Crystal structure of π initiator protein–iteron complex of plasmid R6K: Implications for initiation of plasmid DNA replication

Michael K. Swan, Deepak Bastia[†], and Christopher Davies[†]

Department of Biochemistry and Molecular Biology, Medical University of South Carolina, Charleston, SC 29425

Communicated by Donald R. Helinski, University of California at San Diego, La Jolla, CA, October 18, 2006 (received for review December 17, 2005)

We have determined the crystal structure of a monomeric biologically active form of the π initiator protein of plasmid R6K as a complex with a single copy of its cognate DNA-binding site (iteron) at 3.1-Å resolution. The initiator belongs to the family of winged helix type of proteins. The structure reveals that the protein contacts the iteron DNA at two primary recognition helices, namely the C-terminal $\alpha 4'$ and the N-terminal $\alpha 4$ helices, that recognize the 5' half and the 3' half of the 22-bp iteron, respectively. The base-amino acid contacts are all located in $\alpha 4'$, whereas the $\alpha 4$ helix and its vicinity mainly contact the phosphate groups of the iteron. Mutational analyses show that the contacts of both recognition helices with DNA are necessary for iteron binding and replication initiation. Considerations of a large number of site-directed mutations reveal that two distinct regions, namely $\alpha 2$ and $\alpha 5$ and its vicinity, are required for DNA looping and initiator dimerization, respectively. Further analysis of mutant forms of π revealed the possible domain that interacts with the DnaB helicase. Thus, the structure–function analysis presented illuminates aspects of initiation mechanism of R6K and its control.

initiator structure | replication control | replication initiation

In a major class of drug resistance plasmids, interaction of a plasmid-encoded replication initiator protein with the origin-proximal, cognate, repeated sequences called iterons, initiates Cairns-type replication that generates θ -shaped replication intermediates (1–12). Although a wealth of information exists on the genetics and biochemistry of replication initiation and its control in a few model plasmid systems, there has been a paucity of information on the atomic structure of the plasmid-encoded replication initiator proteins and their detailed structure–function analysis. In fact, the structure of only one plasmid-encoded initiator, namely RepE of F factor, is known to date (13). Unfortunately, the F plasmid system offers very few mutants of the RepE initiator that have been isolated and analyzed to date and very little published information exists on the interaction of the RepE initiator with host-encoded replication proteins.

By contrast, the R6K plasmid system has a rich repertoire of biological and biochemical information on its initiator protein called π (2, 3, 6, 8, 14–19), including its interaction with several host-encoded replication proteins (20–22), and its self interaction, via dimerization (23, 24), dimer–dimer (or dimer–monomer) interactions, that are essential for DNA looping (22). The R6K replicons consist of three replication origins called α , β , and γ . π is a transacting initiator protein that binds to seven tandem iterons at γ (Fig. 1). Dimerization of π negatively controls copy number, as revealed by the fact that mutations that enhance copy number also tend to monomerize the protein upon iteron binding (22, 23, 25). In addition to the dimerization interphase, a second type of π – π contact is required for dimer–dimer or, more likely, monomer–dimer interactions that are essential for activation of replication at a distance at α and β by π -mediated DNA looping (Fig. 1*a* and *b*; refs. 22, 24, and 26). Several mutants that abolish or reduce interaction of π with host-encoded replication proteins are available and some of the mutants forms of the protein have been

Fig. 1. R6K replicons. (a) Diagram of the replication region of plasmid R6K. Note that ori α and β are located $\approx 4,000$ and $1,200$ bp away from ori γ , respectively. π -mediated DNA looping that brings together the single iteron at α and the half iteron at β with the seven iterons at γ , activates ori β and α (see *b*). (b) DNA looping can occur in at least two modes: (i) by iteron-bound monomer pairs bridged by dimeric π or (ii) by two interacting iteron-bound dimers. Dimeric π is remodeled by the DnaK chaperone system into active monomers upon binding to iteron DNA (28). (c) A diagram of the duplex DNA used in the structure determination of π in complex with iteron DNA. The numbering system shown corresponds to that of the model with oligonucleotide A comprising residues 1–23 and oligonucleotide B comprising residues 24–46. The thymines in red were those replaced by bromodeoxyuridine in the oligonucleotides used to solve the structure. Note the overhanging 3' T residues that promoted optimal crystallization.

purified and characterized biochemically (20–22). Recently, we reconstituted the replication of R6K *in vitro* with all purified proteins (27, 28). In plasmid R6K, unlike in the F plasmid, the major replication origins α and β are located ≈ 4 and ≈ 1.2 kb away, respectively, from the primary binding site of the π initiator at the seven iterons of ori γ (Fig. 1*a*). DNA looping mediated by π between γ and α or between γ and β initiates replication from the α and β origins, respectively (24, 29). The dimeric WT π is biologically inert and has to be activated by the DnaK chaperone system into a form that monomerizes upon contact with DNA (Fig. 1*b*; ref. 28).

The compelling reason to solve the crystal structure of the π initiator protein as a complex with a single iteron DNA was to be able to perform structure–function analysis at atomic reso-

Author contributions: D.B. and C.D. designed research; M.K.S. and D.B. performed research; M.K.S., D.B. contributed new reagents/analytic tools; D.B., and C.D. analyzed data; and D.B. and C.D. wrote the paper.

The authors declare no conflict of interest.

Abbreviations: NTD, N-terminal domain; CTD, C-terminal domain.

Data deposition: The coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 2NRA).

[†]To whom correspondence may be addressed. E-mail: bastia@musc.edu or davies@musc.edu.

© 2006 by The National Academy of Sciences of the USA

Swan et al.

NTD, but, most notably, more are with bases. As with the NTD, there are several hydrogen bonds to the phosphate backbone from tyrosines (Tyr-170 and -217) and a serine (Ser-168), as well as some hydrophobic interactions with the ribose moieties involving Phe-220, Pro-221, Ile-222, and Ile-256. There are also two salt bridges, one between Lys-216 and the phosphate of T36 and another between Lys-216 and the phosphate of T5. Mutation of Lys-216 abolished DNA binding and initiation from ori γ , showing that this contact is essential for DNA binding and initiation (Table 4, which is published as supporting information on the PNAS web site). Likewise, residues 250–252 are very close to DNA, especially Glu-252, which is near the phosphate of A3 (and T46), thus explaining the lack of DNA-binding activity of the triple mutant, E250A/K251A/E252A (Table 4).

The most important interactions in this region, however, appear to be those mediated by Arg-225 and Asp-226. Like Ser-71 and Arg-75 in the NTD, these residues penetrate deep within the major groove where they form hydrogen bonds with DNA (Fig. 8*b*, which is published as supporting information on the PNAS web site). Arg-225 forms three contacts with the bases of C37, T38, and G7 (on the opposite strand of DNA). Asp-226 forms a hydrogen bond with both the base of C37 and a salt bridge with Arg-225. When either of these residues is mutated to alanine, π no longer binds DNA in gel-shift assays and does not support replication from ori γ but still binds DnaB (Figs. 9 and 10). Probably critical for this interaction is the unusual conformation of $\alpha 4'$, which follows a wider coil at residue 226 and projects the aspartate closer toward the DNA than would occur in an α -helix. A similar distortion of this helix was also observed in the crystal structure of RepE in complex with DNA (13).

Interestingly, Arg-225 and Asp-226 lie very close to the site in the iteron DNA where the strands are slightly separated, comprising nucleotides 34–36 on the oligonucleotide closest to the protein (Fig. 5). Thus, the contacts made by Asp-226 with C37 and by Arg-225 with the T38/A8 base pair, which are immediately adjacent to this site, may have a role in DNA melting. In addition, those residues contacting the nucleotide directly may also have a role in their separation, including Ser-168, which contacts C35, and Tyr-217, which contacts T36.

Mutant Forms of π . A major advantage of the R6K plasmid system over that of the F plasmid is the large body of biochemical and mutational data that has been accumulated in the former. We have added to this by testing several alanine mutants of π in gel-shift assays to measure DNA binding and in ELISA assays to measure binding to DnaB helicase. The mutations that alter the activity of π are summarized in Table 4 and are mapped onto the structure in Fig. 6. These mutations can now be interpreted in the context of the crystal structure to provide unique insight into the molecular mechanism of initiation mediated by this initiator.

DNA Binding. The mutations that abolish binding with DNA can be grouped into two categories: those within or immediately adjacent to the DNA-binding site (described above) and those located further away, which probably alter the DNA-binding site through a conformational change or by causing misfolding of the protein. The latter group includes K11A/K12A, K14A, E20A, E114A/E115A, and D208A/K209A. The most interesting of these mutations is E114A/E115A. These residues lie immediately after the disordered region between $\alpha 5$ and $\beta 2b$ that contains the high-copy-number mutations but have an opposite activity compared with those mutations, most likely because they cause misfolding in this region.

High-Copy-Number Mutations. The mutations that lead to high copy number are all clustered in a disordered region at the C terminus of $\alpha 5$. Leu-106 is modeled only approximately, and Ser-113 and Phe-107 are absent altogether. These mutations are

Fig. 6. The mapping of various mutants that alter the activity of π . The structure of π is shown as a gray ribbon on which amino acids at sites of mutations are plotted in ball-and-stick form. Important elements of secondary structure are labeled. The coloring scheme for the mutants is as follows: blue, residues that when mutated are replication defective (this study); red, sites of high-copy-number mutants (P106L, F107S, and P113S; ref. 22); orange, the mutation (P42L) that causes a defect in DNA looping (24); green, sites of double mutations that are replication defective and show lowered binding to DnaB helicase but still bind DNA (this study); and purple, mutants that are defective for replication but still bind DNA (23).

present in the protein construct used to determine the crystal structure, because they increase the solubility of π monomers (22), the DNA-binding form of π . Although the effect of these mutations cannot be deduced without a structure of the wild-type protein, it is unlikely that they significantly affect the conformation of $\alpha 5$, because π^* shows normal binding to DnaB (mutations in $\alpha 5$ lower binding to DnaB; see below). If anything, because they promoted the crystallization of π , they may have increased the ordering of this region, and that in turn may explain the slightly higher binding of π^* to DnaB (data not shown). If true, this suggests that $\alpha 5$ and its connecting loops exhibit greater disorder in the WT monomeric form of π .

More importantly, the mapping of high-copy mutants that increase the solubility of the monomeric form of π to a loop adjacent to $\alpha 5$ suggests this region is involved in dimerization. Indeed, in the crystal packing, $\alpha 5$ interacts with its symmetry counterpart, which tentatively suggests a coiled-coil mechanism of dimerization. Dimers of π are repressors of replication (27, 35), and dimer-dimer interactions involving two sets of iterons of a pair of plasmids in trans are known to shut down initiation by handcuffing. Alternatively, interaction of a single array of dimers binding to two sets of iterons and causing coupling or handcuffing of a pair of origins is believed to be a major mechanism of negative control of replication (19, 36).

DNA Looping. Mutation of Pro-42 to leucine abolishes the DNA-looping activity of π (24). This residue is located at the C terminus of $\alpha 2$ and lies within a shallow depression on the surface of the molecule. Interestingly, this mutation may not have altered the structure of the main chain of this region significantly, because the leucine terminates helix $\alpha 2$ in the same way as the proline would be expected to do (and therefore would not affect the putative DnaB-binding region at residues 46–47). However, the presence of the bulky hydrophobic side chain has altered the surface of this region, suggesting this is a second π - π interaction domain.

Unlike the initiator proteins of other plasmids, the dimeric form of π does bind to DNA, albeit less strongly than monomers (7, 37). Therefore, in principle, a DNA loop could be formed by

the interaction between two iteron-bound dimers (Fig. 1*b*). Alternatively, a dimeric π could bridge pairs of iteron-bound monomers (35, 38). The looping models suggest that, besides the dimerization interphase, a second type of protein–protein contact involving a π dimer and a monomer or two dimers would be needed for DNA looping. The P42L looping defective mutation located at the $\alpha 2$ helix probably reveals the second type of contact. This mutation does not cause π dimers to monomerize extensively upon contact with DNA but completely abolishes DNA looping *in vivo* and *in vitro* (24).

Interaction with DnaB. ELISA data showing the physical interactions of the WT and various mutant forms of π with DnaB helicase are shown in Fig. 9. The quadruple mutant of π used to determine the crystal structure (π^*) binds DnaB at a similar level as WT π . The majority of the mutations that lower binding of π to DnaB map to the top of the molecule as viewed in Fig. 6 (colored green), suggesting a DnaB-binding site comprising $\alpha 5$ and the nearby loops between $\alpha 2$ – $\beta 2$ and $\alpha 6$ – $\beta 1'$. Mutations at these sites do not interfere with DNA binding, so the defect in DNA replication of these mutants likely arises by lowered binding to DnaB in the *in vivo* assay used. Because most of these residues are surface-exposed in the structure and apparently available for binding, it seems unlikely their mutation will affect DnaB binding by a conformational effect. Curiously, these sites map very closely to the same region containing the high-copy-number mutations that promotes formation of monomers and yet, in these assays, DNA is not present, and so DnaB must be binding to a dimer of π . Hence, these data argue against $\alpha 5$ itself being directly involved in dimerization, unless the π molecules are packed in a face-to-face fashion, but do not preclude the involvement of the disordered region immediately after $\alpha 5$, which is spatially separate. If so, then the antiparallel coiled-coil arrangement of $\alpha 5$ with its symmetry counterpart in the crystal packing does not represent a biological interaction.

Other Replication-Defective Mutants. Several mutant forms of π are inactive in DNA replication but still bind to DNA (Table 4). One of these, a double-mutant M36A/M38A, yields a stronger dimer of π that binds DNA but does not support replication (23). Met-36 lies in the middle of $\alpha 2$ and projects into the major hydrophobic core of the NTD between $\alpha 1$, $\alpha 2$, $\alpha 6$, $\beta 4$, and $\beta 6$. Met-38 also emanates from $\alpha 2$ and, although it points in an opposite direction, also participates in hydrophobic interactions with $\alpha 2$, $\alpha 3$, $\alpha 7$, and $\beta 2$. Mutation of these residues to alanine would likely weaken these hydrophobic cores and so disrupt the architecture of the protein, but, beyond this, any effect on the structure is hard to predict.

We have constructed the double-mutant K78A/E79A, which, in an *in vivo* assay of the ability of the mutant to support replication (3), shows the same phenotype as M36A/M38A. These residues lie on the DNA-binding helix $\alpha 4$, where they are exposed to solvent, and neither contact DNA directly. Our interpretation of this mutant is that, although contacts of $\alpha 4$ with DNA in the mutants are disrupted, causing loss of replication, there is still net DNA binding because the contacts between $\alpha 4'$ and DNA are maintained. In other words, in dimerization-proficient π , $\alpha 4'$ contacts with DNA should be providing most of the free energy change of DNA–protein interaction as contrasted with that of $\alpha 4$.

In summary, the structure of the monomeric π protein as a complex with a single iteron DNA has provided valuable structural insights into the stages of initiation that precede loading of DnaB, involving DNA bending and melting, as well as revealing potential sites of interactions of π with DnaB and how initiation control is affected by the site-directed mutants. However, because such mutations may affect, for instance, protein–protein interactions by conformational changes rather than being part of

the binding site, additional biochemical and structural work involving cocrystal structures of the ternary complex of π –iteron DNA peptides of DnaB, DnaA, and DnaG might be needed for more definitive answers.

Finally, further understanding of dimerization and DNA looping as processes that are critical for replication control would require solving the structure of the dimerization domain of WT π . To date, the structure of only one dimerization domain of an initiator protein, namely RepA of pPS10, has been reported (39). Although the complete structure of this initiator has not been solved to date, a comparison with the equivalent region in RepE led to the suggestion that the monomer-to-dimer transition might involve the transformation of a loop and β -strand (as seen in monomeric RepE) to an α -helix (as seen in dimeric RepA; ref. 39). By contrast, the disordered loop C-terminal to $\alpha 5$, which is topologically unrelated to the proposed domain in RepE and RepA, is the best candidate for a dimerization domain in π , because it is flexible and contains mutations that promote the formation of monomers and lead to high copy number. We hope to test this hypothesis by determining the crystal structure of π in its dimeric form.

Materials and Methods

Cloning, Expression, and Purification of $\pi^*\Delta 29$. This form of the π initiator protein with the deletion of the last C-terminal 29 amino acids and containing four mutations (P42L, P106L, F107S, and P113S) results in the formation of monomers of π after iteron binding, remains soluble after purification at 50 mg/ml, increases R6K plasmid copy number, and retains full biological activity, as measured by its ability to catalyze DNA replication *in vivo* and *in vitro* (22) and its interaction with the host-encoded proteins DnaA, DnaB, and DnaG (22). The ORF of $\pi^*\Delta 29$ was cloned into the pTXB1-intein vector (NEB, Boston, MA), and the protein was expressed and purified as described (22). Gel mobility-shift and ELISA assays were carried out by using mutants of π as described (20, 22).

Crystallization. Complementary oligonucleotides were designed by using the 22-bp consensus sequence of the seven iterons (14). The duplex used is shown in Fig. 1*c* and comprises the sequence of iteron 7 with the terminal A–T base pairs replaced by G–C to increase the stability of the duplex, with a single thymine overhang at each 3' end. $\pi^*\Delta 29$ was added to duplex DNA in a 1:1 ratio, incubated at room temperature for 20 min, and concentrated to 0.65 mM. Crystals were obtained over wells containing 5% polyethylene glycol 8000/50 mM $MgCl_2$ /100 mM $NH_4H_2PO_4$ and were cryoprotected by adding glycerol to a maximum of 38%. These crystals belonged to the space group P6₁22 with cell dimensions $a = 126.0$ and $c = 138.3$ Å and diffracted to 3.0 Å resolution. There is one π –DNA complex per asymmetric unit, corresponding to a solvent content of 70%.

Structure Determination and Refinement. The structure was solved by multiwavelength anomalous dispersion by using oligonucleotides in which the five thymines at positions 13, 21, 28, 41, and 43 were replaced with 5'-bromo-deoxyuridine (see Fig. 1*c*). Determination of the bromine positions and initial phasing calculations were performed by using SOLVE/RESOLVE (40, 41), but improved phases were later obtained by using SHARP (42) and SOLOMON (43). Model building using XtalView (44) commenced with the initial protein model produced by RESOLVE and B form DNA generated using X3DNA (45). The initial refinement was performed by using CNS (46) but thereafter with REFMAC5 (47). The resolution of the structure did not permit the unambiguous inclusion of any water molecules or magnesium ions. Coordinates and structure factors have been deposited with the Protein Data Bank. Additional experimental

details are described in the *Supporting Text*, which is published as supporting information on the PNAS web site.

We thank S. Zzaman, R. Sharma, J. Reddy, M. Saxena, and Miss Lottje Engle for isolation and characterization of mutant forms of π , *in vitro* replication reconstitutions and gel mobility-shift assays. We thank Dr. Dhruva Chatteraj for useful discussions. We thank the editor of this paper, Dr. Donald Helinski, for continuing interest, encouragement, and many useful discussions. We thank Alexei Soares for assistance

with data collected at NSLS beamline X12B. Data were collected at Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at the Advanced Photon Source, Argonne National Laboratory. Supporting institutions may be found at www.ser-cat.org/members.html. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract W-31-109-Eng-38. This work was supported by grants from National Institute of Allergy and Infectious Diseases and National Institute of General Medical Sciences (to D.B.)

1. Lovett M, Sparks RA, Helinski DR (1975) *Proc Natl Acad Sci USA* 72:2905–2909.
2. Crosa JH, Luttrop L, Falkow S (1978) *J Mol Biol* 124:443–468.
3. Kolter R, Inuzuka M, Helinski DR (1978) *Cell* 15:1199–1258.
4. Crosa JH (1980) *J Biol Chem* 255:11075–11077.
5. Inuzuka N, Inuzuka M, Helinski DR (1980) *J Biol Chem* 255:11071–11074.
6. Germino J, Bastia D (1982) *Proc Natl Acad Sci USA* 79:5475–5479.
7. Germino J, Bastia D (1983) *Cell* 34:125–134.
8. Filutowicz M, McEachern M, Greener A, Mukhopadhyay P, Uhlenhopp E, Durland R, Helinski D (1985) in *Plasmids in Bacteria*, eds Helinski DR, Cohen SN, Clewell DB, Jackson DA, Hollaender A (Plenum, New York), pp 125–140.
9. Cereghino JL, Helinski DR, Toukdarian AE (1994) *Plasmid* 31:89–99.
10. Austin SJ, Mural RJ, Chatteraj DK, Abeles AL (1985) *J Mol Biol* 183:195–202.
11. DasGupta S, Mukhopadhyay G, Papp PP, Lewis MS, Chatteraj DK (1993) *J Mol Biol* 232:23–34.
12. Vocke C, Bastia D (1983) *Cell* 35:495–502.
13. Komori H, Matsunaga F, Higuchi Y, Ishiai M, Wada C, Miki K (1999) *EMBO J* 18:4597–4603.
14. Germino J, Bastia D (1983) *Cell* 34:125–134.
15. Stalker DM, Kolter R, Helinski DR (1982) *J Mol Biol* 161:33–43.
16. Germino J, Bastia D (1983) *Cell* 32:131–140.
17. Germino J, Bastia D (1984) *Proc Natl Acad Sci USA* 81:4692–4696.
18. Greener A, Filutowicz MS, McEachern MJ, Helinski DR (1990) *Mol Gen Genet* 224:24–32.
19. McEachern M, Bott MA, Tooker PA, Helinski DR (1989) *Proc Natl Acad Sci USA* 86:7942–7946.
20. Ratnakar PV AL, Mohanty BK, Lobert M, Bastia D (1996) *Proc Natl Acad Sci USA* 93:5522–5526.
21. Lu YB, Datta HJ, Bastia D (1998) *EMBO J* 17:5192–5200.
22. Abhyankar MM, Reddy JM, Sharma R, Bullesbach E, Bastia D (2004) *J Biol Chem* 279:6711–6719.
23. Kruger R, Konieczny I, Filutowicz M (2001) *J Mol Biol* 306:945–955.
24. Miron A, Mukherjee S, Bastia D (1992) *EMBO J* 11:1205–1216.
25. Miron A, Patel I, Bastia D (1994) *Proc Natl Acad Sci USA* 91:6438–6442.
26. Mukherjee S, Erickson H, Bastia D (1988) *Cell* 52:375–383.
27. Abhyankar MM, Zzaman S, Bastia D (2003) *J Biol Chem* 278:45476–45484.
28. Zzaman S, Reddy JM, Bastia D (2004) *J Biol Chem* 279:50886–50894.
29. Mukherjee S, Erickson H, Bastia D (1988) *Proc Natl Acad Sci USA* 85:6287–6291.
30. Sharma S, Sathyanarayana BK, Bird JG, Hoskins JR, Lee B, Wickner S (2004) *J Biol Chem* 279:6027–6034.
31. Mukherjee S, Patel I, Bastia D (1985) *Cell* 43:189–197.
32. Sharma R, Kachroo A, Bastia D (2001) *EMBO J* 4577–4587.
33. Konieczny I, Doran KS, Helinski DR, Blasina A (1997) *J Biol Chem* 272:20173–20178.
34. Datta HJ, Khatri GS, Bastia D (1999) *Proc Natl Acad Sci USA* 96:73–78.
35. Zzaman S, Bastia D (2005) *Mol Cell* 20:833–843.
36. Pal SK, Chatteraj DK (1988) *J Bacteriol* 170:3554–3560.
37. Germino J, Bastia D (1983) *Cell* 32:131–140.
38. Toukdarian AE, Helinski DR (1998) *Gene* 223:205–211.
39. Giraldo R, Fernandez-Tornero C, Evans PR, Diaz-Orejas R, Romero A (2003) *Nat Struct Biol* 10:565–571.
40. Terwilliger T, Berendzen J (1999) *Acta Crystallogr D* 55:849–861.
41. Terwilliger TC (2002) *Acta Crystallogr D* 58:1937–1940.
42. de la Fortelle E, Bricogne G (1997) *Methods Enzymol* 276:472–494.
43. Abrahams JP, Leslie AGW (1996) *Acta Crystallogr D* 52:30–42.
44. McCree DE (1992) *J Mol Graphics* 10:44–46.
45. Lu, X.-J. & Olson WK (2003) *Nucleic Acids Res* 31:5108–5121.
46. Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, et al. (1998) *Acta Crystallogr D* 54:905–921.
47. Murshudov GN, Vagin AA, Dodson EJ (1997) *Acta Crystallogr D* 53:240–255.