# WORLDWIDE PDB PROTEIN DATA BANK

# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 20, 2023 – 02:51 PM JST

| | | |
|---|---|---|
| PDB ID | : | 7CR6 |
| Title | : | Synechocystis Cas1-Cas2/prespacer binary complex |
| Authors | : | Yu, Y.; Chen, Q. |
| Deposited on | : | 2020-08-12 |
| Resolution | : | 3.72 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 ⓘ 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 ⓘ) were used in the production of this report:

| | | |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Xtriage (Phenix) | : | 1.13 |
| EDS | : | 2.36 |
| 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 |

# 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 3.72 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.
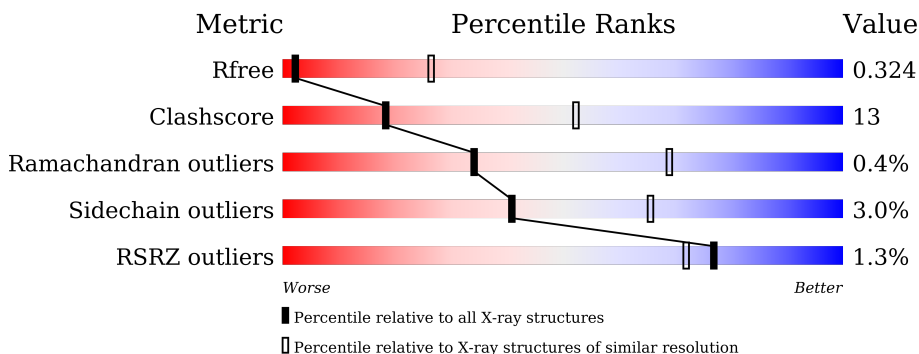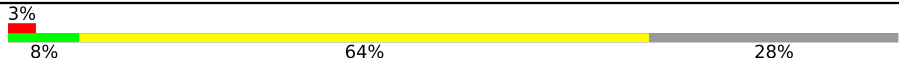


| Metric | Whole archive (#Entries) | Similar resolution (#Entries, resolution range(Å)) |
|---|---|---|
| $R_{free}$ | 130704 | 1089 (3.90-3.54) |
| Clashscore | 141614 | 1012 (3.88-3.56) |
| Ramachandran outliers | 138981 | 1114 (3.90-3.54) |
| Sidechain outliers | 138945 | 1110 (3.90-3.54) |
| RSRZ outliers | 127900 | 1020 (3.90-3.54) |

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 | 336 | 2% ... 74% ... 21% ... |
| 1 | B | 336 | % ... 66% ... 29% ... |
| 1 | C | 336 | % ... 66% ... 29% ... |
| 1 | D | 336 | % ... 69% ... 23% ... 7% |
| 2 | E | 105 | 2% ... 66% ... 21% ... 12% |
| 2 | F | 105 | 67% ... 21% ... 12% |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Length | Quality of chain |
|-----|-------|--------|------------------|
| 3 | G | 36 | 3% 8% 64% 28% |
| 4 | H | 36 | 19% 50% 31% |

# 2  Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 24088 atoms, of which 11235 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 CRISPR-associated endonuclease Cas1.

| Mol | Chain | Residues | Atoms | | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | A | 322 | Total 5072 | C 1661 | H 2484 | N 462 | O 460 | S 5 | 0 | 0 | 0 |
| 1 | B | 323 | Total 5038 | C 1663 | H 2445 | N 461 | O 464 | S 5 | 0 | 0 | 0 |
| 1 | C | 325 | Total 4966 | C 1674 | H 2354 | N 466 | O 466 | S 6 | 0 | 0 | 0 |
| 1 | D | 313 | Total 4970 | C 1622 | H 2445 | N 451 | O 447 | S 5 | 0 | 0 | 0 |

There are 44 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| A | -10 | GLY | - | expression tag | UNP Q6ZEI2 |
| A | -9 | ALA | - | expression tag | UNP Q6ZEI2 |
| A | -8 | SER | - | expression tag | UNP Q6ZEI2 |
| A | -7 | GLY | - | expression tag | UNP Q6ZEI2 |
| A | -6 | SER | - | expression tag | UNP Q6ZEI2 |
| A | -5 | GLY | - | expression tag | UNP Q6ZEI2 |
| A | -4 | THR | - | expression tag | UNP Q6ZEI2 |
| A | -3 | GLY | - | expression tag | UNP Q6ZEI2 |
| A | -2 | SER | - | expression tag | UNP Q6ZEI2 |
| A | -1 | GLY | - | expression tag | UNP Q6ZEI2 |
| A | 0 | SER | - | expression tag | UNP Q6ZEI2 |
| B | -10 | GLY | - | expression tag | UNP Q6ZEI2 |
| B | -9 | ALA | - | expression tag | UNP Q6ZEI2 |
| B | -8 | SER | - | expression tag | UNP Q6ZEI2 |
| B | -7 | GLY | - | expression tag | UNP Q6ZEI2 |
| B | -6 | SER | - | expression tag | UNP Q6ZEI2 |
| B | -5 | GLY | - | expression tag | UNP Q6ZEI2 |
| B | -4 | THR | - | expression tag | UNP Q6ZEI2 |
| B | -3 | GLY | - | expression tag | UNP Q6ZEI2 |
| B | -2 | SER | - | expression tag | UNP Q6ZEI2 |
| B | -1 | GLY | - | expression tag | UNP Q6ZEI2 |

*Continued on next page...*

*Continued from previous page...*

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|---------|-----------|
| B | 0 | SER | - | expression tag | UNP Q6ZEI2 |
| C | -10 | GLY | - | expression tag | UNP Q6ZEI2 |
| C | -9 | ALA | - | expression tag | UNP Q6ZEI2 |
| C | -8 | SER | - | expression tag | UNP Q6ZEI2 |
| C | -7 | GLY | - | expression tag | UNP Q6ZEI2 |
| C | -6 | SER | - | expression tag | UNP Q6ZEI2 |
| C | -5 | GLY | - | expression tag | UNP Q6ZEI2 |
| C | -4 | THR | - | expression tag | UNP Q6ZEI2 |
| C | -3 | GLY | - | expression tag | UNP Q6ZEI2 |
| C | -2 | SER | - | expression tag | UNP Q6ZEI2 |
| C | -1 | GLY | - | expression tag | UNP Q6ZEI2 |
| C | 0 | SER | - | expression tag | UNP Q6ZEI2 |
| D | -10 | GLY | - | expression tag | UNP Q6ZEI2 |
| D | -9 | ALA | - | expression tag | UNP Q6ZEI2 |
| D | -8 | SER | - | expression tag | UNP Q6ZEI2 |
| D | -7 | GLY | - | expression tag | UNP Q6ZEI2 |
| D | -6 | SER | - | expression tag | UNP Q6ZEI2 |
| D | -5 | GLY | - | expression tag | UNP Q6ZEI2 |
| D | -4 | THR | - | expression tag | UNP Q6ZEI2 |
| D | -3 | GLY | - | expression tag | UNP Q6ZEI2 |
| D | -2 | SER | - | expression tag | UNP Q6ZEI2 |
| D | -1 | GLY | - | expression tag | UNP Q6ZEI2 |
| D | 0 | SER | - | expression tag | UNP Q6ZEI2 |

- Molecule 2 is a protein called CRISPR-associated endoribonuclease Cas2 1.

| Mol | Chain | Residues | Atoms | | | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|---|---|---|---|---|---------|---------|-------|
| 2 | E | 92 | Total 1494 | C 485 | H 748 | N 123 | O 135 | S 3 | 0 | 0 | 0 |
| 2 | F | 92 | Total 1505 | C 485 | H 759 | N 123 | O 135 | S 3 | 0 | 0 | 0 |

There are 22 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|---------|-----------|
| E | -10 | GLY | - | expression tag | UNP Q6ZEI1 |
| E | -9 | ALA | - | expression tag | UNP Q6ZEI1 |
| E | -8 | SER | - | expression tag | UNP Q6ZEI1 |
| E | -7 | GLY | - | expression tag | UNP Q6ZEI1 |
| E | -6 | SER | - | expression tag | UNP Q6ZEI1 |
| E | -5 | GLY | - | expression tag | UNP Q6ZEI1 |
| E | -4 | THR | - | expression tag | UNP Q6ZEI1 |

*Continued on next page...*

*Continued from previous page...*

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|---------|-----------|
| E | -3 | GLY | - | expression tag | UNP Q6ZEI1 |
| E | -2 | SER | - | expression tag | UNP Q6ZEI1 |
| E | -1 | GLY | - | expression tag | UNP Q6ZEI1 |
| E | 0 | SER | - | expression tag | UNP Q6ZEI1 |
| F | -10 | GLY | - | expression tag | UNP Q6ZEI1 |
| F | -9 | ALA | - | expression tag | UNP Q6ZEI1 |
| F | -8 | SER | - | expression tag | UNP Q6ZEI1 |
| F | -7 | GLY | - | expression tag | UNP Q6ZEI1 |
| F | -6 | SER | - | expression tag | UNP Q6ZEI1 |
| F | -5 | GLY | - | expression tag | UNP Q6ZEI1 |
| F | -4 | THR | - | expression tag | UNP Q6ZEI1 |
| F | -3 | GLY | - | expression tag | UNP Q6ZEI1 |
| F | -2 | SER | - | expression tag | UNP Q6ZEI1 |
| F | -1 | GLY | - | expression tag | UNP Q6ZEI1 |
| F | 0 | SER | - | expression tag | UNP Q6ZEI1 |

- Molecule 3 is a DNA chain called DNA (36-MER).

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|---|---|---|---|---------|---------|-------|
| | | | Total | C | N | O | P | | | |
| 3 | G | 26 | 526 | 252 | 81 | 167 | 26 | 0 | 0 | 0 |

- Molecule 4 is a DNA chain called DNA (36-MER).

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|---|---|---|---|---------|---------|-------|
| | | | Total | C | N | O | P | | | |
| 4 | H | 25 | 516 | 243 | 102 | 146 | 25 | 0 | 0 | 0 |

- Molecule 5 is water.

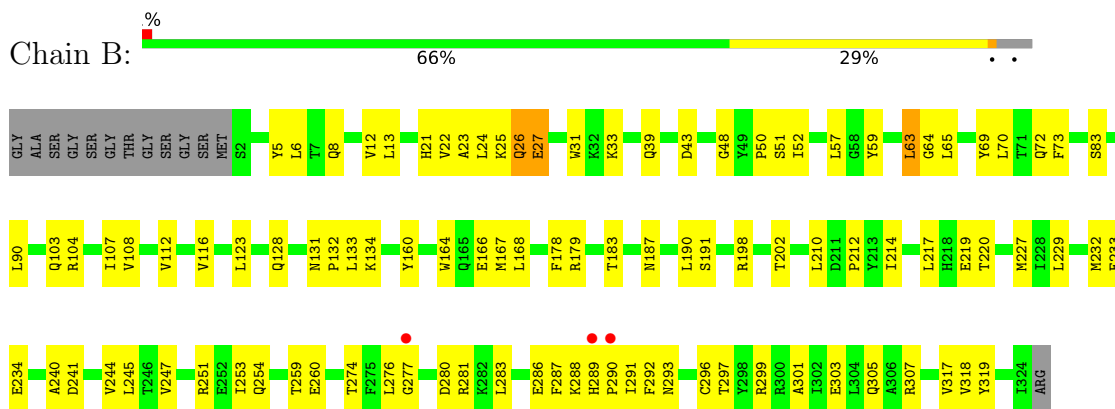| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---------|---------|
| | | | Total | O | | |
| 5 | A | 1 | 1 | 1 | 0 | 0 |

# 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.
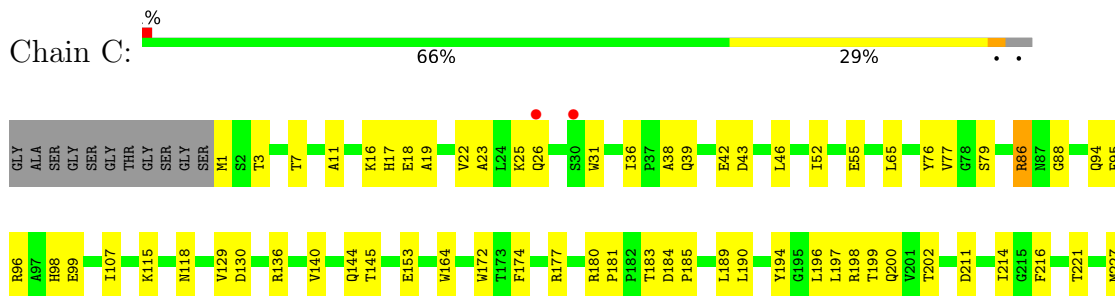
• Molecule 1: CRISPR-associated endonuclease Cas1



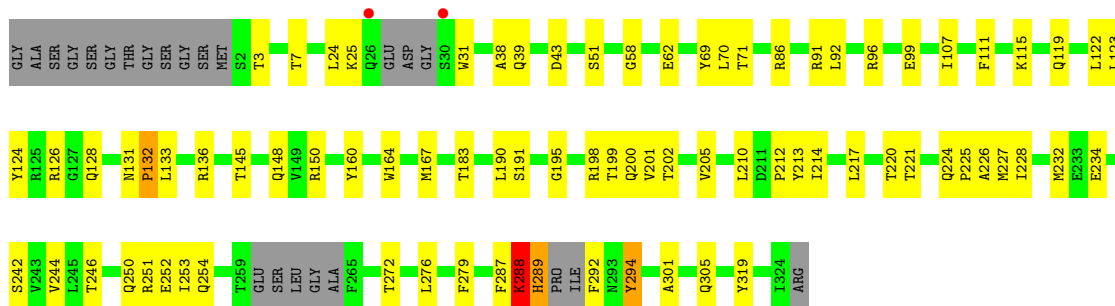• Molecule 1: CRISPR-associated endonuclease Cas1
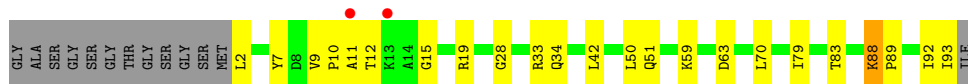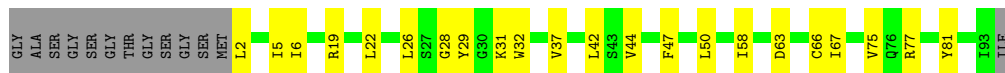


• Molecule 1: CRISPR-associated endonuclease Cas1

I228 L231 M232 E233 R236 A237 L238 V239 A240 D241 S242 V243 V244 V247 Q250 R251 E252 I253 Q254 R255 F258 T259 E260 S261 L262 G263 R266 L267 T268 D269 F275 K282 F287 K288 H289 P290 I291 F292 K295 R299 E303 R307 R311 E315

R325

• Molecule 1: CRISPR-associated endonuclease Cas1

Chain D:

%

69%          23%      7%

GLY ALA SER GLY SER GLY THR GLY SER GLY SER MET S2 T3 T7 L24 K25 Q26 GLU ASP GLY S30 W31 A38 Q39 D43 S51 G58 E62 Y69 L70 T71 R86 R91 L92 R96 E99 I107 F111 K115 Q119 L122 L123

Y124 R125 R126 G127 Q128 N131 P132 L133 R136 T145 Q148 V149 R150 Y160 W164 M167 T183 L190 S191 G195 R198 T199 Q200 V201 T202 V205 L210 D211 P212 Y213 I214 L217 T220 T221 Q224 P225 A226 M227 I228 M232 E233 E234

S242 V243 V244 L245 T246 Q250 R251 E252 I253 Q254 T259 GLU SER LEU GLY ALA F265 T272 L276 F279 F287 K288 H289 PRO ILE F292 N293 Y294 A301 Q305 Y319 I324 ARG

• Molecule 2: CRISPR-associated endoribonuclease Cas2 1

Chain E:

2%

66%          21%    •  12%

GLY ALA SER GLY SER GLY THR GLY SER GLY SER MET L2 Y7 D8 V9 P10 A11 T12 K13 A14 G15 R19 G28 R33 Q34 L42 L50 Q51 K59 D63 L70 I79 T83 K88 P89 I92 I93 ILE

• Molecule 2: CRISPR-associated endoribonuclease Cas2 1

Chain F:

67%          21%      12%

GLY ALA SER GLY SER GLY THR GLY SER GLY SER MET L2 I5 I6 R19 L22 L26 S27 G28 Y29 G30 K31 W32 V37 L42 S43 V44 F47 L50 I58 D63 C66 I67 V75 Q76 R77 Y81 I93 ILE

• Molecule 3: DNA (36-MER)

Chain G:

3%

8%          64%      28%

DT DT DT DT DT T7 T8 G9 T10 G11 C12 C13 C14 C15 T16 G17 G18 C19 G20 G21 T22 C23 G24 C25 T26 T27 T28 C29 T30 T31 T32 DT DT DT

• Molecule 4: DNA (36-MER)

Chain H:

19%          50%      31%

DT DT DT DT DT DT G8 A9 A10 A11 G12 C13 C16 C17 G18 C19 C20 G23 G24 G25 C26 A27 T31 T32 DT DT DT

# 4 Data and refinement statistics   ⓘ

| Property | Value | Source |
|---|---|---|
| Space group | C 2 2 21 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 125.56Å    215.16Å    191.09Å<br>90.00°    90.00°    90.00° | Depositor |
| Resolution (Å) | 33.81   –   3.72<br>33.81   –   3.72 | Depositor<br>EDS |
| % Data completeness<br>(in resolution range) | 91.9 (33.81-3.72)<br>92.0 (33.81-3.72) | Depositor<br>EDS |
| $R_{merge}$ | (Not available) | Depositor |
| $R_{sym}$ | (Not available) | Depositor |
| $< I/\sigma(I) >$ [1] | 2.49 (at 3.76Å) | Xtriage |
| Refinement program | PHENIX 1.15.2_3472 | Depositor |
| R, $R_{free}$ | 0.276   ,   0.324<br>0.276   ,   0.324 | Depositor<br>DCC |
| $R_{free}$ test set | 1058 reflections (4.19%) | wwPDB-VP |
| Wilson B-factor (Å$^2$) | 44.3 | Xtriage |
| Anisotropy | 0.207 | Xtriage |
| Bulk solvent $k_{sol}$(e/Å$^3$), $B_{sol}$(Å$^2$) | 0.29 , 23.9 | EDS |
| L-test for twinning[2] | $< \|L\| > = 0.47, < L^2 > = 0.30$ | Xtriage |
| Estimated twinning fraction | 0.022 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l<br>0.035 for 1/2*h+1/2*k,3/2*h-1/2*k,-l | Xtriage |
| $F_o,F_c$ correlation | 0.79 | EDS |
| Total number of atoms | 24088 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 44.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.73% 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 | 0.25 | 0/2646 | 0.49 | 0/3577 |
| 1 | B | 0.26 | 0/2652 | 0.49 | 0/3589 |
| 1 | C | 0.25 | 0/2671 | 0.49 | 0/3613 |
| 1 | D | 0.25 | 0/2580 | 0.47 | 0/3486 |
| 2 | E | 0.25 | 0/761 | 0.46 | 0/1028 |
| 2 | F | 0.25 | 0/761 | 0.51 | 0/1028 |
| 3 | G | 0.47 | 0/584 | 0.97 | 0/899 |
| 4 | H | 0.49 | 0/580 | 0.80 | 0/893 |
| All | All | 0.28 | 0/13235 | 0.54 | 0/18113 |

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 | D | 0 | 1 |

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

| Mol | Chain | Res | Type | Group |
|-----|-------|-----|------|-------|
| 1 | D | 288 | LYS | Peptide |

## 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 | 2588 | 2484 | 2621 | 48 | 0 |
| 1 | B | 2593 | 2445 | 2618 | 82 | 1 |
| 1 | C | 2612 | 2354 | 2643 | 82 | 1 |
| 1 | D | 2525 | 2445 | 2554 | 70 | 0 |
| 2 | E | 746 | 748 | 767 | 25 | 0 |
| 2 | F | 746 | 759 | 767 | 14 | 0 |
| 3 | G | 526 | 0 | 298 | 29 | 0 |
| 4 | H | 516 | 0 | 279 | 17 | 0 |
| 5 | A | 1 | 0 | 0 | 0 | 0 |
| All | All | 12853 | 11235 | 12547 | 337 | 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 13.

The worst 5 of 337 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) |
|--------|--------|--------------------------|-------------------|
| 4:H:12:DG:H2" | 4:H:13:DC:H5" | 1.31 | 1.07 |
| 1:B:25:LYS:HD2 | 1:B:25:LYS:O | 1.63 | 0.97 |
| 1:D:288:LYS:HD2 | 1:D:289:HIS:H | 1.34 | 0.90 |
| 1:C:46:LEU:HD21 | 1:C:52:ILE:HD11 | 1.52 | 0.90 |
| 1:B:25:LYS:HE2 | 1:B:31:TRP:HE1 | 1.36 | 0.90 |

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:179:ARG:O | 1:C:144:GLN:HE22[4_555] | 1.58 | 0.02 |

## 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 | 318/336 (95%) | 282 (89%) | 34 (11%) | 2 (1%) | 25 | 61 |
| 1 | B | 321/336 (96%) | 282 (88%) | 37 (12%) | 2 (1%) | 25 | 61 |
| 1 | C | 323/336 (96%) | 279 (86%) | 44 (14%) | 0 | 100 | 100 |
| 1 | D | 305/336 (91%) | 268 (88%) | 35 (12%) | 2 (1%) | 22 | 59 |
| 2 | E | 90/105 (86%) | 84 (93%) | 6 (7%) | 0 | 100 | 100 |
| 2 | F | 90/105 (86%) | 80 (89%) | 10 (11%) | 0 | 100 | 100 |
| All | All | 1447/1554 (93%) | 1275 (88%) | 166 (12%) | 6 (0%) | 34 | 69 |

5 of 6 Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | B | 26 | GLN |
| 1 | D | 288 | LYS |
| 1 | D | 132 | PRO |
| 1 | A | 289 | HIS |
| 1 | A | 290 | PRO |

### 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 | 272/280 (97%) | 261 (96%) | 11 (4%) | 31 | 59 |
| 1 | B | 273/280 (98%) | 271 (99%) | 2 (1%) | 84 | 91 |
| 1 | C | 275/280 (98%) | 262 (95%) | 13 (5%) | 26 | 56 |
| 1 | D | 266/280 (95%) | 259 (97%) | 7 (3%) | 46 | 69 |
| 2 | E | 82/89 (92%) | 80 (98%) | 2 (2%) | 49 | 70 |
| 2 | F | 82/89 (92%) | 80 (98%) | 2 (2%) | 49 | 70 |
| All | All | 1250/1298 (96%) | 1213 (97%) | 37 (3%) | 41 | 65 |

5 of 37 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | D | 251 | ARG |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 2 | F | 42 | LEU |
| 1 | D | 279 | PHE |
| 1 | D | 294 | TYR |
| 1 | B | 287 | PHE |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | A | 143 | GLN |
| 1 | A | 148 | GLN |
| 1 | B | 39 | GLN |
| 1 | B | 117 | HIS |
| 1 | C | 289 | HIS |

### 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 ⓘ

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 | 322/336 (95%) | 0.02 | 8 (2%) | 57 | 49 | 11, 43, 83, 107 | 0 |
| 1 | B | 323/336 (96%) | -0.11 | 3 (0%) | 84 | 80 | 14, 38, 71, 100 | 0 |
| 1 | C | 325/336 (96%) | -0.15 | 4 (1%) | 79 | 74 | 9, 30, 72, 120 | 0 |
| 1 | D | 313/336 (93%) | -0.22 | 2 (0%) | 89 | 87 | 12, 32, 62, 79 | 0 |
| 2 | E | 92/105 (87%) | 0.00 | 2 (2%) | 62 | 55 | 22, 45, 73, 89 | 0 |
| 2 | F | 92/105 (87%) | -0.17 | 0 | 100 | 100 | 12, 43, 58, 62 | 0 |
| 3 | G | 26/36 (72%) | 0.47 | 1 (3%) | 40 | 33 | 41, 71, 105, 114 | 0 |
| 4 | H | 25/36 (69%) | 0.30 | 0 | 100 | 100 | 38, 67, 88, 91 | 0 |
| All | All | 1518/1626 (93%) | -0.09 | 20 (1%) | 77 | 72 | 9, 38, 77, 120 | 0 |

The worst 5 of 20 RSRZ outliers are listed below:

| Mol | Chain | Res | Type | RSRZ |
|-----|-------|-----|------|------|
| 1 | C | 30 | SER | 4.6 |
| 1 | B | 289 | HIS | 3.7 |
| 1 | D | 30 | SER | 3.7 |
| 1 | A | 263 | GLY | 3.7 |
| 2 | E | 13 | LYS | 3.2 |

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

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

## 6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

## 6.4 Ligands ⓘ

There are no ligands in this entry.

## 6.5 Other polymers ⓘ

There are no such residues in this entry.