Supporting Online Material

Crystal structure of the long-chain fatty acid transporter FadL

Bert van den Berg[1], Paul N. Black[2], William M. Clemons Jr., and Tom A. Rapoport

Materials and Methods

Fatty acid transport and binding

Cells were grown to mid-log phase (5 X 10^8 cells/ml) in minimal medium containing 5 mM potassium acetate and ampicillin as required, after which arabinose (0.05%) was added for 1 hr. The cells were harvested by centrifugation, resuspended in 0.5 volume medium E containing 200 μg/ml chloramphenicol, and starved for any carbon and energy source for 30 min at 30•C. LCFA transport was carried out as described by Kumar and Black (*25*). Fatty acid transport experiments were done at least six times, in duplicate. LCFA binding experiments were carried out essentially as described by Black (*23*). One volume of starved cells was added to a reaction cocktail resulting in oleate:bovine serum albumin (BSA) ratios of 2.0, 1.0, and 0.5. The BSA concentrations were held constant at

173μM and the oleate (specific activity = 23 Ci/mmol) concentrations varied accordingly. All fatty acid binding experiments were done at least four times, in duplicate. For strain references see (*30* (LS1548), *S1* (LS6164), *14* (LS6949)).

Protein preparation

The gene for *Escherichia coli* FadL including the signal sequence was amplified from genomic DNA by PCR, digested with EcoRI/XbaI, and ligated with the EcoRI/XbaI digested pBAD22 vector (*S2*), which is under the control of the arabinose promotor. A hexa-histidine tag was attached at the C-terminus of FadL in order to facilitate purification. The protein was expressed in *E.coli* C43 (DE3) cells at 37 °C by induction with 0.2% arabinose for 3-4 hrs. The following steps were all done on ice; the procedure describes how the monoclinic crystal form was obtained. After harvesting by centrifugation, the cells were lysed in 20mM Tris-HCl/300 mM NaCl/10% glycerol pH 7.8 (TSG buffer) using a microfluidizer (Microfluidics). Total membranes were collected by centrifugation at 40,000 rpm for 40 min and solubilized by homogenization in 1% LDAO/1% OG (w/v) in TSG buffer followed by stirring for 45 min. After centrifugation for 30 min at 40,000 rpm the supernatant was loaded onto a ~15 ml Nickel column (metal-chelating Sepharose, Pharmacia Biotech). The column was washed with 10 volumes of TSB with 0.2% LDAO in the presence of 10 mM imidazole. His-tagged FadL was eluted with 250 mM imidazole, concentrated to ~10 ml, and loaded onto a Superdex-200 26/60 gel filtration column equilibrated with 10 mM sodium acetate/50 mM NaCl/10% glycerol/0.05% LDAO pH 5.6. The gel filtration step was followed by ion exchange on a Mono-S 10/10 column equilibrated in 10 mM sodium acetate/50 mM

NaCl/10% glycerol/0.45% C_8E_4 pH 5.6. FadL eluted in 2 broad peaks, at ~ 50 mM NaCl (flowthrough) and at ~150 mM NaCl. Only the protein in the flowthrough fraction gave well-diffracting monoclinic crystals. If the LDAO was exchanged thoroughly by washing the column extensively with C_8E_4 buffer, no well-diffracting crystals were obtained. The protein that yielded the hexagonal crystal form was purified in an identical way, with the final ion exchange step performed at pH 4.9, where the protein eluted as a broad peak between 0.1 and 0.4 M NaCl. Seleno-methionine substituted protein was expressed in wild-type C43 cells by inhibition of the methionine biosynthesis pathway (*S3*). Cells were grown in minimal M9 medium supplemented with 0.2% glucose and 5% (v/v) glycerol as carbon sources, and induced with 1% arabinose for 8 hrs. Purification was as described for the native protein, with 0.5 mM Tris-(2-carboxyethyl)phosphine added to buffers to avoid oxidation.

Although the aromatic character of the C-terminal residue of OM proteins is important for stability and integration into the OM (*S4*), the presence of the C-terminal hexameric histidine tag had no effect on the localization or function of FadL. The purified protein was properly folded as indicated by its lower mobility in SDS-gels following heat treatment (data not shown). To test whether the histidine tag affected the function of FadL, the protein was expressed from a plasmid in an *E. coli* strain lacking endogenous FadL, and binding of oleic acid to the cells was determined in strains lacking the downstream enzyme FadD (Table S1). While binding in the absence of FadL was very low (strain LS6949), the histidine-tagged FadL protein allowed binding of oleic acid at the same level as the non-tagged protein. The level of binding was slightly higher than in a control strain containing endogenous FadL (LS1908). Transport assays performed in

strains containing FadD also showed that the level of oleic acid uptake was similar to that in a control strain (LS1548) and much higher than in a strain lacking FadL (LS6164). Thus, the histidine-tagged protein appears to be fully functional. In addition, like the wild-type protein, it functioned as a receptor for bacteriophage T2 (data not shown).

Crystallization

A number of initial crystallization conditions for FadL were found with hanging drop vapor diffusion using a broad screen at 4 °C and 22 °C with protein purified in LDAO, β-OG and C_8E_4. After optimization, the best monoclinic crystals were obtained at 22 °C in 27-32% PEG 4K/50 mM cacodylate/2-5 mM $CuSO_4$ pH 5.3, by mixing 2 µl of an 8 mg/ml protein solution with 1 µl well solution at 22 °C. Bar-shaped crystals appeared overnight and grew to maximum dimensions of 100 x 250 x 700 µM in 4-5 days. They belong to spacegroup $P2_1$, diffract to 2.6 Å, and have 2 molecules in the asymmetric unit (Vm = 3.9 $Å^3$/Da, corresponding to a solvent content of ~59%). Selenomethionine-substituted crystals were obtained under similar conditions (with PEG5K MME instead of PEG4K), but did not diffract beyond 6 Å. The hexagonal crystals were obtained from Crystal Screen I (Hampton research), condition 44 (0.2 M magnesium formate pH 6.8). They appeared after 3-4 days and grew to 300 x 300 x 100 µM in a week. They diffract to 2.8 Å and have 2 molecules in the asymmetric unit (Vm = 3.8 $Å^3$/Da, solvent content ~68%). For heavy atom derivatization, small volumes of heavy atoms solutions in mother liquor were added to the crystallization drop to final concentrations between 0.5 and 5 mM. Soak times varied between 2 and 18 hrs. For cryoprotection of the monoclinic crystals, the glycerol concentration in the drop was gradually increased to ~17% by direct

addition to the drop of mother liquor with 20% glycerol and 0.45% C_8E_4. The hexagonal crystals were cryoprotected by direct transfer into mother liquor with 0.3% C_8E_4 and 50% glycerol. Crystals were flash-frozen in liquid nitrogen.

Attempts to soak LCFAs into the FadL crystals were not successful, as was co-crystallization of FadL in the presence of fatty acids. There may be several reasons for this; first of all, due to the presence of both E319 and R157 as ligands for the zwitterionic LDAO molecule, its affinity for the P1 binding site could be similar or even higher than that of an LCFA, despite the relatively short length of the LDAO alkyl chain. In addition, since the LCFAs partition into the detergent micelles and change their properties, it was not possible to use fatty acid concentrations higher than 0.5-1 mM in the soak and co-crystallization experiments, which might have been too low to displace the LDAO molecule. It also appears that it may be difficult to obtain good crystals without a bound ligand (LDAO or fatty acids) in the high affinity site P1; only poor-quality crystals could be obtained from an LDAO-free preparation (i.e. by extensively exchanging LDAO for C_8E_4).

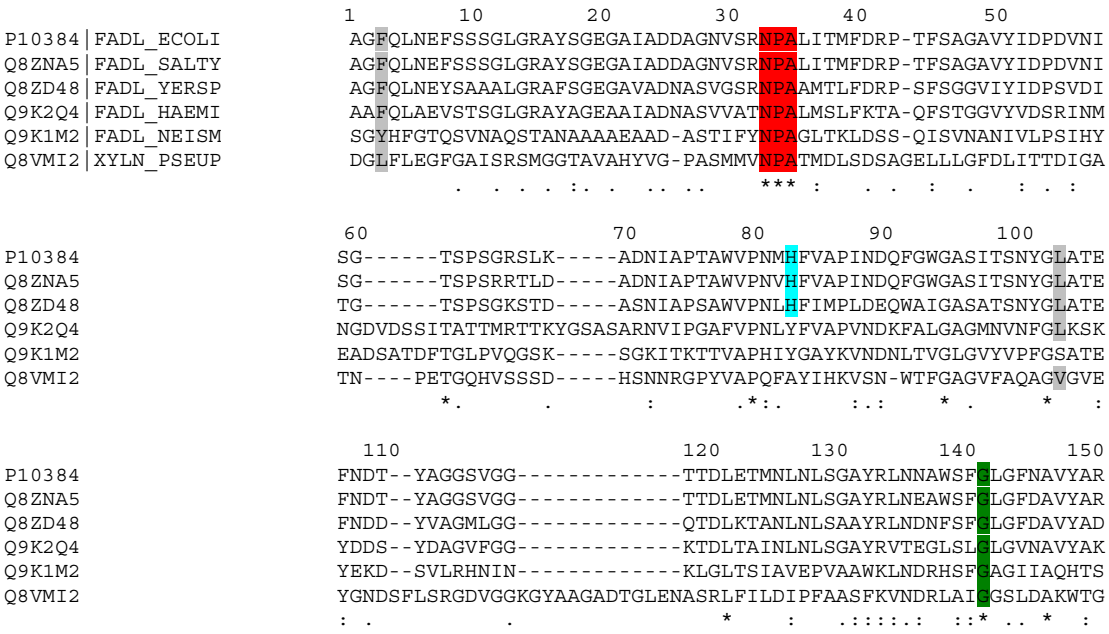
Data collection, structure determination and refinement

Diffraction data were collected at 100K on beamlines at either the National Synchrotron Light Source (NSLS) at Brookhaven National Labs (X25) or at the Advanced Photon Source at Argonne National Labs (8-BM) (Table S2). Data were indexed and scaled with HKL2000 (*S5*). For the monoclinic crystals, MIR and SAD datasets were recorded from $YbCl_3$ and K_2PtCl_4 soaked crystals at the Yb and Pt peak wavelengths. Initial heavy atom positions were obtained with SOLVE (*S6*) and refined using SHARP (*S7*). After density

modification (using DM) with 2-fold NCS averaging, phase extension and solvent flattening, electron density maps were obtained that allowed an initial model to be built with O (*S8*). Iterative cycles of torsion angle refinement and B-factor refinement in CNS (*S9*) followed by model building in O resulted in a final model with an R_{free} of 30.2%, including all residues (1-421) of the protein and the C-terminal hexa-histidine tag. During the later stages of refinement the NCS restraints were relaxed from an initial value of 300 to 10. Phases for the hexagonal crystal form were obtained using MIR datasets recorded for $OsCl_3$, uranyl-acetate and K_2PtCl_4 soaked crystals. The refined monoclinic FadL model was placed manually into the experimental electron density maps of the hexagonal crystal form, followed by rigid-body refinement, torsion angle refinement with NCS restraints and B-factor refinement in CNS, followed by model building in O. The final model includes all residues of the protein (R_{free} 33.1%; in this case no density is observed for the histidine tag. For refinement statistics see Table S2.

Supplementary Figures

Fig. S1 ClustalW sequence alignment of FadL homologues. Sequences from the following organisms were selected for alignment (with their SWISS-PROT accession numbers): P10384; FadL, *Escherichia coli*, Q8ZNA5; FadL, *Salmonella typhimurium*, Q8ZD48; FadL, *Yersinia pestis*, Q9K2Q4; FadL, *Haemophilis influenzae*, Q9K1M2; FadL, *Neisseria menigitidis*, Q8VMI2; toluene transporter XylN from *Pseudomonas putida*. Selected amino acid residues mentioned in the text are shown in colors: the conserved NPA sequence of the hatch in red, glycine residues in green, charged residues in blue, and hydrophobic residues within 4 Å of the high-affinity binding site (Fig. 2B) in gray. The symbols underneath the alignment indicate the degree of conservation: identical residues are shown as "*", highly similar residues are shown as ":", and similar residues are shown as ".".



```
                                 160        170       180       190       200        210
P10384                  AKIERFAGDLGQ-LVAGQIMQSPAGQTQQGQALAATANGIDSNTKIAHLNG-NQWGFGWN
Q8ZNA5                  AKIERFAGDLGQ-LVA-----------AQNPALAPVAGQIPSDTKIAHLNG-NQWGFGWN
Q8ZD48                  AKIVRHLGEAG----------------------G---GLLPANTEAARLEG-TKWGYGWN
Q9K2Q4                  AQVERNAGIIAE-SVKIAQN----AIKTVNPKDKATDYLTSKDKSVVSLQDRAAWGFGWN
Q9K1M2                  AELRKYADWGIK-SKAEILT-----AKPPKPNGVAEAAKIQADG-HADVKG-SDWGFGYQ
Q8VMI2                  LNLDYLLGMNQLGSLAGDGRASGSLMGVIGTLPDPRGVHLSVSKNKEMSSGVDGWGYSAR
                         ::    .                                .      ..   **:. .
                                  220       230       240       250       260
P10384                  ACILYELDKNNRYALTYRSEVKI-DFKGN-YSSDLNRAFNNYGLPIPTATGGATQ--SGY
Q8ZNA5                  ACILYELDKNNRYALTYRSEVKI-DFKGN-YSSDLPIAINRFNLPIPTATGGATQ--SGY
Q8ZD48                  TCILYEIDKENRYSFTYRSEVNI-DFDGD-YSNQLPVIFG--------GLGGKTV--PGS
Q9K2Q4                  ACVMYQFNEANRIGLAYHSKVDI-DFTDR-TATSLEAEVI---------EAGK----KGN
Q9K1M2                  LAWMWDINDRARVGVNYRSKVSH-TLKGD-AEWAADGAAAKAMWSTMLAANGYTANEKAR
Q8VMI2                  LCLLYKVAPTTNVGVSYMFKSHMNDLKGKGTVTAVDGIAGN------VPIEGEVR----F
                         . ::..    . .. *  :     : .                     *

                            270       280       290       300       310       320
P10384                  LTLNLPEMWEVSGYNRVDPQWAIHYSLAYTSWS-QFQQLKATSTSGDTLFQKHEG-----
Q8ZNA5                  LTLNLPEMWEVSGYNRVAPQWAIHYSLAYTSWS-QFQELKAKSTAGDTLFEKHEG-----
Q8ZD48                  LTLNLPAVWEVSGYNKVAPQWAIHYSMAYTTWS-SFKELKATASNGDVLFDKHEG-----
Q9K2Q4                  LTLTLPDYLELSGFHQLTDKFAVHYSYKYTHWS-RLTKLHASFEDGKKAFDKELQ-----
Q9K1M2                  VKIVTPESLSVHGMYKVSDKADLFGDVTWTRHS-RFDKAELVFEKEKTVVKGKSDRTTIT
Q8VMI2                  LDFNTPAKLDVGISHQVTDKWLIAFDVSRVFWKDALKDIKLGFASGMGDVDLKLP-----
                        : :  *   .:    ::  :  :  .   .  .  : . .         .. .

                                    330       340       350       360       370
P10384                  --FKDAYRIALGTTYYYDDNWTFRTGIAFDDSPVP-AQNRSISIPDQDRFWLSAGTTYAF
Q8ZNA5                  --FKDAYRIALGTTYYYDDNWTFRTGIAFDDSPVP-AQNRSISIPDQDRFWLSAGTTYAF
Q8ZD48                  --FRDAYRIALGTTYYYDDNWTFRTGIAFDDSPIP-AGNRSISIPDQDRFWLSAGTTYAF
Q9K2Q4                  --YSNNSRVALGASYNLDEKLTLRAGIAYDQAAS--RHHRSAAIPDTDRTWYSLGATYKF
Q9K1M2                  PNWRNTYKVGFGGSYQISEPLQLRAGIAFDKSPVRNADYRMNSLPDGNRIWFSAGMKYHI
Q8VMI2                  QDAKDQTIMAIGTSYSVTPRLTLRAGYRHATQPFN-DEGLLALIPAVLQDHASLGFSYQL
                            :   :.:* :*       :*:*  .   .          :*   :   * * .* :

                          380       390          400               410        421
P10384                  NKDASVDVGVSYMHGQSVKINE---GP--------YQFESEGKAWLFGTNFNYAF-
Q8ZNA5                  NKDASVDVGVSYMHGQSVKINE---GP--------YQFESEGKAWLFGTNFNYAF-
Q8ZD48                  NKNASVDVGIAYMKGQNVSITEKTPAPS----NTTYEFNSKGSAMLYGVNFNYTF-
Q9K2Q4                  TPNLSVDLGYAYLKGKKVHFKEVQKAVGGF-ITTTANYTSQAHANLYGLNLNYSF-
Q9K1M2                  GKNHVVDAAYTHIHINDTSYRTAKASGNDVDSKGASSARFKNHADIIGLQYTYKFK
Q8VMI2                  SKSGRFDAAYSHAFKESMTNRSAYNTS--------SPVKSSIAQDNFVLAYNYSF-
                          .  .* . ::   :.                       .          .* *
```

Fig. S2 Structural changes in the N-terminus of FadL. Stereoviews of experimental ($2F_o$-F_c) density, contoured at 1σ, of the N-terminus (residues 1-12) of monoclinic (A) and hexagonal (B) FadL, shown with their orientations in the full-length protein. In the models, oxygen atoms are shown in red and nitrogen atoms in blue.

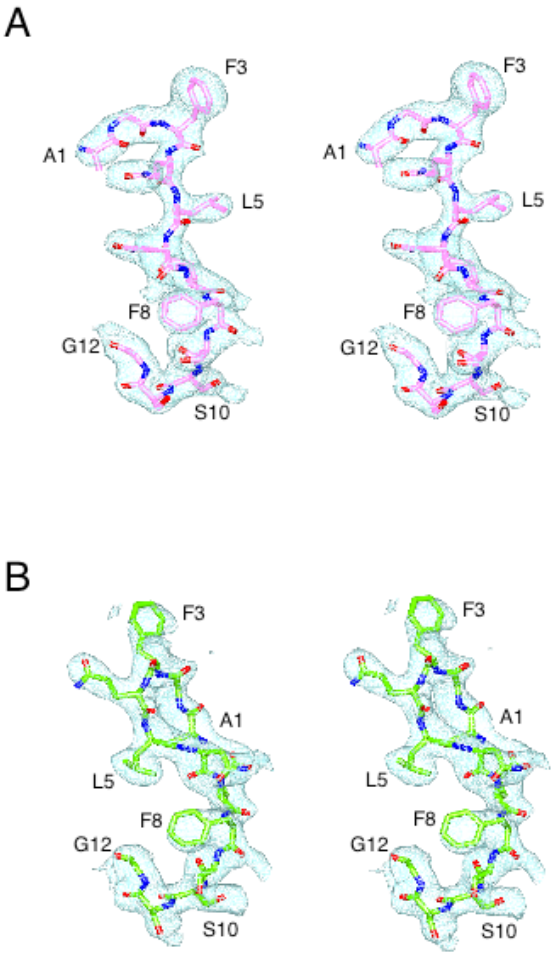
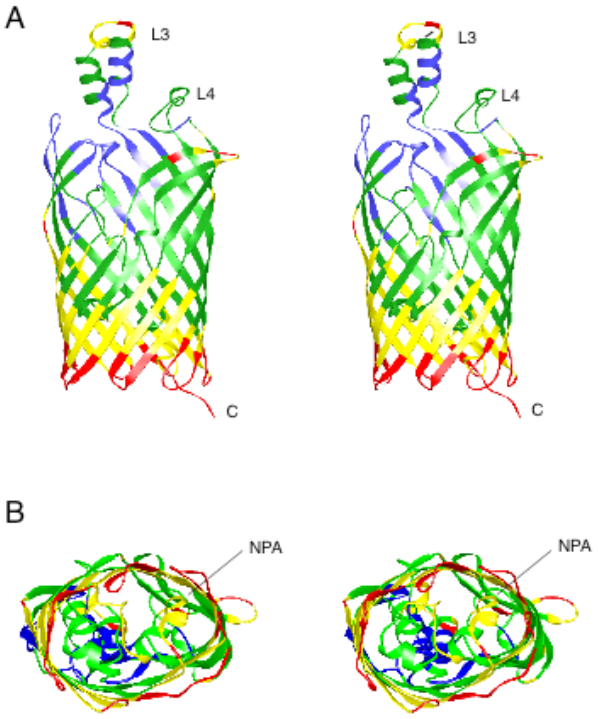


Fig. S3 Ribbon diagrams of a side view (A) and a bottom view (B) of monoclinic FadL showing the C_α B-factors: blue, B < 50 Å^2; green, 50 Å^2 • B < 75 Å^2; yellow, 75 Å^2 • B < 100 Å^2; red, B • 100 Å^2. The location of the NPA sequence in the hatch is shown in (B). The B-factors for the hexagonal crystal form are not shown, since their relative values are very similar to those in the monoclinic crystal form.



Supplementary Tables

Table S1 Fatty acid binding and transport by wild type and C-terminally His-tagged FadL. For details see Materials and Methods

Fatty acid binding (C18:1)[a]

Strain	Genotype	Binding (pmole/mg) (+/- SEM)[b]
LS6949[c]	*fadlΔ/fadDΔ/fadR*	89.9(12.4)
LS1908	*fadDΔ/fadRΔ*	417.8(12.7)
LS6949/pB22[d]	*fadlΔ/fadDΔ/fadR*/FadL^{6xHis}	723.6(23.9)
LS6949/pN130[e]	*fadlΔ/fadDΔ/fadR/*FadL^{+}	673.6(33.4)

Fatty acid transport (C18:1)

Strain	Genotype	Transport (pmole/min/mg protein) (+/- SEM)[b]
LS6164	*fadlΔ/fadR*	9.4(4.4)
LS1548	*fadRΔ*	684.5(60.7)
LS6164/pB22	*fadlΔ/fadR/*FadL^{6xHis}	617.3(52.5)

[a] FFA/BSA ration = 2, with BSA concentration = 173 µM (see Materials and Methods)

[b] SEM: standard error of the mean; n = 6 (transport), n = 4 (binding)

[c] For strain references see materials and Methods

[d] pB22 encodes the C-terminally hexa-histidine tagged FadL (FadL^{6xHis})

[e] pN130 encodes native FadL (FadL^{+})

Table S2 Data collection and refinement statistics

Data collection				
Monoclinic data set ($P2_1$)	Native	$YbCl_3$	K_2PtCl_4	
Wavelength (Å)	1.100	1.386	1.072	
Resolution (Å)	2.6	3.2	3.4	
Completeness (%)	91.1(78.1)	96.8(82.5)	94.5(88.3)	
Redundancy	3.3 (3.0)	7.0(5.8)	4.6(3.8)	
$I/\sigma I$	21.5(2.6)	19.5(2.4)	21.2(3.4)	
R_{sym} (%)[a]	6.9	15.0	9.2	
Hexagonal data set ($P3_121$)	Native	$OsCl_3$	UAc	K_2PtCl_4
Wavelength (Å)	1.008	1.139	1.008	1.070
Resolution (Å)	2.8	3.0	2.9	2.9
Completeness (%)	98.5(88.5)	98.1(99.0)	99.6(97.7)	90.3(85.3)
Redundancy	7.1	10.9	6.6	4.2
$I/\sigma I$	22.2(4.4)	31.8(4.4)	28.5(5.5)	18.7(3.1)
R_{sym} (%)	9.1	9.8	7.6	8.2
Refinement	$P2_1$		$P3_121$	
Resolution range (Å)	10.0-2.6		12-2.8	
R_{work} (%)[b], R_{free}[c]	25.7, 30.2		29.5, 33.1	
R.m.s.d.				
Bond lengths (Å)	0.0083		0.0104	
Bond angles (°)	1.39		1.67	
Protein atoms	6626		6504	
Water molecules	150		106	
Detergent molecules				
LDAO	4		2	
C_8E_4	3		0	
Copper ions	6		0	
Ramachandran statistics (%)				
Most favored, disallowed	82.1, 0.0		74.6, 0.0	

Values in parentheses are for the highest resolution shell

[a] $R_{sym} = \Sigma_{hkl}\Sigma_i|I_i(hkl) - I(hkl)|/\Sigma_{hkl}\Sigma_i|I_i(hkl)|$, where $I(hkl)$ is the average intensity

[b] $R_{work} = \Sigma_{hkl}||F_{obs}| - k|F_{calc}||/\Sigma_{hkl}|F_{obs}|$

c R_{free} = R_{work} for a selected subset (5%) of reflections that was not included in refinement

Methods References

S1. C. L. Ginsburgh, P. N. Black, W. D. Nunn, *J. Biol. Chem.,* **259**, 8437 (1984).

S2. L. M. Guzman, D. Belin, M. J. Carson, J. Beckwith, *Bacteriol*. **177**, 4121 (1995).

S3. G. D. van Duyne, R. F. Standaert, P. L. Karplus, S. L. Schreiber, J. Clardy, *J. Mol. Biol*. **229**, 105 (1993).

S4. M. Struyve, M. Moons, J. Tommassen, *J Mol Biol*., **218**, 141 (1991).

S5. Z. Otwinowski, W. Minor, in *Meth. Enzym*. (eds Carter, W.C.J. and Sweet, R.M., Academic press, New York), 307 (1993).

S6. Terwilliger, T. and Berendzen, J. (1999). Automated MAD and MIR structure determination. *Acta Cryst.* **D50**, 760-763.

S7. E. de la Fortelle, G. Bricogne, Methods Enzymol. 276, 472 (1997).

S8. Jones, T.A. and Kjeldgaard, M. (1997). Electron-density map interpretation. *Meth. Enzym*. **277B**, 173-207.

S9. Brunger, A. T. et al. (1998). Crystallography and NMR system: A new software suite for macromolecular structure determination. *Acta Cryst*. **D54**, 905-921.