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

Jun 18, 2024 – 10:42 PM EDT

| | | |
|---|---|---|
| PDB ID | : | 4NEF |
| Title | : | X-ray structure of human Aquaporin 2 |
| Authors | : | Frick, A.; Eriksson, U.; Mattia, F.D.; Oberg, F.; Hedfalk, K.; Neutze, R.; Grip, W.D.; Deen, P.M.T.; Tornroth-horsefield, S. |
| Deposited on | : | 2013-10-29 |
| Resolution | : | 2.75 Å(reported) |

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

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

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

The following versions of software and data (see references (1)) were used in the production of this report:

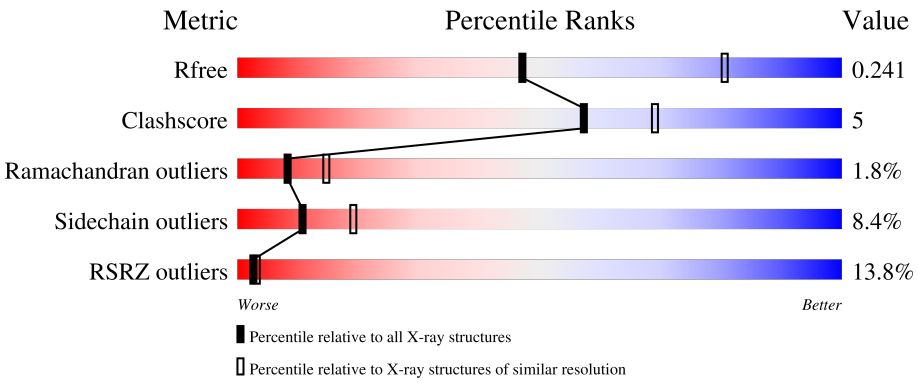
| | | |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Xtriage (Phenix) | : | 1.20.1 |
| EDS | : | 2.37.1 |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| Refmac | : | 5.8.0158 |
| CCP4 | : | 7.0.044 (Gargrove) |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | 2.37.1 |

# 1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.75 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 | 1235 (2.78-2.74) |
| Clashscore | 141614 | 1277 (2.78-2.74) |
| Ramachandran outliers | 138981 | 1257 (2.78-2.74) |
| Sidechain outliers | 138945 | 1257 (2.78-2.74) |
| RSRZ outliers | 127900 | 1207 (2.78-2.74) |

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 | 242 | 11% 80% 17% • • |
| 1 | B | 242 | 13% 81% 15% • • |
| 1 | C | 242 | 7% 81% 13% • • |
| 1 | D | 242 | 22% 78% 16% • • |

# 2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 6986 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 Aquaporin-2.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | | |
| 1 | A | 239 | 1761 | 1150 | 296 | 309 | 6 | 0 | 0 | 0 |
| 1 | B | 233 | 1715 | 1122 | 287 | 300 | 6 | 0 | 0 | 0 |
| 1 | C | 236 | 1736 | 1135 | 290 | 305 | 6 | 0 | 0 | 0 |
| 1 | D | 232 | 1712 | 1116 | 288 | 302 | 6 | 0 | 0 | 0 |

There are 12 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| A | 1 | GLY | - | EXPRESSION TAG | UNP P41181 |
| A | 2 | SER | - | EXPRESSION TAG | UNP P41181 |
| A | 242 | PRO | - | EXPRESSION TAG | UNP P41181 |
| B | 1 | GLY | - | EXPRESSION TAG | UNP P41181 |
| B | 2 | SER | - | EXPRESSION TAG | UNP P41181 |
| B | 242 | PRO | - | EXPRESSION TAG | UNP P41181 |
| C | 1 | GLY | - | EXPRESSION TAG | UNP P41181 |
| C | 2 | SER | - | EXPRESSION TAG | UNP P41181 |
| C | 242 | PRO | - | EXPRESSION TAG | UNP P41181 |
| D | 1 | GLY | - | EXPRESSION TAG | UNP P41181 |
| D | 2 | SER | - | EXPRESSION TAG | UNP P41181 |
| D | 242 | PRO | - | EXPRESSION TAG | UNP P41181 |

- Molecule 2 is CADMIUM ION (three-letter code: CD) (formula: Cd).

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|
| | | | Total | Cd | | |
| 2 | A | 1 | 1 | 1 | 0 | 0 |
| 2 | D | 1 | 1 | 1 | 0 | 0 |

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|
| | | | Total | Zn | | |
| 3 | A | 1 | 1 | 1 | 0 | 0 |

- Molecule 4 is water.

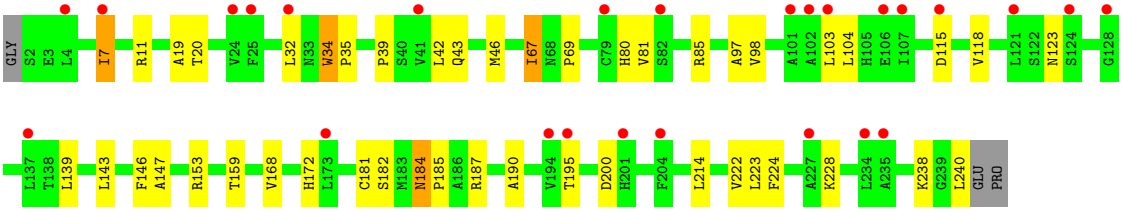
| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|
| | | | Total | O | | |
| 4 | A | 19 | 19 | 19 | 0 | 0 |
| | | | Total | O | | |
| 4 | B | 14 | 14 | 14 | 0 | 0 |
| | | | Total | O | | |
| 4 | C | 13 | 13 | 13 | 0 | 0 |
| | | | Total | O | | |
| 4 | D | 13 | 13 | 13 | 0 | 0 |

# 3 Residue-property plots (i)

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

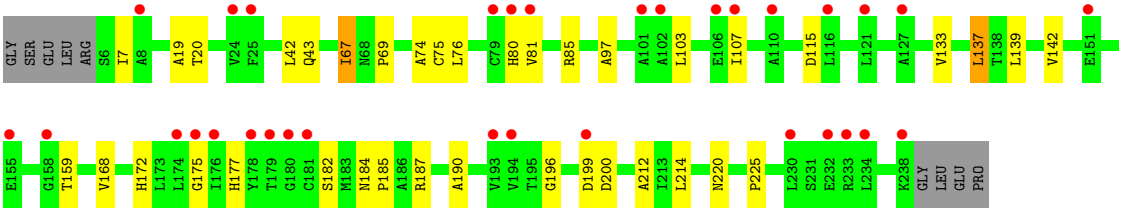
- Molecule 1: Aquaporin-2

Chain A:



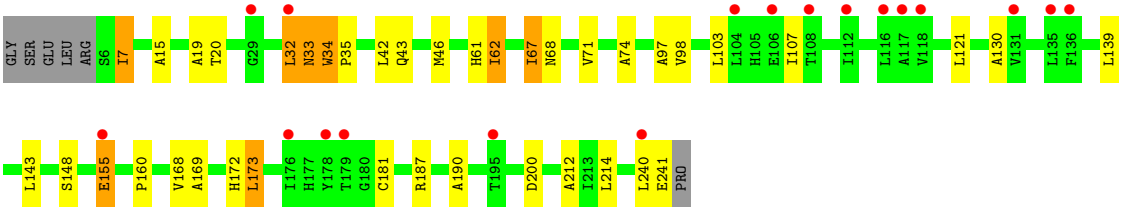

- Molecule 1: Aquaporin-2

Chain B:



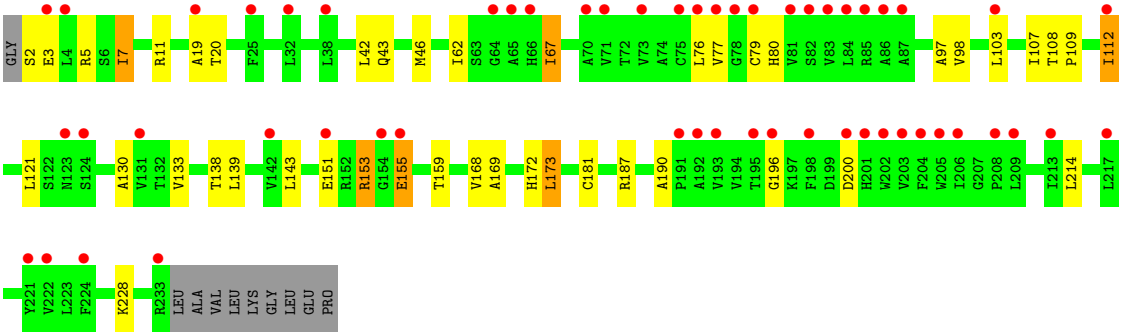

- Molecule 1: Aquaporin-2

Chain C:




- Molecule 1: Aquaporin-2

Chain D:


GLY
S2
E3
L4
R5
S6
I7
R11
A19
T20
F25
L32
L38
L42
Q43
M46
I62
S63
G64
A65
H66
I67
A70
V71
T72
V73
A74
C75
L76
V77
G78
C79
H80
V81
S82
V83
L84
R85
A86
A87
A97
V98
L103
I107
T108
P109
I112
L121
S122
N123
S124
A130
V131
T132
V133
T138
L139
V142
L143
E151
R152
R153
G154
E155
T159
V168
A169
H172
L173
C181
R187
A190
P191
A192
V193
V194
T195
G196
K197
F198
D199
D200
H201
W202
V203
P204
W205
I206
G207
P208
L209
I213
L214
L217
Y221
V222
L223
F224
K228
R233
LEU
ALA
VAL
LEU
LYS
GLY
LEU
GLU
PRO

# 4 Data and refinement statistics (i)

| Property | Value | Source |
|---|---|---|
| Space group | P 42 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 119.11Å 119.11Å 90.62Å<br>90.00° 90.00° 90.00° | Depositor |
| Resolution (Å) | 49.96 – 2.75<br>72.12 – 2.75 | Depositor<br>EDS |
| % Data completeness<br>(in resolution range) | 99.8 (49.96-2.75)<br>99.8 (72.12-2.75) | Depositor<br>EDS |
| $R_{merge}$ | (Not available) | Depositor |
| $R_{sym}$ | 0.09 | Depositor |
| $< I/\sigma(I) >$ [1] | 1.51 (at 2.73Å) | Xtriage |
| Refinement program | BUSTER-TNT, BUSTER 2.10.0 | Depositor |
| R, $R_{free}$ | 0.202 , 0.225<br>0.212 , 0.241 | Depositor<br>DCC |
| $R_{free}$ test set | 1674 reflections (5.06%) | wwPDB-VP |
| Wilson B-factor (Å$^2$) | 64.4 | Xtriage |
| Anisotropy | 0.067 | Xtriage |
| Bulk solvent $k_{sol}$(e/Å$^3$), $B_{sol}$(Å$^2$) | 0.33 , 102.6 | EDS |
| L-test for twinning[2] | $< \|L\| > = 0.49$, $< L^2 > = 0.32$ | Xtriage |
| Estimated twinning fraction | 0.036 for h,-k,-l | Xtriage |
| $F_o$,$F_c$ correlation | 0.91 | EDS |
| Total number of atoms | 6986 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 91.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.46% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

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

# 5 Model quality (i)

## 5.1 Standard geometry (i)

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

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 | 0.46 | 0/1802 | 0.67 | 0/2463 |
| 1 | B | 0.47 | 0/1756 | 0.65 | 0/2402 |
| 1 | C | 0.47 | 0/1777 | 0.70 | 0/2430 |
| 1 | D | 0.47 | 0/1753 | 0.68 | 0/2397 |
| All | All | 0.47 | 0/7088 | 0.67 | 0/9692 |

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 (i)

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

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|---|---|---|---|---|---|---|
| 1 | A | 1761 | 0 | 1805 | 25 | 0 |
| 1 | B | 1715 | 0 | 1757 | 18 | 0 |
| 1 | C | 1736 | 0 | 1775 | 21 | 0 |
| 1 | D | 1712 | 0 | 1741 | 20 | 0 |
| 2 | A | 1 | 0 | 0 | 0 | 0 |
| 2 | D | 1 | 0 | 0 | 0 | 0 |
| 3 | A | 1 | 0 | 0 | 0 | 0 |
| 4 | A | 19 | 0 | 0 | 0 | 0 |
| 4 | B | 14 | 0 | 0 | 0 | 0 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|---|---|---|---|---|---|---|
| 4 | C | 13 | 0 | 0 | 0 | 0 |
| 4 | D | 13 | 0 | 0 | 0 | 0 |
| All | All | 6986 | 0 | 7078 | 69 | 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 5.

All (69) 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:A:222:VAL:O | 1:D:11:ARG:HD3 | 1.97 | 0.64 |
| 1:C:32:LEU:HD12 | 1:C:33:ASN:H | 1.65 | 0.61 |
| 1:B:42:LEU:HB2 | 1:C:43:GLN:HG3 | 1.83 | 0.60 |
| 1:B:69:PRO:HD2 | 1:B:185:PRO:HD2 | 1.87 | 0.57 |
| 1:B:81:VAL:HG23 | 1:B:85:ARG:HD2 | 1.89 | 0.54 |
| 1:D:112:ILE:O | 1:D:112:ILE:HD13 | 2.08 | 0.54 |
| 1:A:80:HIS:HE1 | 1:A:146:PHE:HE1 | 1.56 | 0.53 |
| 1:A:223:LEU:HB3 | 1:D:62:ILE:HD13 | 1.91 | 0.52 |
| 1:A:42:LEU:HB2 | 1:D:43:GLN:HG3 | 1.92 | 0.52 |
| 1:D:19:ALA:HB1 | 1:D:67:ILE:HG12 | 1.93 | 0.50 |
| 1:A:19:ALA:HB1 | 1:A:67:ILE:HG12 | 1.93 | 0.50 |
| 1:A:159:THR:HG21 | 1:D:159:THR:HG22 | 1.94 | 0.50 |
| 1:B:43:GLN:HG3 | 1:D:42:LEU:HB2 | 1.94 | 0.50 |
| 1:B:142:VAL:HG21 | 1:B:212:ALA:HA | 1.94 | 0.49 |
| 1:A:46:MET:HG2 | 1:C:42:LEU:HD11 | 1.94 | 0.49 |
| 1:C:34:TRP:CD1 | 1:C:34:TRP:N | 2.79 | 0.49 |
| 1:A:80:HIS:CE1 | 1:A:146:PHE:HE1 | 2.31 | 0.48 |
| 1:D:169:ALA:O | 1:D:173:LEU:HD22 | 2.13 | 0.48 |
| 1:D:168:VAL:O | 1:D:172:HIS:HD2 | 1.96 | 0.48 |
| 1:A:81:VAL:HG13 | 1:A:85:ARG:HB3 | 1.95 | 0.48 |
| 1:C:74:ALA:HB2 | 1:C:212:ALA:HB1 | 1.96 | 0.48 |
| 1:C:168:VAL:O | 1:C:172:HIS:HD2 | 1.95 | 0.48 |
| 1:B:159:THR:HG22 | 1:D:159:THR:HG21 | 1.94 | 0.48 |
| 1:D:3:GLU:O | 1:D:7:ILE:HG13 | 2.14 | 0.48 |
| 1:B:19:ALA:HB1 | 1:B:67:ILE:HG12 | 1.95 | 0.47 |
| 1:B:168:VAL:O | 1:B:172:HIS:HD2 | 1.97 | 0.47 |
| 1:C:19:ALA:HB1 | 1:C:67:ILE:HG12 | 1.95 | 0.47 |
| 1:D:2:SER:HB3 | 1:D:5:ARG:HG3 | 1.95 | 0.47 |
| 1:A:43:GLN:HG3 | 1:C:42:LEU:HB2 | 1.97 | 0.46 |
| 1:A:168:VAL:O | 1:A:172:HIS:HD2 | 1.98 | 0.46 |
| 1:B:74:ALA:HB2 | 1:B:212:ALA:HB1 | 1.97 | 0.46 |

*Continued on next page...*

*Continued from previous page...*

| Atom-1 | Atom-2 | Interatomic distance (Å) | Clash overlap (Å) |
|---|---|---|---|
| 1:C:61:HIS:H | 1:C:61:HIS:CD2 | 2.33 | 0.46 |
| 1:A:69:PRO:HD2 | 1:A:185:PRO:HD2 | 1.98 | 0.46 |
| 1:B:133:VAL:HG22 | 1:C:107:ILE:HD13 | 1.96 | 0.46 |
| 1:C:148:SER:HB3 | 1:C:160:PRO:HD2 | 1.98 | 0.45 |
| 1:A:34:TRP:HB3 | 1:A:39:PRO:HD3 | 1.99 | 0.44 |
| 1:D:153:ARG:O | 1:D:153:ARG:HG3 | 2.17 | 0.44 |
| 1:B:220:ASN:O | 1:B:225:PRO:HA | 2.17 | 0.44 |
| 1:B:137:LEU:HD21 | 1:B:175:GLY:HA3 | 1.99 | 0.44 |
| 1:C:98:VAL:HG12 | 1:C:190:ALA:HA | 2.00 | 0.44 |
| 1:C:169:ALA:O | 1:C:173:LEU:HD22 | 2.18 | 0.44 |
| 1:B:42:LEU:HD11 | 1:C:46:MET:HG2 | 2.00 | 0.44 |
| 1:A:80:HIS:HE1 | 1:A:146:PHE:CE1 | 2.36 | 0.44 |
| 1:A:97:ALA:HB1 | 1:A:190:ALA:HB2 | 2.01 | 0.43 |
| 1:A:182:SER:C | 1:A:184:ASN:H | 2.22 | 0.43 |
| 1:A:147:ALA:O | 1:A:153:ARG:HD3 | 2.17 | 0.43 |
| 1:A:115:ASP:HB2 | 1:A:118:VAL:HG22 | 2.01 | 0.43 |
| 1:A:32:LEU:HD21 | 1:A:104:LEU:HD21 | 2.01 | 0.43 |
| 1:B:177:HIS:CG | 1:C:33:ASN:HB3 | 2.53 | 0.43 |
| 1:D:121:LEU:HD21 | 1:D:130:ALA:CB | 2.48 | 0.42 |
| 1:A:7:ILE:H | 1:A:7:ILE:HG13 | 1.68 | 0.42 |
| 1:C:121:LEU:HD21 | 1:C:130:ALA:CB | 2.49 | 0.42 |
| 1:B:76:LEU:HA | 1:B:81:VAL:CG1 | 2.50 | 0.42 |
| 1:C:121:LEU:HD21 | 1:C:130:ALA:HB2 | 2.01 | 0.42 |
| 1:A:98:VAL:HG12 | 1:A:190:ALA:HA | 2.02 | 0.42 |
| 1:A:7:ILE:O | 1:A:11:ARG:HG3 | 2.19 | 0.42 |
| 1:D:121:LEU:HD21 | 1:D:130:ALA:HB2 | 2.00 | 0.42 |
| 1:C:15:ALA:HB2 | 1:C:62:ILE:HD11 | 2.01 | 0.41 |
| 1:B:107:ILE:HD13 | 1:D:133:VAL:HG22 | 2.03 | 0.41 |
| 1:C:7:ILE:H | 1:C:7:ILE:HG13 | 1.74 | 0.41 |
| 1:D:98:VAL:HG12 | 1:D:190:ALA:HA | 2.03 | 0.41 |
| 1:B:182:SER:C | 1:B:184:ASN:H | 2.24 | 0.41 |
| 1:A:80:HIS:CE1 | 1:A:146:PHE:CE1 | 3.09 | 0.41 |
| 1:D:97:ALA:HB1 | 1:D:190:ALA:HB2 | 2.01 | 0.41 |
| 1:C:68:ASN:HB3 | 1:C:71:VAL:HB | 2.01 | 0.41 |
| 1:B:97:ALA:HB1 | 1:B:190:ALA:HB2 | 2.02 | 0.41 |
| 1:C:97:ALA:HB1 | 1:C:190:ALA:HB2 | 2.03 | 0.41 |
| 1:A:42:LEU:HD11 | 1:D:46:MET:HG2 | 2.02 | 0.40 |
| 1:A:123:ASN:O | 1:D:109:PRO:HG3 | 2.22 | 0.40 |

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 | 237/242 (98%) | 223 (94%) | 11 (5%) | 3 (1%) | 12 21 |
| 1 | B | 231/242 (96%) | 210 (91%) | 18 (8%) | 3 (1%) | 12 21 |
| 1 | C | 234/242 (97%) | 217 (93%) | 12 (5%) | 5 (2%) | 7 12 |
| 1 | D | 230/242 (95%) | 216 (94%) | 8 (4%) | 6 (3%) | 5 8 |
| All | All | 932/968 (96%) | 866 (93%) | 49 (5%) | 17 (2%) | 8 15 |

All (17) Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | A | 35 | PRO |
| 1 | C | 33 | ASN |
| 1 | C | 35 | PRO |
| 1 | D | 155 | GLU |
| 1 | B | 196 | GLY |
| 1 | C | 155 | GLU |
| 1 | C | 181 | CYS |
| 1 | D | 79 | CYS |
| 1 | D | 80 | HIS |
| 1 | D | 196 | GLY |
| 1 | A | 181 | CYS |
| 1 | D | 181 | CYS |
| 1 | B | 80 | HIS |
| 1 | B | 67 | ILE |
| 1 | C | 67 | ILE |
| 1 | A | 67 | ILE |
| 1 | D | 67 | ILE |

### 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 | 182/184 (99%) | 167 (92%) | 15 (8%) | 11 | 20 |
| 1 | B | 177/184 (96%) | 166 (94%) | 11 (6%) | 18 | 32 |
| 1 | C | 179/184 (97%) | 164 (92%) | 15 (8%) | 11 | 19 |
| 1 | D | 177/184 (96%) | 158 (89%) | 19 (11%) | 6 | 11 |
| All | All | 715/736 (97%) | 655 (92%) | 60 (8%) | 11 | 19 |

All (60) residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | A | 7 | ILE |
| 1 | A | 20 | THR |
| 1 | A | 34 | TRP |
| 1 | A | 103 | LEU |
| 1 | A | 139 | LEU |
| 1 | A | 143 | LEU |
| 1 | A | 184 | ASN |
| 1 | A | 187 | ARG |
| 1 | A | 195 | THR |
| 1 | A | 200 | ASP |
| 1 | A | 214 | LEU |
| 1 | A | 224 | PHE |
| 1 | A | 228 | LYS |
| 1 | A | 238 | LYS |
| 1 | A | 240 | LEU |
| 1 | B | 7 | ILE |
| 1 | B | 20 | THR |
| 1 | B | 75 | CYS |
| 1 | B | 103 | LEU |
| 1 | B | 115 | ASP |
| 1 | B | 137 | LEU |
| 1 | B | 139 | LEU |
| 1 | B | 187 | ARG |
| 1 | B | 199 | ASP |
| 1 | B | 200 | ASP |
| 1 | B | 214 | LEU |
| 1 | C | 7 | ILE |
| 1 | C | 20 | THR |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | C | 32 | LEU |
| 1 | C | 34 | TRP |
| 1 | C | 62 | ILE |
| 1 | C | 103 | LEU |
| 1 | C | 139 | LEU |
| 1 | C | 143 | LEU |
| 1 | C | 155 | GLU |
| 1 | C | 173 | LEU |
| 1 | C | 187 | ARG |
| 1 | C | 200 | ASP |
| 1 | C | 214 | LEU |
| 1 | C | 240 | LEU |
| 1 | C | 241 | GLU |
| 1 | D | 7 | ILE |
| 1 | D | 20 | THR |
| 1 | D | 76 | LEU |
| 1 | D | 77 | VAL |
| 1 | D | 103 | LEU |
| 1 | D | 107 | ILE |
| 1 | D | 108 | THR |
| 1 | D | 112 | ILE |
| 1 | D | 138 | THR |
| 1 | D | 139 | LEU |
| 1 | D | 143 | LEU |
| 1 | D | 151 | GLU |
| 1 | D | 153 | ARG |
| 1 | D | 155 | GLU |
| 1 | D | 173 | LEU |
| 1 | D | 187 | ARG |
| 1 | D | 200 | ASP |
| 1 | D | 214 | LEU |
| 1 | D | 228 | LYS |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (16) such sidechains are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | A | 80 | HIS |
| 1 | A | 93 | GLN |
| 1 | A | 105 | HIS |
| 1 | A | 172 | HIS |
| 1 | A | 184 | ASN |
| 1 | B | 119 | ASN |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | B | 172 | HIS |
| 1 | B | 177 | HIS |
| 1 | C | 61 | HIS |
| 1 | C | 93 | GLN |
| 1 | C | 105 | HIS |
| 1 | C | 172 | HIS |
| 1 | C | 177 | HIS |
| 1 | D | 119 | ASN |
| 1 | D | 172 | HIS |
| 1 | D | 220 | ASN |

#### 5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

### 5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

### 5.6 Ligand geometry (i)

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

### 5.7 Other polymers (i)

There are no such residues in this entry.

## 5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

# 6 Fit of model and data (i)

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

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

| Mol | Chain | Analysed | <RSRZ> | #RSRZ>2 | | | OWAB(Å$^2$) | Q<0.9 |
|---|---|---|---|---|---|---|---|---|
| 1 | A | 239/242 (98%) | 0.90 | 26 (10%) | 5 | 6 | 38, 88, 142, 183 | 0 |
| 1 | B | 233/242 (96%) | 0.93 | 32 (13%) | 3 | 3 | 42, 90, 145, 235 | 1 (0%) |
| 1 | C | 236/242 (97%) | 0.76 | 18 (7%) | 13 | 16 | 43, 81, 124, 193 | 0 |
| 1 | D | 232/242 (95%) | 1.29 | 54 (23%) | 0 | 0 | 38, 91, 133, 190 | 2 (0%) |
| All | All | 940/968 (97%) | 0.97 | 130 (13%) | 2 | 3 | 38, 88, 137, 235 | 3 (0%) |

All (130) RSRZ outliers are listed below:

| Mol | Chain | Res | Type | RSRZ |
|---|---|---|---|---|
| 1 | D | 154 | GLY | 10.3 |
| 1 | D | 84 | LEU | 8.9 |
| 1 | B | 79 | CYS | 6.7 |
| 1 | D | 82 | SER | 6.5 |
| 1 | D | 85 | ARG | 6.1 |
| 1 | D | 83 | VAL | 5.8 |
| 1 | C | 112 | ILE | 5.6 |
| 1 | D | 79 | CYS | 5.3 |
| 1 | D | 193 | VAL | 5.0 |
| 1 | D | 206 | ILE | 4.8 |
| 1 | A | 106 | GLU | 4.7 |
| 1 | D | 204 | PHE | 4.6 |
| 1 | D | 81 | VAL | 4.6 |
| 1 | B | 238 | LYS | 4.4 |
| 1 | B | 178 | TYR | 4.4 |
| 1 | B | 81 | VAL | 4.2 |
| 1 | D | 203 | VAL | 4.1 |
| 1 | D | 196 | GLY | 4.1 |
| 1 | D | 217 | LEU | 4.0 |
| 1 | D | 221 | TYR | 3.9 |
| 1 | D | 201 | HIS | 3.9 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type | RSRZ |
|---|---|---|---|---|
| 1 | B | 175 | GLY | 3.7 |
| 1 | A | 234 | LEU | 3.7 |
| 1 | A | 227 | ALA | 3.7 |
| 1 | D | 4 | LEU | 3.7 |
| 1 | B | 110 | ALA | 3.7 |
| 1 | B | 107 | ILE | 3.6 |
| 1 | B | 180 | GLY | 3.5 |
| 1 | B | 8 | ALA | 3.5 |
| 1 | D | 75 | CYS | 3.5 |
| 1 | B | 25 | PHE | 3.4 |
| 1 | D | 112 | ILE | 3.4 |
| 1 | C | 155 | GLU | 3.4 |
| 1 | D | 76 | LEU | 3.4 |
| 1 | A | 115 | ASP | 3.4 |
| 1 | D | 66 | HIS | 3.3 |
| 1 | D | 103 | LEU | 3.3 |
| 1 | B | 234 | LEU | 3.3 |
| 1 | C | 135 | LEU | 3.2 |
| 1 | D | 131 | VAL | 3.2 |
| 1 | D | 151 | GLU | 3.2 |
| 1 | B | 176 | ILE | 3.2 |
| 1 | A | 124 | SER | 3.1 |
| 1 | B | 24 | VAL | 3.1 |
| 1 | A | 24 | VAL | 3.1 |
| 1 | B | 106 | GLU | 3.1 |
| 1 | D | 77 | VAL | 3.1 |
| 1 | A | 235 | ALA | 3.1 |
| 1 | C | 118 | VAL | 3.1 |
| 1 | C | 116 | LEU | 3.1 |
| 1 | D | 233 | ARG | 3.1 |
| 1 | B | 194 | VAL | 3.1 |
| 1 | D | 198 | PHE | 3.0 |
| 1 | D | 202 | TRP | 3.0 |
| 1 | B | 174 | LEU | 3.0 |
| 1 | C | 117 | ALA | 3.0 |
| 1 | B | 80 | HIS | 2.9 |
| 1 | D | 209 | LEU | 2.9 |
| 1 | B | 151 | GLU | 2.9 |
| 1 | D | 25 | PHE | 2.9 |
| 1 | D | 70 | ALA | 2.9 |
| 1 | B | 155 | GLU | 2.8 |
| 1 | C | 106 | GLU | 2.8 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type | RSRZ |
|---|---|---|---|---|
| 1 | A | 103 | LEU | 2.8 |
| 1 | B | 127 | ALA | 2.8 |
| 1 | D | 222 | VAL | 2.8 |
| 1 | A | 79 | CYS | 2.8 |
| 1 | B | 232 | GLU | 2.7 |
| 1 | A | 41 | VAL | 2.7 |
| 1 | D | 87 | ALA | 2.7 |
| 1 | A | 25 | PHE | 2.7 |
| 1 | D | 3 | GLU | 2.7 |
| 1 | B | 233 | ARG | 2.6 |
| 1 | A | 195 | THR | 2.6 |
| 1 | A | 201 | HIS | 2.6 |
| 1 | C | 240 | LEU | 2.6 |
| 1 | D | 19 | ALA | 2.5 |
| 1 | C | 179 | THR | 2.5 |
| 1 | B | 181 | CYS | 2.5 |
| 1 | D | 64 | GLY | 2.5 |
| 1 | A | 107 | ILE | 2.5 |
| 1 | C | 32 | LEU | 2.5 |
| 1 | D | 213 | ILE | 2.5 |
| 1 | C | 131 | VAL | 2.4 |
| 1 | D | 195 | THR | 2.4 |
| 1 | B | 179 | THR | 2.4 |
| 1 | B | 102 | ALA | 2.4 |
| 1 | A | 82 | SER | 2.4 |
| 1 | D | 224 | PHE | 2.4 |
| 1 | D | 191 | PRO | 2.4 |
| 1 | D | 32 | LEU | 2.4 |
| 1 | D | 200 | ASP | 2.3 |
| 1 | C | 195 | THR | 2.3 |
| 1 | D | 192 | ALA | 2.3 |
| 1 | D | 124 | SER | 2.3 |
| 1 | A | 101 | ALA | 2.3 |
| 1 | A | 137 | LEU | 2.3 |
| 1 | A | 7 | ILE | 2.3 |
| 1 | C | 104 | LEU | 2.3 |
| 1 | D | 65 | ALA | 2.2 |
| 1 | C | 29 | GLY | 2.2 |
| 1 | D | 78 | GLY | 2.2 |
| 1 | B | 101 | ALA | 2.2 |
| 1 | A | 204 | PHE | 2.2 |
| 1 | A | 102 | ALA | 2.2 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type | RSRZ |
|---|---|---|---|---|
| 1 | D | 142 | VAL | 2.2 |
| 1 | A | 194 | VAL | 2.2 |
| 1 | B | 193 | VAL | 2.2 |
| 1 | A | 32 | LEU | 2.2 |
| 1 | B | 116 | LEU | 2.2 |
| 1 | A | 128 | GLY | 2.2 |
| 1 | A | 121 | LEU | 2.1 |
| 1 | A | 173 | LEU | 2.1 |
| 1 | C | 136 | PHE | 2.1 |
| 1 | D | 155 | GLU | 2.1 |
| 1 | B | 199 | ASP | 2.1 |
| 1 | D | 73 | VAL | 2.1 |
| 1 | D | 71 | VAL | 2.1 |
| 1 | D | 208 | PRO | 2.1 |
| 1 | D | 205 | TRP | 2.1 |
| 1 | B | 121 | LEU | 2.1 |
| 1 | D | 86 | ALA | 2.1 |
| 1 | D | 123 | ASN | 2.1 |
| 1 | A | 4 | LEU | 2.0 |
| 1 | B | 230 | LEU | 2.0 |
| 1 | D | 38 | LEU | 2.0 |
| 1 | B | 158 | GLY | 2.0 |
| 1 | C | 108 | THR | 2.0 |
| 1 | C | 178 | TYR | 2.0 |
| 1 | C | 176 | ILE | 2.0 |

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

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

## 6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

## 6.4 Ligands (i)

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

| Mol | Type | Chain | Res | Atoms | RSCC | RSR | B-factors(Å$^2$) | Q<0.9 |
|---|---|---|---|---|---|---|---|---|
| 3 | ZN | A | 302 | 1/1 | 0.93 | 0.29 | 77,77,77,77 | 1 |
| 2 | CD | D | 301 | 1/1 | 0.94 | 0.21 | 87,87,87,87 | 1 |
| 2 | CD | A | 301 | 1/1 | 0.99 | 0.26 | 48,48,48,48 | 0 |

## 6.5 Other polymers (i)

There are no such residues in this entry.