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

May 17, 2020 – 10:24 pm BST

| | | |
|---|---|---|
| PDB ID | : | 6R2L |
| Title | : | NYBR1-A2-SLSKILDTV |
| Authors | : | Rizkallah, P.J.; Cole, D.K.; Sami, M. |
| Deposited on | : | 2019-03-18 |
| Resolution | : | 2.30 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 (i) symbol.

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

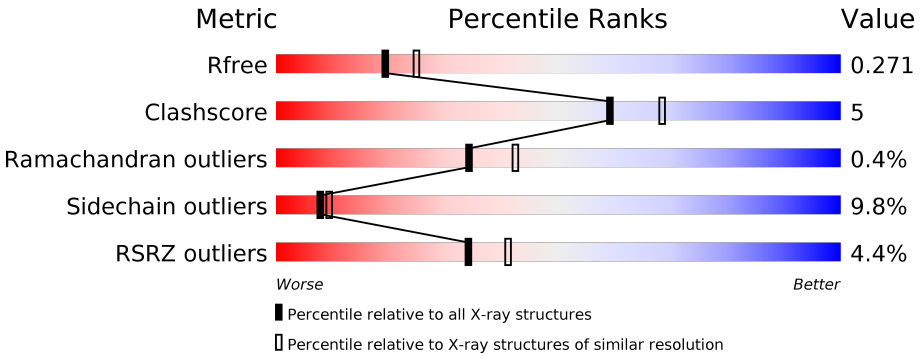
| | | |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Mogul | : | 1.8.5 (274361), CSD as541be (2020) |
| Xtriage (Phenix) | : | 1.13 |
| EDS | : | 2.11 |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| Refmac | : | 5.8.0158 |
| CCP4 | : | 7.0.044 (Gargrove) |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | 2.11 |

# 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 2.30 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 | 5042 (2.30-2.30) |
| Clashscore | 141614 | 5643 (2.30-2.30) |
| Ramachandran outliers | 138981 | 5575 (2.30-2.30) |
| Sidechain outliers | 138945 | 5575 (2.30-2.30) |
| RSRZ outliers | 127900 | 4938 (2.30-2.30) |

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 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 | 276 | 3% 88% 10% • |
| 2 | B | 100 | 83% 16% • |
| 3 | C | 9 | 78% 11% 11% |
| 4 | D | 194 | 13% 79% 15% 5% • |
| 5 | E | 247 | 2% 81% 17% • |

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

| Mol | Type | Chain | Res | Chirality | Geometry | Clashes | Electron density |
|---|---|---|---|---|---|---|---|
| 8 | GOL | A | 304 | - | - | X | - |
| 9 | SO4 | A | 308 | - | - | - | X |
| 9 | SO4 | B | 303 | - | - | - | X |
| 9 | SO4 | E | 302 | - | - | - | X |

## 2 Entry composition (i)

There are 11 unique types of molecules in this entry. The entry contains 7152 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 HLA class I histocompatibility antigen, A-2 alpha chain.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | A | 276 | Total | C | N | O | S | 0 | 1 | 0 |
| | | | 2263 | 1413 | 411 | 430 | 9 | | | |

- Molecule 2 is a protein called Beta-2-microglobulin.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 2 | B | 100 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 837 | 533 | 141 | 159 | 4 | | | |

There is a discrepancy between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| B | 0 | MET | - | initiating methionine | UNP P61769 |

- Molecule 3 is a protein called SER-LEU-SER-LYS-ILE-LEU-ASP-THR-VAL.

| Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 3 | C | 9 | Total | C | N | O | 0 | 0 | 0 |
| | | | 68 | 43 | 10 | 15 | | | |

- Molecule 4 is a protein called T cell receptor alpha variable 22,Human nkt tcr alpha chain.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 4 | D | 194 | Total | C | N | O | S | 0 | 2 | 0 |
| | | | 1535 | 959 | 258 | 309 | 9 | | | |

There are 26 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| D | 1 | MET | - | initiating methionine | UNP A0A0B4J277 |
| D | 54 | TRP | THR | conflict | UNP A0A0B4J277 |
| D | 57 | GLU | ASN | conflict | UNP A0A0B4J277 |

*Continued on next page...*

*Continued from previous page...*

| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| D | 91 | GLY | - | linker | UNP A0A0B4J277 |
| D | 92 | GLY | - | linker | UNP A0A0B4J277 |
| D | 93 | ASN | - | linker | UNP A0A0B4J277 |
| D | 94 | ASP | - | linker | UNP A0A0B4J277 |
| D | 95 | TRP | - | linker | UNP A0A0B4J277 |
| D | 96 | ASN | - | linker | UNP A0A0B4J277 |
| D | 97 | THR | - | linker | UNP A0A0B4J277 |
| D | 98 | ASP | - | linker | UNP A0A0B4J277 |
| D | 99 | LYS | - | linker | UNP A0A0B4J277 |
| D | 100 | LEU | - | linker | UNP A0A0B4J277 |
| D | 101 | ILE | - | linker | UNP A0A0B4J277 |
| D | 102 | PHE | - | linker | UNP A0A0B4J277 |
| D | 103 | GLY | - | linker | UNP A0A0B4J277 |
| D | 104 | THR | - | linker | UNP A0A0B4J277 |
| D | 105 | GLY | - | linker | UNP A0A0B4J277 |
| D | 106 | THR | - | linker | UNP A0A0B4J277 |
| D | 107 | ARG | - | linker | UNP A0A0B4J277 |
| D | 108 | LEU | - | linker | UNP A0A0B4J277 |
| D | 109 | GLN | - | linker | UNP A0A0B4J277 |
| D | 110 | VAL | - | linker | UNP A0A0B4J277 |
| D | 111 | PHE | - | linker | UNP A0A0B4J277 |
| D | 112 | PRO | - | linker | UNP A0A0B4J277 |
| D | 113 | ASN | - | linker | UNP A0A0B4J277 |

- Molecule 5 is a protein called T cell receptor beta variable 11-2,Human nkt tcr beta chain.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 5 | E | 247 | Total | C | N | O | S | 0 | 1 | 0 |
| | | | 1981 | 1264 | 339 | 373 | 5 | | | |

There are 22 discrepancies between the modelled and reference sequences:

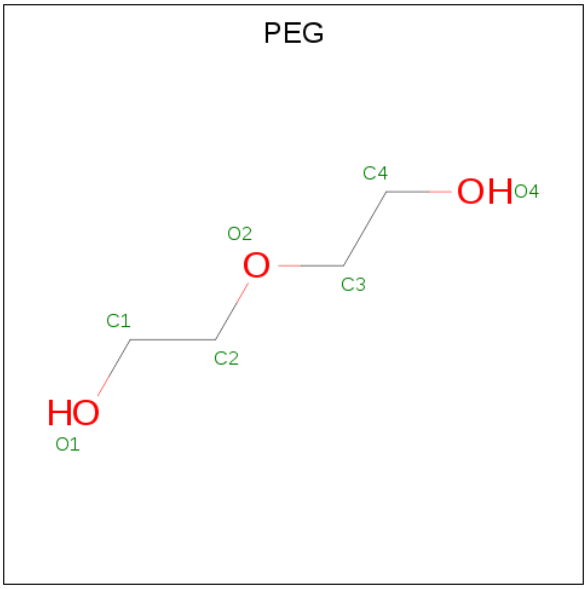
| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| E | 28 | PHE | SER | conflict | UNP A0A584 |
| E | 29 | SER | GLY | conflict | UNP A0A584 |
| E | 31 | PRO | ALA | conflict | UNP A0A584 |
| E | 51 | GLY | GLN | conflict | UNP A0A584 |
| E | 52 | GLY | ASN | conflict | UNP A0A584 |
| E | 53 | TRP | ASN | conflict | UNP A0A584 |
| E | 54 | PRO | GLY | conflict | UNP A0A584 |
| E | 55 | GLY | VAL | conflict | UNP A0A584 |
| E | 90 | LEU | VAL | conflict | UNP A0A584 |

*Continued on next page...*

*Continued from previous page...*

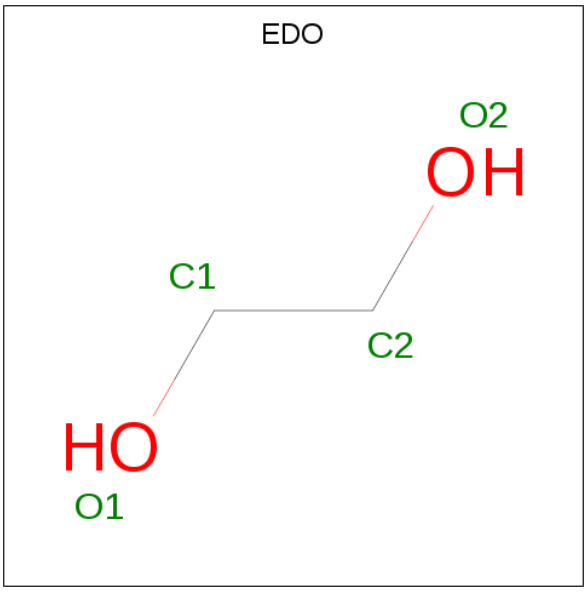
| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| E | 97 | PRO | - | linker | UNP A0A584 |
| E | 98 | LEU | - | linker | UNP A0A584 |
| E | 99 | ASP | - | linker | UNP A0A584 |
| E | 100 | VAL | - | linker | UNP A0A584 |
| E | 101 | SER | - | linker | UNP A0A584 |
| E | 102 | ILE | - | linker | UNP A0A584 |
| E | 103 | SER | - | linker | UNP A0A584 |
| E | 104 | SER | - | linker | UNP A0A584 |
| E | 105 | TYR | - | linker | UNP A0A584 |
| E | 106 | ASN | - | linker | UNP A0A584 |
| E | 109 | PHE | TYR | conflict | UNP K7N5M4 |
| E | 119 | LEU | THR | conflict | UNP K7N5M4 |
| E | 208 | ASP | ASN | conflict | UNP K7N5M4 |

- Molecule 6 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



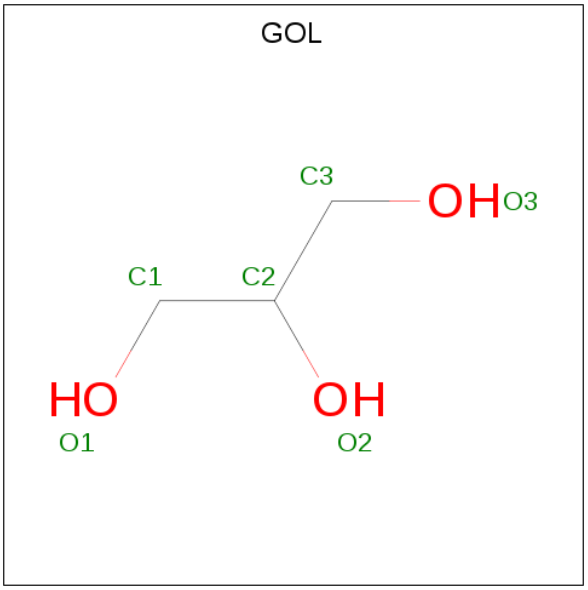

| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |
|---|---|---|---|---|---|
| 6 | A | 1 | Total C O<br>7 4 3 | 0 | 0 |

- Molecule 7 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).




| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |
|---|---|---|---|---|---|
| 7 | A | 1 | Total C O<br>4 2 2 | 0 | 0 |
| 7 | A | 1 | Total C O<br>4 2 2 | 0 | 0 |

- Molecule 8 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



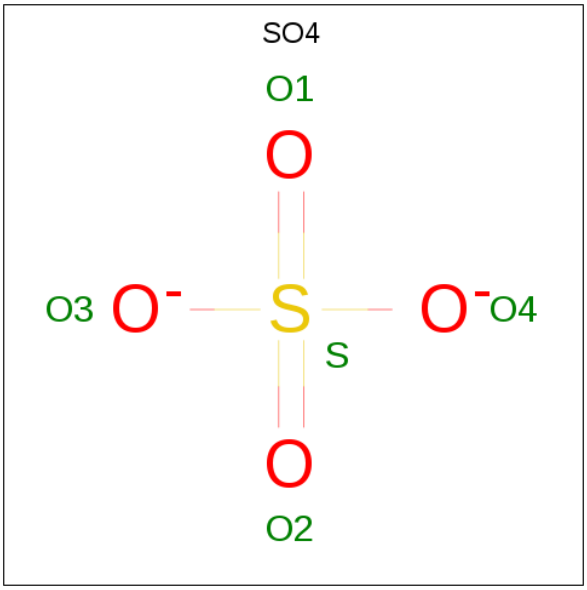

| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |
|---|---|---|---|---|---|
| 8 | A | 1 | Total C O<br>6 3 3 | 0 | 0 |
| 8 | A | 1 | Total C O<br>6 3 3 | 0 | 0 |

*Continued on next page...*

*Continued from previous page...*

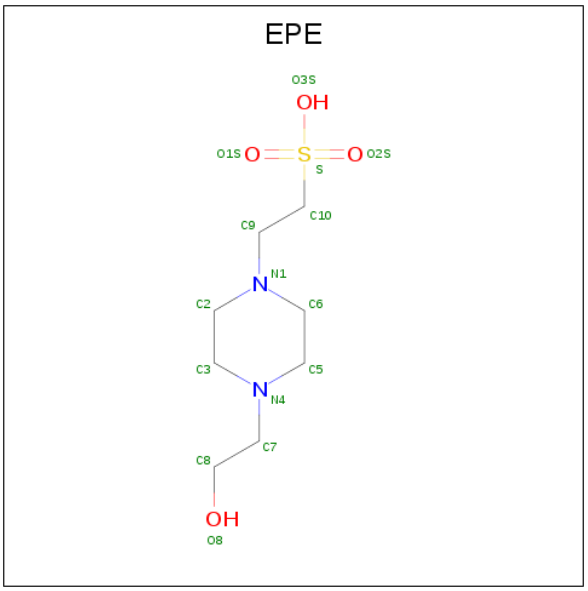
| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |
|---|---|---|---|---|---|
| 8 | E | 1 | Total C O<br>6 3 3 | 0 | 0 |

- Molecule 9 is SULFATE ION (three-letter code: SO4) (formula: $O_4S$).




| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |
|---|---|---|---|---|---|
| 9 | A | 1 | Total O S<br>5 4 1 | 0 | 0 |
| 9 | A | 1 | Total O S<br>5 4 1 | 0 | 0 |
| 9 | A | 1 | Total O S<br>5 4 1 | 0 | 0 |
| 9 | A | 1 | Total O S<br>5 4 1 | 0 | 0 |
| 9 | B | 1 | Total O S<br>5 4 1 | 0 | 0 |
| 9 | B | 1 | Total O S<br>5 4 1 | 0 | 0 |
| 9 | D | 1 | Total O S<br>5 4 1 | 0 | 0 |
| 9 | E | 1 | Total O S<br>5 4 1 | 0 | 0 |

- Molecule 10 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: $C_8H_{18}N_2O_4S$).




| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|---|---|---|
| 10 | B | 1 | Total | C | N | O | S | 0 | 0 |
| | | | 15 | 8 | 2 | 4 | 1 | | |

- Molecule 11 is water.

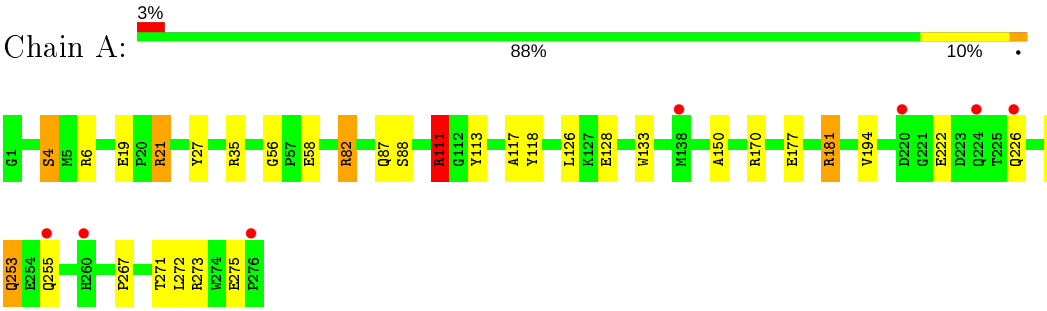
| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|
| 11 | A | 141 | Total | O | 0 | 0 |
| | | | 141 | 141 | | |
| 11 | B | 54 | Total | O | 0 | 0 |
| | | | 54 | 54 | | |
| 11 | C | 9 | Total | O | 0 | 0 |
| | | | 9 | 9 | | |
| 11 | D | 68 | Total | O | 0 | 0 |
| | | | 68 | 68 | | |
| 11 | E | 108 | Total | O | 0 | 0 |
| | | | 108 | 108 | | |

# 3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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.

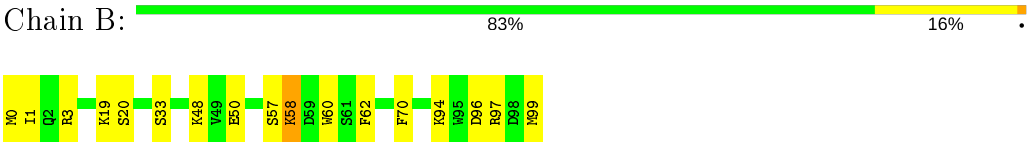
- Molecule 1: HLA class I histocompatibility antigen, A-2 alpha chain

Chain A:



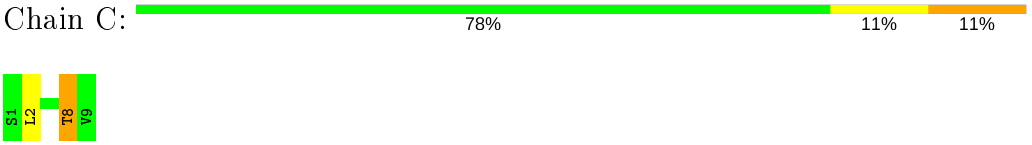

- Molecule 2: Beta-2-microglobulin

Chain B:



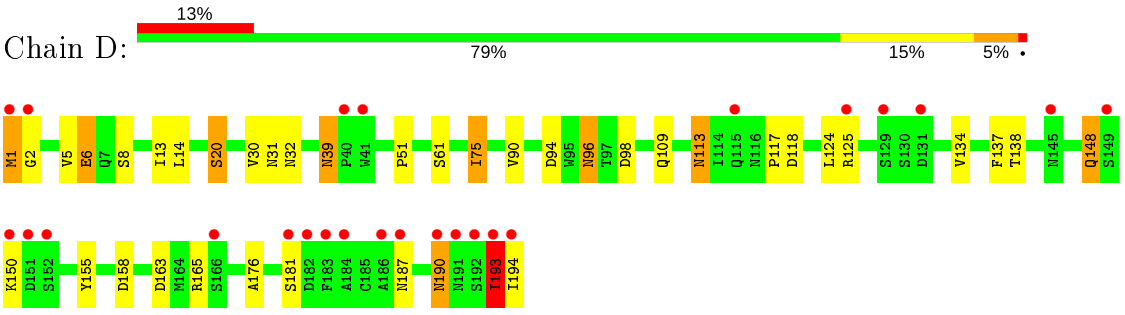

- Molecule 3: SER-LEU-SER-LYS-ILE-LEU-ASP-THR-VAL

Chain C:



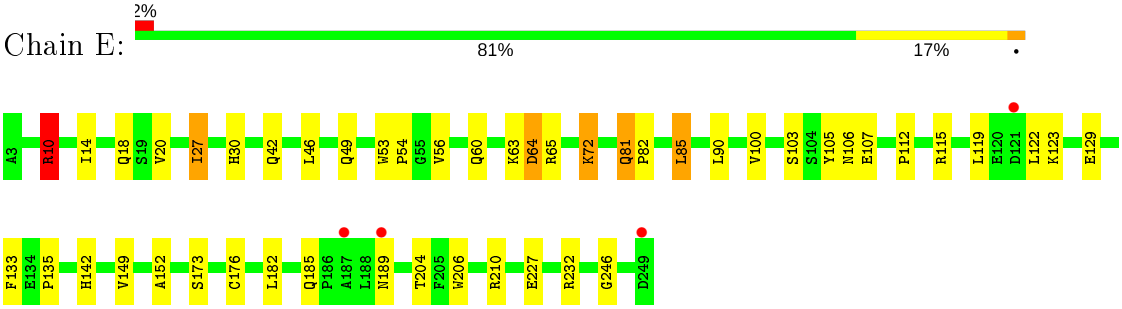

- Molecule 4: T cell receptor alpha variable 22,Human nkt tcr alpha chain

Chain D:




- Molecule 5: T cell receptor beta variable 11-2,Human nkt tcr beta chain

Chain E:


2%
81%
17%
A3
R10
I14
Q18
S19
V20
I27
H30
Q42
L46
Q49
W53
P54
G55
V56
Q60
K63
D64
R65
K72
Q81
P82
L85
L90
V100
S103
S104
Y105
N106
E107
P112
R115
L119
E120
D121
L122
K123
E129
F133
E134
P135
H142
V149
A152
S173
C176
L182
Q185
P186
A187
L188
N189
T204
F205
W206
R210
E227
R232
G246
D249

# 4 Data and refinement statistics (i)

| Property | Value | Source |
|---|---|---|
| Space group | P 1 21 1 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 52.21Å 99.33Å 111.44Å<br>90.00° 90.50° 90.00° | Depositor |
| Resolution (Å) | 52.21 – 2.30<br>52.21 – 2.30 | Depositor<br>EDS |
| % Data completeness<br>(in resolution range) | 99.7 (52.21-2.30)<br>99.7 (52.21-2.30) | Depositor<br>EDS |
| $R_{merge}$ | 0.21 | Depositor |
| $R_{sym}$ | (Not available) | Depositor |
| $< I/\sigma(I) >$ [1] | 1.62 (at 2.29Å) | Xtriage |
| Refinement program | REFMAC 5.8.0238 | Depositor |
| R, $R_{free}$ | 0.214 , 0.265<br>0.219 , 0.271 | Depositor<br>DCC |
| $R_{free}$ test set | 2435 reflections (4.83%) | wwPDB-VP |
| Wilson B-factor (Å$^2$) | 34.5 | Xtriage |
| Anisotropy | 0.206 | Xtriage |
| Bulk solvent $k_{sol}$(e/Å$^3$), $B_{sol}$(Å$^2$) | 0.29 , 31.5 | EDS |
| L-test for twinning[2] | $< \|L\| > = 0.48$, $< L^2 > = 0.31$ | Xtriage |
| Estimated twinning fraction | 0.035 for h,-k,-l | Xtriage |
| $F_o$,$F_c$ correlation | 0.94 | EDS |
| Total number of atoms | 7152 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 43.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.09% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

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

# 5 Model quality

## 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: PEG, GOL, EPE, EDO, SO4

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 | 0.80 | 1/2329 (0.0%) | 1.01 | 6/3161 (0.2%) |
| 2 | B | 0.85 | 0/860 | 1.00 | 2/1162 (0.2%) |
| 3 | C | 1.01 | 0/67 | 1.38 | 1/88 (1.1%) |
| 4 | D | 0.83 | 1/1569 (0.1%) | 0.98 | 0/2133 |
| 5 | E | 0.76 | 0/2039 | 0.99 | 2/2778 (0.1%) |
| All | All | 0.81 | 2/6864 (0.0%) | 1.00 | 11/9322 (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 | 2 |
| 4 | D | 0 | 2 |
| 5 | E | 0 | 1 |
| All | All | 0 | 5 |

All (2) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|---|---|---|---|---|---|---|---|
| 4 | D | 6 | GLU | CD-OE2 | 8.39 | 1.34 | 1.25 |
| 1 | A | 19 | GLU | CD-OE2 | 6.33 | 1.32 | 1.25 |

The worst 5 of 11 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 1 | A | 21 | ARG | NE-CZ-NH1 | 7.78 | 124.19 | 120.30 |
| 1 | A | 21 | ARG | NE-CZ-NH2 | -7.49 | 116.56 | 120.30 |
| 5 | E | 10 | ARG | NE-CZ-NH1 | -6.06 | 117.27 | 120.30 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 1 | A | 6 | ARG | NE-CZ-NH1 | 5.87 | 123.23 | 120.30 |
| 2 | B | 97 | ARG | NE-CZ-NH1 | 5.86 | 123.23 | 120.30 |

There are no chirality outliers.

All (5) planarity outliers are listed below:

| Mol | Chain | Res | Type | Group |
|---|---|---|---|---|
| 1 | A | 27 | TYR | Peptide |
| 1 | A | 4 | SER | Peptide |
| 4 | D | 117 | PRO | Peptide |
| 4 | D | 2 | GLY | Peptide |
| 5 | E | 246 | GLY | 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 | 2263 | 0 | 2108 | 16 | 0 |
| 2 | B | 837 | 0 | 803 | 11 | 0 |
| 3 | C | 68 | 0 | 78 | 0 | 0 |
| 4 | D | 1535 | 0 | 1446 | 21 | 0 |
| 5 | E | 1981 | 0 | 1905 | 22 | 0 |
| 6 | A | 7 | 0 | 10 | 2 | 0 |
| 7 | A | 8 | 0 | 12 | 0 | 0 |
| 8 | A | 12 | 0 | 16 | 4 | 0 |
| 8 | E | 6 | 0 | 8 | 0 | 0 |
| 9 | A | 20 | 0 | 0 | 0 | 0 |
| 9 | B | 10 | 0 | 0 | 0 | 0 |
| 9 | D | 5 | 0 | 0 | 0 | 0 |
| 9 | E | 5 | 0 | 0 | 0 | 0 |
| 10 | B | 15 | 0 | 18 | 0 | 0 |
| 11 | A | 141 | 0 | 0 | 1 | 0 |
| 11 | B | 54 | 0 | 0 | 0 | 0 |
| 11 | C | 9 | 0 | 0 | 1 | 0 |
| 11 | D | 68 | 0 | 0 | 0 | 0 |
| 11 | E | 108 | 0 | 0 | 1 | 0 |
| All | All | 7152 | 0 | 6404 | 62 | 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.

The worst 5 of 62 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:98:ASP:H | 5:E:49:GLN:HE22 | 1.23 | 0.85 |
| 4:D:134:VAL:HG21 | 5:E:149:VAL:HG11 | 1.69 | 0.74 |
| 2:B:58:LYS:H | 2:B:58:LYS:CE | 2.08 | 0.66 |
| 4:D:96:ASN:H | 4:D:96:ASN:HD22 | 1.45 | 0.65 |
| 4:D:124:LEU:HD22 | 5:E:149:VAL:HG12 | 1.77 | 0.65 |

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 | 275/276 (100%) | 266 (97%) | 9 (3%) | 0 | 100 | 100 |
| 2 | B | 98/100 (98%) | 97 (99%) | 1 (1%) | 0 | 100 | 100 |
| 3 | C | 7/9 (78%) | 7 (100%) | 0 | 0 | 100 | 100 |
| 4 | D | 194/194 (100%) | 181 (93%) | 11 (6%) | 2 (1%) | 15 | 17 |
| 5 | E | 246/247 (100%) | 230 (94%) | 15 (6%) | 1 (0%) | 34 | 42 |
| All | All | 820/826 (99%) | 781 (95%) | 36 (4%) | 3 (0%) | 34 | 42 |

All (3) Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 4 | D | 193 | ILE |
| 5 | E | 72 | LYS |
| 4 | D | 181 | SER |

### 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 | 233/232 (100%) | 217 (93%) | 16 (7%) | 15 20 |
| 2 | B | 95/95 (100%) | 87 (92%) | 8 (8%) | 11 13 |
| 3 | C | 9/9 (100%) | 7 (78%) | 2 (22%) | 1 1 |
| 4 | D | 176/174 (101%) | 156 (89%) | 20 (11%) | 5 6 |
| 5 | E | 216/215 (100%) | 190 (88%) | 26 (12%) | 5 5 |
| All | All | 729/725 (101%) | 657 (90%) | 72 (10%) | 8 9 |

5 of 72 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 4 | D | 32 | ASN |
| 4 | D | 150 | LYS |
| 5 | E | 182 | LEU |
| 4 | D | 39 | ASN |
| 4 | D | 96 | ASN |

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

| Mol | Chain | Res | Type |
|---|---|---|---|
| 4 | D | 38 | GLN |
| 4 | D | 39 | ASN |
| 5 | E | 38 | GLN |
| 4 | D | 4 | GLN |
| 4 | D | 32 | 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 carbohydrates in this entry.

## 5.6 Ligand geometry (i)

15 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 | $\#\|Z\| > 2$ |
| 8 | GOL | A | 305 | - | 5,5,5 | 0.19 | 0 | 5,5,5 | 0.46 | 0 |
| 7 | EDO | A | 302 | - | 3,3,3 | 0.04 | 0 | 2,2,2 | 0.13 | 0 |
| 9 | SO4 | A | 307 | - | 4,4,4 | 0.36 | 0 | 6,6,6 | 0.07 | 0 |
| 9 | SO4 | A | 308 | - | 4,4,4 | 0.38 | 0 | 6,6,6 | 0.07 | 0 |
| 9 | SO4 | D | 201 | - | 4,4,4 | 0.36 | 0 | 6,6,6 | 0.08 | 0 |
| 6 | PEG | A | 301 | - | 6,6,6 | 0.23 | 0 | 5,5,5 | 0.23 | 0 |
| 9 | SO4 | B | 303 | - | 4,4,4 | 0.37 | 0 | 6,6,6 | 0.07 | 0 |
| 9 | SO4 | A | 306 | - | 4,4,4 | 0.36 | 0 | 6,6,6 | 0.06 | 0 |
| 9 | SO4 | E | 302 | - | 4,4,4 | 0.37 | 0 | 6,6,6 | 0.05 | 0 |
| 8 | GOL | A | 304 | - | 5,5,5 | 0.21 | 0 | 5,5,5 | 0.64 | 0 |
| 9 | SO4 | B | 302 | - | 4,4,4 | 0.36 | 0 | 6,6,6 | 0.10 | 0 |
| 7 | EDO | A | 303 | - | 3,3,3 | 0.15 | 0 | 2,2,2 | 0.25 | 0 |
| 9 | SO4 | A | 309 | - | 4,4,4 | 0.35 | 0 | 6,6,6 | 0.05 | 0 |
| 8 | GOL | E | 301 | - | 5,5,5 | 0.14 | 0 | 5,5,5 | 0.45 | 0 |
| 10 | EPE | B | 301 | - | 15,15,15 | 2.09 | 1 (6%) | 18,20,20 | 1.54 | 3 (16%) |

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 |
|---|---|---|---|---|---|---|---|
| 8 | GOL | A | 305 | - | - | 0/4/4/4 | - |
| 7 | EDO | A | 302 | - | - | 1/1/1/1 | - |
| 6 | PEG | A | 301 | - | - | 2/4/4/4 | - |
| 8 | GOL | A | 304 | - | - | 2/4/4/4 | - |
| 7 | EDO | A | 303 | - | - | 0/1/1/1 | - |
| 8 | GOL | E | 301 | - | - | 4/4/4/4 | - |
| 10 | EPE | B | 301 | - | - | 4/9/19/19 | 0/1/1/1 |

All (1) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|---|---|---|---|---|---|---|---|
| 10 | B | 301 | EPE | C10-S | -7.63 | 1.66 | 1.77 |

All (3) bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 10 | B | 301 | EPE | O2S-S-C10 | 3.35 | 110.95 | 106.92 |
| 10 | B | 301 | EPE | O3S-S-C10 | 3.08 | 110.75 | 105.77 |
| 10 | B | 301 | EPE | C5-C6-N1 | -2.79 | 104.92 | 110.64 |

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

| Mol | Chain | Res | Type | Atoms |
|---|---|---|---|---|
| 8 | A | 304 | GOL | O1-C1-C2-C3 |
| 8 | E | 301 | GOL | C1-C2-C3-O3 |
| 10 | B | 301 | EPE | S-C10-C9-N1 |
| 10 | B | 301 | EPE | N4-C7-C8-O8 |
| 8 | E | 301 | GOL | O1-C1-C2-O2 |

There are no ring outliers.

2 monomers are involved in 6 short contacts:

| Mol | Chain | Res | Type | Clashes | Symm-Clashes |
|---|---|---|---|---|---|
| 6 | A | 301 | PEG | 2 | 0 |
| 8 | A | 304 | GOL | 4 | 0 |

## 5.7 Other polymers (i)

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.

# 6 Fit of model and data

## 6.1 Protein, DNA and RNA chains

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

| Mol | Chain | Analysed | <RSRZ> | #RSRZ>2 | OWAB(Å$^2$) | Q<0.9 |
|---|---|---|---|---|---|---|
| 1 | A | 276/276 (100%) | 0.05 | 7 (2%) 57 64 | 21, 34, 81, 99 | 0 |
| 2 | B | 100/100 (100%) | -0.24 | 0 100 100 | 23, 32, 57, 74 | 0 |
| 3 | C | 9/9 (100%) | 0.16 | 0 100 100 | 22, 22, 28, 28 | 0 |
| 4 | D | 194/194 (100%) | 0.61 | 25 (12%) 3 5 | 19, 47, 93, 119 | 0 |
| 5 | E | 247/247 (100%) | 0.08 | 4 (1%) 72 77 | 22, 44, 66, 98 | 0 |
| All | All | 826/826 (100%) | 0.16 | 36 (4%) 34 41 | 19, 40, 81, 119 | 0 |

The worst 5 of 36 RSRZ outliers are listed below:

| Mol | Chain | Res | Type | RSRZ |
|---|---|---|---|---|
| 4 | D | 151 | ASP | 9.0 |
| 4 | D | 193 | ILE | 6.5 |
| 4 | D | 194 | ILE | 5.6 |
| 4 | D | 191 | ASN | 5.6 |
| 4 | D | 150 | LYS | 5.5 |

## 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 carbohydrates in this entry.

## 6.4 Ligands

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 |
|---|---|---|---|---|---|---|---|---|
| 6 | PEG | A | 301 | 7/7 | 0.60 | 0.36 | 61,73,77,77 | 0 |
| 9 | SO4 | A | 308 | 5/5 | 0.65 | 0.60 | 148,150,154,155 | 0 |
| 10 | EPE | B | 301 | 15/15 | 0.67 | 0.25 | 72,88,110,115 | 0 |
| 7 | EDO | A | 303 | 4/4 | 0.68 | 0.27 | 55,59,59,61 | 0 |
| 9 | SO4 | E | 302 | 5/5 | 0.72 | 0.51 | 142,144,146,149 | 0 |
| 8 | GOL | A | 305 | 6/6 | 0.75 | 0.34 | 58,61,62,63 | 0 |
| 9 | SO4 | B | 303 | 5/5 | 0.75 | 0.51 | 132,133,137,138 | 0 |
| 9 | SO4 | A | 309 | 5/5 | 0.82 | 0.34 | 104,111,115,121 | 0 |
| 8 | GOL | E | 301 | 6/6 | 0.82 | 0.33 | 63,65,70,71 | 0 |
| 8 | GOL | A | 304 | 6/6 | 0.82 | 0.30 | 44,49,50,51 | 0 |
| 9 | SO4 | B | 302 | 5/5 | 0.89 | 0.32 | 79,91,96,102 | 0 |
| 9 | SO4 | A | 307 | 5/5 | 0.90 | 0.28 | 94,97,101,101 | 0 |
| 7 | EDO | A | 302 | 4/4 | 0.92 | 0.19 | 31,33,35,39 | 0 |
| 9 | SO4 | A | 306 | 5/5 | 0.92 | 0.22 | 91,96,100,104 | 0 |
| 9 | SO4 | D | 201 | 5/5 | 0.97 | 0.17 | 66,67,72,76 | 0 |

## 6.5 Other polymers (i)

There are no such residues in this entry.