Crystal Structures of Human NUDT5 Reveal Insights into the Structural Basis of the Substrate Specificity

ARTICLE *in* JOURNAL OF MOLECULAR BIOLOGY · JANUARY 2007

Impact Factor: 4.33 · DOI: 10.1016/j.jmb.2006.09.078 · Source: PubMed

5 AUTHORS, INCLUDING:

Chen Zhong
Chinese Academy of Sciences
33 PUBLICATIONS **773** CITATIONS
SEE PROFILE

Jianping Ding
Chinese Academy of Sciences
142 PUBLICATIONS **7,884** CITATIONS
SEE PROFILE

doi:10.1016/j.jmb.2006.09.078 J. Mol. Biol. (2006) 364, 1021–1033

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Crystal Structures of Human NUDT5 Reveal Insights into the Structural Basis of the Substrate Specificity

Manwu Zha[1,2], Chen Zhong[1], Yingjie Peng[1], Hongyu Hu[1] and Jianping Ding[1]*

[1]*State Key Laboratory of Molecular Biology, Institute of Biochemistry and Cell Biology Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences 320 Yue-Yang Road Shanghai 200031, China*

[2]*Graduate School of Chinese Academy of Sciences 320 Yue-Yang Road Shanghai 200031, China*

**Corresponding author*

Human NUDT5 (hNUDT5) is an ADP-ribose pyrophosphatase (ADPRase) belonging to the Nudix hydrolase superfamily. It presumably plays important roles in controlling the intracellular level of ADP-ribose (ADPR) to prevent non-enzymatic ADP-ribosylation by hydrolyzing ADPR to AMP and ribose 5′-phosphate. We report here the crystal structures of hNUDT5 in apo form, in complex with ADPR, and in complex with AMP with bound Mg^{2+}. hNUDT5 forms a homodimer with substantial domain swapping and assumes a structure more similar to *Escherichia coli* ADPRase ORF209 than human ADPRase NUDT9. The adenine moiety of the substrates is specifically recognized by the enzyme *via* hydrogen-bonding interactions between N1 and N6 of the base and Glu47 of one subunit, and between N7 of the base and Arg51 of the other subunit, providing the molecular basis for the high selectivity of hNUDT5 for ADP-sugars over other sugar nucleotides. Structural comparisons with *E. coli* ADPRase ORF209 and ADPXase ORF186 indicate that the existence of an aromatic residue on loop L8 in ORF186 seems to be positively correlated with its enzymatic activity on AP_nA, whereas hNUDT5 and ORF209 contain no such residue and thus have low or no activities on AP_nA.

Keywords: Nudix domain; ADPR; ADP-ribose pyrophosphatase; NUDT5; substrate specificity

Introduction

Nudix pyrophosphatases are a superfamily of enzymes widely distributed in nature that catalyze the hydrolysis of a nucleoside diphosphate linked to another moiety X. They are characterized by the presence of a highly conserved amino acid sequence motif $GX_5EX_7REUXEEXGU$ (where U is usually Ile, Leu, or Val).[1] This motif forms a loop-helix-loop structure and is involved in catalysis and particularly in binding of Mg^{2+} or other divalent cations. These enzymes are considered as "sanitizers" for cells because their substrates are either potentially toxic nucleoside diphosphate derivatives or physiological nucleoside metabolites. ADP-ribose (ADPR) is one such substrate of the Nudix enzymes. Free ADPR is a highly active metabolite intermediate generated by turnover of NAD^+, cyclic ADPR, and protein-bound poly and mono-ADPR.[2] High concentration of free ADPR could result in non-enzymatic ADP-ribosylation of proteins,[3] a process implicated to be deleterious by accumulating evidence. For example, non-enzymatic ADP-ribosylation of actin at a cysteine residue can result in inhibition of actin polymerization.[4] Accumulation of ADPR can lead to inhibition of ATP-activated K^+ channels.[5] In addition, the cytotoxic effect of unusual ADP-ribosylation is also suggested by the action of several bacterial toxins (reviewed in references[6,7]). For instance, diphtheria toxin can cause ADP-ribosylation and inactivation of elongation factor EF2, subsequent inhibition of protein synthesis, and ultimate cell death.[8] Thus, precise control of the ADPR level is important in cell maintenance.

Abbreviations used: ADPR, ADP ribose; ADPRase, ADP-ribose pyrophosphatase; EcORF186, *E. coli* ADPXase encoded by *orf186*; EcORF209, *E. coli* ADPRase encoded by *orf209*; hNUDT5, human ADPRase NUDT5; hNUDT9, human ADPRase NUDT9; MtADPRase, *Mycobacterium tuberculosis* ADPRase; TtADPRase, *Thermus thermophilus* ADPRase; R5P, ribose 5′-phosphate; ORF, open reading frame; MAD, multiwavelength anomalous dispersion; MR, molecular replacement.

E-mail address of the corresponding author: jpding@sibs.ac.cn

0022-2836/$ - see front matter

Catabolism of free ADPR occurs by its hydrolysis to AMP and ribose 5′-phosphate (R5P), a process catalyzed by the ADP-ribose pyrophosphatase (ADPRase) subfamily of the Nudix enzymes in the presence of divalent metal ions. The activities of ADPRases have been found in species from archaea, eubacteria to eukaryotes. Biochemical studies have shown that ADPRases from different species exhibit significant divergence in terms of substrate specificity and ion selectivity. For example, *Methanocaldococcus jannaschii* ADPRase can only efficiently hydrolyze ADPR and its derivative 2′-phospho-ADPR and requires strictly Mg^{2+} for optimal activity.[9] In contrast, *Escherichia coli orf186* encodes a Nudix pyrophosphatase (hereinafter referred to as EcORF186) which can take a wide range of ADP-containing nucleoside pyrophosphates as the substrate, such as ADPR, NADH, AP_2A, AP_3A, and AP_37-*methyl*-G (therefore we designate it as ADPXase), and uses Mg^{2+} as well as Zn^{2+} and Mn^{2+} for catalysis.[10] The *E. coli* ADPRase encoded by gene *orf209* (EcORF209) is very specific for ADP-sugars and has about 40% of the maximal activity when Mg^{2+} is replaced by Zn^{2+}. In humans, at least two ADP-sugar pyrophosphatase activities have been observed in erythrocytes[11,12] and placenta.[13] So far, two human cDNA clones, designated as *NUDT5*[14,15] and *NUDT9*,[16] respectively, have been identified to encode proteins with ADPRase activities, which have distinct substrate specificity. Human NUDT5 (hNUDT5) is able to catalyze hydrolysis of a variety of ADP-sugar conjugates with a preference for ADPR, and the optimal Mg^{2+} can be partially substituted by Zn^{2+} and Mn^{2+}.[14,15] Human NUDT9 (hNUDT9) proteins are produced from the major transcript *NUDT9*α and the minor *NUDT9*β. hNUDT9α is highly specific for ADPR and non-physiological IDPR, and Mn^{2+} can partially substitute for Mg^{2+}.[16]

To better understand the underlying mechanism of hydrolysis and substrate specificity of different ADPRases, crystal structures of four ADPRases and one closely related ADPXase have been determined to date, one from *Homo sapiens* (hNUDT9)[17] and four from prokaryotes (*E. coli*, EcORF186 and EcORF209; *Mycobacterium tuberculosis*, MtADPRase; and *Thermus thermophilus*, TtADPRase).[18–21] Though the catalytic domains of all these enzymes contain the core Nudix fold, there are substantial differences in both the overall structure and the interactions between the protein and the substrate. The bacterial enzymes form a homodimer with two equivalent catalytic active sites each comprising residues from both subunits, and the N-terminal domain is involved in dimerization through a domain swap that is essential for substrate recognition and enzymatic activity. However, hNUDT9 is active as a monomer consisting of an N-terminal domain and a C-terminal catalytic domain and the substrate is bound in a cleft between the two domains. Thus, the recognition and binding of the substrate and the mechanism of hydrolysis appear to vary between different ADPRases.[21]

We report here the crystal structures of the full-length hNUTD5 in apo form and a truncated form of the enzyme in complexes with ADPR and AMP with bound Mg^{2+}. Our results indicate that hNUDT5 has a structure similar to those of bacterial enzymes, but different from that of hNUDT9. Structural analysis of hNUDT5 reveals insights into the molecular basis of the substrate specificity. Structural comparisons of hNUDT5 with EcORF209 and EcORF186 also provide a possible explanation for their differing hydrolase activities on AP_nA.

Results and Discussion

Overall structure of hNUDT5

The structure of a truncated hNUDT5 (residues 1–210, ΔhNUDT5) that omitted the C-terminal nine hypervariable residues was first solved at 2.6 Å resolution using the multiple wavelength anomalous dispersion (MAD) method and the structure of the full-length hNUDT5 (residues 1–219) was determined at 2.5 Å resolution using the molecular replacement (MR) method (Table 1). The experimental electron density map is of great quality except that the N-terminal 1–13 residues were disordered without electron density and therefore were omitted in the final model. The apo form full-length hNUDT5 structure contains four hNUDT5 monomers in the asymmetrical unit, forming two homodimers related by a 2-fold non-crystallographic symmetry. The two subunits in the homodimer adopt very similar conformation (RMSD of 0.32 Å for 206 C^{α} atoms). Each hNUDT5 subunit can be divided into an N-terminal domain (residues 14–53), the catalytic Nudix domain (residues 54–210), and an extended α-helix (residues 211–219) at the C terminus (Figure 1). The N-terminal domain comprises three-stranded antiparallel β-sheet from β1 to β3 and is involved in dimerization, which resembles that in the dimeric ADPRases from bacteria,[18–21] but differs from that in the monomeric hNUDT9[17] (see detailed discussion later). The dimer displays substantial domain swapping such that the N-terminal domain of one subunit crosses over to interact with the Nudix domain of the other, forming a major portion of the dimer interface. This part of the interface is stabilized by extensive hydrophobic interactions as well as salt-bridges and hydrogen bonds between secondary elements β1, β2, β3, β6, loop L1, and loop L9 of both subunits. The Nudix domain assumes an α + β fold with mixed β-sheets of β4 to β11 and three flanking α-helices (Figure 1). As in other ADPRases, the highly conserved Nudix motif (residues 97–119) is folded into a typical loop-helix-loop structure. The Nudix domain of one subunit has extensive contacts with the Nudix domain of the other and contributes to dimer formation mainly through hydrogen-bonding interactions between residues of β8, L8, and β10 of one subunit and those of β11* and α3* of the other (for simplicity, residues and secondary structural

Table 1. Statistics of diffraction data and structure refinement

	Se-Met ΔhNUDT5			hNUDT5	ΔhNUDT5 complex	
Data collection						
Data sets	Peak	Edge	Remote	Apo form	ADPR	AMP
Space group		$P6_522$		$P2_1$	$C2$	$C2$
a (Å)		90.7		81.4	112.5	112.1
b (Å)		90.7		70.1	40.6	40.2
c (Å)		207.2		86.8	99.7	100.1
β (°)				100.6	121.4	121.7
Wavelength (Å)	0.9794	0.9797	0.9819	1.0000	1.0000	1.0000
Resolution (Å)[a]	50.0–2.60 (2.69–2.60)	50.0–2.60 (2.69–2.60)	50.0–2.60 (2.69–2.60)	50.0–2.50 (2.59–2.50)	50.0–2.00 (2.07–2.00)	50.0–2.60 (2.69–2.60)
Observed reflections	212,147	212,130	211,421	114,743	80,903	62,858
Unique reflections	16,240	16,215	16,265	32,826	24,484	12,038
Mosaicity	0.20	0.18	0.17	0.78	0.53	1.31
Redundancy	13.1 (8.5)	13.1 (8.5)	13.0 (8.4)	3.5 (2.6)	3.3 (2.1)	5.2 (4.3)
$I/\sigma(I)$	39.2 (2.7)	39.0 (2.5)	40.5 (2.8)	20.1 (1.8)	27.0 (2.4)	18.7 (1.7)
Completeness (%)	99.7 (99.2)	99.4 (97.5)	99.7 (99.4)	96.3 (81.2)	92.8 (62.2)	97.5 (87.0)
R_{merge} (%)[b]	9.1 (39.6)	8.9 (43.7)	8.4 (48.3)	7.0 (41.9)	5.4 (26.7)	11.9 (46.9)
Refinement statistics						
R_{cryst}[c]/R_{free}[d]				21.4/26.2	19.4/22.8	21.5/27.5
Amino acids (atoms)				824 (6440)	410 (3032)	410 (3032)
Number of water molecules				218	291	96
Number of metals					2	6
Number of ligands					2	2
Average B factor of protein atoms				68.8	37.5	60.2
Average B factor of metal ions					34.0	60.9
Average B factor of ligands					45.3	77.4
RMSD bond lengths (Å)				0.007	0.006	0.008
RMSD bond angles (°)				0.84	1.06	0.87
Luzzati atomic positional error (Å)				0.34	0.23	0.35
Ramachandran plot						
Most favored (%)				89.8	90.2	89.3
Allowed (%)				10.2	9.8	10.7

[a] The numbers in parentheses refer to the highest resolution shell.
[b] $R_{merge} = \sum_{hkl}\sum_i | I_{,i}(hkl)_i - \langle I(hkl)\rangle | / \sum_{hkl}\sum_i (hkl)$.
[c] $R_{cryst} = \sum || F_o | - | F_c || / \sum | F_o |$.
[d] R_{free} is monitored with 5% of randomly selected reflections.

elements from the second subunit are indicated with an asterisk hereafter). Together, the dimerization buries 3038 Å^2 (23.8%) of the accessible surface area of each subunit.

Attempts to obtain crystals with good diffraction quality of full-length hNUDT5 in complex with ADPR or AMP by either cocrystallization or soaking of the apo crystals of full-length hNUDT5 with the ligands were unsuccessful. Thus, a truncated form of hNUDT5 with deletion of the C-terminal nine residues (ΔhNUDT5) was constructed and used for the structural studies. The structure of the ΔhNUDT5-ADPR complex was solved at 2.0 Å resolution using the MR method. There are two monomers in the asymmetric unit each containing an ADPR substrate and a tightly bound Mg^{2+} at the active site even though crystals of the complex were grown in the presence of EDTA (Figure 2). This hNUDT5-ADPR complex can be observed without hydrolysis taking place probably because one Mg^{2+} is not sufficient to sustain hydrolysis. The structure of the ΔhNUDT5-AMP complex was obtained by soaking crystals of the ΔhNUDT5-ADPR complex in the presence of Mg^{2+}. The ADPR substrate was hydrolyzed by ΔhNUDT5 to form the AMP product in crystals. The structure of the ΔhNUDT5-AMP complex contains a dimeric ΔhNUDT5, each subunit containing a bound AMP and three Mg^{2+} (Figure 3).

In general, the overall structures of the ΔhNUDT5-ADPR complex and the ΔhNUDT5-AMP complex are very similar to that of the apo form full-length enzyme except that the C-terminal nine residues are non-existent (Figures 2(a) and 3(a)). Superimposition of the apo form hNUDT5 structure with the ΔhNUDT5-ADPR and ΔhNUDT5-AMP complexes yielded an RMSD value of 0.76 Å for 390 C^α atoms and 0.73 Å for 387 C^α atoms, respectively. Conformational differences occur mainly on loops L3, L6, and L9 that are located near the substrate-binding site. Since the C-terminal extended α-helix in the full-length hNUDT5 structure is not involved in dimerization or substrate binding and ΔhNUDT5 has enzymatic activity in crystals, the structures of the ΔhNUDT5-substrate complexes should be regarded as biologically relevant. Therefore, hereafter ΔhNUDT5 will not be distinguished from full-length hNUDT5 unless otherwise specified.

Binding of substrate ADPR

As found in all available structures of bacterial ADPRases,[18–21] in the structure of the hNUDT5-

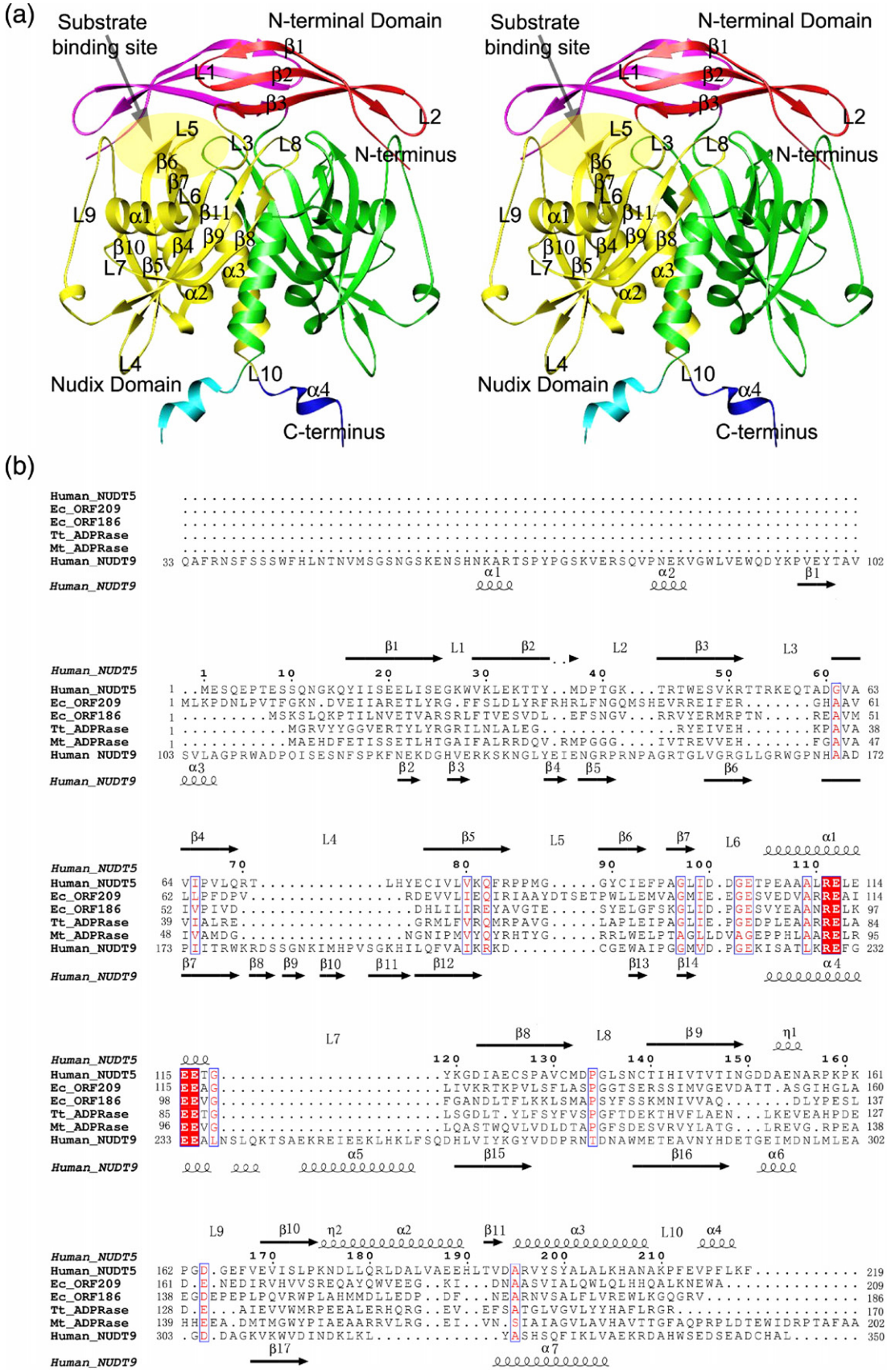

Figure 1 (*legend on opposite page*)

ADPR complex, the ADPR substrate is bound at the dimer interface between the catalytic domain of one subunit and the N-terminal domain of the other (Figure 2(a) and (b)). It assumes a horseshoe-like conformation and has extensive interactions with residues from both subunits (Figure 2(c)). The adenine N1 and N6 are specifically recognized *via* hydrogen-bonding interactions by the amide nitrogen and carbonyl oxygen of the conserved Glu47*, respectively. In addition, the adenine N7 binds directly to one side-chain amino group of Arg51 and indirectly to the other *via* a conserved water molecule (Wat1) through hydrogen bonds. Compared to the apo hNUDT5 structure, conformational changes of the side-chains of Trp28 and Trp46* are observed. In the presence of ADPR, the side-chains of both Trp28 and Trp46* move closer towards ADPR to sandwich the adenine moiety. The indole planes of the two tryptophan residues are almost in parallel to the adenine ring, forming strong π–π stacking interactions (Figure 2(c)). A similar phenomenon has been observed in other proteins such as *Caenorhabditis elegans* diadenosine tetraphosphate hydrolase.[22]

The hydroxyl groups of the adenosyl ribose moiety of ADPR form hydrogen bonds with several water molecules, but have no direct interaction with the protein. The negative charges on the pyrophosphate group of ADPR are partially neutralized by the conserved Arg84. The α-phosphate forms hydrogen bonds with the main-chain amide of Leu98 and several water molecules and the β-phosphate forms hydrogen bonds with the side-chain of Arg84 and the conserved water molecule Wat1. Interestingly, although the crystallization solution contained EDTA, an undetermined metal ion (M) was found to bind at the substrate-binding site with the number of electrons about half of magnesium. This ion is octahedrally coordinated by one α-phosphate oxygen (2.42 Å), one β-phosphate oxygen (2.51 Å), one carboxylate oxygen of Glu116 (2.52 Å), the main-chain carbonyl oxygen of Ala96 (2.32 Å), and two water molecules (2.40 Å and 2.55 Å, respectively) (Figure 2(c)). Normally when Mg^{2+} is present in crystallization solution, Mg ions are observed to bind at the active site in Nudix enzymes. The metal ion binding site in this structure is most likely the magnesium site in an active complex because its coordination is analogous to that of Mg3 in the structure of EcORF209 in complex with an ADPR analog AMPCPR (α, β-methylene-ADP ribose).[23] The terminal ribose is pointing towards the L8* loop, which contains a highly conserved residue Pro134*. Residue Asp133* flanks one side of Pro134* and forms hydrogen bonds with all three hydroxyl groups of the terminal ribose. Residue Gly135* flanks the other side and forms a hydrogen bond with the 2′-OH group of the terminal ribose through the main-chain amide group. In addition, the side-chain of Arg51 of β3 also contributes to the stabilization of the 2′-OH group of the terminal ribose through a direct hydrogen bond.

Previous biochemical data have shown that hNUDT5 has a preference for ADP-sugars over other sugar nucleotides.[15] The specific recognition and binding of the adenine moiety of ADPR with the surrounding residues of hNUDT5 *via* hydrogen-bonding interactions described above provide a structural basis for the high specificity of hNUDT5 for ADP-sugars over other sugar nucleotides. It is also intriguing to note that hNUDT5 has significantly different enzymatic activity for ADPR (100%) and IDPR (9%).[15] IDPR differs from ADPR only at the sixth position of the purine. Since the concentration of the substrate in the experiments is very high (0.3 mM), the activity difference appears to be a k_{cat} effect instead of a K_m effect.

Binding of product AMP

In the presence of Mg^{2+}, ADPR was hydrolyzed by hNUDT5 to form the AMP product in crystals of the hNUDT5-ADPR complex. As shown in Figure 3, the structure of the hNUDT5-AMP complex contains a dimeric hNUDT5, each subunit containing an AMP and three Mg^{2+} at the active site. The adenine and α-phosphate moieties of AMP retain similar positions as those in the ADPR complex and maintain similar interactions with the surrounding residues. The positions of the three Mg^{2+} are similar to those in the structure of EcORF209 in complex with AMPCPR.[23] One Mg^{2+} (Mg1) is located at a position similar to the Mg^{2+} in the ADPR complex structure. This Mg^{2+} is coordinated by one α-phosphate oxygen, the carbonyl oxygen of Ala96, the side-chain of Glu116, and two water molecules. The second Mg^{2+} (Mg2) has coordination ligands from one carboxylate oxygen each of Glu112, Glu116, and Glu166, and one α-phosphate oxygen. The third Mg^{2+} (Mg3) has coordination ligands only from the side-chain of Glu112 and one

Figure 1. Overall structure of the apo form hNUDT5. (a) Stereoview of the full-length hNUDT5 structure. The enzyme forms a homodimer. The N-terminal domains are colored in red and magenta, the Nudix catalytic domains in yellow and green, and the C-terminal nine residues, which form an extended α-helix in blue and cyan, respectively. For clarity, the secondary structure elements are labeled only in one subunit. The substrate-binding site in one subunit is indicated by an ellipsoid in light yellow. (b) Structure-based sequence alignment of different ADPRases. Sequences included are: human_NUDT5 (*Homo sapiens* NUDT5, AF218818), Ec_ORF209 (*E. coli* ADPRase encoded by *Orf209*, P36651), Ec_ORF186 (*E. coli* ADPXase encoded by *Orf186*, P45799), Tt_ADPRase (*T. thermophilus* ADPRase, AB107660), Mt_ADPRase (*M. tuberculosis* ADPRase, O33199), and human_NUDT9 (*H. sapiens* NUDT9, AAK07671). Regions of sequence similarity are denoted with open red boxes and regions of sequence identity with filled red boxes. Secondary structural elements of hNUDT5 are placed on top of the alignment, and those of hNUDT9 at the bottom. The alignment was prepared with ESPript.[34]

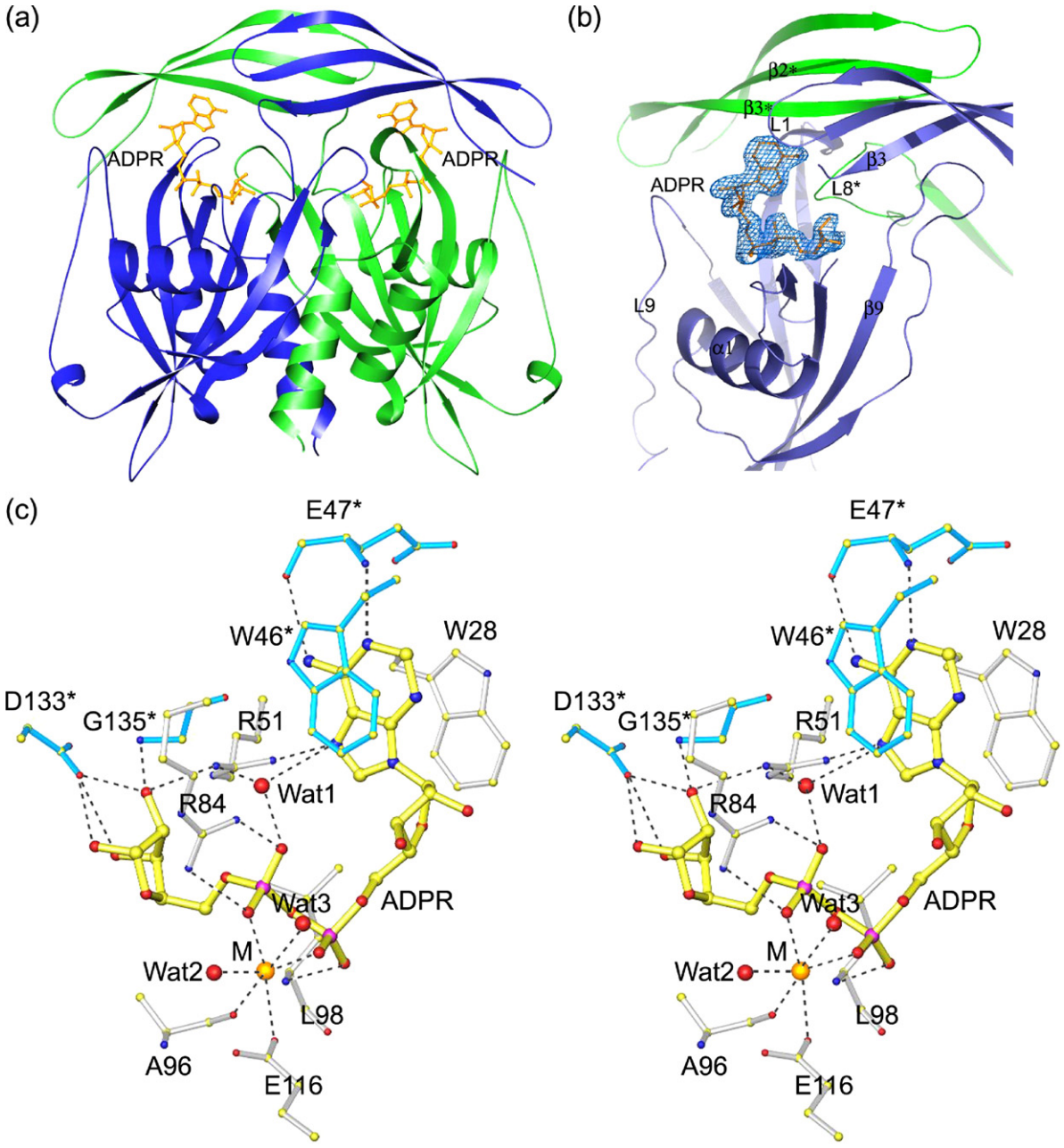

Figure 2. Structure of the ΔhNUDT5-ADPR complex. (a) Overall structure of the ΔhNUDT5-ADPR complex. The substrate-binding pocket is formed by the N-terminal domain of one subunit and the Nudix domain of the other. The bound ADPR substrates are shown with gold ball- and -stick models. (b) Structure of the ADPR-binding pocket. The electron density ($2F_o$–F_c map) for the bound ADPR is shown in slate-blue. The surrounding secondary structural elements from both subunits are shown in and labeled with different colors. (c) Stereoview showing the interactions of ADPR with the surrounding residues. The ADPR is shown as a ball- and -stick model in yellow. The residues are shown with side-chains: these from the first subunit are colored in gray and those from the second subunit in sky blue and indicated with an asterisk. Carbon is shown in yellow, oxygen in red, nitrogen in blue, and phosphorus in magenta. An undetermined metal ion (M) is found to bind at the active site and is shown with a gold sphere. The hydrogen-bonding interactions are indicated with broken lines. The side-chains of W28 and W46* sandwich the adenine moiety of ADPR.

α-phosphate oxygen. The incomplete coordination geometry of these Mg ions might be due to the moderate resolution of the complex structure. These results show that all three metals contribute to bridging the α-phosphate of AMP to the catalytic active site mainly through residues Glu112 and Glu116, which are part of the Nudix signature sequence. Based on these results, we can infer that hNUDT5 may need three metal ions for catalysis and the metal ions may participate in binding of the pyrophosphate of ADPR to the enzyme during the catalytic reaction.

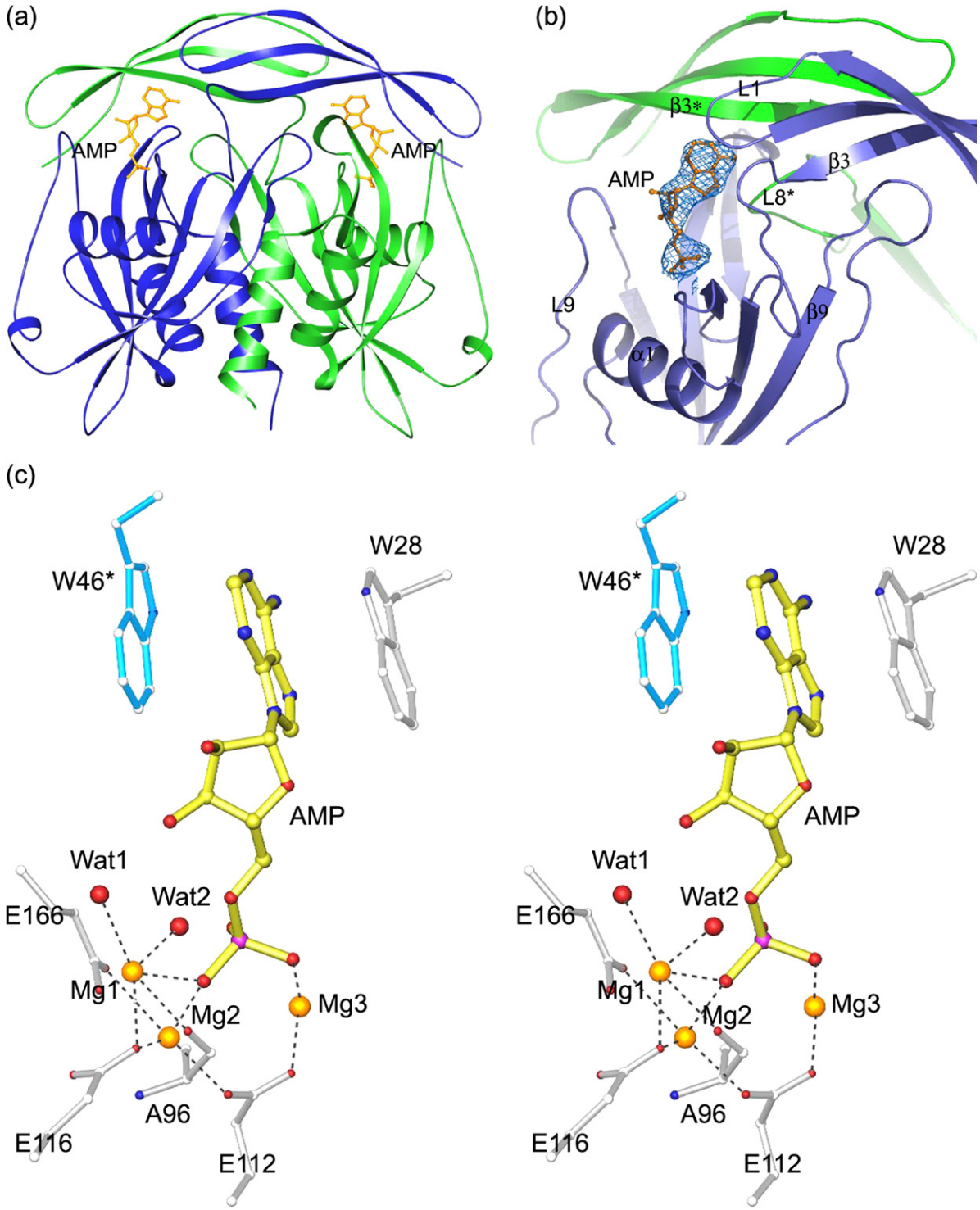

Figure 3. Structure of the ΔhNUDT5-AMP complex. (a) Overall structure of the ΔhNUDT5-AMP complex. The bound AMP substrates are shown with yellow ball- and -stick models. (b) Structure of the AMP-binding pocket. The electron density ($2F_o$–F_c map) for the bound AMP is shown in slate-blue. The surrounding secondary structural elements from both subunits are shown and labeled with different colors. (c) Stereoview showing the interactions of AMP and the Mg ions with the surrounding residues. AMP is shown with a ball- and -stick model, the Mg ions are shown with gold spheres, and the surrounding residues are shown with side-chains. The hydrogen-bonding interactions are indicated with broken lines.

Comparison of hNUDT5 with EcORF209 and EcORF186

Although hNUDT5 shares very low sequence identity with other ADPRases or the ADPXase EcORF186 (16–19% sequence identity) (Figure 1(b)), the overall structure of hNUDT5 is quite similar to that of the four prokaryotic enzymes, but substantially different from that of hNUDT9. On the other hand, there are several notable structural

differences between hNUDT5 and the bacterial proteins, which may account for the versatility of substrate specificity of the ADPRase subfamily. Among the bacterial enzymes, there is not much information about the substrate specificity of *T. thermophilus* ADPRase or *M. tuberculosis* ADPRase, thus, comparison of hNUDT5 was performed primarily with the *E. coli* ADPXase EcORF186 and ADPRase EcORF209 for clues for substrate selectivity. As shown in Figure 4(a), the structures of the Nudix motif and other important secondary elements are quite similar between hNUDT5 and the two *E. coli* enzymes and the major differences reside on the last two α-helices α2 and α3. Helix α2

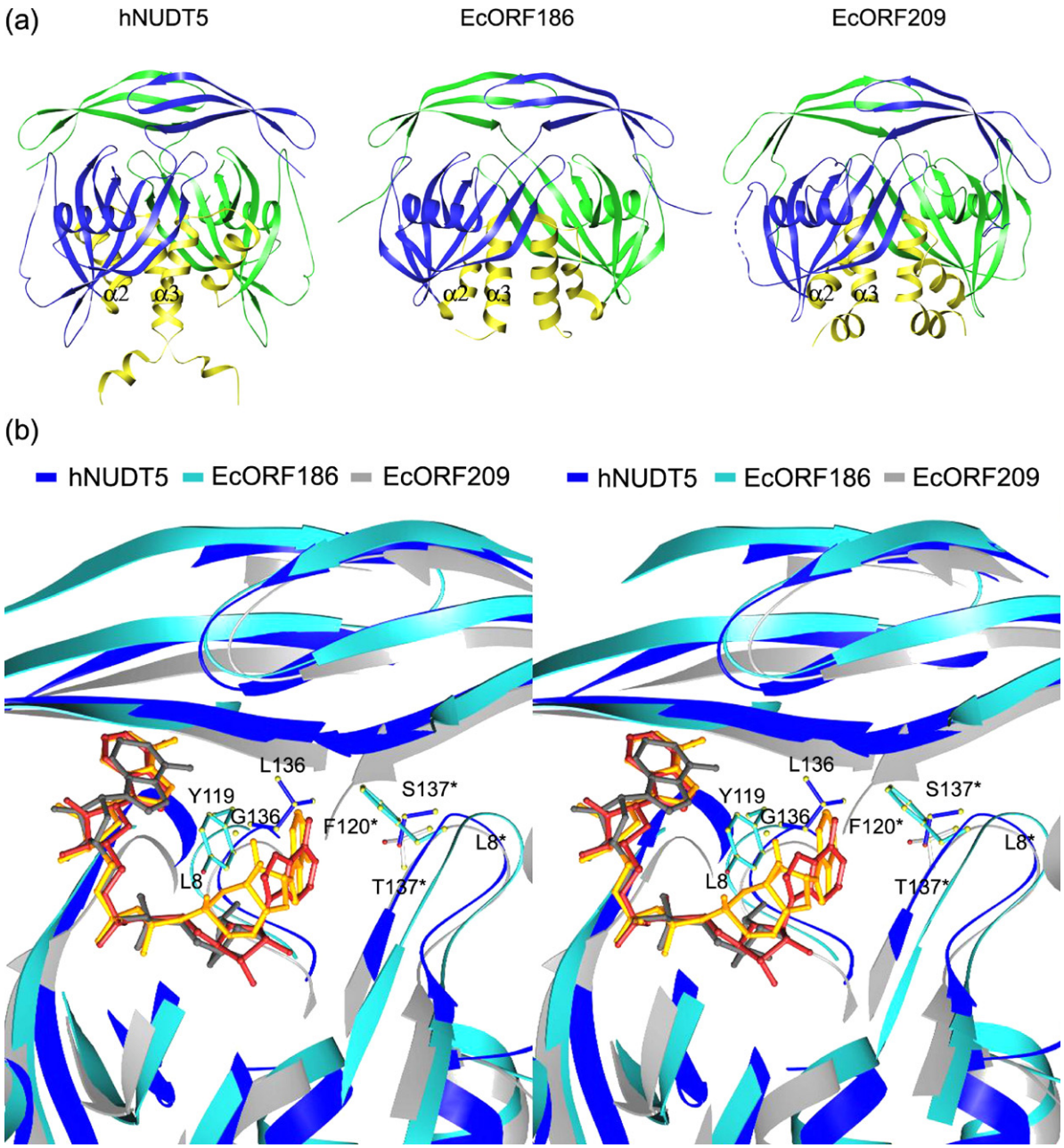

Figure 4. Structural comparison of hNUDT5 with EcORF209 and EcORF186. (a) Overall structures of hNUDT5, EcORF209, and EcORF186. The two subunits of the homodimer are shown in blue and green, respectively. The major differences between the three enzymes reside in the last two α-helices α2 and α3 (in yellow). (b) Superposition of the substrate-binding pocket showing the structural elements and residues involved in the recognition of the terminal moieties of substrates. Structures of the hNUDT5-ADPR (in blue), EcORF209 (in gray), and EcORF186 (in cyan) were superimposed based on alignment of the catalytic domain. ADPR is shown as a ball- and -stick model in gray. AP_2A (in red) and AP_3A (in gold) are docked into the catalytic active site based on the structure of the hNUDT5-ADPR complex. Residues Tyr119 and Phe120* in EcORF186, which are suggested to be involved in the recognition of the terminal adenine moiety of AP_2A and AP_3A, are shown with side-chains in cyan. The equivalent residues in hNUDT5 (Leu136 and Ser137*) and in EcORF209 (Gly136 and Thr137*) are shown in blue and gray, respectively.

is pointing outward and helix α3 protrudes more into the other subunit in the hNUDT5 structures. The protrusion of helix α3 may be caused by extensive hydrogen-bonding interactions between residues of α3 and those of the surrounding β strands, leading to tighter dimeric interactions. The other appreciable difference occurs in the substrate-binding pocket. The substrate-binding pockets in hNUDT5 and EcORF209 are smaller than that in EcORF186. The secondary structural elements β3 and L3 in hNUDT5 and EcORF209 are positioned closer to the substrate-binding pocket compared to those in EcORF186. In addition, loop L3 of hNUDT5 has a short insertion and therefore is longer than its counterparts in the bacterial enzymes, and loop L5 of EcORF209 also has an insertion compared to hNUDT5 and EcORF186. Nevertheless, loops L3 and L5 are not directly involved in the substrate binding. Since EcORF186 has much greater substrate versatility as compared to the other two enzymes, it is possible that these structural differences might play some roles in the substrate selectivity.

For the ADPR pyrophosphatase hydrolase subfamily, it is shown that a highly conserved Pro residue (corresponding to Pro134 of hNUDT5, Pro134 of EcORF209, or Pro117 of EcORF186) is located 15 or 16 amino acid residues downstream of the Nudix motif. In the structures of hNUDT5 and the two *E. coli* enzymes, this conserved Pro residue resides on the tip of loop L8, which flanks the terminal ribose of ADPR (Figure 4(b)). Since loop L8 is involved in interlocking the two subunits and in recognition of the terminal ribose, the presence of a rigid proline residue at this position may be important to maintain the right length and position of L8. Similarly, a conserved Tyr residue is found to locate at 16–18 residues downstream of the Nudix motif in members of the diadenine pyrophosphate hydrolase subfamily, but not in members of the other subfamilies.[24] Among the three proteins, only EcORF186 contains a Tyr residue at the equivalent position (Tyr119), consistent with its enzymatic activities on NADH, AP_2A, and AP_3A.[10] On the other hand, hNUDT5 and EcORF209, which lack an equivalent aromatic residue, have low activity on AP_2A and no activity on AP_3A.[14,15,24] To explore the underlying molecular basis of the differing AP_nA hydrolase activities of EcORF186, hNUDT5, and EcORF209, we docked AP_2A and AP_3A molecules into the catalytic active sites of the three structures based on the position of ADPR in the NUDT5-ADPR complex. As discussed above, the terminal ribose moiety of ADPR is pointing towards loop L8* of the second subunit. In the resulting models, the terminal adenine moieties of AP_2A and AP_3A are localized in the pocket formed by the two L8 loops of the two subunits in the dimeric interface (Figure 4(b)). However, the substrate-binding pocket in hNUDT5 and EcORF209 is smaller than that in EcORF186, and particularly the width of the pocket between the two L8 loops in hNUDT5 and EcORF209 is much smaller than that in EcORF186: the C^α–C^α distance between Leu136 of loop L8 and Cys139 of loop L8* is 9.5 Å; the equivalent distance between Gly136 and Glu139 in EcORF209 is 9.8 Å; and the equivalent distance between Tyr119 and Ser122 in EcORF186 is 13.8 Å, respectively (Figure 4(b)). In both hNUDT5 and EcORF209, the terminal adenine moiety of the docked AP_2A molecule can be accommodated into the narrow pocket between the two L8 loops with minor conformational changes of the side-chains of several surrounding residues (mainly Arg51, Asp133*, and Leu136 of hNUDT5, and Arg56, Glu139, and Ser133* of EcORF209, respectively); however, the terminal adenine moiety of the docked AP_3A molecule cannot be fit into the pocket because the terminal adenine of AP_3A would cause serious steric conflicts with the aforementioned surrounding residues. On the other hand, in EcORF186 the pocket between the two L8 loops is large enough to accommodate the terminal adenine moiety of both AP_2A and AP_3A, and Tyr119 of EcORF186 appears to be in a position to contact with the terminal adenine moiety of AP_2A and AP_3A through π–π stacking interaction (Figure 4(b)). In addition, there is a Phe residue (Phe120 of EcORF186) following this Tyr residue, making it possible for a sandwich-like stacking interaction among Tyr119, the terminal adenine ring, and Phe120*. There seems to be a positive correlation between the AP_nA hydrolase activity of EcORF186 and the presence of an aromatic residue in the L8 loop. Nevertheless, unlike other diadenine pyrophosphate hydrolases that are either monomeric[22,25] or dimeric but with little domain swapping,[26] EcORF186 has no activity on any diadenine pyrophosphate longer than AP_3A, probably due to geometrical constraint of the substrate-binding pocket. Thus, our results support the notion that residues Pro and Tyr of loop L8 are important for the substrate selectivity[24] and the oligomeric state may also play a role in the selectivity.

Comparison of hNUDT5 with hNUDT9

The two human ADPRases hNUDT5 and hNUDT9 display great difference in substrate selectivity: hNUDT9 is highly specific for ADPR and IDPR, whereas hNUDT5 takes a variety of ADP-sugar conjugates as substrates. Structural comparison of the hNUDT5-ADPR complex and the hNUDT9-R5P complex reveals that although both enzymes contain a similar Nudix domain, there are substantial conformational differences in the substrate-binding region (RMSD of 1.87 Å for 89 C^α atoms, Figure 5(a) and (b)). In the hNUDT5-ADPR structure, the adenine moiety is recognized by the enzyme through hydrogen-bonding interactions with several surrounding residues and has π–π stacking interactions with two Trp residues. The residues involved in the interactions of ADPR are not conserved in hNUDT9 (Figure 1(b)). For hNUDT9, ADPR is suggested to bind in

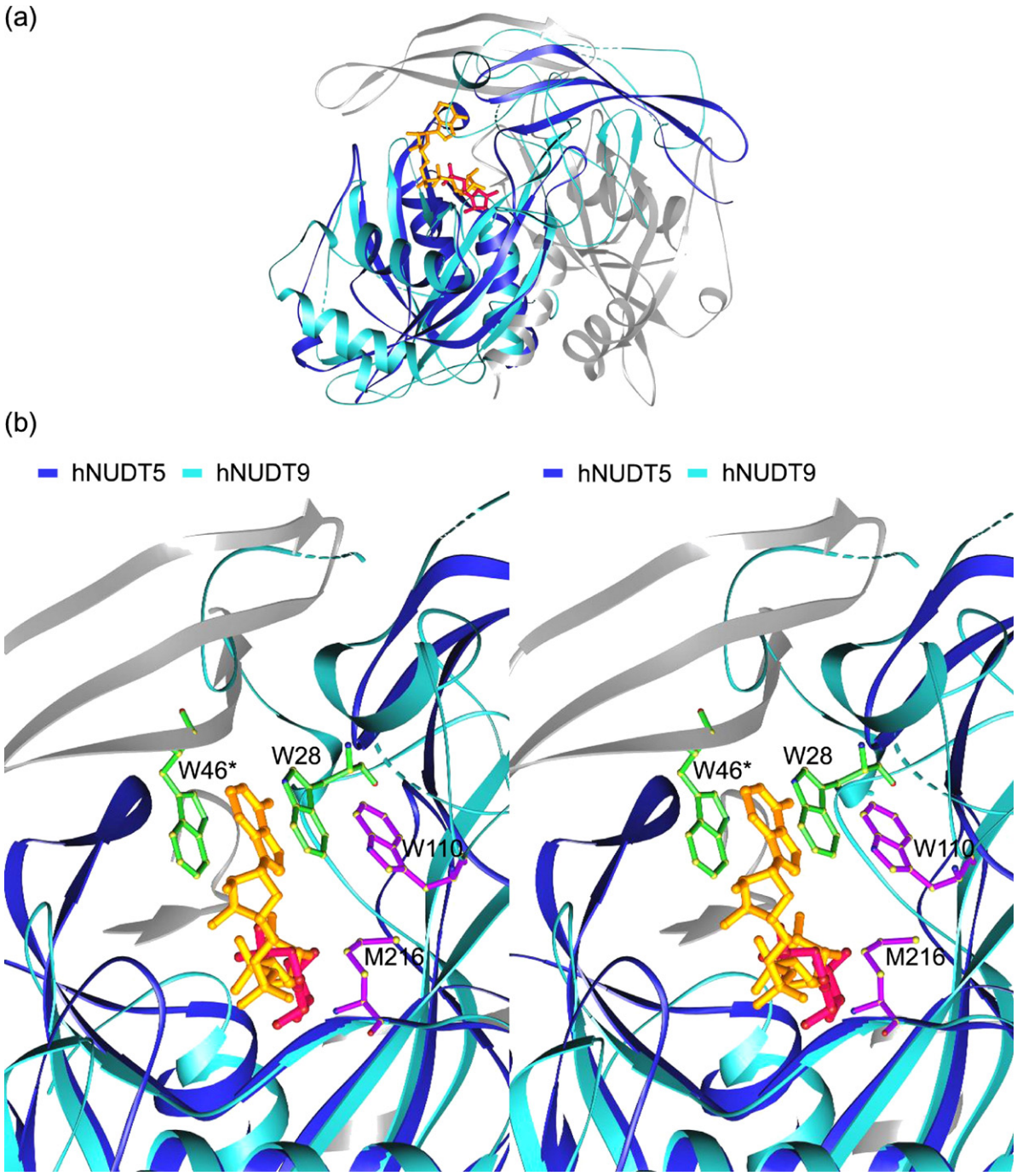

Figure 5. Structural comparison of hNUDT5 with hNUDT9. (a) Superposition of structures of the hNUDT5-ADPR complex (blue for the first subunit and gray for the second) and the hNUDT9-R5P complex (sky blue). The structures were superimposed based on alignment of the catalytic domain. The bound ADPR (in gold) and R5P (in magenta) are shown with ball- and -stick models. (b) Structural comparison in the substrate-binding pocket. The color-coding of the structures is the same as in (a). Residues involved in the recognition of the adenine moiety of ADPR in the hNUDT5-ADPR complex and the hNUDT9-R5P complex are shown with side-chains in green and purple, respectively.

a hydrophobic pocket formed by Trp119 and Met216.[16] These results suggest that the two enzymes appear to have different ADPR binding sites and use different residues in the substrate recognition, and the interactions of the adenine moiety of ADPR with hNUDT5 are much tighter than those with hNUDT9, which might explain the high specificity of hNUDT5 for a broad range of ADP-sugar substrates. The tight binding of the adenine moiety of the substrate with hNUDT5 is also implicated by the result that the AMP product remains bound at the substrate-binding site after ADPR being hydrolyzed by the enzyme in crystals.

Materials and Methods

Cloning, expression, and purification of hNUDT5

The cDNA of the full-length hNUDT5 was amplified by PCR from the cDNA library of human CD34+ haematopoietic stem/progenitor cells and inserted into the NdeI and XhoI restriction sites of the pET-22b(+) expression plasmid (Novagen). The plasmid was transformed into *E. coli* BL21 (DE3) strain (Novagen). The transformed cells were grown in LB media at 37 °C in the presence of 50 μg/ml ampicillin until $A_{600\ nm}$ reached 0.8, and then induced with 0.5 mM IPTG for 12 h at 16 °C. The cells were harvested and lysed by sonication in a lysis buffer (20 mM Tris-HCl (pH 7.0), 300 mM NaCl). Protein purification was carried out by affinity chromatography using a Ni-NTA column (Amersham-Pharmacia) with the lysis buffer supplemented with 20 mM imidazole and 200 mM imidazole serving as washing buffer and elution buffer, respectively. The elution fractions were further purified by Q-Sepharose (Pharmacia LKB) column chromatography and fractions containing the target protein were eluted with a buffer of 250 mM NaCl. After dialysis against a buffer of 20 mM Tris-HCl (pH 7.0), the target protein was concentrated to about 20 mg/ml with Centricon-10 (Millipore) for structural studies. SDS-PAGE analysis of the protein sample showed a single band. The purity and homogeneity of the protein was also confirmed by dynamic light scattering (DLS) analysis.

To obtain crystals of the enzyme in complex with substrate, we constructed a truncated form of *hNUDT5* (residues 1–210, ΔhNUDT5) that omitted the C-terminal nine hypervariable residues to facilitate crystallization. The gene fragment was cloned into the pGEX4T-1 expression plasmid (Novagen), which attaches a glutathione-*S*-transferase (GST) tag at the N terminus of the protein. The plasmid was transformed into *E. coli* DH5α strain (Novagen). Subsequent protein expression was similar to the full-length enzyme as described above. For protein purification, affinity chromatography was carried out on a GSTrap FF 5 ml column (Amersham-Pharmacia). Cleavage of this fusion protein with thrombin protease on the column was conducted at 16 °C for 24 h. Fractions containing the ΔhNUDT5 protein were then loaded onto a Q-Sepharose column (Pharmacia LKB) for further purification and the target protein was eluted with the buffer of 250 mM NaCl. The elution fractions were further purified by gel filtration.

For the preparation of selenomethionine (Se-Met) substituted protein, the expression vector containing Δ*hNUDT5* was transformed into the methionine auxotroph *E. coli* B834 strain. The cells were grown in M9 medium supplemented with 1 mg/l each of riboflavin, niacinamide, pyridoxine mono hydrochloride, and thiamine, 25 mg/l $FeSO_4{\cdot}7H_2O$, 0.4% (w/v) glucose, 2 mM $MgSO_4$, 40 mg/l each of 19 L-amino acids (excluding methionine), and 40 mg/l of L-selenomethionine (Sigma). Purification of the Se-Met ΔhNUDT5 was carried out using the same methods as for the native protein.

Crystallization and diffraction data collection

Crystals of full-length hNUDT5 were grown at 4 °C using the hanging-drop vapor diffusion method. The crystallization solution consists of equal volumes of the protein solution (10 mg/ml) and the reservoir solution containing 160 mM NaAc (pH 5.5), and 25% 2-methyl-2, 4-pentanediol. A cluster of polymorphous crystals grew after a week. Through microseeding we obtained single crystals with a maximum size of 0.2 mm × 0.1 mm × 0.1 mm in three days. Diffraction data were collected to 2.5 Å resolution from a flash-cooled crystal at 100 K using synchrotron radiation at Beijing Synchrotron Radiation Facility. Crystals of Se-Met ΔhNUDT5 were grown from a crystallization solution consisting of equal volumes of the protein solution (10 mg/ml) and the reservoir solution containing 300 mM di-ammonium hydrogen citrate, 6% *n*-propanol, and 15% (w/v) polyethylene glycol 3350. Rod-shaped single crystals grew after four days.MAD diffraction data were collected from a flash-cooled crystal at 100 K using synchrotron radiation at Beijing Synchrotron Radiation Facility. Crystals of the ΔhNUDT5-ADPR complex were grown at 4 °C from a drop consisting of equal volumes of the protein solution containing 10 mg/ml ΔhNUDT5, 20 mM Tris-HCl (pH 7.0), 5 mM EDTA, and 2 mM ADPR, and the reservoir solution containing 200 mM NaAc, 100 mM Tris-HCl (pH 8.0) and 30% polyethylene glycol 4000. Single crystals of the plate shape morphology grew after three days. Diffraction data were collected to 2.0 Å resolution from a flash-cooled crystal at 100 K at beamline BL44B2 of SPring-8, Japan. Crystals of the ΔhNUDT5-AMP complex were obtained at 4 °C by soaking crystals of the ΔhNUDT5-ADPR complex in a mother liquor containing 10 mM $MgCl_2$ for 24 h. Diffraction data were collected to 2.6 Å resolution from a flash-cooled crystal at beamline BL-6A of Photon Factory, Japan. All the diffraction data were processed using the *HKL*2000 program suite.[27] Statistics for data collection and refinement are summarized in Table 1.

Structure determination and refinement

The structure of ΔhNUDT5 was solved using the MAD method implemented in the program SOLVE.[28] The automatic protocol in SOLVE revealed six Se sites in the asymmetric unit. Since ΔhNUDT5 contains four methionine residues (including the first methionine), the SOLVE result suggests that there are two ΔhNUDT5 molecules in the asymmetric unit and the first Met residue of each molecule is likely disordered. The MAD phases were improved by statistical density modification including solvent flattening and histogram matching using the program RESOLVE.[29] RESOLVE automatically built 250 polyalanine residues out of 424 residues and successfully located most of the secondary structural elements. The complete model was manually built using the program O.[30] This model was used as the search model to determine the phases for the 2.5 Å resolution full-length native dataset by molecular replacement (MR) implemented in the program CNS.[31] The structures of ΔhNUDT5 in complexes with ADPR and AMP were solved by MR using the full-length hNUTD5 structure as the search model. There was well-defined electron density for the bound ADPR, AMP, and the metal ions at the active site. However, the bound AMP in the hNUDT5-AMP complex has relatively high *B* factor, suggesting that the ligand might be partially disordered probably because the crystals of the AMP complex were prepared by soaking the crystals of the ADPR complex in the presence of Mg^{2+} and the hydrolysis of ADPR to AMP took place in the crystals. Structure refinement was performed with CNS using standard protocols (energy minimization, simulated annealing, and *B* factor refinement) and the model building was facilitated with program O. The final structure refinement was carried out with the maximum likelihood algorithm

implemented in the program REFMAC5.[32] A bulk solvent correction and a free *R* factor monitor (calculated with 5% of randomly chosen reflections) were applied throughout the refinement and model building. The stereochemical quality of the structure models during the course of refinement and model building was evaluated with the program PROCHECK.[33] Summary of the structure refinement statistics is given in Table 1.

Protein Data Bank accession codes

The structures of the apo form full-length hNUDT5, the ΔhNUDT5-ADPR complex, and the ΔhNUDT5-AMP complex have been deposited in the RCSB Protein Data Bank with accession codes 2DSB, 2DSC, and 2DSD, respectively.

Acknowledgements

We thank the staff members at Beijing Synchrotron Radiation Facility, and Photon Factory and SPring-8, Japan for support in diffraction data collection, and Qiuhua Huang of Shanghai Institute of Hematology, Rui-Jin Hospital for providing the human NUDT5 plasmid. This work was supported by NSFC grants (30125011 and 30570379), MOST grants (2002BA711A13 and 2004CB720102), and CAS grant (KSCX1-SW-17).

References

1. Bessman, M. J., Frick, D. N. & O'Handley, S. F. (1996). The MutT proteins or "Nudix" hydrolases, a family of versatile, widely distributed, "housecleaning" enzymes. *J. Biol. Chem.* **271**, 25059–25062.
2. Olivera, B. M., Hughes, K. T., Cordray, P. & Roth, J. R. (1989). Aspects of NAD metabolism in prokaryotes and eukaryotes. In *ADP-ribose Transfer Reactions, Mechanisms and Biological Significance* (Jacobson, M. K. & Jacobson, E. L., eds), pp. 353–360, Springer Verlag, New York.
3. Jacobson, E. L., Cervantes-Laurean, D. & Jacobson, M. K. (1994). Glycation of proteins by ADP-ribose. *Mol. Cell. Biochem.* **138**, 207–212.
4. Weigt, C., Just, I., Wegner, A. & Aktories, K. (1989). Nonmuscle actin ADP-ribosylated by botulinum C2 toxin caps actin filaments. *FEBS Letters*, **246**, 181–184.
5. Kwak, Y. G., Park, S. K., Kim, U. H., Han, M. K., Eun, J. S., Cho, K. P. & Chae, S. W. (1996). Intracellular ADP-ribose inhibits ATP-sensitive K+ channels in rat ventricular myocytes. *Am. J. Physiol.* **271**, C464–C468.
6. Ueda, K. & Hayaishi, O. (1985). ADP-ribosylation. *Annu. Rev. Biochem.* **54**, 73–100.
7. McDonald, L. J. & Moss, J. (1994). Enzymatic and nonenzymatic ADP-ribosylation of cysteine. *Mol. Cell. Biochem.* **138**, 221–226.
8. Honjo, T., Nishizuka, Y. & Hayaishi, O. (1968). Diphtheria toxin-dependent adenosine diphosphate ribosylation of aminoacyl transferase II and inhibition of protein synthesis. *J. Biol. Chem.* **243**, 3553–3555.
9. Sheikh, S., O'Handley, S. F., Dunn, C. A. & Bessman, M. J. (1998). Identification and characterization of the Nudix hydrolase from the Archaeon, *Methanococcus jannaschii*, as a highly specific ADP-ribose pyrophosphatase. *J. Biol. Chem.* **273**, 20924–20928.
10. O'Handley, S. F., Frick, D. N., Dunn, C. A. & Bessman, M. J. (1998). Orf186 represents a new member of the Nudix hydrolases, active on adenosine(5′)triphospho (5′)adenosine, ADP-ribose, and NADH. *J. Biol. Chem.* **273**, 3192–3197.
11. Kim, J. S., Kim, W. Y., Rho, H. W., Park, J. W., Park, B. H., Han, M. K. *et al.* (1998). Purification and characterization of adenosine diphosphate ribose pyrophosphatase from human erythrocytes. *Int. J. Biochem. Cell. Biol.* **30**, 629–638.
12. Zocchi, E., Guida, L., Franco, L., Silvestro, L., Guerrini, M., Benatti, U. & De Flora, A. (1993). Free ADP-ribose in human erythrocytes: pathways of intra-erythrocytic conversion and non-enzymic binding to membrane proteins. *Biochem. J.* **295**, 121–130.
13. Ribeiro, J. M., Carloto, A., Costas, M. J. & Cameselle, J. C. (2001). Human placenta hydrolases active on free ADP-ribose: an ADP-sugar pyrophosphatase and a specific ADP-ribose pyrophosphatase. *Biochim. Biophys. Acta*, **1526**, 86–94.
14. Yang, H., Slupska, M. M., Wei, Y. F., Tai, J. H., Luther, W. M., Xia, Y. R. *et al.* (2000). Cloning and characterization of a new member of the Nudix hydrolases from human and mouse. *J. Biol. Chem.* **275**, 8844–8853.
15. Gasmi, L., Cartwright, J. L. & McLennan, A. G. (1999). Cloning, expression and characterization of YSA1H, a human adenosine 5′-diphosphosugar pyrophosphatase possessing a MutT motif. *Biochem. J.* **344**, 331–337.
16. Lin, S., Gasmi, L., Xie, Y., Ying, K., Gu, S., Wang, Z. *et al.* (2002). Cloning, expression and characterisation of a human Nudix hydrolase specific for adenosine 5′-diphosphoribose (ADP-ribose). *Biochim. Biophys. Acta*, **1594**, 127–135.
17. Shen, B. W., Perraud, A. L., Scharenberg, A. & Stoddard, B. L. (2003). The crystal structure and mutational analysis of human NUDT9. *J. Mol. Biol.* **332**, 385–398.
18. Badger, J., Sauder, J. M., Adams, J. M., Antonysamy, S., Bain, K., Bergseid, M. G. *et al.* (2005). Structural analysis of a set of proteins resulting from a bacterial genomics project. *Proteins: Struct. Funct. Genet.* **60**, 787–796.
19. Gabelli, S. B., Bianchet, M. A., Bessman, M. J. & Amzel, L. M. (2001). The structure of ADP-ribose pyrophosphatase reveals the structural basis for the versatility of the Nudix family. *Nature Struct. Biol.* **8**, 467–472.
20. Kang, L. W., Gabelli, S. B., Cunningham, J. E., O'Handley, S. F. & Amzel, L. M. (2003). Structure and mechanism of MT-ADPRase, a nudix hydrolase from Mycobacterium tuberculosis. *Structure*, **11**, 1015–1123.
21. Yoshiba, S., Ooga, T., Nakagawa, N., Shibata, T., Inoue, Y., Yokoyama, S. *et al.* (2004). Structural insights into the *Thermus thermophilus* ADP-ribose pyrophosphatase mechanism via crystal structures with the bound substrate and metal. *J. Biol. Chem.* **279**, 37163–37174.
22. Bailey, S., Sedelnikova, S. E., Blackburn, G. M., Abdelghany, H. M., Baker, P. J., McLennan, A. G. & Rafferty, J. B. (2002). The crystal structure of diadenosine tetraphosphate hydrolase from *Caenorhabditis elegans* in free and binary complex forms. *Structure*, **10**, 589–600.
23. Gabelli, S. B., Bianchet, M. A., Ohnishi, Y., Ichikawa, Y., Bessman, M. J. & Amzel, L. M. (2002). Mechanism

of the *Escherichia coli* ADP-ribose pyrophosphatase, a Nudix hydrolase. *Biochemistry*, **41**, 9279–9285.
24. Dunn, C. A., O'Handley, S. F., Frick, D. N. & Bessman, M. J. (1999). Studies on the ADP-ribose pyrophosphatase subfamily of the nudix hydrolases and tentative identification of trgB, a gene associated with tellurite resistance. *J. Biol. Chem.* **274**, 32318–32324.
25. Swarbrick, J. D., Bashtannyk, T., Maksel, D., Zhang, X. R., Blackburn, G. M., Gayler, K. R. & Gooley, P. R. (2000). The three-dimensional structure of the Nudix enzyme diadenosine tetraphosphate hydrolase from *Lupinus angustifolius* L. *J. Mol. Biol.* **302**, 1165–1177.
26. Ranatunga, W., Hill, E. E., Mooster, J. L., Holbrook, E. L., Schulze-Gahmen, U., Xu, W. *et al.* (2004). Structural studies of the Nudix hydrolase DR1025 from *Deinococcus radiodurans* and its ligand complexes. *J. Mol. Biol.* **339**, 103–116.
27. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
28. Terwilliger, T. C. & Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallog. sect. D*, **55**, 849–861.
29. Terwilliger, T. C. (2001). Maximum-likelihood density modification using pattern recognition of structural motifs. *Acta Crystallog. sect. D*, **57**, 1755–1762.
30. Jones, T. A., Zou, J. Y. & Cowan, S. W. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A*, **47**, 110–119.
31. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
32. 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.
33. 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.
34. Gouet, P., Courcelle, E., Stuart, D. I. & Metoz, F. (1999). ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics*, **15**, 305–308.

Edited by I. Wilson

(Received 30 June 2006; received in revised form 27 September 2006; accepted 27 September 2006)
Available online 3 October 2006