# WORLDWIDE PDB PROTEIN DATA BANK

# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 12:53 AM GMT

PDB ID : 2BWE

Title : The crystal structure of the complex between the UBA and UBL domains of Dsk2

Authors : Lowe, E.D.; Hasan, N.; Trempe, J.-F.; Fonso, L.; Noble, M.E.M.; Endicott, J.A.; Johnson, L.N.; Brown, N.R.

Deposited on : 2005-07-13

Resolution : 3.10 Å(reported)

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

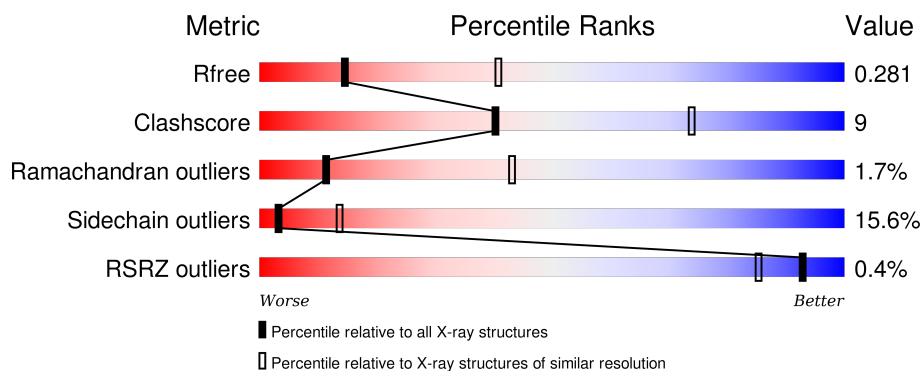| | | |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Mogul | : | 1.7 (RC4), CSD as536be (2015) |
| Xtriage (Phenix) | : | 1.9-1692 |
| EDS | : | rb-20026688 |
| Percentile statistics | : | 20151230.v01 (using entries in the PDB archive December 30th 2015) |
| Refmac | : | 5.8.0135 |
| CCP4 | : | 6.5.0 |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | trunk26865 |

# 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 3.10 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 | Percentile Ranks | Value |
|---|---|---|
| Rfree | | 0.281 |
| Clashscore | | 9 |
| Ramachandran outliers | | 1.7% |
| Sidechain outliers | | 15.6% |
| RSRZ outliers | | 0.4% |

*Worse*                                                                                      *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

| Metric | Whole archive (#Entries) | Similar resolution (#Entries, resolution range(Å)) |
|---|---|---|
| R$_{free}$ | 91344 | 1114 (3.14-3.06) |
| Clashscore | 102246 | 1222 (3.14-3.06) |
| Ramachandran outliers | 100387 | 1174 (3.14-3.06) |
| Sidechain outliers | 100360 | 1174 (3.14-3.06) |
| RSRZ outliers | 91569 | 1119 (3.14-3.06) |

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 on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. 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 | 50 | 70%  18%  •  8% |
| 1 | B | 50 | 68%  18%  10%  • |
| 1 | C | 50 | 64%  22%  10%  • |
| 1 | D | 50 | 66%  20%  6%  •  6% |
| 1 | E | 50 | 68%  16%  8%  8% |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Length | Quality of chain |
|-----|-------|--------|------------------|
| 1 | F | 50 | 2% 72% 10% 8% 10% |
| 1 | G | 50 | 64% 22% 8% |
| 1 | H | 50 | 2% 66% 20% 10% |
| 1 | I | 50 | 72% 16% 10% |
| 1 | J | 50 | 60% 28% 8% |
| 1 | K | 50 | 60% 22% 12% 6% |
| 1 | L | 50 | 68% 12% 10% 8% |
| 1 | M | 50 | 68% 18% 10% |
| 1 | N | 50 | 68% 18% 6% 8% |
| 1 | O | 50 | 68% 14% 8% 8% |
| 1 | P | 50 | 72% 12% 12% |
| 1 | Q | 50 | 66% 22% 6% |
| 1 | R | 50 | 70% 18% 8% |
| 2 | S | 77 | 65% 27% 5% |
| 2 | T | 77 | 66% 26% 5% |
| 2 | U | 77 | 3% 69% 22% 6% |

# 2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 8410 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 DSK2.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | A | 46 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 366 | 224 | 69 | 72 | 1 | | | |
| 1 | B | 48 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 378 | 232 | 71 | 74 | 1 | | | |
| 1 | C | 48 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 382 | 233 | 71 | 77 | 1 | | | |
| 1 | D | 47 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 374 | 230 | 70 | 73 | 1 | | | |
| 1 | E | 46 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 366 | 224 | 69 | 72 | 1 | | | |
| 1 | F | 45 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 358 | 218 | 68 | 71 | 1 | | | |
| 1 | G | 46 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 366 | 224 | 69 | 72 | 1 | | | |
| 1 | H | 45 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 358 | 218 | 68 | 71 | 1 | | | |
| 1 | I | 45 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 358 | 218 | 68 | 71 | 1 | | | |
| 1 | J | 46 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 366 | 224 | 69 | 72 | 1 | | | |
| 1 | K | 47 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 374 | 228 | 70 | 75 | 1 | | | |
| 1 | L | 46 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 366 | 224 | 69 | 72 | 1 | | | |
| 1 | M | 45 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 358 | 218 | 68 | 71 | 1 | | | |
| 1 | N | 46 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 366 | 224 | 69 | 72 | 1 | | | |
| 1 | O | 46 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 366 | 224 | 69 | 72 | 1 | | | |
| 1 | P | 44 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 354 | 216 | 67 | 70 | 1 | | | |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|---|---|---|---|---------|---------|-------|
| 1 | Q | 47 | Total<br>374 | C<br>230 | N<br>70 | O<br>73 | S<br>1 | 0 | 0 | 0 |
| 1 | R | 46 | Total<br>366 | C<br>224 | N<br>69 | O<br>72 | S<br>1 | 0 | 0 | 0 |

- Molecule 2 is a protein called DSK2.

| Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|---|---|---|---------|---------|-------|
| 2 | S | 73 | Total<br>573 | C<br>360 | N<br>102 | O<br>111 | 0 | 0 | 0 |
| 2 | T | 73 | Total<br>573 | C<br>360 | N<br>102 | O<br>111 | 0 | 0 | 0 |
| 2 | U | 72 | Total<br>567 | C<br>357 | N<br>101 | O<br>109 | 0 | 0 | 0 |

- Molecule 3 is water.

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---------|---------|
| 3 | A | 7 | Total<br>7 | O<br>7 | 0 | 0 |
| 3 | B | 5 | Total<br>5 | O<br>5 | 0 | 0 |
| 3 | C | 9 | Total<br>9 | O<br>9 | 0 | 0 |
| 3 | D | 10 | Total<br>10 | O<br>10 | 0 | 0 |
| 3 | E | 5 | Total<br>5 | O<br>5 | 0 | 0 |
| 3 | F | 5 | Total<br>5 | O<br>5 | 0 | 0 |
| 3 | G | 2 | Total<br>2 | O<br>2 | 0 | 0 |
| 3 | H | 3 | Total<br>3 | O<br>3 | 0 | 0 |
| 3 | I | 1 | Total<br>1 | O<br>1 | 0 | 0 |
| 3 | J | 6 | Total<br>6 | O<br>6 | 0 | 0 |
| 3 | K | 8 | Total<br>8 | O<br>8 | 0 | 0 |
| 3 | L | 10 | Total<br>10 | O<br>10 | 0 | 0 |

*Continued from previous page...*

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---------|---------|
| 3 | M | 3 | Total 3 | O 3 | 0 | 0 |
| 3 | N | 1 | Total 1 | O 1 | 0 | 0 |
| 3 | O | 2 | Total 2 | O 2 | 0 | 0 |
| 3 | P | 4 | Total 4 | O 4 | 0 | 0 |
| 3 | Q | 1 | Total 1 | O 1 | 0 | 0 |
| 3 | R | 4 | Total 4 | O 4 | 0 | 0 |
| 3 | S | 9 | Total 9 | O 9 | 0 | 0 |
| 3 | T | 6 | Total 6 | O 6 | 0 | 0 |

# 3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.
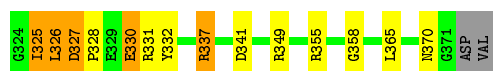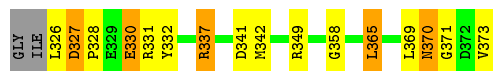
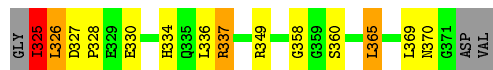- Molecule 1: DSK2

Chain A:



- Molecule 1: DSK2

Chain B:



- Molecule 1: DSK2

Chain C:



- Molecule 1: DSK2

Chain D:



- Molecule 1: DSK2

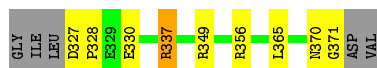Chain E:



- Molecule 1: DSK2

Chain F:

GLY ILE LEU D327 E330 R331 Y332 E333 H334 R337 Q338 L339 R349 G358 L365 N370 G371 ASP VAL

• Molecule 1: DSK2

Chain G:

| 64% | 22% | • • | 8% |

GLY ILE L326 D327 P328 E329 E330 R331 Y332 R337 D341 R349 R356 S357 G358 V361 L365 N370 G371 ASP VAL

• Molecule 1: DSK2

Chain H:

2%
| 66% | 20% | • | 10% |

GLY ILE LEU D327 P328 E329 E330 R331 Y332 R337 D341 R349 G358 L365 L368 L369 N370 G371 ASP VAL

• Molecule 1: DSK2

Chain I:

| 72% | 16% | • | 10% |

GLY ILE LEU D327 P328 E329 E330 R337 R349 R356 L365 N370 G371 ASP VAL

• Molecule 1: DSK2

Chain J:

| 60% | 28% | • • | 8% |

GLY ILE L326 D327 P328 E329 E330 R331 Y332 R337 F344 R349 R355 R356 S357 G358 L365 L369 N370 G371 ASP VAL

• Molecule 1: DSK2

Chain K:

| 60% | 22% | 12% | 6% |

GLY ILE L326 D327 P328 E329 E330 R331 Y332 R337 D341 M342 R349 R356 S357 G358 Q362 L365 L369 N370 G371 D372 VAL

• Molecule 1: DSK2

Chain L:

| 68% | 12% | 10% | • | 8% |

GLY ILE L326 D327 E330 H334 R337 D346 R349 S360 V361 Q362 L365 N370 G371 ASP VAL

• Molecule 1: DSK2

Chain M:

| 68% | 18% | • | 10% |

GLY ILE LEU D327 P328 E329 E330 R331 Y332 R337 D341 R349 R356 L365 N370 G371 ASP VAL

● Molecule 1: DSK2

Chain N:

68%　　18%　　6%　　8%

GLY ILE L326 D327 P328 E329 E330 R331 Y332 R337 R349 R356 S357 G358 L365 N370 G371 ASP VAL

● Molecule 1: DSK2

Chain O:

68%　　14%　　8%　　·　8%

GLY ILE L326 D327 P328 E329 E330 H334 R337 R349 R356 S357 G358 L365 N370 G371 ASP VAL
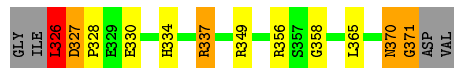
● Molecule 1: DSK2

Chain P:

72%　　12%　　·　12%

GLY ILE LEU D327 E330 H334 R337 D341 R349 G358 L365 N370 GLY ASP VAL

● Molecule 1: DSK2

Chain Q:

66%　　22%　　·　·　6%

GLY I325 L326 D327 E330 R331 Y332 Q335 L336 R337 M342 R349 G358 G359 S360 L365 N370 G371 ASP VAL

● Molecule 1: DSK2

Chain R:

70%　　18%　　·　8%

GLY ILE L326 D327 P328 R331 R337 F344 R349 R355 L365 D366 L369 N370 G371 ASP VAL

● Molecule 2: DSK2

Chain S:

65%　　27%　　·　5%

LEU ASP MET S2 S9 G10 K13 W14 E15 P20 E21 Q26 N35 G36 I37 N41 Q42 R43 Y46 S47 I50 L51 Y60 H66 L70 V71 K72 S73 Q74 PRO

● Molecule 2: DSK2

Chain T:

66%　　26%　　·　5%

LEU ASP MET S2 S9 G10 K13 W14 E15 P20 E21 Q26 N35 G36 I37 P38 V39 A40 N41 Q42 R43 S47 I50 L51 Y60 L70 V71 K72 S73 Q74 PRO

● Molecule 2: DSK2

3%

Chain U:

69%　　22%　　··　6%
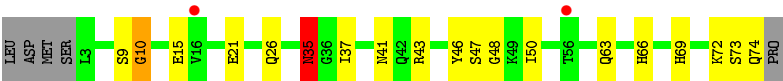
# 4 Data and refinement statistics (i)

| Property | Value | Source |
|---|---|---|
| Space group | P 1 21 1 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 78.36Å    88.85Å    141.50Å<br>90.00°    106.09°    90.00° | Depositor |
| Resolution (Å) | 136.08   –   3.10<br>59.29   –   3.10 | Depositor<br>EDS |
| % Data completeness<br>(in resolution range) | 98.5 (136.08-3.10)<br>98.5 (59.29-3.10) | Depositor<br>EDS |
| $R_{merge}$ | 0.10 | Depositor |
| $R_{sym}$ | (Not available) | Depositor |
| $< I/\sigma(I) >$ [1] | 3.19 (at 3.13Å) | Xtriage |
| Refinement program | REFMAC 5.2.0005 | Depositor |
| R, $R_{free}$ | 0.239   ,   0.267<br>0.281   ,   0.281 | Depositor<br>DCC |
| $R_{free}$ test set | 1707 reflections (5.35%) | DCC |
| Wilson B-factor (Å$^2$) | 66.1 | Xtriage |
| Anisotropy | 0.579 | Xtriage |
| Bulk solvent $k_{sol}$(e/Å$^3$), $B_{sol}$(Å$^2$) | 0.33 , 36.8 | EDS |
| Estimated twinning fraction | 0.459 for h,-k,-h-l | Xtriage |
| L-test for twinning[2] | $< |L| > = 0.48$, $< L^2 > = 0.31$ | Xtriage |
| Outliers | 1 of 33640 reflections (0.003%) | Xtriage |
| $F_o$,$F_c$ correlation | 0.88 | EDS |
| Total number of atoms | 8410 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 82.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.17% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

# 5 Model quality (i)

## 5.1 Standard geometry (i)

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 | 0.79 | 0/371 | 0.87 | 1/497 (0.2%) |
| 1 | B | 0.78 | 0/383 | 0.79 | 0/513 |
| 1 | C | 0.91 | 0/387 | 0.89 | 0/518 |
| 1 | D | 0.88 | 0/379 | 0.89 | 0/508 |
| 1 | E | 0.71 | 0/371 | 0.85 | 0/497 |
| 1 | F | 0.60 | 0/363 | 0.73 | 0/486 |
| 1 | G | 0.78 | 0/371 | 0.89 | 1/497 (0.2%) |
| 1 | H | 0.57 | 0/363 | 0.63 | 0/486 |
| 1 | I | 0.74 | 1/363 (0.3%) | 0.67 | 0/486 |
| 1 | J | 0.80 | 0/371 | 0.81 | 0/497 |
| 1 | K | 0.89 | 0/379 | 0.84 | 0/508 |
| 1 | L | 0.85 | 0/371 | 0.83 | 1/497 (0.2%) |
| 1 | M | 0.79 | 0/363 | 0.78 | 1/486 (0.2%) |
| 1 | N | 0.68 | 0/371 | 0.78 | 1/497 (0.2%) |
| 1 | O | 2.08 | 2/371 (0.5%) | 0.91 | 2/497 (0.4%) |
| 1 | P | 0.63 | 0/359 | 0.67 | 0/481 |
| 1 | Q | 0.72 | 0/379 | 0.80 | 1/508 (0.2%) |
| 1 | R | 0.80 | 0/371 | 0.89 | 2/497 (0.4%) |
| 2 | S | 0.64 | 0/583 | 0.77 | 0/787 |
| 2 | T | 0.63 | 0/583 | 0.76 | 0/787 |
| 2 | U | 0.44 | 0/577 | 0.67 | 0/779 |
| All | All | 0.84 | 3/8429 (0.0%) | 0.80 | 10/11309 (0.1%) |

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 | 1 |
| 1 | D | 0 | 1 |
| 1 | E | 0 | 1 |
| 1 | G | 0 | 1 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | #Chirality outliers | #Planarity outliers |
|-----|-------|---------------------|---------------------|
| 1 | J | 0 | 1 |
| 1 | L | 0 | 2 |
| 1 | O | 0 | 2 |
| 1 | Q | 0 | 2 |
| All | All | 0 | 11 |

All (3) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|-----|-------|-----|------|-------|---|-------------|----------|
| 1 | O | 371 | GLY | C-O | 37.37 | 1.83 | 1.23 |
| 1 | O | 371 | GLY | CA-C | 7.64 | 1.64 | 1.51 |
| 1 | I | 371 | GLY | C-O | 6.76 | 1.34 | 1.23 |

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

| Mol | Chain | Res | Type | Atoms | Z | Observed(°) | Ideal(°) |
|-----|-------|-----|------|-------|---|-------------|----------|
| 1 | O | 371 | GLY | CA-C-O | -10.45 | 101.78 | 120.60 |
| 1 | G | 341 | ASP | CB-CG-OD2 | -7.56 | 111.49 | 118.30 |
| 1 | Q | 326 | LEU | N-CA-C | -6.19 | 94.28 | 111.00 |
| 1 | O | 326 | LEU | CA-CB-CG | 5.75 | 128.53 | 115.30 |
| 1 | M | 356 | ARG | NE-CZ-NH1 | -5.32 | 117.64 | 120.30 |

There are no chirality outliers.

5 of 11 planarity outliers are listed below:

| Mol | Chain | Res | Type | Group |
|-----|-------|-----|------|-------|
| 1 | A | 326 | LEU | Peptide |
| 1 | D | 325 | ILE | Peptide |
| 1 | E | 370 | ASN | Peptide |
| 1 | G | 326 | LEU | Peptide |
| 1 | J | 326 | LEU | 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.

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|
| 1 | A | 366 | 0 | 343 | 6 | 0 |
| 1 | B | 378 | 0 | 357 | 9 | 0 |
| 1 | C | 382 | 0 | 356 | 16 | 0 |
| 1 | D | 374 | 0 | 354 | 11 | 0 |
| 1 | E | 366 | 0 | 343 | 10 | 0 |
| 1 | F | 358 | 0 | 332 | 5 | 0 |
| 1 | G | 366 | 0 | 343 | 7 | 0 |
| 1 | H | 358 | 0 | 332 | 7 | 0 |
| 1 | I | 358 | 0 | 332 | 2 | 0 |
| 1 | J | 366 | 0 | 343 | 8 | 0 |
| 1 | K | 374 | 0 | 347 | 13 | 0 |
| 1 | L | 366 | 0 | 343 | 13 | 0 |
| 1 | M | 358 | 0 | 332 | 4 | 0 |
| 1 | N | 366 | 0 | 343 | 5 | 0 |
| 1 | O | 366 | 0 | 343 | 4 | 0 |
| 1 | P | 354 | 0 | 329 | 4 | 0 |
| 1 | Q | 374 | 0 | 354 | 11 | 0 |
| 1 | R | 366 | 0 | 343 | 6 | 0 |
| 2 | S | 573 | 0 | 570 | 11 | 0 |
| 2 | T | 573 | 0 | 570 | 11 | 0 |
| 2 | U | 567 | 0 | 565 | 10 | 0 |
| 3 | A | 7 | 0 | 0 | 1 | 0 |
| 3 | B | 5 | 0 | 0 | 0 | 0 |
| 3 | C | 9 | 0 | 0 | 0 | 0 |
| 3 | D | 10 | 0 | 0 | 1 | 0 |
| 3 | E | 5 | 0 | 0 | 1 | 0 |
| 3 | F | 5 | 0 | 0 | 1 | 0 |
| 3 | G | 2 | 0 | 0 | 1 | 0 |
| 3 | H | 3 | 0 | 0 | 1 | 0 |
| 3 | I | 1 | 0 | 0 | 0 | 0 |
| 3 | J | 6 | 0 | 0 | 0 | 0 |
| 3 | K | 8 | 0 | 0 | 1 | 0 |
| 3 | L | 10 | 0 | 0 | 1 | 0 |
| 3 | M | 3 | 0 | 0 | 0 | 0 |
| 3 | N | 1 | 0 | 0 | 0 | 0 |
| 3 | O | 2 | 0 | 0 | 0 | 0 |
| 3 | P | 4 | 0 | 0 | 0 | 0 |
| 3 | Q | 1 | 0 | 0 | 0 | 0 |
| 3 | R | 4 | 0 | 0 | 0 | 0 |
| 3 | S | 9 | 0 | 0 | 0 | 0 |
| 3 | T | 6 | 0 | 0 | 0 | 0 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|
| All | All | 8410 | 0 | 7874 | 140 | 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 9.

The worst 5 of 140 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:O:371:GLY:C | 1:O:371:GLY:O | 1.83 | 1.17 |
| 1:E:367:SER:O | 1:E:371:GLY:HA3 | 1.43 | 1.16 |
| 1:L:326:LEU:HA | 1:L:327:ASP:HB2 | 1.37 | 1.04 |
| 1:D:325:ILE:N | 1:D:326:LEU:HB3 | 1.81 | 0.96 |
| 1:C:369:LEU:HD11 | 1:K:365:LEU:HD13 | 1.47 | 0.94 |

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 | A | 44/50 (88%) | 41 (93%) | 2 (4%) | 1 (2%) | 8 | 35 |
| 1 | B | 46/50 (92%) | 42 (91%) | 3 (6%) | 1 (2%) | 8 | 36 |
| 1 | C | 46/50 (92%) | 44 (96%) | 1 (2%) | 1 (2%) | 8 | 36 |
| 1 | D | 45/50 (90%) | 42 (93%) | 3 (7%) | 0 | 100 | 100 |
| 1 | E | 44/50 (88%) | 42 (96%) | 1 (2%) | 1 (2%) | 8 | 35 |
| 1 | F | 43/50 (86%) | 40 (93%) | 3 (7%) | 0 | 100 | 100 |
| 1 | G | 44/50 (88%) | 41 (93%) | 2 (4%) | 1 (2%) | 8 | 35 |
| 1 | H | 43/50 (86%) | 42 (98%) | 1 (2%) | 0 | 100 | 100 |
| 1 | I | 43/50 (86%) | 41 (95%) | 2 (5%) | 0 | 100 | 100 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Analysed | Favoured | Allowed | Outliers | Percentiles | |
|-----|-------|----------|----------|---------|----------|-------------|---|
| 1 | J | 44/50 (88%) | 43 (98%) | 1 (2%) | 0 | 100 | 100 |
| 1 | K | 45/50 (90%) | 42 (93%) | 1 (2%) | 2 (4%) | 3 | 18 |
| 1 | L | 44/50 (88%) | 42 (96%) | 1 (2%) | 1 (2%) | 8 | 35 |
| 1 | M | 43/50 (86%) | 42 (98%) | 1 (2%) | 0 | 100 | 100 |
| 1 | N | 44/50 (88%) | 41 (93%) | 2 (4%) | 1 (2%) | 8 | 35 |
| 1 | O | 44/50 (88%) | 41 (93%) | 2 (4%) | 1 (2%) | 8 | 35 |
| 1 | P | 42/50 (84%) | 41 (98%) | 1 (2%) | 0 | 100 | 100 |
| 1 | Q | 45/50 (90%) | 42 (93%) | 3 (7%) | 0 | 100 | 100 |
| 1 | R | 44/50 (88%) | 43 (98%) | 1 (2%) | 0 | 100 | 100 |
| 2 | S | 71/77 (92%) | 62 (87%) | 7 (10%) | 2 (3%) | 6 | 30 |
| 2 | T | 71/77 (92%) | 62 (87%) | 6 (8%) | 3 (4%) | 3 | 19 |
| 2 | U | 70/77 (91%) | 62 (89%) | 6 (9%) | 2 (3%) | 6 | 29 |
| All | All | 1005/1131 (89%) | 938 (93%) | 50 (5%) | 17 (2%) | 11 | 43 |

5 of 17 Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | C | 327 | ASP |
| 1 | G | 327 | ASP |
| 1 | L | 327 | ASP |
| 1 | N | 327 | ASP |
| 1 | A | 327 | 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 | A | 38/41 (93%) | 32 (84%) | 6 (16%) | 3 | 13 |
| 1 | B | 39/41 (95%) | 31 (80%) | 8 (20%) | 1 | 6 |
| 1 | C | 40/41 (98%) | 35 (88%) | 5 (12%) | 6 | 22 |
| 1 | D | 39/41 (95%) | 31 (80%) | 8 (20%) | 1 | 6 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Analysed | Rotameric | Outliers | Percentiles | | |
|-----|-------|----------|-----------|----------|-----|-----|---|
| 1 | E | 38/41 (93%) | 33 (87%) | 5 (13%) | 5 | 20 | |
| 1 | F | 37/41 (90%) | 30 (81%) | 7 (19%) | 2 | 8 | |
| 1 | G | 38/41 (93%) | 31 (82%) | 7 (18%) | 2 | 9 | |
| 1 | H | 37/41 (90%) | 32 (86%) | 5 (14%) | 5 | 20 | |
| 1 | I | 37/41 (90%) | 31 (84%) | 6 (16%) | 3 | 12 | |
| 1 | J | 38/41 (93%) | 31 (82%) | 7 (18%) | 2 | 9 | |
| 1 | K | 39/41 (95%) | 31 (80%) | 8 (20%) | 1 | 6 | |
| 1 | L | 38/41 (93%) | 33 (87%) | 5 (13%) | 5 | 20 | |
| 1 | M | 37/41 (90%) | 31 (84%) | 6 (16%) | 3 | 12 | |
| 1 | N | 38/41 (93%) | 32 (84%) | 6 (16%) | 3 | 13 | |
| 1 | O | 38/41 (93%) | 30 (79%) | 8 (21%) | 1 | 6 | |
| 1 | P | 37/41 (90%) | 31 (84%) | 6 (16%) | 3 | 12 | |
| 1 | Q | 39/41 (95%) | 33 (85%) | 6 (15%) | 3 | 14 | |
| 1 | R | 38/41 (93%) | 33 (87%) | 5 (13%) | 5 | 20 | |
| 2 | S | 64/69 (93%) | 56 (88%) | 8 (12%) | 6 | 22 | |
| 2 | T | 64/69 (93%) | 56 (88%) | 8 (12%) | 6 | 22 | |
| 2 | U | 63/69 (91%) | 56 (89%) | 7 (11%) | 8 | 29 | |
| All | All | 876/945 (93%) | 739 (84%) | 137 (16%) | 3 | 14 | |

5 of 137 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | J | 356 | ARG |
| 1 | L | 370 | ASN |
| 2 | T | 41 | ASN |
| 1 | J | 370 | ASN |
| 1 | K | 365 | LEU |

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

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | Q | 370 | ASN |
| 1 | R | 334 | HIS |
| 2 | T | 35 | ASN |
| 1 | M | 338 | GLN |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1   | O     | 338 | GLN  |

### 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 carbohydrates in this entry.

## 5.6  Ligand geometry (i)

There are no ligands in this entry.

## 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(Å²) | Q<0.9 |
|-----|-------|----------|--------|---------|-----|-----|----------|-------|
| 1 | A | 46/50 (92%) | 0.22 | 0 | 100 | 100 | 79, 83, 86, 89 | 0 |
| 1 | B | 48/50 (96%) | 0.25 | 0 | 100 | 100 | 81, 83, 87, 89 | 0 |
| 1 | C | 48/50 (96%) | 0.43 | 0 | 100 | 100 | 81, 83, 86, 88 | 1 (2%) |
| 1 | D | 47/50 (94%) | 0.45 | 0 | 100 | 100 | 81, 83, 85, 88 | 0 |
| 1 | E | 46/50 (92%) | 0.29 | 0 | 100 | 100 | 80, 83, 85, 88 | 0 |
| 1 | F | 45/50 (90%) | 0.17 | 1 (2%) | 65 | 42 | 81, 83, 86, 88 | 0 |
| 1 | G | 46/50 (92%) | 0.05 | 0 | 100 | 100 | 81, 83, 86, 88 | 0 |
| 1 | H | 45/50 (90%) | 0.08 | 1 (2%) | 65 | 42 | 81, 83, 86, 88 | 0 |
| 1 | I | 45/50 (90%) | 0.13 | 0 | 100 | 100 | 81, 83, 85, 88 | 0 |
| 1 | J | 46/50 (92%) | 0.36 | 0 | 100 | 100 | 81, 83, 86, 88 | 0 |
| 1 | K | 47/50 (94%) | 0.35 | 0 | 100 | 100 | 72, 83, 85, 88 | 1 (2%) |
| 1 | L | 46/50 (92%) | 0.36 | 0 | 100 | 100 | 81, 83, 86, 88 | 0 |
| 1 | M | 45/50 (90%) | 0.27 | 0 | 100 | 100 | 80, 83, 86, 88 | 0 |
| 1 | N | 46/50 (92%) | 0.09 | 0 | 100 | 100 | 81, 83, 85, 88 | 0 |
| 1 | O | 46/50 (92%) | 0.16 | 0 | 100 | 100 | 81, 83, 86, 88 | 0 |
| 1 | P | 44/50 (88%) | 0.04 | 0 | 100 | 100 | 81, 83, 85, 88 | 1 (2%) |
| 1 | Q | 47/50 (94%) | 0.13 | 0 | 100 | 100 | 51, 83, 87, 91 | 0 |
| 1 | R | 46/50 (92%) | 0.30 | 0 | 100 | 100 | 79, 83, 86, 88 | 0 |
| 2 | S | 73/77 (94%) | 0.06 | 0 | 100 | 100 | 72, 81, 85, 86 | 0 |
| 2 | T | 73/77 (94%) | 0.11 | 0 | 100 | 100 | 72, 81, 85, 86 | 0 |
| 2 | U | 72/77 (93%) | 0.37 | 2 (2%) | 56 | 32 | 72, 81, 85, 86 | 0 |
| All | All | 1047/1131 (92%) | 0.22 | 4 (0%) | 93 | 85 | 51, 83, 86, 91 | 3 (0%) |

All (4) RSRZ outliers are listed below:

| Mol | Chain | Res | Type | RSRZ |
|---|---|---|---|---|
| 1 | H | 368 | LEU | 2.2 |
| 2 | U | 16 | VAL | 2.2 |
| 2 | U | 56 | THR | 2.1 |
| 1 | F | 339 | LEU | 2.1 |

## 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 carbohydrates in this entry.

## 6.4 Ligands (i)

There are no ligands in this entry.

## 6.5 Other polymers (i)

There are no such residues in this entry.