Extension of resolution and oligomerization-state studies of 2,4'-dihydroxyacetophenone dioxygenase from *Alcaligenes* sp. 4HAP

J. Guo,^a P. Erskine,^a A. R. Coker,^a J. Gor,^b S. J. Perkins,^b S. P. Wood^a and J. B. Cooper^{a*}

Received 29 May 2015

Accepted 25 August 2015

Edited by R. A. Pauptit, Macclesfield, England

Keywords: 2,4'-dihydroxyacetophenone dioxygenase; protein X-ray structure; oligomerization; analytical ultracentrifugation; gel filtration; docking; dioxygenase.

PDB reference: 2,4'-dihydroxyacetophenone dioxygenase, 5bpx

Supporting information: this article has supporting information at journals.iucr.org/f

^aLaboratory of Protein Crystallography, Centre for Amyloidosis and Acute Phase Proteins, UCL Division of Medicine (Royal Free Campus), Rowland Hill Street, London NW2 3PF, England, and ^bDepartment of Structural and Molecular Biology, Division of Biosciences, University College London, Darwin Building, Gower Street, London WC1E 6BT, England. *Correspondence e-mail: jon.cooper@ucl.ac.uk

The enzyme 2,4'-dihydroxyacetophenone dioxygenase (DAD) catalyses the conversion of 2,4'-dihydroxyacetophenone to 4-hydroxybenzoic acid and formic acid. This enzyme is a very unusual dioxygenase in that it cleaves a C—C bond in a substituent of the aromatic ring rather than within the ring itself. Whilst it has been shown that DAD is a tetramer in solution, the recently solved crystal structure of the *Alcaligenes* sp. 4HAP enzyme was in fact dimeric rather than tetrameric. Since the use of limited chymotrypsinolysis, which apparently results in removal of the first 20 or so N-terminal residues of DAD, was necessary for crystallization of the protein, it was investigated whether this was responsible for the change in its oligomerization state. Gel-filtration and analytical ultracentrifugation studies were conducted, which confirmed that chymotrypsinolysed DAD has an apparent molecular weight of around 40 kDa, corresponding to a dimer. In contrast, the native enzyme has a molecular weight in the 70–80 kDa region, as expected for the tetramer. The structural basis for tetramerization has been investigated by the use of several docking servers, and the results are remarkably consistent with the tetrameric structure of a homologous cupin protein from *Ralstonia eutropha* (PDB entry 3ebr).

1. Introduction

The enzyme 2,4'-dihydroxyacetophenone dioxygenase (DAD; EC 1.13.11.41), which was first identified by Hopper (1986) in a strain of the Gram-negative, aerobic soil bacterium *Alcaligenes* sp. 4HAP, catalyses a reaction in which 2,4'-dihydroxyacetophenone is converted to 4-hydroxybenzoic acid and formic acid (Fig. 1). The metabolism of aromatic ring compounds that is catalysed by dioxygenases usually involves the incorporation of two O atoms from a single molecule of O₂ into the aromatic ring of the substrate (Hayaishi *et al.*, 1955). An abundance of aromatic compounds occur naturally in the environment, including lignin, which is the precursor of 2,4'-dihydroxyacetophenone, as well as pollutants which are generated synthetically (Díaz *et al.*, 2001). Dioxygenases such as DAD, expressed in appropriately engineered microorganisms, have the potential to use these aromatic pollutants as a carbon source and thus remove them from the environment (Bugg *et al.*, 2011). Indeed, the bacterial species from which the enzyme originates (*Alcaligenes*) has been used industrially for biotransformations in the production of unusual amino acids, polyhydroxybutanol and polyhydroxybutyrates for use in biodegradable plastics (Nair *et al.*, 2009).

© 2015 International Union of Crystallography

DAD from *Alcaligenes* sp. 4HAP has been shown to be a homotetramer in which each subunit has a molecular mass of 20.3 kDa and consists of 177 amino acids (Hopper & Kaderbhai, 1999). Recently, structural studies have shown that it is a member of the cupin superfamily, which is characterized by a conserved β -barrel fold (Beaven *et al.*, 2014; Keegan *et al.*, 2014). The core cupin fold of DAD is formed by residues 60–132, which assemble into two four-stranded antiparallel β -sheets, forming a double-stranded β -helix that is extended at one end by an additional β -hairpin in each sheet. The enzyme possesses a catalytic ferrous ion that is tightly bound in the active site at the centre of the barrel by the three invariant histidine residues: His76, His78 and His114. The putative binding site for the aromatic substrate is buried and is mainly composed of conserved hydrophobic residues. The two β -hairpins forming one end of the barrel are somewhat splayed apart, exposing an unusual hydrophobic patch which might be the entry point for substrate (Beaven *et al.*, 2014; Keegan *et al.*, 2014).

In this paper, we report the refinement of the *Alcaligenes* sp. 4HAP DAD structure to a higher resolution than has been achieved previously. These new data suggest that an unidentified ligand which coordinates the catalytic iron is probably

an acetate group rather than a carbonate dianion as had been suggested earlier. An additional ligand, which binds in the position that the aromatic substrate would be expected to occupy, appears to be a glycerol molecule, although this assignment remains slightly ambiguous. We present evidence that the limited chymotrypsinolysis which is needed to crystallize the enzyme changes it from a tetramer to a dimer. Finally, docking studies and inspection of the structure of a known homologue of DAD suggest a model of the tetrameric form of the enzyme.

2. Crystallization of DAD

The expression of *Alcaligenes* sp. 4HAP DAD in *Escherichia coli* BL21 (DE3) cells and nickel-affinity purification were conducted as previously described by Beaven *et al.* (2014). The protein was concentrated to 5 mg ml⁻¹ using a Vivaspin centrifugal concentrator (10 kDa cutoff, GE Healthcare) in 50 mM Tris pH 7.3, 100 mM NaCl. The concentrated protein had a grey colour, which was also found with *Burkholderia* sp. AZ11 DAD (Enya *et al.*, 2014). Whilst no efforts were made to specifically remove the His tag using thrombin, it had previously been found that it was necessary to add chymotrypsin to DAD in a ratio of 1:100(w:w) before setting up the crystallization trials. Screening for crystallization conditions was then performed using the hanging-drop method at 4, 12 and 21°C with the Structure Screen 1 + 2 kits from Molecular Dimensions. Needle-like crystals started to appear after 3 d in several conditions. Subsequent optimization based on condition 12 from Structure Screen 1 (0.1 M sodium cacodylate pH 6.5, 1.4 M sodium acetate) revealed that better crystals than had previously been obtained could be grown reproducibly in 0.1 M sodium cacodylate pH 5.9–6.8 and 1.1–1.6 M sodium acetate. Further optimization by incubation of the enzyme with different concentrations of the substrate (2,4'-dihydroxyacetophenone), the product (4-hydroxybenzoic acid) and 4-aminobenzamidine (a structural analogue of the substrate) prior to setting up the hanging-drop plates indicated that better crystals could be obtained by the addition of 2 mM 4-hydroxybenzoic acid or 0.1 mM 4-aminobenzamidine (Fig. 2). The choice of 4-aminobenzamidine was driven by our

Figure 1
The reaction catalysed by 2,4'-dihydroxyacetophenone dioxygenase (DAD). The enzyme has a high affinity for oxygen, which is used for C–C bond cleavage of the substrate 2,4'-dihydroxyacetophenone, yielding 4-hydroxybenzoic acid and formic acid.

Figure 2
Crystals of chymotrypsinolysed *Alcaligenes* sp. 4HAP DAD. The needle-like crystals are up to 1.5 mm in length and are around 20 μ m thick.

Table 1

Data-collection and processing statistics for *Alcaligenes* sp. 4HAP DAD.

Values in parentheses are for the outer resolution shell.

Beamline	I04-1, DLS
Wavelength (Å)	0.92
Space group	<i>P</i> 6 ₅ 22
Unit-cell parameters	
<i>a</i> = <i>b</i> (Å)	83.25
<i>c</i> (Å)	114.28
Mosaic spread (°)	0.085
Resolution (Å)	41.62–1.88 (1.93–1.88)
<i>R</i> _{merge} [†] (%)	12.2 (80.0)
<i>R</i> _{meas} [‡] (%)	12.5 (82.0)
CC _{1/2} [§] (%)	99.9 (92.7)
Completeness (%)	99.5 (99.0)
Average <i>I</i> /σ(<i>I</i>)	21.8 (5.0)
Multiplicity	20.3 (19.5)
No. of observed reflections	396011 (26985)
No. of unique reflections	19556 (1385)
Wilson-plot <i>B</i> factor (Å ²)	20.5
Solvent content (%)	56.3
<i>R</i> factor (%)	14.2
Free <i>R</i> factor (%)	17.2
R.m.s.d., bond lengths (Å)	0.025
R.m.s.d., bond angles (°)	2.233
No. of reflections in working set	19555
No. of reflections in test set	940
Mean protein <i>B</i> factor (Å ²)	24.7

[†] $R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$, where $\langle I(hkl) \rangle$ is the mean intensity of the *i* observations *I_i*(*hkl*) of each unique reflection *hkl* after scaling. [‡] $R_{\text{meas}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$, where $\langle I(hkl) \rangle$ is the mean intensity of the *i* observations *I_i*(*hkl*) of each unique reflection *hkl* after scaling. [§] CC_{1/2} is the half-set correlation coefficient, as described by Karplus & Diederichs (2012).

earlier finding that the thrombin-cleaved His-tagged protein could not be eluted from a benzamidine column in nondenaturing conditions. Selected crystals were mounted in loops with approximately 30% glycerol as the cryoprotectant before flash-cooling in a nitrogen-gas cryostream (Oxford Cryosystems cryocooler) at 100 K.

3. X-ray data collection and processing

X-ray data collection was performed at station I04-1 at the Diamond Light Source (DLS), Didcot, England. Compared with the DAD crystals reported by Beaven *et al.* (2014), which generally diffracted synchrotron radiation to at best 3.0–2.5 Å resolution, with one favourable case giving diffraction to 2.2 Å resolution, many of the new crystals diffracted to resolutions in the range 2.5–1.8 Å. For the crystal giving the best data set described in this paper, 190° of data were collected using a wavelength of 0.92 Å and an exposure time of 1 s per image. The data were processed to a resolution of 1.88 Å using *iMosflm* (Battye *et al.*, 2011; Powell *et al.*, 2013), *SCALA* (Evans, 2006) and other programs in the *CCP4* suite (Winn *et al.*, 2011). This revealed that the crystal was hexagonal, with unit-cell parameters *a* = *b* = 83.3, *c* = 114.3 Å, and belonged to the same space group (*P*6₅22) as the previously published structure. The data set had overall *R*_{merge} and *R*_{meas} values of 12.2 and 12.8%, respectively. The program *MATTHEWS_COEF* (Matthews, 1968; Kantardjiev & Rupp, 2003) revealed that the solvent content was 56.3%. Other statistics are shown in Table 1.

4. Structure determination and analysis

Since the crystal form was isomorphous with that reported previously, the structure was subjected to several rounds of refinement and manual rebuilding with the new data set using *REFMAC* (Murshudov *et al.*, 2011) and *Coot* (Emsley *et al.*, 2010), which decreased the final *R* factor and *R*_{free} values to 14.2 and 17.2%, with 143 water molecules built in the model. Least-squares superposition of the new structure with the previous structure (Fig. 3) gave an r.m.s. deviation of 0.27 Å for the 150 equivalent C^α atoms.

The new data set allowed three more residues (Leu24, Pro25 and Glu26) at the N-terminal end of the protein to be built into the electron-density map. The preceding residues are all presumed to be disordered or missing owing to the digestion caused by the chymotrypsin which was present during crystallization. The newly visible residues were all well defined by the electron-density map for the new higher resolution structure. For the remainder of the molecule, continuous electron density of high quality was visible up to and including the C-terminal Arg176. Residues 139–143 which connect β-strands C1 and C2 were poorly defined in the previous electron-density map at 2.2 Å resolution. In contrast, in the higher resolution structure they appear to be extremely well delineated as a three-residue β-hairpin turn with glycine at the third residue position.

5. Oligomerization-state determination

5.1. Gel filtration

The oligomeric state of both full-length DAD and chymotrypsinolysed DAD in solution was determined by gel-filtration chromatography with a Superdex 75 column (GE Healthcare) and the collected fractions were analysed by SDS-PAGE. The chymotrypsinolysed DAD samples were

Figure 3

A superposition of the 1.88 Å resolution structure (yellow) with the previous model solved at 2.2 Å resolution (green). The arrow indicates the only region (residues 139–143) where there are significant differences, which are probably owing to this loop being poorly defined in the previous structure.

prepared by incubation of DAD with chymotrypsin at a 100:1(*w:w*) ratio at room temperature overnight. Three proteins were used as molecular-weight standard markers to calibrate the column: albumin (66 kDa), ovalbumin (44 kDa) and horse heart cytochrome *c* (13.4 kDa) (Amersham Biosciences). The full-length DAD protein was found to elute at approximately the same point as albumin, presumably because the resolution of the column was not high enough to separate these proteins (see Supporting Information). However, it is clear that the full-length DAD enzyme elutes much earlier than ovalbumin, thus confirming earlier findings that it is likely to be a tetramer of approximately 80 kDa in solution (Hopper & Kaderbhai, 1999; Enya *et al.*, 2014). In contrast, the chymotrypsinolysed DAD protein eluted from the column between albumin and cytochrome *c*, at approximately the same position as ovalbumin. This suggests that the chymotrypsinolysed DAD is actually a dimer in solution, which is in accord with the previously published X-ray crystal structure (Keegan *et al.*, 2014) and the higher resolution structure reported in this paper.

5.2. Analytical ultracentrifugation

Analytical ultracentrifugation data were collected using a Beckman ProteomeLab XL-I ultracentrifuge (Beckman Coulter, High Wycombe, England) with an An-50 Ti rotor. Following extensive dialysis of the stock protein solutions with 50 mM Tris, 100 mM NaCl pH 7.3, six different concentrations of native and chymotrypsinolysed DAD (0.1, 0.2, 0.4, 0.6, 0.8 and 1.0 mg ml⁻¹) and a buffer blank were prepared. Sedimentation-velocity data collection was then undertaken at 20°C using a rotor speed of 50 000 rev min⁻¹. 100 absorbance and interference scans were obtained from each sample in two-sector cells with column heights of 12 mm. Sedimentation analysis was undertaken using *SEDFIT* v.14.81 (Schuck, 2000) using direct boundary Lamm fits. The buffer density of 0.99823 g ml⁻¹ and viscosity of 0.01002 cP were calculated using the *SEDNTERP* web server (Laue *et al.*, 1992). Optimization of the size-distribution analysis *c(s)* fit was performed for each sample until the overall r.m.s.d. and visual appearance indicated that a satisfactory fit had been obtained. A representative fit which demonstrates the good agreement between the experimental boundary scans and the fitted lines is presented in the Supporting Information. Data analysis by *SEDFIT* reveals a major *c(s)* peak for native DAD at about 4.8 S, which corresponds to a molecular mass of 64.4 ± 0.7 kDa. However, the major *c(s)* peak of the chymotrypsinolysed DAD is at around 3.0 S, which yields a molecular mass of 31.5 ± 0.9 kDa. The results were consistently seen at the different concentrations and their mean value gave final sedimentation coefficients of $s_{20,w}^0 = 4.86 \pm 0.03$ S for native DAD and $s_{20,w}^0 = 2.96 \pm 0.06$ S for chymotrypsinolysed DAD. This outcome indicates that the native DAD is a tetramer whilst the chymotrypsinolysed enzyme is actually a dimer, although the molecular masses obtained are both somewhat smaller than expected. This could be owing to a difference in the volume and the shape of native and chymotrypsinolysed

DAD, which in turn affects the partial specific volume v (0.729287). The small peaks at about 5.0 and 2.7 S in Fig. 4 indicate that small amounts of tetramer and dimer might be present in the chymotrypsinolysed and native DAD, respectively.

6. Quaternary-structure prediction

The crystallographically observed dimer of DAD involves a hydrophobic interface composed of ten β -strands, five in each monomer, which contribute a large number of aromatic residues to the monomer interface (Keegan *et al.*, 2014). Since the dimeric structure of DAD is quite similar to the constituent dimers of the tetramer formed by the RmlC-like protein from *Ralstonia eutropha* (PDB entry 3ebr; Joint Center for Structural Genomics, unpublished work), it is possible that the full-length DAD enzyme also adopts a similar tetrameric structure. To explore this possibility, three docking servers, *ZDOCK*

Figure 4 The size-distribution analysis *c(s)* shows clearly that the native DAD is a tetramer with a larger sedimentation coefficient $s_{20,w}^0$ of 4.86 S, while the chymotrypsinolysed DAD is a dimer with a smaller $s_{20,w}^0$ of 2.96 S. The lower peaks indicate that a small amount of the dimer and tetramer might still exist in the native and chymotrypsinolysed DAD, respectively.

(Pierce *et al.*, 2014), *SymmDock* (Schneidman-Duhovny *et al.*, 2005*a,b*) and *ClusPro* (Comeau *et al.*, 2004*a,b*; Kozakov *et al.*, 2006, 2013), were used for theoretical docking studies using a pair of DAD dimers. Intriguingly, the top solution with the lowest energy from each server was found to have a similar tetrameric structure to that of PDB entry 3ebr (Fig. 5). These models of the DAD tetramer were analysed using the *PDBEPIA* server (Krissinel & Henrick, 2005, 2007; Krissinel, 2010), as was the 3ebr structure. The interface areas, Δ^iG values and Δ^iG *P*-values of the chains of each dimer (*A* and *B*, and *C* and *D*) suggest that there is a stable tetramer interface formed by chains *A* and *D* and by chains *B* and *C* in the

Table 2
Interface parameters for the putative DAD tetramer obtained from the *PDBEPIA* server.

The values are given for three docking models and the homologous structure with PDB code 3ebr. Note that the equivalent *A–B* and *C–D* interfaces are those already present within the crystallographically observed dimer of DAD. The putative tetramer interface is formed by the *A* and *D* chains and by the *B* and *C* chains. Note that the *A–B* and *C–D* interface areas are appreciably different in the *ClusPro* model since this is the only docking program of the three used here which energy-minimizes the final model of the complex.

Models	Chains	Interface area (Å ²)	Δ^iG (kcal mol ^{−1})	Δ^iG <i>P</i> -value	<i>N</i> _{HB} [†]	<i>N</i> _{SB} [‡]
<i>ClusPro</i>	<i>A–B</i> , <i>C–D</i>	1366	−17.1	0.343	28	9
	<i>A–D</i> , <i>B–C</i>	642	−4.3	0.683	6	0
<i>ZDOCK</i>	<i>A–B</i> , <i>C–D</i>	1431	−14.0	0.198	24	8
	<i>A–D</i> , <i>B–C</i>	718	−3.7	0.539	7	1
<i>SymmDock</i>	<i>A–B</i> , <i>C–D</i>	1439	−14.0	0.201	24	8
	<i>A–D</i> , <i>B–C</i>	611	−4.4	0.451	3	0
3ebr	<i>A–B</i> , <i>C–D</i>	1322	−23.0	0.104	16	4
	<i>A–D</i> , <i>B–C</i>	695	−4.3	0.699	12	4

[†] Number of hydrogen bonds. [‡] Number of salt bridges.

docked structures, albeit with fewer hydrogen bonds than in the 3ebr tetramer (Table 2). This putative tetrameric interface of the DAD enzymes involves a number of conserved, hydrophobic residues.

It is possible that the 20 or so N-terminal residues of DAD which are not visible in the structure of the chymotrypsinolyzed enzyme form a number of hydrogen bonds and/or salt bridges which contribute to maintaining the tetrameric structure of the full-length DAD. Thus, the loss of these residues during chymotrypsinolysis, which appears to be essential for crystallization, might cause the disassembly of the DAD tetramer into two dimers. Although the first few N-terminal residues in 3ebr do not contribute hydrogen bonds or salt

Figure 5
(*a*) A superposition of the models of the tetramer obtained by docking with the structure of the RmlC-like cupin protein from *R. eutropha* (PDB entry 3ebr). The structures are coloured as follows: red, PDB entry 3ebr; green, *ClusPro* model; blue, *SymmDock* model; yellow, *ZDOCK* model. (*b*) The tetrameric structure of DAD generated using the *ClusPro* docking server with pairs of monomers forming the constituent dimers coloured similarly. (*c*) A surface representation of the docked DAD dimers forming the putative tetramer. Again, the two monomers forming each constituent dimer are coloured similarly in green and faint green and in red and faint red.

Figure 6

The electron density in the active-site region of the DAD. The iron-binding histidines and the acetate ligand are shown along with a selection of the key aromatic residues forming the substrate-binding site, which is occupied by a putative glycerol molecule. The $2F_o - F_c$ electron density is contoured at 1.0 r.m.s..

bridges to the tetramer interface, there are many more residues in the longer N-terminal region of the DAD enzyme and these could therefore play an important role in tetramerization.

7. Catalytic site

The active site of the newly refined structure has well defined electron density for most of the residues which form the iron-binding and substrate-binding sites, where only minor conformational adjustments to the side chains of the earlier model were required. The catalytic iron is bound by three histidines (His76, His78 and His114) along with a putative anionic group that we have previously speculated might be a carbonate dianion (Keegan *et al.*, 2014). Following refinement with various possible anions in the same position, we have reinterpreted this group as being an acetate anion, since this gives a very satisfactory fit to the high-resolution electron-density maps (Fig. 6) and acetate is present in high concentration in the crystallization conditions. In addition, the methyl group of the acetate is located in a hydrophobic pocket formed by Leu65, Leu116 and Phe129. In the new structure, there is improved electron density for Glu108, which we speculated may be involved in the entry of the diatomic oxygen substrate (Keegan *et al.*, 2014). There is an additional ligand in the pocket where the aromatic substrate moiety probably binds, which we have modelled as a glycerol molecule, although this interpretation is somewhat speculative since its electron density is slightly sparse. The higher reso-

lution structure and data set have been deposited in the PDB as entry 5bpx.

Acknowledgements

We gratefully acknowledge the Diamond Light Source for X-ray beam time and travel support for data collection (award MX-8922).

References

- Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* **D67**, 271–281.
- Beaven, G., Bowyer, A., Erskine, P., Wood, S. P., McCoy, A., Coates, L., Keegan, R., Lebedev, A., Hopper, D. J., Kaderbhai, M. A. & Cooper, J. B. (2014). *Acta Cryst.* **F70**, 823–826.
- Bugg, T. D., Ahmad, M., Hardiman, E. M. & Singh, R. (2011). *Curr. Opin. Biotechnol.* **22**, 394–400.
- Comeau, S. R., Gatchell, D. W., Vajda, S. & Camacho, C. J. (2004a). *Bioinformatics*, **20**, 45–50.
- Comeau, S. R., Gatchell, D. W., Vajda, S. & Camacho, C. J. (2004b). *Nucleic Acids Res.* **32**, W96–W99.
- Díaz, E., Ferrández, A., Prieto, M. A. & García, J. L. (2001). *Microbiol. Mol. Biol. Rev.* **65**, 523–569.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Enya, M., Aoyagi, K., Hishikawa, Y., Yoshimura, A., Mitsukura, K. & Maruyama, K. (2014). *Biosci. Biotechnol. Biochem.* **76**, 567–574.
- Evans, P. (2006). *Acta Cryst.* **D62**, 72–82.
- Hayaishi, O., Katagiri, M. & Rothberg, S. (1955). *J. Am. Chem. Soc.* **77**, 5450–5451.
- Hopper, D. J. (1986). *Biochem. J.* **239**, 469–472.
- Hopper, D. J. & Kaderbhai, M. A. (1999). *Biochem. J.* **344**, 397–402.
- Kantardjieff, K. A. & Rupp, B. (2003). *Protein Sci.* **12**, 1865–1871.
- Karplus, P. A. & Diederichs, K. (2012). *Science*, **336**, 1030–1033.
- Keegan, R., Lebedev, A., Erskine, P., Guo, J., Wood, S. P., Hopper, D. J., Rigby, S. E. J. & Cooper, J. B. (2014). *Acta Cryst.* **D70**, 2444–2454.
- Kozakov, D., Beglov, D., Bohnuud, T., Mottarella, S. E., Xia, B., Hall, D. R. & Vajda, S. (2013). *Proteins*, **81**, 2159–2166.
- Kozakov, D., Brenke, R., Comeau, S. R. & Vajda, S. (2006). *Proteins*, **65**, 392–406.
- Krissinel, E. (2010). *J. Comput. Chem.* **31**, 133–143.
- Krissinel, E. & Henrick, K. (2005). *Computational Life Sciences*, edited M. R. Berthold, R. C. Glen, K. Diederichs, O. Kohlbacher & I. Fischer, pp. 163–174. Berlin, Heidelberg: Springer.
- Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
- Laue, T. M., Shah, B. D., Ridgeway, T. M. & Pelletier, S. L. (1992). *Analytical Ultracentrifugation in Biochemistry and Polymer Science*, pp. 90–125. Cambridge: The Royal Society of Chemistry.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.
- Nair, I. C., Pradeep, S., Ajayan, M. S., Jayachandran, K. & Shashidhar, S. (2009). *Appl. Biochem. Biotechnol.* **159**, 545–552.
- Pierce, B. G., Wiehe, K., Hwang, H., Kim, B. H., Vreven, T. & Weng, Z. (2014). *Bioinformatics*, **30**, 1771–1773.
- Powell, H. R., Johnson, O. & Leslie, A. G. W. (2013). *Acta Cryst.* **D69**, 1195–1203.
- Schneidman-Duhovny, D., Inbar, Y., Nussinov, R. & Wolfson, H. J. (2005a). *Nucleic Acids Res.* **33**, W363–W367.
- Schneidman-Duhovny, D., Inbar, Y., Nussinov, R. & Wolfson, H. J. (2005b). *Proteins*, **60**, 224–231.
- Schuck, P. (2000). *Biophys. J.* **78**, 1606–1619.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.