University of Kent

Kent Academic Repository

Nichols, C.E., Ren, J., Leslie, Kris D., Dhaliwal, B., Lockyer, M., Charles, I., Hawkins, A.R. and Stammers, D.K. (2004) *Comparison of Ligand-induced Conformational Changes and Domain Closure Mechanisms, Between Prokaryotic and Eukaryotic Dehydroquinate Synthases.* Journal of Molecular Biology, 343 (3). pp. 533-546. ISSN 0022-2836.

Downloaded from

https://kar.kent.ac.uk/115036/ The University of Kent's Academic Repository KAR

The version of record is available from

https://doi.org/10.1016/j.jmb.2004.08.039

This document version

Publisher pdf

DOI for this version

Licence for this version

UNSPECIFIED

Additional information

Versions of research works

Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in ***Title of Journal*** , Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries

If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies).

doi:10.1016/j.jmb.2004.08.039 J. Mol. Biol. (2004) 343, 533–546

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Comparison of Ligand-induced Conformational Changes and Domain Closure Mechanisms, Between Prokaryotic and Eukaryotic Dehydroquinate Synthases

C. E. Nichols[1], J. Ren[1], K. Leslie[1], B. Dhaliwal[1], M. Lockyer[2], I. Charles[2] A. R. Hawkins[3] and D. K. Stammers[1]*

[1]*Division of Structural Biology The Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK*

[2]*Arrow Therapeutics Ltd Britannia House, Trinity Street Borough, London SE1 1DA UK*

[3]*School of Cell and Molecular Biology, Medical School Catherine Cookson Building Framlington Place, University of Newcastle-upon-Tyne Newcastle NE2 4HH, UK*

**Corresponding author*

Dehydroquinate synthase (DHQS) is a potential target for the development of novel broad-spectrum antimicrobial drugs, active against both prokaryotes and lower eukaryotes. Structures have been reported for *Aspergillus nidulans* DHQS (*An*DHQS) in complexes with a range of ligands. Analysis of these *An*DHQS structures showed that a large-scale domain movement occurs during the normal catalytic cycle, with a complex series of structural elements propagating substrate binding-induced conformational changes away from the active site to distal locations. Compared to corresponding fungal enzymes, DHQS from bacterial species are both mono-functional and significantly smaller. We have therefore determined the structure of *Staphylococcus aureus* DHQS (*Sa*DHQS) in five liganded states, allowing comparison of ligand-induced conformational changes and mechanisms of domain closure between fungal and bacterial enzymes. This comparative analysis shows that substrate binding initiates a large-scale domain closure in both species' DHQS and that the active site stereochemistry, of the catalytically competent closed-form enzyme thus produced, is also highly conserved. However, comparison of *An*DHQS and *Sa*DHQS open-form structures, and analysis of the putative dynamic processes by which the transition to the closed-form states are made, shows a far lower degree of similarity, indicating a significant structural divergence. As a result, both the nature of the propagation of conformational change and the mechanical systems involved in this propagation are quite different between the DHQSs from the two species.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: DHQS; crystal structure; domain closure; conformational changes; X-ray crystallography

Introduction

The shikimate pathway of bacteria and microbial eukaryotes[1] is absent in mammals. Pathogenic bacteria mutant in the pathway are attenuated for virulence.[2] Enzymes of the shikimate pathway are thus attractive targets for the development of antimicrobial drugs. One such potential target is the enzyme dehydroquinate synthase (DHQS), which catalyses the conversion of 3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP) to dehydroquinate (DHQ).

DHQS from both eukaryotic and prokaryotic species has been thoroughly characterised biochemically and there are significant structural differences.[3–6] Steps 2–6 of the shikimate pathway in fungi, such as *Aspergillus nidulans*, are carried out by the pentafunctional AROM protein (with DHQS forming the N-terminal domain), whilst in bacterial species the corresponding enzymes, including DHQS, are encoded on separate genes. Bacterial

Abbreviations used: CBP, carbaphosphonate; DHQ, dehydroquinate; DHQS, dehydroquinate synthase; *Sa*DHQS, *Staphlococcus aureus* DHQS; *An*DHQS, *Aspergillus nidulans* DHQS; SE(s), structural element(s); SSE(s), secondary structure element(s) e.g. β-strands or α-helices; Z1–Z7, zones of difference 1–7; PE1–PE3, proximal element 1–3; DE1–DE6, distal element 1–6; AE0–AE6, additional element 0–6; TA(s), torsion angle(s); TAD(s), torsion angle difference(s); TADP(s), torsion angle difference peak(s).

E-mail address of the corresponding author: daves@strubi.ox.ac.uk

0022-2836/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.

DHQSs are also significantly smaller than their eukaryotic orthologues.

More recently, the solution of the X-ray crystal structure of *A. nidulans* DHQS (*An*DHQS),[7] as a ternary complex with cofactor NAD and substrate analogue carbaphosphonate (CBP),[6] has led to the proposal of a catalytic mechanism. An understanding of domain closure in *An*DHQS was achieved by determining the structure of the enzyme in eight complexes with combinations of ligands including: NAD; the cofactor fragment ADP and the substrate analogue CBP.[8,9] A further preliminary report of an open form binary complex of *An*DHQS with NAD has also been published,[10] as well as a high-resolution structure at 1.7 Å.[11] Our analysis of 21 *An*DHQS structures revealed that substrate binding induces a large-scale domain movement, with the N and C-terminal domains showing a relative rotation of 11–13° about a two-point hinge, approximately parallel with the bound co-factor NAD.[9] The domain movement closes the active site cleft, excludes bulk solvent, and creates the necessary stereochemical environment for catalysis, which precludes the enantiomer formation seen with *in vitro* experiments.[12]

The mechanical basis of the *An*DHQS domain rearrangement was found to involve the complex interplay of an array of structural elements (SEs) comprising: (1) a "rigid-core" showing little structural change during domain closure; (2) two hinge points (HP1–HP2), forming a two-point single axis hinge about which the two domains undergo their relative rotation; (3) a dimer interface region (DI); (4) eight elements showing ligand-state-dependent conformation, subdivided into three proximal elements (PE1–PE3) containing residues contacting the substrate in the closed form and five distal elements (DE1–DE5) whose conformation showed variations dependent on the liganded state of the enzyme but which make no direct contact with the substrate; and (5) six additional surface elements (AE1–AE6) showing variable conformation independent of state, but which probably have flexibility to allow the accommodation of stresses generated by the relative movement of other structural elements (SEs).

Substrate binding induces conformational changes directly in the PEs, which then propagate through the protein structure *via* connecting elements, affecting the conformation of DEs, and collectively inducing an ordered sequence of rearrangements leading to domain closure. The results also indicated that the closed complex can be seen as a system under tension, held closed only by the strength of the bonds made to the bound substrate. Thus, at the completion of the reaction cycle, when many of these linkages are cleaved, the restoring forces exerted by the elements under strain become greater than the forces holding the domains closed. The opening and closing of the active site is thus driven in both directions, rather than being passive events.[9] The domain closure mechanism in *An*DHQS is thus a complex system and the necessity of such a large-scale movement for the formation of the active site raises the intriguing possibility that it might be feasible to develop inhibitors that function by modulating *An*DHQS's domain movement, rather than by competing directly with the substrate.

The aim of the current work is to study DHQS from a second species, widely separated from *A. nidulans* in evolutionary terms, in order to assess the degree of structural conservation of the active site and the mechanical systems involved in effecting the relative movement of domains. We therefore selected DHQS from a prokaryotic species, *Staphylococcus aureus* (*Sa*DHQS), as the source for this comparison because: (1) in prokaryotes, such as *S. aureus*, steps 2–6 of the shikimate pathway are catalysed by separate stand-alone enzymes, whilst the eukaryotic *An*DHQS previously investigated is derived from one domain of the penta-domain AROM complex; (2) the *Sa*DHQS enzyme is significantly smaller than the *An*DHQS enzyme, with a chain length of 354 residues rather than 393; and (3) the overall amino acid identity of the best sequence match is 28% (rising to 39% identity between a core 228 residue fragment with the N and C termini omitted; Figure 1). The sequence identity is thus above the threshold correlation for prediction of a similar folding pattern, but still low enough to indicate significant evolutionary divergence between the two species. Thus any features conserved between them are also likely to be common to most other species. We report the crystal structure of five *Sa*DHQS complexes and a detailed comparison of *An*DHQS and *Sa*DHQS domain closure mechanisms.

Results and Discussion

Structure determination and preliminary analysis of *Sa*DHQS–ligand complexes

*Sa*DHQS was crystallised in an unliganded form as well as in complexes with NAD, carbaphosphonate (CBP) and NAD/CBP. An additional complex involved soaking CBP into crystals of the binary NAD complex (NAD/CBPsk). In contrast to *An*DHQS, where ten crystal forms were identified,[9,13] only two closely related crystal forms were observed for *Sa*DHQS. To improve diffraction quality, some crystals were subjected to partial dehydration by addition of increasing concentrations of DMSO and ethylene glycol (complexes and datasets derived from these dehydration protocols are denoted here by subscript d, see also Experimental Procedures). Data from an untreated NAD complex crystal were partially refined (R_{free} ≈35%) and the structure indicated that conformational changes compared to the NAD binary complex as a result of the application of the dehydration process appeared relatively minor.

*Sa*DHQS form A dataset I was solved by

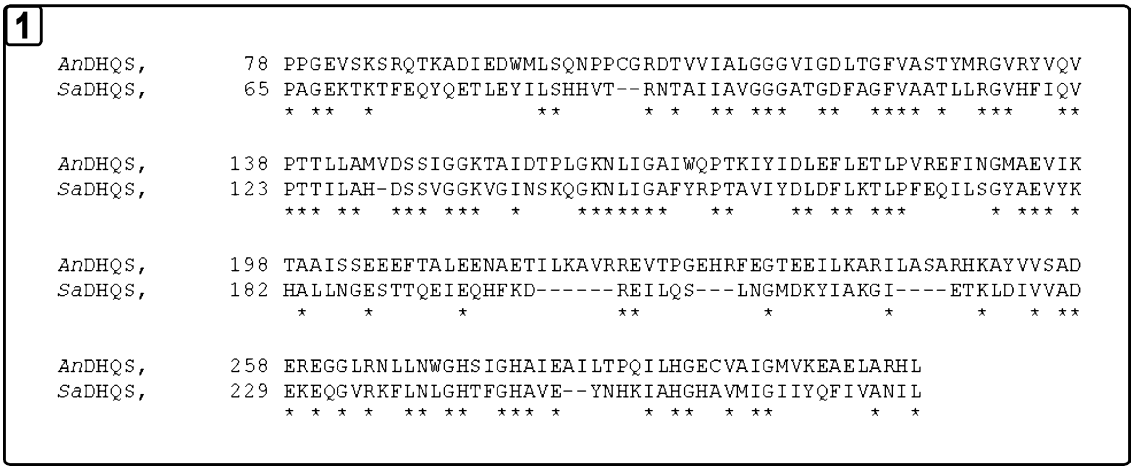


Figure 1. Sequence alignment of the core 228 residue fragment of *Sa*DHQS and *An*DHQS.

molecular replacement using *An*DHQS PDB model 1NR5 and refined to R_{work} of 0.197 (R_{free} = 0.269). The *Sa*DHQS model was used to solve form B *Sa*DHQS complexes by low-resolution, rigid-body refinement, indicating forms A and B share essentially the same packing. Conformational changes in surface features explain the unit cell *c*-axis decrease induced by the partial dehydration process. Final *R*-factors for the five structures were all less than 22% with the retention of reasonable stereochemistry (Table 1). Representative electron density is shown in Figure 2a. Figure 2c and d show STRIDE-delineated, cartoon-format representations of the open form $NAD_{(d)}$ and closed form NAD/CBP structures, respectively.

Analysis of overlays of *Sa*DHQS and *An*DHQS shows the general structural organisation to be similar, with the boundary between the N and C-terminal domains forming a cleft in the absence of CBP (Figure 2c). When CBP binds, the cleft closes, excluding bulk solvent and generating the catalytically competent active site (Figure 2d). The

Table 1. Data collection and refinement statistics for *Sa*DHQS complexes

Crystal	I	II	III	IV	V	VI
Ligands	NAD^+, CBP	NAD^+	CBP	NAD^+, CBP	NAD^+, CBP soak[a]	NAD^+
Partial dehydration	No	Yes	Yes	Yes	Yes	No
No. crystals	1	4	1	2	2	1
Space group	$P4_32_12$	$P4_3$	$P4_3$	$P4_3$	$P4_3$	$P4_3$
Unit cell parameters						
a, *b* (Å)	55.26	54.90	53.29	54.74	55.10	55.62
c (Å)	232.87	229.11	240.07	231.16	232.32	235.22
Crystal form	*A*	*B*	*B*	*B*	*B*	*B*
Resolution range (Å)	30.00–2.45 (2.54–2.45)	30.00–2.20 (2.24–2.20)	30.00–2.30 (2.34–2.30)	30.00–2.35 (2.40–2.35)	30.00–2.80 (2.90–2.80)	30.00–3.00 (3.11–3.00)
Redundancy	6.7 (4.6)	12.5 (11.2)	3.1 (2.7)	3.5 (3.5)	4.8 (3.8)	4.3 (3.8)
Completeness (%)	99.2 (98.8)	99.5 (100.0)	97.7 (96.8)	97.6 (97.8)	97.8 (94.2)	98.8 (88.7)
R_{merge}[b]	0.120 (0.358)	0.100 (0.559)	0.044 (0.247)	0.051 (0.235)	0.080 (0.481)	0.086 (0.732)
$I/\sigma I$	18.89 (2.69)	26.22 (2.38)	23.46 (3.69)	20.43 (2.88)	12.85 (1.57)	16.2 (1.22)
Asymmetric unit (subunits)	1	2	2	2	2	2
R_{work}[c,d] (%)	19.7	20.9	21.0	21.7	19.7	35.5[c]
R_{free}[c,d] (%)	26.9	28.1	31.3	31.0	26.4	–
Mean *B*-factors (Å^2)						
All atoms	48.4	51.8	61.5	58.2	62.9	–
Protein (main-chain)	47.0	49.5	60.3	57.1	61.0	–
Protein (side-chain)	50.1	54.6	63.7	60.4	65.6	–
Water	47.4	48.1	52.1	49.9	47.6	–
Ligands	42.4	42.8	44.5	50.5	67.1	–
r.m.s.d. values						
Bond lengths (Å)	0.006	0.008	0.008	0.006	0.006	–
Bond angles (deg.)	1.37	1.37	1.28	1.38	1.33	–

Figures in parentheses are for outer-shell data.
[a] 4 mM soak, 24 hours.
[b] $R_{merge} = \Sigma|I_{obs} - \langle I\rangle|/\Sigma\langle I\rangle$.
[c] All atoms, rigid-body refinement only.
[d] $R = \Sigma_{hkl}|F_o(hkl) - F_c(hkl)|/\Sigma_{hkl}F_o(hkl)$.

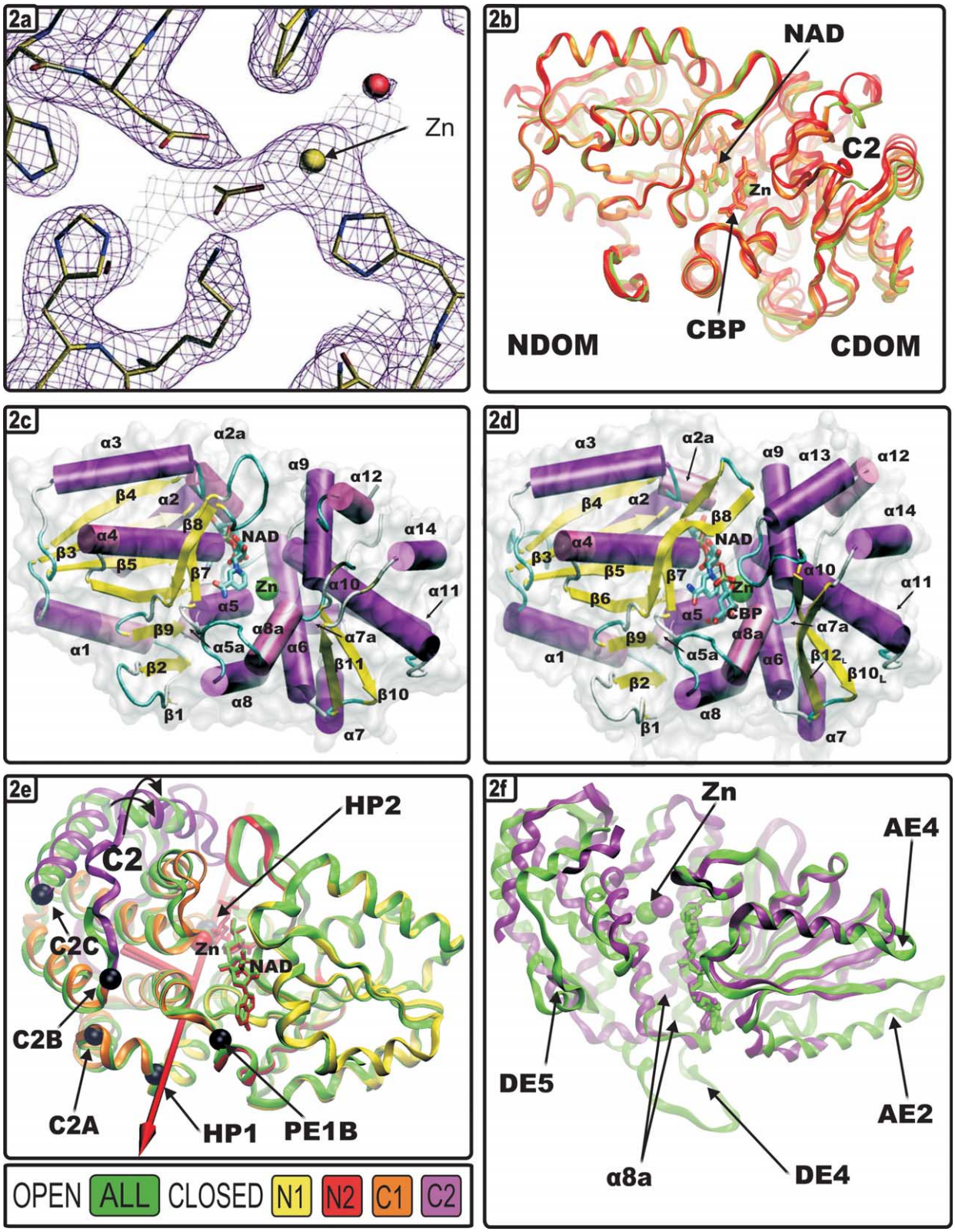


Figure 2. a, Representative $2F_o - F_c$ electron density of the active site of the *Sa*DHQS $NAD_{(d)}$ structure; cf. Nichols *et al*.[9] b, Ribbon-format overlay of *Sa*DHQS structures: $NAD_{(d)}$ (green), $NAD/CBP_{(d)}$ (orange) and NAD/CBP (red), illustrating the partial relaxation of the C2 sub-domain induced by the crystal dehydration process. c and d, Combination diagram of transparent SURF-generated molecular surface (1.5 Å probe) and cartoon-format secondary structure definitions for *Sa*DHQS open form (c) and closed form (d) structures. e, Ribbon-format overlay of *Sa*DHQS open and closed forms, illustrating HINGEFIND results and *Sa*DHQS domain structure. f, Ribbon-format overlay of *An*DHQS NAD (green) and *Sa*DHQS $NAD_{(d)}$ (purple) structures, illustrating reorientation of helix α8a and truncation of loops AE2, AE4 and DE4–DE5 in *Sa*DHQS.

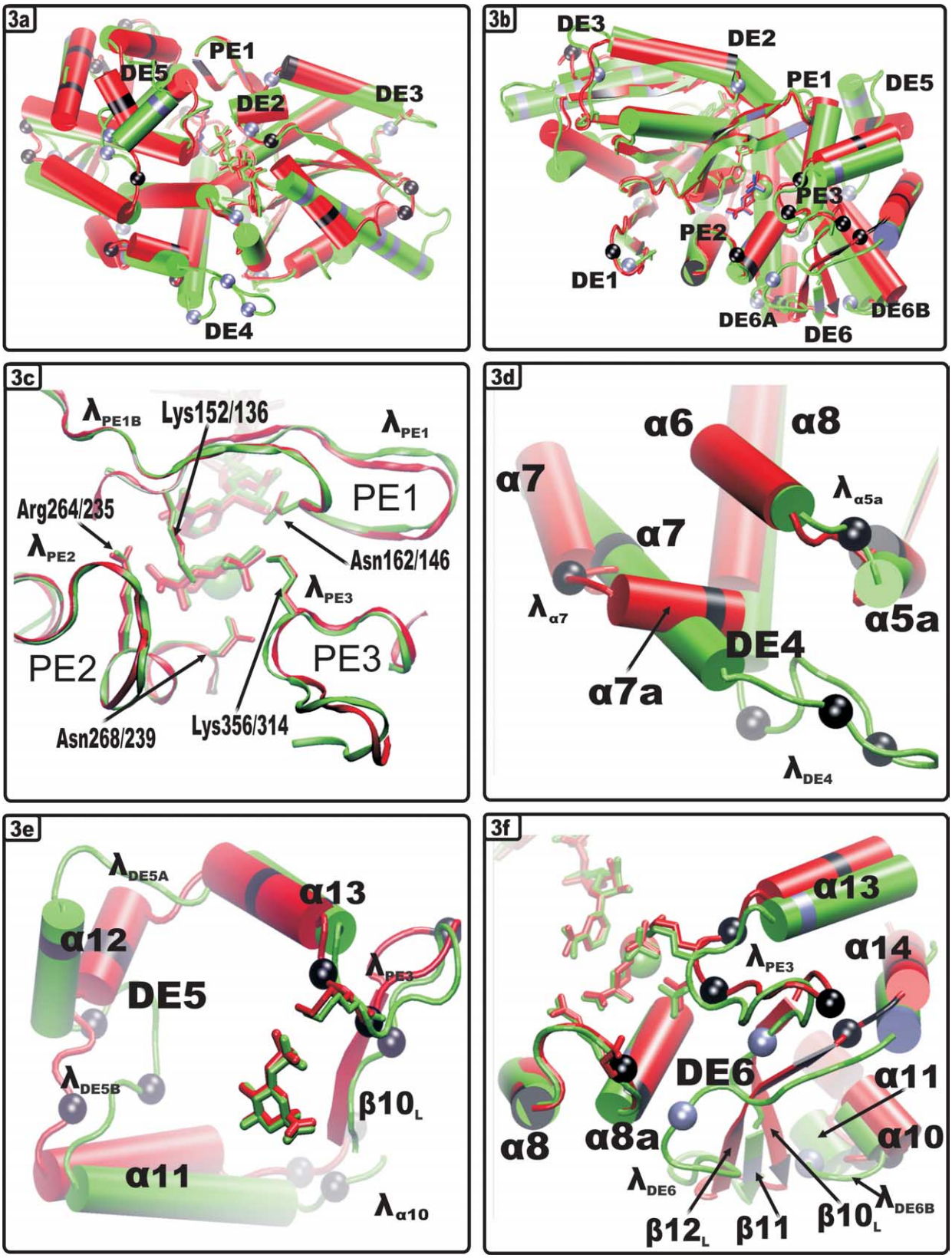


Figure 3. a and b, Distal (a) and proximal (b) views of cartoon-format overlays of *An*DHQS NAD/CBP (green) and *Sa*DHQS NAD/CBP (red) structures, with TADP marked, respectively, as black and violet spheres or bands. c, Ribbon-format overlay of *Sa*DHQS (red) and *An*DHQS (green) active site regions illustrating conservation of active site and proximal elements. d–f, Cartoon-format overlay of *Sa*DHQS (red) and *An*DHQS (green) structures illustrating: d, splitting of helix α7 in *Sa*DHQS; e, truncation of DE5 and re-orientation of helix α12 in *Sa*DHQS; and f, variation in DE6 secondary structure between *Sa*DHQS and *An*DHQS.

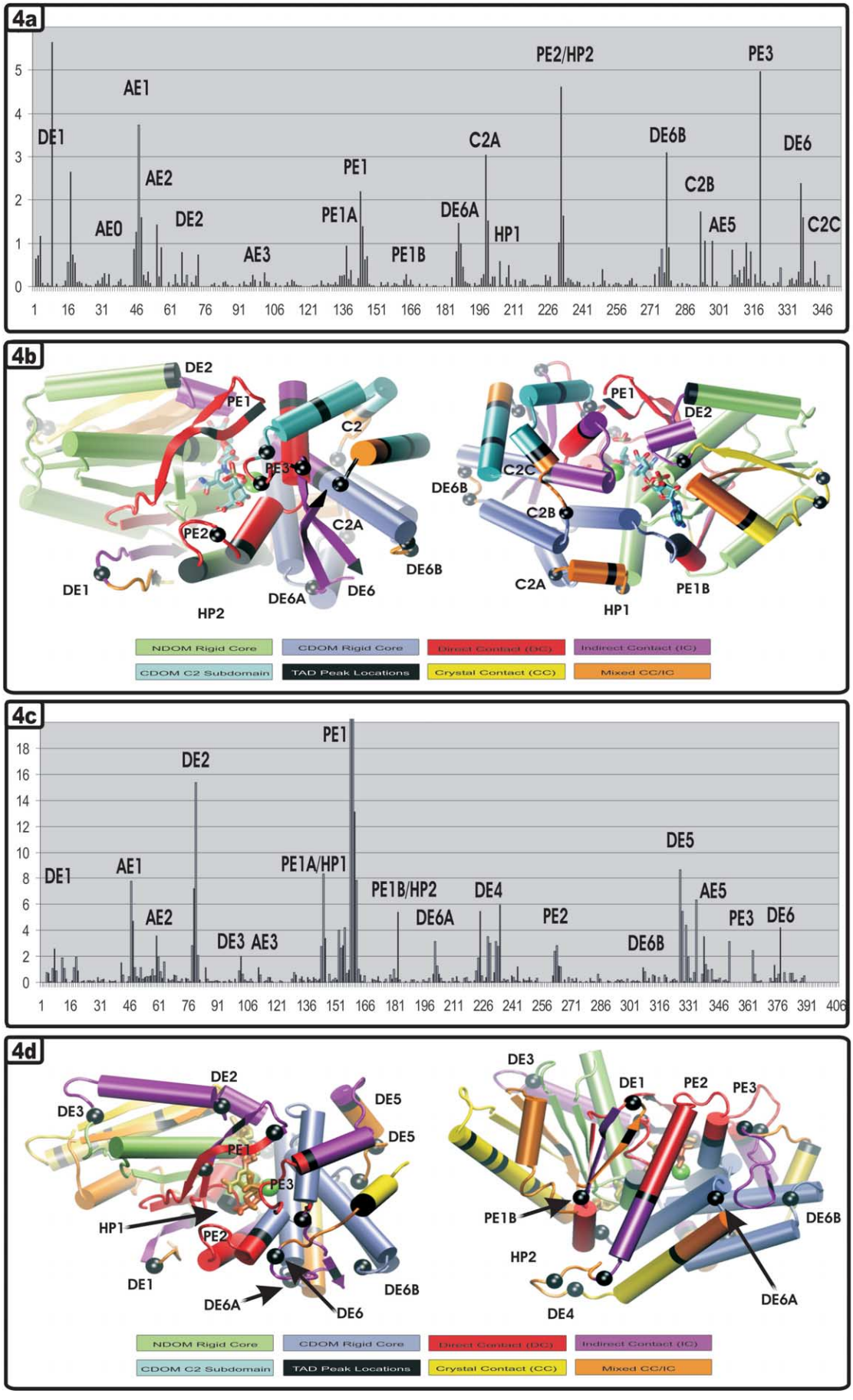


Figure 4 (*legend opposite*)

underlying process of rearrangement is also similar with substrate binding causing induced-fit conformational changes in proximal elements, which propagate through the structure and cause further changes in distal elements. The two sets of changes then act concertedly to initiate an approximately rigid-body counter-rotation of the two main parent domains about an axis parallel with the NAD-binding site, thereby yielding the closed conformation.

*An*DHQS structures partition to two main states: "open form" (apo, NAD) and the CBP-dependent "closed form" (CBP, NAD/CBP).[9] With *Sa*DHQS we observe three states, an open form (NAD, $NAD_{(d)}$), a closed form (NAD/CBP) and an "inter-intermediary form" (Figure 2b). In the latter case the PE3 Lys314 contact detaches from CBP and the PE3/C2 and PE1 elements relax back away from the active site to a partially open conformation ($CBP_{(d)}$, $NAD/CBP_{(d)}$ and $NAD/CBPsk_{(d)}$). This may be as a result of the CBP-bound closed form being in a strained conformation and the partial dehydration releasing the weakest link (Lys314). Outward movement of the equivalent *An*DHQS residue (Lys356) was hypothesised to give the final stage of the domain closure. Inward movement of Lys356 was suggested as the first part of domain opening, since cleavage of the substrate phosphate releases it once more. The $NAD/CBP_{(d)}$ complex can be considered as the final stage intermediate of the complete domain closure sequence.

In spite of the use of the partial dehydration process there were some regions where certain *Sa*DHQS side-chains were not clearly defined in electron density maps, indicating some inherent flexibility. As a result, the mean torsional angle difference (TAD) backgrounds from comparison of individual pairs of ligand sets are rather high at 5–11°. The noise level can, however, be greatly reduced by generating TAD sets for every available pair of *Sa*DHQS structures, including between chains of the same crystal, and cross-averaging them to create a "composite flexion plot". This process then yields data with a substantially reduced TAD background (2–3°), and clear peaks related to known structural features from comparison with *An*DHQS (Figure 4a–d), (see also Experimental Procedures, section Model analysis). It was therefore possible to analyse conformational changes in *Sa*DHQS on ligand binding with some degree of confidence.

Comparative analysis of *An*DHQS and *Sa*DHQS structural data

The *Sa*DHQS protein chain is significantly shorter than that of *An*DHQS with 354, rather than 393, residues. The two species' primary sequences also show relatively low homology, and an initial BLAST alignment identifies only a 228 residue overlap region with 39% sequence identity (Figure 1). Comparing overlays of the two species' open and closed form complexes clearly reveals that the closed forms overlap well (Figure 3a and b), CA atom r.m.s.d. 1.1 Å for 292 residue core overlap), whilst the open ones do not (Figure 2f, CA atom r.m.s.d. 1.75 Å for 275 residue core overlap). This is indicative of the fact that the stereo-chemical mechanism of catalysis is highly conserved between *An*DHQS and *Sa*DHQS, whilst the mechanical basis of domain closure is not.

For reasons of space, the wealth of data generated from comparison of the more than 30 available *An*DHQS and *Sa*DHQS structures cannot all be presented here. The major points can however be summarised as follows. (1) The closed form active sites overlay very tightly with substrate contacting residues showing 100% amino acid identity (Figure 3c). (2) Substrate contacting elements PE1–PE3 are highly conserved (Figure 3c). (3) Distal elements DE1 and DE2 exist in both species but show some alteration in secondary structure and low amino acid sequence identity (Figure 3a and b). (4) The putative "springy" loop regions DE3–DE4 and DE5, which are believed to assist in the reopening of the active site in *An*DHQS, are substantially truncated in *Sa*DHQS (Figure 2f), with a concomitant re-orientation of helix α12 bringing it closer to top of the cleft (Figure 3e). (5) *An*DHQS elements AE4 and AE6 were originally designated thus because conformational changes could only be clearly separated from the background in some comparisons between open and closed complexes, but not others. However, with the newer high-resolution *An*DHQS TAD analysis, it can clearly be seen that both regions undergo torsional changes due to substrate binding (Figure 4a–d). Additionally, the TAD peak for AE6 also exists in the SaDHQS flexion plot, the region of the C domain lying between helices α6, α8a and α10 is thus re-designated as DE6 and the two loops λ_{DE6A} and λ_{DE6B}, which link helices α7 and α10 to α8 and α11, respectively, are designated as DE6A and DE6B, as they show conformational changes associated with those of DE6, but their TADP are at different locations in the flexion plots (Figure 4a–d). (6) The secondary structure is largely conserved between *An*DHQS and *Sa*DHQS, with just the following three changes going from *An*DHQS to *Sa*DHQS. Firstly, the loosely helical *An*DHQS region α4a is deleted from the *Sa*DHQS sequence; secondly, helix α7 has split into two short helices, designated α7 and α7a, with a flexible linking loop region, designated $\lambda\alpha_7$ (Figure 3d); and thirdly, the consecutive β-strand regions β10–β11 and β12–β13 each become conjoined to single elongated β-strand regions ($\beta10_L$ and $\beta12_L$,

Figure 4. a and c, TAD flexion plots scaled by 2σ for *Sa*DHQS and *An*DHQS, respectively. b and d, Cartoon-format colour-coded delineations of structural elements with representative TADP marked for *Sa*DHQS and *An*DHQS, respectively.

respectively; Figure 3f). The difference in chain length is thus largely attributable to the major truncation of loop regions AE2, AE4 and DE4–DE5 (Figure 2f). (7) In all *Sa*DHQS open form crystal structures, direct contacts are retained between the N and C faces of the active site cleft, even in the open form, thereby constraining the maximal angle of hinge opening to ~8° rather than the larger 11–13° seen with *An*DHQS (Figure 2f). (8) The hinge axis still lies approximately parallel with the NAD-binding groove, but has translocated 2–3 Å towards the C face of the cleft, and now passes through helix α8a adjacent to helix α7 rather than α8 as it does with *An*DHQS. (9) The C-domain is more flexible with *Sa*DHQS and a section of it, denoted C2, curls over the rest rotating towards PE1 during domain closure (Figure 2e).

Of the above observations, the most important for understanding the underlying reasons why the DHQS distal elements and domain closure mechanisms show such variation for the two species, whilst the active sites and proximal elements are conserved, is that regarding the retention of contact between domain faces in *Sa*DHQS' open form. (1) It reduces the maximal angle of domain opening, thereby substantially affecting the nature of substrate binding and the initiation of domain closure. (2) Combined with the re-orientation of helix α12 (Figure 3e) and changes to the structure of DE1–DE3, it changes the linkages between structural elements and thus affects the propagation of strain through the protein structure and the possible mechanics of domain closure. (3) It creates an anchor point between PE1/DE2 of the N face and PE3 of the C face. Thus, as their respective parent domains rotate away from one another during domain opening these elements are tensioned by the retained contacts. Therefore, since these elements are also inherently flexible and their basal contact-points, within their respective rigid domain cores, are distributed asymmetrically with respect to these contacts across the cleft, the tensions generated by the re-opening of the active site also cause them to undergo an additional rotation relative to one another about an axis approximately orthogonal to that of the main rigid-body motion. As a result, *Sa*DHQS hinge opening twists and stretches PE1, PE3 and helix α13, reducing their level of order, directly affecting the antiparallel β-sheet portion of PE1 (β7–β8) and indirectly affecting the antiparallel β-sheet portion of DE6, which is continuous to the C-terminal end of PE3 ($\beta10_L$–$\beta12_L$). In both cases this causes the associated open form regions to show a reduction in apparent secondary structure (Figure 2c and d).

As a final point, it is worth noting that throughout our analysis of *Sa*DHQS, it is assumed that the retention of contacts between the N and C faces, in the crystal structure of the *Sa*DHQS open form, is truly representative of the majority conformation in solution. This is obviously a difficult assumption to test and it might be argued that crystallisation samples one of a number of different open conformations in one DHQS, and a different one in the other. However, despite crystallising in many different forms, with very different crystal contacts, all *An*DHQS open form structures overlay very well. Added to this, as is demonstrated by our analysis, detailed comparison of the two species' DHQS structures allows one to explain almost all of the observed structural differences as directly resulting from the effects of retained contacts in the open form of *Sa*DHQS. It therefore seems most likely that the conformations observed do, indeed, each represent the majority conformation in solution for their respective species.

Substrate binding and active site differences

Since the angle of domain movement in *Sa*DHQS is significantly smaller than that observed with *An*DHQS, C-face and N-face substrate binding residues are, in the open form, mostly still very close to their closed form positions. As such, the *Sa*DHQS open-form active site has a smaller more tightly defined pocket than that of *An*DHQS. It is therefore most likely that the substrate DHAP, or substrate analogue CBP, would interact with both faces simultaneously as they first diffuse into the active site. Also, despite the low overall amino acid sequence identity, there is a 100% identity between substrate-contacting residues, with the atoms making contact showing no greater than 0.2 Å separation between *Sa*DHQS and *An*DHQS closed-form overlap positions. The active site stereochemical environment created by the transition to the closed form is thus essentially identical between the two species. It would therefore seem that the structural requirements for effective catalysis are very tightly defined, with a high probability that DHQSs from all species will thus also show structural conservation of their active sites.

Initiation of domain closure

The motive force driving substrate-induced conformational change is generated by an attraction between the substrate and the adjacent residues of the active site. As substrate binds, the attraction induces local conformational changes, pulling the residues into their closed-form positions and generating torsional stresses within the adjacent protein chain. Stresses then propagate through the structure, causing additional alterations of state at distal sites. The two sets of changes then act concertedly to close the active site cleft. PE3 relaxes further than the other proximal elements in the open form and the evidence suggests that its inward rotation is the final stage of completing the active site with conformational changes propagating to it, and affecting it, after the initial binding. Thus, domain closure initiation is largely a function of induced fit conformational changes in PE1 and PE2. Considering the two elements together, both have 100% amino acid sequence identity with *An*DHQS

for their respective contact residues, and C^{α} traces of closed-form structures overlay almost perfectly (Figure 3(c)). PE1 and PE2 also have the same separation distances between their pairs of contacts with substrate. A pattern of torsional stresses similar to those generated between the *An*DHQS residue pairs Lys152/Asn162 and Arg264/Asn268 would therefore also be expected to be generated between the equivalent *Sa*DHQS pairs Lys136/Asn146 and Arg235/Asn239. The first stage of domain closure initiation is thus likely to be very similar between the two species.

Propagation of conformational changes

Distal elements can be divided into two classes, primary DEs, which contact the PE sections that actually move to make contact with the substrate: DE1, DE2 and DE6 (Figures 3a and b, and 4b and d) and secondary DEs which are located further away from the active site. The latter show conformational changes mediated either by the pivoting of a linking SSE: DE3, DE4 and DE5 (Figures 3a and b, and 4b and d), or by the movement of a primary DE to which they are attached: DE6A, DE6B (Figure 4b and d). Considering the primary DEs first, equivalent structures must exist in both *An*DHQS and *Sa*DHQS as the associated PEs undergo non-rigid body motion relative to the rest of their parent domains. As a result, the attached SSEs are necessarily tensioned, engendering a corresponding conformational change irrespective of their actual type/structure. Thus, corresponding TAD peaks can always be delineated for the equivalent sections of the two enzymes' flexion plots even for DE6, despite the actual structures being very different (Figure 4a and c). By contrast, *An*DHQS secondary DE structures show TAD peaks because torsional stresses are able to propagate to them through extended systems, in counterpoint to one another, thereby creating a tensioned system. Such tensioning is likely to exert a restoring force to reopen the pocket at the completion of the reaction cycle. As such, they are entirely dependent on the form of linkages between elements and the way in which this affects the localisation of strain induced by substrate binding. Thus, since this pattern is radically altered by the retention of upper cleft N and C face contacts in the *Sa*DHQS open-form structure, and alterations to the structure of DE1, DE2 and DE6, most of the domain closure associated tensioning now localises to the PE1/DE2 *versus* PE3/C2 counter-rotating system, rather than the smaller *Sa*DHQS sections equivalent to DE3–DE5. TADP are therefore only discernible for these elements in *An*DHQS and not *Sa*DHQS (Figure 4a and c).

Comparison of overall hinge motions

With *An*DHQS, HINGEFIND[14] identifies a two-point hinge axis with an effective rotation arc of 11–13° approximately parallel with the NAD binding site/helix α8 and passing through Met144 in the DI region ($HP1_{AnDHQS}$) and Leu183 ($HP2_{AnDHQS}$) of $\lambda_{\alpha 5a}$ at the distal end of the pocket. This analysis also showed that the domain structure was relatively simple, with $HP2_{AnDHQS}$ marking the boundary between the N and C-terminal domains. Thus, given that CBP binds to the C face of the pocket first, the major rigid-body motion is a simple rotation of the N domain towards the C domain. Single-domain C^{α}-trace overlays also revealed that PE1-3 undergo additional non-rigid body "over-bending" towards the bound substrate to yield the final closed conformation. The tension generated in PE2 by such movement thus causes the "transfer helix" α8 to pivot about Lys250, transferring conformational change to λ_{DE4} at the distal end of the pocket, which is itself cross-linked to $HP2_{AnDHQS}$. Given that open-form structures show very little difference in the angle of opening, it was thus postulated that, in the absence of bound substrate, cross-linking between λ_{DE4}, α8, $\lambda_{\alpha 5a}$ and α5a holds the hinge open. The PE2 change induced movement of α8 then released this constraint and the hinge is able to close.

With *Sa*DHQS, the situation is very different. The N and C domains do not fully detach in the open form, remaining connected at the top of the cleft. Domain opening therefore strains retained N and C face contacts, generating an additional set of tension vectors between contacting elements. This strain is then accommodated partly by local distortion, giving rise to TADP for DE2, PE1 and PE3 (Figure 4a and b) and, in the case of PE1 and PE3, also partly by movement relative to adjacent structural elements, as they counter-rotate opening out the active site (Figure 2e). The sub-domain structure is therefore also affected and four regions are now delineated: (1) N1, the main N-terminal rigid-body domain; (2) N2, the extended PE1 element, running from the active site to the distal hinge point, which shows TADP at either end as it undergoes an additional rotation into the active site in response to substrate binding (PE1A-B, Figure 4a and b); (3) C1, the main C-terminal rigid-body domain, which rotates towards the N-terminal domain about the axis shown in Figure 2e, giving rise to TADP at the relative hinge-points (HP1-2, Figure 4a and b) and (4) C2, the PE3-containing sub-domain, which curls over the surface of C2, but is constrained by linkages to PE1 causing it to undergo an additional rotation around the fixed axis generated by the retention of contact with PE1, hence giving rise to flexion TADP at the distal sites marked C2A, C2B and C2C (Figures 2e and 4a and b). This probably also explains the splitting of helix α7, allowing better accommodation of the stresses thus engendered in the distal region connected to the C2 sub-domain.

With quantitative HINGEFIND analysis, care had to be taken in the selection of partition values, but by keeping them relatively high a two-domain single-axis result was obtained, indicating the gross hinge movement is of significantly greater

magnitude than the smaller changes in orientation between the sub-domains. At first glance, the results of this two-domain analysis also appear similar to those of *An*DHQS, with a hinge axis running approximately parallel with the NAD-binding site (Figure 2e). However, the effective arc of rotation between domains with *Sa*DHQS is only about 8°, as compared to 11–13° with *An*DHQS. The axis has shifted substantially across the pocket towards the C face, with concomitant changes in overall domain structure (helix α7a and α8 (residues 199–229) now falling within the N-terminal rather than C-terminal domain). Additionally, the axis now passes through the protein chain at three points rather than two. (1) At the proximal end of the cleft the axis passes through helix α8a, which undergoes direct conformational change when substrate binds such that the HP2 and PE2 TADP co-locate in *Sa*DHQS's flexion plot (Figure 4a). (2) At the distal end of the cleft the axis passes through the C-terminal end of helix α7a, giving rise to TADP in the composite flexion plot (HP1, Figure 4a). (3) In the middle of the cleft the axis passes through the centre of helix α6, but α6 is almost orthogonal to the axis, and the corresponding TADP occur at the end of the helix where it connects to helix α8a and DE6 (DE6A, Figure 4a and b), not where the axis passes through α6. Such features therefore indicate that this region forms a central pivot, about which α6 rotates during domain closure, rather than a true hinge-point.

Evolutionary differences between prokaryotic and eukaryotic DHQS and significance for drug design

Overall then, we see that both *An*DHQS and *Sa*DHQS undergo domain closure and have evolved highly complex mechanical systems to facilitate the rearrangements involved. The two systems are also similar, in that the closed form represents a strained conformation with both the bound substrate and protein structural elements being under tension. In the former case, the patterns of tension are the same between the two species, lending additional weight to the hypothesis that the application of these mechanical tensions to the bound substrate may actually increase turnover rates by directly lowering the energy required for bond cleavage. In the latter case, however, whilst the restoring forces from PE1 and PE2 will be similar, those associated with other elements will not. The counter-rotation of PE1/DE2 and PE3/C2 proximal to the active site probably exerts a restoring force in *Sa*DHQS, but not in *An*DHQS. Conversely, the tensioning of DE3–DE5 will create a restoring force in *An*DHQS, but not in *Sa*DHQS. Such differences in mechanical control of hinge operation arise naturally from the retention of N and C face contacts in the *Sa*DHQS open form, which changes the way strain can propagate through the protein structure. Thus, the point of evolutionary divergence from their common ancestor was probably either mutation preventing separation, as in *Sa*DHQS, or mutation allowing separation, as in *An*DHQS, with the two species then evolving divergent systems to optimise their function. Obviously it is possible that this represents a random phenomenon with no particular advantage, or disadvantage, being derived from the differences in mechanical systems. However, if we consider the following observations, an alternative hypothesis presents itself. Firstly, if we examine the potential effect of mutations on the functioning of the two species' mechanical elements, the *An*DHQS DE3–DE5 elements are further away from the point where conformational change is initiated and propagation must therefore occur over significantly greater distances. This greater separation might therefore be expected to have been selected against by genetic drift, as mutations preventing effective domain closure would cripple the resultant progeny and, due to the increased separation, there are simply more vulnerable residues in the *An*DHQS structure than in that of *Sa*DHQS. Such selection has obviously not occurred, as all three elements are at the extremes of spatial separation from the substrate-binding site, suggesting that an additional selective pressure has operated to cause the changes in organisation, leading to these distal elements moving further away from the active site, and the active site opening out to a far greater degree.

Secondly, in *A. nidulans*, fivefold gene fusion[13] has led to DHQS becoming one part of the penta-domain AROM complex rather than a stand-alone enzyme as it is in *S. aureus* and, since the active site of DHQS is only complete as a dimer, it must also be dimerised within the active AROM complex.[15] In its native state, *An*DHQS therefore probably has numerous contacts to other domains of AROM, which will naturally affect both the form of conformational changes that can possibly be accommodated and the way in which these changes can then propagate through *An*DHQS's structure. As such, it is a reasonable hypothesis that the observed structural and functional differences between *An*DHQS and *Sa*DHQS have arisen, and been maintained, as a result of *An*DHQS evolving interactions with other portions of AROM in a way that *Sa*DHQS does not interact with the equivalent stand-alone shikimate pathway enzymes. It is obviously difficult to be certain of the consequences of these interactions without a solved structure of the full AROM complex but certain possibilities do suggest themselves in the light of these and other data.

Firstly, if we consider the metabolic flux analysis experiments conducted with *A. nidulans* AROM,[16] they show that there is a channelling effect but that the flow of material through the AROM complex is generally leaky as an up-regulation of catabolic dehydroshikimate dehydratase can siphon DHQ from the pathway. Thus, in the converse case where suitable metabolites are available to the host, DHQ produced as part of the quinate pathway would also be expected to be able to diffuse into the AROM

complex. There is therefore a need for the cell to prevent substantial quantities of this material flowing through the remainder of the shikimate, rather than quinate, pathway and the mechanical interlinking of different domains such that their cycling is co-dependent would provide a very elegant solution. In this context, it is also interesting to note that EPSPS (the adjacent enzyme by primary sequence, which also forms a sub-domain with DHQS in the *A. nidulans* AROM complex,[15] and DHQase,[17] both also show substrate-induced conformational changes. DHQase and EPSPS are therefore both potential candidates for a "see-saw" interaction with DHQS. In the former case, this would affect the flux through DHQase directly, by reducing its turnover rate unless DHQS was also cycling. In the latter case the effect would be similar, but achieved more indirectly, as retardation of EPSPS would lead to an accumulation of earlier metabolites and hence also retardation of the shikimate pathway enzymes intermediary to DHQS and EPSPS.

Secondly, although the flow of material through AROM was found to be generally leaky, substrate channelling was demonstrated at a level that it is believed would be physiologically significant in times of poor nutrient and oxygen supply.[16] Thus, the differences in structure might also reflect the evolution of this channelling effect with the positive benefit of increased metabolic control in times of stress providing a selective pressure to drive their evolutionary divergence.

The implications of this comparison, of *An*DHQS and *Sa*DHQS, for drug design are also significant. Any drug designed to bind at the distal end of the pocket, or that works by interrupting the propagation of conformational change would be unlikely to have broad-spectrum activity, as these regions are highly divergent between *An*DHQS and *Sa*DHQS. Conversely, any drug binding in the proximal portion of the active site might be expected to have broad-spectrum activity. Additionally, if the hypothesis regarding the interlinking of turnover of separate domains of AROM by contacts between their respective elements that undergo conformational change is correct, then combination therapies attacking both AROM targets simultaneously might not be substantially more effective than single therapy, as their binding might thus be mutually exclusive.

As can be seen across the preceding analysis, despite the mediocre quality of the initial crystals of *Sa*DHQS, a great deal of useful structural information could still be obtained. The technique of post crystal growth dehydration to substantially improve data quality has again shown its general utility, as with our earlier studies,[18,19] to aid effective generation of structural data. The use of multiply averaged TAD flexion plots, as a guide to delineation of structural elements, was critical to the success of the structural analyses reported here and follows similar lines to our earlier analyses of *An*DHQS alone[9] and *E. coli* aspartate β-semialdehyde dehydrogenase.[18] It can therefore be seen that this kind of detailed TAD analysis provides an extremely powerful tool for elucidating the often complex interplay of structural elements that occur during the cycles of domain movement concomitant to an enzyme catalytic cycle.

Experimental Procedures

Cloning and protein purification

The *S. aureus* DHQS coding sequence was PCR amplified and sub-cloned into the *E. coli* expression vector pET3d to yield the recombinant plasmid pMUT27, which was transformed into the *E. coli* expression strain BL21 DE3 pLysS. Transformants were selected using ampicillin. For large-scale expression, cultures were grown in Luria broth at 37 °C and induced with IPTG for five hours. Cells were harvested by centrifugation, resuspended in phosphate (pH 7.0), 1 mM DTT, 10 μM $ZnSO_4$, 1 mM benzamidine and sonicated with cell debris removed by centrifugation. All subsequent steps were carried out at 4 °C. The protein was purified by Q-Sepharose column chromatography, ammonium sulphate fractionation between 30% and 55% saturation, followed by Sephacryl S-300 gel-filtration and separation on hydroxyapatite columns. Active fractions in the eluate were analysed by SDS-PAGE using a 10% (w/v) separating gel, and suitable fractions pooled.

Crystallisation and post-growth processing

For all *Sa*DHQS crystallisations, aliquots were concentrated and buffer-exchanged into 10 mM Tris (pH 7.4), 40 mM KCl using Vivascience Vivaspin2 concentrators with polyethersulphone membranes. Enzyme was filtered through Amersham NAP™ 25 columns, re-concentrated to 40 mg ml^{-1}, flash-frozen and stored at −80 °C. For some experiments involving CBP-only co-crystallisation, residual co-purified NAD levels were reduced by incubating *Sa*DHQS at 0.5 mg ml^{-1} for 12 hours with 1 mM ADP, 20% (v/v) glycerol, 0.1 M Tris (pH 7.4), 250 mM KCl, followed by buffer exchange to remove salts, glycerol and nucleotides.

For crystallisation trials, sparse matrix and Grid screening kits were used, i.e. Crystal Screen I, Crystal Screen II, Crystal Screen Cryo, PEG/Ion, Natrix, MembFac, PEG/LiCl Grid, NaCl Grid, PEG6000 Grid and A/S Grid (Hampton Research Corporation), with *Sa*DHQS at 20–40 mg ml^{-1}, a 1 : 1 mixing ratio and 6 μl total initial droplet size. Sitting-drop, vapour-diffusion crystallisations were set up at 277 K utilising microbridges.[20] Optimisation of hits used finer intervals of pH/precipitant concentration and the introduction of additives.

With the exception of the single MPD-derived crystal used in dataset I, all *Sa*DHQS crystals used were grown from 7–18% PEG3350, 0.2 M KCl or K_2SO_4 and 0.1 M Hepes or Tris at pH 7–8. After

initial growth, crystals were partially dehydrated by transfer to their well solution, followed by successive washes with increasing DMSO and ethylene glycol (final concentrations being 15–30% DMSO and 7.5–15% ethylene glycol). This protocol improved diffraction limits and scaling statistics. Datasets were collected in-house using a Mar345 image plate system on a Rigaku generator equipped with a Cu anode and Osmic multilayer optics to give Kα radiation (λ=1.5418 Å) and at ESRF: beamline ID14EH4 (λ=0.9333 Å), in all cases at 100 K.

Data processing, molecular replacement and model refinement

Indexing, integration and merging of data images used DENZO and SCALEPACK.[21,22] In some instances datasets collected from more than one crystal of a particular complex were merged to improve data quality. Rotation function and translation searches together with rigid-body refinement were carried out using CNS.[23–25] Molecular replacement solutions were checked using O.[26] Residues 70–330 of *An*DHQS (pdb 1NR5) were used as the search model for rotation function searches with dataset I. CNS was also used for further refinement of models using positional, *B*-factor and simulated annealing refinement and initial water picking. Manual model rebuilding was carried out using O. The overall scheme of structure solution was: two to four rounds of rigid-body refinement with increasing resolution range and subdivision into smaller fragments, followed by positional refinement; *B*-factor refinement; simulated annealing; fitting of ligands and 6–12 rounds of rebuilding in O, alternated with positional and *B*-factor refinement. Dataset III was solved using the structure determined from dataset I. Datasets IV and V were similarly solved using the structure determined from dataset III. Auto-indexing with DENZO indicated tetragonal symmetry for all *Sa*DHQS crystals. The initial dataset (dataset I) was solved and refined satisfactorily as $P4_32_12$. However, some later datasets showed the presence of $h00$ reflections with odd indices, indicative of a lower-symmetry space group. Each dataset was then processed in the various space group options from $P1$ to $P4_32_12$ and the separation of *R*-factors and quality of final model geometry assessed for each, leading to the assignment of form *A* as $P4_32_12$ and form *B* as $P4_3$.

Model analysis

A full analysis of the *Sa*DHQS domain closure mechanism required careful comparison with data from *An*DHQS, using a combination of data from several sources. (1) Model fitness was evaluated during rebuilding using PROCHECK,[27] r.m.s.d. variation between structures was evaluated with the program TOP3D[28] and EDPDB was used to evaluate mean *B*-factors.[29] (2) The VMD STRIDE[30] plugin was used to delineate SSEs for both proteins and the results visualised in VMD for comparison.[31] Where possible, *Sa*DHQS β-strands and αα-helices were given the same designations as in *An*DHQS. Cartoon-format representations of open and closed form *Sa*DHQS C^{α}-traces, showing SSE designations together with their "all-atom" SURF plots are shown in Figure 1d and e, respectively. (3) TOP3D was also used extensively to align different structures with different ligand sets from both proteins and the resultant overlays analysed visually using VMD and O. (4) C^{α} torsion angles (TAs) for each residue i, defined as the dihedral angle $C^{\alpha}(i-1)-C^{\alpha}(i)-C^{\alpha}(i+1)-C^{\alpha}(i+2)$, were calculated as described by Flocco & Mowbray.[32] Using programs supplied by the authors, TAD sets were then calculated for every available pair of chain comparisons, including those within the same crystal, with the following direction of comparison to prevent cancellation of domain movement-related TAD peaks: NAD/CBP→NAD/$CBP_{(d)}$→CBP soak of $NAD_{(d)}$→$CBP_{(d)}$→$NAD_{(d)}$, (the first structure in the list was used as the reference, the second as the comparative and final mean TAD values for each set ranged from 5–11°). All TAD sets were then averaged together to obtain a composite flexion plot, flattening lattice strain-induced distortions of normally rigid sections, variations due to inherent flexibility and positional uncertainties (final mean TAD for composite flexion plot= 2.13°). The same general approach was also applied to the *An*DHQS TAD data to facilitate direct comparison and, since the full set of TADP for all *An*DHQS comparisons was already known, to test the concept that all flexion points should be visible in a single plot of this kind. Additional flattening was seen for the inclusion of TADs from the comparison of chains within the same crystal in Figure 3c but here, the higher resolution, lower overall flexibility and greater magnitude of change gives a much better signal-to-noise ratio anyway and the TADP are completely interpretable, even for single-chain comparisons. (5) Quantitative assessment of the effective arcs of rotation and hinge-axis points in *Sa*DHQS's substrate binding-induced domain closure was carried out with the TCL script HINGEFIND (final statistics presented are calculated with a partition value=1.0 and max_domains=2, where partition value is a user-determined variable controlling the allowed level of deviation of an individual residue's arc of movement from that calculated for the bulk domain, before it is no longer counted as being part of that bulk domain), run and visualised with the graphics package VMD.[31] (6) Crystal contacts were identified with the program contacts, cluster groupings evaluated with the program cluster, (kindly supplied by R. Esnouf), and visualised with VMD. (7) Solvent-accessible surfaces were calculated with the SURF[33] plugin and also visualised in VMD.

With the exception of Figures 1 and 2a (screenshots from NCBI blast-p alignment, and O, respectively), and 4a and c (graph outputs from EXCEL), all other Figures feature modified VMD

screenshots. All final Figures were arranged using Corel Graphics Suite 11.

Protein Data Bank accession codes

Coordinates for crystals I to V have been deposited with the RCSB Protein Data Bank under accession codes 1XAG, 1XAH, 1XAI, 1XAJ and 1XAL, respectively.

Conflict of Interest Statement

A.R.H., I.C. and D.K.S. are co-founders of Arrow Therapeutics Ltd, for whom A.R.H. and D.K.S. also act as consultants.

Acknowledgements

We thank the staff at the ESRF, Grenoble, France and at SRS for help with the synchrotron data collection. Dr R. Esnouf, Mr A. Turner and Ms J. Dong are acknowledged for computer support. We thank Arrow Therapeutics Ltd for financial support.

References

1. Bentley, R. (1990). The shikimate pathway – a metabolic tree with many branches. *Crit. Rev. Biochem. Mol. Biol.* **25**, 307–384.
2. Gunel-Ozcan, A., Brown, K. A., Allen, A. G. & Maskell, D. J. (1997). *Salmonella typhimurium* aroB mutants are attentuated in BALB/c mice. *Microb. Pathog.* **23**, 311–316.
3. Widlanski, T., Bender, S. L. & Knowles, J. R. (1989). Dehydroquinate synthase: the use of substrate analogues to probe the late steps of the catalyzed reaction. *Biochemistry*, **28**, 7572–7582.
4. Srinivasan, P. R., Rothchild, J. & Sprinson, D. B. (1963). The enzymatic conversion of 3-deoxy-D-arabino-heptulosinic acid 7-phosphate to 5-dehydroquinate. *J. Biol. Chem.* **238**, 3176–3182.
5. Hawkins, A. R., Lamb, H. K., Moore, J. D., Charles, I. G. & Roberts, C. F. (1993). The pre-chorismate (shikimate) and quinate pathways in filamentous fungi: theoretical and practical aspects. *J. Gen. Microbiol.* **139**, 2891–2899.
6. Bender, S. L., Widlanski, T. & Knowles, J. R. (1989). Dehydroquinate synthase: the use of substrate analogues to probe the early steps of the catalyzed reaction. *Biochemistry*, **28**, 7560–7572.
7. Carpenter, E., Hawkins, A., Frost, J. & Brown, K. (1998). Structure of dehydroquinate synthase reveals an active site capable of multistep catalysis. *Nature*, **394**, 299–302.
8. Nichols, C. E., Ren, J., Lamb, H., Haldane, F., Hawkins, A. R. & Stammers, D. K. (2001). Identification of many crystal forms of *Aspergillus nidulans* dehydroquinate synthase. *Acta Crystallog. sect. D*, **57**, 306–309.
9. Nichols, C. E., Ren, J., Lamb, H. K., Hawkins, A. R. & Stammers, D. K. (2003). Ligand-induced conformational changes and a mechanism for domain closure in *Aspergillus nidulans* dehydroquinate synthase. *J. Mol. Biol.* **327**, 129–144.
10. Brown, K. A., Carpenter, E. P., Watson, K. A., Coggins, J. R., Hawkins, A. R., Koch, M. H. & Svergun, D. I. (2003). Twists and turns: a tale of two shikimate-pathway enzymes. *Biochem. Soc. Trans.* **31**, 543–547.
11. Nichols, C., Hawkins, A. & Stammers, D. (2004). Structure of the open form of *Aspergillus nidulans* 3-dehydroquinate synthase, at 1.7 Å resolution, from crystals grown following enzyme turnover. *Acta Crystallog. sect, D*, **60**, 971–973.
12. Bartlett, P. A., McLaren, K. L. & Marx, M. A. (1994). Divergence between the enzyme-catalysed and non-catalysed synthesis of 3-dehydroquinate. *J. Org. Chem.* **59**, 2082–2085.
13. Hawkins, A. R. (1987). The complex Arom locus of *Aspergillus nidulans*. Evidence for multiple gene fusions and convergent evolution. *Curr. Genet*, **11**, 491–498.
14. Wriggers, W. & Schulten, K. (1997). Protein domain movements: detection of rigid domains and visualization of hinges in comparisons of atomic coordinates. *Proteins: Struct. Funct. Genet.* **29**, 1–14.
15. Hawkins, A. R. & Smith, M. (1991). Domain structure and interaction within the pentafunctional arom polypeptide. *Eur. J. Biochem.* **196**, 717–724.
16. Lamb, H. K., van der Hombergh, J. P. T. W., Newton, G. H., Moore, J. D., Roberts, C. F. & Hawkins, A. R. (1992). Differential flux through the quinate and shikimate pathways. Implications for the channelling hypothesis. *Biochem. J.* **284**, 181–187.
17. Lee, W. H., Perles, L. A., Nagem, R. A., Shrive, A. K., Hawkins, A., Sawyer, L. & Polikarpov, I. (2002). Comparison of different crystal forms of 3-dehydroquinase from *Salmonella typhi* and its implication for the enzyme activity. *Acta Crystallog. sect. D*, **58**, 798–804.
18. Nichols, C. E., Dhaliwal, B., Lockyer, M., Hawkins, A. R. & Stammers, D. K. (2004). High-resolution structures reveal details of domain closure and "half-of-sites-reactivity" in *Escherichia coli* aspartate beta-semialdehyde dehydrogenase. *J. Mol. Biol.* **341**, 797–806.
19. Nichols, C. E., Lockyer, M., Hawkins, A. R. & Stammers, D. K. (2004). Crystal structures of *Staphylococcus aureus* type I dehydroquinase from enzyme turnover experiments. *Proteins: Struct. Funct. Genet.* **56**, 625–628.
20. Harlos, K. (1992). Microbridges for sitting drop crystallisations. *J. Appl. Crystllog.* **25**, 536–538.
21. Otwinowski, Z. & Minor, W. (1996). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
22. Otwinowski, Z. (1993). *Data Collection and Processing*, Daresbury Laboratory, Warrington, UK.
23. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse, K. R. W. *et al.* (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
24. Brunger, A. T., Krukowski, A. & Erickson, J. (1990). Slow-cooling protocols for crystallographic refinement by simulated annealing. *Acta Crystallog. sect. A*, **46**, 585–593.
25. Brunger, A. T. (1990). Extension of molecular replacement: a new search strategy based on Patterson correlation coefficient refinement. *Acta Crystallog. sect. A*, **46**, 46–57.
26. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M.

(1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A*, **47**, 110–119.
27. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.
28. Collaborative Computational Project Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
29. Zhang, X. J. & Matthews, B. W. (1995). A multifunctional tool for protein structure analysis. *J. Appl. Crystallog.*, 624–630.
30. Frishman, D. & Argos, P. (1995). Knowledge-based protein secondary structure assignment. *Proteins: Struct. Funct. Genet.* **23**, 566–579.
31. Humphrey, W., Dalke, A. & Schulten, K. (1996). VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 27–28.
32. Flocco, M. M. & Mowbray, S. L. (1995). C alpha-based torsion angles: a simple tool to analyze protein conformational changes. *Protein Sci.* **4**, 2118–2122.
33. Nicholls, A., Bharadwaj, R. & Honig, B. (1993). GRASP: graphical representation and analysis of surface properties. *Biophys. J.* **64**, 166–170.

Edited by R. Huber

(Received 30 April 2004; received in revised form 6 August 2004; accepted 11 August 2004)