Purification, crystallization and preliminary X-ray analysis of two nudix hydrolases from *Deinococcus radiodurans*

Elizabeth Litzenberger Holbrook,[a,b] Ursula Schulze-Gahmen,[a] Garry W. Buchko,[c] Shuisong Ni,[c] Michael A. Kennedy[c] and Stephen R. Holbrook[a]*

[a]Structural Biology Department, Physical Biosciences Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA, [b]Chemistry Department, University of California at Berkeley, Berkeley, CA 94720, USA, and [c]Fundamental Sciences, Biology Sciences Division, Pacific Northwest National Laboratory, Richland, WA 99352, USA

Correspondence e-mail: srholbrook@lbl.gov

Two nudix hydrolases from *Deinococcus radiodurans* have been purified and crystallized. Diffraction data have been collected to 1.4 and 1.9 Å resolution for DR1025 and DR0079, respectively. DR1025 belongs to space group $P4_12_12/P4_32_12$, with unit-cell parameters $a = b = 53.2$, $c = 122.6$ Å (unit-cell volume 346 883 Å^3, $V_M = 2.5$ Å^3 Da^{-1}, solvent content 50.2%). DR0079 belongs to space group $C222_1$, with unit-cell parameters $a = 34.1$, $b = 157.2$, $c = 126.5$ Å (unit-cell volume 677 308 Å^3, $V_M = 2.2$ Å^3 Da^{-1}, solvent content 44.0%). The calculated cell content of DR1025 indicates the presence of one molecule in the asymmetric unit. Dynamic light scattering and gel filtration suggest it to be a dimer in solution. The space group and unit-cell parameters of DR0079 indicate the presence of two molecules per asymmetric unit. Gel filtration and NMR spectroscopy suggest it to be a monomer in solution.

1. Introduction

The nudix hydrolases represent a class of proteins identified in many species including archaea, bacteria and eukarya (Xu *et al.*, 2001) that are known to catabolyse substrates that are *nu*cleoside *di*phosphates linked to a moiety *X* (Bessman *et al.*, 1996). They can be sub-divided functionally according to their preference for dinucleoside polyphosphates, ADP-ribose, nucleotide sugars, NADH or various (deoxy)ribonucleoside triphosphates (Dunn *et al.*, 1999). Nudix hydrolases are characterized by the presence of the highly conserved sequence motif (the nudix box) $GX_5EX_7REUXEEXGU$ (where *U* is usually I, V or L and *X* can be any amino acid) that contributes to the catalytic and substrate-binding sites of the enzyme (Abeyguna-wardana *et al.*, 1995). The prototypical member of this family is the *Escherichia coli* MutT protein, which catalyzes the hydrolysis of nucleotide triphosphates with a preference for 8-oxo-dGTP, implying a role in DNA maintenance and repair.

Structures of functionally distinct nudix hydrolases from a variety of species have been solved (Abeygunawardana *et al.*, 1995; Lin *et al.*, 1997; Swarbrick *et al.*, 2000; Bonanno *et al.*, 2001; Durbecq *et al.*, 2001; Gabelli *et al.*, 2001; Bailey *et al.*, 2002; Fletcher *et al.*, 2002; Wang *et al.*, 2002). Some of these structures are monomers and some are dimers. All of them conserve a structural core of a central, mainly antiparallel, β-sheet and two helices (of which one is part of the loop–helix–loop containing the nudix box). The structural similarity of the core is extensive despite very low

sequence identities; alignment of Ap$_4$A (diadenosine 5′,5′′′-P^1,P^4-tetraphosphate) hydrolase sequences from animal and plant species reveals that even they have very low sequence identities (Swarbrick *et al.*, 2000; Bailey *et al.*, 2002; Fletcher *et al.*, 2002).

We are investigating the extremely radio-resistant bacterium *Deinococcus radiodurans* (Mattimore & Battista, 1996), which has been shown to encode at least 23 hypothetical nudix hydrolases, the largest number of any organism, which have diverse enzymatic activities and sizes (Makarova *et al.*, 2000; Xu *et al.*, 2001). As an initial step toward a structure–function analysis of the *D. radiodurans* nudix family, we report the expression, purification, crystallization and diffraction data collection for two members, DR1025 and DR0079. A multiple sequence alignment is shown in Fig. 1, comparing the *E. coli* MutT protein with DR1025 and DR0079 and illustrating the high sequence diversity outside the nudix box, even though a larger structural core is expected to be conserved for substrate binding and catalysis.

2. Materials and methods

2.1. Cloning, expression and purification

2.1.1. DR1025. The strain expressing DR1025 (pET24a expression vector in the Tuner cell background; Novagen) was obtained from Dr Maurice Bessman (Xu *et al.*, 2001). Cells were grown in Luria broth at 310 K with 30 µg ml^{-1} kanamycin to an OD$_{600}$ of ~0.8. Protein expression was induced by addition of isopropyl-β-D-galactopyranoside (IPTG) to a
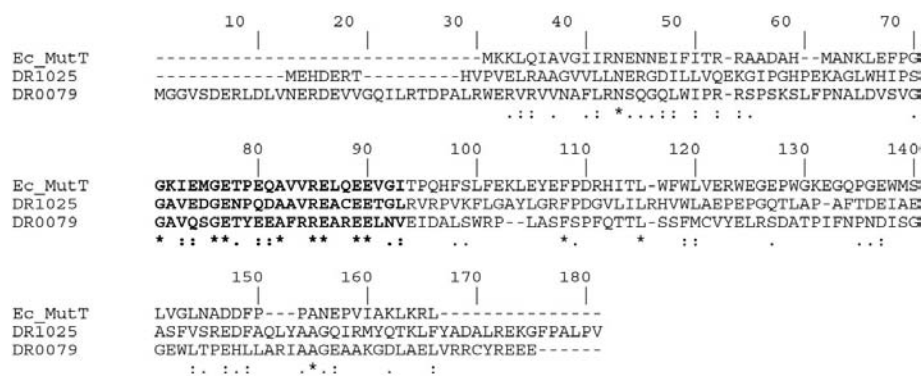
crystallization papers

Figure 1
ClustalW multiple sequence alignment of *E. coli* MutT, DR1025 and DR0079. As shown ('*', identities; '.', similar; ':', very similar), only the nudix box (bold) is conserved (8/23 identical). The rest of the sequences demonstrate the diversity of this protein family.

concentration of 0.5 mM. Growth continued at 310 K for 3 h. Cells were harvested, rinsed and stored at 193 K. 3 g of cells was resuspended in 30 ml lysis buffer (50 mM Tris pH 7.5, 1 mM phenylmethylsulfonyl fluoride (PMSF), 10 µg ml^{-1} DNAse, 5 µg ml^{-1} leupeptin, 10 µg ml^{-1} antipain, 0.7 µg ml^{-1} pepstatin, 0.3 µg ml^{-1} chymostatin, 1 mM DTT) and lysed by sonication (pulsing) for 6 min. The cleared supernatant formed after centrifugation (12 000g, 20 min) was treated with streptomycin sulfate (1/2 volume of 10% solution) in order to precipitate nucleic acids. The resulting supernatant after centrifugation (12 000g, 20 min) was brought to 2.4 M ammonium sulfate to precipitate the proteins. The pellet formed after centrifugation (13 200g, 50 min) was resuspended in 25 mM Tris pH 7.5, 1 mM EDTA and dialyzed (8 kDa molecular-weight cutoff) against the same to remove salt. Anion-exchange chromatography (protein pI 4.76) was performed on a Pharmacia Hi-Trap Q HP column (5 ml) using the BioCAD Sprint system. Protein was eluted with a salt gradient from 0 to 1 M NaCl. DR1025-containing fractions (~0.4 M NaCl) were pooled, concentrated and desalted through buffer exchange using an Ultrafree centrifuge filter (5 kDa molecular-weight cutoff). SDS–PAGE showed the protein to be at least 98% pure and of proper molecular weight. Yield was approximately 14 mg pure protein from 1.5 g of cells (wet weight; 250 ml growth media) or 56 mg per litre of cells. Dynamic light-scattering experiments on a DynaPro99 (Protein Solutions) showed a radius of 2.52 nm and polydispersity of 42%. This is consistent with the protein being a monomer with a calculated molecular weight of 17 569 Da (159 amino acids).

2.1.2. DR0079. The DR0079 gene was amplified from the genomic DNA of *D. radiodurans* (ATCC) with the primers

5′-GACGGTCTCCCATGGGGGGGCGTG-AGC and 5′-GACGAGGAGGAATTCT-TATTCTTCCTCCCTGTAGCAGC using standard PCR methods. The resulting PCR product was digested with *Nco*I and *Eco*RI and cloned into the pET30b expression vector (Novagen). DNA sequencing indicated that the sequence of the clone was identical to the reported sequence. The plasmid containing DR0079 was transformed into *E. coli* strain BL21(DE3) (Novagen) and cells were grown at 310 K to an OD_{600} of ~0.8 in 750 ml Luria broth supplemented with 34 µg ml^{-1} kanomycin. Protein expression was induced by addition of IPTG to a concentration of 1.0 mM. Following addition of 50 µl of 0.5 M zinc acetate, the cells were shaken for 4 h at 301 K, harvested (OD_{600} of ~1.4) and frozen at 193 K. Thawed cells were resuspended in 40 ml lysis buffer (0.3 M NaCl, 50 mM potassium phosphate, 10 mM imidazole pH 7.5) and brought to 0.2 µM in PMSF prior to three passes through a French press (SLM Instruments). Following 1 min of sonication, the cell debris was spun at 24 000g for 45 min. The supernatant was loaded onto a 20 ml Ni–NTA affinity column (Qiagen) and washed stepwise with buffer (0.3 M NaCl, 50 mM potassium phosphate pH 8.1) containing increasing concentrations of imidazole (10–500 mM). The DR0079 fraction (~100 mg, 250 mM imidazole) was dialyzed (8 kDa molecular-weight cutoff) at 277 K overnight (50 mM NaCl, 20 mM Tris–HCl pH 7.6). Following volume reduction to ~3.0 ml (Amicon Centriprep-10), the N-terminal polyhistidine tag was removed at room temperature with 25 µl enterokinase (Novagen) and 6 µl of 1.0 M CaCl$_2$. After overnight incubation, an aliquot of the reaction mixture was loaded onto a Bio-Rad Econo-Pak High-Q cartridge (5 ml) using the BioCAD Sprint System (protein pI 4.85) to monitor the progress of the reaction. The

column was washed (50 mM Tris–HCl, 50 mM Tris-bis propane pH 7.5) and eluted with a linear gradient from 0 to 150 mM NaCl over 20 column volumes. Under these conditions, cut DR0079 eluted with 55 mM NaCl and its fusion protein eluted with 90 mM NaCl. If the reaction was not over 90% complete, additional enterokinase was added.

Approximately 40 mg protein fractions were then loaded onto the High-Q cartridge. The peak containing DR0079 was collected, pooled and the volume reduced to ~2 ml. DR0079 was further purified on a Pharmacia Superdex75 HiLoad column that simultaneously exchanged it into crystallization buffer (20 mM Tris–HCl, 100 mM KCl, 1 mM DTT pH 7.1). Using a flow rate of 1.5 ml min^{-1}, approximately 18 mg of DR0079 eluted at 52 min, a retention time characteristic of a protein of size consistent with the calculated monomeric molecular weight of 19 283 Da (171 amino acids). The wide chemical shift dispersion and well defined cross peaks in the $^{15}N/^1H$ HSQC NMR spectrum also suggest DR0079 to be a monomer in solution. SDS–PAGE showed the protein to be greater than 98% pure.

2.2. Crystallization and soaking

2.2.1. DR1025. Vapor-diffusion crystallization trials using hanging drops were set up using screens from Hampton Research. Large single crystals formed at room temperature after 1–2 d upon mixing equal volumes of protein (15 mg ml^{-1} in 25 mM Tris pH 7.5, 1 mM EDTA) and reservoir (100 mM sodium acetate pH 4.5, 2 M sodium formate). The protein formed bipyramidal crystals of 200–500 µm in the longest dimension. A large single crystal is shown in Fig. 2(a). Crystals were prepared for freezing in liquid nitrogen by successively soaking for 5–10 min in reservoir solutions containing 10, 20 and 30% glycerol. Samarium-derivative crystals were prepared by soaking in 250 mM Sm(NO$_3$)$_3$, 0.74 M sodium acetate pH 4.5, 1.4 M sodium formate, 30% glycerol for 10 min or in 0.5 mM SmCl$_3$, 0.74 M sodium acetate pH 4.5, 1.4 M sodium formate for 24 h.

2.2.2. DR0079. Crystallization screening of DR0079 was performed at room temperature using the vapour-diffusion hanging-drop method and screens from Hampton Research. 2 µl of DR0079 (29 mg ml^{-1}) in crystallization buffer was mixed with 2 µl reservoir (30% PEG 4000, 0.1 M sodium citrate pH 5.6, 0.2 M ammonium acetate or 30% PEG 4000, 0.1 M Tris–HCl pH 8.5, 0.2 M lithium sulfate). Crystals

formed as 'coffin-shaped' plates after approximately four weeks and are shown in Fig. 2(*b*). Crystals were frozen in liquid nitrogen in the absence of additional cryo-protectant.

2.3. Diffraction data collection and reduction

All diffraction data were collected at the Macromolecular Crystallography Facility of the Advanced Light Source (ALS) synchrotron at Lawrence Berkeley National Laboratory. Heavy-atom incorporation into the DR1025 crystals was verified by fluorescence scanning. Images were integrated and scaled using the *HKL* and *HKL*2000 packages of data-reduction software (Otwinowski, 1993). Crystal lattices and space groups for both proteins were determined by autoindexing and examination of systematic absences in the reduced data.

3. Results

3.1. Data collection for DR1025

Data from a native crystal were initially recorded to 1.6 Å resolution on ALS beamline 5.0.1. Subsequently, data were collected from a crystal that had been soaked for 24 h in $SmCl_3$ but did not show incorporation of Sm by fluorescence scans. This crystal diffracted to beyond 1.4 Å on ALS beamline 5.0.2. This data set is summarized in Table 1 and will be referred
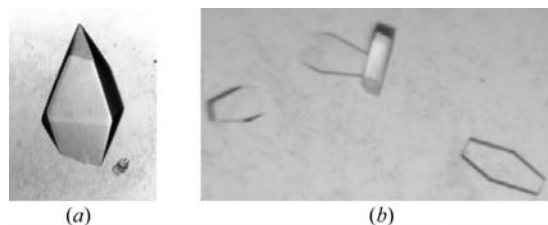
Figure 2
(*a*) Tetragonal crystal of DR1025. The longest dimension is approximately 500 µm. (*b*) Orthorhombic crystals of DR0079.
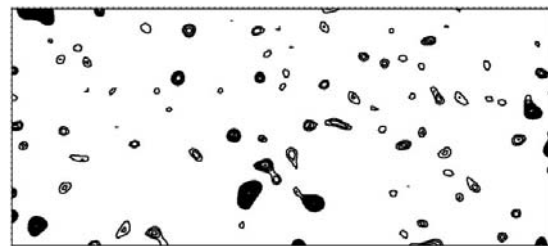
Figure 3
Harker section of anomalous difference Patterson for DR1025 samarium derivative. The *xz* plane is shown with the vertical axis *x* and the horizontal axis *z* in fractional coordinates from 0.0 to 0.5 for both axes. The three largest peaks of 21.8σ (0.5, 0.5, 0.0171), 11.1σ (0.1105, 0.5, 0.2194) and 10.6σ (0.0899, 0.5, 0.2794) can be seen.

Table 1
Summary of crystallographic data.

Values in parentheses refer to the highest resolution shell (1.39–1.44 Å for DR1025 native; 1.80–1.84 Å for DR1025 Sm; 1.90–1.97 Å for DR0079).

	DR1025 native	DR1025 Sm	DR0079
Wavelength (Å)	1.0	1.2	1.0
Crystal system	Tetragonal	Tetragonal	Orthorhombic
Space group	$P4_12_12/P4_32_12$	$P4_12_12/P4_32_12$	$C222_1$
Unit-cell parameters (Å)	$a = b = 53.2$,	$a = b = 53.2$,	$a = 34.1$, $b = 157.2$,
	$c = 122.6$	$c = 122.0$	$c = 126.5$
Resolution (Å)	1.39	1.80	1.90
Total No. of observations	879484	453884	642643
Total No. of unique reflections	35874	30510†	26124
Coverage completeness (%)	98.4 (89.5)	98.3 (98.2)	95.4 (77.8)
R_{sym}‡ (%)	5.8 (34.4)	4.8 (9.6)	8.3 (36.7)
Average $I/\sigma(I)$	56.1 (4.4)	43.2 (18.9)	22.7 (2.2)

† Anomalous reflections not merged. ‡ $R_{sym} = \sum_{hkl} \sum_i |I_i - \langle I \rangle| / \sum \langle I \rangle$, where I_i is the intensity of the *i*th measurement of reflection *hkl* and $\langle I \rangle$ is the mean intensity of reflection *hkl*.

to as the 'native' data set. Assuming a monomer in the asymmetric unit, the estimated solvent content is 50.2% and the Matthews coefficient (V_M) is calculated to be 2.5 $Å^3$ Da^{-1}.

Diffraction data were also collected from a crystal soaked in $Sm(NO_3)_3$ (see above). This crystal diffracted to 1.80 Å resolution with an R_{sym} of 4.8%, as shown in Table 1 (DR1025 Sm). Friedel pairs were collected by the inverse-beam procedure and reduced for SAD phasing. The presence of samarium was indicated by a fluorescence scan. The average anomalous difference for this data set was 5.0 and 7.3% (on F_{obs}) for the highest resolution range. A Harker section from the anomalous difference Patterson is shown in Fig. 3, illustrating several strong peaks from samarium sites.

3.2. Data collection for DR0079

Data from a native crystal were measured to 1.9 Å resolution on ALS beamline 5.0.2. The space group, unit-cell parameters and data-reduction statistics are given in Table 1. The diffraction indicated increased mosaicity of the crystals as data collection progressed. Assuming two protein molecules in the asymmetric unit, the estimated solvent content is 44.0% and the Matthews coefficient (V_M) is calculated to be 2.2 $Å^3$ Da^{-1}, similar to that of DR1025.

4. Conclusions

In *D. radiodurans*, the nudix hydrolase family may contain up to 24 proteins, the most known

for any organism. In light of the extreme radioresistance of this bacterium and the activity of the prototypical nudix protein MutT in DNA repair, we are seeking to characterize the structure–function relationships of proteins of this family.

High-resolution native diffraction data were collected on two nudix family proteins from *D. radiodurans*, DR0079 and DR1025. Anomalous diffraction data were collected on a potential derivative of DR1025.

Dynamic light scattering and gel filtration suggest that DR1025 is a dimer and DR0079 is a monomer in solution, even though DR0079 has two molecules per asymmetric unit. Previous structural studies have found nudix hydrolases both as monomers with one or two molecules per asymmetric unit (Swarbrick *et al.*, 2000; Bailey *et al.*, 2002) and as standard (Wang *et al.*, 2002) or domain-swapped dimers (Gabelli *et al.*, 2001). Solution-activity screens suggest that DR1025 may have activity towards dGTP and Ap_4A substrates, while the specificity of DR0079 is not apparent (Xu *et al.*, 2001). The three-dimensional structures of these proteins and their substrate complexes can be used to further characterize the functions of these proteins and the evolutionary relationships of this diverse family, as well as their role in the maintenance of cellular and genomic integrity. The data collected from the crystals described here is the initial step toward these goals.

crystallization papers

References

Abeygunawardana, C., Weber, D. J., Gittis, A. G., Frick, D. N., Lin, J., Miller, A. F., Bessman, M. J. & Mildvan, A. S. (1995). *Biochemistry*, **34**, 14997–15005.

Bailey, S., Sedelnikova, S. E., Blackburn, G. M., Abdelghany, H. M., Baker, P. J., McLennan, A. G. & Rafferty, J. B. (2002). *Structure*, **10**, 589–600.

Bessman, M. J., Frick, N. & O'Handley, S. F. (1996). *J. Biol. Chem.* **271**, 25059–25062.

Bonanno, J. B., Edo, C., Eswar, N., Pieper, U., Romanowski, M. J., Ilyin, V., Gerchman, S. E., Kycia, H., Studier, F. W., Sali, A. & Burley, S. K. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 12896–12901.

Dunn, C. A., O'Handley, S. F., Frick, D. N. & Bessman, M. J. (1999). *J. Biol. Chem.* **274**, 32318–32324.

Durbecq, V., Sainz, G., Oudjama, Y., Clantin, B., Bompard-Gilles, C., Tricot, C., Caillet, J., Stalon, V., Droogmans, L. & Villeret, V. (2001). *EMBO J.* **20**, 1530–1537.

Fletcher, J. I., Swarbrick, J. D., Maksel, D., Gayler, K. R. & Gooley, P. R. (2002). *Structure*, **10**, 205–213.

Gabelli, S. B., Bianchet, M. A., Bessman, M. J. & Amzel, L. M. (2001). *Nature Struct. Biol.* **8**, 467–472.

Lin, J., Abeygunawardana, C., Frick, D. N., Bessman, M. J. & Mildvan, A. S. (1997). *Biochemistry*, **36**, 1199–1211.

Makarova, K. S., Aravind, L., Daly, M. J. & Koonin, E. V. (2000). *Genetica*, **108**, 25–34.

Mattimore, V. & Battista, J. R. (1996). *J. Bacteriol.* **178**, 633–637.

Otwinowski, Z. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.

Swarbrick, J. D., Bashtannyk, T., Maksel, D., Zhang, X. R., Blackburn, G. M., Gayler, K. R. & Gooley, P. R. (2000). *J. Mol. Biol.* **302**, 1165–1177.

Wang, S., Mura, C., Sawaya, M. R., Cascio, D. & Eisenberg, D. (2002). *Acta Cryst.* D**58**, 571–578.

Xu, W., Shen, J., Dunn, C. A., Desai, S. & Bessman, M. J. (2001). *Mol. Microbiol.* **39**, 286–290.