participation of acetyl-coenzyme A (GlmU).^{11,12} In eukaryotic organisms GlcNAc6P is directly isomerized into GlcNAc1P;¹³ this metabolic difference between animals and bacteria points toward bacterial NAGPases as potential drug targets. *Ec*NAGPase has been obtained pure from an overproducing *E. coli* strain and shown to be a tetramer of identical subunits of 41 kDa and 382 residues. Its kinetics was studied both in the forward and the backward directions of the reaction and a kinetic mechanism was proposed.¹⁴ Although

the reverse acetylation reaction was demonstrated *in vitro*, there is no evidence that it could have a physiological role *in vivo*. The crystal structures of the NAGPases from *Thermotoga maritima* (PDB code 1O12, 2.5 Å resolution) and *Bacillus subtilis* (PDB code 1UN7, 2.0 Å resolution)¹⁵ have been deposited into the Protein Data Bank (PDB). Simultaneously with the deposition of the present structure (PDB code 1YRR, 2.0 Å resolution), a new entry for the *E. coli* deacetylase also appeared (PDB code 1YMY, 2.6 Å resolution). The *B. subtilis* enzyme contains a binuclear iron center, whereas the *T. maritima* structure has only one iron atom bound. These structures show that the GlcNAc6P deacetylases belong to the structural superfamily of metal-dependent amidohydrolases,¹⁵ whose members can bind one, two or three metal atoms in the active site, with a wide repertoire of possible metals, including Fe²⁺, Zn²⁺, Co²⁺ and Cu²⁺, amongst others.¹⁶ The enzymes from different organisms show variable oligomeric organization. As in *E. coli*, the NAGPases purified from the bacteria *Thermus caldophilus*,¹⁷ *V. cholerae*,¹⁸ and *Alteromonas* sp.¹⁹ are tetrameric proteins, in contrast with *B. subtilis* deacetylase that has been shown to be a dimer.¹⁵ NAGPases from *B. subtilis* and *T. maritima* are iron-proteins and there is indirect evidence that those from *Alteromonas* sp., *V. cholerae* and *T. caldophilus* are metalloproteins, although the involved metal was not identified. Here, we present the crystallographic structure of GlcNAc6P deacetylase from *E. coli* at 2.0 Å resolution as a tetrameric particle and in the metal-free state. However, both experiments, photon-induced X-ray emission (PIXE) and inductively coupled plasma emission spectroscopy (IC-PES), show that the native *Ec*NAGPase contains Zn atoms. Enzyme inactivation by metal-sequestering agents and subsequent reactivation by several divalent cations were used to explore the role of the metal ligands in catalysis by the enzyme.

Results and Discussion

The quality of the model

The crystal structure of *Ec*NAGPase was determined by the single isomorphous replacement with anomalous scattering (SIRAS) method using low-resolution (2.9 Å) anomalous scattering of iodine. The positions of 94.7% of the expected non-hydrogen protein atoms were identified and refined up to 2.0 Å resolution (Table 1A). Model building and refinement improved the *R* and *R*_{free} factors from 30.3% and 26.7% to 19.9% and 15.9%, respectively. The model presented very good stereochemical parameters (Table 1B). The r.m.s. deviation for bond distances was 0.022 Å and for bond angles was 1.89°. The average *B*-factor over all atoms was 21.2 Å². Most of the amino acid residues have their main-chain Φ/Ψ angles lying in the favored regions of the Ramachandran plot²⁰ (Table 1C). Only four amino acid residues (Asp70A, Asp70B, Val138A, and Tyr218B) were

Table 1. Diffraction data processing, refinement statistics and model validation

	I/NSLS	Cs/APS
A. Data processing		
Space group	$P2_12_12$	$P2_12_12$
a (Å)	81.95	82.30
b (Å)	113.74	114.25
c (Å)	80.76	80.43
Soaking	0.5 M NaI	0.5 M CsCl
Soaking time (min)	5	10
Wavelength (Å)	1.722	1.77120
Resolution range (Å)	20.00–2.45	14.00–1.85
No. total/unique refl.	171,278/23,748	578,400/54,996
Completeness (%)	83.7(47.5)	84.4(44.5)
Multiplicity	7.2(3.4)	10.5(3.4)
$\langle I/\sigma \rangle$	14.4(4.7)	24.9(8.0)
R_{sym}	0.094(0.190)	0.062(0.140)
B. Refinement statistics		
Resolution (Å)	20.00–2.00	
No. of total/expected		
Residues	716/764 (93.7%)	
Atoms (including H)	5914/6221	
Protein atoms	5441/5748	
Phosphate ions	2	
Glycerol molecules	3	
Water molecules	435	
R	0.1592 (48,144 refl.)	
R_{free} (5%)	0.1987 (2540 refl.)	
r.m.s. deviation from ideality		
Bond lengths (Å)	0.022	
Bond angles (deg.)	1.887	
Torsion angles (deg.)	7.751	
B -factor (Å ²)	Min.	Av. Max.
Protein	2.00	18.97 63.95
Main-chain	4.25	17.67 59.93
Side-chain	2.00	20.42 63.95
Phosphate ions	39.45	53.23 66.96
Glycerol molecules	69.87	49.59 82.69
Water molecules	22.04	46.48 76.78
C. Model validation		
Ramachandran ϕ/ψ distribution (residues \neq G or P)		
Most favored regions	547 (88.9%)	
Additional allowed regions	55 (8.9%)	
Generously allowed regions	9 (1.5%)	

found in the disallowed regions. The hydrogen bond interactions between Asp70 O^{δ2} and the Ala288 main-chain amide (2.91 Å) and between Tyr218 hydroxyl and Asn256 main-chain oxygen (2.67 Å) may explain the observed energetically unfavorable conformation of those residues. The Val138 residue was found in a high B -factor region of the structure where the electron density map was poorly interpreted. Double conformations were assigned to the side-chains of Met1, Arg8, Arg112, Arg115, Leu128, Asp152, Arg208, Tyr223, Asp240, Arg262, Glu322, Arg348 from chain A and Met1, Tyr223 and Asp329 from chain B. The peptide bond between residues Gly132 and Pro133 of both chains was present in the *cis* configuration.

The crystal asymmetric unit comprises a homodimer (chains A and B), two phosphate ions, three glycerol molecules and 435 water molecules. About 94% (716/764) of the amino acid residues were modeled in both chains. Residue 141 from chain A and residues 137–152 and 276–306 from chain B were excluded from the coordinates set due to poor or not interpretable electron density on the Fourier

map of coefficients $2m|F_o| - \Delta|F_c|$. The chain A missing residue 141 belongs to a disordered loop that was difficult to model, showing high atomic B -factors on refinement. This loop is even more disordered in the chain B, in which residues 137–152 could not be unambiguously positioned. The remaining B chain missing residues 276–306 are located in a flexible region affected by steric hindrance, due to lattice packing contacts. In general, monomer A is well packed in the crystal, establishing extensive contacts with six neighbor molecules. In contrast, monomer B is loosely packed with contacts restricted to one of its domains and two other symmetrically related molecules.

The tertiary structure

The deacetylase monomer folds into two domains, a pseudo (β/α)₈ barrel (α -domain) common to members of the metal-dependent amidohydrolases superfamily,¹⁶ and a small β -sandwich (β -domain) made up from secondary

structure elements contributed by the N and C termini. The topology and tertiary structure of the *Ec*NAGPase are illustrated in Figure 1, in which the same nomenclature for the *Bs*NAGPase topology¹⁵ was adopted. The polypeptide chain begins in the first β -sheet (magenta) of the β -domain, passes to the second one (cyan), turns to the first β -sheet again, builds up the complete α -domain, finally turning back to the β -domain, ending the last three β -strands.

In *Ec*NAGPase the β -domain appears to play a structural role in the quaternary organization of the protein by coupling subunits within the tetrameric particle. A DALI search²¹ with this domain showed that it most closely resembles the β -domains of the isoaspartyl dipeptidase,²² the cytosine deaminase,²³ the D-aminoacylase²⁴ and other metal-dependent aminohydrolases in which the β -domain plays a similar oligomerization role.

The α -domain contains 294 residues belonging to the central portion of the protein primary sequence (residues 55–350), and forms an oblate ellipsoid comprising ten α -helices and three β -sheets, one of them parallel (red) and the other two antiparallel (yellow and green). In the α -domain eight twisted β -strands (red) are surrounded by eight α -helices (blue) folded in a pseudo-TIM barrel enclosing

the funnel-like catalytic site of the enzyme. The interior of this barrel is mainly filled with bulky polar and hydrophobic side-chains. The α -domain barrel is somewhat distorted due to helix α 10, thus preventing the formation of hydrogen bonds between residues of the β 1 and β 8 strands to complete the internal wall of a classical TIM barrel. In space, the β -domain is laterally attached to the $(\beta/\alpha)_8$ barrel *via* two short linkers formed by Pro53-Gly54 and Gly350-Thr351. The relative position of both domains is stabilized by multiple hydrogen bonds and hydrophobic interactions between residues of the helices α 9 and α 10 (purple) and residues of the C terminus β -sheet strands (cyan; Figure 1).

The α -domain topology is well conserved in proteins of the urease-related superfamily.²⁵ The main differences between the topology of the *Bs*NAGPase and the *Ec*NAGPase are in the β -domain. The U-shaped β -barrel of the *Bs*NAGPase structure gives place to a β -sandwich in the *E. coli* enzyme, where the β A and β B strands of the *Tm*NAGPase topology are divided, respectively, into β A and β A1 and into β B1 and β B strands in *Ec*NAGPase. A second minor difference is the presence of a fourth β -strand (β J1) in the green antiparallel β -sheet of the α -domain.

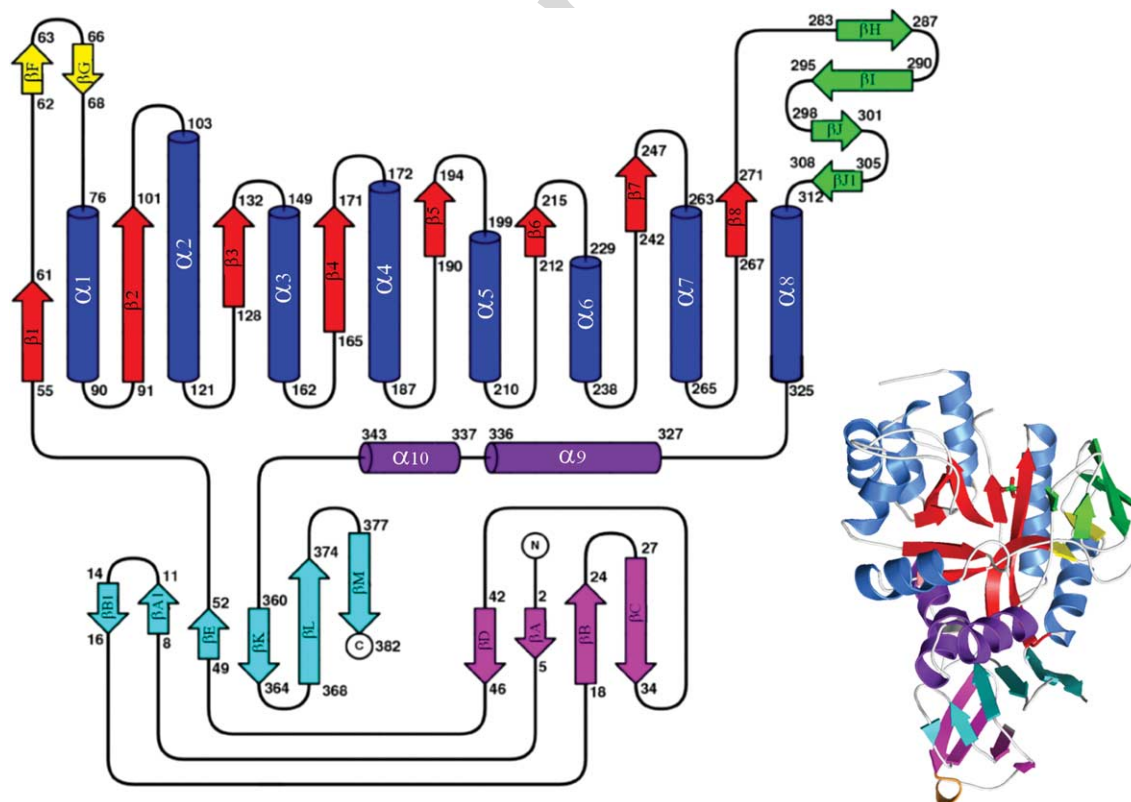


Figure 1. The topology of the GlcNAc6P deacetylase from *E. coli*. The secondary structure assignment is according to PROMOTIF.⁵⁰ The numbers are the amino acid positions of the primary sequence of the enzyme. This Figure was drawn using TOPDRAW.⁵⁵ On the right, the ribbons representation of the structure, adopting the same color scheme. The phosphate ion can be observed into the active site cleft. This Figure was drawn using PyMOL (<http://www.pymol.sourceforge.net>).

Comparison of *Ec*NAGPase and its orthologous *Bs* and *Tm* enzymes

The search in the DALI server²¹ to compare *Ec*NAGPase monomer with all known 3-D structures in the PDB, revealed that *Ec*NAGPase has considerable structural similarity with the catalytic domain of the proteins from the amidohydrolases superfamily. In particular, the overall fold of the enzyme is highly similar to the *Tm* and *Bs* NAGPases, as expected from the structural superposition of the three enzymes shown in Figure 2. Indeed, *Ec* and *Bs* NAGPases share 31% identity (Z-score 50.9) and the r.m.s. deviation of the 336 equivalent C α positions is 1.36 Å for the monomers of both structures, and 1.82 Å for the dimers (586 C α positions). In the *Ec* and *Tm* NAGPases superposition, which share 28% sequence identity (Z-score 44.3), the r.m.s. deviation of the 327 equivalent C α positions is 1.53 Å for the monomers and 1.64 Å for the dimers (596 C α positions). The sequence alignment derived from the structure superposition of the three enzymes is shown in Figure 3.

The main differences between the three known deacetylase structures are observed at the active site region, where loops $\beta 3$ - $\alpha 3$ (loop I) and βJ - $\beta J1$ (loop II) adopt a closed conformation in the *Bs* and *Tm* NAGPases and an open conformation in the *E. coli*

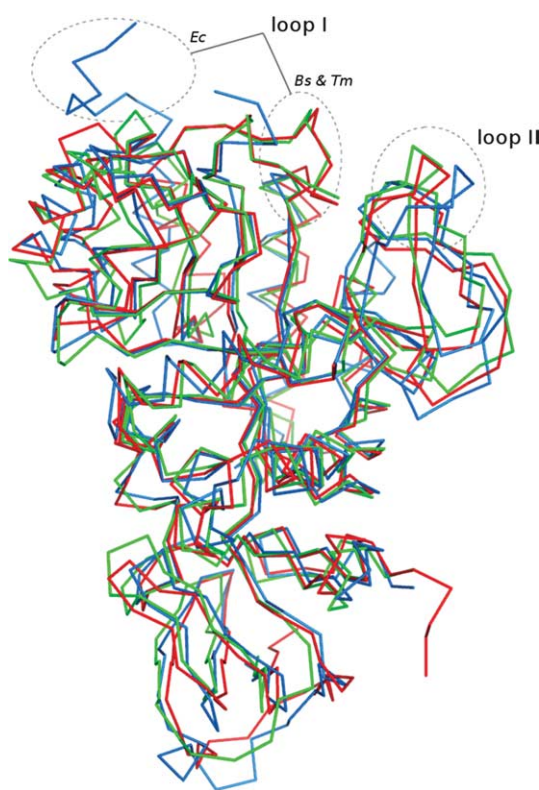


Figure 2. Superposition of the *Ec*NAGpase (blue), *Bs*NAGpase (red) and *Tm*NAGpase (green). Loops I and II, which are involved in metal binding and adopt a closed conformation in the *Bs* and *Tm* structures and opened and disordered in the *Ec* structure, are highlighted. This Figure was drawn using PyMOL.

apoenzyme (Figures 2 and 4). The loop I in the *Ec*NAGPase appears in a different position compared to the other two deacetylases. In the *Tm* and *Bs* enzymes, loop I is well ordered and points toward the active site cleft, while in *Ec*NAGPase it is disordered with missing residues in the electron density maps. Extensive model building and refinement cycles of the *E. coli* structure allowed us to build some residues of loop I, with its N-terminal side (residues 133–140) laying on the pocket entry, whereas the C-terminal side (residues 142–148) is wide open and exposed to the solvent. These two stretches of loop I are only partially occupied, as substantiated by high atomic temperature factors of their residues. It is worth noting that helix $\alpha 3$ in the *Ec*NAGPase structure has one additional turn when compared with the *Bs* or *Tm* NAGPases and it is farther away from the catalytic site entry (Figure 4). These conformational particularities of *Ec*NAGPase must be related to the absence of metal ligands in the active site. The *Tm*NAGPase structure has one Fe atom in the active site, while the *Bs*NAGPase presents two Fe atoms in a binuclear arrangement and also contains a GlcNAc6P molecule, which is one of the enzyme reaction products. In contrast, the *Ec*NAGPase was crystallized in high-concentration phosphate buffer at low pH, presenting a clearly identified phosphate ion in the active site close to two water molecules (Figure 5). In the *Tm* and *Bs* NAGPases the loops I and II are held closed by several direct or bridged interactions involving the metal ligands and, in the case of the *B. subtilis* enzyme, through a bridging water molecule, serving as an anchor for loop I. Equivalent residues are present in the *Ec*NAGPase and therefore it is not unreasonable to hypothesize that the *E. coli* apoenzyme could undergo an extensive conformational change upon metal binding. In this case, as the helix $\alpha 3$ of the *Bs* and *Tm* NAGPases structures is smaller and inclined when compared to that of *E. coli*, and the loop I has three residues less than their homologues, one might suggest that the segment corresponding to the top three residues of the *E. coli* $\alpha 3$ helix should unfold upon metal binding to allow the β -hairpin of loop I to approach the substrate on the active site. Further structural studies would be needed to support this hypothesis.

The quaternary structure

The two polypeptide chains in the asymmetric unit are related by a non-crystallographic 2-fold symmetry axis parallel with the *ab* face of the crystal unit cell, forming a 16° angle with the crystallographic axis *b*. The superposition of chains A and B reveals their structural equivalence with the r.m.s. distance between the main-chain atoms of both chains being 0.53 Å (max. 2.52 Å) for 335 equivalent residues.

The monomer–monomer interface is composed by interacting residues from the loop $\beta 6$ - $\alpha 6$, from the helix $\alpha 6$ and the helix $\alpha 7$ of the adjacent

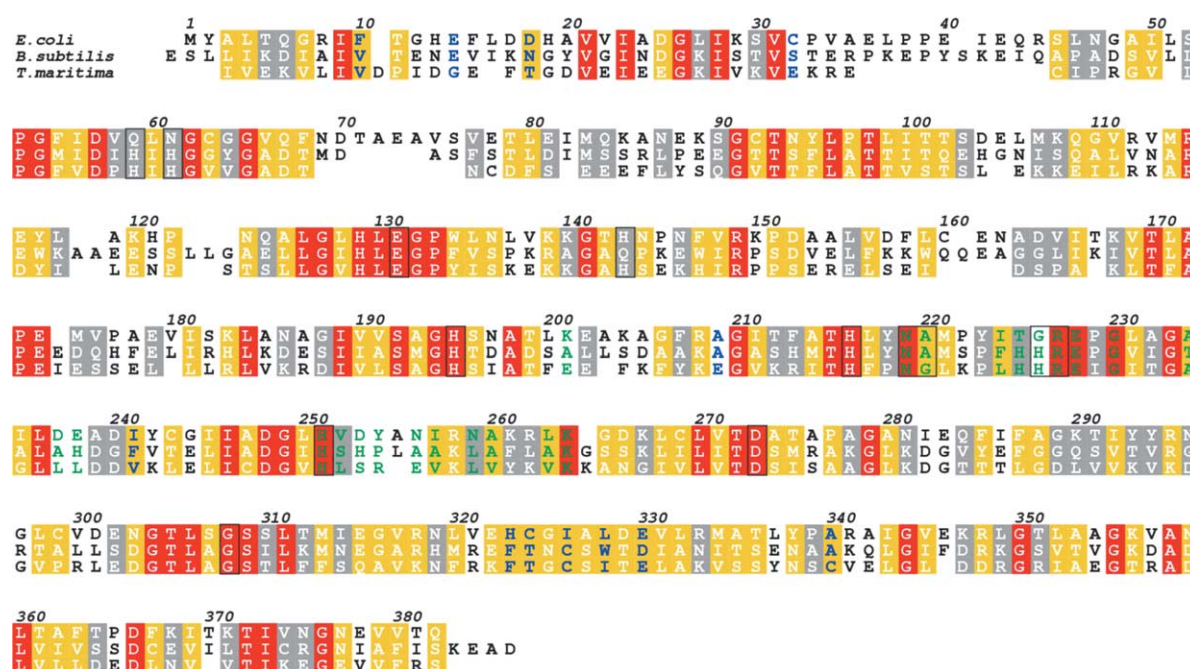


Figure 3. Primary sequence alignment based on the structure superposition of *Ec*, *Bs* and *Tm* NAGPases (see STAMP⁴⁹). Red shades are identical residues, gold shades are considerably conserved residues (more than 60% of similarity), and gray shades are poorly conserved residues (up to 60% of similarity). The green and the blue letters are, respectively, the monomer-monomer and the dimer-dimer interface interacting residues, and the boxed letters are the residues of the active site.

subunit. The loop $\beta 6-\alpha 6$ projects an arm over the helix $\alpha 7$, clasping the contiguous subunit and going back to the helix $\alpha 6$ of the first chain. Under this arm-like loop, the helices $\alpha 6$ from the first subunit and $\alpha 7$ from the second are tightly packed, providing an additional stability to the dimer. The monomer-monomer interface is composed of 54.1% of non-polar residues. Furthermore, there are two salt-bridges between Asp237 and Arg258, 16 hydrogen bonds, six bridging water molecules and many hydrophobic interactions. The shape of the interface is relatively flat with an r.m.s. of the best fit plane of 3.68 Å. The excluded solvent-accessible area of the dimer interface is 1053 Å², corresponding to 6.8% of the total monomer area. The interacting residues of the monomer-monomer interface (water excluded) are shown in green letters in the primary sequence alignment of the three bacterial enzymes, showing good conservation (Figure 3).

The crystallographic symmetry applied to the asymmetric unit, produces a homotetramer (Figure 6), which is expected to be identical with that found in solution. The analysis of the other possible crystallographic tetramer configurations revealed a smaller number of contacts. In the dimer-dimer interface of the most probable crystallographic tetramer, four intersubunit salt-bridges were found between Arg8 and Glu322, and between Asp240 and Arg341. Additionally, there are 16 hydrogen bonds with no bridging water molecules and several hydrophobic interactions mainly from the inner interface residues (Gly325, Ala327 and Asp329). The dimers are twisted

together across an elongated interface. The excluded solvent-accessible area of the dimer-dimer interface is 5343 Å², more than the commonly found value for interfaces²⁶ and corresponding to 18.4% of the total dimer area. The shape correlation statistics (*S_c*) or complementarity²⁷ obtained for the crystallographic tetramer interface is 0.76, which is in the upper limit for oligomeric proteins.

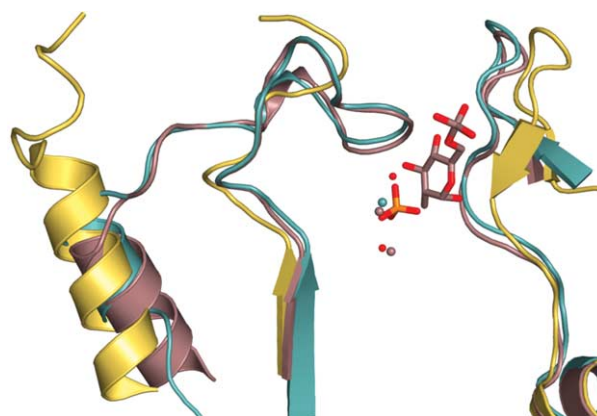


Figure 4. The relative positions of the loops I and II of the *Ec* NAGPase (gold), including the bound phosphate ion and water molecules (red spheres) and its equivalent counterparts from *Bs* (brown, with two Fe ions) and *Tm* (cyan, with one Fe ion). The O2 of the phosphate ion in the *Ec* NAGPase and the Fe ions in the *Bs* and *Tm* structures occupy almost the same position in the active site. The second Fe ion from the *B. subtilis* structure occupies the same position as the HOH257 molecule in the *E. coli* structure. This Figure was drawn using PyMOL.

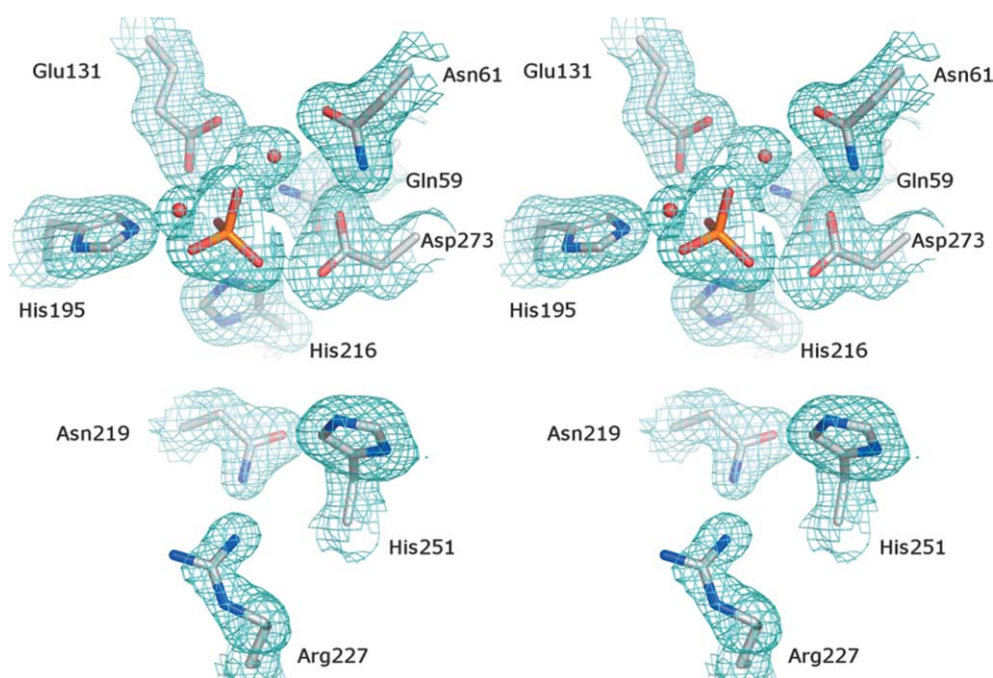


Figure 5. Stereoscopic projection of the Fourier map with coefficients $2m|F_o| - \Delta|F_c|$, contoured at 1.2σ . The electron density at the active site of the *EcNAGPase* was modeled and refined as a phosphate ion and two water molecules, HOH257 (back) and HOH259 (front). This Figure was drawn using PyMOL.

The prominent hydrophobic patch of the dimer-dimer interface was observed with some charged residues in both complementary surfaces, comprising about 45.4% of non-polar amino acid residues. The interacting residues of the dimer-dimer interface are shown in blue letters in the primary sequence alignment (Figure 3). In spite of the fact that the residues involved in salt-bridging are not conserved in the other two deacetylase structures, more than half of the interface residues are conserved and some residues are identical, predominantly in a well-conserved region of the sequence. The tetrameric crystal structure of *EcNAGPase* was confirmed by SAXS in the solution phase. The molecular mass derived from the experimental data was 160(5) kDa, which is in excellent agreement with the value obtained from its amino acid sequence. Comparison of the theoretical and experimental scattering profiles of the crystallographic tetramer resulted in a very good χ fitting of 1.47 (Figure 7(a)). These results suggest that the protein quaternary structure in solution is essentially identical with that observed in the crystal. The distribution function of the distances does not exhibit negative values (Figure 7(b)), so the protein can be considered as present in dilute solution, thus validating the subsequent molecular shape analysis. The values of D_{\max} and R_g for the tetramer are reported in Table 2. All of the *ab initio* models obtained from SAXS data yielded similar physical parameters. A representative envelope model is presented in the Figure 7(c) superposed to the *EcNAGPase* X-ray crystal structure, showing an excellent agreement between both models.

The active site

The active site of the enzyme is located on the bottom of the α -domain cavity. The sequence alignment based on structural superposition of the three known deacetylase structures, shows that all the amino acid residues involved in metal and substrate binding are essentially conserved (Figure 3). Most of the active site residues appear in similar positions in the three structures (Figure 8).

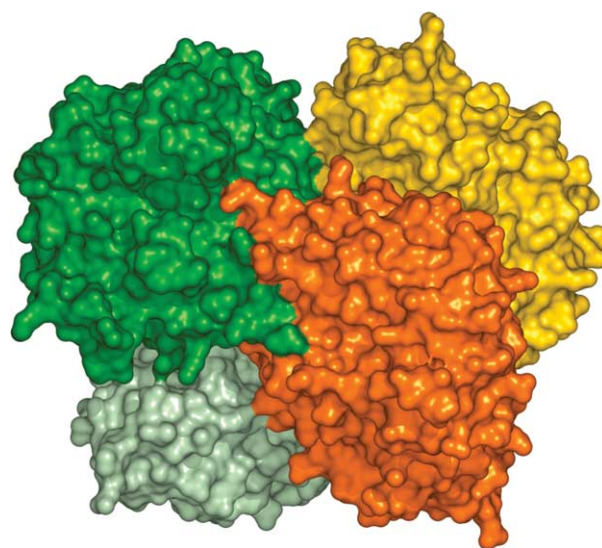


Figure 6. The surface representation of the *EcNAGPase* tetrameric structure, each monomer represented with a different color. This Figure was drawn using PyMOL.

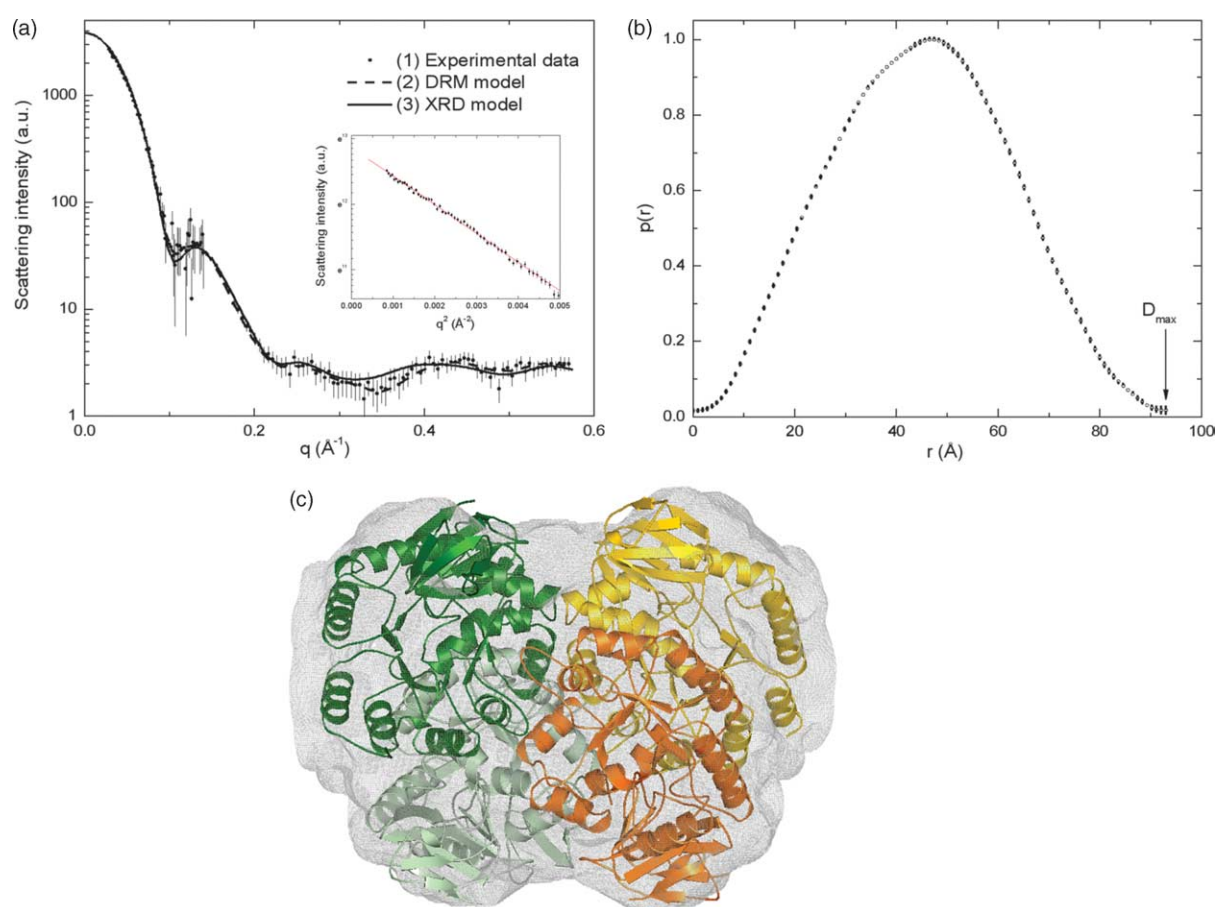


Figure 7. Experimental SAXS curves of the *Ec*NAGPase tetramer in solution and results of the fitting procedures. (a) $\log I$ versus q experimental curve of *Ec*NAGPase in solution; (1) experimental curve; (2) scattering from the dummy residues model (DRM) obtained by GASBOR; and (3) scattering intensity obtained by the crystallographic model (XRD). On the right, details of the correspondent Guinier plots ($\log I$ versus q^2). (b) Distance distribution functions of the *Ec*NAGPase; arrows point out the maximum diameter D_{\max} . (c) Low-resolution SAXS envelope of the *Ec*NAGPase tetramer superposed to the X-ray structure.

A noticeable difference is the replacement of His63 and His65 in *Bs*NAGPase (and the equivalent residues His49 and His51 in *Tm*NAGPase) by Gln59 and Asn61, respectively, in *Ec*NAGPase.

Table 2. Structural parameters corresponding to the *E. coli* deacetylase protein in solution studied by SAXS, and mathematical GNOM and GASBOR output parameters

Parameter	<i>Ec</i> NAGPase tetramer	
	Exp. ^a	DRM ^b
Mass (kDa)	160(1)	
D_{\max} (Å)	93(2)	93(2)
R_g (Å)	34.1(1)	34.2(1)
V (nm ³)		4390
Free parameters	17.2 ^c	1528
Discrepancy χ	1.47	0.76(3)
Resolution (Å) ^d	10.8	10.8

The protein molecular mass was obtained using an indirect method from $I(0)$ intensity. Errors are under 5%.

^a Exp., calculated from the experimental data.

^b DRM, parameters of the dummy residue models.

^c Number of Shannon channels is given by $N_s = D_{\max}(q_{\max} - q_{\min})/\pi$.

^d Given by $R = 2\pi/q_{\max}$.

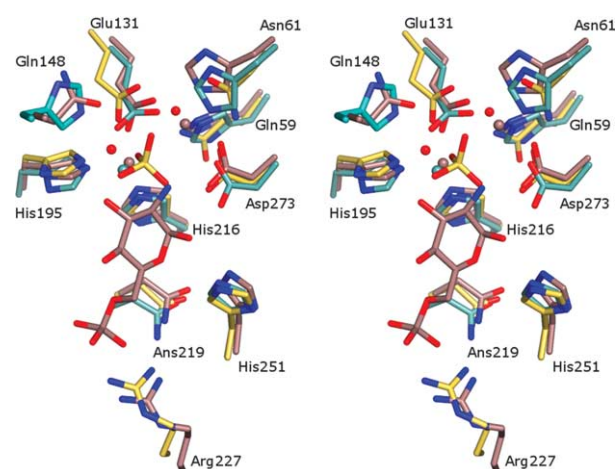


Figure 8. Stereoscopic projection of the superposition of active site residues of the three known NAGPase structures *Ec* (gold), *Bs* (brown) and *Tm* (cyan). Most of the active site residues are conserved and occupy similar positions in the three structures. This Figure was drawn using PyMOL.

The active site residues span over three distinct regions. The first one comprises Gln59 and Asn61 from the α - β -domains linker, Glu131 from the β 2 strand, and Asp273 from the loop β 8- β H. This first region is occupied by the water molecule HOH259 in *Ec*NAGPase, by a Fe ion in *Bs*NAGPase and is empty in *Tm*NAGPase. The second region of the active site comprises the residue Glu131 from the β 2 strand, His195 from the loop β 5- α 5, and His216 from the loop β 6- α 6. In *Ec*NAGPase this second site is occupied by a phosphate ion, while the equivalent positions in *Bs*NAGPase and *Tm*NAGPase contain a Fe ion. The third region of the active site is formed by residues from the β -hairpin of the loop β 3- α 3, by residues from the loop β J1- α 8 and by some residues of the loop β 6- α 6 from the second subunit. Residues in these segments provide the interactions that hold the sugar moiety of the substrate, mainly through hydrogen bonds. No metal ion was found at the active site in the *Ec*NAGPase structure and no peak appears in the anomalous difference Fourier map of the active site cleft. The remaining electron density in the catalytic site region was satisfactorily modeled as a single phosphate ion, which was the dominant component of the crystallization solution (1.85 M). However, in the 1YMY structure there is a non-modeled strong electron density peak (5σ), appearing in the same position as the phosphate ion in the present structure. The pH of the crystallization solution in this study was about 4.0, while 1YMY was crystallized at pH 8.5. It is then possible that the active site metal in *Ec*NAGPase has been displaced by protonation of His195 and His216 imidazole groups, allowing the phosphate ion to bind, even with low affinity, as a result of its high concentration. Additionally to the phosphate ion, the metal-binding site of *Ec*NAGPase shows two water molecules that were absent from the other two NAGPase structures. These water molecules build hydrogen bonds with the phosphate ion and other active site residues, providing stability to the whole cavity of the apoenzyme. The hydrogen bond distances between the two water molecules (HOH257 and HOH259), the phosphate ion and some active site residues, are depicted in Figure 9.

*Ec*NAGPase is a zinc metalloprotein

To ascertain if the absence of metal ions in the present crystal structure could be a consequence of the crystallization conditions, PIXE and ICP-AES spectra were obtained using enzyme samples that had no contact with chelating agents or added metal ions during their purification procedure. The PIXE spectrum shows that the Zn emission lines $K\alpha_1$ (8.64 keV) and $K\beta_1$ (9.57 keV) are the dominant features; a much smaller line from Cu atoms (Cu $K\alpha_1$ line, 8.05 meV) is also observed (Figure 10). This result unambiguously shows that Zn is the physiological metal ion of *Ec*NAGPase. The presence of zinc in *Ec*NAGPase was also confirmed by ICP-AES. Taking the 16 sulfur atoms of *Ec*NAGPase as an internal standard, a stoichiometry of $1.4(\pm 0.1)$

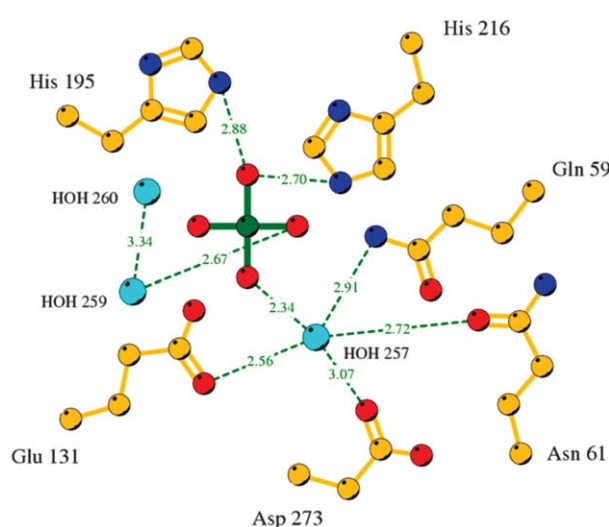


Figure 9. The interacting residues of the *Ec*NAGPase active site with the phosphate ion and water molecules. The bond distances in green are in Å.

Zn atoms per subunit was calculated. This non-integer value suggests that the enzyme can accommodate more than one atom and that the second Zn-binding site (equivalent to metal site I in *Bs*NAGPase) could have a high dissociation constant and as a result, it is partially lost during the enzyme purification in the absence of added Zn. Binuclear Zn centers are characteristic of metallo- β -lactamases, and crystallographic structures with one or two Zn atoms bound were reported.²⁸ Non-integer stoichiometries are frequent; for instance, the binuclear Zn center of *B. cereus* β -lactamase binds the metal with negative cooperativity, making it difficult to reach the theoretical stoichiometry of two Zn atoms per site.²⁹ The role of Zn^{2+} and other

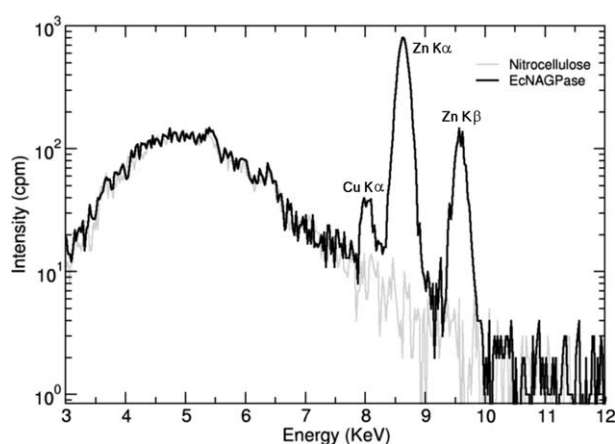


Figure 10. The PIXE spectrum shows that the Zn emission lines $K\alpha_1$ (8.64 keV) and $K\beta_1$ (9.57 keV) are predominant in the *Ec*NAGPase sample. The gray line on the back is the spectrum of the nitrocellulose membrane. PIXE and ICP-AES results suggest that native *Ec*NAGPase is a zinc metalloprotein.

cations for *Ec*NAGPase function was also explored through inactivation experiments using metal-chelating agents. Enzyme samples were exhaustively dialyzed against EDTA or 1,10-phenanthroline and their activity evaluated after desalting, by determination of the apparent k_{cat} values. The effect of EDTA was milder and slower than 1,10-phenanthroline, requiring higher concentrations and longer dialysis times to reduce activity. Dialysis against 40 mM EDTA for 24 h produced 72% inactivation and the subsequent treatment of the enzyme with an excess of zinc acetate (500 times the enzyme concentration) recovered its activity up to 82% of the control value (Table 3). It was not possible to obtain the inactive apoenzyme in solution, even after exhaustive dialysis. When an excess of Zn^{2+} was added to the untreated native enzyme its apparent k_{cat} value did not change. These results support the assertion that *Ec*NAGPase is a metallo-enzyme containing Zn, and if it binds a second atom, this is not essential for catalysis. The enzyme samples partially inactivated by extensive dialysis against EDTA (Table 3, sample B) were unstable and become inactive after three days at 4 °C, while the non-dialyzed enzyme can be stored for several days under these conditions without activity loss.

Other metal ions can replace Zn in the reactivation experiments of the enzyme partially inactivated by chelating agents. The conditions that produce partial inactivation of deacetylase by EDTA and the restoration of its activity by the addition of Fe^{2+} , Mn^{2+} and Co^{2+} , are summarized in Table 3. Note that Co^{2+} strongly activates the *Ec*NAGPase, which had been partially deprived of its metal (Table 3, sample B), yielding values twice those for the control enzyme (Table 3, sample A). In other Zn-aminoacylases, their metal can be replaced by Cd or Co,³⁰ and in some cases higher activities are obtained when Co replaces Zn. This effect is particularly conspicuous in *N*-acetylornithine deacetylase (ArgE) from *E. coli*, in which cobalt can produce an activity increase of one order of magnitude.³¹

Implications on catalytic mechanism of *Ec*NAGPase

Even though we report the structure of the apoenzyme of *Ec*NAGPase, in which a phosphate ion occupies the position equivalent to metal site II in *Bs*NAGPase, the studies performed on the native enzyme indicate that native *Ec*NAGPase is a zinc-containing metalloenzyme. The structure of the active site is strongly conserved in all known NAGPases; although native *Ec*NAGPase contains Zn, the enzyme is also able to bind Fe^{2+} . As shown in Table 3, the apoprotein regains the same activity after re-incorporation of either Zn^{2+} or Fe^{2+} . The stoichiometry of more than one Zn^{2+} atom per site of the native enzyme suggests that the *E. coli* enzyme may be able to bind up to two equivalents of Zn per subunit. The presence of two zinc sites in close proximity is characteristic of all metallo- β -lactamases, although the specific role of the two metal-binding sites and their interaction in catalysis is still unclear.³² The reaction mechanisms of many well-studied zinc-dependent hydrolases share several common features. A Zn-coordinated water molecule or hydroxide ion is the nucleophile in many of these hydrolysis reactions. A single Zn atom can polarize a bound water molecule for nucleophilic attack on the carbonyl carbon of the substrate molecule; this mechanism requires the proximity of a general base, a role that could be played in *Ec*NAGPase by Glu131, which could abstract a proton from water. A binuclear Zn center could hold an OH^- anion that bridges the two zinc atoms.²⁸ The variation of kinetic parameters with pH reported in the literature¹⁴ did not reveal any pK_a value on the alkaline side of the explored range (up to pH 9.5), thus favoring the suggestion that the nucleophile can be a high pK_a species. Further crystallographic and physico-chemical research will be needed to characterize the Zn binding in *Ec*NAGPase and its role in catalysis and stability of the native enzyme.

Materials and Methods

Protein expression and purification

*Ec*NAGPase was expressed and purified as described in the literature,¹⁴ except that EDTA was absent from the solution used to suspend the cells for sonic disruption. The hydrophobic interaction step was skipped and the supernatant of the ammonium sulfate fractionation in the range 35–60% saturation, was dialyzed against 25 mM Tris-HCl buffer (pH 8.0) and submitted to two successive anion-exchange chromatography steps. The dialyzed sample was loaded onto a 2.5 mm \times 100 mm Mono-Q column (Amersham-Pharmacia) equilibrated with the same Tris buffer, and eluted with a ionic strength gradient of KCl in the range 0.1–0.3 M. The fraction containing the enzyme was eluted from a small Protein-Pack Q column using a similar protocol. The pure enzyme exhibited a k_{cat} value of 118 s⁻¹, which is a similar value to that reported (105 s⁻¹).¹⁴

Table 3. *E. coli* *N*-acetylglucosamine-6-phosphate deacetylase activity assays

Sample	Description	% Activity ^a
A	Dialyzed against buffer A ^b	100
B	Dialyzed against buffer A + 25 mM EDTA ^c	27.5
C	Sample A, + 25 μM ZnSO_4	104
D	Sample B, + 25 μM ZnSO_4	84
E	Sample B, + 25 μM cobalt-II acetate	226.0
F	Sample B, + 25 μM MnSO_4	92
G	Sample B, + 25 μM ferrous citrate complex	66

^a V_{max} values were obtained from the kinetic analysis of each sample and the corresponding relative activities are expressed with respect to the V_{max} of sample A.

^b Buffer A is 50 mM Tris-HCl (pH 7.5) at 4 °C.

Activity assay

Two chelating agents were used in an attempt to obtain the metal-free enzyme, sodium EDTA and 1,10-phenanthroline. In all cases the *Ec*NAGPase (3–5 μ M) was dialyzed against a 2000-fold volume excess of 40 mM EDTA or 5 mM 1,10-phenanthroline, in 50 mM Tris–HCl (pH 7.5) (buffer A) at 4 °C. The chelating agent was removed by extensive dialysis against three changes of 400-fold volumes of buffer A. Reconstitution experiments were performed by adding the metal salt to the treated *Ec*NAGPase and pre-incubating it at room temperature for 1 h, before assay. *Ec*NAGPase activity was calculated from the decrease of the absorbance of the carboxamido group of GlcNAc6P at 215 nm, at 30 °C in buffer A, using a double-beam spectrophotometer.¹⁴ GlcNAc6P UV spectra under these conditions show strong hypochromic and bathochromic effects on increasing amino sugar concentration. However, at 215 nm there is an isosbestic point that, in spite of its low molar absorptivity (370 mol^{−1} l cm^{−1}), can be successfully used for the continuous and direct assay of *Ec*NAGPase, provided that a good spectrophotometer is used. The curves of initial velocities *versus* substrate concentration (range 0–5 mM) shows strong substrate inhibition. Data were fitted to the following equation: $v_o = (V[S])/(K_m + [S] + S^2/K_i)$, which introduces K_i as a substrate inhibition constant.¹⁴ The program GraphPad Prism 4.0b was used for the non-linear regression analyses.

Crystallization, data collection and processing

Deacetylase crystals belonging to the primitive orthorhombic $P2_12_12$ space group were obtained by the hanging-drop, vapor-diffusion technique at 18 °C, using 1.85 M NaH₂PO₄ as precipitant.³³ After soaking, the crystals used for diffraction data collection were flash-frozen at cryogenic temperature (100 K). They were initially transferred for a few seconds to a conditioning solution consisting of the well solution with 20% (v/v) glycerol and directly frozen in a liquid nitrogen gaseous stream. The first monochromatic data set (Cs/APS, λ = 1.1808 Å) was collected at the Advanced Photon Source (APS), Argonne, USA, at the beamline 17-ID on an ADSC Quantumq210 CCD detector. The crystal was transferred for 5 min to the well solution with 0.5 M CsCl. The second monochromatic data set (I/NSLS, λ = 1.7220 Å) was collected at the National Synchrotron Light Source (NSLS), Brookhaven, USA, at the beamline X6A on an ADSC Quantumq210 CCD detector. The crystal was transferred for 10 min to the well solution with 0.5 M NaI.³⁴ The observed intensities were indexed and integrated using MOSFLM³⁵ and scaled using SCALA³⁶ from the CCP4 suite.³⁷ Data processing statistics of both crystals are summarized in Table 1A.

Structure determination, refinement and analysis

Deacetylase crystal structure was solved by using the SIRAS technique. The structure factors of the isomorphous and anomalous differences of both data sets were calculated using SHELXC.³⁸ No useful anomalous information was provided by the Cs/APS data set, thus it was used as reference for the I/NSLS. Both data sets were shown to be more isomorphous to each other than when compared to the native data set (Nat/NSLS) collected at the Brazilian National Synchrotron Laboratory, the data set³³

that the structure was refined against. For the purpose of heavy-atom location with SHELXD,³⁸ the I/NSLS data set was truncated at 2.9 Å resolution. Twenty-three iodine sites of partial occupation (1.0–0.3) were found and the output correlation coefficient between the observed and calculated structure factor differences were 37.16% (all data) and 27.71% (weak data). The first structure factors phase estimation and the heavy atoms refinement were performed with SHELXE.³⁹ Further structure factors phase calculations were carried out with SOLVE.⁴⁰ The phase extension and the succeeding density modification steps (solvent flattening, NCS averaging and pattern matching) were performed with RESOLVE.⁴¹ Phasing statistics are summarized in Table 1B. As the resolution of the native data set extended to 2.0 Å and it was 97.7% complete, automatic model building with refinement was attempted with ARP/WARP.⁴² An initial hybrid model was built by merging amino acid residues traced in different runs and sub-cycles of model building. The remaining expected amino acid residues, most of them from loop regions, were manually built by using XFIT from the XTALVIEW suite.⁴³ The structure refinement was carried out with REFMAC5⁴⁴ using bulk solvent correction and TLS parameters refinement⁴⁵ for each domain of both chains. Non-crystallographic symmetry restraints, used in the early stages of refinement, were gradually relaxed. Water molecules were gradually added to the model during the later stages of refinement and subsequently analyzed for rejection. The agreement between the deposited model coordinates and their diffraction data was assessed by using SFCHECK⁴⁶ and the model geometrical and stereo chemical quality was assessed by using PROCHECK⁴⁷ and WHAT-CHECK.⁴⁸ The primary sequence alignment based on structure superposition was performed with the program STAMP.⁴⁹ The secondary structure assignments for the *Ec*NAGPase were made with the program PROMOTIF.⁵⁰ To identify the interacting residues between the subunits in the tetramer, hydrogen bonds were calculated using the program HBPLUS.⁵¹ The buried area of the solvent-accessible surface of both the monomer and the dimer were calculated with the program SURFACE from the CCP4 suite.³⁷ The shape correlation statistic S_c ²⁷ for the crystallographic tetramer was calculated by using the program SC, also from the CCP4 suite.³⁷

SAXS data collection and processing

Small-angle X-ray scattering experiments with synchrotron radiation were carried out on the storage ring DORIS III of the Deutsches Elektronen Synchrotron (DESY). Data were collected using standard procedures on the X33 camera of the European Molecular Biology Laboratory (EMBL) and multiwire proportional chambers with delay line readout. Samples at concentrations between 2 mg/ml and 15 mg/ml were measured at 0.15 nm wavelength for sample–detector distances of 2.9 m and 1.4 m, covering the momentum transfer ranges $0.020 < q < 0.20$ Å^{−1} and $0.035 < q < 0.56$ Å^{−1}, respectively ($q = 4\pi \sin(\theta)/\lambda$, where 2θ is the scattering angle). The data were normalized to the intensity of the incident beam and corrected for the detector response, the scattering of the buffer was subtracted, and the difference curves were scaled for protein concentration. All procedures involved statistical error propagation using the SAPOKO program (unpublished). The curves recorded at a sample–detector distance of 2.9 m were extrapolated to zero concentration and merged with the data obtained at 1.4 m. The corrected and normalized SAXS curves of the *Ec*NAGPase tetramer are shown in Figure 7.

The molecular mass of the tetramer in solution was verified by using an indirect method for solution scattering profile data. The theoretical scattering profile of the EcNAGPase tetramer was also calculated and compared to the experimental one. In order to determine the distance distribution function and the related physical parameters of the tetramer, such as the radius of gyration R_g and the maximum diameter D_{max} , data were analyzed with the program GNOM.⁵² The distance distribution function $p(r)$ obtained is presented in Figure 7. Finally, the dummy residues models (DRM) were built from the experimental data at low resolution by using the program GASBOR,⁵³ assuming a tetramer 4-fold symmetry axis to reduce the number of free parameters of the models. The tetramer starting volume corresponded to a sphere of 93 Å diameter, containing 1528 dummy residues with an excluded volume of 27.73 Å³ per radius. Ten independent *ab initio* GASBOR simulations were performed in order to verify the uniqueness of the resulting shapes, being the models restored from different starting conditions. The dummy residues models were superimposed to each other and to the crystallographic structure with the SUPCOMB program.⁵⁴ The SAXS envelope (Figure 7) was built from a dummy residues model with the programs NCSMASK and MAPMASK from the CCP4 suite.³⁶

Metal analyses: PIXE and ICP-AES

PIXE experiments were carried out in a tandem Van de Graaff electrostatic accelerator, using a 3 MeV proton beam of 1 mm diameter. The enzyme sample was dried on a nitrocellulose membrane. X-ray emission was detected in two solid-state detectors, a Si-pin Amptek of 180 eV-resolution, operated under helium stream; and for heavier elements a LE Ge detector was used. Spectra were collected for 8 min, at least in three places in each sample. The metal content of the purified and untreated enzyme was determined using ICP-AES. The analyzed samples (9 µmol l⁻¹) were digested with HNO₃, diluted to a final volume of 3 ml using 18 megohm-cm purity water and analyzed for S, Zn, Fe and Co in an IC-PES Leeman PS3000 spectrometer. Sulfur was taken as an internal standard, taking into account that the enzyme subunit contains 16 S atoms per polypeptide chain, from its six methionine and six cysteine residues.

Protein Data Bank accession code

The structure coordinate set has been deposited in the RCSB Protein Data Bank with accession code 1YRR.

Acknowledgements

F.M.F.G.O. thank FAPESP for financial support (PhD grant 99/12029-4 and CBME grant 98/14138-2), Dr Vivian Stojanoff for the beam time access at the NSLS/X6A, Dr Andrew Howard for the beam time access at the APS/17BM during the American Crystallographic Association Protein Crystallography Summer Course 2003. We also thank Marc Malfois and Dmitri Svergun for assistance and beam time for SAXS data collection at DESY. G.M.H., M.C. and M.L.C. acknowledge support

from CONACYT (grant 41328-Q) and funds received from PAPIIT-UNAM. We also thank Dr Margarita Gutiérrez-Ruiz (UNAM) for ICP-AES analyses and Dr José L. Ruvalcaba-Sil for PIXE analyses.

References

- White, R. J. (1968). Control of amino sugar metabolism in *Escherichia coli* and isolation of mutants unable to degrade amino sugars. *Biochem. J.* **106**, 847–858.
- Plumbridge, J. A. (1991). Repression and induction of the *nag* regulon of *Escherichia coli* K12: the roles of *nagC* and *nagA* in maintenance of the uninduced state. *Mol. Microbiol.* **5**, 2053–2062.
- Midelfort, C. F. & Rose, I. A. (1977). Studies on the mechanism of *Escherichia coli* glucosamine-6-phosphate isomerase. *Biochemistry*, **19**, 1590–1596.
- Singh, P., Ghosh, S. & Datta, A. (2001). Attenuation of virulence and changes in morphology in *Candida albicans* by disruption of the N-acetylglucosamine catabolic pathway. *Infect. Immun.* **69**, 7898–7903.
- Calcagno, M., Campos, P. J., Mulliert, G. & Suástegui, J. (1984). Purification, molecular and kinetic properties of glucosamine-6-phosphate isomerase (deaminase) from *E. coli*. *Biochim. Biophys. Acta*, **787**, 165–173.
- Montero-Morán, G. M., Horjales, E., Calcagno, M. L. & Altamirano, M. M. (1998). Tyr254 hydroxyl group acts as a two-way switch mechanism in the coupling of heterotropic and homotropic effects in *Escherichia coli* glucosamine-6-phosphate deaminase. *Biochemistry*, **37**, 7844–7849.
- Alvarez-Añorve, L. I., Calcagno, M. L. & Plumbridge, J. (2005). Why does *Escherichia coli* grow more slowly on glucosamine than on N-acetylglucosamine? Effects of enzyme levels and allosteric activation of GlcN6P deaminase on growth rates. *J. Bacteriol.* **187**, 2974–2982.
- Park, J. T. (2001). Identification of a dedicated recycling pathway for anhydro-N-acetylmuramic acid and N-acetylglucosamine derived from *Escherichia coli* cell wall murein. *J. Bacteriol.* **183**, 3842–3847.
- Uehara, T., Suefuji, K., Valbuena, N., Meeham, B., Donegan, M. & Park, J. T. (2005). Recycling of the anhydro-N-acetylmuramic acid derived from cell wall murein involves a two-step conversion to N-acetylglucosamine-phosphate. *J. Bacteriol.* **187**, 3643–3649.
- Uehara, T. & Park, J. T. (2004). The N-acetyl-D-glucosamine kinase of *Escherichia coli* and its role in murein recycling. *J. Bacteriol.* **186**, 7273–7279.
- Mengin-Lecreulx, D. & van Heijenoort, J. (1994). Copurification of glucosamine-1-phosphate acetyltransferase and N-acetylglucosamine-1-phosphate uridylyltransferase activities of *Escherichia coli*: characterization of the *glmU* gene product as a bifunctional enzyme catalyzing two subsequent steps in the pathway for UDP-N-acetylglucosamine synthesis. *J. Bacteriol.* **176**, 5788–5795.
- Mengin-Lecreulx, D. & van Heijenoort, J. (1996). Characterization of the essential gene *glmM* encoding phosphoglucosamine mutase in *Escherichia coli*. *J. Biol. Chem.* **271**, 32–39.
- Mio, T., Yabe, T., Arisawa, M. & Yamada-Okade, H. (1998). The eukaryotic UDP-N-acetylglucosamine

- pyrophosphorylases. Gene cloning, protein expression, and catalytic mechanism. *J. Biol. Chem.* **273**, 14392–14397.
14. Souza, J. M., Plumbridge, J. A. & Calcagno, M. L. (1997). *N*-Acetyl-D-glucosamine-6-phosphate deacetylase from *Escherichia coli*: purification and molecular and kinetic characterization. *Arch. Biochem. Biophys.* **340**, 338–346.
 15. Vincent, F., Yates, D., Garman, E., Davies, G. J. & Brannigan, J. A. (2004). The three-dimensional structure of the *N*-acetylglucosamine-6-phosphate deacetylase, NagA, from *Bacillus subtilis*. *J. Biol. Chem.* **279**, 2809–2816.
 16. Holm, L. & Sander, C. (1997). An evolutionary treasure: unification of a broad set of amidohydrolases related to urease. *Proteins: Struct. Funct. Genet.* **28**, 72–82.
 17. Yamano, N., Matushita, Y., Kamada, Y., Fujishima, S. & Arita, M. (1996). Purification and characterization of *N*-acetylglucosamine 6-phosphate deacetylase with activity against *N*-acetylglucosamine, from *Vibrio cholerae*. *Biosci. Biotech. Biochem.* **60**, 1320–1323.
 18. Shin, H. J., Kim, M. Y. & Lee, D. S. (1999). Purification and characterization of *N*-acetylglucosamine 6-phosphate deacetylase from *Thermus caldophilus*. *J. Biosci. Bioeng.* **8**, 319–322.
 19. Yamano, N., Higashida, N., Endo, C., Sakata, N., Fujishima, S., Maruyama, A. & Higashihara, T. (2000). Purification and characterization of *N*-acetylglucosamine-6-phosphate deacetylase from a psychrotrophic marine bacterium, *Alteromonas* sp. *Mar. Biotechnol.* **2**, 57–64.
 20. Ramakrishnan, C. & Ramachandran, G. N. (1965). Stereochemical criteria for polypeptide and protein chain conformations. 2. Allowed conformations for a pair of peptide units. *Biophys. J.* **5**, 909–933.
 21. Holm, L. & Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123–138.
 22. Jozic, D., Kaiser, J. T., Huber, R., Bode, W. & Maskos, K. (2003). X-ray structure of isoaspartyl dipeptidase from *E. coli*: a dinuclear zinc peptidase evolved from amidohydrolases. *J. Mol. Biol.* **332**, 243–256.
 23. Ireton, G. C., McDermott, G., Black, M. E. & Stoddard, B. L. (2002). The structure of *Escherichia coli* cytosine deaminase. *J. Mol. Biol.* **315**, 687–697.
 24. Liaw, S. H., Chen, S. J., Ko, T. P., Hsu, C. S., Chen, C. J., Wang, A. H. J. & Tsai, Y. C. (2003). Crystal structure of D-aminoacylase from *Alcaligenes faecalis* DA1. *J. Biol. Chem.* **278**, 4957–4962.
 25. Jabri, E., Carr, M. B., Hausinger, R. P. & Karplus, P. A. (1995). The crystal structure of urease from *Klebsiella aerogenes*. *Science*, **268**, 998–1004.
 26. Jones, S. & Thornton, J. (1995). Protein–protein interactions: a review of protein dimer structures. *Prog. Biophys. Mol. Biol.* **63**, 31–65.
 27. Lawrence, M. C. & Colman, P. M. (1993). Shape complementarity at protein/protein interfaces. *J. Mol. Biol.* **234**, 946–950.
 28. Coleman, J. E. (1998). Zinc enzymes. *Curr. Opin. Chem. Biol.* **2**, 222–234.
 29. de Seny, D., Heinz, U., Wommer, S., Kiefer, M., Meyer-Klaucke, W., Galleni, M. *et al.* (2001). Metal ion binding and coordination geometry for wild type and mutants of metallo-lactamase from *Bacillus cereus* 569/H/9 (BcII). A combined thermodynamic, kinetic, and spectroscopic approach. *J. Biol. Chem.* **276**, 45065–45078.
 30. Bicknell, R., Hanson, G. R., Holmquist, B. & Little, C. (1986). A spectral study of Cobalt(II)-substituted *Bacillus cereus* phospholipase-C. *Biochemistry*, **25**, 7208–7215.
 31. Javid-Majd, F. & Blanchard, J. S. (2000). Mechanistic analysis of the argE-encoded *N*-acetylornithine deacetylase. *Biochemistry*, **39**, 1285–1293.
 32. Wang, Z., Fast, W., Valentine, A. M. & Benkovic, S. J. (1999). Metallo- β -lactamase: structure and mechanism. *Curr. Opin. Chem. Biol.* **3**, 614–622.
 33. Ferreira, F. M., Mendoza-Hernandez, G., Calcagno, M. L., Minauro, F., Delboni, L. F. & Oliva, G. (2000). Crystallization and preliminary crystallographic analysis of *N*-acetylglucosamine 6-phosphate deacetylase from *Escherichia coli*. *Acta Crystallog. sect. D*, **56**, 670–672.
 34. Dauter, Z., Dauter, M. & Rajashankar, K. R. (2000). Novel approach to phasing proteins: derivatization by short cryo-soaking with halides. *Acta Crystallog. sect. D*, **56**, 232–237.
 35. Leslie, A. G. W. (1992). Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 and ESF-EAMCB Newsletter on Protein Crystallography*, **26**.
 36. Evans, P. R. (1997). Scala. *Joint CCP4 and ESF-EAMCB Newsletter on Protein Crystallography*, **33**, 22–24.
 37. Collaborative Computational Project, Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
 38. Schneider, T. R. & Sheldrick, G. M. (2002). Substructure solution with SHELXD. *Acta Crystallog. sect. D*, **58**, 1772–1779.
 39. Sheldrick, G. M. (2002). Macromolecular phasing with SHELXE. *Z. Kristallog.* **217**, 644–650.
 40. Terwilliger, T. C. & Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallog. sect. D*, **55**, 849–861.
 41. Terwilliger, T. C. (2002). Automated main-chain model-building by template-matching and iterative fragment extension. *Acta Crystallog. sect. D*, **59**, 34–44.
 42. Perrakis, A., Morris, R. & Lamzin, V. S. (1999). Automated protein model building combined with iterative structure refinement. *Nature Struct. Biol.* **6**, 458–463.
 43. McRee, D. E. (1999). XtalView/Xfit—a versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* **125**, 156–165.
 44. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog. sect. D*, **53**, 240–255.
 45. Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallog. sect. D*, **57**, 122–133.
 46. Vaguine, A. A., Richelle, J. & Wodak, S. J. (1999). SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallog. sect. D*, **55**, 191–205.
 47. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.
 48. Hoof, R. W., Vriend, G., Sander, C. & Abola, E. E. (1996). Errors in protein structures. *Nature*, **381**, 272.
 49. Russel, R. B. & Barton, G. J. (1992). Multiple protein sequence alignment from tertiary structure comparisons. *Proteins: Struct. Funct. Genet.* **14**, 309–323.

50. Hutchinson, E. G. & Thornton, J. M. (1996). PROMO-TIF—a program to identify and analyze structural motifs in proteins. *Protein Sci.* **5**, 212–220.
51. McDonald, I. K. & Thornton, J. M. (1994). Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* **238**, 777–793.
52. Svergun, D. I. (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallog.* **25**, 495–503.
53. Svergun, D. I., Petoukhov, M. V. & Koch, M. H. J. (2001). Determination of domain structure of proteins from X-ray solution scattering. *Biophys. J.* **80**, 2946–2953.
54. Volkov, V. V. & Svergun, D. I. (2003). Uniqueness of *ab initio* shape determination in small-angle scattering. *J. Appl. Crystallog.* **36**, 860–864.
55. Bond, C. S. (2003). TopDraw: a sketchpad for protein structure topology cartoons. *Bioinformatics*, **19**, 311–312.

Edited by M. Guss

(Received 7 February 2006; received in revised form 9 March 2006; accepted 12 March 2006)

Available online 29 March 2006