Acta Crystallographica Section D
Biological Crystallography
ISSN 1399-0047

The first crystal structure of NAD-dependent 3-dehydro-2-deoxy-D-gluconate dehydrogenase from *Thermus thermophilus* HB8

Kudigana J. Pampa,[a]* Neratur K. Lokanath,[b] Naoki Kunishima[c] and Ravishankar Vittal Rai[a]

[a]Department of Studies in Microbiology, University of Mysore, Manasagangotri, Mysore 570 006, India, [b]Department of Studies in Physics, University of Mysore, Manasagangotri, Mysore 570 006, India, and [c]Advanced Protein Crystallography Research Group, RIKEN SPring-8 Center, Harima Institute, 1-1-1 Kouto, Sayo-cho, Sayo-gun, Hyogo 679-5148, Japan

Correspondence e-mail: sagarikakj@gmail.com

Received 27 August 2013
Accepted 30 December 2013

PDB references: DDGDH, native form I, 4jp2; native form II, 2ekq; cofactor complex, 2ekp; citrate complex, 4jp3

2-Keto-3-deoxygluconate (KDG) is one of the important intermediates in pectin metabolism. An enzyme involved in this pathway, 3-dehydro-3-deoxy-D-gluconate 5-dehydrogenase (DDGDH), has been identified which converts 2,5-diketo-3-deoxygluconate to KDG. The enzyme is a member of the short-chain dehydrogenase (SDR) family. To gain insight into the function of this enzyme at the molecular level, the first crystal structure of DDGDH from *Thermus thermophilus* HB8 has been determined in the apo form, as well as in complexes with the cofactor and with citrate, by X-ray diffraction methods. The crystal structures reveal a tight tetrameric oligomerization. The secondary-structural elements and catalytically important residues of the enzyme were highly conserved amongst the proteins of the NAD(P)-dependent SDR family. The DDGDH protomer contains a dinucleotide-binding fold which binds the coenzyme NAD^+ in an intersubunit cleft; hence, the observed oligomeric state might be important for the catalytic function. This enzyme prefers NAD(H) rather than NADP(H) as the physiological cofactor. A structural comparison of DDGDH with mouse lung carbonyl reductase suggests that a significant difference in the α–loop–α region of this enzyme is associated with the coenzyme specificity. The structural data allow a detailed understanding of the functional role of the conserved catalytic triad (Ser129–Tyr144–Lys148) in cofactor and substrate recognition, thus providing substantial insights into DDGDH catalysis. From analysis of the three-dimensional structure, intersubunit hydrophobic interactions were found to be important for enzyme oligomerization and thermostability.

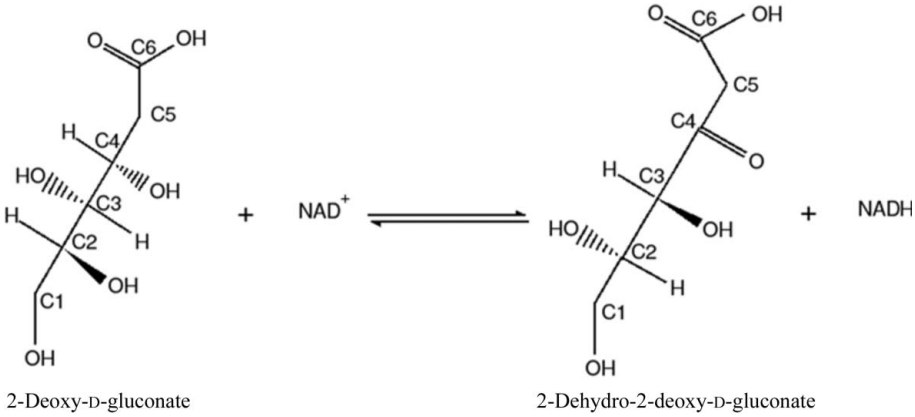
1. Introduction

NAD/$NAD(P)^+$-dependent glucose dehydrogenase (GlcDH) is an enzyme distributed in a wide variety of prokaryotic organisms. Based on sequence and structural similarities, two unique families of GlcDHs corresponding to short-chain dehydrogenases/reductases (SDRs) and medium-chain dehydrogenases/reductases (MDRs) have been identified. Members of the MDR superfamily, with a protomer molecular weight of approximately 40 kDa, contain structural and catalytic zinc ions (Jörnvall *et al.*, 1987; John *et al.*, 1994). In contrast, proteins of the SDR superfamily have protomers of molecular weight between 25 and 30 kDa (Jörnvall *et al.*, 1995; Oppermann *et al.*, 2003). The MDR and SDR enzymes are structurally related and have a common dinucleotide-binding motif (Rossmann fold; Rossmann *et al.*, 1973); however, the active-site structure and reaction mechanism are dissimilar, indicating functional divergence. SDR enzymes have received much attention as they constitute a large and diverse group of NAD(H)-dependent or NADP(H)-dependent oxidoreductase

protein family enzymes with low amino-acid sequence similarity and are found in all life forms, with several genes being found in the human genome (Jörnvall *et al.*, 1995; Oppermann *et al.*, 2001; Kallberg *et al.*, 2002; Korozowski, 1992). All available three-dimensional structures display a highly similar β/α-fold pattern. Most of them are homotetrameric or homodimeric in quaternary structure (Varughese *et al.*, 1992; Ghosh *et al.*, 1994; Tanaka *et al.*, 1996), with the exception of porcine carbonyl reductase, which is monomeric (Ghosh *et al.*, 2001).

With over 2500 sequences annotated in databases and about 540 crystal structures determined, a general picture of the SDR architecture and catalytic mechanism is emerging. Considering all of the sequences, no strict positional conservation is noted. However, multiple sequence alignments revealed several consensus motifs, the most conserved being the N-terminal TG*XXX*G*X*G motif starting from Thr8 that forms part of the nucleotide-binding fold. The active-site Ser–Tyr–Lys triad, identified through structural alignments and functional analyses, reveals the critical involvement of conserved elements for coenzyme binding, maintenance of the SDR scaffold and catalysis. Some structures of proteins unrelated in sequence but displaying the SDR fold have considerably extended the understanding of the structure–function relationship. Thus, the SDR domain structure appears to be a generic scaffold that includes not only dehydrogenase/reductase, lyase, epimerase and hydratase activities, but also RNA-binding proteins, kinases and transcription factors (Stammers *et al.*, 2001). Greater understanding of the mechanistic and structural principles governing the SDR architecture will reveal novel protein–substrate and protein–protein interactions and will facilitate the development of inhibitors directed against biologically relevant SDR targets.

2-Keto-3-deoxygluconate (KDG) is one of the important intermediates in pectin metabolism. We identified an enzyme involved in this pathway, 3-dehydro-3-deoxy-D-gluconate 5-dehydrogenase (DDGDH), which converts 2,5-diketo-3-deoxygluconate (DDG) to KDG, as a member of the SDR family (Fig. 1). To gain insight into the function of this enzyme, its catalytic mechanism and its structural rearrangement, which is presumably important for catalysis, we determed the crystal structure of DDGDH from *Thermus thermophilus* HB8. We describe here for the first time the crystal structure of DDGDH from *T. thermophilus* HB8 in two apo forms and in NAD^+-bound and citrate-bound forms at atomic resolution. To our knowledge, this is the first detailed three-dimensional structural characterization of any DDGDH. Unexpectedly, we found that *T. thermophilus* DDGDH binds NAD rather than NADP as a cofactor at an intersubunit cleft.



2-Deoxy-D-gluconate 2-Dehydro-2-deoxy-D-gluconate

Figure 1
Reaction scheme of DDGDH.

2. Materials and methods

2.1. Protein expression and purification

The plasmid encoding the DDGDH protein, provided by the RIKEN Genomic Sciences Center, was digested with *Nde*I and *Bgl*II and the fragment was inserted into the expression vector pET-11a linearized with *Nde*I and *Bam*HI. The recombinant plasmid was transformed into *Escherichia coli* BL21(DE3) cells and was grown at 310 K in Luria–Bertani medium containing 509 µg ml^{-1} ampicillin for 20 h. The cells were harvested by centrifugation at 6500 rev min^{-1} for 5 min at 277 K. The cell pellet was suspended in 20 m*M* Tris–HCl pH 8.0 containing 0.5 *M* sodium chloride and 5 m*M* β-mercaptoethanol and was then homogenized by ultrasonication. The supernatant was heated at 343 K for 12 min and the cell debris and denaturated protein were then removed by centrifugation (14 000 rev min^{-1}, 30 min); the supernatant solution was used as the crude extract for purification. The crude extract was desalted using a HiPrep 26/10 desalting column and applied onto a Super Q Toyopearl 650M column equilibrated with 20 m*M* Tris–HCl pH 8.0. The fraction containing DDGDH was eluted with a linear gradient of 0–0.3 *M* sodium chloride. The protein was then dialyzed against 20 m*M* Tris–HCl pH 8.0 and subjected to a Resource Q column (Amersham Biosciences) equilibrated with 20 m*M* Tris–HCl pH 8.0. Fractions containing protein were again eluted with a linear gradient of 0–0.3 *M* sodium chloride. The protein was desalted using a HiPrep 26/10 desalting column with 10 m*M* sodium phosphate pH 7.0 and applied onto a Bio-Scale CHT20-I column (Bio-Rad) equilibrated with 10 m*M* sodium phosphate pH 7.0. DDGDH was eluted with a linear gradient of 10–150 m*M* sodium phosphate and was subsequently desalted using a HiPrep 26/10 desalting column with 20 m*M* Tris–HCl pH 8.0 containing 0.05 *M* sodium chloride and applied onto a Mono Q column (Amersham Biosciences) equilibrated with 20 m*M* Tris–HCl pH 8.0 containing 0.05 *M* sodium chloride. The fraction containing DDGDH was eluted with a linear gradient of 0–0.5 *M* sodium chloride. The fraction containing the protein was cooled, concentrated by ultrafiltration (Vivaspin, 10 kDa cutoff) and loaded onto a HiLoad 16/60

Table 1
Crystal data, data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

	Apo form I	Apo form II	NAD^+-bound form	Citrate-bound form
Data statistics				
Space group	$P4_22_12$	$P4_12_12$	$I222$	$I222$
Unit-cell parameters (Å)	$a = b = 62.9$, $c = 114.3$	$a = b = 110.5$, $c = 200.8$	$a = 66.8$, $b = 87.2$, $c = 92.1$	$a = 66.9$, $b = 86.9$, $c = 91.7$
V_M (Å^3 Da^{-1})	2.2	3.0	2.6	2.6
Content of the asymmetric unit	Protomer	Tetramer	Protomer	Protomer
Resolution (Å)	50.0–1.15 (1.19–1.15)	50.0–1.80 (1.86–1.80)	50.0–1.15 (1.19–1.15)	50.0–1.50 (1.55–1.50)
Reflections (measured/unique)	763174/79896	628072/114641	502279/94728	278041/42860
R_{merge}† (%)	7.1 (14.8)	6.6 (24.4)	6.2 (24.9)	6.1 (29.1)
Completeness (%)	99.8 (99.7)	99.3 (100.0)	99.1 (98.1)	99.4 (96.4)
$\langle I/\sigma(I)\rangle$	16.6 (13.7)	10.5 (6.4)	10.1 (6.6)	8.5 (2.6)
Multiplicity	9.6	5.5	5.3	6.5
Refinement statistics				
Resolution (Å)	40.0–1.40	40.0–1.80	40.0–1.15	40.0–1.50
R_{cryst} (%)	18.9	18.4	16.8	16.3
R_{free} (%)	19.9	20.8	16.8	17.8
R.m.s.d., bond lengths (°)	0.004	0.004	0.005	0.004
R.m.s.d., bond angles (Å)	1.15	1.20	1.10	1.20
Ramachandran plot (%)				
Most favoured	89.9	90.3	90.5	89.9
Additionally allowed	10.1	9.7	9.5	10.1
Generously allowed	0.0	0.0	0.0	0.0

Superdex 75 pg column (Amersham Biosciences) equilibrated with 20 m*M* Tris–HCl pH 8.0 containing 0.05 *M* sodium chloride. The homogeneity and the identity of the purified sample were estimated by SDS–PAGE (Laemmli, 1970) and N-terminal sequence analysis. Protein concentrations were determined using the UV method and a Bio-Rad protein assay based on the Bradford dye-binding procedure using bovine serum albumin as a standard. Finally, the purified DDGDH was concentrated to 20.0 mg ml^{-1} by ultrafiltration and was stored at 203 K.

The oligomeric state of purified DDGDH was examined by a dynamic light-scattering experiment performed using a DynaPro MS/X instrument (Protein Solutions). The measurements were obtained at 291 K using the purified protein at 1 mg ml^{-1} in buffer solution consisting of 20 m*M* Tris–HCl, 200 m*M* sodium chloride. Several measurements were taken and analyzed by the *DYNAMICS* software (v.3.30; Protein Solutions). A bimodal analysis resulted in a molecular weight of 98 kDa, which is consistent with a tetrameric form of the protein in solution.

2.2. Crystallization

Diffraction-quality crystals were grown at room temperature by the sitting-drop method by mixing equal volumes (1 µl) of protein solution and reservoir solution. Initial crystallization conditions were established using screening kits from Hampton Research (Jancarik & Kim, 1991). Apo forms I and II were obtained with solutions consisting of 0.1 *M* sodium acetate pH 4.8, 2.0 *M* sodium formate and 0.1 *M* citrate–HCl pH 5.7, 0.2 *M* ammonium sulfate, 16.5%(*w*/*v*) polyethylene glycol (PEG) 20 000, respectively. Crystals of the DDGDH complex with cofactor were successfully obtained with the protein sample adjusted to 15.0 mg ml^{-1} and were incubated with 5 m*M* β-NAD^+ (Sigma) overnight. Crystals of the DDGDH–NAD^+ complex were grown using a reservoir solution consisting of 0.1 *M* Bicine pH 9.0, 2% dioxane, 10% PEG 20 000. We also attempted to cocrystallize DDGDH with KDG by incubating the protein sample with 5 m*M* KDG overnight, but obtained only citrate-bound crystals. Citrate-bound crystals were grown using a reservoir solution containing 1.6 *M* trisodium citrate dihydrate pH 6.5. All of the above crystals appeared within a week and grew to typical dimensions of 0.3 × 0.2 × 0.3 mm.

2.3. Data collection, processing and characterization of crystals

All X-ray diffraction data sets were collected on beamline BL26B1 at SPring-8, Japan under cryogenic conditions. Crystals were flash-cooled with liquid nitrogen at 100 K in their respective mother liquor or in a soaking solution containing 30%(*v*/*v*) glycerol as the cryoprotectant. A Rigaku image-plate detector was used for data collection. The data were indexed and scaled with the *HKL* package (Otwinowski & Minor, 1997). The data sets were completed by including all possible *hkl* and R_{free} columns using *UNIQUE* from the *CCP*4 suite (Winn *et al.*, 2011). Care was taken to keep the same R_{free} flags in all data sets. Structure factors were calculated with the program *TRUNCATE* from the *CCP*4 suite. Data-collection parameters and processing statistics are given in Table 1.

2.4. Structure determination

Molecular replacement (Brünger *et al.*, 1998) was performed using a trimmed model of the structure of *Thermotoga maritima* DDGDH (PDB entry 1vl8; Joint Center for Structural Genomics, unpublished data) as the search model. The sequence identity between the DDGDHs from *Thermus thermophilus* and *Thermotoga maritima* is 37%. To minimize the bias caused by the search model, we modified the search model based on sequence alignment. Low sequence similarity and sequence gaps were eliminated from the model. The orientation of the search model was determined by a fast rotation function calculated using *CNS* (Brünger *et al.*, 1998) using data between 15 and 4 Å resolution. The best solution from the cross-rotation search scored 2.15σ above the mean. This peak gave a clear solution in the translation search in space group $P4_22_12$ for the apo form I crystal. Subsequently, a refined monomer of the apo form I of DDGDH was used as the phasing model in a molecular-replacement search for the apo form II, cofactor-complex and citrate-complex crystals. The cofactor-bound and citrate-bound crystals were

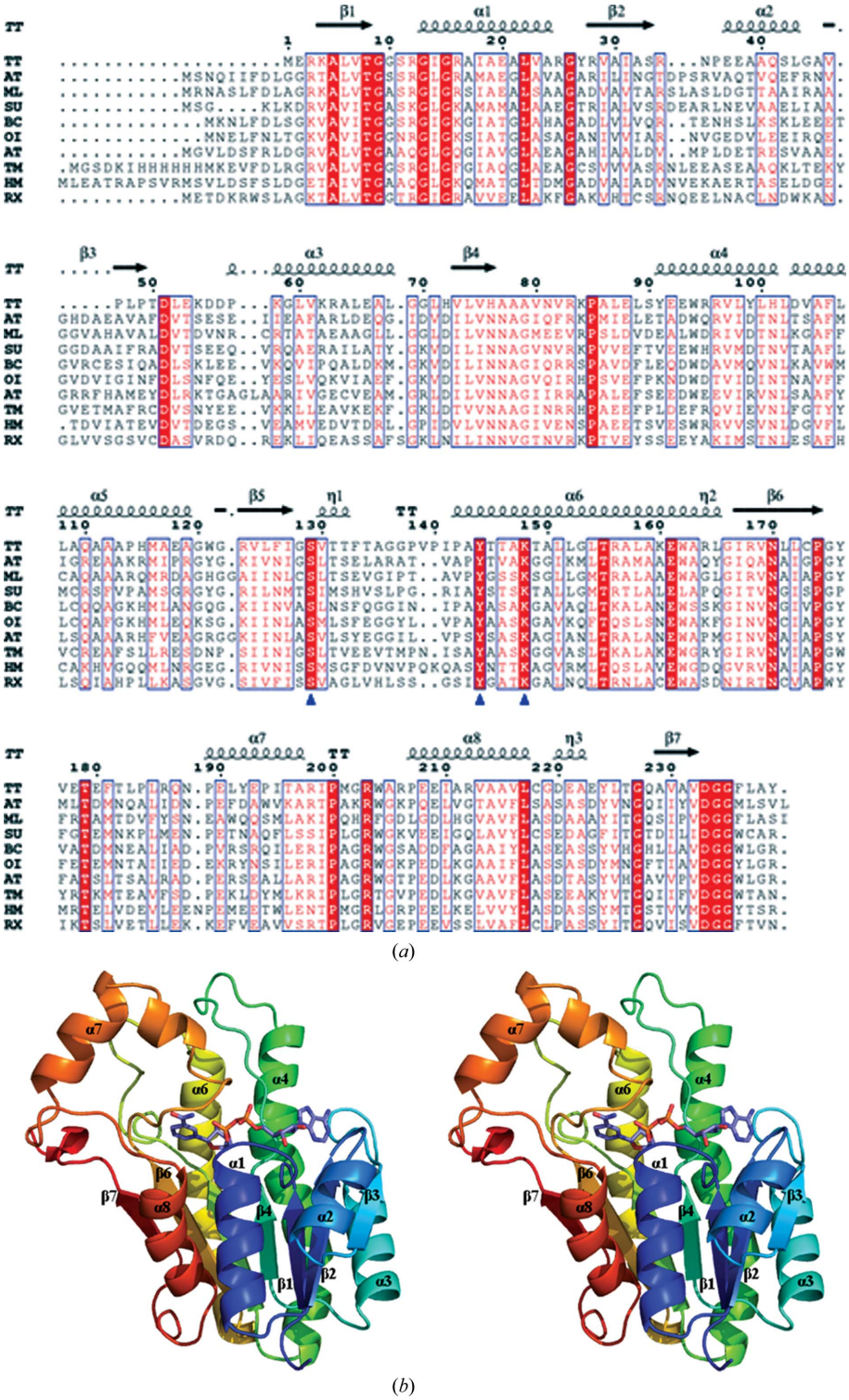


Figure 2
Sequence alignment of 3-dehydro-2-deoxy-D-gluconate dehydrogenase orthologues. (*a*) The secondary-structural elements of *T. thermophilus* DDGDH are shown above the alignment. Strictly conserved residues are highlighted. The blue triangles below the sequences mark the key residues of the catalytic triad (Ser129–Tyr144–Lys148) of the SDR family. TT, *Thermus thermophilus* HB8; AT, *Arabidopsis thaliana*; ML, *Mesorhizobium loti* MAFF; SU, *Solibacter usitatus*; BC, *Bacillus clausii* KSM-K16, OI, *Oceanobacillus iheyensis* HTE831; AT, *Agrobacterium tumefaciens* C58; TM, *Thermotoga maritima* MSB8; HM, *Haloarcula marismortui* ATCC; RX, *Rubrobacter xylanophilus* DSM. This figure was drawn with *ClustalX* (Thompson *et al.*, 1997) and *ESPript* (Gouet *et al.*, 1999). (*b*) Stereoview of the ribbon representation of the DDGDH protomer with the secondary-structural elements indicated. The N- and C-termini are coloured blue and red, respectively. The cofactor NAD^+ is shown as a ball-and-stick model.

isomorphous, even though they were grown in different crystallization conditions.

2.5. Crystallographic refinement, density modification and model building

The initial *R* factor of the apo form I DDGDH structure was 52.5% ($R_{\rm free}$ = 55.1%). Subsequently, the auto-building program *ARP*/*wARP* was used to build ~86% of the protein model. Further model building of the apo form I structure was achieved using the program *QUANTA*. Model refinement was performed using *CNS* (Brünger *et al.*, 1998). A randomly chosen 5% of all of the measured reflections were set aside for cross-validation. After iterative rounds of model fitting and refinement, solvent atoms were located manually using *QUANTA*. Model refinement of the apo form II, NAD^+-complex and citrate-complex structures was also carried out using *CNS*. Clear electron density was observed for the cofactor or citrate bound to the enzyme. After corrections of the protein model were almost complete, cofactor or citrate molecules were added to the respective model and were refined using *CNS*. For the apo form II structure, a glycerol molecule and sulfate ions were added near the active site. The topology and parameter files for NAD^+, citrate and glycerol were obtained from the Hetero-compound Information Centre – Uppsala (HIC-Up; Kleywegt & Jones, 1998). The stereochemical quality of all of the models was checked with *PROCHECK*. The refinement statistics are summarized in Table 1.

Electron density is continuous and well defined in all the crystal structures, but in the apo form II crystal structure amino acids Asp180–Asp189 (the loop between β6 and α7) in protomer *D* had weak density or no density at all. A residue at the N-terminus was also left out in all structures, as no well defined electron density revealing its position and conformation was observed.

2.6. Model analyses

The *DALI* server was utilized for a similarity search against all known

structures deposited in the PDB. Superimposition of protein models was performed using *LSQKAB* (Kabsch, 1976). *CNS* and the *Protein–Protein Interaction Server* were used for the calculation of solvent-accessible surface areas (http://www.biochem.ucl.ac.uk/bsm/PP/server/). Publication-quality pictures were generated using *PyMOL* (Schrödinger; http://www.pymol.org).

2.7. Protein Data Bank accession codes

The atomic coordinates and structure-factor amplitudes for native form I, native form II, cofactor-bound and citrate-bound forms of DDGDH have been deposited in the RCSB Protein Data Bank (http://www.rcsb.org/pdb) with accession codes 4jp2 (1x1e), 2ekq, 2ekp and 4jp3, respectively.

3. Results and discussion

3.1. Structure determination and model quality

The crystal structure of NAD-dependent DDGDH from *T. thermophilus* HB8 in apo form I has been determined at

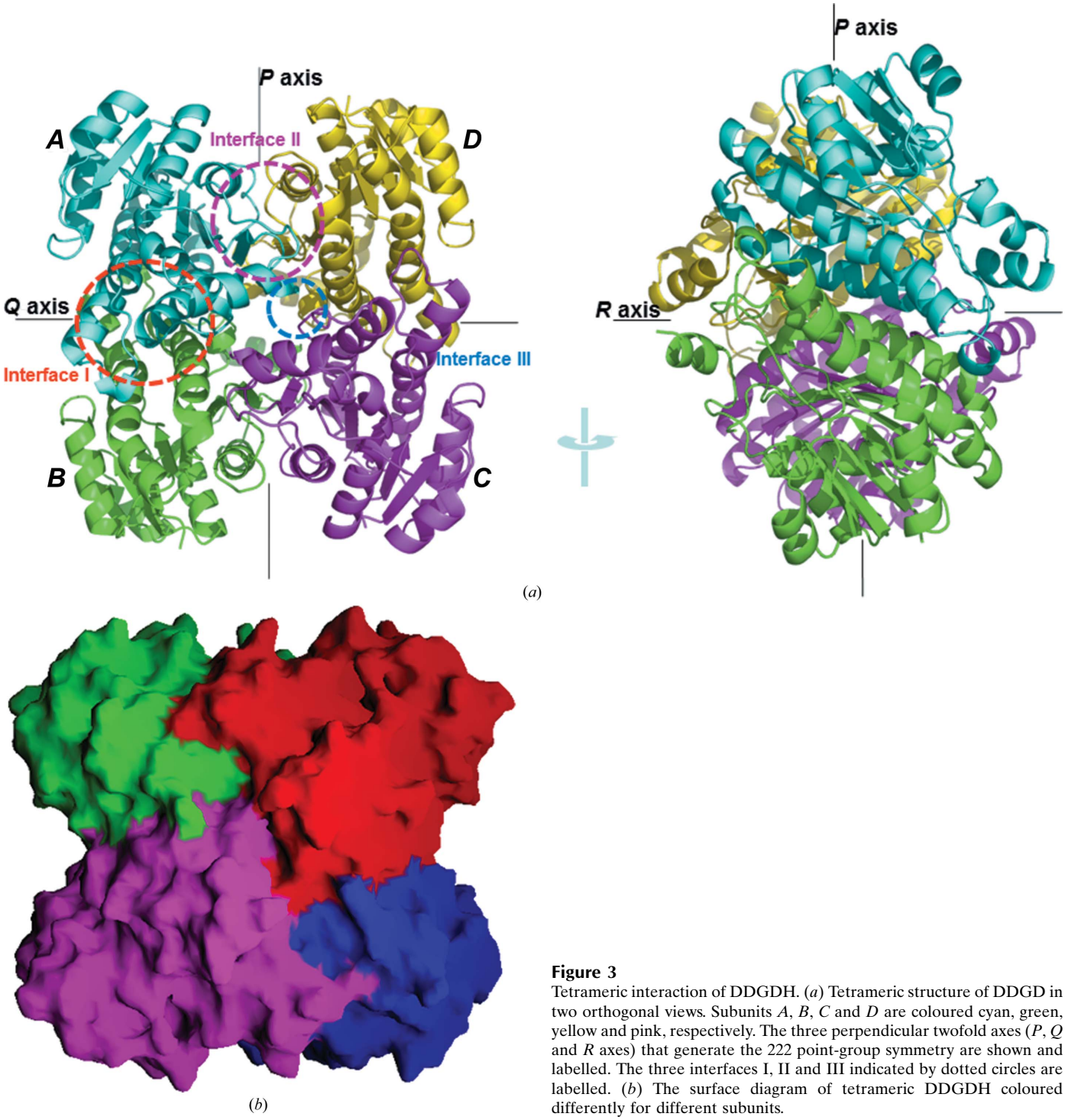


Figure 3
Tetrameric interaction of DDGDH. (*a*) Tetrameric structure of DDGD in two orthogonal views. Subunits *A*, *B*, *C* and *D* are coloured cyan, green, yellow and pink, respectively. The three perpendicular twofold axes (*P*, *Q* and *R* axes) that generate the 222 point-group symmetry are shown and labelled. The three interfaces I, II and III indicated by dotted circles are labelled. (*b*) The surface diagram of tetrameric DDGDH coloured differently for different subunits.

1.15 Å resolution by the molecular-replacement method (Brünger *et al.*, 1998). Subsequently, the structures of apo form II, the cofactor complex and the citrate complex were determined by the molecular-replacement method using the apo form I coordinates as a search model and were refined to 1.80, 1.15 and 1.50 Å resolution, respectively. The asymmetric units of the apo form I, NAD-complex and citrate-complex crystals consist of a single protomer with a calculated Matthews coefficient (Matthews, 1968) V_M of 2.5 Å^3 Da^{-1}, whereas the apo form II asymmetric unit consists of four subunits with a Matthews coefficient V_M of 2.9 Å^3 Da^{-1}. Continuous significant electron density was observed for main-chain atoms, and the atomic model for residues 2–239 could easily be established. Analysis of the stereochemistry using *PROCHECK* (Laskowski *et al.*, 1993) shows that 89.9% of the residues are in the favoured regions and 10.1% are in additionally allowed regions of the Ramachandran plot. No residues lie in the generously allowed or disallowed regions. The refined model has very low *R* values and excellent geometry. Refinement statistics and the model quality are summarized in Table 1.

3.2. Overall structure of DDGDH protomer

The DDGDH protomer forms a globular structure consisting of seven β-strands (named $\beta1$–$\beta7$) and eight α-helices (named $\alpha1$–$\alpha8$) (Figs. 2*a* and 2*b*). The protomer of DDGDH contains a slightly modified Rossmann-fold motif, with a twisted parallel seven-stranded β-sheet ($\beta1$–$\beta7$) flanked on both sides by four parallel α-helices ($\alpha1$, $\alpha2$, $\alpha3$, $\alpha8$ and $\alpha4$–$\alpha7$) located on each side of the β-sheet, and can be described as a doubly wound β/α structure. This folding pattern has been referred to as a basic β/α SDR fold (Varughese *et al.*, 1992). Three longer helices ($\alpha4$, $\alpha5$ and $\alpha6$) form the interfacial four-helix bundle typical of dimeric and tetrameric SDR proteins. The conservation of the overall fold of SDR family proteins in DDGDH, despite the low sequence similarity, suggests that the overall fold may be conserved among all SDR family proteins.

The apo-form structure could be superimposed on the cofactor-complex and citrate-complex structures with a root-mean-square deviation (r.m.s.d) of less than 0.28 Å for the main-chain atoms. These low r.m.s.d. values indicate that binding of the cofactor or citrate does not lead to significant conformational changes in the enzyme.

3.3. Tetrameric quaternary structure

Crystallographic analyses reveal that DDGDH exists as an apparent tetramer stabilized by extensive interfaces in the crystal. Notably, all of the crystal forms contain globally identical tight tetramers with approximate dimensions of 73.5 × 68.5 × 60.5 Å (Fig. 3*a*). The observed tetramers are consistent with the results of our dynamic light-scattering measurements, which suggested that DDGDH is predominantly tetrameric in solution (see §2). These observations suggest that DDGDH functions as a homotetramer (Fig. 3*b*). The asymmetric unit of the apo form II crystal contains four subunits of DDGDH related by noncrystallographic 222 point-group symmetry mediated by three perpendicular twofold axes that are conventionally termed the *P*, *Q* and *R* axes (Rossmann *et al.*, 1973; Fig. 3*a*; the subunits are denoted as chains *A*, *B*, *C* and *D*). The four subunits exhibit essentially identical overall structures, with root-mean-square deviation (r.m.s.d.) values between equivalent C$^\alpha$ atoms of subunits *A*, *B*, *C* and *D* of about 0.28 Å. We define the interface between subunits *A* and *B* as interface I, that between subunits *A* and *C* as interface II and that between subunits *A* and *D* as interface III (Fig. 3*a*). The surface areas that are buried on the formation of interfaces I and II are 1470 and 1492 Å^2, corresponding to 13.7 and 13.9% of the total protomer surface area, respectively. Interface III buries a minor area of 354 Å^2. The solvent-accessible surface areas for the isolated protomer, dimer and tetramer were calculated to be 10 853, 18 551 and 30 225 Å^2, respectively. The major interface I mediated by residues of the $\alpha5$ and $\alpha6$ helices is in the vicinity of the active site (catalytic triad). The interacting surface at interfaces I and II are essentially hydrophobic. Interface I consists of aromatic interactions and is surrounded by 18 polar interactions. On the other hand, interfaces II and III (residues between $\beta5$ and $\alpha6$, between $\alpha7$ and $\alpha8$, and between $\alpha8$ and residues at the C-termini) are dominated by eight and four hydrogen bonds, respectively. We mapped the highly conserved residues in DDGDH orthologues onto the molecular surface, which clearly revealed a conserved pattern (not shown). Interfaces I and II include the invariant residues Glu161, Pro200 and Arg203 (Fig. 2*a*). This result may indicate the biological importance of these residues in the catalytic function or in maintaining the integrity of the *AB* and *AC* dimeric interfaces. The minor interface does not contain invariant residues, indicating that residue conservation is not necessarily required. Invariant residues might contribute to the thermostability of the protein, as reported for the dimeric/tetrameric 2-deoxyribose-5-phosphate aldolase (Lokanath *et al.*, 2004).

The association of monomers into a dimer is accomplished by the four-helix bundle which is formed by helices $\alpha5$ and $\alpha6$ of the respective subunits. The $\alpha5$ and $\alpha6$ helices from one protomer are in contact with the corresponding helices from the other protomer, such that the helices from different protomers are nearly antiparallel.

3.4. Structural comparison with other dehydrogenases

A three-dimensional structural similarity search utilizing the *DALI* server (Holm & Sander, 1993) was performed between the refined models of the DDGDH protomers and the coordinates available in the Protein Data Bank (PDB). The search revealed several enzymes belonging to the SDR family fold with *Z*-scores of 32.2–15.2 and with r.m.s.d. values of 1.5–2.4 Å. The enzyme structurally most similar to DDGDH is carbonyl reductase from mouse lung (MLCR; Tanaka *et al.*, 1996; 28% sequence identity; PDB entry 1cyd; *Z*-score 32.2; r.m.s.d. of 1.5 Å for 233 C$^\alpha$ atoms). Fig. 4 shows a superimposition of the DDGDH and MLCR structures. Most of the secondary-structural elements are conserved between these two enzymes, although they show a few apparent

Table 2
R.m.s. deviations compared with DDGDH (PDB entry 4jp2).

Enzyme†	*Z*-score	R.m.s.d (Å)	Residues fitted	Sequence identity (%)
1cyd	32.2	1.5	233	28
1ybv	29.8	2.1	229	28
1e6w	28.4	1.9	216	25
1e7w	28.1	2.1	227	31
1bdb	27.5	2.0	222	26
1zmt	26.8	2.4	225	24
1gz6	26.6	2.3	207	28

† The enzymes are listed by their PDB entries: 1cyd, carbonyl reductase from mouse (Tanaka *et al.*, 1996); 1ybv, trihydroxynaphthalene reductase from *Magnaporthe grisea* (Andersson *et al.*, 1996); 1e6w, 3-hydroxyacyl-CoA dehydrogenase from *Rattus norvegicus* (Powell *et al.*, 2000); 1e7w, pteridine reductase from *Leishmania major* (Gourley *et al.*, 2001); 1bdb, *cis*-biphenyl-2,3-dihydrodiol-2,3-dehydrogenase from *Pseudomonas* sp. (Hulsmeyer *et al.*, 1998); 1zmt, haloalcohol dehalogenase from *Agrobacterium radiobacter* (de Jong *et al.*, 2005; 1gz6, (3*R*)-hydroxyacyl-CoA dehydrogenase type 2 (from rat; Haapalainen *et al.*, 2003).

differences. The main structural differences are observed at helix $\alpha3$, between $\beta2$ and $\alpha2$, and in the loop between $\beta6$ and $\alpha6$. Thus, our result is consistent with the assumption that DDGDH belongs to the SDR family. Similar structures with a nucleotide-binding domain obtained in the *DALI* search are listed in Table 2.

3.5. Cofactor NAD binding at the intersubunit cleft

Most of the NAD/NADP-dependent dehydrogenases of the SDR family share a common framework (a Rossmann fold) for binding the dinucleotide cofactor (Rossmann *et al.*, 1974), including an extended consensus sequence TG*XXX*G*X*G at the N-terminus. We analyzed the crystal structure of DDGDH complexed with NAD^+ determined by the molecular-replacement method at 1.15 Å resolution. Thanks to the high resolution of the structure, we could establish a complete atomic model of the bound NAD^+ (Fig. 5*a*) and all of the residues and solvent molecules that directly interact with NAD^+ (Fig. 5*b*). The NAD^+ molecule is bound in the intersubunit cleft between the two protomers of the dimer, and is placed at the top of the parallel seven-stranded β-sheet in an extended conformation. The cofactor interacts directly with the adjacent subunit (Fig. 5*c*). Thus, NAD^+ plays an important role in the oligomeric state of DDGDH. The adenine and nicotinamide rings are oriented roughly perpendicular to the planes of the riboses, resulting in an *anti* conformation for the adenine ring and a *syn* conformation for the nicotinamide ring. The high-resolution electron-density map allowed the unambiguous identification of the C2′-*endo* conformation for both of the ribose rings. The distance between C6 of the adenine ring and C2 of the nicotinamide ring is 14.4 Å, which is very close to those found in various other SDR enzymes. The conformational properties of the bound NAD^+ are also common to most SDR enzymes (Didierjean *et al.*, 1997).

Amino-acid residues of the conserved SDR motifs, such as Asp51, Ser129, Tyr144 and Lys148, are in close contact with NAD^+. Furthermore, Ser11 and Glu118 (from the other subunit) form hydrogen bonds to the O3′ and O2′ hydroxyl groups of the adenine ribose moiety of the dinucleotide, with bond lengths of 2.7 and 2.6 Å, respectively (Fig. 5*c*). This region determines the specificity towards NAD(H) over NADP(H), since the side chains of Ser11 and Glu118 (from the other subunit) would come into unacceptably close contact with the 2′-phosphate group of the adenosine moiety of NADP(H). Our crystal structure indicates that *T. thermophilus* DDGDH prefers NAD rather than NADP. The cofactor involved in hydrogen bonding to the adjacent subunit is highly specific to DDGDH, indicating a novel cofactor interaction of the SDR family. Of the catalytic triad Ser129–Tyr144–Lys148, Tyr144 and Lys148 hydrogen-bonded to O2′ and O3′ of the nicotinamide ribose moiety of the dinucleotide, with bond lengths of 2.69 and 2.90 Å, respectively. Strikingly, the residues responsible for cofactor recognition are not well conserved among the DDGDH orthologues except for the catalytic triad residues Tyr144, Lys148 and Asp51. These residues are important determinants of the interaction with NAD^+.

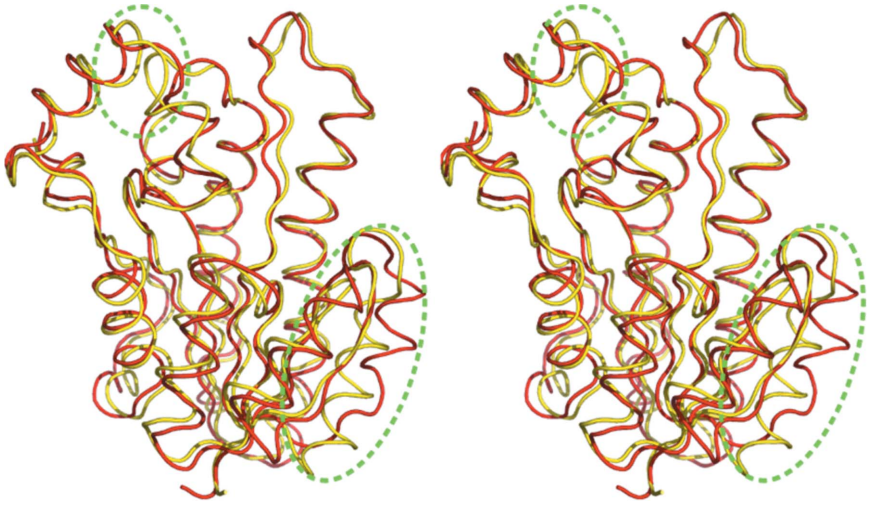


Figure 4
Stereoview of a superimposition of the DDGDH and MLCR subunit structures. DDGDH is coloured yellow and MLCR red. Differences at the main-chain level between the two structures are circled with green broken lines.

3.6. The active site (citrate, glycerol and sulfate binding)

Attempts to co-crystallize DDGDH with KDG yielded crystals of the citrate complex. The reservoir solution used for crystallization contained 0.5 *M* citrate–HCl. We analyzed the crystal structure of DDGDH complexed with the citrate molecule determined by the molecular-replacement method at 1.50 Å resolution. In this complex, the asymmetric unit contains a protomer of DDGDH. Superposition of the protomer structures of the citrate complex and apo form I results in an r.m.s.d. of 0.28 Å, indicating that the structures are very similar in spite of their different packing in the crystals. The citrate molecule

is bound to the active site of DDGDH (Fig. 6*a*) and makes hydrogen bonds to the active-site (catalytic triad) residues (Fig. 6*d*). Therefore, it is likely that the citrate molecule and the substrate DDG share the same binding pocket. The citrate molecule is associated with one subunit. The chemical structures of citrate and DDG are very similar (Fig. 7*a*), and thus citrate may be regarded as a substrate analogue. The citrate molecule interacts with the side chains of Asn81, Arg83, Ser129 and Tyr144 and the main-chain O atom of Gly175 (Fig. 6*d*). Of the three carboxyl groups of the citrate, the C1 carboxylate group makes hydrogen bonds to the side chains of Asn81 and Arg83, while the C2 carboxyl group makes hydrogen bonds to the side chains of the catalytic triad residues Ser129 and Tyr144, and the C3 carboxyl group makes hydrogen bonds to the main-chain O atom of Gly175. The orientation of the citrate molecule is optimal for catalysis because the catalytically important Ser129 and Tyr144 interact with the C2 carboxyl group. To attempt to determine the DDG binding mode in the substrate-binding pocket unambiguously, the equivalent substrate KDG was added to mother liquor containing crystals of the apo forms, cofactor complex or citrate complex of DDGDH. However, the crystals were severely damaged upon addition of the substrate, implying substantial ligand-induced conformational changes.

In the apo form II crystal, the active site is occupied by glycerol and sulfate molecules (Fig. 6*b*). The reservoir solution used for crystallization contained 1.0 *M* ammonium sulfate and the crystal was cryoprotected using 30% glycerol. The glycerol molecule interacts with the side chain of Ser129 of the catalytic triad and the sulfate molecule. The O2 atom of the glycerol molecule makes a hydrogen bond to the side chain of Ser129 and the O4 atom of the sulfate molecule (Fig. 6*e*). The sulfate molecule interacts with the side chain of Ser129, Tyr144, the main chain of Gly128 and the glycerol molecule. The O4 atom of the sulfate molecule makes a hydrogen bond to the side chain of Ser129, Tyr144, the main-chain O atom of Gly128 and the O2 atom of the glycerol molecule. We also observed a strong positive peak in the $F_o - F_c$ electron-density

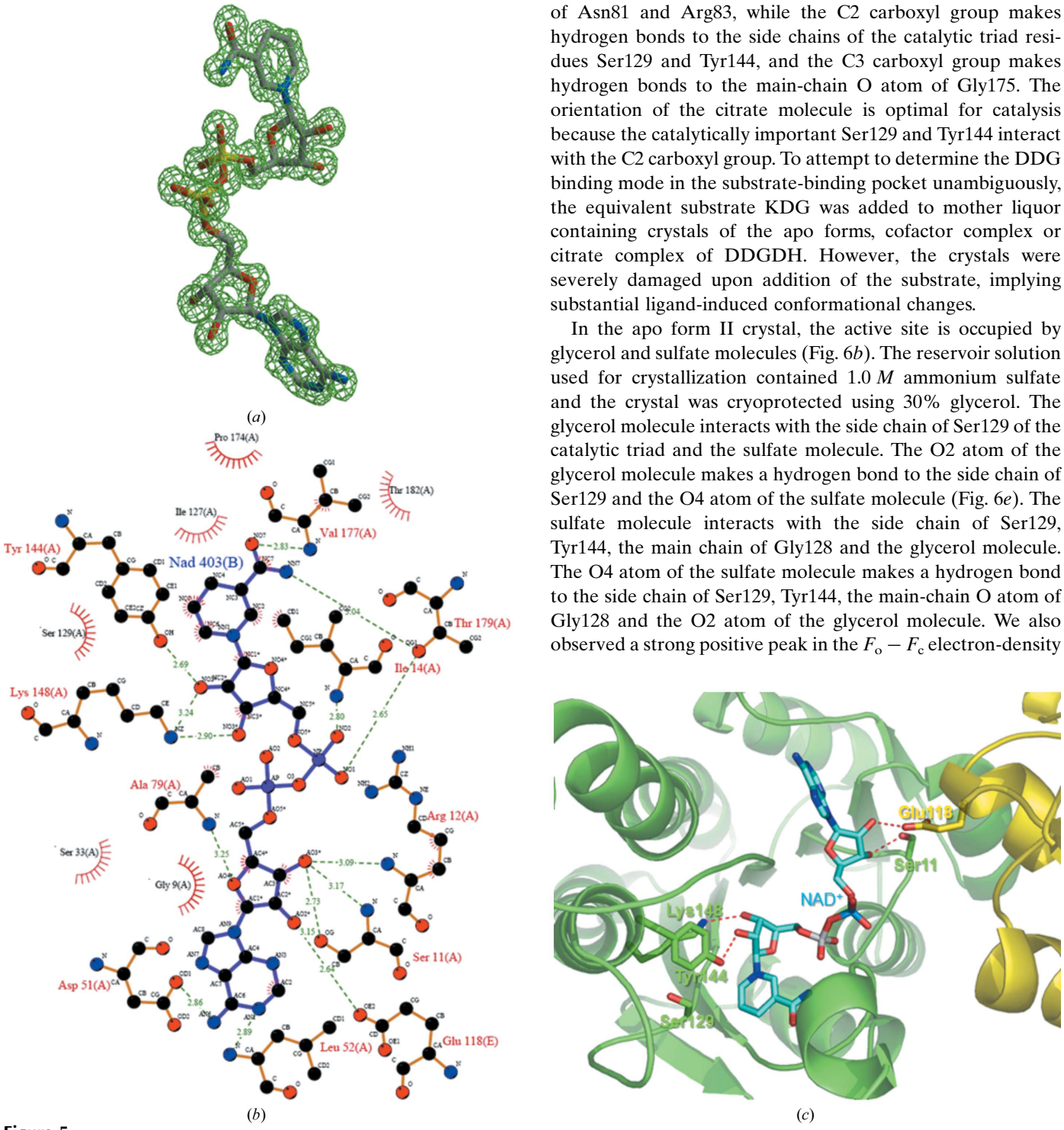


Figure 5
Binding of the cofactor near the active site of DDGDH. (*a*) $2F_o - F_c$ electron-density map of NAD^+ at 1.15 Å resolution at a 1.2σ contour level. (*b*) Structural surroundings of the NAD^+ molecule. (*c*) NAD^+ molecule located at the intersubunit cleft. Ser129, Tyr144 and Lys148 from the catalytic triad required for the hydride-transfer reaction. In the structure of DDGDH, the close contact of Ser11 and Glu118 from the other subunit with the hydroxyl groups of adenine ribose does not leave sufficient space for the phosphate group of NADP(H). This guides the specificity towards NAD(H). The subunits shown are coloured green and yellow, respectively.

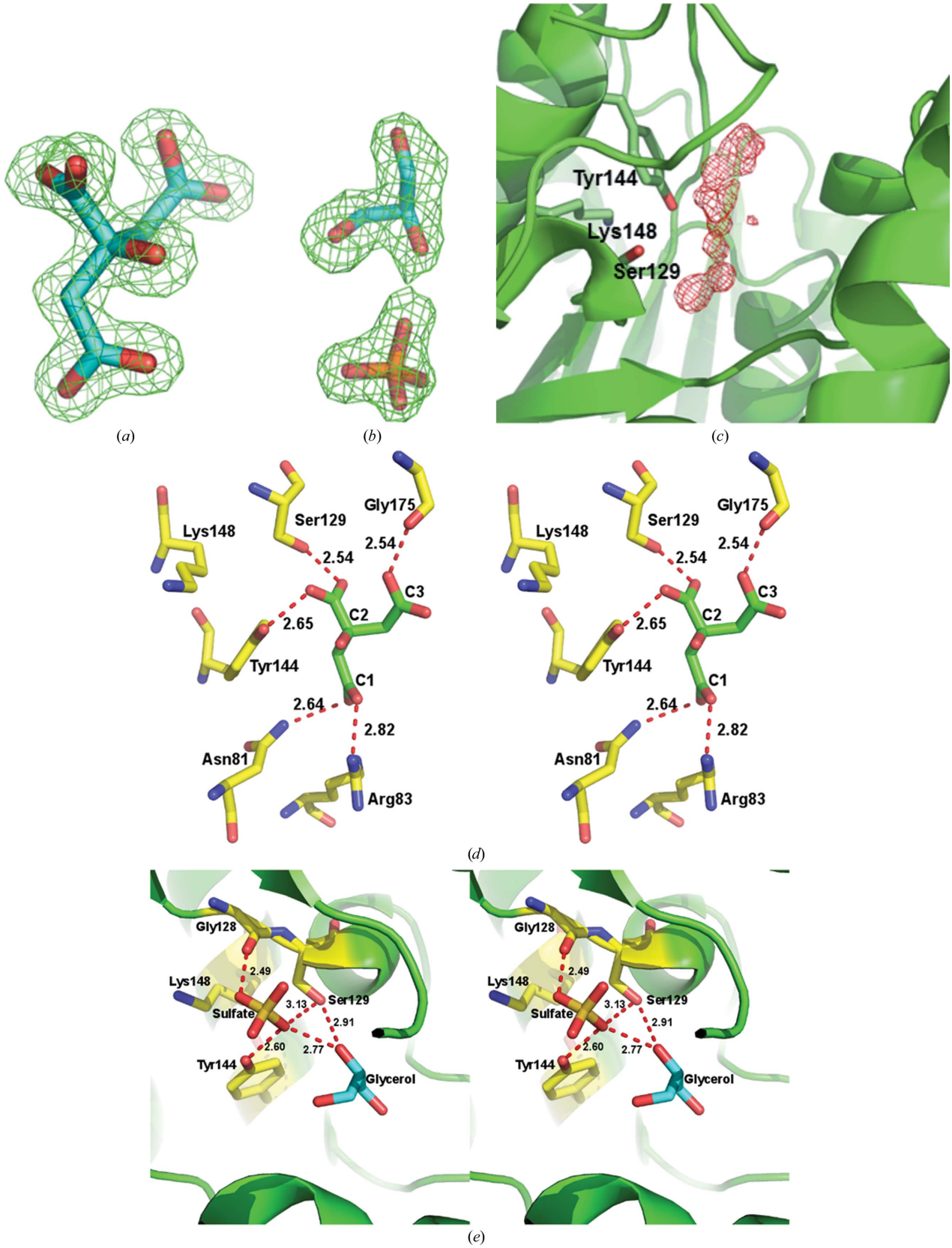


Figure 6
Binding of the citrate, glycerol and sulfate molecules at the active site of DDGDH. (*a*) Simulated OMIT map of citrate (bound at the active site) at 1.5 Å resolution at a 1.0σ contour level. (*b*) Simulated OMIT map of glycerol and sulfate ion (bound to the apo form II crystal) at 1.8 Å resolution at a 0.9σ contour level. (*c*) $F_o - F_c$ negative electron-density map at the active site of the apo form I crystal at 1.15 Å resolution. Catalytic triad residues are labelled. (*d*) Residues around the citrate molecule. Important hydrogen bonds are indicated by broken lines. (*e*) Residues around the glycerol and sulfate ion (bound to apo form I). Catalytic triad residues are labelled.

map in the vicinity of the active site of apo form I, indicating the presence of substrate (Fig. 6*c*). However, we could not model the substrate (only the hydroxyl group was very clear) owing to a lack of clear density for the substrate.

3.7. Substrate recognition and catalytic machinary

To understand how the substrate is recognized by *T. thermophilus* DDGDH, we made an attempt to fit the substrate KDG (the chemical structures of DDG and KDG are highly similar) into the active-site pocket on the basis of the binding of citrate, glycerol and sulfate molecules. In our modelling efforts, it was clear that the substrate fits well into the substrate-binding pocket (Ohshima *et al.*, 2004). The dehydrogenase activity of the SDR family enzymes has been well studied (Jörnvall *et al.*, 1995; Ghosh *et al.*, 2001; Filling *et al.*, 2002). There is a conserved catalytic triad, Ser129–Tyr144–Lys148, in the catalytic site, with the tyrosine residue (Tyr144) functioning as the catalytic base abstracting a proton from the substrate. The serine (Ser129) plays a role in the stabilization of the bound substrate, and the lysine (Lys148) directly interacts with the hydroxyl of the tyrosine residue through the 2′-hydroxyl of the nicotinamide ribose of the cofactor and lowers its pK_a value (Jörnvall *et al.*, 1995; Ghosh *et al.*, 2001; Filling *et al.*, 2002). This catalytic triad is structurally well conserved in DDGDH, suggesting that DDGDH follows a similar catalytic mechanism. We were not able to obtain the complex with KDG. However, citrate, glycerol and sulfate molecules binding at the active site provided the structural environment at the substrate-binding site. Since no structural changes were observed between the native structure and those of the NAD^+ complex and the citrate complex, we could construct a putative model of the DDGDH–citrate–NAD^+ ternary complex by superimposing the two structures (the r.m.s.d is 0.19 Å). Moreover, the cofactor-complex and citrate-complex crystals are isomorphous and share the same space group, even though they were crystallized in different conditions. The chemical structures of citrate and DDG are highly similar (Fig. 7*a*), and thus citrate may be regarded as a substrate. The initiation of the catalytic reaction is the abstraction of a proton from the C1 hydroxyl of the citrate by Tyr144 in combination with Ser129 and Lys148. Hydride transfer subsequently occurs from the C1 hydroxyl of the citrate molecule to C4 of the nicotinamide ring of the cofactor NAD^+ directly (Fig. 7*b*). In the putative model of the ternary complex, the distance between the C1 hydroxyl of the citrate molecule and C4 of the nicotinamide ring is approximately 3.9 Å and the N—C4—C1 angle is 115°. These values are consistent with those obtained from theoretical considerations of hydride transfer and those observed in a number of struc-

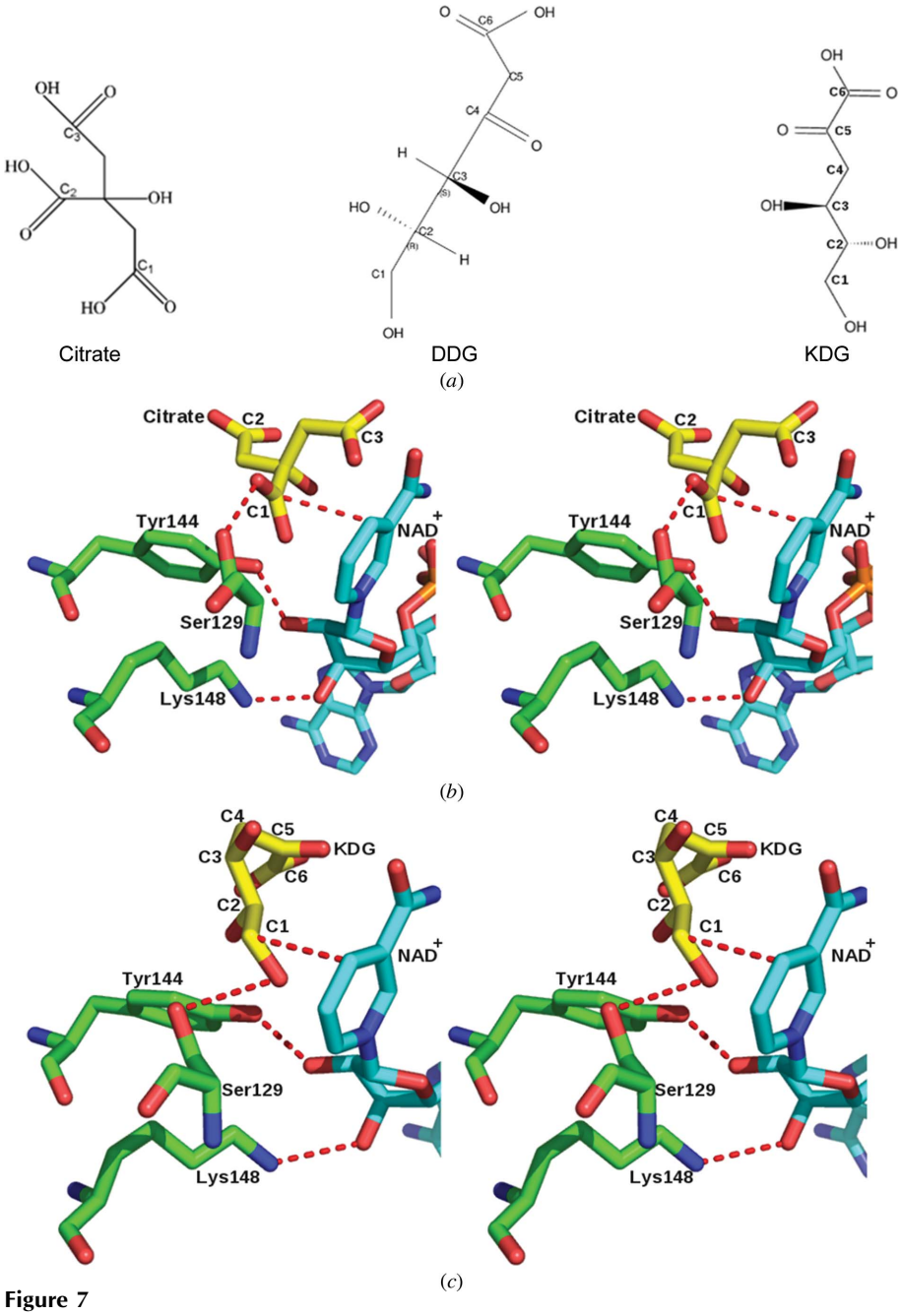


Figure 7
A proposed model for substrate binding to *T. thermophilus* DDGDH. (*a*) Chemical structures of citrate, 2,5-diketo-3-deoxygluconate (DDG) and 2-keto-3-deoxygluconate (KDG). (*b*) Stereoview (modelled) of the binding of citrate and NAD^+ to DDGDH. (*c*) Stereoview (modelled) of the binding of 2-keto-3-deoxygluconate to NAD-dependent DDGDH. Ser129, Tyr144 and Lys148 in this figure constitute the catalytic triad in *T. thermophilus* DDGDH.

tures of NAD(P)-dependent enzymes (Wu & Houk, 1991). In a similar fashion, our modelled substrate KDG also has identical bond lengths and angles to the cofactor and the catalytic triad, suggesting a similar catalytic mechanism (Fig. 7*c*).

In this study, we determined the three-dimensional crystal structure of *T. thermophilus* NAD-dependent DDGDH and its complex with citrate for the first time. The cofactor NAD^+ molecule was bound at the cleft between the two adjacent subunits. The citrate molecule is located in the active site (catalytic triad). It is likely that citrate and the substrate KDG share the same binding pocket. The observed intersubunit active site comprising both subunits of the *AB* dimer suggests the catalytic importance of the oligomeric state. The active site and the cofactor-binding site have been identified deep in the cleft formed between the intersubunit interface. The identification of the active site and structural comparison with a homologous enzyme suggests that the SDR scaffold with the Ser129–Tyr144–Lys148 catalytic triad is conserved. The observed binding mode of the citrate molecule and cofactor provide insights into the substrate recognition and catalytic machinery of DDGDH.

The authors would like to thank IOE, University of Mysore for providing the X-ray diffraction facility. The authors would also like to thank the staff of the RIKEN Genomic Sciences Center for providing the plasmid and the beamline staff for assistance during data collection at beamline BL26B1 of SPring-8, Japan.

References

Andersson, A., Jordan, D., Schneider, G. & Lindqvist, Y. (1996). *Structure*, **4**, 1161–1170.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Didierjean, C., Rahuel-Clermont, S., Vitoux, B., Dideberg, O., Branlant, G. & Aubry, A. (1997). *J. Mol. Biol.* **268**, 739–759.
Filling, C., Berndt, K. D., Benach, J., Knapp, S., Prozorovski, T., Nordling, E., Ladenstein, R., Jörnvall, H. & Oppermann, U. (2002). *J. Biol. Chem.* **277**, 25677–25684.
Ghosh, D., Sawicki, M., Pletnev, V., Erman, M., Ohno, S., Nakajin, S. & Duax, W. L. (2001). *J. Biol. Chem.* **276**, 18457–18463.
Ghosh, D., Wawrzak, Z., Weeks, C. M., Duax, W. L. & Erman, M. (1994). *Structure*, **2**, 629–640.
Gouet, P., Courcelle, E., Stuart, D. I. & Métoz, F. (1999). *Bioinformatics*, **15**, 305–308.
Gourley, D. G., Schüttelkopf, A., Leonard, G., Luba, J., Hardy, L., Beverley, S. & Hunter, W. N. (2001). *Nature Struct. Biol.* **8**, 521–525.
Haapalainen, A. M., Koski, M. K., Qin, Y.-M., Hiltunen, J. K. & Glumoff, T. (2003). *Structure*, **11**, 87–97.
Holm, L. & Sander, C. (1993). *J. Mol. Biol.* **233**, 123–138.
Hulsmeyer, M., Hecht, H.-J., Niefind, K., Hofer, B., Eltis, L. D., Timmis, K. N. & Schomburg, D. (1998). *Protein Sci.* **7**, 1286–1293.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
John, J., Crennell, S. J., Hough, D. W., Danson, M. J. & Taylor, G. L. (1994). *Structure*, **2**, 358–393.
Jong, R. M. de, Tiesinga, J. J. W., Villa, A., Tang, L., Janssen, D. B. & Dijkstra, B. W. (2005). *J. Am. Chem. Soc.* **127**, 13338–13343.
Jörnvall, H., Persson, B. & Jeffery, J. (1987). *Eur. J. Biochem.* **167**, 195–201.
Jörnvall, H., Persson, B., Krook, M., Atrian, S., Gonzàlez-Duarte, R., Jeffery, J. & Ghosh, D. (1995). *Biochemistry*, **34**, 6003–6013.
Kabsch, W. (1976). *Acta Cryst.* A**32**, 922–923.
Kallberg, Y., Oppermann, U., Jörnvall, H. & Persson, B. (2002). *Protein Sci.* **11**, 636–641.
Kleywegt, G. J. & Jones, T. A. (1998). *Acta Cryst.* D**54**, 1119–1131.
Korozowski, Z. (1992). *Mol. Cell. Endocrinol.* **84**, C25–C31.
Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lokanath, N. K., Shiromizu, I., Ohshima, N., Nodake, Y., Sugahara, M., Yokoyama, S., Kuramitsu, S., Miyano, M. & Kunishima, N. (2004). *Acta Cryst.* D**60**, 1816–1823.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Ohshima, N., Inagaki, E., Yasuike, K., Takio, K. & Tahirov, T. H. (2004). *J. Mol. Biol.* **340**, 477–489.
Oppermann, U., Filling, C., Hult, M., Shafqat, N., Wu, X., Lindh, M., Shafqat, J., Nordling, E., Kallberg, Y., Persson, B. & Jörnvall, H. (2003). *Chem. Biol. Interact.* **143–144**, 247–253.
Oppermann, U. C., Filling, C. & Jörnvall, H. (2001). *Chem. Biol. Interact.* **130–132**, 699–705.
Otwinowski, Z. & Minor, M. (1997). *Methods Enzymol.* **276**, 307–326.
Powell, A. J., Read, J. A., Banfield, M. J., Gunn-Moore, F., Yan, S. D., Lustbader, J., Stern, A. R., Stern, D. M. & Brady, R. L. (2000). *J. Mol. Biol.* **303**, 311–327.
Rossmann, M. G., Adams, M. J., Buehner, M., Ford, G. C., Hackert, M. L., Liljas, A., Rao, S. T., Banaszak, L. J., Hill, E., Tsernoglou, D. & Webb, L. (1973). *J. Mol. Biol.* **76**, 533–537.
Rossmann, M. G., Moras, D. & Olsen, K. W. (1974). *Nature (London)*, **250**, 194–199.
Stammers, D. K., Ren, J., Leslie, K., Nichols, C. E., Lamb, H. K., Cocklin, S., Dodds, A. & Hawkins, A. R. (2001). *EMBO J.* **20**, 6619–6626.
Tanaka, N., Nonaka, T., Nakanishi, M., Deyashiki, Y., Hara, A. & Mitsui, Y. (1996). *Structure*, **4**, 33–45.
Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). *Nucl. Acids Res.* **25**, 4876–4882.
Varughese, K. I., Skinner, M. M., Whiteley, J. M., Matthews, D. A. & Xuong, N. H. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 6080–6084.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Wu, Y.-D. & Houk, K. N. (1991). *J. Am. Chem. Soc.* **113**, 2353–2358.