# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 1, 2023 – 05:31 PM JST

| | | |
|---|---|---|
| PDB ID | : | 8W84 |
| Title | : | HLA-DQ2.5-alpha2 gliadin peptide in complex with DQN0344AE02 |
| Authors | : | Irie, M.; Tsushima, T.; Teranishi-Ikawa, Y.; Takahashi, N.; Ishii, S.; Okura, Y.; Fukami, T.A.; Torizawa, T. |
| Deposited on | : | 2023-08-31 |
| Resolution | : | 2.10 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 ⓘ 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:

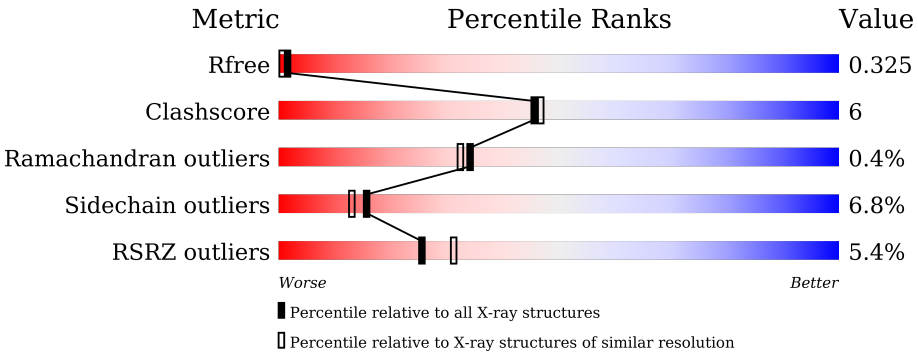
| | | |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Mogul | : | 1.8.5 (274361), CSD as541be (2020) |
| Xtriage (Phenix) | : | 1.13 |
| EDS | : | 2.36 |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| Refmac | : | 5.8.0158 |
| CCP4 | : | 7.0.044 (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 (i)

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

The reported resolution of this entry is 2.10 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 | Whole archive (#Entries) | Similar resolution (#Entries, resolution range(Å)) |
|---|---|---|
| $R_{free}$ | 130704 | 5197 (2.10-2.10) |
| Clashscore | 141614 | 5710 (2.10-2.10) |
| Ramachandran outliers | 138981 | 5647 (2.10-2.10) |
| Sidechain outliers | 138945 | 5648 (2.10-2.10) |
| RSRZ outliers | 127900 | 5083 (2.10-2.10) |

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 | 228 | 5% 75% 18% • 5% |
| 2 | B | 215 | 12% 81% 16% •• |
| 3 | C | 189 | 3% 82% 15% • • |
| 4 | D | 226 | % 69% 17% 14% |

## 2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 6415 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 DQN0344AE02 Fab heavy chain.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | A | 217 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 1594 | 1009 | 264 | 312 | 9 | | | |

- Molecule 2 is a protein called DQN0344AE02 Fab light chain.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 2 | B | 212 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 1580 | 988 | 258 | 329 | 5 | | | |

- Molecule 3 is a protein called HLA class II histocompatibility antigen, DQ alpha 1 chain.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 3 | C | 188 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 1460 | 946 | 239 | 273 | 2 | | | |

There are 7 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| C | 47 | SER | CYS | engineered mutation | UNP P01909 |
| C | 184 | LEU | - | expression tag | UNP P01909 |
| C | 185 | GLU | - | expression tag | UNP P01909 |
| C | 186 | VAL | - | expression tag | UNP P01909 |
| C | 187 | LEU | - | expression tag | UNP P01909 |
| C | 188 | PHE | - | expression tag | UNP P01909 |
| C | 189 | GLN | - | expression tag | UNP P01909 |

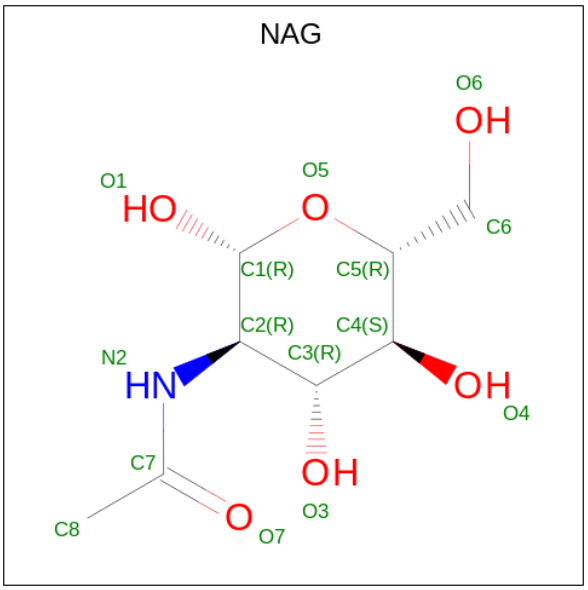
- Molecule 4 is a protein called MHC class II HLA-DQ-beta-1 - alpha2 gliadin peptide chimeric protein.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 4 | D | 195 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 1586 | 1011 | 280 | 288 | 7 | | | |

There are 23 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| D | 984 | GLY | - | linker | PDB ? |
| D | 985 | SER | - | linker | PDB ? |
| D | 986 | GLY | - | linker | PDB ? |
| D | 987 | GLY | - | linker | PDB ? |
| D | 988 | GLY | - | linker | PDB ? |
| D | 989 | GLY | - | linker | PDB ? |
| D | 990 | SER | - | linker | PDB ? |
| D | 991 | ILE | - | linker | PDB ? |
| D | 992 | GLU | - | linker | PDB ? |
| D | 993 | GLY | - | linker | PDB ? |
| D | 994 | ARG | - | linker | PDB ? |
| D | 995 | GLY | - | linker | PDB ? |
| D | 996 | SER | - | linker | PDB ? |
| D | 997 | GLY | - | linker | PDB ? |
| D | 998 | GLY | - | linker | PDB ? |
| D | 999 | GLY | - | linker | PDB ? |
| D | 1000 | SER | - | linker | PDB ? |
| D | 1191 | LEU | - | expression tag | UNP O19712 |
| D | 1192 | GLU | - | expression tag | UNP O19712 |
| D | 1193 | VAL | - | expression tag | UNP O19712 |
| D | 1194 | LEU | - | expression tag | UNP O19712 |
| D | 1195 | PHE | - | expression tag | UNP O19712 |
| D | 1196 | GLN | - | expression tag | UNP O19712 |

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |
|---|---|---|---|---|---|
| 5 | C | 1 | Total C N O<br>14 8 1 5 | 0 | 0 |
| 5 | D | 1 | Total C N O<br>14 8 1 5 | 0 | 0 |

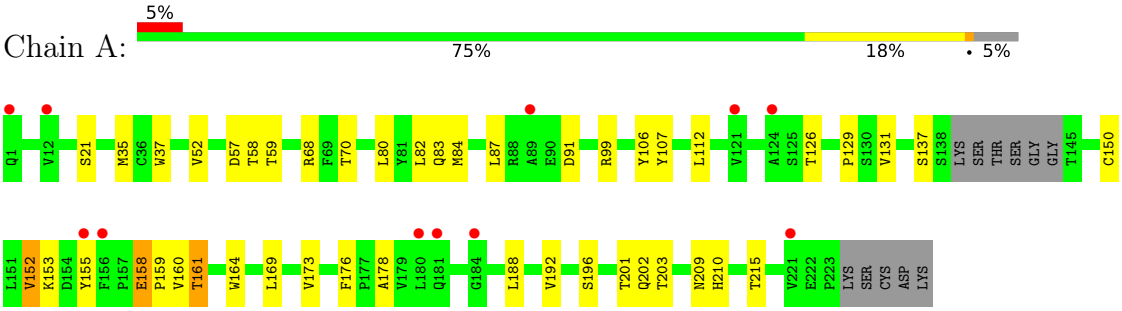
- Molecule 6 is water.

| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |
|---|---|---|---|---|---|
| 6 | A | 31 | Total O<br>31 31 | 0 | 0 |
| 6 | B | 30 | Total O<br>30 30 | 0 | 0 |
| 6 | C | 42 | Total O<br>42 42 | 0 | 0 |
| 6 | D | 64 | Total O<br>64 64 | 0 | 0 |

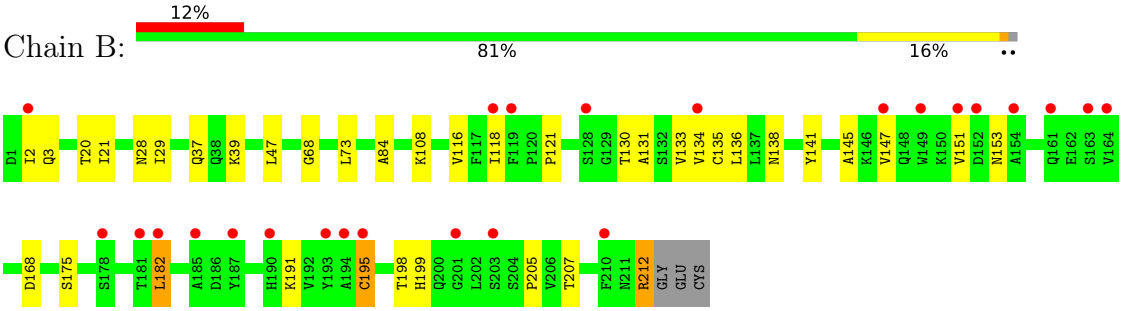
# 3 Residue-property plots (i)

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.

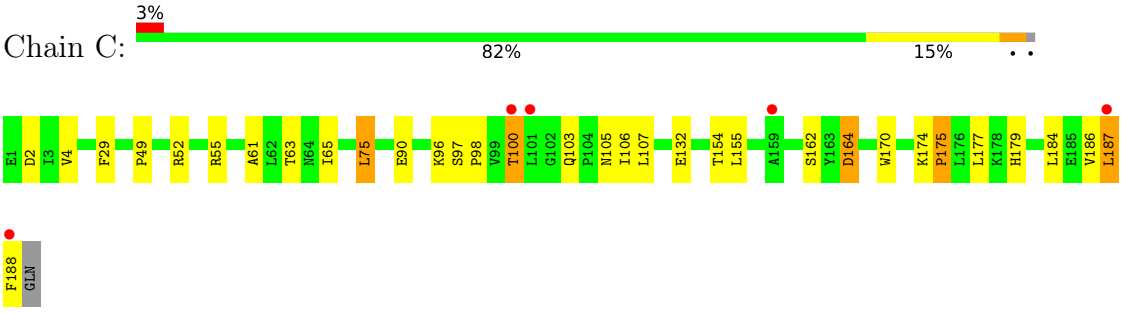
• Molecule 1: DQN0344AE02 Fab heavy chain



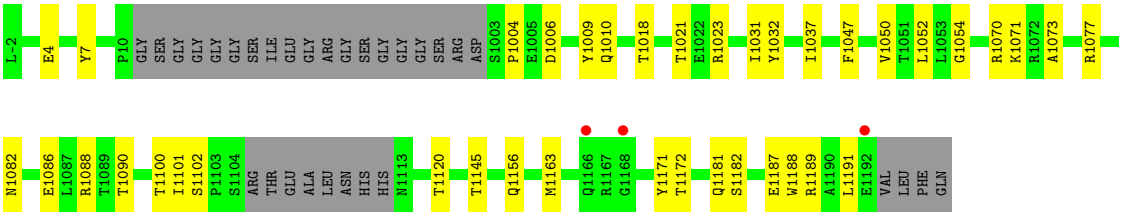

• Molecule 2: DQN0344AE02 Fab light chain




• Molecule 3: HLA class II histocompatibility antigen, DQ alpha 1 chain




• Molecule 4: MHC class II HLA-DQ-beta-1 - alpha2 gliadin peptide chimeric protein


L-2
E4
Y7
P10
GLY SER GLY GLY GLY GLY SER ILE GLU GLY ARG GLY SER GLY GLY GLY SER ARG ASP
S1003
P1004
E1005
D1006
Y1009
Q1010
T1018
T1021
E1022
R1023
I1031
Y1032
I1037
F1047
V1050
T1051
L1052
L1053
G1054
R1070
K1071
R1072
A1073
R1077
N1082
E1086
L1087
R1088
T1089
T1090
T1100
I1101
S1102
P1103
S1104
ARG THR GLU ALA LEU ASN HIS HIS
N1113
T1120
T1145
Q1156
M1163
Q1166
R1167
G1168
Y1171
T1172
Q1181
S1182
E1187
W1188
R1189
A1190
L1191
E1192
VAL LEU PHE GLN

# 4 Data and refinement statistics (i)

| Property | Value | Source |
|---|---|---|
| Space group | I 21 21 21 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 82.86Å 138.71Å 201.93Å<br>90.00° 90.00° 90.00° | Depositor |
| Resolution (Å) | 114.33 – 2.10<br>114.33 – 2.11 | Depositor<br>EDS |
| % Data completeness<br>(in resolution range) | 58.1 (114.33-2.10)<br>58.1 (114.33-2.11) | Depositor<br>EDS |
| $R_{merge}$ | 0.08 | Depositor |
| $R_{sym}$ | (Not available) | Depositor |
| $< I/\sigma(I) >$ [1] | 1.86 (at 2.10Å) | Xtriage |
| Refinement program | BUSTER 2.11.8 | Depositor |
| R, $R_{free}$ | 0.260 , 0.325<br>0.253 , 0.325 | Depositor<br>DCC |
| $R_{free}$ test set | 1916 reflections (4.87%) | wwPDB-VP |
| Wilson B-factor (Å$^2$) | 48.1 | Xtriage |
| Anisotropy | 0.085 | Xtriage |
| Bulk solvent $k_{sol}$(e/Å$^3$), $B_{sol}$(Å$^2$) | 0.30 , 41.2 | EDS |
| L-test for twinning[2] | $< \|L\| > = 0.49$, $< L^2 > = 0.32$ | Xtriage |
| Estimated twinning fraction | No twinning to report. | Xtriage |
| $F_o$,$F_c$ correlation | 0.92 | EDS |
| Total number of atoms | 6415 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 52.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.43% 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: NAG

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 | 0.40 | 0/1638 | 0.59 | 0/2246 |
| 2 | B | 0.43 | 0/1615 | 0.61 | 0/2207 |
| 3 | C | 0.43 | 0/1503 | 0.63 | 0/2059 |
| 4 | D | 0.46 | 0/1626 | 0.66 | 0/2216 |
| All | All | 0.43 | 0/6382 | 0.62 | 0/8728 |

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 | 1594 | 0 | 1465 | 29 | 0 |
| 2 | B | 1580 | 0 | 1459 | 14 | 0 |
| 3 | C | 1460 | 0 | 1386 | 23 | 0 |
| 4 | D | 1586 | 0 | 1549 | 21 | 0 |
| 5 | C | 14 | 0 | 13 | 4 | 0 |
| 5 | D | 14 | 0 | 13 | 0 | 0 |
| 6 | A | 31 | 0 | 0 | 0 | 0 |
| 6 | B | 30 | 0 | 0 | 0 | 0 |
| 6 | C | 42 | 0 | 0 | 0 | 0 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|---|---|---|---|---|---|---|
| 6 | D | 64 | 0 | 0 | 0 | 0 |
| All | All | 6415 | 0 | 5885 | 77 | 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 6.

All (77) 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:D:1172:THR:HG22 | 4:D:1187:GLU:HG2 | 1.69 | 0.73 |
| 1:A:52:VAL:HG23 | 1:A:59:THR:HG22 | 1.71 | 0.72 |
| 4:D:1073:ALA:HB1 | 4:D:1077:ARG:HE | 1.55 | 0.70 |
| 3:C:100:THR:HB | 3:C:103:GLN:CB | 2.25 | 0.66 |
| 3:C:170:TRP:CD1 | 5:C:1000:NAG:H83 | 2.31 | 0.65 |
| 3:C:107:LEU:HD13 | 3:C:155:LEU:HD21 | 1.80 | 0.62 |
| 3:C:75:LEU:HD13 | 4:D:1009:TYR:HE2 | 1.65 | 0.62 |
| 3:C:162:SER:HB2 | 3:C:179:HIS:HE1 | 1.64 | 0.61 |
| 1:A:160:VAL:HG23 | 1:A:210:HIS:HD2 | 1.66 | 0.61 |
| 3:C:170:TRP:NE1 | 5:C:1000:NAG:H83 | 2.16 | 0.60 |
| 1:A:70:THR:HB | 1:A:83:GLN:HB2 | 1.82 | 0.60 |
| 1:A:178:ALA:HB2 | 1:A:188:LEU:HD23 | 1.84 | 0.58 |
| 3:C:106:ILE:HG12 | 3:C:154:THR:HG22 | 1.87 | 0.56 |
| 1:A:84:MET:HB3 | 1:A:87:LEU:HD21 | 1.86 | 0.56 |
| 3:C:164:ASP:HB3 | 3:C:177:LEU:HD12 | 1.88 | 0.55 |
| 4:D:1070:ARG:HH12 | 4:D:1077:ARG:CZ | 2.20 | 0.55 |
| 2:B:37:GLN:HB2 | 2:B:47:LEU:HD11 | 1.90 | 0.54 |
| 2:B:198:THR:HG23 | 2:B:205:PRO:HG3 | 1.89 | 0.54 |
| 1:A:164:TRP:HD1 | 1:A:173:VAL:HG11 | 1.73 | 0.54 |
| 2:B:191:LYS:O | 2:B:212:ARG:HB2 | 2.08 | 0.54 |
| 3:C:187:LEU:HG | 4:D:1145:THR:HG22 | 1.90 | 0.54 |
| 1:A:164:TRP:HB3 | 1:A:169:LEU:HD23 | 1.90 | 0.53 |
| 3:C:187:LEU:HG | 4:D:1145:THR:CG2 | 2.38 | 0.53 |
| 1:A:160:VAL:CG2 | 1:A:210:HIS:HD2 | 2.21 | 0.53 |
| 1:A:129:PRO:HB3 | 1:A:155:TYR:HB3 | 1.92 | 0.52 |
| 2:B:121:PRO:HD3 | 2:B:133:VAL:HG22 | 1.92 | 0.52 |
| 4:D:1163:MET:HE3 | 4:D:1171:TYR:CE1 | 2.44 | 0.52 |
| 1:A:160:VAL:CG2 | 1:A:210:HIS:CD2 | 2.93 | 0.52 |
| 1:A:99:ARG:HB3 | 1:A:112:LEU:HB3 | 1.93 | 0.51 |
| 3:C:170:TRP:CD1 | 5:C:1000:NAG:C8 | 2.94 | 0.51 |
| 1:A:107:TYR:O | 4:D:1070:ARG:HD3 | 2.12 | 0.50 |
| 3:C:49:PRO:HA | 3:C:52:ARG:HD2 | 1.92 | 0.50 |

*Continued on next page...*

*Continued from previous page...*

| Atom-1 | Atom-2 | Interatomic distance (Å) | Clash overlap (Å) |
|---|---|---|---|
| 2:B:2:ILE:HD13 | 2:B:29:ILE:HG22 | 1.94 | 0.50 |
| 2:B:21:ILE:HD12 | 2:B:73:LEU:HD23 | 1.92 | 0.50 |
| 3:C:174:LYS:O | 3:C:175:PRO:O | 2.30 | 0.49 |
| 2:B:39:LYS:HG2 | 2:B:84:ALA:HB2 | 1.94 | 0.49 |
| 3:C:184:LEU:HD12 | 3:C:187:LEU:HB2 | 1.93 | 0.49 |
| 2:B:145:ALA:HB2 | 2:B:199:HIS:HD2 | 1.77 | 0.49 |
| 1:A:37:TRP:CE2 | 1:A:82:LEU:HB2 | 2.48 | 0.49 |
| 1:A:169:LEU:HD21 | 1:A:192:VAL:HG21 | 1.95 | 0.49 |
| 1:A:210:HIS:HB3 | 1:A:215:THR:HB | 1.94 | 0.48 |
| 4:D:1082:ASN:O | 4:D:1086:GLU:HG2 | 2.14 | 0.48 |
| 4:D:1047:PHE:CZ | 4:D:1071:LYS:HG3 | 2.49 | 0.47 |
| 1:A:202:GLN:HG3 | 1:A:203:THR:N | 2.30 | 0.47 |
| 4:D:1037:ILE:O | 4:D:1054:GLY:HA3 | 2.16 | 0.46 |
| 1:A:106:TYR:O | 4:D:1077:ARG:HD3 | 2.16 | 0.46 |
| 2:B:28:ASN:OD1 | 2:B:68:GLY:HA2 | 2.16 | 0.46 |
| 2:B:131:ALA:HB3 | 2:B:182:LEU:HD12 | 1.99 | 0.45 |
| 1:A:160:VAL:HG23 | 1:A:210:HIS:CD2 | 2.48 | 0.45 |
| 1:A:131:VAL:HG22 | 1:A:152:VAL:HG12 | 1.99 | 0.44 |
| 4:D:1189:ARG:HH21 | 4:D:1191:LEU:HD21 | 1.82 | 0.44 |
| 3:C:170:TRP:HZ3 | 4:D:1006:ASP:HB2 | 1.82 | 0.44 |
| 3:C:96:LYS:HD3 | 3:C:106:ILE:HD12 | 2.00 | 0.44 |
| 3:C:162:SER:HB2 | 3:C:179:HIS:CE1 | 2.50 | 0.44 |
| 4:D:1101:ILE:HG21 | 4:D:1188:TRP:HB2 | 1.99 | 0.43 |
| 4:D:1010:GLN:HB2 | 4:D:1031:ILE:HB | 2.01 | 0.43 |
| 1:A:35:MET:HB3 | 1:A:80:LEU:HD22 | 2.01 | 0.43 |
| 1:A:164:TRP:CD1 | 1:A:173:VAL:CG1 | 3.02 | 0.42 |
| 2:B:108:LYS:HA | 2:B:141:TYR:OH | 2.18 | 0.42 |
| 1:A:158:GLU:HG2 | 1:A:159:PRO:HA | 2.01 | 0.42 |
| 3:C:61:ALA:O | 3:C:65:ILE:HG12 | 2.19 | 0.42 |
| 3:C:170:TRP:HE1 | 5:C:1000:NAG:H83 | 1.80 | 0.42 |
| 1:A:164:TRP:HD1 | 1:A:173:VAL:CG1 | 2.32 | 0.41 |
| 1:A:57:ASP:O | 1:A:59:THR:HG23 | 2.21 | 0.41 |
| 3:C:29:PHE:CD2 | 4:D:1090:THR:HG21 | 2.56 | 0.41 |
| 4:D:1120:THR:HG22 | 4:D:1156:GLN:HB2 | 2.02 | 0.41 |
| 1:A:160:VAL:HG22 | 1:A:210:HIS:HA | 2.02 | 0.41 |
| 3:C:75:LEU:HD22 | 4:D:1032:TYR:HB2 | 2.03 | 0.41 |
| 2:B:116:VAL:HA | 2:B:136:LEU:O | 2.20 | 0.41 |
| 1:A:68:ARG:NH2 | 1:A:91:ASP:OD2 | 2.54 | 0.41 |
| 3:C:170:TRP:CE2 | 4:D:1004:PRO:HD2 | 2.56 | 0.41 |
| 4:D:7:TYR:CE1 | 4:D:1071:LYS:HE2 | 2.56 | 0.41 |
| 1:A:153:LYS:HZ2 | 1:A:153:LYS:HG2 | 1.67 | 0.41 |

*Continued on next page...*

*Continued from previous page...*

| Atom-1 | Atom-2 | Interatomic distance (Å) | Clash overlap (Å) |
|---|---|---|---|
| 1:A:161:THR:HG22 | 1:A:209:ASN:HB2 | 2.03 | 0.41 |
| 1:A:176:PHE:HE1 | 2:B:175:SER:O | 2.04 | 0.40 |
| 2:B:118:ILE:HD12 | 2:B:195:CYS:HB2 | 2.02 | 0.40 |
| 3:C:97:SER:HB2 | 3:C:98:PRO:HD2 | 2.03 | 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 | 213/228 (93%) | 199 (93%) | 12 (6%) | 2 (1%) | 17 12 |
| 2 | B | 210/215 (98%) | 196 (93%) | 14 (7%) | 0 | 100 100 |
| 3 | C | 186/189 (98%) | 178 (96%) | 7 (4%) | 1 (0%) | 29 26 |
| 4 | D | 189/226 (84%) | 179 (95%) | 10 (5%) | 0 | 100 100 |
| All | All | 798/858 (93%) | 752 (94%) | 43 (5%) | 3 (0%) | 34 32 |

All (3) Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 3 | C | 175 | PRO |
| 1 | A | 137 | SER |
| 1 | A | 58 | THR |

### 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 | 168/193 (87%) | 160 (95%) | 8 (5%) | 25 24 |
| 2 | B | 172/188 (92%) | 158 (92%) | 14 (8%) | 11 8 |
| 3 | C | 160/173 (92%) | 147 (92%) | 13 (8%) | 11 8 |
| 4 | D | 175/200 (88%) | 164 (94%) | 11 (6%) | 18 15 |
| All | All | 675/754 (90%) | 629 (93%) | 46 (7%) | 16 13 |

All (46) residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | A | 21 | SER |
| 1 | A | 126 | THR |
| 1 | A | 150 | CYS |
| 1 | A | 152 | VAL |
| 1 | A | 158 | GLU |
| 1 | A | 161 | THR |
| 1 | A | 196 | SER |
| 1 | A | 201 | THR |
| 2 | B | 3 | GLN |
| 2 | B | 20 | THR |
| 2 | B | 130 | THR |
| 2 | B | 134 | VAL |
| 2 | B | 135 | CYS |
| 2 | B | 138 | ASN |
| 2 | B | 147 | VAL |
| 2 | B | 151 | VAL |
| 2 | B | 153 | ASN |
| 2 | B | 168 | ASP |
| 2 | B | 182 | LEU |
| 2 | B | 195 | CYS |
| 2 | B | 207 | THR |
| 2 | B | 212 | ARG |
| 3 | C | 2 | ASP |
| 3 | C | 4 | VAL |
| 3 | C | 55 | ARG |
| 3 | C | 63 | THR |
| 3 | C | 75 | LEU |
| 3 | C | 90 | GLU |
| 3 | C | 100 | THR |
| 3 | C | 105 | ASN |
| 3 | C | 132 | GLU |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type |
|---|---|---|---|
| 3 | C | 164 | ASP |
| 3 | C | 186 | VAL |
| 3 | C | 187 | LEU |
| 3 | C | 188 | PHE |
| 4 | D | 4 | GLU |
| 4 | D | 1018 | THR |
| 4 | D | 1021 | THR |
| 4 | D | 1023 | ARG |
| 4 | D | 1050 | VAL |
| 4 | D | 1052 | LEU |
| 4 | D | 1088 | ARG |
| 4 | D | 1100 | THR |
| 4 | D | 1102 | SER |
| 4 | D | 1181 | GLN |
| 4 | D | 1182 | SER |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

#### 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

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the

expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

| Mol | Type | Chain | Res | Link | Bond lengths | | | Bond angles | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Counts | RMSZ | $\#\|Z\| > 2$ | Counts | RMSZ | $\#\|Z\| > 2$ |
| 5 | NAG | D | 2000 | 4 | 14,14,15 | 0.32 | 0 | 17,19,21 | 1.45 | 3 (17%) |
| 5 | NAG | C | 1000 | 3 | 14,14,15 | 0.25 | 0 | 17,19,21 | 0.59 | 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 |
|---|---|---|---|---|---|---|---|
| 5 | NAG | D | 2000 | 4 | - | 0/6/23/26 | 0/1/1/1 |
| 5 | NAG | C | 1000 | 3 | - | 0/6/23/26 | 0/1/1/1 |

There are no bond length outliers.

All (3) bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 5 | D | 2000 | NAG | O5-C1-C2 | 4.25 | 118.00 | 111.29 |
| 5 | D | 2000 | NAG | C1-C2-N2 | -2.62 | 106.02 | 110.49 |
| 5 | D | 2000 | NAG | C1-O5-C5 | 2.52 | 115.61 | 112.19 |

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 4 short contacts:

| Mol | Chain | Res | Type | Clashes | Symm-Clashes |
|---|---|---|---|---|---|
| 5 | C | 1000 | NAG | 4 | 0 |

## 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

## 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 | 217/228 (95%) | 0.34 | 11 (5%) 28 33 | 27, 62, 87, 104 | 0 |
| 2 | B | 212/215 (98%) | 0.67 | 25 (11%) 4 5 | 31, 59, 89, 93 | 0 |
| 3 | C | 188/189 (99%) | 0.13 | 5 (2%) 54 60 | 26, 49, 78, 84 | 0 |
| 4 | D | 195/226 (86%) | 0.02 | 3 (1%) 73 77 | 23, 39, 60, 69 | 0 |
| All | All | 812/858 (94%) | 0.30 | 44 (5%) 25 31 | 23, 51, 84, 104 | 0 |

All (44) RSRZ outliers are listed below:

| Mol | Chain | Res | Type | RSRZ |
|---|---|---|---|---|
| 2 | B | 181 | THR | 7.1 |
| 2 | B | 185 | ALA | 7.0 |
| 2 | B | 151 | VAL | 5.6 |
| 2 | B | 194 | ALA | 4.7 |
| 2 | B | 147 | VAL | 3.8 |
| 3 | C | 188 | PHE | 3.8 |
| 1 | A | 184 | GLY | 3.6 |
| 1 | A | 221 | VAL | 3.4 |
| 2 | B | 195 | CYS | 3.3 |
| 2 | B | 118 | ILE | 3.3 |
| 2 | B | 149 | TRP | 3.3 |
| 2 | B | 119 | PHE | 3.3 |
| 2 | B | 178 | SER | 3.2 |
| 2 | B | 2 | ILE | 3.0 |
| 1 | A | 124 | ALA | 3.0 |
| 2 | B | 134 | VAL | 3.0 |
| 2 | B | 187 | TYR | 3.0 |
| 2 | B | 210 | PHE | 2.9 |
| 2 | B | 154 | ALA | 2.8 |
| 2 | B | 182 | LEU | 2.8 |
| 3 | C | 187 | LEU | 2.7 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type | RSRZ |
|---|---|---|---|---|
| 2 | B | 152 | ASP | 2.6 |
| 1 | A | 155 | TYR | 2.6 |
| 2 | B | 193 | TYR | 2.6 |
| 2 | B | 128 | SER | 2.6 |
| 4 | D | 1168 | GLY | 2.6 |
| 2 | B | 203 | SER | 2.5 |
| 3 | C | 100 | THR | 2.5 |
| 1 | A | 181 | GLN | 2.5 |
| 3 | C | 159 | ALA | 2.5 |
| 3 | C | 101 | LEU | 2.4 |
| 1 | A | 89 | ALA | 2.3 |
| 2 | B | 201 | GLY | 2.3 |
| 1 | A | 180 | LEU | 2.3 |
| 2 | B | 161 | GLN | 2.2 |
| 2 | B | 190 | HIS | 2.2 |
| 1 | A | 12 | VAL | 2.2 |
| 1 | A | 121 | VAL | 2.1 |
| 4 | D | 1166 | GLN | 2.1 |
| 4 | D | 1192 | GLU | 2.0 |
| 1 | A | 156 | PHE | 2.0 |
| 2 | B | 163 | SER | 2.0 |
| 1 | A | 1 | GLN | 2.0 |
| 2 | B | 164 | VAL | 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.

*Continued on next page...*

*Continued from previous page...*

| Mol | Type | Chain | Res | Atoms | RSCC | RSR | B-factors(Å$^2$) | Q<0.9 |
|---|---|---|---|---|---|---|---|---|

| Mol | Type | Chain | Res | Atoms | RSCC | RSR | B-factors(Å$^2$) | Q<0.9 |
|---|---|---|---|---|---|---|---|---|
| 5 | NAG | D | 2000 | 14/15 | 0.86 | 0.16 | 75,76,77,77 | 0 |
| 5 | NAG | C | 1000 | 14/15 | 0.87 | 0.11 | 76,77,77,77 | 0 |

## 6.5 Other polymers (i)

There are no such residues in this entry.