# WORLDWIDE wwPDB PROTEIN DATA BANK

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

Oct 23, 2021 – 09:48 AM EDT

|  |  |  |
|---|---|---|
| PDB ID | : | 1QSE |
| Title | : | STRUCTURE OF HUMAN A6-TCR BOUND TO HLA-A2 COMPLEXED WITH ALTERED HTLV-1 TAX PEPTIDE V7R |
| Authors | : | Ding, Y.H.; Baker, B.M.; Garboczi, D.N.; Biddison, W.E.; Wiley, D.C. |
| Deposited on | : | 1999-06-21 |
| Resolution | : | 2.80 Å(reported) |

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

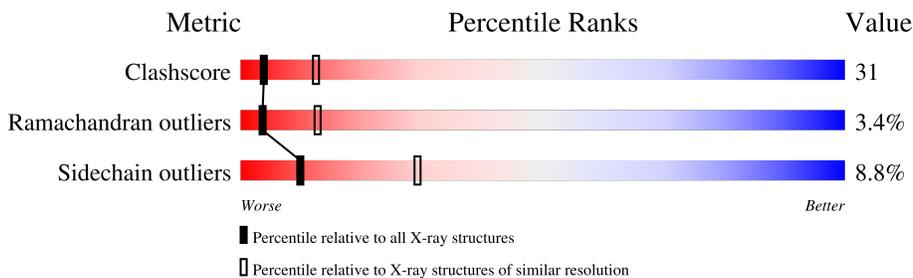|  |  |  |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Xtriage (Phenix) | : | NOT EXECUTED |
| EDS | : | NOT EXECUTED |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | 2.23.2 |

# 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 2.80 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.

| Metric | Percentile Ranks | Value |
|---|---|---|
| Clashscore | | 31 |
| Ramachandran outliers | | 3.4% |
| Sidechain outliers | | 8.8% |

*Worse*        *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

| Metric | Whole archive (#Entries) | Similar resolution (#Entries, resolution range(Å)) |
|---|---|---|
| Clashscore | 141614 | 3569 (2.80-2.80) |
| Ramachandran outliers | 138981 | 3498 (2.80-2.80) |
| Sidechain outliers | 138945 | 3500 (2.80-2.80) |

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%

Note EDS was not executed.

| Mol | Chain | Length | Quality of chain |
|---|---|---|---|
| 1 | A | 274 | 63%   32%   5% |
| 2 | B | 100 | 58%   36%   6% |
| 3 | C | 9 | 33%   56%   11% |
| 4 | D | 200 | 44%   50%   6% |
| 5 | E | 243 | 47%   45%   8% |

# 2  Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 6500 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 PROTEIN (MHC class I HLA-A).

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | A | 274 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 2201 | 1377 | 398 | 417 | 9 | | | |

- Molecule 2 is a protein called PROTEIN (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 | ALA | engineered mutation | UNP P61769 |

- Molecule 3 is a protein called Tax Peptide V7R.

| Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 3 | C | 9 | Total | C | N | O | 0 | 0 | 0 |
| | | | 81 | 57 | 12 | 12 | | | |

- Molecule 4 is a protein called PROTEIN (human T-Cell receptor).

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 4 | D | 200 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 1502 | 941 | 246 | 309 | 6 | | | |

- Molecule 5 is a protein called PROTEIN (human T-Cell receptor).

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 5 | E | 243 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 1879 | 1182 | 330 | 359 | 8 | | | |

WORLDWIDE wwPDB PROTEIN DATA BANK
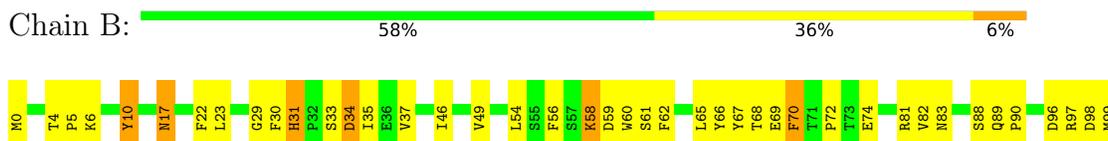
# 3 Residue-property plots ⓘ

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. 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. 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.

Note EDS was not executed.
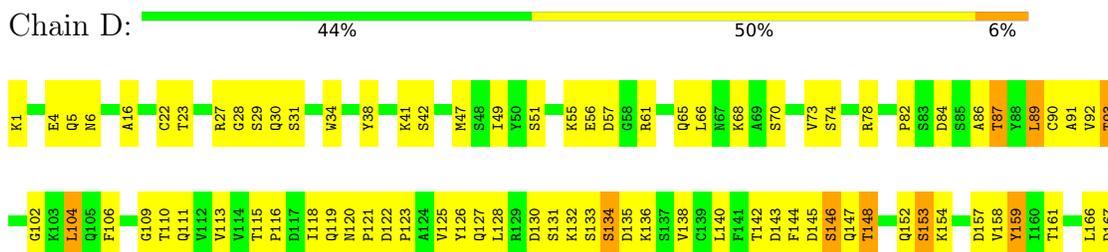
• Molecule 1: PROTEIN (MHC class I HLA-A)
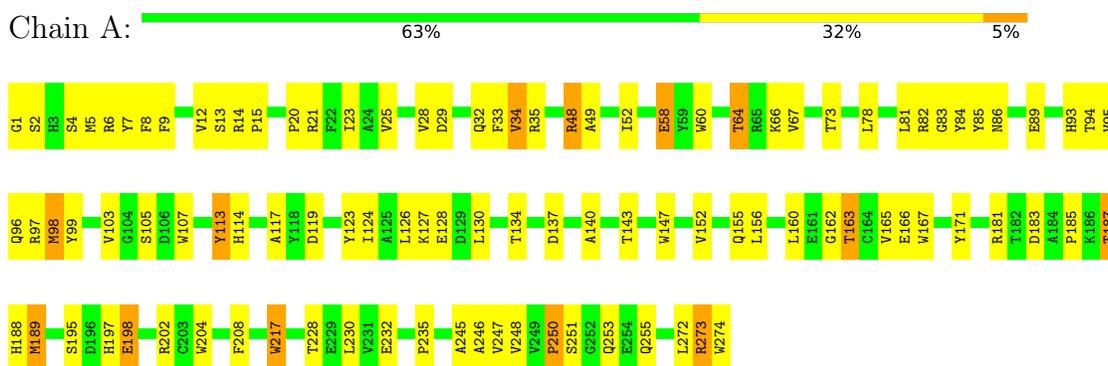
Chain A:



• Molecule 2: PROTEIN (beta-2 microglobulin)

Chain B:



• Molecule 3: Tax Peptide V7R

Chain C:



• Molecule 4: PROTEIN (human T-Cell receptor)

Chain D:

M168 F173 K174 S175 V179 A180 W181 S182 N183 K184 S185 D186 C189 A190 F193 N194 N195 S196 I197 I198 P199 E200 D201 T202 F203 F204 P205 S206

• Molecule 5: PROTEIN (human T-Cell receptor)

Chain E:

47%        45%        8%

G3 Q6 T7 P8 Q11 V12 L13 K14 T15 L21 Q25 Y31 M32 S33 W34 Q37 D38 P39 G40 M41 S49 T55 N66 V67 T71 D74 L77 R78 L79 A82 A83 P84 S85 Q86 T87 S88 V89 Y90 F91 C92 R95 L98 P103 E105

Q106 Y107 F108 G109 P110 G111 T112 R113 L114 T115 V116 T116A L119 K120 N121 V122 F123 P124 P125 E126 V127 A128 V129 F130 E131 P132 S133 T140 Q141 K142 A143 T144 L145 V146 C147 L148 F152 Y153 P154 D155 E158 L159 S160 W161 W162 V163 N164 V168 H169 S170 G171 V172 S173 T174

L179 K180 E181 Q182 P183 A184 L185 N186 D187 S188 R189 Y190 A191 L192 S193 S194 R195 L196 R197 V198 S199 A200 T201 F202 W203 Q204 N205 P206 R207 N208 H209 F210 R211 C212 Q213 V214 Q215 F216 Y217 G218 L219 S220 E221 N222 D223 E224 W225 T226 Q227 D228 R229 A230 K231 P232 V233 T234 Q235 I236 V237 S238

A239 G243 D246

# 4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

| Property | Value | Source |
|---|---|---|
| Space group | C 1 2 1 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 224.86Å 48.89Å 94.46Å<br>90.00° 90.49° 90.00° | Depositor |
| Resolution (Å) | 25.00 – 2.80 | Depositor |
| % Data completeness<br>(in resolution range) | (Not available) (25.00-2.80) | Depositor |
| $R_{merge}$ | 0.07 | Depositor |
| $R_{sym}$ | (Not available) | Depositor |
| Refinement program | CNS | Depositor |
| R, $R_{free}$ | 0.243 , 0.290 | Depositor |
| Estimated twinning fraction | No twinning to report. | Xtriage |
| Total number of atoms | 6500 | wwPDB-VP |
| Average B, all atoms (Å²) | 66.0 | wwPDB-VP |

## 5 Model quality (i)

### 5.1 Standard geometry (i)

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

| Mol | Chain | Bond lengths | | Bond angles | |
|---|---|---|---|---|---|
| | | RMSZ | #$\|Z\|$ >5 | RMSZ | #$\|Z\|$ >5 |
| 1 | A | 0.52 | 0/2264 | 0.73 | 0/3076 |
| 2 | B | 0.61 | 0/860 | 0.75 | 0/1162 |
| 3 | C | 0.78 | 0/84 | 1.05 | 0/112 |
| 4 | D | 0.52 | 0/1535 | 0.76 | 0/2093 |
| 5 | E | 0.54 | 0/1932 | 0.78 | 1/2637 (0.0%) |
| All | All | 0.54 | 0/6675 | 0.76 | 1/9080 (0.0%) |

There are no bond length outliers.

All (1) bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 5 | E | 187 | ASP | N-CA-C | -5.46 | 96.26 | 111.00 |

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 | 2201 | 0 | 2037 | 105 | 0 |
| 2 | B | 837 | 0 | 803 | 52 | 0 |
| 3 | C | 81 | 0 | 83 | 8 | 0 |
| 4 | D | 1502 | 0 | 1363 | 118 | 0 |
| 5 | E | 1879 | 0 | 1755 | 151 | 0 |
| All | All | 6500 | 0 | 6041 | 386 | 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 31.

The worst 5 of 386 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:119:GLN:O | 4:D:121:PRO:HD3 | 1.54 | 1.07 |
| 4:D:197:ILE:HD13 | 4:D:197:ILE:H | 1.06 | 1.06 |
| 5:E:55:THR:HG21 | 5:E:67:VAL:HG13 | 1.40 | 1.04 |
| 1:A:188:HIS:HB3 | 1:A:204:TRP:HB2 | 1.41 | 1.02 |
| 1:A:202:ARG:HH12 | 2:B:99:MET:HG2 | 1.24 | 1.02 |

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 | 272/274 (99%) | 245 (90%) | 25 (9%) | 2 (1%) | 22 | 53 |
| 2 | B | 98/100 (98%) | 91 (93%) | 4 (4%) | 3 (3%) | 4 | 14 |
| 3 | C | 7/9 (78%) | 5 (71%) | 1 (14%) | 1 (14%) | 0 | 0 |
| 4 | D | 198/200 (99%) | 167 (84%) | 19 (10%) | 12 (6%) | 1 | 4 |
| 5 | E | 241/243 (99%) | 205 (85%) | 26 (11%) | 10 (4%) | 3 | 9 |
| All | All | 816/826 (99%) | 713 (87%) | 75 (9%) | 28 (3%) | 3 | 13 |

5 of 28 Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | A | 195 | SER |
| 4 | D | 6 | ASN |
| 4 | D | 132 | LYS |
| 4 | D | 134 | SER |
| 4 | D | 135 | ASP |

### 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 | 222/230 (96%) | 202 (91%) | 20 (9%) | 9 | 28 |
| 2 | B | 95/95 (100%) | 90 (95%) | 5 (5%) | 22 | 54 |
| 3 | C | 8/8 (100%) | 8 (100%) | 0 | 100 | 100 |
| 4 | D | 160/178 (90%) | 146 (91%) | 14 (9%) | 10 | 29 |
| 5 | E | 198/208 (95%) | 177 (89%) | 21 (11%) | 6 | 20 |
| All | All | 683/719 (95%) | 623 (91%) | 60 (9%) | 10 | 29 |

5 of 60 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 4 | D | 57 | ASP |
| 5 | E | 185 | LEU |
| 4 | D | 159 | TYR |
| 5 | E | 159 | LEU |
| 5 | E | 236 | ILE |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 4 | D | 127 | GLN |
| 5 | E | 11 | GLN |
| 5 | E | 6 | GLN |
| 5 | E | 37 | GLN |
| 1 | A | 218 | GLN |

### 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 ⓘ

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

## 6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

## 6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

## 6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.