FolX from *Pseudomonas aeruginosa* is octameric in both crystal and solution

Mads Gabrielsen[1a], Katherine S. H. Beckham[1a], Richard J. Cogdell[b], Olwyn Byron[c], Andrew J. Roe[a*]

[a]Institute of Infection, Immunity and Immunology, [b]Institute of Molecular, Cell and Systems Biology, [c]School of Life Sciences, College of Medical, Veterinary and Life Sciences, Sir Graeme Davies Building, University of Glasgow, Glasgow, G12 8QQ, UK

[1]The authors contributed equally to this work.

[*]To whom correspondence should be addressed

Sir Graeme Davies Building

120 University Place

University of Glasgow

Glasgow

G12 8QQ

UK

Andrew.roe@glasgow.ac.uk

Phone: (+44) 141 330 2980

Fax: (+44) 141 330 4600

URL: http://www.gla.ac.uk:443/ibls/staff/staff.php?who=Pee~Pd

Abbreviations: FolX, 7,8-dihydroneopterin-triphosphate-epimerase; AUC, analytical ultracentrifugation; SAXS, small-angle X-ray scattering, SV, sedimentation velocity; SE, sedimentation equilibrium

Abstract

FolX encodes an epimerase that forms one step of the tetrahydrofolate biosynthetic pathway, which is of interest as it is an established target for important drugs. Here we report the crystal structure of FolX from the bacterial opportunistic pathogen *Pseudomonas aeruginosa,* as well as a detailed analysis of the protein in solution, using analytical ultracentrifugation (AUC) and small-angle X-ray scattering (SAXS). In combination, these techniques confirm that the protein is an octamer both in the crystal structure, and in solution.

Highlights

The structure of FolX from *P.aeruginosa* was investigated using X-ray crystallography, AUC and SAXS. This study represents an interesting and unusual case study where the combination of AUC and SAXS allows a limited x-ray structure to be 'improved' so that the protein's oligomeric form can be unequivocally determined. These biophysical techniques show definitively that the protein is an octamer in both crystal and in solution. Despite only a small proportion of the two tetramers contributing to the interfaces, they consist of a network of 16 hydrogen bonds and 8 salt bridges providing the necessary stability for the octamer.

Introduction

The tetrahyrofolate biosynthetic pathway is an established target for important antibiotics including trimethoprim that inhibits the activity of bacterial dihydrofolate reductase [1-3]. However, with the ever-increasing spread of antibiotic resistance in

bacteria, there is an urgent need to explore alternative approaches to controlling pathogens. Recent work has elucidated the target proteins of one group of so-called "anti-virulence" compounds, the salicylidene acylhydrazides. In this published study, several target proteins were identified including FolX, encoded by a gene only carried by the *Gammaproteobacteria*. FolX has been shown to act as an epimerase (Figure 1) in conjunction with FolM, a reductase, during the conversion of dihydroneopterin triphosphate to tetrahyrobiopterin [4]. Therefore, FolX and FolM are essential for tetrahydomonapterin synthesis in species such as *Escherichia coli* and *Pseudomonas aeruginosa*.

Here we report the crystal structure of FolX from *P. aeruginosa* as well as a detailed analysis of the protein in solution using analytical ultracentrifugation (AUC) and small-angle X-ray scattering (SAXS). The combination of these techniques confirms that the octameric crystal structure is consistent with the biological state of FolX in solution.

Materials and methods

Protein expression and purification

The gene encoding FolX was amplified from *P. aeruginosa* genomic DNA and cloned into the vector pET-151 (Invitrogen) [5]. The expression and purification was performed following previously published methods [6].

Crystallisation

Purified FolX was dialysed against 20 mM Tris pH 7.5, 50 mM NaCl and concentrated to approximately 6 mg ml^{-1}, based on the absorbance at 280 nm (A$_{280}$) and an extinction coefficient of 5960 M cm^{-1} derived from the sequence composition. Screens were set up using commercially available crystallisation kits, using vapour

diffusion, with drops consisting of 500 nl protein solution and 500 nl reservoir. Cubes (0.2 x 0.2 x 0.2 mm) appeared in conditions, containing 40% (v/v) 1,2-propanediol, 100 mM HEPES pH 7.5, within a week, at room temperature.

Data collection, processing and structure solution

Crystals were flash-frozen in liquid nitrogen, with no further cryo-protection, and brought to Diamond Light Source, station I03. Data were collected on a PILATES 6M detector, at a wavelength of 0.97625 Å. A total of 127º of data were collected using an increment of 0.15º. Data were processed using MOSFLM [7] and scaled in SCALA [8]. The structure was determined using the Balbes molecular replacement server [9], which identified pdb entry 1B9L [10] as the best search model. The solved structure was refined with BUSTER [11], using TLS parameterisation and torsion restraints from the search model, and inspected and altered when required, using COOT [12]. Waters were added using BUSTER. The geometry of the structure was validated by MOLPROBITY [13]. The data showed signs of anisotropy and were processed to 3.0 Å, based on the statistical factors presented in Table 1, although diffraction spots could be detected closer to 2.5 Å.

Analytical ultracentrifugation

Purified FolX was dialysed against 20 mM Tris pH 7.5, 150 mM NaCl and concentrated to approximately 10 mg ml^{-1}, based on the A$_{280}$. Analytical ultracentrifugation (AUC) was carried out in a Beckman Coulter (Palo Alto, CA) Optima XL-I analytical ultracentrifuge. Sedimentation velocity (SV) experiments were performed at 4°C at a rotor speed of 49 krpm. 360 µl of sample, at concentrations of FolX ranging between 0.2 and 10 mg ml^{-1}, were loaded into double sector centrepieces. Data were acquired every 7 minutes with interference and absorbance optics and were subsequently analysed using SEDFIT [14]. The partial specific volume of FolX (0.737 g ml^{-1} / 0.743 g ml^{-1}), the buffer density (1.00677 g ml^{-1} /

1.00499 g ml^{-1}) and viscosity (0.0156 P / 0.0102 P) at 4°C and 20°C respectively, were all calculated using the program SEDNTERP [15]. Sedimentation equilibrium (SE) experiments were carried out with the same range of FolX concentrations using 90 µl of sample with a rotor speed of 23 krpm. Scans were taken every 3 h until analysis of the scans, using WinMATCH (Jeffrey Lary, University of Connecticut, Storrs, CT, USA), indicated that equilibrium had been reached. SE data were analysed using Origin and were fitted with a tetramer-octamer model with non-ideality. The dissociation constant was calculated using the method described by Solovyova *et al*. [16].

Small angle X-ray scattering measurements
SAXS data were collected at the ESRF ID14 EH3 beamline for samples with a range of FolX concentrations of between 0.5 and 6 mg ml^{-1} in 20 mM Tris pH 7.5, 150 mM NaCl. No concentration dependence effects were observed, therefore data from the highest concentration (6 mg ml^{-1}) were processed for further analysis. Initial processing of the data was done using PRIMUS and p(r) analysis was carried out using GNOM [17]. The resolution of the data was calculated from the highest angle at which useable scattering data were recorded [16] *Ab initio* models of FolX were generated from the experimental data using DAMMIF [18]. Twenty DAMMIF models were superimposed and averaged using DAMAVER [19] and the averaged model was superimposed onto the crystal structure using SUPCOMB [20].

Figures were made using ALINE [21] and PyMOL [22].

Results and Discussion
Monomeric structure

The monomeric structure of FolX comprised a four-stranded antiparallel sheet, composed of β1 (residues 10-12 and 16-20), β2 (residues 33-42), β3 (residues 98-106) and β4 (residues 114-121) (Figure 2A). The broken β1 strand, caused by a small kink introduced in the strand, was also observed in the structure of FolX from *E. coli* [10]. A short α-helix (α1, 4 residues) is located in the loop between sheets β1 and β2, whereas α-helices α2 and α3 are nestled against the sheet on the concave side. The structure is not complete, as only residues 6-46 and 55-122 can be observed in the electron density. There are a number of side chains that cannot be observed in the electron density, and have, accordingly, been cut back to the last ordered atom.

The sequence and structure of FolX are highly conserved throughout the gammaproteobacteria, and the sequence identity and similarity between FolX from *P. aeroginosa* and *E. coli* is 60% and 78% respectively (based on a level of 0.7, using the ALSCRIPT algorithm [23]) (Figure 2B). The structures of the two homologues superpose with a root mean square deviation (r.m.s.d.) of 0.99 Å for 100 Cα. The main differences are mostly limited to the loop between α2 and β3. These differences are not localised near the interfaces, or the putative active site.

Active site

Ploom et al. [10] suggested a putative active site for FolX from *E. coli* involving residues that are mostly conserved between *E. coli* and *P. aeroginosa* with only 2 substitutions (Asn to Glu and Lys to Arg, respectively) (Figure 2B). Most of these residues are disordered and not present in the electron density, implying that the active site of *P. aeruginosa* FolX is flexible in the absence of substrate, or exhibits multiple conformations.

Quaternary structure

FolX crystallised in space group *I*432, with a single chain in the asymmetric unit. The Protein Interfaces, Surfaces and Assemblies (PDBePISA) server [24] suggests that the oligomeric state, based on the crystallographic symmetry, is tetrameric (Figure 3A), consisting of a circle made up of the convex sides of the sheets of the subunits facing each other. The interfaces between the subunits are made up by strands $\beta 4$ on one subunit and $\beta 1$ on the next in the circle. The contacts comprise 20 residues on each strand, accounting for ~18% of the total solvent accessible area of the tetramer. The interface is made up by hydrogen bonds formed between a number of residues, mostly involving main chain nitrogen and carboxyl groups (Supplementary Table 1).

Previous studies have suggested that FolX is an octamer [10, 25]. However, when the crystal structure is analysed by PDBePISA, it is predicted to be a tetramer. The only interactions between the tetramers, in what would be a dimer of tetramers (Figure 3B), involve 7% of the solvent accessible area compared with 18% involved in stabilising the tetramer itself. Analysis of the interface between the tetrameric rings, using PDBePISA, indicates that there are 16 hydrogen bonds and 8 salt bridges connecting the two tetramers (Supplementary Table 2), which may suggest that this interaction is significant despite the relatively small interface area.

In order to try to get conclusive evidence for the quaternary state of FolX, analytical ultracentrifugation was performed. Sedimentation velocity experiments revealed that FolX was present as a single species in solution, as evidenced by a single peak in the concentration distribution of the apparent $s_{20,w}$ (c(s), Figure 4A). The infinite dilution sedimentation coefficient ($s^0_{20,w}$) of FolX, derived from the concentration dependence of $s_{20,w}$, determined by fitting the data with a non-interacting discrete

species model, is 6.09 ± 0.03 S. This corresponds with the value of $s^0_{20,w}$ computed (using SOMO [26]) for the octamer crystal structure (5.97 S) and not with that computed for the tetramer (3.62 S).   Sedimentation equilibrium data fitted with a single species model indicated the presence of a species with a mass of 141,500 ± 6,659 Da at infinite dilution (M^0). This value is slightly lower than the calculated octamer mass of 143,864 Da, therefore, in order to improve the fit of the model parameters, the effects of non-ideality and the presence of a tetramer-octamer equilibrium were introduced into the data analysis. Addition of non-ideality improved the χ^2 of the global fit from 0.01890 to 0.00316. Extending the model to include a tetramer-octamer equilibrium further improved the fit to a χ^2 of 0.00297 and gave a K_d of 0.887 μM. The fit to the data along with the resultant residuals is shown in Figure 4B&C.

SAXS structure

To further confirm the oligomeric state of FolX, the solution structure was determined using SAXS. Indirect Fourier transformation of the data using GNOM [17] indicated a D_{max} of 177 Å and an Rg of 44.11 ± 0.6 Å. An *ab initio* model of FolX was generated using DAMMIF [18], imposing *P*4 symmetry, based on the crystal structure, and the fit of the final averaged model to the experimental data is shown in Figure 5A. We have determined the K_d of FolX to be 0.887 μM, therefore at the protein concentration used in the SAXS study 99.5% (by mass) of the protein would have been fully octamerised. The crystal structure of the octamer was superposed onto the 10.5 Å resolution envelope of FolX in solution, (Figure 5B). The octamer crystal structure fits well into the envelope confirming that this is the true solution oligomeric state.

Here we present the structure of FolX from *P. aeruginosa*, which is an octamer in both the crystal and in solution (in equilibrium with its tetrameric form). Despite only

7% of the surface area of the two tetramers being involved in the interface these interactions are stabilised by 16 hydrogen bonds and 8 salt bridges. Consequently the association of the two tetrameric rings is strong, with a K_d of 0.887 µM indicating a high affinity association.

The active site of FolX is predicted to comprise residues from two adjacent subunits, which suggests that the tetramer is essential for the activity of the enzyme. The functional need for an octameric form is not apparent from the previously published biochemical data.

FIGURES

Figure 1: Chemdraw representation of the epimerase reaction catalysed by FolX, where dihydroneopterin is converted to dihydromonapterin triphosphate.

FIGURE 2: (A) Cartoon representation of a subunit of FolX. The helices are coloured in purple and the strands are coloured in teal. The N- and C-termini have been labelled. (B) Structural alignment of FolX from *P. aeruginosa* and *E. coli*, with secondary structure elements coloured as above. Highlighted in yellow are the

residues involved in the putative active site. Residues in grey indicate residues not modelled in the electron density.

FIGURE 3: Oligomerisation of FolX. The tetramer ring formed by FolX (A) dimerises to form an octamer (B).

FIGURE 4: AUC analysis of FolX. (A) c(S) distributions derived via SEDFIT from SV data for varying concentrations of FolX are dominated by a peak at $s_{20,w} \approx 4$ S. Different concentrations are shown in different colours: 0.2 mg ml^{-1}, light pink; 0.5 mg ml^{-1}, purple; 1 mg ml^{-1}, grey; 2.5 mg ml^{-1}, green; 5 mg ml^{-1}, pink, 7.5 mg ml^{-1}, blue; 10 mg ml^{-1}, orange. (B & C) The global fit to SE data using a non-ideal tetramer-octamer self-association model. The experimental data and fits are shown in (B); the residuals for each fit are plotted in (C). Different protein concentrations are represented by the colour scheme used in (A).

FIGURE 5: Solution structure of FolX. (A) The experimental scattering data (grey) with the fit of the DAMMIF model to the data shown in pink. Inset is the pairwise distribution p(r) function of the data. (B) The octamer crystal structure superimposed onto the DAMMIF model shown as side and top view.

TABLE

Table 1: Data collection, refinement and model quality statistics for FolX. Values in brackets represent the values of the highest resolution shell. Rmsd is root-mean square deviation.
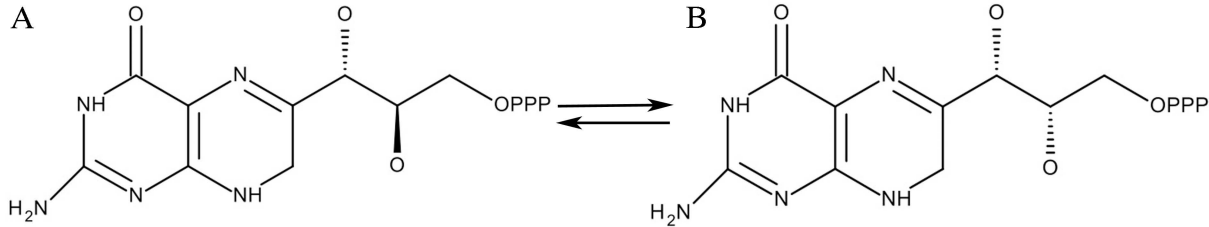
PDB code	4AEY
Space group	*I*432

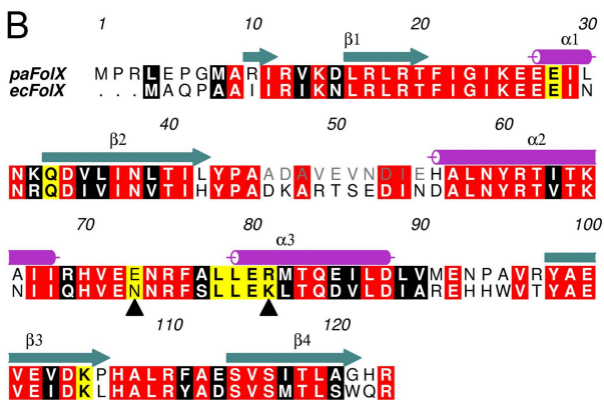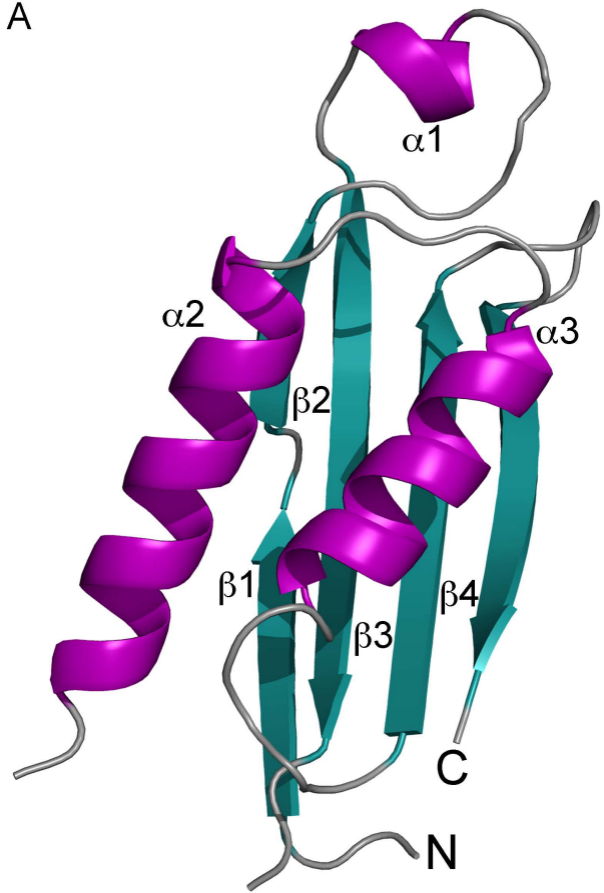|	
Unit Cell (Å)	a = 133.98
Resolution (Å)	94.74 - 3.00 (3.16-3.00)
Observed reflections	103928
Unique reflections	4383
Multiplicity	23.7 (25.1)
Completeness (%)	100.0 (100.0)
R_{meas} (%)	5.1 (141.6)
R_{pim} (%)	1.1 (38.4)
I/σI	34.3 (2.9)
Wilson B (Å^2)	115.3
Protein residues / atoms	107 / 754
Water molecules	8
Rwork (%)	28.3
Rfree (%)	31.1
R.m.s.d. for bond lengths (Å) / angles (°)	0.01 / 1.49
Molprobity score	13.9 [97th percentile]

REFERENCES

1.	Achari A, Somers DO, Champness JN, Bryant PK, Rosemond J & Stammers DK (1997) Crystal structure of the anti-bacterial sulfonamide drug target dihydropteroate synthase. *Nat Struct Biol* **4**, 490-497.

2.	Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM, et al. (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**, 496-512.

3.	Spitzy KH (1991) Use of antibacterial chemotherapy. A historical comparison. *Infection* **19 Suppl 1**, S52-56.

4.	Pribat A, Blaby IK, Lara-Nunez A, Gregory JF, 3rd, de Crecy-Lagard V & Hanson AD (2010) FolX and FolM are essential for tetrahydromonapterin synthesis in *Escherichia coli* and *Pseudomonas aeruginosa*. *J Bacteriol* **192**, 475-482.

5.	Wang D, Zetterstrom CE, Gabrielsen M, Beckham KS, Tree JJ, Macdonald SE, Byron O, Mitchell TJ, Gally DL, Herzyk P, et al. (2011) Identification of bacterial target proteins for the salicylidene acylhydrazide class of virulence-blocking compounds. *J Biol Chem* **286**, 29922-29931.

6.	Gabrielsen M, Zetterstrom CE, Wang D, Beckham KS, Elofsson M, Isaacs NW & Roe AJ (2010) Expression, purification, crystallization and initial X-ray diffraction analysis of thiol peroxidase from *Yersinia pseudotuberculosis*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **66**, 1606-1609.

7.	Leslie AGW (1992) Recent changes to the MOSFLM package for processing film and image plate data *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography* **26**.

8.	Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* **62**, 72-82.

9.   Long F, Vagin AA, Young P & Murshudov GN (2008) BALBES: a molecular-replacement pipeline. *Acta Crystallogr D Biol Crystallogr* **64**, 125-132.

10.   Ploom T, Haussmann C, Hof P, Steinbacher S, Bacher A, Richardson J & Huber R (1999) Crystal structure of 7,8-dihydroneopterin triphosphate epimerase. *Structure* **7**, 509-516.

11.   Bricogne G, Blanc E, Brandl M, Flensburg C, Keller P, Paciorek W, Roversi P, Sharff A, Vonrhein C & Womack TO (2011) BUSTER version 2.8.0. In. Global Phasing Ltd, Cambridge, UK.

12.   Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501.

13.   Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB, 3rd, Snoeyink J, Richardson JS, et al. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* **35**, W375-383.

14.   Schuck P (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys J* **78**, 1606-1619.

15.   Laue TM, Shah, B.D., Ridgeway, T.M., Peletier, S.L. (1992) Computer-aided interpretation of analytical sedimentation data for proteins. In *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding SE RA, Horton JC, editors., ed, pp. 99-125. Society of Chemistry, Cambridge.

16.   Solovyova AS, Nöllmann M, Mitchell TJ & Byron O (2004) The solution structure and oligomerization behavior of two bacterial toxins: pneumolysin and perfringolysin O. *Biophys J* **87**, 540-552.

17.   Svergun DI (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Crystallogr* **25**, 495-503.

18.   Franke D & Svergun DI (2009) DAMMIF, a program for rapid *ab initio* shape determination in small-angle scattering. *J Appl Crystallogr* **42**, 342-346.

19.   Volkov VV & Svergun DI (2003) Uniqueness of *ab initio* shape determination in small-angle scattering. *J Appl Crystallogr* **36**, 860-864.

20.   Kozin MB & Svergun DI (2001) Automated matching of high- and low-resolution structural models. *J Appl Crystallogr* **34**, 33-41.

21.   Bond CS & Schüttelkopf AW (2009) ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments. *Acta Crystallogr D Biol Crystallogr* **65**, 510-512.

22.   DeLano (2009) The PYMOL molecular graphics system. In. DeLano Scientific, Palo Alto, CA.

23.   Barton GJ (1993) ALSCRIPT a tool to format multiple sequence alignments. *Protein Engineering* **6**, 37-44.

24.   Krissinel E & Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-797.

25.   Haussmann C, Rohdich F, Schmidt E, Bacher A & Richter G (1998) Biosynthesis of pteridines in *Escherichia coli*. Structural and mechanistic similarity of dihydroneopterin-triphosphate epimerase and dihydroneopterin aldolase. *J Biol Chem* **273**, 17418-17424.

26.   Rai N, Nöllmann M, Spotorno B, Tassara G, Byron O & Rocco M (2005) SOMO(SOlution MOdeler): Differences between X-ray- and NMR-derived bead models suggest a role for side chain flexibility in protein hydrodynamics. *Structure* **13**, 723-734.
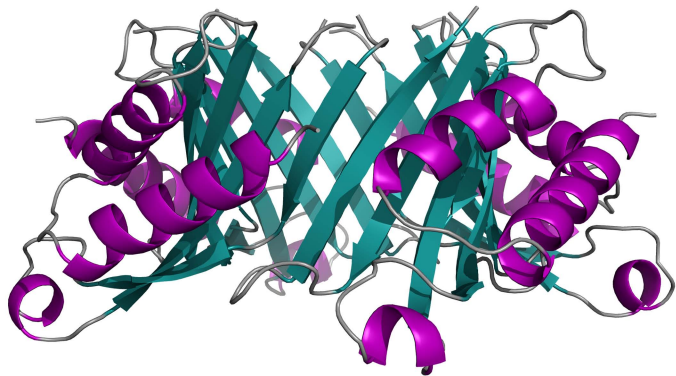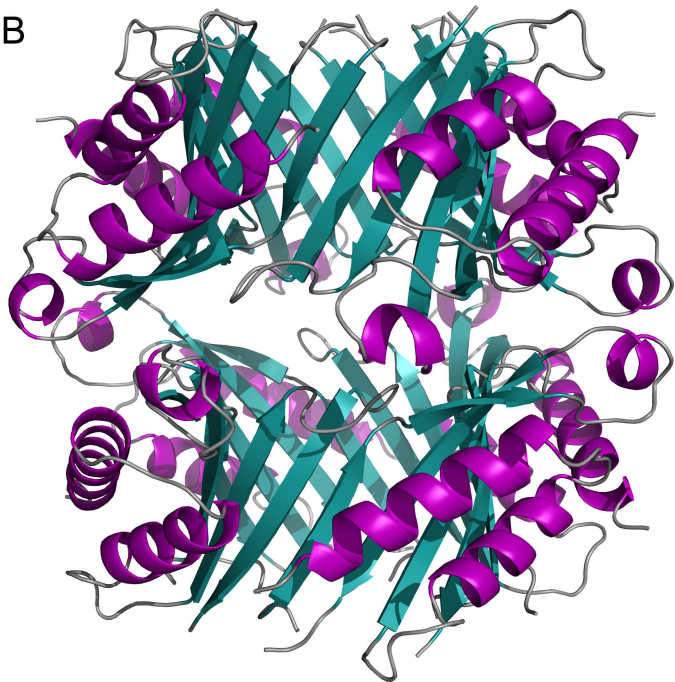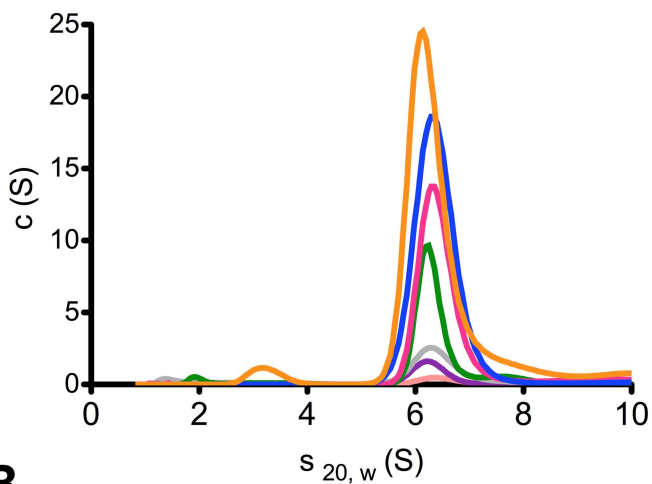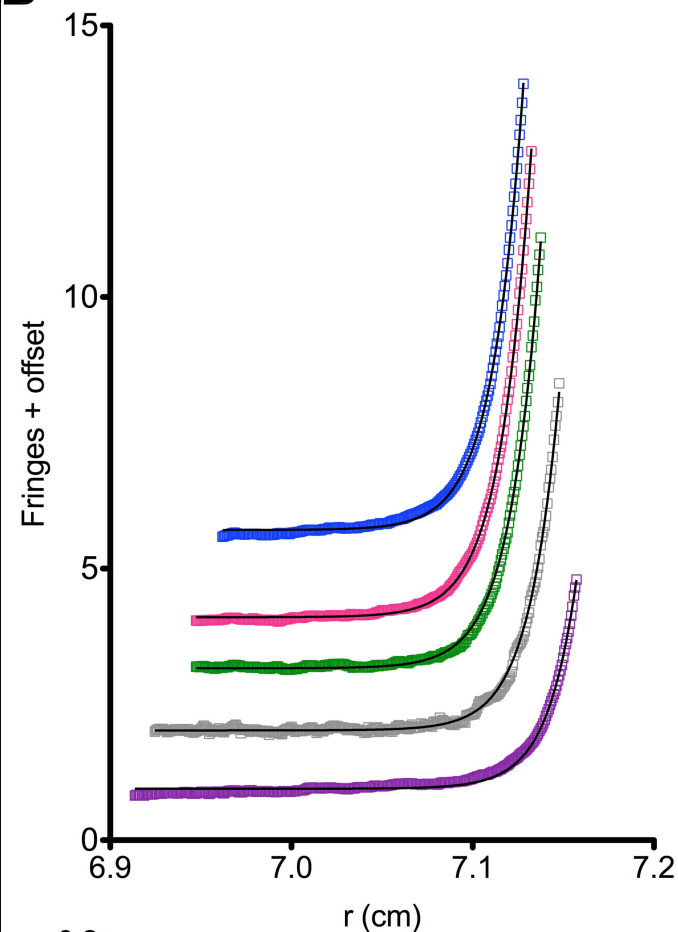
A

α1, α2, α3, β1, β2, β3, β4, N, C

B

paFolX
ecFolX

```
           1                    10              20                    30
                                              β1                      α1
paFolX  M P R L E P G M A R I R V K D L R L R T F I G I K E E E I L
ecFolX  . . . M A Q P A A I I K N L R L R T F I G I K E E E I N

              β2              40              50              60          α2
        N K Q D V L I N L T I L Y P A A D A V E V N D I E H A L N Y R T I T K
        N R Q D I V I N V T I H Y P A D K A R T S E D I N D A L N Y R T V T K

              70                  80              α3          90              100
        A I I R H V E E N R F A L L E R M T Q E I L D L V M E N P A V R V Y A E
        N I I Q H V E N N R F L L E K L T Q D V L D I A R E H H W V T Y A E
                        ▲           ▲
                               110                   120
              β3                              β4
        V E V D K P H A L R F A E S V S I T L A G H R
        V E I D K L H A L R Y A D S V S M T L S W Q R
```
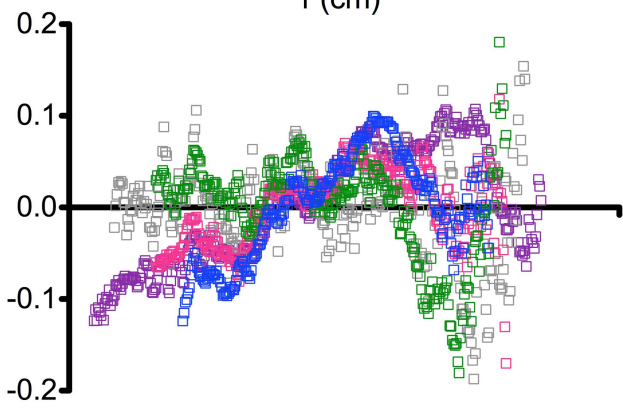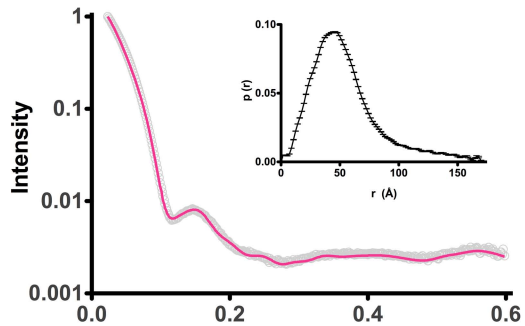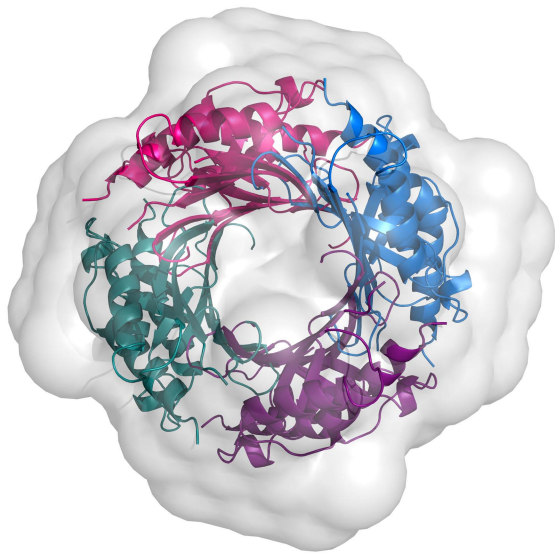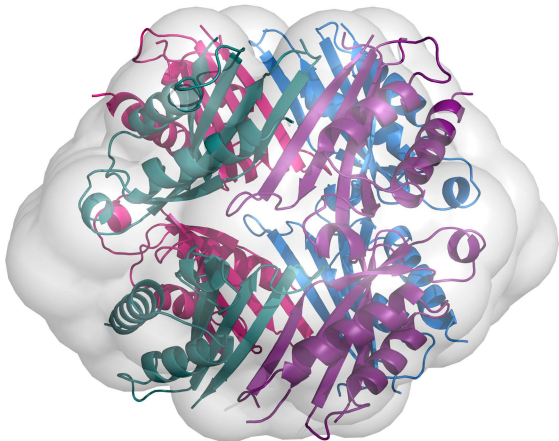
Supplementary Table 1: Hydrogen bonding involved in the interfaces of tetrameric FolX, as determined by PISA.

Subunit A	Distance (Å)	Subunit B
Glu113 (N)	3.80	Asp15 (O)
Glu113 (N)	2.75	Lys14 (O)
Ser 114 (N)	3.63	Lys14 (O)
Ser116 (N)	3.06	Arg12 (O)
Thr118 (N)	2.82	Arg10 (O)
Ala120 (N)	3.11	Met8 (O)
Thr118 (O)	3.22	Arg10 (N)
Ser116 (O)	3.03	Arg12 (N)
Ser114 (O)	3.02	Lys14 (N)
Phe111 (O)	3.18	Arg17 (N)

Supplementary Table 2: Hydrogen bonding and salt bridges involved in the interfaces of octameric FolX, as determined by PISA.

Tetramer A	Distance (Å)	Tetramer B
Hydrogen bonds		
Arg19 (NH2)	2.58	Arg17 (O)
Arg19 (NH1)	3.42	Glu72 (OE2)
Arg17 (O)	2.58	Arg19 (NH2)
Glu72 (OE2)	3.42	Arg219 (NH1)
Salt bridges		
Arg19 (NH1)	3.42	Glu72 (OE2)
Glu72 (OE2)	3.42	Arg219 (NH1)