

The restriction enzyme *SgrAI*: structure solution via combination of poor MIRAS and MR phases

Pete W. Dunten,^{a*} Elizabeth J. Little^b and Nancy C. Horton^b

^aStanford Synchrotron Radiation Laboratory, Stanford University, Menlo Park, CA 94025, USA, and ^bDepartment of Biochemistry and Molecular Biophysics, University of Arizona, Tucson, AZ 85721, USA

Correspondence e-mail: pete@slac.stanford.edu

Received 30 October 2008

Accepted 27 January 2009

Uninterpretable electron-density maps were obtained using either MIRAS phases or MR phases in attempts to determine the structure of the type II restriction endonuclease *SgrAI* bound to DNA. While neither solution strategy was particularly promising (map correlation coefficients of 0.29 and 0.22 with the final model, respectively, for the MIRAS and MR phases and *Phaser Z* scores of 4.0 and 4.3 for the rotation and translation searches), phase combination followed by density modification gave a readily interpretable map. MR with a distantly related model located a dimer in the asymmetric unit and provided the correct transformation to use in averaging electron density between *SgrAI* subunits. MIRAS data sets with low substitution and MR solutions from only distantly related models should not be ignored, as poor-quality starting phases can be significantly improved. The bootstrapping strategy employed to improve the initial MIRAS phases is described.

1. Introduction

While the current focus of structure-solution pipelines is on MAD phasing *via* selenomethionine (SeMet) substitution of proteins expressed in *Escherichia coli*, this approach is not always feasible and one must fall back on heavy-atom soaking methods or, in suitable cases, molecular-replacement methods. The restriction enzyme *SgrAI* is one such example: crystals grown from SeMet-substituted protein diffracted only poorly, while unsubstituted protein yielded crystals diffracting to better than 2 Å resolution. In favorable cases of the heavy-atom soaking method, a single derivative may be sufficient to solve a new structure. Full substitution of occupied sites, anomalous data and isomorphism are all helpful. Similarly, in favorable cases a correct molecular-replacement solution may be identified and prove sufficient for determination of a novel structure. In the case described here, phases from two weak derivatives and a poor molecular-replacement model were combined to achieve the desired result of an interpretable map and finally an atomic model (Dunten *et al.*, 2008). A similar strategy, also employing electron density as a molecular-replacement search model, was used in the structure determination of UDP-galactopyranose mutase (Sanders *et al.*, 2001).

The restriction enzyme *SgrAI* is active as a dimer and converts to higher molecular-weight forms upon DNA binding (Daniels *et al.*, 2003). Because the cleavage site recognized by *SgrAI* is eight base pairs in length, the enzyme is useful for genomic mapping studies. *SgrAI* recognizes the cleavage-site sequence CR|CCGGYG (where | denotes the cut site), which is related to the shorter cleavage sites of *NgomIV* (G|CCGGC), *Cfr10I* (R|CCGGY) and *Bse634I* (R|CCGGY). The X-ray structures of the latter three enzymes are all known (Deibert *et al.*, 2000; Bozic *et al.*, 1996; Grazulis *et al.*, 2002). The primary sequences of *Bse634I* and *Cfr10I* share only 21% and 18% sequence identity with *SgrAI*, yet were the most promising molecular-replacement models available in the Protein Data Bank.

Table 1

Diffraction data.

Values in parentheses are for the highest resolution shell.

	Form 1			Form 2	Form 3
	Native	K ₂ PtCl ₆	Hg(SCN) ₂		
Beamline	SSRL BL7-1	SSRL BL1-5	SSRL BL9-1	SSRL BL9-2	SSRL BL9-2
DNA [nucleotides]	Noncognate [17]	Noncognate [17]	Noncognate [17]	Cognate [18]	Cognate [18]
Wavelength (Å)	1.075	1.0716	0.984	1.009	1.001
Space group	<i>P</i> ₂ ₁	<i>P</i> ₂ ₁	<i>P</i> ₂ ₁	<i>P</i> ₂ ₁	<i>C</i> 222 ₁
Unit-cell parameters (Å, °)	<i>a</i> = 69.73, <i>b</i> = 78.62, <i>c</i> = 86.83, β = 109.76	<i>a</i> = 70.04, <i>b</i> = 79.38, <i>c</i> = 86.69, β = 109.71	<i>a</i> = 69.68, <i>b</i> = 78.30, <i>c</i> = 86.91, β = 109.82	<i>a</i> = 59.14, <i>b</i> = 119.25, <i>c</i> = 123.80, β = 95.60	<i>a</i> = 122.95, <i>b</i> = 49.69, <i>c</i> = 122.45
<i>Sgr</i> AI monomers per ASU	2	2	2	4	1
Matthews coefficient (Å ³ Da ⁻¹)	2.60	2.60	2.60	2.49	2.14
Resolution (Å)	1.91 (1.95–1.91)	2.50 (2.56–2.50)	2.24 (2.31–2.24)	1.89 (1.96–1.89)	3.00 (3.08–3.00)
Total observations	157240 (7049)	96133 (3864)	91302 (8523)	429421 (34710)	29191 (2096)
Unique observations	66598 (3420)	29836 (1716)	39972 (3695)	126082 (10971)	7806 (562)
Completeness (%)	96.1 (66.8)	96.1 (76.2)	94.2 (95.1)	92.8 (81.0)	99.7 (99.9)
<i>I</i> / <i>σ</i> (<i>I</i>)	6.8 (1.5)	5.3 (0.4)	6.1 (1.9)	6.5 (2.2)	4.7 (0.7)
Multiplicity	2.4 (2.1)	3.2 (2.3)	2.3 (2.3)	3.4 (3.2)	3.7 (3.7)
<i>R</i> _{merge} [†]	0.071 (0.483)	0.126 (1.48)	0.073 (0.386)	0.120 (0.532)	0.141 (1.05)
PDB code	3dpg			3dvo	

[†] $R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$, where $\langle I(hkl) \rangle$ is the mean intensity over all symmetry-related and equivalent reflections for reflection *hkl* and the $I_i(hkl)$ are the observed intensities for reflection *hkl*.

This level of sequence conservation is well into the ‘twilight zone’ of molecular replacement, where success is not guaranteed and recognizing correctly placed solutions can be difficult.

2. Methods

MOLEMAN2 (Kleywegt, 1996) was used to generate polyserine versions of the molecular-replacement models. The *Cfr10I* model included residues 55–96, 99–139, 143–211, 215–220 and 223–283. MR searches were performed using *Phaser* (McCoy *et al.*, 2007). *LSQMAN* (Kleywegt & Jones, 1994) was used for structure superpositions and calculation of r.m.s.d. values. Crystals were grown *via* hanging-drop vapor-phase diffusion as described in Dunten *et al.* (2008). The *Sgr*AI concentration was 10–30 mg ml⁻¹ in a storage buffer consisting of 10 mM HEPES pH 7.5, 150 mM NaCl, 1 mM EDTA, 1 mM DTT. The DNA oligonucleotides were self-complementary, 17–18 residues in length and included the *Sgr*AI recognition sequence. The protein was mixed with DNA and 5–10 mM CaCl₂, MnCl₂ or MgCl₂ to give a 1:2 molar ratio of *Sgr*AI dimer:DNA duplex. The hanging drops contained 1.5–3.0 μl protein:DNA mixture and 1.0–1.5 μl precipitating solution [25–21% PEG 4K, 0.15–0.20 M NaCl, 0.1 M buffer (imidazole pH 6.5 or HEPES pH 7.5)] and were equilibrated over 1 ml precipitating solution at 290 K. Heavy-atom soaks were performed by transferring crystals to a stabilization solution (25% PEG 4K, 0.3 M NaCl, 0.1 M buffer) containing 1 mM heavy-atom compound for 12–24 h. Crystals were then cryoprotected by transfer to 25% PEG 4000, 0.3 M NaCl, 0.1 M buffer, 30% glycerol and frozen in liquid nitrogen prior to data collection. Attempting to soak crystals in higher concentrations of the heavy-atom compounds resulted in cracked crystals. Data were collected from the Hg and Pt derivatives at energies just above the *L*_{III} edges of the metals. Data integration and scaling were performed with *HKL-2000* (Otwinowski & Minor, 1997), *d*TREK* (Pflugrath, 1999) or *MOSFLM* (Leslie, 1992) and *SCALA*. The direction of the twofold noncrystallographic symmetry axis was determined with *POLARRFN* from the *CCP4* package (Collaborative Computational Project, Number 4, 1994). An initial averaging operator consistent with the self-rotation function result was determined with *FIND2FOLDS* (Dunten & Hennig, 2002). Difference Patterson maps were calculated for potential derivatives using data from 12 to 5 Å resolution. Refinement of heavy-atom

parameters and calculation of phases were performed using *MLPHARE* from the *CCP4* program suite (Otwinowski, 1991). The initial phase estimates were improved *via* density modification using *DM* from the *CCP4* suite (Cowtan, 1994). The DNA strands were built manually with *O* (Jones & Kjeldgaard, 1997). Refinement was performed with *REFMAC* (Murshudov *et al.*, 1997) and *phenix.refine* (Adams *et al.*, 2002). Fig. 2 was created with *MAPROT* (Stein *et al.*, 1994) and *MAPSLICER* from the *CCP4* program suite. Figs. 3 and 4 were created with *CCP4MG* (Potterton *et al.*, 2004).

3. Results and discussion

3.1. Multiple isomorphous replacement with anomalous scattering (MIRAS)

Multiple crystal forms of *Sgr*AI bound to DNA were obtained with different DNA sequences and divalent cations (Ca²⁺, Mn²⁺ or Mg²⁺), all of which were consistent with a dimeric form of the enzyme (Table 1). A number of heavy-atom compounds could be soaked into the form 1 *P*₂₁ crystals of *Sgr*AI without loss of diffraction and data sets were collected from several mercury, platinum and lead compounds. With two molecules per asymmetric unit in the *P*₂₁ crystal form, a pseudo-Harker plane occurs normal to the direction of the noncrystallographic twofold symmetry axis (Figs. 1 and 2*b*). A cross-peak in this plane was used as a starting point to locate two Hg atoms in the Hg(SCN)₂ derivative using *RSPS* from the *CCP4* program suite (Knight, 2000). Phases from the Hg derivative were used to identify two Pt sites in the K₂PtCl₆ derivative and two additional Pt sites with lower occupancy were subsequently identified when combined MIRAS plus molecular-replacement phases were available. The phasing power of the two derivatives was modest (Table 2). One hand of the heavy-atom sites gave a clearly better map than the other after density-modification with *DM* (Fig. 3*a*). Although the solvent boundary was clear, the map was not of sufficient quality to interpret. As described below, the initial choice of averaging transformation for averaging the electron density between monomers was not correct. Without an accurate description of the noncrystallographic symmetry, the optimal density-modification protocol could not be used to improve the MIRAS phases.

MR was pursued at this point, since potential solutions could be confirmed by verifying that the molecular-replacement phases could

locate the known heavy-atom sites in $F_{\text{deri}} - F_{\text{nati}}$ difference maps. Molecular replacement was not the preferred route initially, given the low sequence identity between *SgrAI* and the potential structural homologs.

3.2. Molecular replacement (MR)

The choice of MR models was made based on threading results returned from the threading server at <http://inub.cse.buffalo.edu> (Fischer & Eisenberg, 1996). The server determines which folds in the Protein Data Bank are most compatible with the input sequence and generates an alignment of the input sequence on the identified folds. Molecular replacement was pursued in both the $P2_1$ crystal form with two molecules per asymmetric unit (form 1 in Table 1) and a $C222_1$ form with a single molecule per asymmetric unit (form 3). While a solution was expected to be easier to obtain in the $C222_1$ form with one molecule per asymmetric unit, a solution in the $P2_1$ form was more desirable given the limited extent of diffraction from the $C222_1$ crystals. A candidate solution was obtained in the $C222_1$ crystal form with a 219-residue model from PDB entry 1cfr (Bozic *et al.*, 1996). The model accounts for less than half of the mass of the protein–nucleic acid complex in the asymmetric unit. The Z scores from *Phaser* for the rotation and translation functions were 4.0 and 4.3, respectively. The R factor after rigid-body refinement in *REFMAC* was 50%. The crystallographic twofold axis along $[1, 0, 0]$ in the $C222_1$ cell generates the complete dimer of *SgrAI*. The corresponding *Cfr10I* dimer was placed in the $C222_1$ cell by superpositioning a single chain of the dimer with the *SgrAI* model. Comparison of the *Cfr10I* dimer with the *SgrAI* dimer revealed that the twofold axis of the *Cfr10I* dimer was misaligned by 5.4° with the $C222_1$ crystallographic twofold axis responsible for generating the *SgrAI* dimer. Hence, a dimeric *Cfr10I* search model was unlikely to be useful for MR in the *SgrAI* $P2_1$ crystal forms. Instead, the electron density encompassing a complete dimer was used as a search model in *Phaser* and a

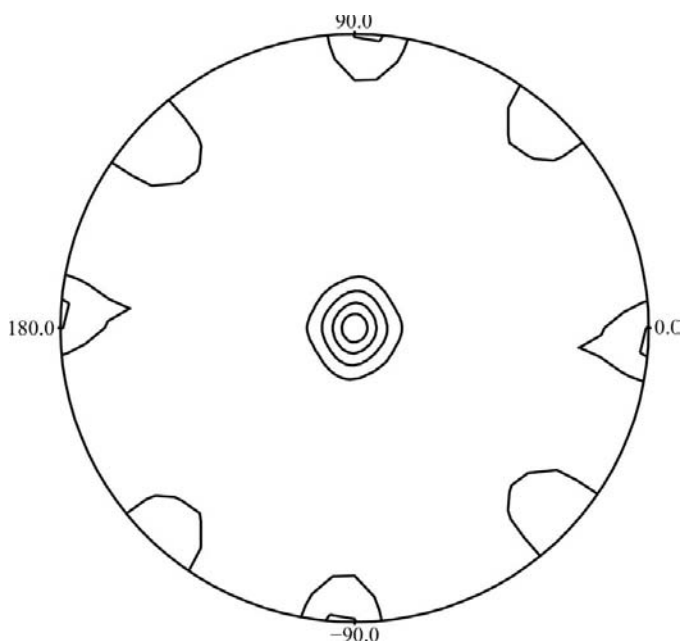


Figure 1
Self-rotation function. A stereographic projection of the $\kappa = 180^\circ$ section is shown for the $P2_1$ (form 1) native data. The crystallographic 2_1 axis generates the peak in the center of the projection at $\omega = 0^\circ$. The peaks arising from noncrystallographic twofold symmetry appear at $\omega = 90^\circ$ along the periphery of the projection, at φ values close to 0° and 90° . Contours are shown at intervals of 1σ starting at the 2σ level.

molecular-replacement solution was obtained in the $P2_1$ crystal form used for the heavy-atom soaks. The Z scores for the rotation and translation functions were 8.7 and 3.5, respectively. The correctness of the MR solution was confirmed by using the MR phases to locate the Hg sites in a difference Fourier synthesis. This step also brought the MR model and the heavy-atom coordinates into register along the polar 2_1 axis. The map calculated with MR phases alone had a correlation coefficient of 0.22 with the final model (Fig. 3*b*). The final *SgrAI* model and the 219-residue *Cfr10I* molecular-replacement model can be aligned over 184 residues, with an r.m.s.d. for C^α positions of 1.63 Å (with the requirement that matching C^α atoms lie within 3.5 Å of one another). Superposition of the final *SgrAI* model with the full 283-residue *Cfr10I* structure does not match any addi-

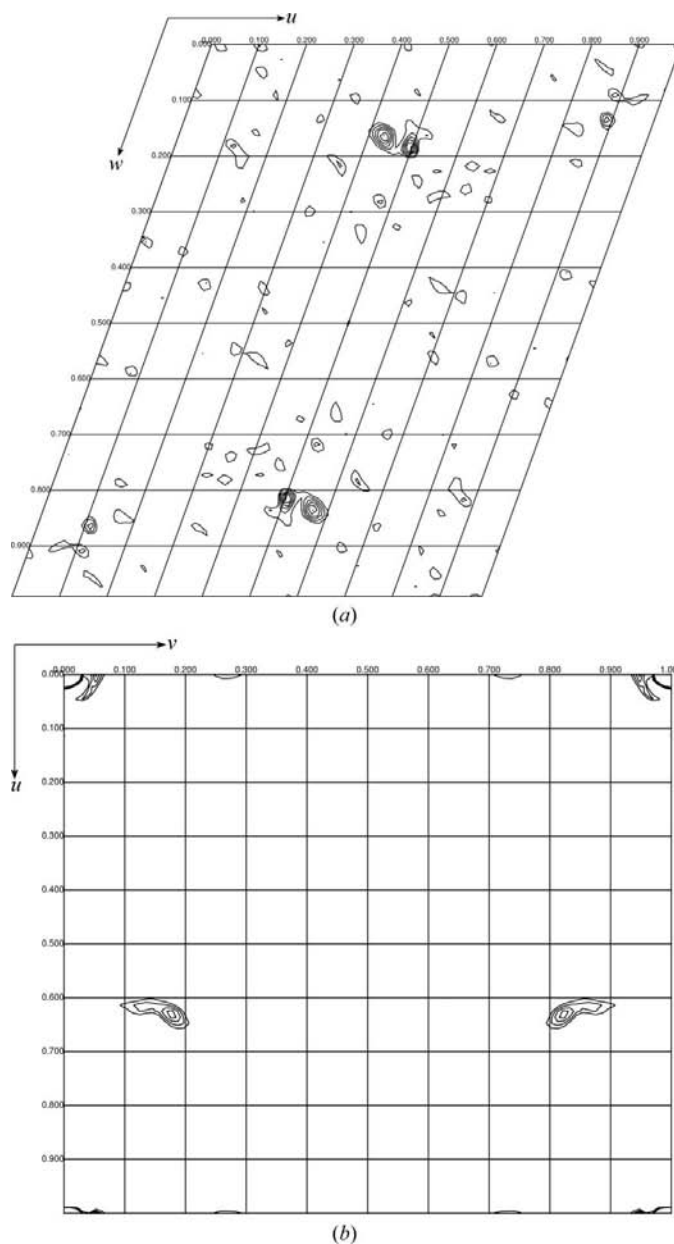


Figure 2
Patterson map sections. (a) The Harker section at $\nu = \frac{1}{2}$. (b) The pseudo-Harker section normal to the noncrystallographic twofold symmetry axis. The difference Patterson map for the $\text{Hg}(\text{SCN})_2$ derivative was calculated with data from 12 to 5 Å resolution. The map was skewed to allow plotting of the pseudo-Harker section. Contours are shown at intervals of 1σ starting at the 2σ level.

Table 2
Heavy-atom phasing statistics.

	Hg(SCN) ₂	K ₂ PtCl ₆
No. of sites	3 (2 modeled)	4
Resolution† (Å)	2.45, 4.0	2.50, 4.0
<i>R</i> _{Cullis} (acentric)	0.91, 0.82	0.96, 0.91
<i>R</i> _{Cullis} (centric)	0.80, 0.70	0.90, 0.82
<i>R</i> _{Cullis} (anomalous)		1.00, 0.98
Phasing power (acentric)	0.77, 1.14	0.43, 0.75
Phasing power (centric)	0.63, 0.90	0.37, 0.51
FOM (acentric)	0.24, 0.38	
FOM (centric)	0.45, 0.61	
FOM (all)	0.25, 0.40	

† Statistics are given twice, first for data to the high-resolution limit stated and then for data to 4.0 Å.

tional secondary-structural elements. Hence, the threading procedure appears to have returned a nearly optimal MR model.

3.3. Phase combination

The MIRAS solution was brought to the same origin as the MR solution and into register along the polar 2₁ axis by applying a fractional shift of (0.5, -0.1, -0.5) to the heavy-atom coordinates. Phases from the MR solution (extending to 3.16 Å resolution) were combined with MIRAS phases (extending to 2.45 Å resolution) using *SIGMAA* (Read, 1986) and the combined phases were improved and extended to 1.9 Å resolution *via* density modification including twofold averaging with *DM* (Cowtan, 1994). The averaging protocol started with data extending to 5 Å resolution and gradually proceeded to the high-resolution limit of the native data over 200 steps. The averaging mask was automatically determined by *DM*. The map calculated with combined density-modified phases had a correlation coefficient of 0.46 with the final model and was used to initiate building (Fig. 3c). *resolve_build* (Terwilliger, 2003) and *Buccaneer* (Cowtan, 2006) gave an initial model consisting of 85 residues. Phases from the initial model were combined with MIRAS phases using *SIGMAA* and the combined phases were improved and extended with *DM*. Further cycles of building and phase combination led to the final model comprising all 338 residues in each *SgrAI* monomer, 17 base pairs, four Ca²⁺ ions and 470 solvent molecules (Dunten *et al.*, 2008). At the end of refinement the *R* and *R*_{free} values were 18.3% and 22.6%, respectively.

3.4. Retrospective analysis of the heavy-atom derivatives

Analysis of the heavy-atom derivatives after the structure had been solved showed that the Hg compound substituted Cys19 in both monomers of the *SgrAI* dimer. Refinement against the data for the Hg derivative showed that the Hg sites were partially occupied and that heavy-atom binding induced local changes in the protein structure. While Hg binding occurred at a single site in one of the monomers, in the other monomer the Hg bound to two sites separated by 3.18 Å. The two adjacent sites effectively merge together at 5 Å resolution, which was the limiting resolution used to calculate the difference Patterson map. The fact that the heavy-atom model included only two Hg sites may have contributed to the poor phasing power of the Hg compound, particularly in the higher resolution shells. Additional factors limiting the usefulness of the Hg derivative were the local changes in the structure induced by heavy-atom binding and the partial occupancy of the binding sites. The Pt compound substituted sites near the S atoms of Met62 and Met203 on the surface of both monomers with partial occupancy. Refinement of the heavy-atom parameters with *MLPHARE* had shown that the occu-

pancy of all the Pt sites was less than that of the Hg sites, which was consistent with the lower phasing power of the Pt derivative. The Pt derivative's data quality was poor in the highest resolution bins and to assess whether truncating the data set at a lower resolution would help, the phasing in *MLPHARE* was repeated with data from the Pt derivative limited to 2.99 Å resolution, where the merging *R* factor is 0.317 and $\langle I/\sigma(I) \rangle$ is 2.2. Although the quality of the Pt-derivative data in the highest resolution bins was not high, limiting the Pt data to 2.99 Å did not significantly change the results of the heavy-atom parameter refinement in *MLPHARE*, nor did it improve the quality

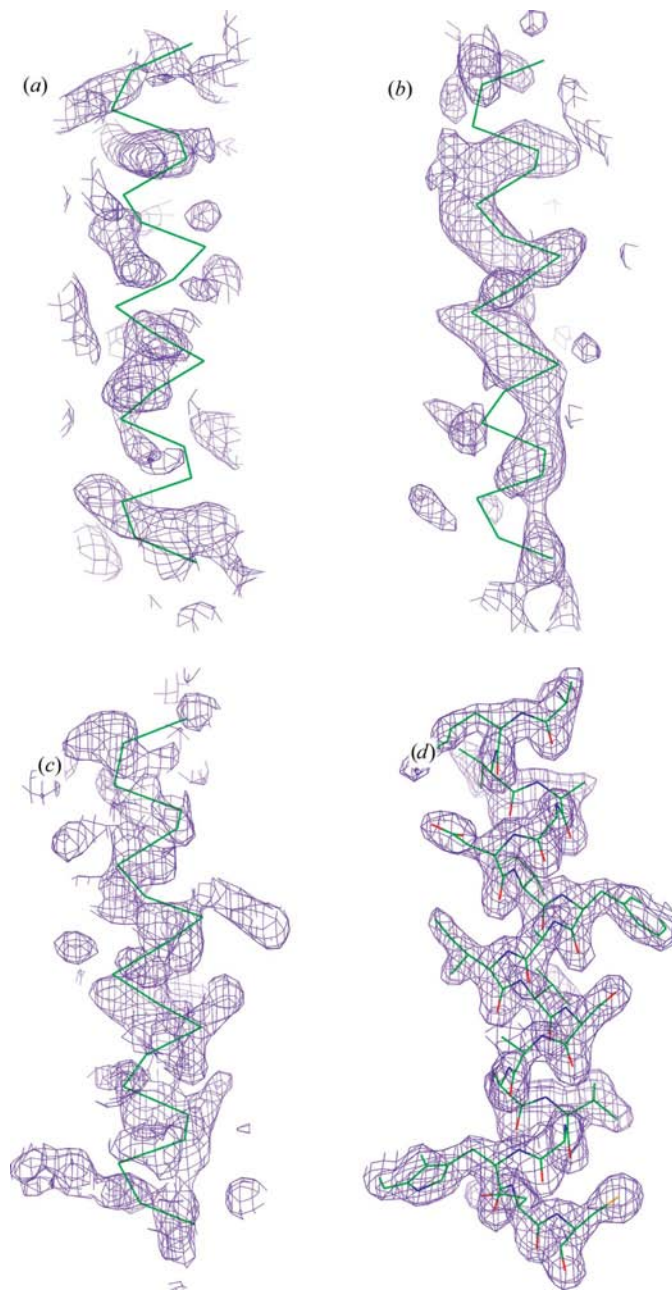


Figure 3
Experimental electron-density maps. The electron density for the longest α -helix in the structure is shown, together with the C α trace in (a)–(c) and the all-atom model in (d). (a) MIRAS phases after averaging with the wrong transformation (2.5 Å), (b) MR phases only (3.1 Å), (c) combined phases after density modification (1.9 Å) and (d) the final $2mF_o - DF_c$ map (1.9 Å).

of the map produced by MIRAS phasing and subsequent density modification.

3.5. Retrospective analysis of the noncrystallographic symmetry

The averaging transformation applied during density modification was initially deduced from the positions of the two Hg atoms in the Hg derivative. Interestingly, the map based on MIRAS phases before averaging has a higher correlation coefficient with the final model than the correlation coefficient of the averaged map with the final model (0.29 *versus* 0.13). To specify the relationship between the two subunits of the *SgrAI* dimer, a set of trial operators was constructed that would superpose the two Hg atoms *via* a 180° rotation. Among this set was one operator consistent with the noncrystallographic symmetry evident in the self-rotation function. Consistency with the self-rotation function was judged by comparing the direction cosines of the twofold axis determined by the self-rotation function (0.923, 0, 0.386) with those of the twofold axis determined using the positions of the Hg atoms (0.936, -0.113, 0.333). Using this operator to perform twofold averaging as part of the density-modification procedure degraded the quality of the MIRAS phases, as shown by the drop in the map correlation coefficient. The crux of the problem lies in the assumption that the two Hg atoms are bound to subunits of one dimer rather than neighboring dimers. As Fig. 4 shows, a twofold axis can be placed between the Hg bound to two subunits of one *SgrAI* dimer or between the Hg bound to subunits of adjacent *SgrAI* dimers. The Hg atoms are separated by 43 Å in the dimer and by 45 Å for the two adjacent dimers. The twofold axes relating the Hg positions are approximately parallel to each other, as can be seen in Fig. 4. To successfully average the electron density in adjacent dimers, a small (3.3 Å) translation parallel to the rotation axis must be taken into account. Given only the Hg-atom positions, it is not possible to determine this translation. Once a molecular-replacement solution with a dimeric model was in hand, the correct choice of the averaging transformation was clear. A new averaging transformation was calculated by (i) generating a dimeric version of the *Cfr10I* molecular-replacement model in the $C222_1$ cell, (ii) placing it in the $P2_1$ cell using the rotation and translation determined by *Phaser* and (iii) determining the rotation and translation needed to superimpose one monomer of the *Cfr10I* dimer onto the other monomer with *LSQMAN*. Repeating the MIRAS phasing and density modification

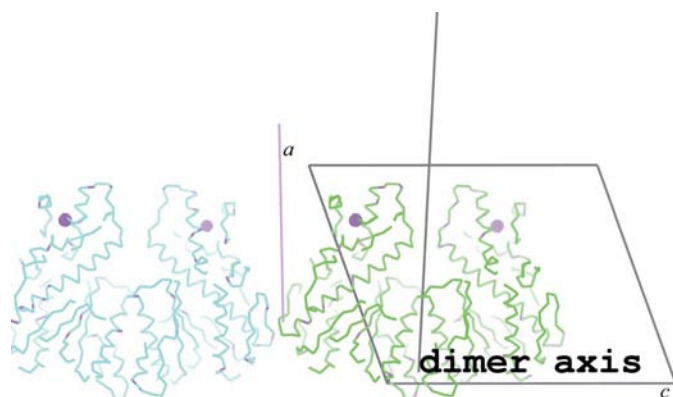


Figure 4

Hg positions in neighboring dimers in the $P2_1$ (form 1) cell. Hg atoms are shown as colored spheres. The green and blue *SgrAI* dimers are related by a translation (0, 0, 1) parallel to the cell edge *c*. Because the noncrystallographic twofold axis is nearly perpendicular to the cell edge *c*, the Hg atoms in neighboring dimers are also approximately related by a twofold rotation axis perpendicular to *c* (shown as a colored line segment).

with the correct averaging transformation gave a map with roughly the same correlation coefficient to the final model's electron density as the combined-phase (MIRAS and MR) map used for building.

4. Summary and conclusions

The quaternary structure of *SgrAI* in all the crystal forms of Table 1 consists of a dimer of the protein bound to duplex DNA. In contrast, the restriction enzymes with closely related sequence specificity (*Cfr10I*, *Bse634I* and *NgoMIV*) are all tetramers. As this work employing *Cfr10I* as a molecular-replacement model shows, the best model was a monomer of *Cfr10I* and not the dimer which binds duplex DNA. The final solution of the *SgrAI* structure depended on a bootstrapping procedure. The initial MR solution was obtained in a crystal form with one monomer per asymmetric unit. That a correct solution had been discovered was not obvious. Although the packing was feasible, the *Z* scores were low and the electron density in the region expected to be occupied by DNA was not recognizable as such. This crystal form diffracted to only 3 Å resolution, had a relatively low solvent content and did not offer the possibility of phase improvement *via* application of noncrystallographic symmetry averaging. We were able to switch to a more favorable crystal form by using the poor electron density corresponding to a (crystallographic) dimer of *SgrAI* as a search model. A correct solution of the MR problem was then relatively easy to recognize by verifying that the MR phases could locate the Hg atoms of the Hg(SCN)₂ derivative in a difference Fourier synthesis. The MR solution immediately provided the correct averaging transformation for averaging electron density within the noncrystallographic dimer, rather than between neighboring dimers. Combination of the MR and MIRAS phases and subsequent phase refinement *via* density modification led to the first map with interpretable features. At that point, the routine bootstrapping approach of building a partial model and combining the partial model phases with experimental phases gave an improved map and after a few cycles a complete model was built.

Portions of this research were carried out at the Stanford Synchrotron Radiation Laboratory, a national user facility operated by Stanford University on behalf of the US Department of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research and by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program and the National Institute of General Medical Sciences. The projects described were partially supported by grant No. 5P41RR001209 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH), and the contents are solely the responsibility of the authors and do not necessarily represent the official view of NCRR or NIH. This work was also supported by National Institutes of Health grant No. 5R01GM066805 (to NCH).

References

- Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger, T. C. (2002). *Acta Cryst.* **D58**, 1948–1954.
- Bozic, D., Grazulis, S., Siksnys, V. & Huber, R. (1996). *J. Mol. Biol.* **255**, 176–186.
- Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.
- Cowan, K. (1994). *Int. CCP4/ESF-EACBM Newsl. Protein Crystallogr.* **31**, 34–38.

- Cowtan, K. (2006). *Acta Cryst.* **D62**, 1002–1011.
- Daniels, L. E., Wood, K. M., Scott, D. J. & Halford, S. E. (2003). *J. Mol. Biol.* **327**, 579–591.
- Deibert, M., Grazulis, S., Sasnauskas, G., Siksny, V. & Huber, R. (2000). *Nature Struct. Biol.* **7**, 792–799.
- Dunten, P. & Hennig, M. (2002). *Acta Cryst.* **A58**, C76.
- Dunten, P. W., Little, E. J., Gregory, M. T., Manohar, V. M., Dalton, M., Hough, D., Bitinaite, J. & Horton, N. C. (2008). *Nucleic Acids Res.* **36**, 5405–5416.
- Fischer, D. & Eisenberg, D. (1996). *Protein Sci.* **5**, 947–955.
- Grazulis, S., Deibert, M., Rimseliene, R., Skirgaila, R., Sasnauskas, G., Lagunavicius, A., Repin, V., Urbanke, C., Huber, R. & Siksny, V. (2002). *Nucleic Acids Res.* **30**, 876–885.
- Jones, T. A. & Kjeldgaard, M. (1997). *Methods Enzymol.* **277**, 173–208.
- Kleywegt, G. J. (1996). *Jnt CCP4/ESF-EACBM Newsl. Protein Crystallogr.* **32**, 32–36.
- Kleywegt, G. J. & Jones, T. A. (1994). *Jnt CCP4/ESF-EACBM Newsl. Protein Crystallogr.* **31**, 9–14.
- Knight, S. D. (2000). *Acta Cryst.* **D56**, 42–47.
- Leslie, A. G. W. (1992). *Jnt CCP4/ESF-EACBM Newsl. Protein Crystallogr.* **26**.
- 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., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* **D53**, 240–255.
- Otwinowski, Z. (1991). *Proceedings of the CCP4 Study Weekend. Isomorphous Replacement and Anomalous Scattering*, edited by W. Wolf, P. R. Evans & A. G. W. Leslie, pp. 80–86. Warrington: Daresbury Laboratory.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Pflugrath, J. W. (1999). *Acta Cryst.* **D55**, 1718–1725.
- Potterton, L., McNicholas, S., Krissinel, E., Gruber, J., Cowtan, K., Emsley, P., Murshudov, G. N., Cohen, S., Perrakis, A. & Noble, M. (2004). *Acta Cryst.* **D60**, 2288–2294.
- Read, R. J. (1986). *Acta Cryst.* **A42**, 140–149.
- Sanders, D. A. R., McMahon, S. A., Leonard, G. L. & Naismith, J. H. (2001). *Acta Cryst.* **D57**, 1415–1420.
- Stein, P. E., Boodhoo, A., Armstrong, G. D., Cockle, S. A., Klein, M. H. & Read, R. J. (1994). *Structure*, **2**, 45–47.
- Terwilliger, T. C. (2003). *Acta Cryst.* **D59**, 38–44.