Published in final edited form as:

Proteins. 2012 December ; 80(12): 2804–2809. doi:10.1002/prot.24182.

Crystal structure of the *Campylobacter jejuni* Cj0090 protein reveals a novel variant of the immunoglobulin fold among bacterial lipoproteins

Seonghee Paek, Fumihiko Kawai, Kyoung-Jae Choi, and Hye-Jeong Yeo*

Department of Biology and Biochemistry, University of Houston, Houston, TX 77204

Abstract

Bacterial lipoproteins play an important role in bacterial pathogenesis and physiology. The genome of *Campylobacter jejuni*, a major foodborne pathogen, is predicted to contain over 20 lipoproteins. However, the functions of the majority of *C. jejuni* lipoproteins remain unknown. The Cj0090 protein is encoded by a lipoprotein operon composed of *cj0089*, *cj0090*, and *cj0091*. Here, we report the crystal structure of Cj0090 at 1.9 Å resolution, revealing a novel variant of the immunoglobulin fold with β -sandwich architecture. The structure suggests that Cj0090 may be involved in protein-protein interactions, consistent with a possible role for bacterial lipoproteins.

Keywords

bacterial lipoprotein; *Campylobacter jejuni*; crystal structure; Cj0090; immunoglobulin fold

INTRODUCTION

Campylobacter jejuni is a Gram-negative commensal bacterium present in the intestinal tracts of a wide range of animals. It is an important food-borne pathogen and the leading cause of bacterial gastroenteritis in humans.¹ Moreover, severe post-infectious complications include Guillain-Barré syndrome, a significant cause of acute neuromuscular paralysis, and irritable bowel syndrome.^{2,3} Despite being a major diarrheal pathogen and the availability of multiple genome sequences, the pathogenesis of *C. jejuni* in humans is not well understood. This is in part because *C. jejuni* largely lacks homologs of virulence factors found in other pathogens.¹ Along with the significant emergence of antibiotic-resistant strains, the incidence of human *C. jejuni* infection is noticeably increasing. Thus, there is a growing interest in investigating novel virulence factors to better understand the pathogenic process of *C. jejuni* as a key step toward controlling the disease.

Bacterial lipoproteins are universal components of bacterial membranes, and are anchored to membranes via fatty acids, which covalently modify the N-terminal Cys residue of the mature protein. The ability of lipoproteins to reside in the bacterial membranes provides for diverse essential structural and functional roles in host-pathogen interactions, from surface adhesion to translocation of virulence factors into host cells.⁴ Given the broad occurrence of lipoproteins among bacteria and their structural characteristics, it is not surprising that at least one of the human toll-like receptors (TLRs), especially TLR2 in cooperation with TLR1 and TLR6, detects bacterial lipoproteins as an innate immune response to bacterial pathogens.⁵ Due to the ability to trigger the host inflammatory response, lipoproteins are

*Corresponding author: Tel: (1) 713 743 8377; Fax: (1) 713 743 8351; hyeo@uh.edu.

also involved in disease pathogenesis. Furthermore, surface-exposed lipoproteins that are crucial for survival in the host are of considerable interest as potential vaccine targets.

The genome of *C. jejuni* appears to contain over 20 lipoproteins.⁶ However, the functions of most *C. jejuni* lipoproteins remain unknown. To date, only one surface-exposed lipoprotein, JlpA, has a known crystal structure in *C. jejuni*.⁷ While *C. jejuni* belongs to a distinct clan of proteobacteria, predicted lipoproteins of *C. jejuni* often have no homologues in the sequence data banks. Yet, the signal peptide ‘lipobox’ region recognized by lipid modification enzymes appears to be similar to the *Escherichia coli* signal peptide. Recently, we initiated structure-function studies on putative lipoproteins of *C. jejuni* in order to discover their potential as vaccine candidates and their contribution to the pathogenesis. Among our target lipoproteins, the Cj0090 protein is encoded within lipoprotein gene cluster composed of *cj0089*, *cj0090*, and *cj0091*. A recent study reported the function and regulation of the *cj0089-cj0090-cj0091* operon, demonstrating that this lipoprotein operon is activated and directly regulated by CmeR, a pleiotropic transcription regulator modulating the expression of multiple genes including the multidrug efflux pump CmeABC.⁸ However, the function of Cj0090 remains undetermined.

Here, we report the crystal structure of Cj0090 refined to 1.9 Å resolution, revealing a novel form of the immunoglobulin fold that implies a role for Cj0090 in protein-protein interactions.

MATERIALS AND METHODS

Cloning, expression, and purification of Cj0090

The DNA fragment encoding Cj0090 without the signal peptide (Δ^{1-15} Cj0090) was amplified by PCR using genomic DNA from *Campylobacter jejuni* NCTC 11168 (ATCC 700819), *Pfu* DNA polymerase (Stratagene), and primers Cj0090f (5′-CCG GGT ACC TGT GCA CCA AGT TAT CAA ATA AAT TC-3′) and Cj0090r (5′-CCG GAG CTC TTA ATT TTT TGC TTT AAT TTC TAA TC-3′). For the cloning purpose, KpnI and SacI sites (underlined) were introduced to the forward and reverse primers, respectively. The PCR product was inserted into KpnI / SacI sites of the pET45b(+) (Novagen) to express Δ^{1-15} Cj0090 as an N-terminal His-tagged protein. The cloned gene sequence was confirmed by DNA sequencing.

For expression, the recombinant plasmid pET45b:: Δ^{1-15} Cj0090 was transformed into BL21(DE3). The cells were grown at 37°C in 1 L of Luria-Bertani media supplemented with 100 µg/ml ampicillin. When the culture reached an A_{600} 0.8, expression was induced with 0.3 mM IPTG at 23°C for 3 h. Subsequently, bacterial cells were collected by centrifugation at 4,000g (Sorvall) for 20 minutes, resuspended in 40 mL of buffer-A (20 mM Tris-HCl pH 8.0, 250 mM NaCl) containing 5 mM β-mercaptoethanol, 0.1% Triton-X and EDTA-free Protease Inhibitor Cocktail Tablets (Roche), and kept frozen at −20°C until use.

For purification, the frozen cells were thawed at room temperature and further lysed using sonication. After centrifugation at 40,000g for 20 min (Sorvall), supernatants were subjected to a binding reaction with Ni²⁺-NTA (Qiagen) resin for 30 min in a batch purification procedure. Protein-resin complexes were then packed onto a column and washed with buffer-A containing 25 mM imidazole. The proteins were eluted using a step gradient method with 75, 125, and 250 mM imidazole in buffer-A. Fractions containing the target protein were combined, concentrated using a Vivaspin concentrator, and further purified using a Sephacryl™ S-200 HR HiPrep™ 16/60 gel filtration column (GE Healthcare) equilibrated with buffer-B (50 mM HEPES pH 7.0, 200 mM NaCl, 0.1 mM EDTA, 5% glycerol). The highly pure elution fractions from gel filtration were combined and

concentrated to ~12 mg/mL for crystallization. The protein concentration was determined by UV spectroscopy or by Bradford assay (Bio-Rad).

Crystallization and data collection

The concentrated protein sample (~12 mg/mL in buffer-B) was screened for crystallization using commercially available screen kits. Small Cj0090 crystals were observed with several crystallization conditions. Following extensive optimization trials, the best crystals were obtained in a 1:1 mixture of protein and reservoir solution containing 0.1 M HEPES pH 7.0, 0.8 M NaCl using the hanging drop vapor diffusion method at 17°C. Lozenge-shaped crystals grew to typical dimensions of 150 μm $\times$ 75 μm $\times$ 50 μm for two weeks.

Native Cj0090 crystals were treated with cryoprotectant (0.1 M HEPES pH 7.0, 0.8 M NaCl, and 24% glycerol) and flash cooled in liquid N₂. For phasing, Br-derivatized crystals were prepared by soaking native Cj0090 crystals in cryoprotectant containing 1M KBr for 1 min and flash cooled in liquid N₂. Diffraction data were collected at 0.91956 Å for the Br-SAD datasets and at 0.97921 Å for native datasets on beamline 19ID at the Advanced Photon Source (Argonne), using an ADSC Quantum 315 CCD detector. Data sets were processed with HKL3000.⁹ Cj0090 crystals belonged to space group $P4_12_12$ with unit cell dimensions of $a=b=51.15$ Å, $c=186.79$ Å.

Structure determination and refinement

The structure was solved by SAD phasing and three Br sites were located using the program AutoSol from the PHENIX package.¹⁰ The initial model was generated by combining automatic building and manual adjustment with COOT¹¹ and was refined using the PHENIX package. The final model of the Cj0090 structure was obtained after iterative cycles of manual model building with COOT and refinement with REFMAC5.¹² Solvent molecules were placed at positions where spherical electron density peaks were found above 1.5 σ in the $|2F_o - F_c|$ map and above 3.5 σ in the $|F_o - F_c|$ map, and where stereo-chemically reasonable hydrogen bonds were allowed. Structural analysis of the final model using the Protein Data Bank (PDB) validation suite indicated almost all residues are in the most favored regions of the Ramachandran plot (Table 1). The coordinates and structure factors of the Cj0090 structure have been deposited at PDB ID code 4GIO.

RESULTS AND DISCUSSION

Structure determination and molecular architecture of Cj0090

The Cj0090 protein was annotated as a putative lipoprotein with the characteristic signal peptide features such as hydrophobic regions and lipobox sequence.¹³ As a first step to examine signal peptide processing of Cj0090, we constructed *Escherichia coli* expression systems for a full length Cj0090 protein and for a Cj0090 variant without the N-terminal signal peptide (Δ^{1-15} Cj0090: amino acid residues 16 to 122). The latter serves as a control, since it is over-expressed in the cytoplasm and cannot be processed to the cell surface. Membrane fractionation and glycine-acid extraction experiments indicate that the signal peptide of Cj0090 is efficiently processed to present Cj0090 on the *E. coli* cell surface and to the extracellular medium of *E. coli* (Fig. S1). While this observation does not necessarily translate to what happens in *Campylobacter jejuni*, the signal peptide of Cj0090 is functional in *E. coli* (a heterologous system), which increased our motivation and paved the way for the structural study. For crystallization, the predicted mature form of Cj0090 (Δ^{1-15} Cj0090) was produced as His-tagged protein in *E. coli*. The Δ^{1-15} Cj0090 sequence does not contain any Met residue, and Br-derivatized crystals were prepared by soaking native crystals for the phasing purpose. The structure was solved by single wavelength anomalous dispersion method at the Br edge. The Cj0090 crystals contain 2 molecules in the asymmetric unit. The

final model contains residues 26 to 121 for both chain A and chain B (Table 1). No electron density was observed for N-terminal residues 16 to 25, indicating that this region is likely flexible or unstructured, although the protein recovered from crystals showed the same size as the purified protein sample.

The crystal structure of Cj0090 reveals a single domain structure composed of seven β -strands forming two β sheets (β sandwich), one containing 4 strands (EDGA) and the other containing 3 strands (FCB), with a Greek key motif (Fig. 1A and B). Thus, the Cj0090 structure represents an immunoglobulin (Ig)-like β sandwich fold, yet differs from the most common Ig-like superfamily folds referred to as C-, V-, H-, and I-type folds.^{14,15} Current classification systems of Ig-like folds are mostly based on the number of β strands distributed in two sheets with characteristic topology and connectivity. In the classical C-type as occurring in the constant domains of immunoglobulin, seven β strands are arranged in two β sheets (sheet 1, strands DEBA; and sheet 2, strands GFC) (Fig. S2). On the other hand, in the other variants, V-, H-, and I-type, additional strands are formed between strands C and D. Over the recent years, the variation in number and position, with respect to the classic Ig fold, of additional β strands has been observed and defined topologically distinct subtypes of Ig-like fold such as C1-type and C2-type (Fig. S2). According to the classification of Ig-like fold, Cj0090 is a variant of type-C that is distinct from the type-V in the number of strands (typically, 7 strands for type-C and 9 strands for type-V). Contrast to the C1-type (that is the classical C-type), the C2-type (also referred to as S-type¹⁴) fold has three strands for sheet 1 and four strands for sheet 2. A particular case of C2-type is referred to as the fibronectin type III (Fn3) fold that is found in many animal proteins and involved in ligand binding, since the Fn3 structure does not contain the conventional disulfide bond commonly present in other C2-type Ig related proteins.^{14,15} Thus, the fold of Cj0090 can be seen as a circular permutation product of the C2-type fold (see below). Indeed, if strand A were to be at the C-terminus by permutation, strands of Cj0090 would occur as EBA (from FCB) and DCFG (from EDGA), which is a typical C2-type topology (or Fn3 fold) (Fig. S2). Taken together, the fold of Cj0090 is unique and a novel variant of Ig-like superfamily.

Two protomers in the asymmetric unit are very similar and superimposed with an r.m.s.d. of 0.29 Å (aligning all Ca atoms). Localized conformation differences between two protomers are observed for several regions within connecting loops, including the connecting loop between strands B and C (⁴⁸NSNG⁵¹), and a portion of the connecting loop between strands D and E (⁸²RDVLNE⁸⁷) (Fig. 1A). The crystallographic interface between two protomers in the asymmetric unit involves strand A, the N-terminus of the connecting loop between strands A and B, the connecting loop between strands D and E, and strand G, resulting in ~1600 Å² of molecular surface becoming buried (per subunit). Based on size-exclusion chromatography, Cj0090 was found in a monomer-dimer equilibrium in solution (Fig. S3). Both forms of the proteins crystallized in very similar conditions, but we only obtained high quality diffraction data from crystals produced from the dimer solution. Monomer-dimer equilibria have been reported for numerous proteins, especially proteins with binding activities.¹⁶ While the specific function of Cj0090 remains elusive, we speculate that Cj0090 is involved in molecular interactions as a lipoprotein.

Structural comparison

At the primary structure level, Cj0090 is highly conserved within *Campylobacter* spp. and has homologs with low (below 36%) sequence identities from various organisms. The function of these proteins is unknown. Structure homology searches using the DALI¹⁷ server revealed a large number of solved structures (Z score over 5) belonging to the Ig superfamily. However, none of these proteins has sequence identity over 17% with Cj0090. Among these, two bacterial proteins were noteworthy: (1) uncharacterized protein referred to as Shew_1734 (PDB 3o0l; Z score 9.9 and r.m.s.d. 2.3) and (2) ApaG (PDB 1tza; Z score

6 and r.m.s.d. 2.6). Both proteins are similar to Cj0090 in size. The crystal structure of Shew_1734 from *Shewanella loihica* (solved by the Joint Center for Structural Genomics) represents the closest structural homolog of the Cj0090 structure. Given low sequence identity between Cj0090 and Shew_1734 (15% identical excluding signal peptide), it was not surprising that BLAST searches failed to identify Shew_1734 as a homolog. Analogous to Cj0090, Shew_1734 is a putative bacterial lipoprotein and contains the signature sequence peptide (Fig. 1A). Despite the low sequence identity, Cj0090 and Shew_1734 clearly share the same topology except for the N-terminus (strands A₁ and A₂) (Fig. 2A and B). Interestingly, strands A₁ and A₂ contribute to a 5-stranded β sheet with strands FCB in Shew_1734, whereas strand A contributes to a 4-stranded β sheet with strands GDE in Cj0090. Thus, this novel variant of the Ig fold observed in Cj0090 may be conserved with minor differences among bacterial lipoproteins. The next level of structural homology, albeit the low sequence identity of 6%, was found in the ApaG proteins (~14 kDa) that belong to a conserved protein family from Gram-negative bacteria. While the function of the ApaG proteins is unknown, it has been suggested that ApaG may be involved in the metabolism of nucleotide polyphosphates.¹⁸ Unlike Cj0090 and Shew_1734, the ApaG proteins do not contain a signal peptide and thus are not lipoproteins. The structures of four ApaG proteins from *Xanthomonas citri*,¹⁸ *Vibrio cholera*, *Shewanella oneidensis*, and *Bordetella pertussis* have been determined. All the structures are very similar and characterized by seven antiparallel β strands forming two β sheets, one with three strands (ABE) and the other with four strands (GFCD) (Fig. 2C). Thus, ApaG adopts a Fn3 fold that is a particular case of C2-type, as described above. The topology diagrams in Figure 2 indicate the fold relationship between Cj0090-like proteins and ApaG-like proteins, which is best described as circular permutations. Among different mechanisms for the fold change in evolution of protein structures, circular permutations are changes in protein connectivity that can be visualized through ligation of the termini and cleavage at another site. The N and C termini of protein domain structures are often placed in close proximity. Such an arrangement favors circular permutations.¹⁹ Clearly, the Cj0090 fold and the ApaG fold can be switched by circular permutation (Fig. S2).

In summary, the Cj0090 structure reveals a new variant of Ig superfamily fold that appears to be conserved among bacterial lipoproteins. Having an Ig-like fold suggests that Cj0090 is involved in mediating protein-protein interactions, consistent with a possible role for bacterial lipoproteins.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Results shown in this report are derived from work performed at Argonne National Laboratory, Structural Biology Center (19ID) at the Advanced Photon Source. Argonne is operated by UChicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357. We thank Dr. Bill Widger for valuable discussion. This work was supported in part by NIH grant AI068943 and the R. Welch Foundation grant E-1616 (to H.J.Y.).

REFERENCES

1. Poly F, Guerry P. Pathogenesis of *Campylobacter*. Curr Opin Gastroenterol. 2008; 24:27–31. [PubMed: 18043229]
2. Nachamkin I, Liu J, Li M, Ung H, Moran AP, Prendergast MM, Sheikh K. *Campylobacter jejuni* from patients with Guillain-Barré syndrome preferentially expresses a GD(1a)-like epitope. Infect Immun. 2002; 70:5299–5303. [PubMed: 12183587]

3. Zilbauer M, Dorrell N, Wren BW, Bajaj-Elliott M. *Campylobacter jejuni*-mediated disease pathogenesis: an update. *Trans R Soc Trop Med Hyg.* 2008; 102:123–129. [PubMed: 18023831]
4. Tokuda H, Matsuyama S. Sorting of lipoproteins to the outer membrane in *E. coli*. *Biochim Biophys Acta.* 2004; 1693:5–13. [PubMed: 15276320]
5. Aliprantis AO, Yang RB, Mark MR, Suggett S, Devaux B, Radolf JD, Klimpel GR, Godowski P, Zychlinsky A. Cell activation and apoptosis by bacterial lipoproteins through toll-like receptor-2. *Science.* 1999; 285:736–739. [PubMed: 10426996]
6. Babu MM, Priya ML, Selvan AT, Madera M, Gough J, Aravind L, Sankaran K. A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins. *J Bacteriol.* 2006; 188:2761–2773. [PubMed: 16585737]
7. Kawai F, Paek S, Choi KJ, Prouty M, Kanipes MI, Guerry P, Yeo HJ. Crystal structure of JlpA, a surface-exposed lipoprotein adhesin of *Campylobacter jejuni*. *J Struct Biol.* 2012; 177:583–588. [PubMed: 22245776]
8. Oakland M, Jeon B, Sahin O, Shen Z, Zhang Q. Functional characterization of a lipoprotein-encoding operon in *Campylobacter jejuni*. *PLoS One.* 2011; 6:e20084. [PubMed: 21625391]
9. Minor W, Cymborowski M, Otwinowski Z, Chruszcz M. HKL-3000: the integration of data reduction and structure solution—from diffraction images to an initial model in minutes. *Acta Crystallogr D Biol Crystallogr.* 2006; 62:859–866. [PubMed: 16855301]
10. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr.* 2010; 66:213–221. [PubMed: 20124702]
11. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr.* 2004; 60:2126–32. [PubMed: 15572765]
12. The CCP4 Suite: Programs for Protein Crystallography. *Acta Crystallogr D Biol Crystallogr.* 1994; 50:760–763. [PubMed: 15299374]
13. Babu MM, Priya ML, Selvan AT, Madera M, Gough J, Aravind L, Sankaran K. A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins. *J Bacteriol.* 2006; 188:2761–2773. [PubMed: 16585737]
14. Bork P, Holm L, Sander C. The immunoglobulin fold. Structural classification, sequence patterns and common core. *J Mol Biol.* 1994; 242:309–320. [PubMed: 7932691]
15. Halaby DM, Poupon A, Mornon J. The immunoglobulin fold family: sequence analysis and 3D structure comparisons. *Protein Eng.* 1999; 12:563–571. [PubMed: 10436082]
16. Woodbury RL, Hardy SJ, Randall LL. Complex behavior in solution of homodimeric SecA. *Protein Sci.* 2002; 11:875–882. [PubMed: 11910030]
17. Holm L, Rosenström P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* 2010; 38:W545–549. [PubMed: 20457744]
18. Cicero DO, Contessa GM, Pertinhez TA, Gallo M, Katsuyama AM, Paci M, Farah CS, Spisni A. Solution structure of ApaG from *Xanthomonas axonopodis* pv. citri reveals a fibronectin-3 fold. *Proteins.* 2007; 67:490–500. [PubMed: 17256769]
19. Grishin NV. Fold change in evolution of protein structures. *J Struct Biol.* 2001; 134:167–185. [PubMed: 11551177]

Figure 1.

Structure of Cj0090. **(A)** Sequence alignment of Cj0090 and representative putative lipoproteins: Cj, *Campylobacter jejuni* (accession#: YP_002343550); Sl, *Shewanella loihica* (accession#: YP_001093860); Vc, *Vibrio cholerae* (accession#: NP_231529); Se, *Salmonella enterica* (accession#: NP_460176); and Ec, *Escherichia coli* (accession#: NP_287238). The signal peptides are indicated in red box. The Cj0090 secondary structures are indicated above the primary sequences. Identical residues, strongly conserved residues, and weakly conserved residues are highlighted in blue hues. **(B)** Cartoon representation of the Cj0090 structure containing two β -sheets (blue, AGDE; magenta, BCF). The β strands are labeled as A to G. **(C)** The Cj0090 dimer present in the asymmetric unit (subunit A, blue/magenta; and subunit B, pale blue/magenta).

Figure 2.

Topology diagrams of Cj0090 and close structural homologs with Ig fold variants. The corresponding cartoon representation of the structure is shown to the right. The conserved core strands are labeled and colored in gold and silver. The structurally variable strands are highlighted in red. (A) The Cj0090 structure. (B) Uncharacterized putative lipoprotein Shew_1734 (PDB id: 3o0l). (C) ApaG protein (PDB id: 1tza).

Table I**Data collection, Phasing, and refinement Statistics**

Data Collection	Native	Br -SAD
Wavelength (Å)	0.97921	0.91956 (peak)
Space group	$P_{41}2_12$	$P_{41}2_12$
Cell dimensions <i>a</i> , <i>b</i> , <i>c</i> (Å)	51.15, 51.15, 186.79	51.14, 51.14, 186.51
Resolution (Å) ^a	20.00 – 1.90 (1.97 – 1.90)	50.00–2.20 (2.28–2.20)
Reflections (Total / Unique)	159523 / 19705	112775 / 13060
Completeness (%) ^a	95.5 (88.8)	95.6 (87.1)
R_{merge} (%) ^{a, c}	7.1 (31.3)	8.4 (32.5)
$\langle I/\sigma(I) \rangle$ ^a	25.3 (3.7)	21.2 (2.2)
Redundancy ^a	8.1 (4.0)	8.7 (3.4)
Sigma cutoff	0.3	0.0
Phasing		
Mean FOM after AutoSol ^b		0.62
Structure Refinement		
Resolution (Å)	20-1.90	
No. reflections	18693	
R/R_{free} ^d (%)	19.6 / 24.4	
No. atoms (Protein / Water)	1571/ 133	
Protein molecules	2	
Average B factors		
Protein / Water / Br (Å ²)	25.3 / 29.7 / 28.1	
rms deviation		
Bonds (Å)	0.023	
Angles (°)	1.847	
Ramachandran plot (%)		
favored region	100	
allowed region	0	
outlier region	0	

^aThe values in parentheses are for the highest resolution shells.

^bThe figure of merit (FOM) = $|\bar{F}_{\text{best}}| \cdot |\bar{F}|$.

^c $R_{\text{merge}} = \sum |I_j - \langle I \rangle| / \sum I_j$, where I_j is the intensity of an observation, $\langle I \rangle$ is the mean value for that reflection, and the summations are over all equivalents.

^d R -factor = $\sum h |F_O(h) - |F_C(h)|| / \sum h F_O(h)$, where F_O and F_C are the observed and calculated structure factor amplitudes, respectively. R_{free} was calculated with 5% of the data excluded from the refinement.