# wwPDB X-ray Structure Validation Summary Report (i)

Jun 11, 2024 – 05:38 PM EDT

| | | |
|---|---|---|
| PDB ID | : | 2FSF |
| Title | : | Escherichia coli SecA, the preprotein translocase dimeric ATPase |
| Authors | : | Papanikolau, Y.; Petratos, K.; Economou, A. |
| Deposited on | : | 2006-01-23 |
| Resolution | : | 2.00 Å(reported) |

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

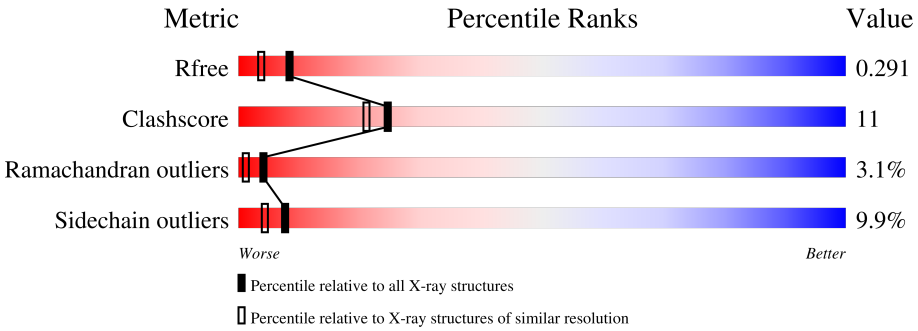
| | | |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Xtriage (Phenix) | : | 1.20.1 |
| EDS | : | 2.36.2 |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| Refmac | : | 5.8.0158 |
| CCP4 | : | 7.0.044 (Gargrove) |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | 2.36.2 |

# 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.




| Metric | Whole archive (#Entries) | Similar resolution (#Entries, resolution range(Å)) |
|---|---|---|
| $R_{free}$ | 130704 | 8085 (2.00-2.00) |
| Clashscore | 141614 | 9178 (2.00-2.00) |
| Ramachandran outliers | 138981 | 9054 (2.00-2.00) |
| Sidechain outliers | 138945 | 9053 (2.00-2.00) |

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

| Mol | Chain | Length | Quality of chain |
|---|---|---|---|
| 1 | A | 853 | 61% 15% • • 20% |
| 1 | B | 853 | 55% 21% 7% • 15% |

## 2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 11637 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Preprotein translocase secA subunit.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | A | 686 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 5462 | 3426 | 966 | 1044 | 26 | | | |
| 1 | B | 723 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 5741 | 3600 | 1009 | 1104 | 28 | | | |

- Molecule 2 is water.

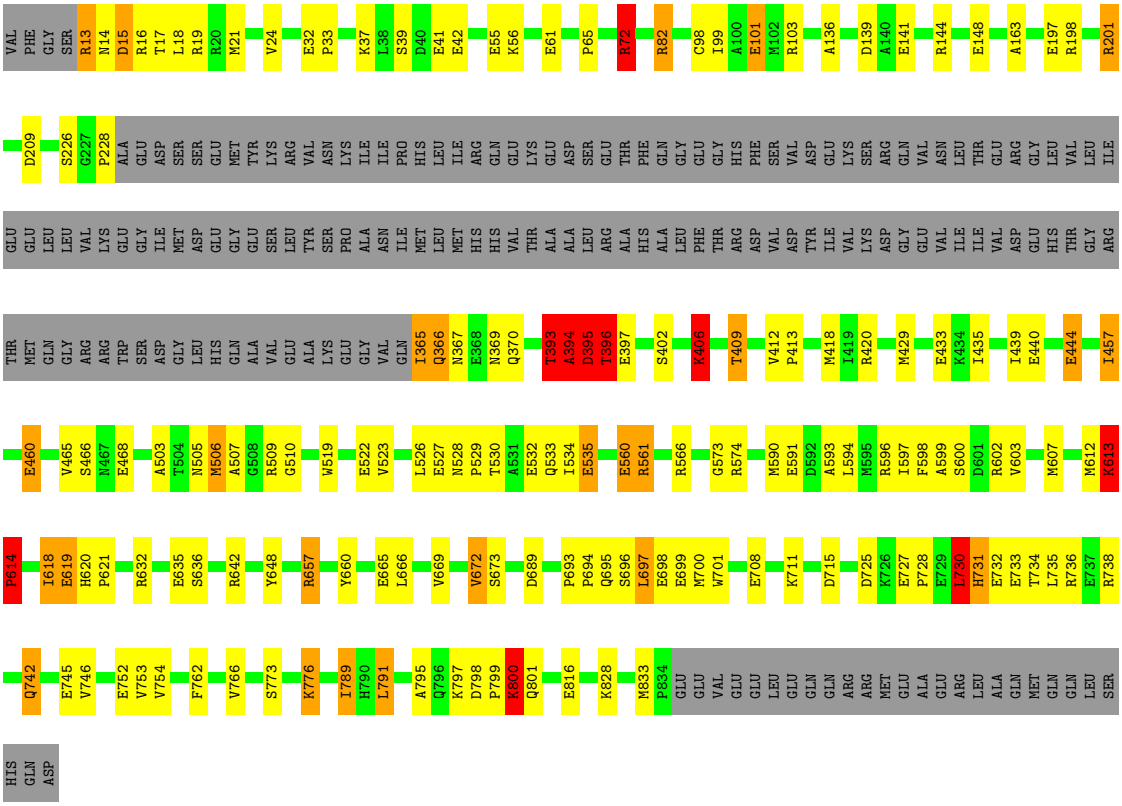
| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|
| 2 | A | 177 | Total | O | 0 | 0 |
| | | | 177 | 177 | | |
| 2 | B | 257 | Total | O | 0 | 0 |
| | | | 257 | 257 | | |

# 3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

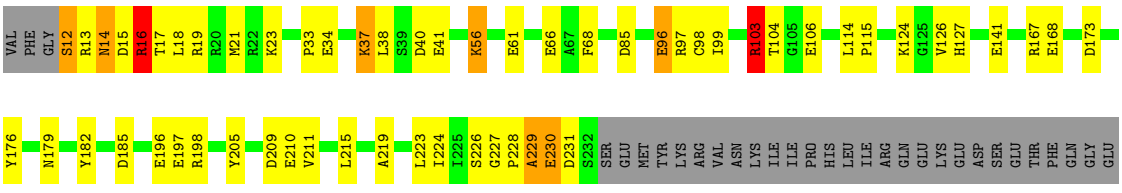
- Molecule 1: Preprotein translocase secA subunit

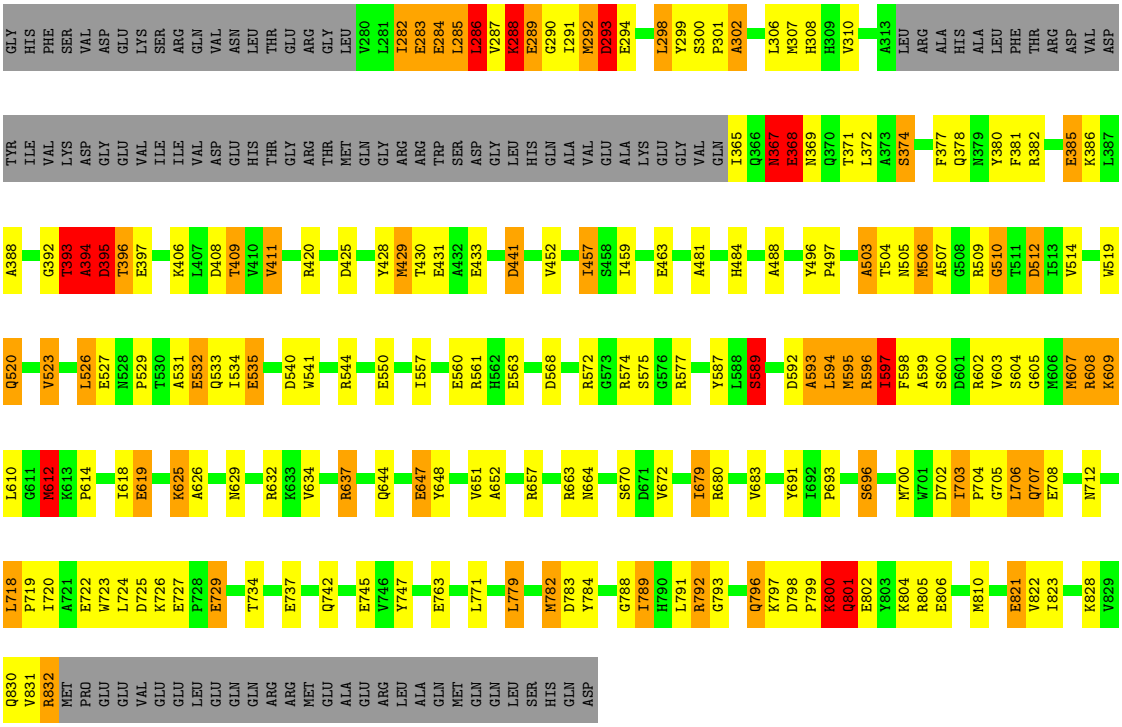
Chain A:

- Molecule 1: Preprotein translocase secA subunit

Chain B:


GLY HIS PHE SER VAL ASP GLU LYS SER ARG GLN VAL ASN LEU THR GLU ARG GLY LEU V280 L281 I282 E283 E284 L285 L286 V287 K288 E289 G290 I291 M292 D293 E294 L298 Y299 S300 P301 A302 L306 M307 H308 H309 V310 A313 LEU ARG ALA HIS ALA LEU PHE THR ARG ASP VAL ASP
TYR ILE VAL LYS ASP GLY GLU VAL ILE ILE VAL ASP GLU HIS THR GLY ARG ARG THR MET GLN GLY ARG ARG TRP SER ASP GLY LEU HIS GLN ALA VAL GLU ALA LYS GLU GLY VAL GLN I365 Q366 N367 E368 N369 Q370 T371 L372 A373 S374 F377 Q378 N379 Y380 F381 R382 E385 K386 L387
A388 G392 T393 A394 D395 T396 E397 K406 L407 D408 T409 V410 V411 R420 D425 Y428 M429 T430 E431 A432 E433 D441 V452 I457 S458 I459 E463 A481 H484 A488 Y496 P497 A503 T504 N505 M506 A507 G508 R509 G510 T511 D512 I513 V514 W519
Q520 V523 L526 E527 N528 P529 T530 A531 E532 Q533 I534 E535 D540 W541 R544 E550 I557 E560 R561 H562 E563 D568 R572 G573 R574 S575 G576 R577 Y587 L588 S589 D592 A593 L594 M595 R596 I597 F598 A599 S600 D601 R602 V603 S604 G605 M606 M607 R608 K609
L610 G611 K612 K613 P614 I618 E619 K625 A626 N629 R632 K633 V634 R637 Q644 E647 Y648 V651 A652 R657 R663 N664 S670 D671 V672 I679 R680 V683 Y691 I692 P693 S696 M700 W701 D702 I703 P704 G705 L706 Q707 E708 N712
L718 P719 I720 A721 E722 W723 L724 D725 K726 E727 P728 E729 T734 E737 Q742 E745 V746 Y747 E763 L771 L779 M782 D783 Y784 G788 I789 H790 L791 R792 G793 Q796 K797 D798 P799 K800 Q801 E802 Y803 K804 R805 E806 M810 E821 V822 I823 K828 V829
Q830 V831 R832 MET PRO GLU GLU VAL GLU GLU LEU GLU GLN GLN ARG ARG MET GLU ALA GLU ARG LEU ALA GLN MET GLN GLN LEU SER HIS GLN ASP

# 4 Data and refinement statistics (i)

| Property | Value | Source |
|---|---|---|
| Space group | P 1 21 1 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 75.03Å 90.17Å 163.05Å<br>90.00° 100.48° 90.00° | Depositor |
| Resolution (Å) | 19.98 – 2.00<br>19.98 – 2.00 | Depositor<br>EDS |
| % Data completeness<br>(in resolution range) | 97.1 (19.98-2.00)<br>97.1 (19.98-2.00) | Depositor<br>EDS |
| $R_{merge}$ | 0.04 | Depositor |
| $R_{sym}$ | (Not available) | Depositor |
| $< I/\sigma(I) >$ [1] | 1.82 (at 2.01Å) | Xtriage |
| Refinement program | REFMAC 5.2.0005 | Depositor |
| R, $R_{free}$ | 0.213 , 0.261<br>0.260 , 0.291 | Depositor<br>DCC |
| $R_{free}$ test set | 7010 reflections (5.01%) | wwPDB-VP |
| Wilson B-factor (Å$^2$) | 35.8 | Xtriage |
| Anisotropy | 0.109 | Xtriage |
| Bulk solvent $k_{sol}$(e/Å$^3$), $B_{sol}$(Å$^2$) | 0.39 , 64.3 | EDS |
| L-test for twinning[2] | $< \|L\| > = 0.50$, $< L^2 > = 0.33$ | Xtriage |
| Estimated twinning fraction | 0.015 for h,-k,-h-l | Xtriage |
| $F_o$,$F_c$ correlation | 0.93 | EDS |
| Total number of atoms | 11637 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 57.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.93% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

# 5 Model quality

## 5.1 Standard geometry

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

| Mol | Chain | Bond lengths | | Bond angles | |
|---|---|---|---|---|---|
| | | RMSZ | #$\|Z\|$ >5 | RMSZ | #$\|Z\|$ >5 |
| 1 | A | 1.34 | 35/5552 (0.6%) | 1.10 | 19/7491 (0.3%) |
| 1 | B | 1.49 | 61/5833 (1.0%) | 1.20 | 36/7870 (0.5%) |
| All | All | 1.42 | 96/11385 (0.8%) | 1.15 | 55/15361 (0.4%) |

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

| Mol | Chain | #Chirality outliers | #Planarity outliers |
|---|---|---|---|
| 1 | A | 0 | 4 |
| 1 | B | 0 | 9 |
| All | All | 0 | 13 |

The worst 5 of 96 bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|---|---|---|---|---|---|---|---|
| 1 | A | 98 | CYS | CB-SG | -13.49 | 1.59 | 1.82 |
| 1 | A | 698 | GLU | CD-OE2 | 13.10 | 1.40 | 1.25 |
| 1 | A | 698 | GLU | CD-OE1 | 10.85 | 1.37 | 1.25 |
| 1 | A | 560 | GLU | CG-CD | 9.58 | 1.66 | 1.51 |
| 1 | A | 697 | LEU | C-N | 9.56 | 1.56 | 1.34 |

The worst 5 of 55 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 1 | B | 167 | ARG | NE-CZ-NH1 | 9.57 | 125.09 | 120.30 |
| 1 | B | 425 | ASP | CB-CG-OD2 | -9.12 | 110.09 | 118.30 |
| 1 | B | 425 | ASP | CB-CG-OD1 | 9.05 | 126.45 | 118.30 |
| 1 | A | 72 | ARG | NE-CZ-NH2 | 8.68 | 124.64 | 120.30 |
| 1 | B | 561 | ARG | NE-CZ-NH1 | 8.38 | 124.49 | 120.30 |

There are no chirality outliers.

5 of 13 planarity outliers are listed below:

| Mol | Chain | Res | Type | Group |
|---|---|---|---|---|
| 1 | A | 13 | ARG | Peptide |
| 1 | A | 393 | THR | Peptide |
| 1 | A | 395 | ASP | Peptide |
| 1 | A | 697 | LEU | Mainchain |
| 1 | B | 229 | ALA | Peptide |

## 5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|---|---|---|---|---|---|---|
| 1 | A | 5462 | 0 | 5460 | 90 | 0 |
| 1 | B | 5741 | 0 | 5731 | 155 | 1 |
| 2 | A | 177 | 0 | 0 | 12 | 0 |
| 2 | B | 257 | 0 | 0 | 30 | 0 |
| All | All | 11637 | 0 | 11191 | 245 | 1 |

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 11.

The worst 5 of 245 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

| Atom-1 | Atom-2 | Interatomic distance (Å) | Clash overlap (Å) |
|---|---|---|---|
| 1:B:96:GLU:HG3 | 2:B:1033:HOH:O | 1.46 | 1.13 |
| 1:A:799:PRO:HA | 2:A:1026:HOH:O | 1.50 | 1.08 |
| 1:B:800:LYS:O | 1:B:801:GLN:HB2 | 1.61 | 1.01 |
| 1:B:103:ARG:HG3 | 1:B:103:ARG:HH11 | 1.26 | 0.97 |
| 1:A:799:PRO:C | 1:A:800:LYS:O | 1.95 | 0.93 |

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

| Atom-1 | Atom-2 | Interatomic distance (Å) | Clash overlap (Å) |
|---|---|---|---|
| 1:B:600:SER:OG | 1:B:830:GLN:O[1_655] | 1.97 | 0.23 |

## 5.3 Torsion angles (i)

### 5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

| Mol | Chain | Analysed | Favoured | Allowed | Outliers | Percentiles |
|---|---|---|---|---|---|---|
| 1 | A | 682/853 (80%) | 635 (93%) | 31 (4%) | 16 (2%) | 6 2 |
| 1 | B | 717/853 (84%) | 650 (91%) | 39 (5%) | 28 (4%) | 3 1 |
| All | All | 1399/1706 (82%) | 1285 (92%) | 70 (5%) | 44 (3%) | 4 1 |

5 of 44 Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | A | 394 | ALA |
| 1 | A | 396 | THR |
| 1 | A | 507 | ALA |
| 1 | A | 599 | ALA |
| 1 | A | 614 | PRO |

### 5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

| Mol | Chain | Analysed | Rotameric | Outliers | Percentiles |
|---|---|---|---|---|---|
| 1 | A | 582/728 (80%) | 532 (91%) | 50 (9%) | 10 6 |
| 1 | B | 613/728 (84%) | 545 (89%) | 68 (11%) | 6 3 |
| All | All | 1195/1456 (82%) | 1077 (90%) | 118 (10%) | 8 4 |

5 of 118 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | B | 230 | GLU |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | B | 783 | ASP |
| 1 | B | 368 | GLU |
| 1 | B | 779 | LEU |
| 1 | B | 696 | SER |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | B | 664 | ASN |
| 1 | B | 761 | HIS |
| 1 | B | 742 | GLN |
| 1 | B | 528 | ASN |
| 1 | B | 644 | GLN |

#### 5.3.3 RNA

There are no RNA molecules in this entry.

### 5.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

### 5.5 Carbohydrates

There are no monosaccharides in this entry.

### 5.6 Ligand geometry

There are no ligands in this entry.

### 5.7 Other polymers

There are no such residues in this entry.

### 5.8 Polymer linkage issues

There are no chain breaks in this entry.

# 6 Fit of model and data

## 6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

## 6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

## 6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

## 6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

## 6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.