Structure of a fungal form of aspartate semialdehyde dehydrogenase from *Cryptococcus neoformans*

Gopal Dahal and Ronald E. Viola*

Department of Chemistry and Biochemistry, University of Toledo, Toledo, OH 43606, USA. *Correspondence e-mail: ron.viola@utoledo.edu

Received 31 July 2015

Accepted 18 September 2015

Edited by W. N. Hunter, University of Dundee, Scotland

Keywords: aspartate semialdehyde dehydrogenase; *Cryptococcus neoformans*; fungal enzyme; drug target.

PDB reference: aspartate semialdehyde dehydrogenase, 5cef

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

Aspartate semialdehyde dehydrogenase (ASADH) functions at a critical junction in the aspartate-biosynthetic pathway and represents a valid target for antimicrobial drug design. This enzyme catalyzes the NADPH-dependent reductive dephosphorylation of β -aspartyl phosphate to produce the key intermediate aspartate semialdehyde. Production of this intermediate represents the first committed step in the biosynthesis of the essential amino acids methionine, isoleucine and threonine in fungi, and also the amino acid lysine in bacteria. The structure of a new fungal form of ASADH from *Cryptococcus neoformans* has been determined to 2.6 Å resolution. The overall structure of CnASADH is similar to those of its bacterial orthologs, but with some critical differences both in biological assembly and in secondary-structural features that can potentially be exploited for the development of species-selective drugs.

1. Introduction

The aspartate-biosynthetic pathway produces essential metabolites that are required for numerous important cellular functions in microorganisms, including the synthesis of the essential amino acids threonine, methionine, isoleucine and lysine (Viola *et al.*, 2011). Aspartate semialdehyde dehydrogenase (ASADH; EC 1.2.1.11) lies at the first branch point of this biosynthetic pathway, and catalyzes NADPH-dependent reductive dephosphorylation of aspartyl phosphate to produce aspartate semialdehyde (Black & Wright, 1955). The catalytic functional groups and substrate-binding groups are highly conserved in all of the ASADHs studied to date, despite the sequence homologies in this enzyme family ranging from greater than 90% to less than 10% (Gao *et al.*, 2010). Because this entire pathway is absent in mammals, but is absolutely essential for microbial survival, the enzymes that catalyze the reactions in this pathway are attractive potential targets for novel antimicrobial drug design (Pavlovsky *et al.*, 2012). ASADHs from a wide range of bacterial species have been biochemically, mechanistically and structurally characterized (Hadfield *et al.*, 2001; Blanco *et al.*, 2003; Faehnle *et al.*, 2006; Viola *et al.*, 2008). However, only a single structure from a fungal species has been reported to date (Arachea *et al.*, 2010), leaving many unanswered questions about the key differences between this enzyme in bacterial and fungal species. Intriguingly, despite the highly conserved set of active-site functional groups, significant differences in inhibitor-binding specificity have been observed between the only well characterized fungal ASADH from *Candida albicans* and its bacterial counterparts (Gao *et al.*, 2010). Advances have continued to be made in the development of selective inhibitors of the bacterial ASADHs (Pavlovsky *et al.*, 2014), but there has not

been any significant progress in the identification of fungal ASADH inhibitors.

We have now cloned, purified and determined the structure of another fungal form of ASADH from *Cryptococcus neoformans* (CnASADH). *C. neoformans* is an encapsulated basidiomycetous fungal pathogen that is responsible for causing cryptococcosis (Rodrigues *et al.*, 1999) and cryptococcal meningoencephalitis in immunocompromised, as well as immunocompetent, human populations (Kent & Juneann, 1998). Infections by *C. neoformans* can be lethal if untreated, and even treatment with the most potent antifungals can be fatal if the patient does not have sufficient T-cell-dependent immune function (Kent & Juneann, 1998). Some structural similarities and a few critical differences have been observed among this expanding group of fungal ASADHs, including the recently determined structure from *Trichophyton rubrum* (Q. Jin & S. Cui, personal communication).

2. Materials and methods

2.1. Cloning, expression and enzyme purification

The gene sequence encoding the full-length CnASADH enzyme from *C. neoformans* (Gene ID 3253595) was codon-optimized for bacterial expression and synthesized by Invitrogen (Life Technologies) with NcoI and XhoI sites at the N-terminus and C-terminus, respectively. The optimized gene was ligated into a pET-28a(+) vector (Novagen, USA) with a carboxy-terminal hexahistidine tag using these NcoI and XhoI restriction sites (Supplementary Table 1). The ligated plasmid was transformed into the DH5 α *Escherichia coli* cell line and the construct was extracted and purified from kanamycin-resistant colonies using a Miniprep plasmid-extraction kit (Qiagen, USA). The DNA sequence was confirmed by restriction digestion and by DNA sequencing. *E. coli* cell line BL21 (DE3) (Novagen, USA) was transformed with the gene-containing plasmid for expression of the recombinant CnASADH enzyme. The transformed cells were grown at 310 K in LB medium containing 30 $\mu\text{g ml}^{-1}$ kanamycin until an A_{600} of 0.6–0.8 was obtained. The cells were then induced by the addition of 0.5 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) and were grown for an additional 16 h at 293 K (optimum conditions for protein expression). The cells were harvested by centrifugation at 17 000g for 10 min. The cell pellet (5 g) was resuspended in buffer A consisting of 50 mM potassium phosphate pH 7.2, 200 mM NaCl, 25 mM imidazole, homogenized using a hand homogenizer and subsequently lysed by ultrasonication for a total of 8 min at 30 s intervals with a Fisher model 550 sonic dismembrator. Cell debris was removed by centrifugation at 12 000g for 30 min and the supernatant was filtered through a 28 mm 0.45 μm membrane syringe filter. The clarified soluble lysate was loaded onto a nickel–nitrilotriacetic acid (Ni–NTA) column and the target protein was eluted using a 25–400 mM imidazole gradient. Eluted fractions showing catalytic activity and a single band on a Coomassie-stained polyacrylamide gel were pooled and dialyzed overnight against a buffer consisting

of 20 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.0, 50 mM NaCl, 0.5 mM EDTA, 1 mM DTT. The dialyzed sample was concentrated to 10 mg ml $^{-1}$ by ultracentrifugation using a 10 kDa molecular-weight cutoff spin concentrator (Millipore). The identity of the enzyme was confirmed by MALDI-MS (described below) and an anti-His Western blot protocol. The protein concentration was measured using a Nanodrop 2000 spectrophotometer and the enzyme was stored at 253 K until use. Cloning details are summarized in the Supporting Information.

2.2. Protein characterization

The protein band corresponding to the expected size of ASADH was excised from the SDS gel, reduced in-gel, alkylated, destained and proteolytically cleaved with TPCK-treated trypsin (Invitrogen) following a standard protocol (Shevchenko *et al.*, 2007). The cleaved peptide masses were identified by MALDI-TOF mass spectrometry using a Bruker ultrafleXtreme MALDI-TOF/TOF spectrometer, using a Mascot database search (Perkins *et al.*, 1999) for confirmation of the protein identity.

2.3. Sequence alignment of different ASADHs

The reference amino-acid sequences for fungal, archaeal and bacterial ASADHs were obtained from the UniProt database (<http://www.uniprot.org>). ClustalW2 was used for sequence alignment, and a secondary-structure alignment of these sequences was constructed using ESPript 3.0 (Robert & Gouet, 2014).

2.4. Biochemical assay

Owing to the instability of the substrate aspartyl phosphate, the enzymatic activity assay was carried out in the non-physiological direction to measure the increase in absorbance of NADPH at 340 nm ($\epsilon_{340} = 6220 \text{ M}^{-1} \text{ cm}^{-1}$) using a Cary UV–visible spectrophotometer. The assay was performed at 298 K in a 1 ml quartz cuvette using a buffer consisting of final concentrations of 20 mM phosphate, 1.5 mM NADP, 0.3 mM aspartate semialdehyde (ASA), 120 mM *N*-cyclohexyl-2-aminoethanesulfonic acid (CHES) buffer pH 8.6 and 200 mM KCl with a final enzyme concentration of 0.05 mg ml $^{-1}$ (Gao *et al.*, 2010). The increase in absorbance of NADPH correlates with the formation of the product aspartyl phosphate obtained by phosphorylation of the substrate ASA.

2.5. Crystallization

Initial screening of crystallization conditions for the apo form of CnASADH were performed in 24-well plates using the hanging-drop vapor-diffusion method with commercial crystallization kits from Jena Bioscience as well as in-house custom screening kits. The best initial hit was found from the JB Basic 2 screen: 10% PEG 8000, 8% ethylene glycol, 100 mM HEPES pH 7.5. This initial crystal hit was extensively optimized by grid screening around this starting condition, and the best crystals were grown at 293 K by using a 2:1 ratio of protein solution (9 mg ml $^{-1}$) to well solution. Rod-shaped

Table 1

Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	APS beamline 21-ID-G
Wavelength (Å)	0.9786
Temperature (K)	100
Detector	MAR Mosaic CCD
Crystal-to-detector distance (mm)	250
Rotation range per image (°)	1
Total rotation range (°)	360
Space group	$P2_1$
a, b, c (Å)	76.59, 130.96, 91.94
α, β, γ (°)	90, 92.87, 90
Resolution range (Å)	66.05–2.60
Total No. of reflections	405413
No. of unique reflections	55730
Completeness (%)	100 (100)
R_{merge}	0.094 (0.445)
Multiplicity	7.3 (7.1)
$\langle I/\sigma(I) \rangle$	15.2 (4.1)
Overall B factor from Wilson plot (Å ²)	39.1

crystals (Supplementary Fig. S1) were obtained after 48 h using 10% PEG 8000, 6% ethylene glycol, 0.1 M HEPES pH 7.5 as the well solution. In parallel, the recombinant enzyme was co-crystallized in the presence of 4 mM NADP to obtain crystals of the enzyme–cofactor complex. The enzyme–NADP crystals were also soaked with a 10 mM solutions of several of the moderate inhibitors but, unfortunately, none of these crystals diffracted to even moderate resolution.

2.6. Data collection and processing

The optimized crystals were cryoprotected by briefly soaking them in the crystallization buffer containing 25%(v/v) ethylene glycol and were flash-cooled in liquid nitrogen. These crystals were preliminarily screened in-house with a Rigaku FR-E rotating-anode generator equipped with an R-Axis IV image-plate detector. A complete data set was subsequently collected from a single cryocooled crystal of *CnASADH* at 100 K on beamline 22-ID-G of LS-CAT at the Advanced Photon Source (Argonne National Laboratory). The diffraction images were processed and scaled using *iMosflm* (Battye *et al.*, 2011) and *AIMLESS* (Evans & Murshudov, 2013) from the *CCP4* suite (Winn *et al.*, 2011). Data-collection statistics for this data set are summarized in Table 1.

2.7. Structure solution and refinement

CnASADH shares 54% sequence identity with the ASADH from *C. albicans* (*CaASADH*; PDB entry 3hsk; Arachea *et al.*, 2010), with a *BLAST* score of 395. The atomic coordinates of *CaASADH* were used as the search model for molecular replacement using *Phaser* (McCoy *et al.*, 2007) to determine the initial phases. Deleting chain *B* from the coordinates of this search model gave a solution with a translation-function *Z*-score (TFZ) of 16.8. This resulting model was then subjected to 20 cycles of rigid-body refinement using the medium-resolution data to 3.5 Å with *REFMAC5* (Murshudov *et al.*, 2011), followed by restrained-body refinement at the highest resolution (2.6 Å). This refinement gave initial R_{work} and R_{free} values of 0.31 and 0.37, respectively. The density was

further modified with *Parrot* and the structure was built using *Buccaneer* (Cowtan, 2006), an automated protein model-building program from the *CCP4* suite. The program automatically located most of the protein residues, and the model that was generated showed significant improvement in the R values ($R_{\text{work}} = 0.26$ and $R_{\text{free}} = 0.32$). The structure was further built by a combination of iterative rounds of restrained refinement after manual building in *Coot* (Emsley & Cowtan, 2004). The stereochemical quality of the final model was checked by *PROCHECK*. The final structure was visualized and analyzed using *PyMOL* (DeLano, 2002) and surface calculations were conducted using *PDBEPIA* (Krissinel &

Figure 1

(a) Overall structure of aspartate semialdehyde dehydrogenase from *C. neoformans*. Each monomeric unit is shown in dark and light colors, with the dimers colored blue and green and the N- and C-terminus of each monomer labeled. (b) Superposition of a dimer of *CnASADH* (blue) with the *CaASADH* (green) structure. For *CnASADH* the coenzyme-binding domain is shown in light blue and the catalytic domain is shown in dark blue. The active site is labeled with a red star.

Table 2
Structure refinement.

Values in parentheses are for the outer shell.

Resolution range (Å)	66.05–2.60 (2.67–2.60)
Completeness (%)	100.0 (100.0)
No. of reflections, working set	52836 (3875)
No. of reflections, test set	2849 (234)
Final R_{cryst}	0.181 (0.248)
Final R_{free}	0.220 (0.337)
No. of non-H atoms	
Protein	10796
Ligand	4
Solvent	619
Total	11419
R.m.s. deviations	
Bonds (Å)	0.012
Angles (°)	1.513
Average B factors (Å ²)	
Overall	45.0
Protein	45.3
Ligand	40.7
Waters	49.0
Ramachandran plot	
Most favored region (%)	97.0
Allowed region (%)	2.8
Outlier region (%)	0.2

Henrick, 2007). Atomic coordinates and structure factors have been deposited in the Protein Data Bank with accession code 5cef after revalidation with *PDB_REDO* (Joosten *et al.*, 2009). The final statistics of the structure refinement are summarized in Table 2.

3. Results and discussion

3.1. Overall structure of *C. neoformans* ASADH

Full-length *Cn*ASADH contains 365 amino acids and the enzyme crystallized in the monoclinic space group $P2_1$ with four molecules in the asymmetric unit, giving a Matthews coefficient of $2.96 \text{ Å}^3 \text{ Da}^{-1}$ and a solvent content of $\sim 58\%$.

The final electron-density map was of good quality and allowed near-complete building of the structure for residues 4–189 and 193–365 of each monomer, with each of the monomers lacking interpretable electron density for the first three amino acids and for residues 190–192. The four monomers in the asymmetric unit represent the functional dimer of dimers for this enzyme (Fig. 1*a*). These dimers are nearly identical, with r.m.s.d.s of 0.19 Å between chains *A* and *B* and 0.11 Å between chains *C* and *D*. This oligomeric assembly is further supported by native PAGE and gel-filtration chromatography (data not shown), which suggest a molecular weight of about 160 kDa, which is consistent with this assembly. The structure of *Cn*ASADH belongs to the Rossmann-fold superfamily of pyridine-linked dehydrogenases and shares the same overall monomeric structural features as the other ASADHs for which structures have been determined. Each monomer is composed of an N-terminal coenzyme-binding domain assembled from residues 1–154 and 346–361, and a C-terminal catalytic and dimerization domain consisting of residues 155–345 (Fig. 2). The NADP-binding domain consists of mixed β -strands flanked on both sides with α -helices, with seven β -strands, five α -helices and two 3_{10} -helices, while the C-terminal domain consists of six β -strands, six α -helices and a single 3_{10} -helix that represent the catalytic site and the dimerization interface. This fungal ortholog most closely resembles the *Ca*ASADH structure (Arachea *et al.*, 2010), with an overall r.m.s.d. of 0.85 Å for the backbone atoms and only some minor structural differences (Fig. 1*b*).

3.2. Secondary-structure arrangement and comparison

The structure of the coenzyme-binding domain in *Cn*ASADH starts with a β -strand followed by an α -helix connected to two short β -strands (β_1 – α_1 – β_2), representing the β – α – β motif of the Rossmann fold present in the typical ASADH family. This motif is then linked to a 45-residue

Figure 2

Topology diagram showing the secondary-structural arrangement of a *Cn*ASADH monomer with the β -sheets inside light blue shaded boxes and the α -helices outside. Each β -sheet and α -helix is numbered in increasing order starting from the N-terminus.

Figure 3

Sequence alignment of representative ASADH enzymes from fungal (*Cryptococcus neoformans*), yeast (*Candida albicans*), archaeal (*Methanocaldococcus jannaschii*), Gram-positive (*Streptococcus pneumoniae*) and Gram-negative (*Vibrio cholerae*) bacterial species. Fully conserved residues are shown with a red background, residues with conservative mutations are shown in red letters and active-site residues are marked with orange (Cys156) and blue (His255) stars. The helical subdomain region found in the bacterial ASADHs is shown inside a blue rectangle.

connecting region composed of two short helices, $\alpha 2$ and $\alpha 3$, one short β -strand, β_3 , and a 3_{10} helix, η_1 , located primarily on the surface and also involved in cofactor binding (Fig. 2). There is a second β - α - β motif after this surface loop (β_4 - α_4 - β_5) that is highly conserved among all ASADHs. *CnASADH* also contains a third β - α - β motif (β_6 - η_2 - α_5 - β_7), where an extra 3_{10} -helix, η_2 , is present between β_6 and α_5 , and this motif is less conserved in the different ASADHs. This β_7 sheet is then connected to a conserved helix α_6 and is directed towards the first β -strand (β_8) of the dimerization domain. The helical subdomain found in bacterial ASADHs (helix–turn–helix) is absent in *CnASADH* (Fig. 3) and this region is replaced by two residues without significant electron density, suggesting the presence of a somewhat flexible region. This flexible region is followed by two helices, $\alpha 7$ and $\alpha 8$, that are similar to the *CaASADH* structure. The helix $\alpha 8$ is connected to an 18-residue surface linker which is structured into two short β -strands in *CaASADH*, and the remaining structural arrangement is similar to *CaASADH*, with minor differences in the length of the helix and the β -strands.

3.3. Comparison to related ASADH structures

There is only a single published structure of a fungal ASADH with a bound cofactor (Arachea *et al.*, 2010) and, despite significant efforts to grow NADP-bound crystals or crystals soaked with several inhibitors of *CnASADH*, none of these crystals diffracted to better than 5 Å resolution. The loop that was not modeled in this structure of *CnASADH* because of missing density was also absent in the previously determined *CaASADH* structure, suggesting that this is an intrinsically flexible and disordered region in the absence of bound substrate. The $\alpha 5$ helix is four residues shorter than the comparable helix in *CaASADH*, but is similar in length to those in the other bacterial and archaeal orthologs. The active-site cysteine nucleophile in *CnASADH* is found to be present in a single conformation, positioned at the end of a conserved helix ($\alpha 5$), similar to the orientation observed in the bacterial ASADHs. In contrast, in the other fungal ASADH from *CaASADH* this nucleophile is seen to be more flexible and is present in multiple conformations in this structure.

While the critical catalytic residues in the active site are highly conserved among the ASADHs, the catalytic properties of *CnASADH* shows significant differences between fungal and some bacterial ASADHs. The k_{cat} of *CnASADH* is $10.2 \pm 0.6 \text{ s}^{-1}$ and the K_{m} for the native ASA substrate is $0.34 \pm 0.05 \text{ mM}$, leading to a catalytic efficiency ($k_{\text{cat}}/K_{\text{m}}$) of $3 \times 10^4 \text{ s}^{-1} \text{ M}^{-1}$. In contrast, the k_{cat} for *CaASADH* is only $0.12 \pm 0.01 \text{ s}^{-1}$ (Arachea *et al.*, 2010). Thus, *CnASADH* has an 85-fold higher k_{cat} , a value that is comparable to some of the Gram-positive bacterial orthologs (Faehnle *et al.*, 2006). This higher catalytic activity of *CnASADH* suggests that there may be some particularly favorable interactions in the active site when the substrate is bound and may reflect a specific growth requirement for this particular fungal pathogen.

3.4. Positioning of the conserved active-site residues

As stated above, the identities of the critical catalytic functional groups are highly conserved among the different members of the ASADH family. Superposition of the residues around the *CnASADH* active site with yeast *CaASADH* shows that these conserved residues are each located in the same position (Fig. 4). The only significant active-site differences in the apoenzyme are the orientation of the cysteine nucleophile in a single conformation in *CnASADH* and that Ser341 of *CaASADH* is replaced by an isoleucine in this new structure. The distance between the S atom of the Cys156 nucleophile and the N $^{\epsilon}$ atom of His255 is 5.1 Å in this *CnASADH* structure (Fig. 4), whereas these same atoms are reported to be about 2 Å closer in *CaASADH* (Arachea *et al.*, 2010). This increased distance in the *C. neoformans* structure suggests the need for a conformational change in the active site in response to substrate binding that would move these active-site residues into an optimal position to catalyze this reaction. While Ser341 has been replaced by isoleucine in this fungal enzyme form, this residue is involved in making the same hydrogen-bonding interaction between its backbone amide group and the backbone carbonyl O atom of His255 as was observed in *CaASADH*.

3.5. Binding of an exogenous ligand

The apo form of *CnASADH* was crystallized, and the structure was subsequently determined in the absence of any added ligands. However, after final refinement of the structure the difference Fourier map at the 3σ level showed additional unassigned electron density near the active-site base His255, but only in chain A. A search among the possible solution components identified ethylene glycol (EG) as the most likely bound ligand, and this density nicely accommodates this molecule. The EG could have come from the crystallization precipitant (6% EG) or from the cryoprotectant solution

Figure 4

Superposition of the active-site residues involved in substrate binding and catalysis between *CnASADH* (cyan) and *CaASADH* (green). Unlike in the *C. albicans* structure, a single conformation was observed for the active-site cysteine of *CnASADH*, with a greater distance between the S atom of the cysteine nucleophile and the N $^{\epsilon}$ atom of the histidine acid/base catalyst.

(25% EG) during flash-cooling for data collection. This ligand is anchored at the active site through hydrogen-bonding interactions between its O atom and the carbonyl O atom of His255 and also with the S atom of Cys338 (Supplementary Fig. S2). In addition, there are several hydrophobic interactions with other residues in the active-site cavity. The *B* factors for EG are double those for the interacting enzyme side chains, and this is not unexpected for the fortuitous binding of a non-optimized ligand. However, the nature of the ligand interactions within the active site can serve as an additional guide for the design of inhibitors of fungal ASADHs. In fact, a structurally related molecule, 3-bromopyruvate, was found to be a selective and moderate inhibitor of only a fungal ASADH, with a K_i value of 0.34 mM for *Ca*ASADH and no inhibition observed for bacterial ASADHs (Gao *et al.*, 2010).

3.6. Tetramer interface and oligomeric arrangement

*Cn*ASADH has been shown to exist as a tetramer in solution, similar to that observed for the fungal ASADH from *T. rubrum* (Q. Jin & S. Cui, personal communication). The four molecules in the asymmetric unit share an extensive contact interface that buries a total of 9940 Å² of surface area, utilizing 36 residues within each dimer (*AB* and *CD*) that can provide up to 14 hydrogen bonds and three salt bridges between chains *A* and *B* and a similar set of contacts between chains *C* and *D*. There are also cross-interactions between chains *A* and *C*, with eight residues making four hydrogen bonds, whereas chains *B* and *D* interact in a different fashion, with nine residues from chain *B* interacting with 11 residues from chain *D*, producing one salt bridge and four hydrogen-bonding interactions. The three pairs of salt bridges, Glu211–Arg325, Glu211–Arg327 and Glu215–Lys335, that are formed between chains *A* and *B* and chains *C* and *D*, as well as an additional salt bridge between Lys55 (chain *B*) and Asp198 (chain *D*), are not present in the *Ca*ASADH structure. This suggests that the additional salt-bridge interactions from different chains assist in the formation and stabilization of the dimers of dimers found in the *Cn*ASADH structure.

4. Conclusions

*Cn*ASADH is a potential antifungal drug target. The structure of *Cn*ASADH in complex with ethylene glycol has been determined to 2.6 Å resolution. *Cn*ASADH possess a similar overall structure to the most closely related yeast ASADH (*Ca*ASADH). There is a difference in biological assembly between *Cn*ASADH (a dimer of dimer) and *Ca*ASADH (a

dimer), and this information will be helpful in designing a drug targeting the interface of the dimers.

Acknowledgements

The authors wish to thank the staff members at beamline 21-ID-G of LS-CAT at the Advanced Photon Source for their assistance with data collection and Dr Ronan Keegan (STFC Rutherford Appleton Laboratory) and the 2015 CCP4 Summer School for guidance in obtaining the optimized structural solution. This work was supported by a grant from the National Institutes of Health (AI077720).

References

- Arachea, B. T., Liu, X., Pavlovsky, A. G. & Viola, R. E. (2010). *Acta Cryst.* **D66**, 205–212.
- Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* **D67**, 271–281.
- Black, S. & Wright, N. G. (1955). *J. Biol. Chem.* **213**, 39–50.
- Blanco, J., Moore, R., Kabaleeswaran, V. & Viola, R. (2003). *Protein Sci.* **12**, 27–33.
- Cowtan, K. (2006). *Acta Cryst.* **D62**, 1002–1011.
- DeLano, W. L. (2002). *PyMOL*. <http://www.pymol.org>.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* **D69**, 1204–1214.
- Faehnle, C. R., Le Coq, J., Liu, X. & Viola, R. E. (2006). *J. Biol. Chem.* **281**, 31031–31040.
- Gao, G., Liu, X., Pavlovsky, A. & Viola, R. E. (2010). *J. Biomol. Screen.* **15**, 1042–1050.
- Hadfield, A. T., Shammass, C., Kryger, G., Ringe, D., Petsko, G. A., Ouyang, J. & Viola, R. E. (2001). *Biochemistry*, **40**, 14475–14483.
- Joosten, R. P. *et al.* (2009). *J. Appl. Cryst.* **42**, 376–384.
- Kent, L. B. & Juneann, W. M. (1998). *Emerg. Infect. Dis.* **4**, 71–83.
- Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- 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.
- Pavlovsky, A. G., Liu, X., Faehnle, C. R., Potente, N. & Viola, R. E. (2012). *Chem. Biol. Drug Des.* **79**, 128–136.
- Pavlovsky, A. G., Thangavelu, B., Bhansali, P. & Viola, R. E. (2014). *Acta Cryst.* **D70**, 3244–3252.
- Perkins, D. N., Pappin, D. J. C., Creasy, D. M. & Cottrell, J. S. (1999). *Electrophoresis*, **20**, 3551–3567.
- Robert, X. & Gouet, P. (2014). *Nucleic Acids Res.* **42**, W320–W324.
- Rodrigues, M. L., Alviano, C. S. & Travassos, L. R. (1999). *Microbes Infect.* **1**, 293–301.
- Shevchenko, A., Tomas, H., Havliš, J., Olsen, J. V. & Mann, M. (2007). *Nature Protoc.* **1**, 2856–2860.
- Viola, R. E., Faehnle, C. R., Blanco, J., Moore, R. A., Liu, X., Arachea, B. T. & Pavlovsky, A. G. (2011). *J. Amino Acids*, **2011**, 1–11.
- Viola, R. E., Liu, X., Ohren, J. F. & Faehnle, C. R. (2008). *Acta Cryst.* **D64**, 321–330.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.