Journal of Structural Biology 173 (2011) 294–302

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Crystallographic analysis of a thermoactive nitrilase

Joanna E. Raczynska [a,1], Constantinos E. Vorgias [b], Garabed Antranikian [c], Wojciech Rypniewski [a,*]

[a] Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan, Poland
[b] Department of Biochemistry and Molecular Biology, National and Kapodistrian University of Athens, Athens, Greece
[c] Hamburg University of Technology, Institute of Technical Microbiology, Hamburg, Germany

ARTICLE INFO

Article history:
Received 12 July 2010
Received in revised form 4 November 2010
Accepted 15 November 2010
Available online 21 November 2010

Keywords:
Nitrilase
Carboxylate group binding
X-ray structure
Thermophilic enzyme

ABSTRACT

The nitrilase superfamily is a large and diverse superfamily of enzymes that catalyse the cleavage of various types of carbon–nitrogen bonds using a Cys–Glu–Lys catalytic triad. Thermoactive nitrilase from *Pyrococcus abyssi* (PaNit) hydrolyses small aliphatic nitriles like fumaro- and malononitryl. Yet, the biological role of this enzyme is unknown. We have analysed several crystal structures of PaNit: without ligands, with an acetate ion bound in the active site and with a bromide ion in the active site. In addition, docking calculations have been performed for fumaro- and malononitriles. The structures provide a proof for specific binding of the carboxylate ion and a general affinity for negatively changed ligands. The role of residues in the active site is considered and an enzymatic reaction mechanism is proposed in which Cys146 acts as the nucleophile, Glu42 as the general base, Lys113/Glu42 as the general acid, WatA as the hydrolytic water and Nζ_Lys113 and N_Phe147 form the oxyanion hole.

1. Introduction

Nitrilases are a large and diverse superfamily of enzymes that utilize a Cys–Glu–Lys catalytic triad to hydrolyse non-peptide carbon–nitrogen bonds. Based on the substrate specificity they have been subdivided into thirteen branches, of which only one possesses a true nitrilase activity. The most common reaction catalyzed by the family members is the hydrolysis of amides and N-carbamyl amides, such as N-carbamyl-D-amino acids or β-ureidopropionate. The nitrilase superfamily also contains N-acyltransferases, which catalyse the amidase reaction in reverse, condensing a fatty acid to an N-terminal cysteine residue (Pace and Brenner, 2001; Brenner, 2002). All of these reactions are initiated by nucleophilic attack by the active site cysteine residue on the cyano or carbonyl carbon atom. The glutamate is likely to play the role of a catalytic base while the lysine provides electrostatic stabilization as a part of the nitrilase equivalent of the oxyanion hole (Nakai et al., 2000). The first crystal structures of the superfamily members were that of N-carbamyl-D-amino acid amidohydrolase (DCase) from *Agrobacterium* (1erz, Nakai et al., 2000) and a NitFhit protein from *C. elegans*, (1ems, Pace et al., 2000). Both proteins exist as tetramers. The majority of microbial nitrilases form large homooligomeric spirals or helices (Thuku et al., 2009). Probably the best example is the nitrilase from *Rhodococcus rhodochrous* J1 which was shown to form active 'c' shaped oligomers of 480 kDa that rearrange themselves into long helices upon the truncation of a 39-amino-acid long, C-terminal fragment, most likely due to auto-proteolysis (Thuku et al., 2007).

Structures of enzymes from the nitrilase branch of the superfamily are few and the details of the catalytic mechanism remain elusive. Existing knowledge indicates a three-phase reaction with thioimidate and thiol-acylenzyme intermediates (Pace and Brenner, 2001) (Fig. 1). As can be seen, the process requires the addition of two water molecules (as opposed to only one in the case of amides). Another difference in comparison to amide hydrolysis is the need for the transfer of three protons to the terminal nitrogen atom before it is free to dissociate as ammonia. Whether acid catalysis is performed by the same amino acid residue as base catalysis (the glutamate), is not known. Kimani et al. (2007) from their studies of an amidase from *Geobacillus pallidus* concluded that another conserved glutamate residue (Glu142 in the *G. pallidus* enzyme) plays the role of a catalytic base for the hydrolytic water molecule.

We have tried to address these questions based on the structural analysis of a nitrilase from *Pyrococcus abyssi* (PaNit), which has been shown to hydrolyse small aliphatic dinitriles to the corresponding monoacid mononitriles. Fumaro- and malononitrile were identified as substrates with K_M values of 9.48 and 3.47 mM, respectively (Mueller et al., 2006). The enzyme exhibits maximum activity at 80 °C, which makes it an attractive candidate for

Abbreviations: PaNit, *Pyrococcus abyssi* nitrilase; DCase, N-carbamyl-D-amino acid amidohydrolase; pdb, protein data bank; r.m.s.d., root mean square deviation.

* Corresponding author. Address: Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznań, Poland. Fax: +48 61 852 05 32.
E-mail address: wojtekr@ibch.poznan.pl (W. Rypniewski).
[1] Present address: University of Texas Southwestern Medical Center at Dallas, USA.

1047-8477/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2010.11.017

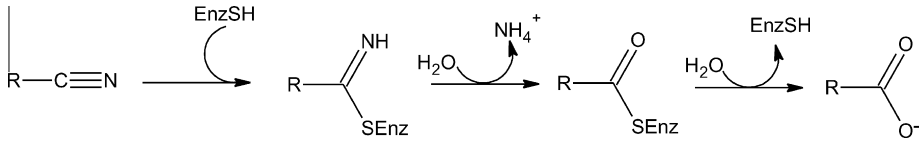

Fig. 1. A scheme of enzymatic hydrolysis of nitriles.

industrial applications. Still, the biological role of PaNit is unknown and so is the identity of amino acid residues responsible for the substrate binding. Here we report the first, to our knowledge, crystal structures of an enzyme from the nitrilase superfamily exhibiting true nitrilase activity. It is the only identified microbial nitrilase active as a dimer and not as a larger assembly.

2. Materials and methods

2.1. Materials

All enzymes used in the cloning procedures were supplied by New England Biolabs and the pcr2.1™ cloning system from InVitrogen. Column chromatography media were from Pharmacia and all the other chemicals from Sigma, Merck or Fluka in the highest analytical grade.

2.2. Overproduction and purification of recombinant NIT

The gene for the *nit-30* from *P. abyssi* was amplified from the chromosomal DNA of the archeon and cloned as described in Mueller et al. (2006). The cloned gene in the plasmid pET-11a was introduced in BL21(DE3) and the cells were grown in LB medium in the presence of ampicillin at 37 °C. The clone was induced at OD_{600nm} 0.8 by the addition of 1 mM IPTG. For preparative purposes 5 L culture was used and after 3 h induction the cells were collected by low-speed centrifugation. The bacterial pellet was lysed by sonication in an ice-water bath for 20 min with 30 s intervals in buffer A (20 mM Na-phosphate pH 8.0, 1 mM EDTA, 100 mM NaCl, 0.1% PMSF, 0.5% TX-100). Usually, 10 ml buffer A per 1 g bacterial paste was used. The total bacterial extract was clarified by centrifugation in a SS-34 Sorvall rotor at 20,000 rpm for 20 min. The extractable protein material was further fractionated by ammonium sulphate between 30% and 50% saturation, at 4 °C. The partially purified Nit was further purified on a Phenyl-Sepharose CL-6B column in buffer B (20 mM Na-phosphate pH 8.0, 1 mM EDTA, 1 M ammonium sulphate). Bound Nit was eluted via a descending gradient between 1 M and 0 M ammonium sulphate in buffer B. Fractions containing highly enriched Nit were combined, adjusted in buffer C (20 mM Na-phosphate pH 8.0, 1 mM EDTA) and directly applied on a 5 ml Q-Sepharose Fast Flow column equilibrated in buffer C. Nit was eluted from the column by a 100 ml linear gradient between 0 and 500 mM NaCl buffer C. In all cases the fractions were analysed by 0.1% SDS-10% PAGE. Pure Nit fractions were pooled, concentrated and further used for activity or crystallization experiments.

2.3. Protein crystallization

The purified PaNit was crystallized by the hanging drop/vapour diffusion method from 35% PEG 550MME, 0.2 M $MgCl_2$ or $Mg(CH_3COOH)_2$ and 0.2 M TrisHCl pH 7.5. Four structures were obtained: unliganded form (*Nit*), two complexes with acetate (*Nit–Acy–P2₁* and *Nit–Acy–P4₁*), and a structure with Br^- ions (*Nit–Br*). The diffraction data for the *Nit* structure were collected at ESRF, Grenoble, for the remaining structures at EMBL, Hamburg.

2.4. Data processing, structure solution and refinement

Data for the *Nit* and *Nit–Br* structures were merohedrally twinned, as indicated by the cumulative intensity distribution graphs produced by *Truncate* (Collaborative Computational Project, Number 4, 1994) when converting the intensities to the structure factors. Also, for these structures, problems existed with overlapping reflections from the neighbouring zones owing to a long unit cell axis (~130 Å, see Table 1). It was not possible to mount the crystals in such a way to avoid the unfavorable orientation with the long axis aligned with the beam. As a result the high resolution range for these data is largely incomplete, reducing the effective resolution of the structures. However, because the signal-to-noise ratio was still above 2 in the highest resolution shell these data were included in the refinement. The scaling and merging statistics are given for the full data. All data sets were processed and scaled with *HKL2000* (Otwinowski and Minor, 1997).

In *Nit* the twinning fraction estimated by the Yeates twin server (Yeates, 1997) was 0.48, which is too high for detwinning. Instead, the twin-related reflections were averaged to generate a perfectly twinned data set and thus avoid uncertainties in the twinning fraction value, as suggested in Yeates (1997). This was done as follows: the data were first scaled in the higher symmetry space group – $P4_122$ – and then duplicated by applying the twinning operator $k, h, -l$. The free set of reflections was chosen so that the twin-related reflections are both either in the test or in the working set. The structure was solved by molecular replacement with *Phaser* (McCoy et al., 2007) using the PH062 protein from *Pyrococcus horikoshi* as a model (pdb id 1j31, Sakai et al., 2004). The refinement was conducted in *Refmac* (v. 5.5.0092) Murshudov et al. (1997) against the twinned intensities.

For the *Nit–Acy–P4₁* structure, which has the same crystal form as Nit, the cumulative intensity distribution did not indicate twinning. The data were checked with the Yeates twin server which estimated the twinning fraction to be ~0.01. *Refmac* run with the twin keyword found no twin domains. In addition to the standard geometric restraints, NCS restraints were used. The *Nit–Acy–P2₁* structure was also refined in *Refmac*. Real space model building was done in *Coot* (Emsley and Cowtan, 2004) and figures were prepared with *PyMol* (DeLano). The electrostatic potential was calculated with APBS (Baker et al., 2001) using a PDB2PQR package (Dolinsky et al., 2004) to generate the .pqr file.

2.5. Nit–Br structure

This structure was prepared by 1-min soaking of PaNit crystals in a cryo-solution containing 1 M KBr for the purpose of a MAD experiment, as described by Dauter et al. (2000). Experimental phasing was unsuccessful, probably due to the twinning. The twinning fraction was 0.42 and the twin-related reflections were not averaged. The anomalous signal in the dataset collected at the Br absorption edge was of sufficient strength to locate the Br^- ions in the anomalous difference map. The same data set, with Friedel opposites merged, was used in the refinement, which was performed in *Refmac* against the twinned intensities. Temperature factors of the bound Br^- ions were refined anisotropically.

Table 1
X-ray data and refinement statistics.

	Nit	Nit–Acy–$P2_1$	Nit–Acy–$P4_1$	Nit–Br
Space group	$P4_1$	$P2_1$	$P4_1$	$P4_1$
Cell parameters	$a = b = 59.2$ $c = 127.2$	$a = 58.4$ $b = 67.2$ $c = 59.2$ $\beta = 90.8$	a = b = 59.2 c = 129.0	a = b = 59.3 c = 129.2
X-ray data				
No. observations	284,825	348,065	191,353	228,477
No. unique data	51,424	60,039	40,685	37,990
Resolution range [Å]	1.57 (1.59–1.57)[a]	1.60 (1.63–1.60)	1.80 (1.83–1.80)	1.76 (1.79–1.76)
Effective resolution [Å]	1.71[b]	1.60	1.80	1.84
$I/\sigma(I)$	11.8 (1.7)	20 (4)	16 (3)	32(2)
Redundancy	5.5 (1.4)	5.8 (4.9)	4.7 (3.5)	6 (2.1)
R_{merge} [%]	6.8 (19.3)	7.9 (37.2)	9.5 (48.3)	5.8 (43.1)
Completeness [%]	83.6 (16.0)	100.0 (99.4)	99.1 (92.2)	85.0 (30.1)
Refinement				
R_{work}/R_{free} [%]	13.1/17.6	14.5/18.7	15.9/20.6	16.2/20.4
r.m.s. d.				
Bond lengths [Å]	0.015	0.017	0.017	0.018
Angles [°]	1.6	1.7	1.6	1.8
No. protein atoms	4199	4318	4263	4195
No. ligand atoms	–	16	12	8
No. ions	1	2	1	15
No. waters	126	388	274	218
Average B factor				
Protein	23.7	18.2	19.4	28.2
Ligands	–	20.8	25.9	25.2
Ions	22.7	22.8	–	55.8
Waters	28.8	26.2	27.59	35.2
Residues in Ramachandran plot [%]				
Favored	98.2	98.2	97.7	97.9
Outliers	0	0	0	0

[a] Numbers in brackets indicate highest resolution shell.
[b] Calculated in the program DATAMAN by counting the number of reflections and estimating the resolution at which this number would correspond to a 100% complete data set; only reflections with $F \geqslant 2\sigma(F)$ were included.

2.6. Docking

The substrates, fumaro- and malononitrile, were docked into the active site of PaNit using the AutoDock4 program suite (Goodsell et al., 1996). The calculations were performed using a Lamarckian Genetic Algorithm (GA) with default parameters except for the number of runs, which was in the range of 20–50, and the population size, which was 200. The grid box was approximately a cube (edge length ~14 Å) containing the binding cavity. One of the water molecules (the binding water, see 'Acetate binding') was included in the protein model and Cys146 was changed to alanine.

2.7. Crystal packing

Cell parameters of both crystal forms are shown in Table 1. Only one parameter differs significantly between the two forms: in the $P4_1$ crystals the unit cell length along the symmetry axis (z) is 129 Å while in $P2_1$ form the parameter along the symmetry axis (y) equals 67 Å. The remaining cell parameters, including the β angle which is close to 90°, are roughly the same. As a consequence, the unit cell volume for the tetragonal crystals is approximately twice as large as that of the monoclinic form. In both crystal forms the molecules are arranged in layers running perpendicular to the symmetry axis and the intermolecular interactions within a layer are the same. In the tetragonal crystals, the molecules from the adjacent layers are rotated by 90° and form a cross-like arrangement. In the monoclinic structure the dimers from adjacent layers have their long axes approximately parallel to each other. A comparison between the sequences of PaNit and *R. rhodochrous* J1 nitrilase reveals that the residues engaged in interactions with neighbouring molecules within a layer approximately correspond to the areas responsible for formation of the spiral structures (C and D surfaces) (Thuku et al., 2007, 2009).

2.8. Validation and deposition

All four structures were validated with *Procheck* (Collaborative Computational Project, Number 4, 1994) and *Molprobity* (Davis et al., 2007). The coordinates and structure factors were deposited in the Protein Data Bank with pdb ids 3ivz (*Nit*), 3ki8 (*Nit–Acy–$P2_1$*), 3iw3 (*Nit–Acy–$p4_1$*) and 3klc (*Nit–Br*).

3. Results and discussion

3.1. Presented structures

We present a structure of the unliganded form of the enzyme, which crystallized in $P4_1$ space group (*Nit*) and three structures with an acetate ion bound near the active site: *Nit–Acy–$P4_1$* and *Nit–Br*, having the same crystal packing as *Nit*, and a monoclinic crystal form (*Nit–Acy–$P2_1$*). *Nit–Br* additionally contains Br^- ions and was originally prepared for a MAD experiment. Two of the analysed structures, *Nit* and *Nit–Br* are merohedrally twinned.

3.2. Overall structure

In both crystal forms, a dimer possessing 2-fold molecular symmetry is present in the asymmetric unit. Each subunit contains 262 residues and has an αββα sandwich fold, the same as described earlier for the nitrilase superfamily members (Nakai et al., 2000; Pace et al., 2000). When the two subunits associate, a 'super-sandwich' (αββααββα) structure is formed (Fig. 2A). The C-terminal part of

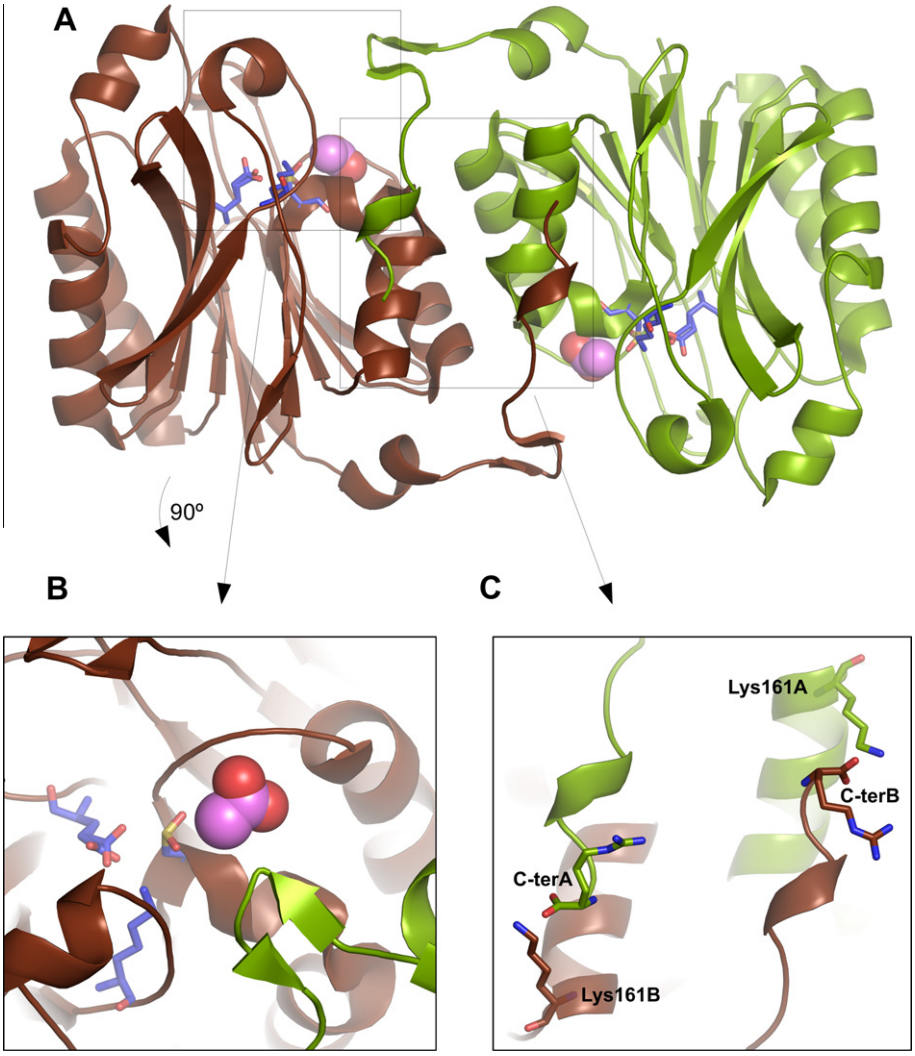

Fig. 2. Overall structure. (A) Nitrilase dimer shown as a cartoon model, view along the non-crystallographic twofold axis. The bound acetate ions are shown as spheres while the catalytic residues (Cys146, Glu42, Lys113) as sticks. (B) A close-up view of the binding site, the scene is rotated 90° downwards compared to A. (C) A detailed view of the C-termini, interacting residues are shown as sticks.

each subunit extends away from the core and interacts with the other subunit thus playing an important role in protein dimerization (Fig. 2). In all but one structures the carboxyl end forms a salt bridge with Lys161 of the other chain. The exception is the *Nit–Br* structure where the carboxylate group of Arg262A interacts with Arg255B (see 'Crystal packing' for details). The dimer interface contains hydrophobic as well as charged residues. Salt bridges between arginine and glutamate residues constitute a considerable part of the interactions responsible for the dimer formation.

3.3. Characterization of the binding site

The binding pocket, located near the inter-subunit interface, is lined with hydrophobic residues, mostly Phe. The bottom of the pocket is formed by the side chains of Cys146 and Leu172. The walls are built by the aromatic rings of Phe117, Phe147, Phe150 and Trp149, and the side chain of Val173, while Met174 adheres to the ligand from the outside. Three of the analysed structures contain an acetate ion bound near the active site and in each case it binds to the enzyme *via* its carboxylate group (Fig. 3A and B). The binding is accomplished by specific interactions with the main chain atoms of residues 173–177, having sequence VMPYA, described here as the binding loop. Each oxygen atom of the carboxylate group is an acceptor of two H-bonds: from the main chain nitrogen atoms of Val173 and Met174 for one O atom, and from Trp149 indole nitrogen and a water molecule for the other O atom. The water ('binding water') further interacts with N_Tyr176, N_Ala177 and O_Met174 from the binding loop. In *Nit* (which doesn't contain acetate ions), a water molecule is found in the same position. Additionally, two water molecules are present approximately in the same place as oxygen atoms of the carboxylate group in the acetate-containing structures (Fig. 3C). In Nit–Br the acetate ions bound by both protein subunits were modeled with partial occupancy (0.6 in chain A 0.5 in chain B). The anomalous difference map indicated the presence of Br^- ions in the same position, they were also refined with partial occupancy (0.4 in chain A and 0.5 in chain B). The electron density maximum in the anomalous map in chain A was 4.6σ high and in chain B 5.8σ. This difference can be explained by crystal contacts: molecule A is very closely surrounded by the symmetry neighbours while molecule B is generally more solvent exposed with easier access to the binding pocket. Strongest peaks in the map (>10σ) correspond to two Br^- ions bound on the surface of chain B. Penetration of Br^- ions deeply into the active site of PaNit during a short soak demonstrates its high affinity for negatively charged ions. No positively charged residues are present in this vicinity

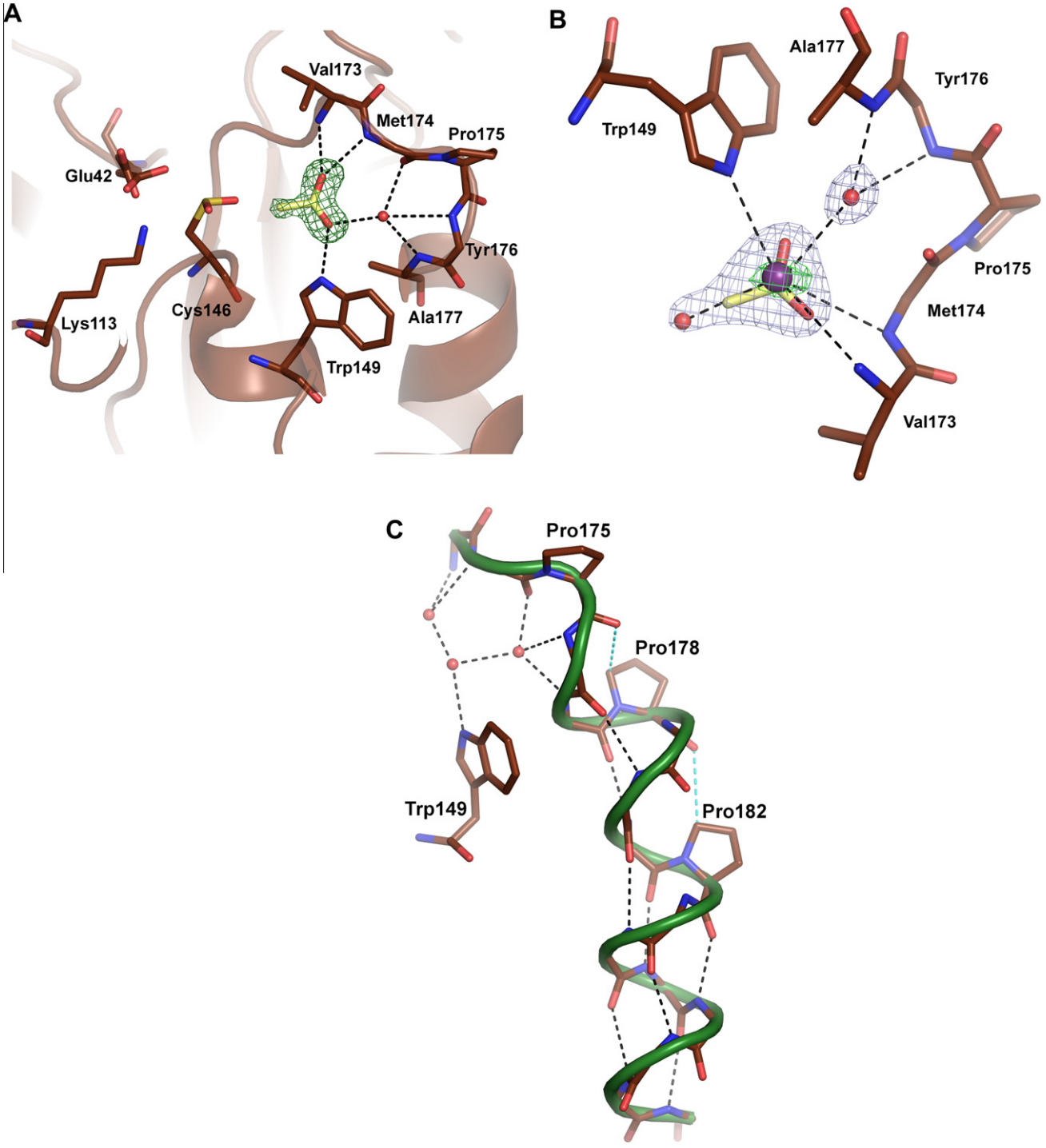

Fig. 3. The binding pocket of PaNit. (A) *Nit–Acy–P2$_1$* structure, the view is from the entrance to the cavity towards the core. The acetate ion is shown in the sigma-A weighted mF_o–DF_c omit electron density map, contoured at 4σ. Some side chains are omitted for clarity. The catalytic residues Cys146 (oxidized), Lys113 and Glu42 are also shown. (B) *Nit–Br* structure, the acetate is shown as sticks, the Br^- ion is coloured dark violet. The view is rotated 180 ° about the horizontal axis compared to A. In light blue the $2mF_o$–DF_c electron density map contoured at 1.5σ is shown, while in green is the anomalous difference map contoured at 5σ. (C) The binding loop and the supporting helix; the side chains have been shown only for the three Pro residues and Trp149. The two non-conventional C–H···O hydrogen bonds O_Pro175–Cδ_Pro178 (~3.4 Å) and O_Pro178–Cδ_Pro182 (~3.2 Å) are coloured cyan. After the second of these, there is a break in the H-bonding network (O_Arg179 has its hydrogen bonding potential unsatisfied) that corresponds to a kink in the helix. O_Tyr176, as well as O_Ala180, form two H-bonds with the N atoms of residues $n+3$ and $n+4$. The Figure was prepared using the unliganded *Nit* structure, which contains three water molecules interacting with the binding loop and Trp149. The two water molecules on the left are replaced by the carboxylate group in both *Nit–Acy* structures.

and the main source of the positive potential is the dipole moment of a three-turn helix, which begins at the end of the binding loop (Fig. 3C). The helix is kinked due to the presence of two additional (to the one found in the loop) Pro residues in positions 178 and 182. The beginning of the helix is slightly uncoiled, which is likely to be caused by the simultaneous effect of Pro175, Pro178 and Pro182. It seems that the binding loop is kept in place by this proline-rich motif.

3.4. Homologous structures

The most similar structure found in the PDB is that of PH0642 protein from *P. horikoshi* (annotated as 'hypothetical protein') (Sakai et al., 2004) whose sequence identity with PaNit is 86%, with r.m.s.d. between the two structures being 0.45 Å. Other proteins found by the EMBL–EBI server are: a pyrimidine degrading enzyme from *Drosophila melanogaster* (r.m.s.d. 1.42 Å, 2vhh, (Lundgren et al., 2008), Nit3 protein from *Saccharomyces cerevisiae* (r.m.s.d. 1.61 Å, 1f89, (Kumaran et al., 2003), aliphatic amidase from *Pseudomonas aeruginosa* (r.m.s.d. 1.62 Å, 2uxy, (Andrade et al., 2007), amidase from *Nesterenkonia sp*. (1.63 Å, 3hkx), a CN-hydrolase family protein from *Xanthomonas campestris* (r.m.s.d. 1.66, 2e11, (Chin et al., 2007), *Helicobacter pylori* formamidase (r.m.s.d. 1.71 Å, 2e2 l, (Hung et al., 2007), amidase from *G. pallidus* (r.m.s.d. 1.72 Å, 2plq, (Kimani et al., 2007), Nit2 from *Mus musculus* (r.m.s.d. 1.73 Å, 2w1v, (Barglow et al., 2008), N-carbamyl-D-amino acid amidohydrolase (DCase) from *Agrobacterium radiobacter* (r.m.s.d. 1.76–1.93 Å for various structures). They all share the same overall fold of a subunit but differ in terms of quaternary structure, some of them being tetramers (e.g. DCase) or hexamers (amidases from *P. aeruginosa* or *G. pallidus)*. The basic building block of these oligomers always consists of the two subunits corresponding to the nitrilase dimer (the super-sandwich structure). In all of these proteins the C-terminal part, of varying length, is engaged in the formation of this super-sandwich structure.

Structures of an inactive mutant (C161A) of DCase with several bound substrates, D-amino acids, reveal a different mode of the carboxylate group binding than observed in the presented structures. On the other hand, an acetate ion found in the structure of PH0642 protein has exactly the same interactions as in PaNit. In none of the other homologous proteins present in the PDB is the structure of the binding loop conserved. Analysis of homologous sequences (data not shown) revealed that most contained 8–11 amino acid insertions in the binding loop region. The observed binding site architecture seems to exist only in a small number of closely related proteins.

3.5. The active site and implications for the catalytic mechanism

Sequence and structure comparison with other proteins from the nitrilase superfamily revealed that the catalytic residues in PaNit are Glu42, Lys113 and Cys146. Another highly conserved amino acid residue in this region is Glu120. The carboxylate groups of both glutamate residues interact with the Nζ atom of Lys113. Also, the Oε2 atom of Glu42 interacts with Sγ_Cys146. The details of these interactions vary between structures in terms of distances, interacting atoms and positions of the water molecules. The active site hydrogen bond network is shown in Fig. 4 while all of the distances are given in the Supplementary Table. Glu42 has been modeled in two alternative conformations in the *Nit–Acy* structures. In one of them, the carboxylate group is held in place by interactions with Nδ2_Asn97 and N_Asn171, and in *Nit–Acy–P2$_1$* the contact between Oε2_Glu42 and the Sγ atom of the active cysteine is very short: 2.59 Å in chain B and 2.63 Å in chain A. This observation supports the role of Glu42 as the catalytic base responsible for activation of the nucleophile (Cys146), as suggested by Nakai et al. (2000) and Yeates (1997).

The catalytic cysteine was oxidized, with either one (*Nit* and *Nit–Br*) or two (both *Nit–Acy* structures) oxygen atoms connected to the Sγ atom, each engaged in H-bonds. In some aspects these oxygen atoms resemble the reaction intermediates and when their interactions are investigated, insight into the catalytic mechanism can be gained. And so the Oδ2 atom interacts only with N_Asn171, while the Oδ1 atom (as well as the Oδ atom from the singly oxidized cysteine) forms three hydrogen bonds: with N_Phe147, Nζ_Lys113 and with a water molecule, WatA, which further interacts with Oη_Tyr48 and O_Asn171. N_Phe147 and Nζ_Lys113 were previously thought to constitute the nitrilase equivalent of the oxyanion hole and our structures confirm this proposal. The binding of WatA seems to be promoted by the Oδ/Oδ2 atom of Cys146, because no water is found in this place in the structure of PH0642 protein from *P. horikoshi*, where the catalytic cysteine is not oxidized. This information, however, cannot be verified because the structure factors for the PH0642 protein have not been released. Still, it seems to be a reasonable position for the actual hydrolytic water molecule. Another important observation is that the oxygen atom bound in the oxyanion hole (Oδ1 or Oδ in the analysed structures) is located ~3.7 Å from the carboxylate group of Glu42. In the nitrile hydrolysis reaction the creation of the thioimidate requires a proton to be donated to the terminal nitrogen

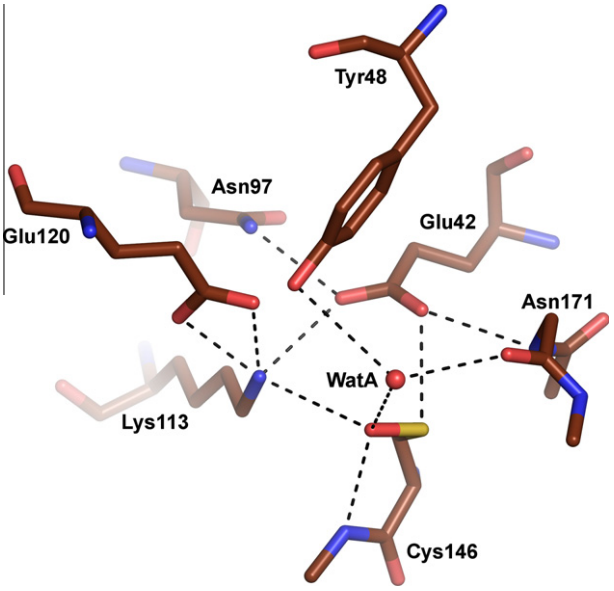

Fig. 4. Hydrogen bonding network in the active site of Nit structure, where the catalytic cysteine is only singly oxidized. An ordered water molecule WatA is shown as red sphere.

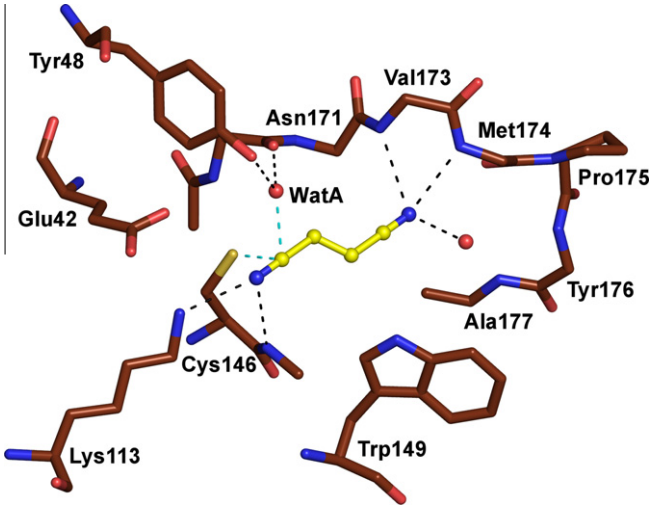

Fig. 5. Fumaronitrile (yellow) docked into the active site of PaNit. One end of the ligand interacts with the binding loop and the binding water in a similar way as the acetate ions. The to be hydrolysed cyano carbon atom is in a close contact with the nucleophile (Cys146) and the putative hydrolytic water molecule WatA; both contacts indicated by the cyan-coloured dashed lines. The N atom is located in the oxyanion hole.

atom of the nitrile substrate, and such a distance to Glu42 would not allow for the acid catalysis. In our opinion the likeliest possibility for the catalytic acid is Lys113. In this case, the proton transfer from Nζ_Lys113 to the substrate could occur concurrently to the nucleophilic attack and the Cys146 → Glu42 proton transfer. It would also explain the high conservation of Glu120 because its presence enhances the pKa of Lys113 ensuring that it will be positively charged and ready to donate a proton to the substrate.

3.6. Possible binding mode of the substrates

Attempts to obtain structures with fumaro- and malononitrile (or the products of their hydrolysis) were unsuccessful and the exact binding mode of these two ligands remains unknown. It is likely that both ligands would interact with the binding loop. The binding cavity is rich in hydrogen donors (Nε1_Trp149, N_Val113, N_Met174, N_Tyr176 and N_Ala177) and each can be possibly engaged in the interaction with one of the CN groups. The nitrile binding can also be accomplished *via* a water molecule, as is in the case of the carboxylate group. Yet only such a binding mode, in which the carbon atom from the to-be-hydrolysed CN group is in close contact with Cys146, would be productive. To gain more insight into the binding mode of the substrates they have been docked into the active site of PaNit using AutoDock4 program suite. Since the used docking software is not able to predict covalent linkages that may form between the protein and the ligand, the distances between non-bonded atoms cannot be lower than the sum of their Van der Waals radii. However, it excludes finding

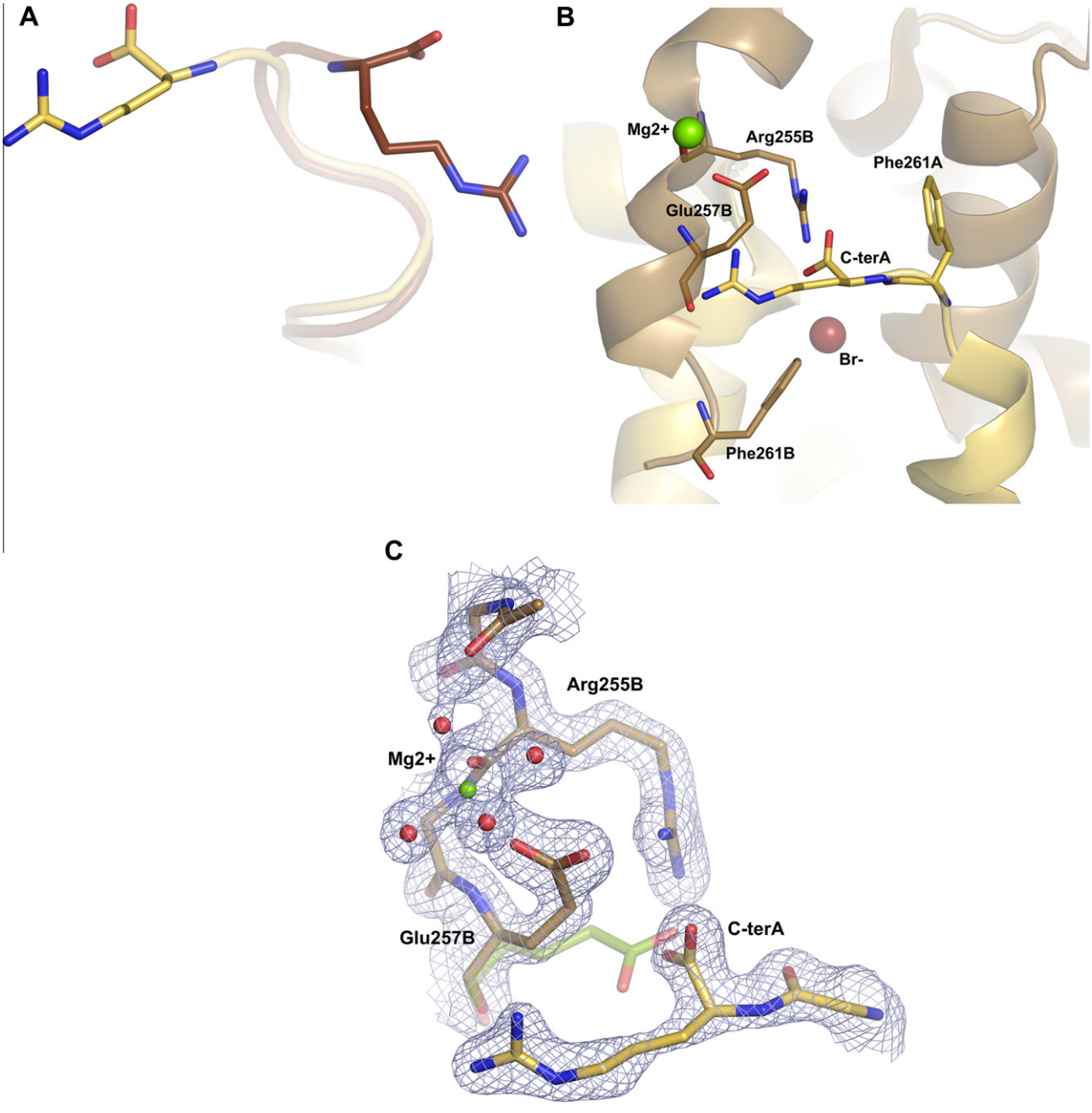

Fig. 6. The C-terminus in *Nit–Br* structure. (A) Comparison with the *Nit–Acy–P2₁* structure (brown). (B) Details of the interactions involving the C-terminus of chain A, Arg255B, Glu257B and Mg^{2+} (green sphere), as well as a Br^{-} ion and Phe261B. (C) The above residues in the $2mF_o–DF_c$ electron density map (1.1σ); green sticks represent Glu257B from *Nit–Acy–P2₁* , which is identical as in the remaining structures (chains).

a productive binding conformation of the substrates which requires a close contact between Sγ_Cys146 and one of the nitrile carbons. A solution to that problem is to mutate the cysteine into alanine in the input pdb file. That way, the results can be 'cross-validated' by analyzing interactions with the Sγ atom of Cys146 which was not included in the calculations. After the docking, the Sγ-$C_{nitrile}$ distances are ~2.3 Å for both ligands. Additionally, the nitriles are in a close contact (2.4 Å) with a bound water molecule (WatA, not included in the calculations). The final result for fumaronitrile is shown in Fig. 5.

3.7. Mg^{2+} ions, Br^- ions and the C-terminus

Crystal packing is influenced by Mg^{2+} ions which are present in each of the analysed structures. In the monoclinic crystal form of PaNit, two magnesium ions are found in positions not related by the non-crystallographic dyad. The first one is coordinated by six water molecules, which are H-bonded to the carbonyl oxygen atoms of Ile63A and Lys90A, to the carboxylate groups of Glu65A and Asp91A, and to Oη_Tyr176B from a molecule related by the $-x, y - 1/2, -z$ symmetry operation. Mg^{2+} bound in chain B interacts with five water molecules and Asn49B. An acetate ion is found close and it interacts with two coordinating water molecules.

In *Nit* and *Nit–Acy–P4₁* the ion is found in the same position, coordinated by the carbonyl oxygen atom of Met222A and water molecules, five in *Nit*, and only three in *Nit–Acy–P4₁*. Two of the water molecules further interact with the carboxylate groups of Glu229A and Glu258A[$-y, x, z + 1/4$]. Surprisingly, in Nit–Br, which has the same crystal packing, the ion is found in a different position. It is coordinated by the Oε2 atom of Glu257B, O_Arg255B and four water molecules, one of which is hydrogen-bonded to O_Val219B[$y + 1, -x + 1, z - 1/4$]. As can be deduced from the identifiers of the interacting residues (Glu258A and Glu257B), the location of the ions within a single subunit is similar and, taking into account the crystal symmetry, their positions in these structures are approximately related by the non-crystallographic two-fold axis. Due to the interaction with Mg^{2+} the side chain of Glu257B is displaced relative to the remaining structures, where it interacts with the guanidinium group of Arg255 from the same chain. This results in a different arrangement of the C-terminus of chain A in *Nit–Br* structure (Fig. 6) such that it stabilizes the unsatisfied positive charge of Arg255B.

The conformational change of the C-terminus in *Nit–Br* structure is associated with unusual position of Phe261A phenyl ring on the protein surface. The side chain itself is disordered and not well defined by the electron density map. If Phe261A occupied the same position as in the remaining structures the phenyl ring would be ~3.5 Å away from a bound Br^- ion (268A), which could have been unfavorable. However, the ion is located on the non-crystallographic dyad and it has a 3.5 Å contact with Phe261B (Fig. 6). Still, the different arrangement of the C-terminus, together with the change in position of the magnesium ion in *Nit-Br* structure, are most likely caused by the soaking.

Comparison with the structure of PH0642 protein from *P. horikoshi* reveals a different crystal packing ($P2_1$ space group with four molecules in the asymmetric unit) associated with yet another arrangement of the C-terminus. The fact that three possible arrangements of the C-terminus were observed indicates its flexibility. This behaviour can be the cause, or one of the causes, of the protein's tendency to aggregation.

4. Conclusions

The structures provide a proof for specific binding of the carboxylate group, as well as a more general electrostatic preference for negatively charged ligands revealed by binding of the Br^- ions. The key element in PaNit structure is the binding loop (residues 173–177) whose main chain atoms are H-bonded to the COO^- group or Br^- directly or *via* a water molecule. The side chain of Trp149 also takes part in the binding. The binding site has a positive electrostatic potential which suggests that other negatively charged groups could bind in a similar fashion.

Since all interactions between the binding loop and the ligands (direct or water-mediated) are with the protein main chain, point mutations within the loop should not impede PaNit specificity. The exception is Pro175 which seems to be important for keeping the binding loop in the proper conformation; Pro178 and Pro182 from the helix following the loop are likely to play similar role and mutating each of them would probably alter the observed loop-helix architecture. It would be interesting to see how modifications such as insertions of one or two amino acids in the loop region would affect the specificity of PaNit.

Nitrilases are not very well studied and the description of the mechanism is speculative. But structural analysis is quite informative in this respect and often allows inferring the dynamics of the chemical reaction from the basically static picture. Particularly, it can be said that the importance of Lys113 arises from its ability to act as an acid and donate the proton to the substrate. A concerted acid–base catalysis by Lys113 and Glu42 is the most likely mechanism of the nitrilase reaction.

Acknowledgments

J.R. and W.R. acknowledge support of the Ministry of Science and Higher Education (Poland, N N301 071534), the European Community Research Infrastructure Action under the FP6 'Structuring the European Research Area Programme' (contract number RII3/CT/2004/5060008). J.R., C.V., W.R. received support from the Polish–Greek bilateral cooperation program. J.R. thanks D. Borek and Z. Otwinowski for help in dealing with the twinned diffraction data, and K. Henderson for reading the manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jsb.2010.11.017.

References

Andrade, J., Karmali, A., Carrondo, M.A., Frazão, C., 2007. Structure of amidase from Pseudomonas aeruginosa showing a trapped acyl transfer reaction intermediate state. J. Biol. Chem. 282, 19598–19605.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., McCammon, J.A., 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl. Acad. Sci. USA 98, 10037–10041.

Barglow, K.T., Saikatendu, K.S., Bracey, M.H., Huey, R., Morris, G.M., Olson, A.J., Stevens, R.C., Cravatt, B.F., 2008. Functional proteomic and structural insights into molecular recognition in the nitrilase family enzymes. Biochemistry 47, 13514–13523.

Brenner, C., 2002. Catalysis in the nitrilase superfamily. Curr. Opin. Struct. Biol. 12, 775–782.

Chin, K., Tsai, Y., Chan, N., Huang, K., Wang, A.H., Chou, S., 2007. The crystal structure of XC1258 from *Xanthomonas campestris*: a putative procaryotic Nit protein with an arsenic adduct in the active site. Proteins 69, 665–671.

Collaborative Computational Project, Number 4, 1994. The CCP4 suite: programs for protein crystallography. D Biol. Crystallogr. 50, 760–763.

Dauter, Z., Dauter, M., Rajashankar, K.R., 2000. Novel approach to phasing proteins: derivatization by short cryo-soaking with halides. Acta Crystallogr. D Biol. Crystallogr. 56, 232–237.

Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B., Snoeyink, J., Richardson, J.S., Richardson, D.C., 2007. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res. 35, W375–83.

DeLano, W.L. The PyMOL Molecular Graphics System. DeLano Scientific LLC, San Carlos, CA, USA. <http://www.pymol.org>.

Dolinsky, T.J., Nielsen, J.E., McCammon, J.A., Baker, N.A., 2004. PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations. Nucleic Acids Res. 32, W665–7.

Emsley, P., Cowtan, K., 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2232.

Goodsell, D.S., Morris, G.M., Olson, A.J., 1996. Automated docking of flexible ligands: applications of AutoDock. J. Mol. Recognit. 9, 1–5.

Hung, C., Liu, J., Chiu, W., Huang, S., Hwang, J., Wang, W., 2007. Crystal structure of *Helicobacter pylori* formamidase AmiF reveals a cysteine–glutamate–lysine catalytic triad. J. Biol. Chem. 282, 12220–12229.

Kimani, S.W., Agarkar, V.B., Cowan, D.A., Sayed, M.F.R., Sewell, B.T., 2007. Structure of an aliphatic amidase from *Geobacillus pallidus* RAPc8. Acta Crystallogr. D Biol. Crystallogr. 63, 1048–1058.

Kumaran, D., Eswaramoorthy, S., Gerchman, S.E., Kycia, H., Studier, F.W., Swaminathan, S., 2003. Crystal structure of a putative CN hydrolase from yeast. Proteins 52, 283–291.

Lundgren, S., Lohkamp, B., Andersen, B., Piskur, J., Dobritzsch, D., 2008. The crystal structure of beta-alanine synthase from Drosophila melanogaster reveals a homooctameric helical turn-like assembly. J. Mol. Biol. 377, 1544–1559.

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

Mueller, P., Egorova, K., Vorgias, C.E., Boutou, E., Trauthwein, H., Verseck, S., Antranikian, G., 2006. Cloning, overexpression, and characterization of a thermoactive nitrilase from the hyperthermophilic archaeon *Pyrococcus abyssi*. Protein Expr. Purif. 47, 672–681.

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

Nakai, T., Hasegawa, T., Yamashita, E., Yamamoto, M., Kumasaka, T., Ueki, T., Nanba, H., Ikenaka, Y., Takahashi, S., Sato, M., Tsukihara, T., 2000. Crystal structure of N-carbamyl-D-amino acid amidohydrolase with a novel catalytic framework common to amidohydrolases. Structure 8, 729–737.

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

Pace, H.C., Brenner, C. The nitrilase superfamily: classification, structure and function. Genome Biol. 2 (2001) REVIEWS0001.

Pace, H.C., Hodawadekar, S.C., Draganescu, A., Huang, J., Bieganowski, P., Pekarsky, Y., Croce, C.M., Brenner, C., 2000. Crystal structure of the worm NitFhit Rosetta stone protein reveals a Nit tetramer binding two Fhit dimers. Curr. Biol. 10, 907–917.

Sakai, N., Tajika, Y., Yao, M., Watanabe, N., Tanaka, I., 2004. Crystal structure of hypothetical protein PH0642 from *Pyrococcus horikoshii* at 1.6 Å resolution. Proteins 57, 869–873.

Thuku, R.N., Weber, B.W., Varsani, A., Sewell, B.T., 2007. Post-translational cleavage of recombinantly expressed nitrilase from *Rhodococcus rhodochrous* J1 yields a stable active helical form. FEBS J. 274, 2099–2108.

Thuku, R.N., Brady, D., Benedik, M.J., Sewell, B.T., 2009. Microbial nitrilases: versatile, spiral forming, industrial enzymes. J. Appl. Microbiol. 106, 703–727.

Yeates, T.O., 1997. Detecting and overcoming crystal twinning. Methods Enzymol. 276, 344–358.