# WORLDWIDE PDB
## PROTEIN DATA BANK

# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 25, 2023 – 07:28 PM EDT

| | | |
|---|---|---|
| PDB ID | : | 6BQT |
| Title | : | Complex of 14-3-3 theta with an IRSp53 peptide doubly-phosphorylated at T340 and T360 |
| Authors | : | Kast, D.J.; Dominguez, R. |
| Deposited on | : | 2017-11-28 |
| Resolution | : | 2.80 Å(reported) |

---

# 1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 2.80 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.
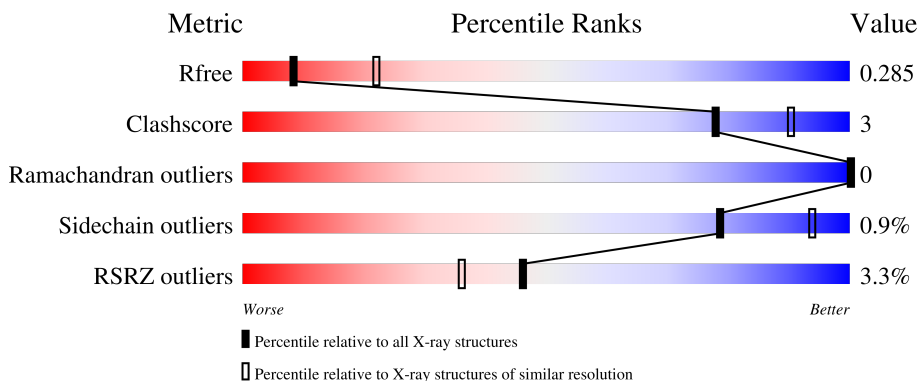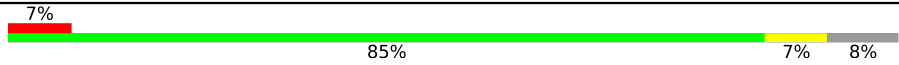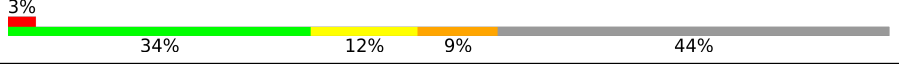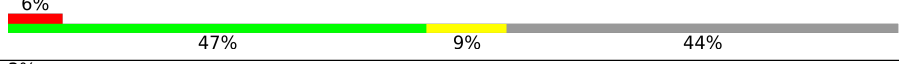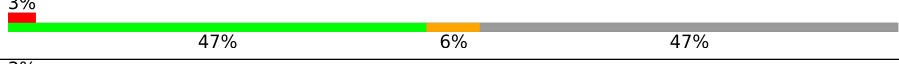


| Metric | Whole archive (#Entries) | Similar resolution (#Entries, resolution range(Å)) |
|---|---|---|
| $R_{free}$ | 130704 | 3140 (2.80-2.80) |
| Clashscore | 141614 | 3569 (2.80-2.80) |
| Ramachandran outliers | 138981 | 3498 (2.80-2.80) |
| Sidechain outliers | 138945 | 3500 (2.80-2.80) |
| RSRZ outliers | 127900 | 3078 (2.80-2.80) |

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 | A | 245 | % 88% 6% 6% |
| 1 | B | 245 | 87% 6% 7% |
| 1 | D | 245 | 3% 89% 5% 6% |
| 1 | E | 245 | 3% 87% 7% 6% |
| 1 | G | 245 | 2% 86% 8% 6% |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Length | Quality of chain |
|-----|-------|--------|------------------|
| 1 | H | 245 | 7% / 85% / 7% / 8% |
| 1 | J | 245 | 86% / 8% / 6% |
| 1 | K | 245 | 7% / 85% / 8% / 7% |
| 2 | C | 32 | 3% / 34% / 12% / 9% / 44% |
| 2 | F | 32 | 6% / 47% / 9% / 44% |
| 2 | I | 32 | 3% / 47% / 6% / 47% |
| 2 | L | 32 | 3% / 34% / 16% / • / 47% |

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

| Mol | Type | Chain | Res | Chirality | Geometry | Clashes | Electron density |
|-----|------|-------|-----|-----------|----------|---------|------------------|
| 9 | PEG | E | 302 | - | - | - | X |

# 2   Entity composition (i)

There are 10 unique types of molecules in this entry. The entry contains 31276 atoms, of which 15611 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 14-3-3 protein theta.

| Mol | Chain | Residues | Atoms | | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | A | 230 | Total 3711 | C 1166 | H 1855 | N 307 | O 373 | S 10 | 0 | 2 | 0 |
| 1 | B | 227 | Total 3679 | C 1156 | H 1839 | N 307 | O 367 | S 10 | 0 | 3 | 0 |
| 1 | D | 230 | Total 3708 | C 1163 | H 1856 | N 310 | O 369 | S 10 | 0 | 1 | 0 |
| 1 | E | 230 | Total 3702 | C 1162 | H 1850 | N 309 | O 371 | S 10 | 0 | 2 | 0 |
| 1 | G | 230 | Total 3746 | C 1174 | H 1873 | N 314 | O 375 | S 10 | 0 | 5 | 0 |
| 1 | H | 225 | Total 3614 | C 1136 | H 1809 | N 301 | O 358 | S 10 | 0 | 0 | 0 |
| 1 | J | 230 | Total 3687 | C 1158 | H 1843 | N 307 | O 369 | S 10 | 0 | 0 | 0 |
| 1 | K | 227 | Total 3662 | C 1152 | H 1828 | N 306 | O 368 | S 8 | 0 | 1 | 0 |

- Molecule 2 is a protein called Insulin receptor substrate protein of 53 kDa, peptide (IRSp53).

| Mol | Chain | Residues | Atoms | | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | C | 18 | Total 301 | C 90 | H 147 | N 28 | O 34 | P 2 | 0 | 0 | 0 |
| 2 | F | 18 | Total 301 | C 90 | H 147 | N 28 | O 34 | P 2 | 0 | 0 | 0 |
| 2 | I | 17 | Total 290 | C 87 | H 142 | N 27 | O 32 | P 2 | 0 | 0 | 0 |
| 2 | L | 17 | Total 290 | C 87 | H 142 | N 27 | O 32 | P 2 | 0 | 0 | 0 |

- Molecule 3 is 2-{2-[2-(2-{2-[2-(2-ETHOXY-ETHOXY)-ETHOXY]-ETHOXY}-ETHOXY)-ETHOXY]-ETHOXY}-ETHANOL (three-letter code: PE4) (formula: $C_{16}H_{34}O_8$).
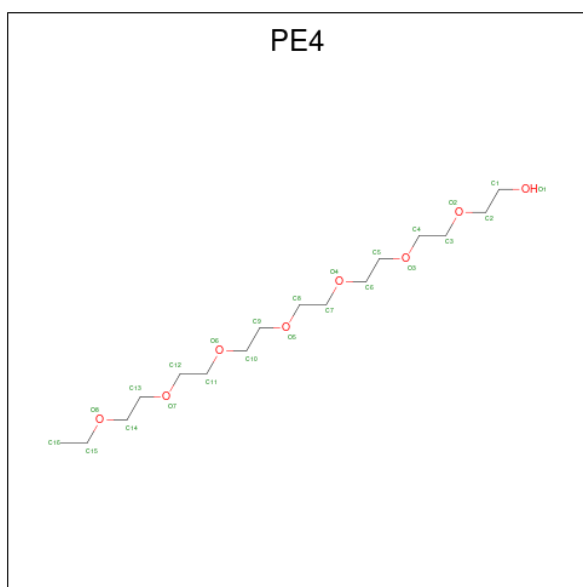
| Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---|---|---------|---------|
| 3 | A | 1 | Total 58 | C 16 | H 34 | O 8 | 0 | 0 |

- Molecule 4 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula: $C_{10}H_{22}O_6$).
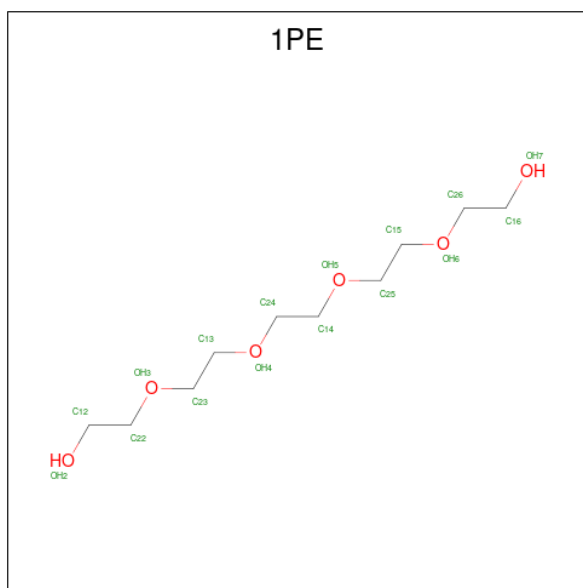


| Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---|---|---------|---------|
| 4 | A | 1 | Total 37 | C 10 | H 21 | O 6 | 0 | 0 |
| 4 | A | 1 | Total 37 | C 10 | H 21 | O 6 | 0 | 0 |
| 4 | B | 1 | Total 37 | C 10 | H 21 | O 6 | 0 | 0 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|---|---|
| 4 | G | 1 | Total | C | H | O | 0 | 0 |
| | | | 37 | 10 | 21 | 6 | | |
| 4 | H | 1 | Total | C | H | O | 0 | 0 |
| | | | 37 | 10 | 21 | 6 | | |
| 4 | J | 1 | Total | C | H | O | 0 | 0 |
| | | | 37 | 10 | 21 | 6 | | |
| 4 | J | 1 | Total | C | H | O | 0 | 0 |
| | | | 37 | 10 | 21 | 6 | | |

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).
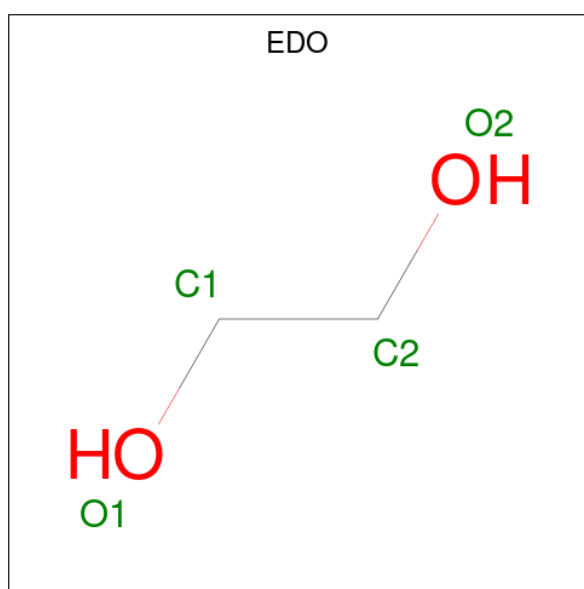


| Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|---|---|
| 5 | B | 1 | Total | C | H | O | 0 | 0 |
| | | | 9 | 2 | 5 | 2 | | |
| 5 | D | 1 | Total | C | H | O | 0 | 0 |
| | | | 9 | 2 | 5 | 2 | | |

- Molecule 6 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).

| Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---|---|---------|---------|
| 6 | B | 1 | Total 30 | C 8 | H 17 | O 5 | 0 | 0 |
| 6 | D | 1 | Total 30 | C 8 | H 17 | O 5 | 0 | 0 |
| 6 | E | 1 | Total 30 | C 8 | H 17 | O 5 | 0 | 0 |

- Molecule 7 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).
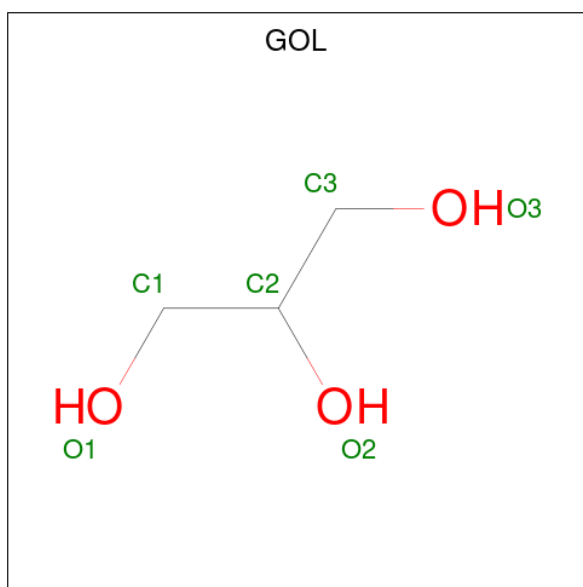


| Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---|---|---------|---------|
| 7 | B | 1 | Total 14 | C 3 | H 8 | O 3 | 0 | 0 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---|---|---------|---------|
| | | | Total | C | H | O | | |
| 7 | I | 1 | 14 | 3 | 8 | 3 | 0 | 0 |

- Molecule 8 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$).
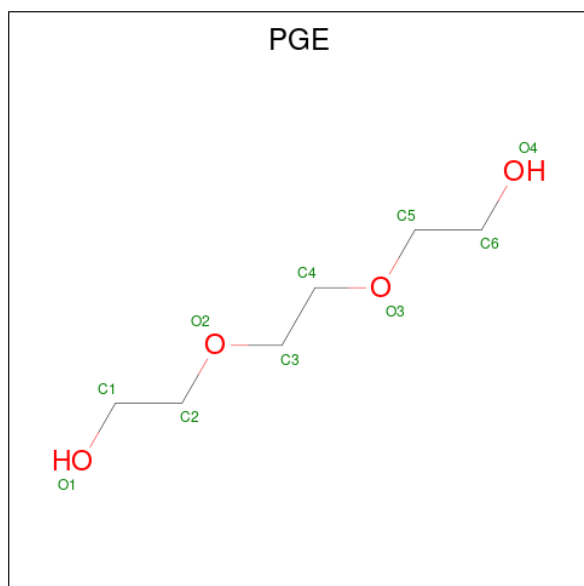


| Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---|---|---------|---------|
| | | | Total | C | H | O | | |
| 8 | E | 1 | 23 | 6 | 13 | 4 | 0 | 0 |

- Molecule 9 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).
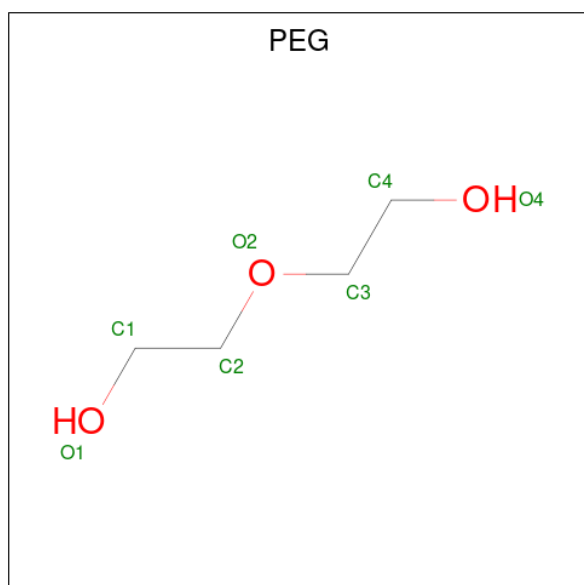
| Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---|---|---------|---------|
| 9 | E | 1 | Total 16 | C 4 | H 9 | O 3 | 0 | 0 |

- Molecule 10 is water.

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---------|---------|
| 10 | A | 17 | Total 17 | O 17 | 0 | 0 |
| 10 | B | 21 | Total 21 | O 21 | 0 | 0 |
| 10 | C | 5 | Total 5 | O 5 | 0 | 0 |
| 10 | D | 10 | Total 10 | O 10 | 0 | 0 |
| 10 | E | 12 | Total 12 | O 12 | 0 | 0 |
| 10 | F | 8 | Total 8 | O 8 | 0 | 0 |
| 10 | G | 7 | Total 7 | O 7 | 0 | 0 |
| 10 | H | 4 | Total 4 | O 4 | 0 | 0 |
| 10 | I | 3 | Total 3 | O 3 | 0 | 0 |
| 10 | J | 6 | Total 6 | O 6 | 0 | 0 |

# 3  Residue-property plots ⓘ

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: 14-3-3 protein theta
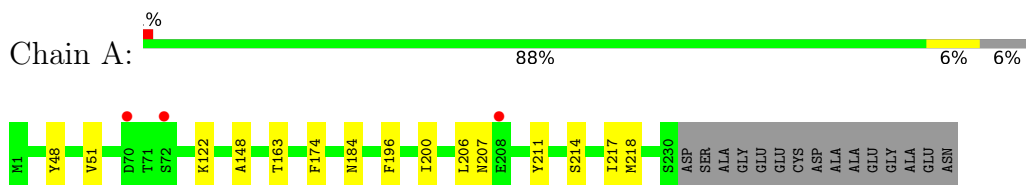
Chain A:

- Molecule 1: 14-3-3 protein theta
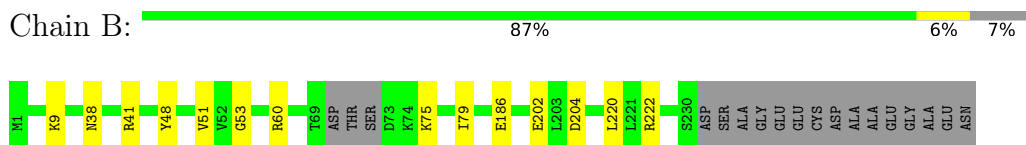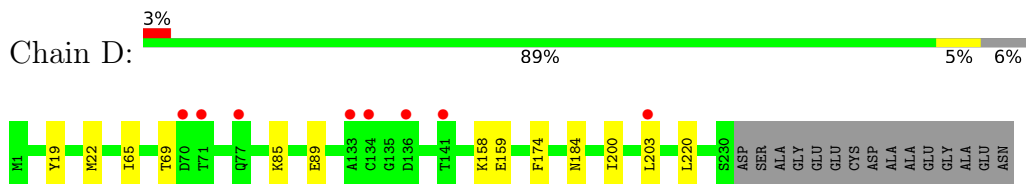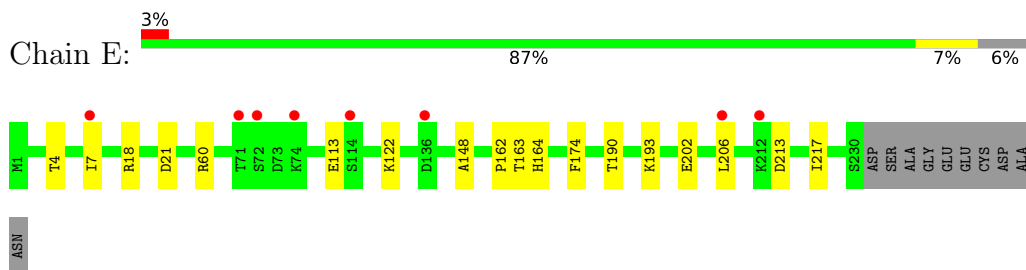
Chain B:

- Molecule 1: 14-3-3 protein theta

Chain D:

- Molecule 1: 14-3-3 protein theta

Chain E:

- Molecule 1: 14-3-3 protein theta

Chain G:

ASN

- Molecule 1: 14-3-3 protein theta

7%

Chain H:

85% 7% 8%

M1 E2 K3 T4 E5 L6 I7 R18 D21 E40 K49 R56 V61 I65 E66 Q67 K68 T69 D70 T71 SER D73 K74 K75 L78 R83 E89 T110 D124 L129 V132 A133 C134 G135 D136 D143 Y149 Q150 E151 F174 T205 LEU

ASN GLU ASP S210 S230 ASP SER ALA GLY GLU GLU CYS ASP ALA ALA GLU GLY ALA GLU ASN
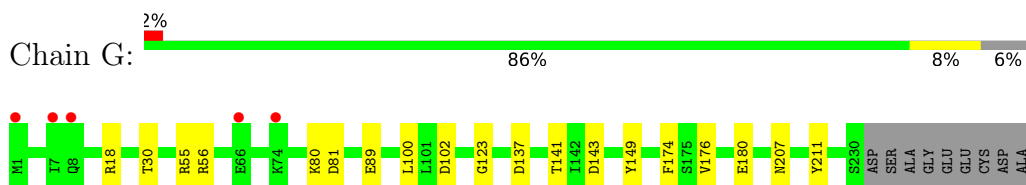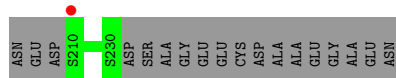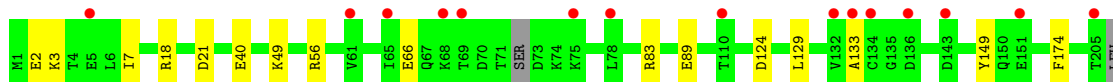
- Molecule 1: 14-3-3 protein theta

Chain J:

86% 8% 6%

M1 L12 N38 E39 Y48 V51 R55 R56 V86 E89 L90 R91 S92 I93 D124 V132 Y149 G169 D209 I217 R222 L225 T229 S230 ASP SER ALA GLY GLU GLU CYS ASP ALA ALA GLU GLY ALA GLU ASN
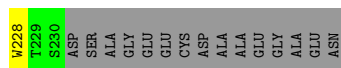
- Molecule 1: 14-3-3 protein theta

7%

Chain K:

85% 8% 7%

MET E2 E5 R18 N38 R41 K49 I65 K68 T69 D70 T71 S72 D73 Q77 L78 I79 K80 T96 N111 S114 F126 A133 CYS GLY D136 D137 T141 I142 D143 F174 T194 A195 T205 L206 N207 Y211 D223 T226 L227

W228 T229 S230 ASP SER ALA GLY GLU GLU CYS ASP ALA ALA GLU GLY ALA GLU ASN

- Molecule 2: Insulin receptor substrate protein of 53 kDa, peptide (IRSp53)

3%

Chain C:

34% 12% 9% 44%

ASP SER Y337 S338 N339 T340 R344 K345 SER VAL THR PRO LYS ASN SER TYR ALA THR THR E357 N358 K359 T360 L361 S365 SER

- Molecule 2: Insulin receptor substrate protein of 53 kDa, peptide (IRSp53)

6%

Chain F:

47% 9% 44%

ASP SER Y337 T340 K345 SER VAL THR PRO LYS ASN SER TYR ALA THR THR E357 T360 S365 SER

- Molecule 2: Insulin receptor substrate protein of 53 kDa, peptide (IRSp53)

3%
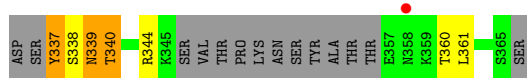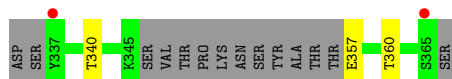
Chain I:

47% 6% 47%
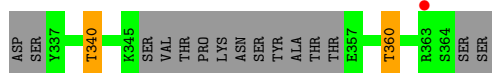
• Molecule 2: Insulin receptor substrate protein of 53 kDa, peptide (IRSp53)
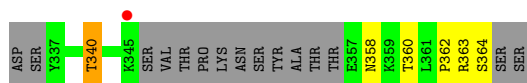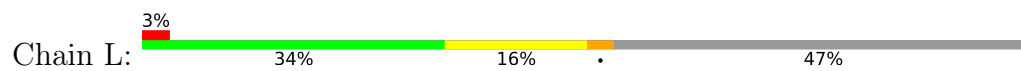
# 4 Data and refinement statistics ⓘ

| Property | Value | Source |
|:---:|:---:|:---:|
| Space group | P 1 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 61.10Å    69.57Å    150.47Å<br>99.16°    92.98°    115.98° | Depositor |
| Resolution (Å) | 42.98  –  2.80<br>49.05  –  2.80 | Depositor<br>EDS |
| % Data completeness<br>(in resolution range) | 96.0 (42.98-2.80)<br>96.0 (49.05-2.80) | Depositor<br>EDS |
| $R_{merge}$ | 0.14 | Depositor |
| $R_{sym}$ | (Not available) | Depositor |
| $< I/\sigma(I) >$ [1] | 2.55 (at 2.81Å) | Xtriage |
| Refinement program | PHENIX 1.12_2829 | Depositor |
| R, $R_{free}$ | 0.234  ,  0.287<br>0.235  ,  0.285 | Depositor<br>DCC |
| $R_{free}$ test set | 2001 reflections (3.88%) | wwPDB-VP |
| Wilson B-factor (Å$^2$) | 30.9 | Xtriage |
| Anisotropy | 0.439 | Xtriage |
| Bulk solvent $k_{sol}$(e/Å$^3$), $B_{sol}$(Å$^2$) | 0.39 , 53.2 | EDS |
| L-test for twinning[2] | $< |L| > = 0.49, < L^2 > = 0.32$ | Xtriage |
| Estimated twinning fraction | 0.004 for -h,-k,h+k+l | Xtriage |
| $F_o$,$F_c$ correlation | 0.88 | EDS |
| Total number of atoms | 31276 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 41.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.47% 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 ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: PEG, 1PE, TPO, PG4, PE4, EDO, PGE, GOL

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 | 0.28 | 0/1887 | 0.41 | 0/2544 |
| 1 | B | 0.27 | 0/1872 | 0.41 | 0/2523 |
| 1 | D | 0.28 | 0/1880 | 0.42 | 0/2534 |
| 1 | E | 0.28 | 0/1885 | 0.43 | 0/2542 |
| 1 | G | 0.48 | 0/1912 | 0.48 | 0/2579 |
| 1 | H | 0.26 | 0/1828 | 0.40 | 0/2461 |
| 1 | J | 0.27 | 0/1869 | 0.40 | 0/2520 |
| 1 | K | 0.40 | 0/1858 | 0.47 | 0/2506 |
| 2 | C | 0.27 | 0/131 | 0.53 | 0/170 |
| 2 | F | 0.25 | 0/131 | 0.42 | 0/170 |
| 2 | I | 0.23 | 0/125 | 0.41 | 0/162 |
| 2 | L | 0.26 | 0/125 | 0.50 | 0/162 |
| All | All | 0.32 | 0/15503 | 0.43 | 0/20873 |

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts ⓘ

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 | 1856 | 1855 | 1855 | 9 | 0 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|
| 1 | B | 1840 | 1839 | 1827 | 12 | 0 |
| 1 | D | 1852 | 1856 | 1856 | 6 | 0 |
| 1 | E | 1852 | 1850 | 1838 | 14 | 0 |
| 1 | G | 1873 | 1873 | 1868 | 13 | 0 |
| 1 | H | 1805 | 1809 | 1809 | 11 | 0 |
| 1 | J | 1844 | 1843 | 1843 | 12 | 0 |
| 1 | K | 1834 | 1828 | 1828 | 13 | 0 |
| 2 | C | 154 | 147 | 147 | 7 | 0 |
| 2 | F | 154 | 147 | 147 | 3 | 0 |
| 2 | I | 148 | 142 | 142 | 3 | 0 |
| 2 | L | 148 | 142 | 142 | 4 | 0 |
| 3 | A | 24 | 34 | 34 | 0 | 0 |
| 4 | A | 32 | 42 | 44 | 0 | 0 |
| 4 | B | 16 | 21 | 22 | 1 | 0 |
| 4 | G | 16 | 21 | 22 | 3 | 0 |
| 4 | H | 16 | 21 | 22 | 0 | 0 |
| 4 | J | 32 | 42 | 44 | 0 | 0 |
| 5 | B | 4 | 5 | 6 | 0 | 0 |
| 5 | D | 4 | 5 | 6 | 0 | 0 |
| 6 | B | 13 | 17 | 18 | 0 | 0 |
| 6 | D | 13 | 17 | 18 | 0 | 0 |
| 6 | E | 13 | 17 | 18 | 0 | 0 |
| 7 | B | 6 | 8 | 8 | 0 | 0 |
| 7 | I | 6 | 8 | 8 | 0 | 0 |
| 8 | E | 10 | 13 | 14 | 0 | 0 |
| 9 | E | 7 | 9 | 10 | 0 | 0 |
| 10 | A | 17 | 0 | 0 | 0 | 0 |
| 10 | B | 21 | 0 | 0 | 1 | 0 |
| 10 | C | 5 | 0 | 0 | 0 | 0 |
| 10 | D | 10 | 0 | 0 | 0 | 0 |
| 10 | E | 12 | 0 | 0 | 0 | 0 |
| 10 | F | 8 | 0 | 0 | 0 | 0 |
| 10 | G | 7 | 0 | 0 | 0 | 0 |
| 10 | H | 4 | 0 | 0 | 0 | 0 |
| 10 | I | 3 | 0 | 0 | 0 | 0 |
| 10 | J | 6 | 0 | 0 | 0 | 0 |
| All | All | 15665 | 15611 | 15596 | 89 | 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 3.

The worst 5 of 89 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:E:162:PRO:HD2 | 1:E:202:GLU:OE1 | 1.61 | 0.99 |
| 1:E:60:ARG:HD2 | 2:F:357:GLU:OE2 | 1.71 | 0.90 |
| 1:E:60:ARG:CD | 2:F:357:GLU:OE2 | 2.25 | 0.85 |
| 1:H:66:GLU:OE1 | 1:H:83:ARG:NE | 2.23 | 0.72 |
| 1:K:137:ASP:O | 1:K:141:THR:HG22 | 1.93 | 0.69 |

There are no symmetry-related clashes.

## 5.3   Torsion angles ⓘ

### 5.3.1   Protein backbone ⓘ

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 | 230/245 (94%) | 226 (98%) | 4 (2%) | 0 | 100 | 100 |
| 1 | B | 226/245 (92%) | 215 (95%) | 11 (5%) | 0 | 100 | 100 |
| 1 | D | 229/245 (94%) | 223 (97%) | 6 (3%) | 0 | 100 | 100 |
| 1 | E | 230/245 (94%) | 225 (98%) | 5 (2%) | 0 | 100 | 100 |
| 1 | G | 233/245 (95%) | 228 (98%) | 5 (2%) | 0 | 100 | 100 |
| 1 | H | 219/245 (89%) | 214 (98%) | 5 (2%) | 0 | 100 | 100 |
| 1 | J | 228/245 (93%) | 222 (97%) | 6 (3%) | 0 | 100 | 100 |
| 1 | K | 224/245 (91%) | 213 (95%) | 11 (5%) | 0 | 100 | 100 |
| 2 | C | 12/32 (38%) | 12 (100%) | 0 | 0 | 100 | 100 |
| 2 | F | 12/32 (38%) | 12 (100%) | 0 | 0 | 100 | 100 |
| 2 | I | 11/32 (34%) | 11 (100%) | 0 | 0 | 100 | 100 |
| 2 | L | 11/32 (34%) | 11 (100%) | 0 | 0 | 100 | 100 |
| All | All | 1865/2088 (89%) | 1812 (97%) | 53 (3%) | 0 | 100 | 100 |

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains ⓘ

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 | 205/212 (97%) | 203 (99%) | 2 (1%) | 76 | 93 |
| 1 | B | 203/212 (96%) | 203 (100%) | 0 | 100 | 100 |
| 1 | D | 204/212 (96%) | 201 (98%) | 3 (2%) | 65 | 89 |
| 1 | E | 205/212 (97%) | 204 (100%) | 1 (0%) | 88 | 96 |
| 1 | G | 208/212 (98%) | 205 (99%) | 3 (1%) | 67 | 90 |
| 1 | H | 198/212 (93%) | 197 (100%) | 1 (0%) | 88 | 96 |
| 1 | J | 203/212 (96%) | 202 (100%) | 1 (0%) | 88 | 96 |
| 1 | K | 202/212 (95%) | 201 (100%) | 1 (0%) | 88 | 96 |
| 2 | C | 16/29 (55%) | 13 (81%) | 3 (19%) | 1 | 5 |
| 2 | F | 16/29 (55%) | 16 (100%) | 0 | 100 | 100 |
| 2 | I | 15/29 (52%) | 15 (100%) | 0 | 100 | 100 |
| 2 | L | 15/29 (52%) | 15 (100%) | 0 | 100 | 100 |
| All | All | 1690/1812 (93%) | 1675 (99%) | 15 (1%) | 78 | 94 |

5 of 15 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | D | 220 | LEU |
| 1 | J | 222 | ARG |
| 1 | E | 174 | PHE |
| 1 | K | 174 | PHE |
| 1 | G | 174 | PHE |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | E | 161 | GLN |
| 2 | F | 358 | ASN |
| 1 | J | 8 | GLN |

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

8 non-standard protein/DNA/RNA residues 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 | #$|Z| > 2$ |
| 2 | TPO | I | 360 | 2 | 8,10,11 | 1.12 | 0 | 10,14,16 | 1.49 | 2 (20%) |
| 2 | TPO | C | 360 | 2 | 8,10,11 | 1.60 | 1 (12%) | 10,14,16 | 1.61 | 1 (10%) |
| 2 | TPO | I | 340 | 2 | 8,10,11 | 1.08 | 0 | 10,14,16 | 1.61 | 1 (10%) |
| 2 | TPO | F | 360 | 2 | 8,10,11 | 1.61 | 1 (12%) | 10,14,16 | 1.51 | 1 (10%) |
| 2 | TPO | C | 340 | 2 | 8,10,11 | 0.92 | 0 | 10,14,16 | 1.82 | 1 (10%) |
| 2 | TPO | F | 340 | 2 | 8,10,11 | 1.15 | 0 | 10,14,16 | 1.61 | 2 (20%) |
| 2 | TPO | L | 340 | 2 | 8,10,11 | 1.04 | 0 | 10,14,16 | 1.56 | 1 (10%) |
| 2 | TPO | L | 360 | 2 | 8,10,11 | 1.11 | 0 | 10,14,16 | 1.56 | 1 (10%) |

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

| Mol | Type | Chain | Res | Link | Chirals | Torsions | Rings |
|---|---|---|---|---|---|---|---|
| 2 | TPO | I | 360 | 2 | - | 1/9/11/13 | - |
| 2 | TPO | C | 360 | 2 | - | 4/9/11/13 | - |
| 2 | TPO | I | 340 | 2 | - | 1/9/11/13 | - |
| 2 | TPO | F | 360 | 2 | - | 3/9/11/13 | - |
| 2 | TPO | C | 340 | 2 | - | 1/9/11/13 | - |
| 2 | TPO | F | 340 | 2 | - | 2/9/11/13 | - |
| 2 | TPO | L | 340 | 2 | - | 3/9/11/13 | - |
| 2 | TPO | L | 360 | 2 | - | 1/9/11/13 | - |

All (2) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|-----|-------|-----|------|-------|------|-------------|----------|
| 2 | F | 360 | TPO | P-O1P | 3.41 | 1.61 | 1.50 |
| 2 | C | 360 | TPO | P-O1P | 3.34 | 1.61 | 1.50 |

The worst 5 of 10 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|-----|-------|-----|------|-----------|-------|----------------|-------------|
| 2 | C | 340 | TPO | P-OG1-CB | -5.23 | 107.40 | 123.21 |
| 2 | C | 360 | TPO | P-OG1-CB | -4.36 | 110.04 | 123.21 |
| 2 | I | 340 | TPO | P-OG1-CB | -4.32 | 110.17 | 123.21 |
| 2 | L | 360 | TPO | P-OG1-CB | -4.13 | 110.75 | 123.21 |
| 2 | L | 340 | TPO | P-OG1-CB | -4.07 | 110.90 | 123.21 |

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

| Mol | Chain | Res | Type | Atoms |
|-----|-------|-----|------|----------------|
| 2 | C | 340 | TPO | O-C-CA-CB |
| 2 | L | 360 | TPO | O-C-CA-CB |
| 2 | F | 360 | TPO | CB-OG1-P-O1P |
| 2 | C | 360 | TPO | CB-OG1-P-O3P |
| 2 | C | 360 | TPO | C-CA-CB-CG2 |

There are no ring outliers.

4 monomers are involved in 5 short contacts:

| Mol | Chain | Res | Type | Clashes | Symm-Clashes |
|-----|-------|-----|------|---------|--------------|
| 2 | I | 360 | TPO | 2 | 0 |
| 2 | I | 340 | TPO | 1 | 0 |
| 2 | C | 340 | TPO | 1 | 0 |
| 2 | L | 340 | TPO | 1 | 0 |

## 5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

17 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 | $\#\|Z\| > 2$ |
| 4 | 1PE | B | 302 | - | 15,15,15 | 0.53 | 0 | 14,14,14 | 0.25 | 0 |
| 4 | 1PE | A | 302 | - | 15,15,15 | 0.53 | 0 | 14,14,14 | 0.27 | 0 |
| 3 | PE4 | A | 301 | - | 23,23,23 | 0.54 | 0 | 22,22,22 | 0.27 | 0 |
| 5 | EDO | D | 301 | - | 3,3,3 | 0.48 | 0 | 2,2,2 | 0.27 | 0 |
| 5 | EDO | B | 301 | - | 3,3,3 | 0.49 | 0 | 2,2,2 | 0.25 | 0 |
| 6 | PG4 | B | 303 | - | 12,12,12 | 0.52 | 0 | 11,11,11 | 0.29 | 0 |
| 6 | PG4 | D | 300 | - | 12,12,12 | 0.55 | 0 | 11,11,11 | 0.28 | 0 |
| 4 | 1PE | J | 301 | - | 15,15,15 | 0.52 | 0 | 14,14,14 | 0.33 | 0 |
| 7 | GOL | I | 401 | - | 5,5,5 | 0.87 | 0 | 5,5,5 | 1.02 | 0 |
| 8 | PGE | E | 301 | - | 9,9,9 | 0.51 | 0 | 8,8,8 | 0.29 | 0 |
| 7 | GOL | B | 304 | - | 5,5,5 | 1.03 | 0 | 5,5,5 | 0.89 | 0 |
| 4 | 1PE | J | 300 | - | 15,15,15 | 0.52 | 0 | 14,14,14 | 0.49 | 0 |
| 9 | PEG | E | 302 | - | 6,6,6 | 0.49 | 0 | 5,5,5 | 0.29 | 0 |
| 4 | 1PE | A | 303 | - | 15,15,15 | 0.51 | 0 | 14,14,14 | 0.42 | 0 |
| 4 | 1PE | H | 300 | - | 15,15,15 | 0.52 | 0 | 14,14,14 | 0.31 | 0 |
| 6 | PG4 | E | 300 | - | 12,12,12 | 0.53 | 0 | 11,11,11 | 0.26 | 0 |
| 4 | 1PE | G | 300 | - | 15,15,15 | 0.53 | 0 | 14,14,14 | 0.36 | 0 |

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

| Mol | Type | Chain | Res | Link | Chirals | Torsions | Rings |
|---|---|---|---|---|---|---|---|
| 4 | 1PE | B | 302 | - | - | 4/13/13/13 | - |
| 4 | 1PE | A | 302 | - | - | 3/13/13/13 | - |
| 3 | PE4 | A | 301 | - | - | 6/21/21/21 | - |
| 5 | EDO | D | 301 | - | - | 0/1/1/1 | - |
| 5 | EDO | B | 301 | - | - | 1/1/1/1 | - |
| 6 | PG4 | B | 303 | - | - | 3/10/10/10 | - |
| 6 | PG4 | D | 300 | - | - | 2/10/10/10 | - |
| 4 | 1PE | J | 301 | - | - | 5/13/13/13 | - |
| 7 | GOL | I | 401 | - | - | 0/4/4/4 | - |

*Continued on next page...*

*Continued from previous page...*

| Mol | Type | Chain | Res | Link | Chirals | Torsions | Rings |
|-----|------|-------|-----|------|---------|----------|-------|
| 8 | PGE | E | 301 | - | - | 2/7/7/7 | - |
| 7 | GOL | B | 304 | - | - | 0/4/4/4 | - |
| 4 | 1PE | J | 300 | - | - | 7/13/13/13 | - |
| 9 | PEG | E | 302 | - | - | 0/4/4/4 | - |
| 4 | 1PE | A | 303 | - | - | 8/13/13/13 | - |
| 4 | 1PE | H | 300 | - | - | 2/13/13/13 | - |
| 6 | PG4 | E | 300 | - | - | 2/10/10/10 | - |
| 4 | 1PE | G | 300 | - | - | 6/13/13/13 | - |

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 51 torsion outliers are listed below:

| Mol | Chain | Res | Type | Atoms |
|-----|-------|-----|------|-------|
| 8 | E | 301 | PGE | O2-C3-C4-O3 |
| 3 | A | 301 | PE4 | O6-C11-C12-O7 |
| 4 | H | 300 | 1PE | OH7-C16-C26-OH6 |
| 3 | A | 301 | PE4 | O6-C10-C9-O5 |
| 4 | A | 302 | 1PE | OH7-C16-C26-OH6 |

There are no ring outliers.

2 monomers are involved in 4 short contacts:

| Mol | Chain | Res | Type | Clashes | Symm-Clashes |
|-----|-------|-----|------|---------|--------------|
| 4 | B | 302 | 1PE | 1 | 0 |
| 4 | G | 300 | 1PE | 3 | 0 |

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

| Ligand PE4 A 301 | |
|---|---|
|   Bond lengths |   Bond angles |
|   Torsions |   Rings |
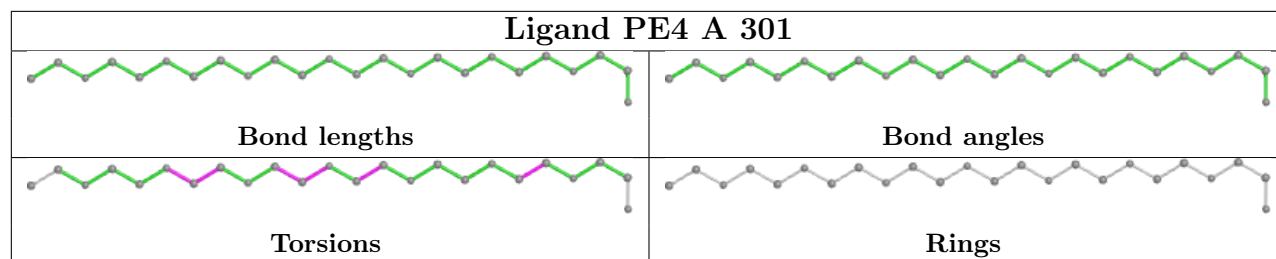
## 5.7 Other polymers ⓘ

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 ⓘ

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 | A | 230/245 (93%) | 0.05 | 3 (1%) | 77 | 72 | 9, 25, 62, 79 | 0 |
| 1 | B | 227/245 (92%) | -0.02 | 0 | 100 | 100 | 9, 23, 55, 75 | 0 |
| 1 | D | 230/245 (93%) | 0.16 | 8 (3%) | 44 | 34 | 10, 26, 66, 88 | 0 |
| 1 | E | 230/245 (93%) | 0.27 | 8 (3%) | 44 | 34 | 14, 32, 72, 97 | 0 |
| 1 | G | 230/245 (93%) | 0.17 | 5 (2%) | 62 | 52 | 13, 35, 67, 91 | 0 |
| 1 | H | 225/245 (91%) | 0.49 | 16 (7%) | 16 | 9 | 24, 46, 77, 89 | 0 |
| 1 | J | 230/245 (93%) | 0.25 | 1 (0%) | 92 | 91 | 14, 37, 69, 88 | 0 |
| 1 | K | 227/245 (92%) | 0.53 | 16 (7%) | 16 | 9 | 22, 49, 79, 98 | 0 |
| 2 | C | 16/32 (50%) | 0.70 | 1 (6%) | 20 | 12 | 23, 46, 62, 62 | 0 |
| 2 | F | 16/32 (50%) | 0.75 | 2 (12%) | 3 | 2 | 24, 49, 71, 76 | 0 |
| 2 | I | 15/32 (46%) | 0.65 | 1 (6%) | 17 | 10 | 23, 42, 76, 84 | 0 |
| 2 | L | 15/32 (46%) | 0.68 | 1 (6%) | 17 | 10 | 32, 45, 67, 72 | 0 |
| All | All | 1891/2088 (90%) | 0.25 | 62 (3%) | 46 | 36 | 9, 34, 71, 98 | 0 |

The worst 5 of 62 RSRZ outliers are listed below:

| Mol | Chain | Res | Type | RSRZ |
|---|---|---|---|---|
| 1 | D | 134 | CYS | 4.6 |
| 1 | K | 70 | ASP | 4.2 |
| 1 | K | 69 | THR | 4.1 |
| 1 | H | 136 | ASP | 4.1 |
| 1 | H | 210 | SER | 4.0 |

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

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 |
|-----|------|-------|-----|-------|------|------|------------------|-------|
| 2 | TPO | L | 340 | 11/12 | 0.95 | 0.18 | 26,26,31,31 | 0 |
| 2 | TPO | L | 360 | 11/12 | 0.95 | 0.14 | 35,36,43,43 | 0 |
| 2 | TPO | F | 340 | 11/12 | 0.96 | 0.18 | 20,22,25,25 | 0 |
| 2 | TPO | I | 360 | 11/12 | 0.96 | 0.18 | 32,33,41,41 | 0 |
| 2 | TPO | I | 340 | 11/12 | 0.97 | 0.19 | 22,23,28,28 | 0 |
| 2 | TPO | C | 360 | 11/12 | 0.97 | 0.19 | 24,26,30,30 | 0 |
| 2 | TPO | C | 340 | 11/12 | 0.97 | 0.15 | 16,23,28,28 | 0 |
| 2 | TPO | F | 360 | 11/12 | 0.97 | 0.18 | 32,33,39,40 | 0 |

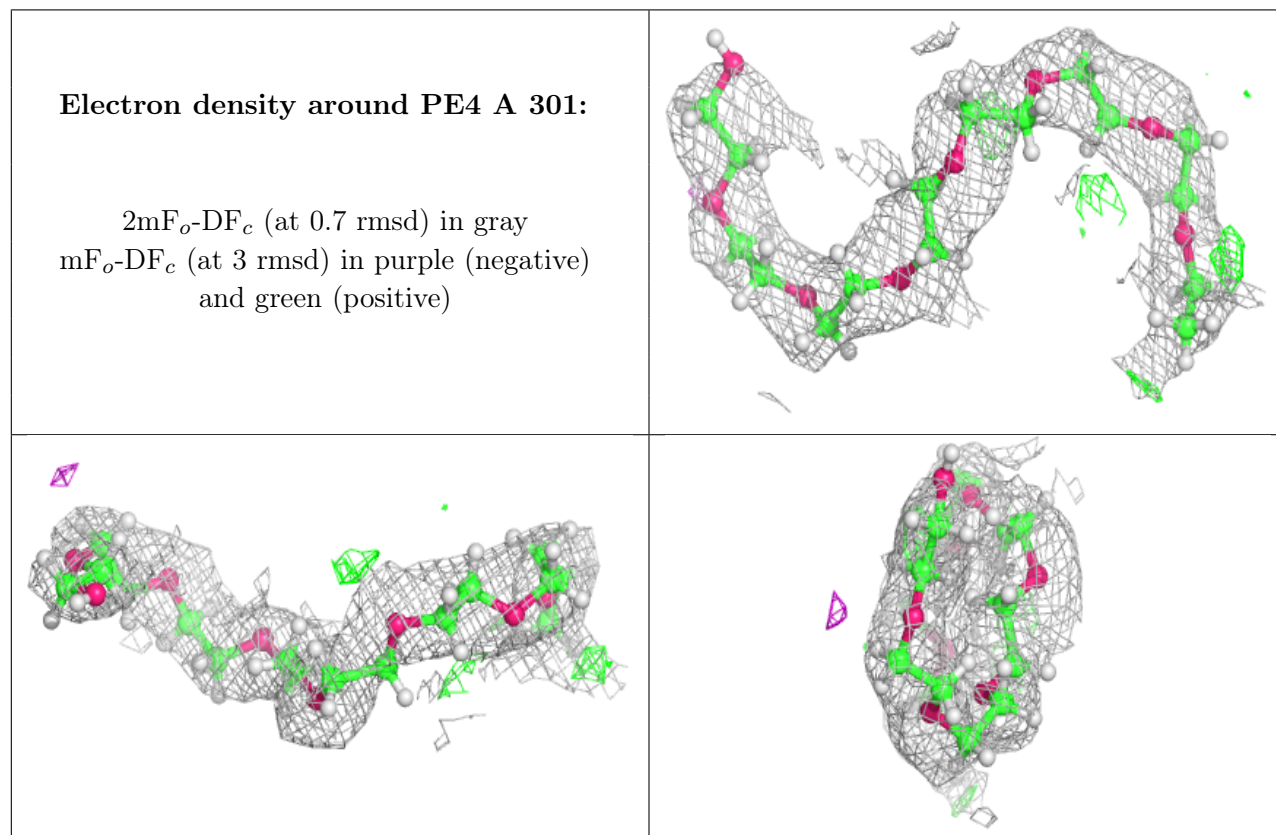## 6.3 Carbohydrates (i)

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 |
|-----|------|-------|-----|-------|------|------|------------------|-------|
| 5 | EDO | B | 301 | 4/4 | 0.68 | 0.25 | 33,41,49,49 | 0 |
| 4 | 1PE | H | 300 | 16/16 | 0.71 | 0.31 | 59,74,91,96 | 0 |
| 9 | PEG | E | 302 | 7/7 | 0.74 | 0.44 | 48,58,67,72 | 0 |
| 7 | GOL | I | 401 | 6/6 | 0.76 | 0.28 | 41,56,67,68 | 0 |
| 4 | 1PE | J | 300 | 16/16 | 0.76 | 0.27 | 33,51,72,73 | 0 |
| 4 | 1PE | A | 303 | 16/16 | 0.77 | 0.29 | 44,62,86,87 | 0 |
| 4 | 1PE | G | 300 | 16/16 | 0.80 | 0.27 | 30,60,86,88 | 0 |
| 8 | PGE | E | 301 | 10/10 | 0.81 | 0.22 | 48,57,67,68 | 0 |
| 4 | 1PE | J | 301 | 16/16 | 0.81 | 0.21 | 49,62,76,78 | 0 |
| 7 | GOL | B | 304 | 6/6 | 0.84 | 0.23 | 39,51,61,64 | 0 |
| 4 | 1PE | B | 302 | 16/16 | 0.85 | 0.23 | 19,40,53,53 | 0 |
| 4 | 1PE | A | 302 | 16/16 | 0.85 | 0.24 | 31,45,67,73 | 0 |
| 6 | PG4 | D | 300 | 13/13 | 0.86 | 0.24 | 17,37,47,50 | 0 |
| 6 | PG4 | B | 303 | 13/13 | 0.87 | 0.19 | 35,47,58,59 | 0 |
| 3 | PE4 | A | 301 | 24/24 | 0.87 | 0.20 | 17,46,58,58 | 0 |
| 5 | EDO | D | 301 | 4/4 | 0.92 | 0.23 | 16,19,23,23 | 0 |
| 6 | PG4 | E | 300 | 13/13 | 0.92 | 0.21 | 22,33,49,49 | 0 |

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around PE4 A 301:**

$2mF_o$-$DF_c$ (at 0.7 rmsd) in gray
$mF_o$-$DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)



## 6.5　Other polymers ⓘ

There are no such residues in this entry.