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

Jun 12, 2024 – 07:36 AM EDT

| | | |
|---|---|---|
| PDB ID | : | 1SKY |
| Title | : | CRYSTAL STRUCTURE OF THE NUCLEOTIDE FREE ALPHA3BETA3 SUB-COMPLEX OF F1-ATPASE FROM THE THERMOPHILIC BACILLUS PS3 |
| Authors | : | Shirakihara, Y.; Leslie, A.G.W.; Abrahams, J.P.; Walker, J.E.; Ueda, T.; Sekimoto, Y.; Kambara, M.; Saika, K.; Kagawa, Y.; Yoshida, M. |
| Deposited on | : | 1997-02-26 |
| Resolution | : | 3.20 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:

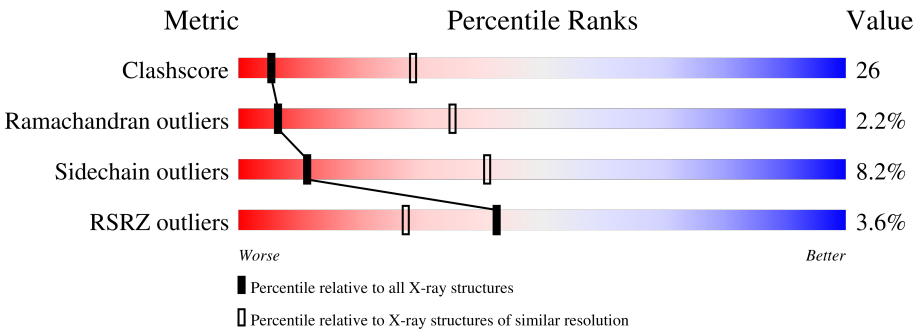
| | | |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Mogul | : | 2022.3.0, CSD as543be (2022) |
| Xtriage (Phenix) | : | 1.20.1 |
| EDS | : | 2.36.2 |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| Refmac | : | 5.8.0158 |
| CCP4 | : | 7.0.044 (Gargrove) |
| 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 3.20 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)) |
|---|---|---|
| Clashscore | 141614 | 1253 (3.20-3.20) |
| Ramachandran outliers | 138981 | 1234 (3.20-3.20) |
| Sidechain outliers | 138945 | 1233 (3.20-3.20) |
| RSRZ outliers | 127900 | 1095 (3.20-3.20) |

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% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

| Mol | Chain | Length | Quality of chain |
|---|---|---|---|
| 1 | B | 502 | 5%; 49%; 43%; • • 5% |
| 2 | E | 473 | 2%; 53%; 40%; 6% • |

# 2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 7299 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.

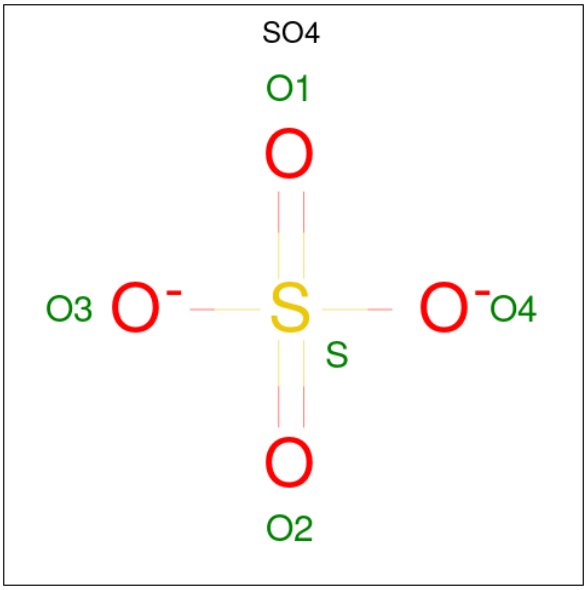
- Molecule 1 is a protein called F1-ATPASE.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | | |
| 1 | B | 479 | 3660 | 2312 | 640 | 697 | 11 | 0 | 0 | 0 |

- Molecule 2 is a protein called F1-ATPASE.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | | |
| 2 | E | 470 | 3629 | 2289 | 630 | 696 | 14 | 0 | 0 | 0 |

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: $O_4S$).



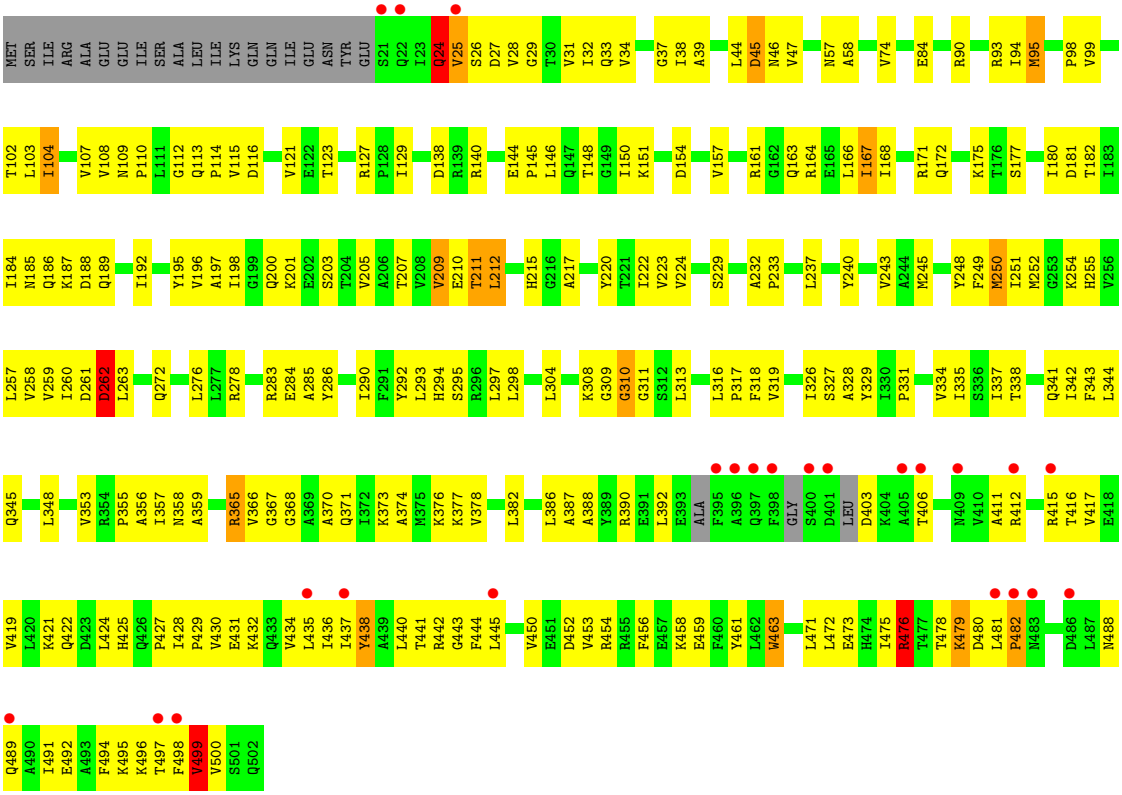

| Mol | Chain | Residues | Atoms | | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|---|
| | | | Total | O | S | | |
| 3 | B | 1 | 5 | 4 | 1 | 0 | 0 |
| 3 | E | 1 | 5 | 4 | 1 | 0 | 0 |

# 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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). 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.

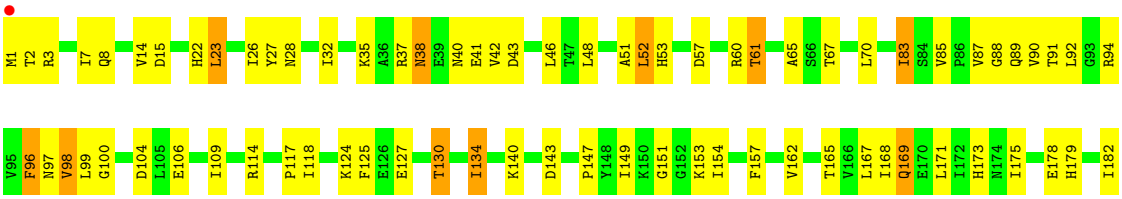
- Molecule 1: F1-ATPASE

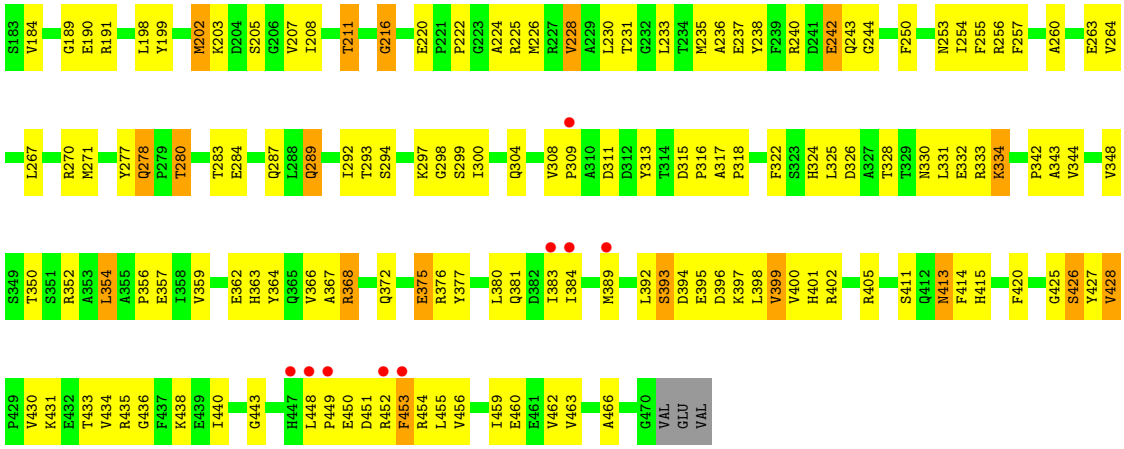
Chain B:




- Molecule 2: F1-ATPASE

Chain E:




S183 V184 G189 E190 R191 L198 Y199 M202 K203 D204 S205 G206 V207 I208 T211 G216 E220 P221 P222 G223 A224 R225 M226 R227 V228 A229 L230 T231 G232 L233 T234 M235 A236 E237 Y238 F239 R240 D241 E242 Q243 G244 F250 N253 I254 F255 R256 F257 A260 E263 V264
L267 R270 M271 Y277 Q278 P279 T280 T283 E284 Q287 L288 Q289 I292 T293 S294 K297 G298 S299 I300 Q304 V308 P309 A310 D311 D312 Y313 T314 D315 P316 A317 P318 F322 S323 H324 L325 D326 A327 T328 T329 N330 L331 E332 R333 K334 P342 A343 V344 V348
S349 T350 S351 R352 A353 L354 A355 P356 E357 I358 V359 E362 H363 Y364 Q365 V366 A367 R368 Q372 E375 R376 Y377 L380 Q381 D382 I383 I384 M389 L392 S393 D394 E395 D396 K397 L398 V399 V400 H401 R402 R405 S411 Q412 N413 F414 H415 F420 G425 S426 Y427 V428
P429 V430 K431 E432 T433 V434 R435 G436 F437 K438 E439 I440 G443 H447 L448 P449 E450 D451 R452 F453 R454 L455 V456 I459 E460 E461 V462 V463 A466 G470 VAL GLU VAL

# 4 Data and refinement statistics (i)

| Property | Value | Source |
|---|---|---|
| Space group | P 21 3 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 159.50Å 159.50Å 159.50Å<br>90.00° 90.00° 90.00° | Depositor |
| Resolution (Å) | 6.00 – 3.20<br>48.09 – 3.00 | Depositor<br>EDS |
| % Data completeness<br>(in resolution range) | 99.1 (6.00-3.20)<br>98.6 (48.09-3.00) | Depositor<br>EDS |
| $R_{merge}$ | 0.07 | Depositor |
| $R_{sym}$ | (Not available) | Depositor |
| $< I/\sigma(I) >$ [1] | 1.71 (at 3.01Å) | Xtriage |
| Refinement program | X-PLOR | Depositor |
| R, $R_{free}$ | 0.222 , 0.299<br>0.218 , (Not available) | Depositor<br>DCC |
| $R_{free}$ test set | No test flags present. | wwPDB-VP |
| Wilson B-factor (Å$^2$) | 76.0 | Xtriage |
| Anisotropy | 0.000 | Xtriage |
| Bulk solvent $k_{sol}$(e/Å$^3$), $B_{sol}$(Å$^2$) | 0.29 , 77.0 | EDS |
| L-test for twinning[2] | $< \|L\| > = 0.51$, $< L^2 > = 0.35$ | Xtriage |
| Estimated twinning fraction | 0.020 for l,-k,h | Xtriage |
| $F_o$,$F_c$ correlation | 0.92 | EDS |
| Total number of atoms | 7299 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 67.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.01% 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 (i)

## 5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: SO4

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 | B | 0.54 | 0/3716 | 0.76 | 2/5035 (0.0%) |
| 2 | E | 0.54 | 0/3692 | 0.74 | 2/4998 (0.0%) |
| All | All | 0.54 | 0/7408 | 0.75 | 4/10033 (0.0%) |

There are no bond length outliers.

All (4) bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 1 | B | 365 | ARG | NE-CZ-NH2 | 7.45 | 124.02 | 120.30 |
| 2 | E | 216 | GLY | N-CA-C | -5.29 | 99.87 | 113.10 |
| 1 | B | 262 | ASP | N-CA-C | 5.23 | 125.12 | 111.00 |
| 2 | E | 98 | VAL | CB-CA-C | -5.19 | 101.55 | 111.40 |

There are no chirality outliers.

There are no planarity outliers.

## 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 | B | 3660 | 0 | 3718 | 203 | 0 |
| 2 | E | 3629 | 0 | 3644 | 191 | 0 |
| 3 | B | 5 | 0 | 0 | 0 | 0 |
| 3 | E | 5 | 0 | 0 | 0 | 0 |
| All | All | 7299 | 0 | 7362 | 384 | 0 |

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 26.

The worst 5 of 384 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:167:ILE:HG22 | 1:B:342:ILE:HB | 1.40 | 1.02 |
| 2:E:240:ARG:HG3 | 2:E:299:SER:N | 1.90 | 0.85 |
| 1:B:475:ILE:HG12 | 1:B:481:LEU:HD21 | 1.59 | 0.84 |
| 2:E:254:ILE:O | 2:E:257:PHE:HB3 | 1.79 | 0.82 |
| 2:E:178:GLU:HG3 | 2:E:427:TYR:CE2 | 2.16 | 0.80 |

There are no symmetry-related clashes.

## 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 | B | 471/502 (94%) | 397 (84%) | 60 (13%) | 14 (3%) | 4 28 |
| 2 | E | 468/473 (99%) | 407 (87%) | 54 (12%) | 7 (2%) | 10 44 |
| All | All | 939/975 (96%) | 804 (86%) | 114 (12%) | 21 (2%) | 6 35 |

5 of 21 Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | B | 24 | GLN |
| 1 | B | 310 | GLY |
| 1 | B | 411 | ALA |
| 1 | B | 499 | VAL |
| 2 | E | 394 | ASP |

### 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 | B | 391/410 (95%) | 359 (92%) | 32 (8%) | 11 | 41 |
| 2 | E | 387/390 (99%) | 355 (92%) | 32 (8%) | 11 | 40 |
| All | All | 778/800 (97%) | 714 (92%) | 64 (8%) | 11 | 41 |

5 of 64 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 2 | E | 399 | VAL |
| 2 | E | 420 | PHE |
| 1 | B | 386 | LEU |
| 1 | B | 382 | LEU |
| 2 | E | 426 | SER |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 2 | E | 278 | GLN |
| 2 | E | 372 | GLN |
| 2 | E | 423 | GLN |
| 2 | E | 407 | GLN |
| 2 | E | 330 | ASN |

### 5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry (i)

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

| Mol | Type | Chain | Res | Link | Bond lengths | | | Bond angles | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Counts | RMSZ | $\#\|Z\| > 2$ | Counts | RMSZ | $\#\|Z\| > 2$ |
| 3 | SO4 | E | 474 | - | 4,4,4 | 0.56 | 0 | 6,6,6 | 0.37 | 0 |
| 3 | SO4 | B | 503 | - | 4,4,4 | 0.51 | 0 | 6,6,6 | 0.81 | 0 |

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers (i)

There are no such residues in this entry.

## 5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

# 6 Fit of model and data (i)

## 6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, $95^{th}$ percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

| Mol | Chain | Analysed | <RSRZ> | #RSRZ>2 | | | OWAB(Å$^2$) | Q<0.9 |
|---|---|---|---|---|---|---|---|---|
| 1 | B | 479/502 (95%) | 0.03 | 24 (5%) | 28 | 16 | 22, 54, 155, 249 | 0 |
| 2 | E | 470/473 (99%) | -0.15 | 10 (2%) | 63 | 49 | 26, 51, 121, 192 | 0 |
| All | All | 949/975 (97%) | -0.06 | 34 (3%) | 42 | 27 | 22, 53, 152, 249 | 0 |

The worst 5 of 34 RSRZ outliers are listed below:

| Mol | Chain | Res | Type | RSRZ |
|---|---|---|---|---|
| 1 | B | 401 | ASP | 17.3 |
| 1 | B | 397 | GLN | 15.1 |
| 1 | B | 21 | SER | 8.4 |
| 1 | B | 482 | PRO | 7.7 |
| 1 | B | 396 | ALA | 7.7 |

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

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

## 6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

## 6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, $95^{th}$ percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

| Mol | Type | Chain | Res | Atoms | RSCC | RSR | B-factors(Å$^2$) | Q<0.9 |
|---|---|---|---|---|---|---|---|---|
| 3 | SO4 | E | 474 | 5/5 | 0.88 | 0.17 | 135,135,135,136 | 0 |
| 3 | SO4 | B | 503 | 5/5 | 0.95 | 0.23 | 63,63,64,64 | 0 |

## 6.5 Other polymers (i)

There are no such residues in this entry.