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

Feb 4, 2024 – 04:39 AM EST

| | | |
|---|---|---|
| PDB ID | : | 1PMA |
| Title | : | PROTEASOME FROM THERMOPLASMA ACIDOPHILUM |
| Authors | : | Loewe, J.; Stock, D.; Jap, B.; Zwickl, P.; Baumeister, W.; Huber, R. |
| Deposited on | : | 1994-12-19 |
| Resolution | : | 3.40 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 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:

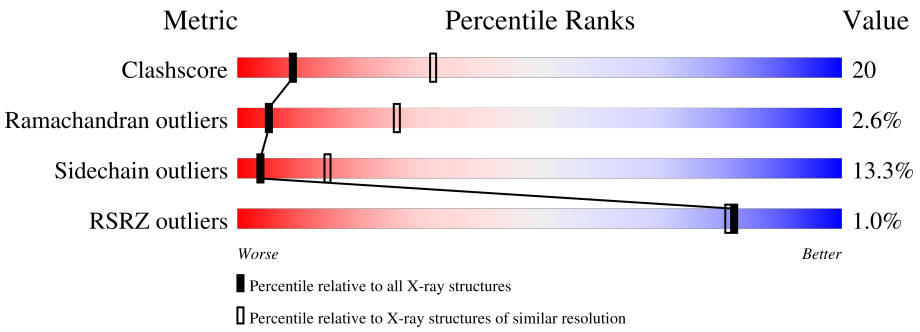
| | | |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Xtriage (Phenix) | : | 1.13 |
| EDS | : | 2.36 |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| Refmac | : | 5.8.0158 |
| CCP4 | : | 7.0.044 (Gargrove) |
| Ideal geometry (proteins) | : | Engh & 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 3.40 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)) |
|---|---|---|
| Clashscore | 141614 | 1055 (3.48-3.32) |
| Ramachandran outliers | 138981 | 1038 (3.48-3.32) |
| Sidechain outliers | 138945 | 1038 (3.48-3.32) |
| RSRZ outliers | 127900 | 2173 (3.50-3.30) |

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 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 | 233 | 2%, 48%, 42%, 5%, 5% |
| 1 | C | 233 | 51%, 35%, 8%, 5% |
| 1 | D | 233 | 2%, 49%, 39%, 7%, 5% |
| 1 | E | 233 | 2%, 49%, 37%, 9%, 5% |
| 1 | F | 233 | 2%, 45%, 42%, 7%, 5% |
| 1 | G | 233 | 2%, 44%, 44%, 7%, 5% |
| 1 | H | 233 | 2%, 48%, 41%, 5%, 5% |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Length | Quality of chain |
|---|---|---|---|
| 1 | I | 233 | 2% 49% 38% 7% 5% |
| 1 | J | 233 | 3% 51% 37% 6% 5% |
| 1 | K | 233 | 2% 54% 35% 6% 5% |
| 1 | L | 233 | % 51% 37% 6% 5% |
| 1 | M | 233 | 2% 51% 39% 5% 5% |
| 1 | N | 233 | 2% 55% 32% 7% 5% |
| 1 | O | 233 | % 50% 36% 9% 5% |
| 2 | 1 | 211 | 54% 34% 9% |
| 2 | 2 | 211 | 55% 33% 8% |
| 2 | B | 211 | 51% 37% 9% |
| 2 | P | 211 | 57% 31% 8% |
| 2 | Q | 211 | 58% 31% 7% |
| 2 | R | 211 | 51% 37% 8% |
| 2 | S | 211 | 52% 36% 8% |
| 2 | T | 211 | 59% 31% 7% |
| 2 | U | 211 | 54% 35% 8% |
| 2 | V | 211 | 53% 36% 8% |
| 2 | W | 211 | 54% 34% 9% |
| 2 | X | 211 | 57% 33% 6% |
| 2 | Y | 211 | 58% 30% 8% |
| 2 | Z | 211 | 55% 35% 7% |

# 2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 56294 atoms, of which 10402 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 PROTEASOME.

| Mol | Chain | Residues | Atoms | | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | H | N | O | S | | | |
| 1 | A | 221 | 2113 | 1092 | 393 | 290 | 335 | 3 | 0 | 0 | 0 |
| 1 | C | 221 | 2113 | 1092 | 393 | 290 | 335 | 3 | 0 | 0 | 0 |
| 1 | D | 221 | 2113 | 1092 | 393 | 290 | 335 | 3 | 0 | 0 | 0 |
| 1 | E | 221 | 2113 | 1092 | 393 | 290 | 335 | 3 | 0 | 0 | 0 |
| 1 | F | 221 | 2113 | 1092 | 393 | 290 | 335 | 3 | 0 | 0 | 0 |
| 1 | G | 221 | 2113 | 1092 | 393 | 290 | 335 | 3 | 0 | 0 | 0 |
| 1 | H | 221 | 2113 | 1092 | 393 | 290 | 335 | 3 | 0 | 0 | 0 |
| 1 | I | 221 | 2113 | 1092 | 393 | 290 | 335 | 3 | 0 | 0 | 0 |
| 1 | J | 221 | 2113 | 1092 | 393 | 290 | 335 | 3 | 0 | 0 | 0 |
| 1 | K | 221 | 2113 | 1092 | 393 | 290 | 335 | 3 | 0 | 0 | 0 |
| 1 | L | 221 | 2113 | 1092 | 393 | 290 | 335 | 3 | 0 | 0 | 0 |
| 1 | M | 221 | 2113 | 1092 | 393 | 290 | 335 | 3 | 0 | 0 | 0 |
| 1 | N | 221 | 2113 | 1092 | 393 | 290 | 335 | 3 | 0 | 0 | 0 |
| 1 | O | 221 | 2113 | 1092 | 393 | 290 | 335 | 3 | 0 | 0 | 0 |

- Molecule 2 is a protein called PROTEASOME.

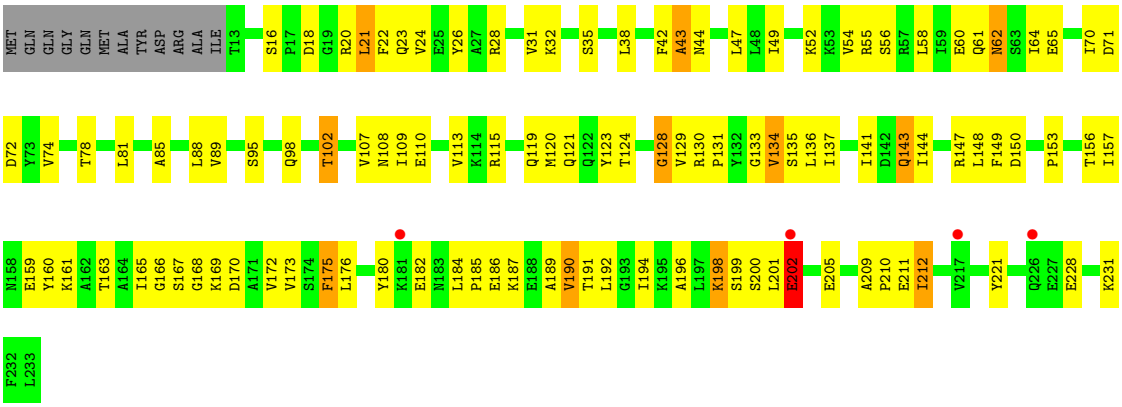
| Mol | Chain | Residues | Atoms | | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | H | N | O | S | | | |
| 2 | B | 203 | 1908 | 985 | 350 | 264 | 298 | 11 | 0 | 0 | 0 |
| 2 | P | 203 | 1908 | 985 | 350 | 264 | 298 | 11 | 0 | 0 | 0 |
| 2 | Q | 203 | 1908 | 985 | 350 | 264 | 298 | 11 | 0 | 0 | 0 |
| 2 | R | 203 | 1908 | 985 | 350 | 264 | 298 | 11 | 0 | 0 | 0 |
| 2 | S | 203 | 1908 | 985 | 350 | 264 | 298 | 11 | 0 | 0 | 0 |
| 2 | T | 203 | 1908 | 985 | 350 | 264 | 298 | 11 | 0 | 0 | 0 |
| 2 | U | 203 | 1908 | 985 | 350 | 264 | 298 | 11 | 0 | 0 | 0 |
| 2 | V | 203 | 1908 | 985 | 350 | 264 | 298 | 11 | 0 | 0 | 0 |
| 2 | W | 203 | 1908 | 985 | 350 | 264 | 298 | 11 | 0 | 0 | 0 |
| 2 | X | 203 | 1908 | 985 | 350 | 264 | 298 | 11 | 0 | 0 | 0 |
| 2 | Y | 203 | 1908 | 985 | 350 | 264 | 298 | 11 | 0 | 0 | 0 |
| 2 | Z | 203 | 1908 | 985 | 350 | 264 | 298 | 11 | 0 | 0 | 0 |
| 2 | 1 | 203 | 1908 | 985 | 350 | 264 | 298 | 11 | 0 | 0 | 0 |
| 2 | 2 | 203 | 1908 | 985 | 350 | 264 | 298 | 11 | 0 | 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.

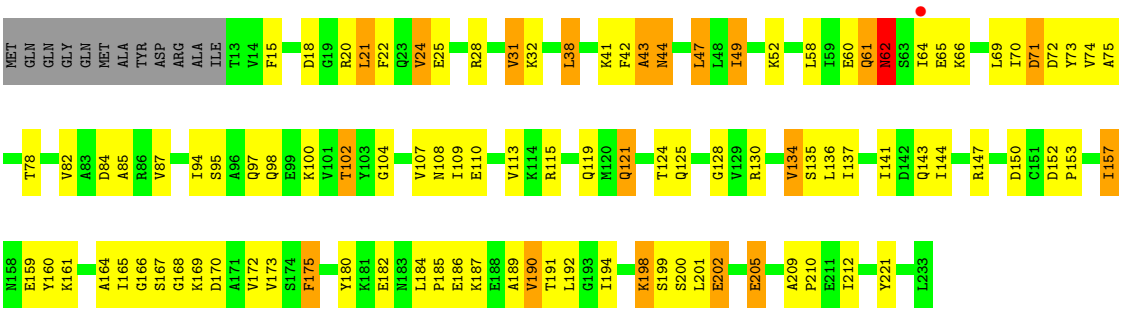
• Molecule 1: PROTEASOME

Chain A: 2% 48% 42% 5% 5%



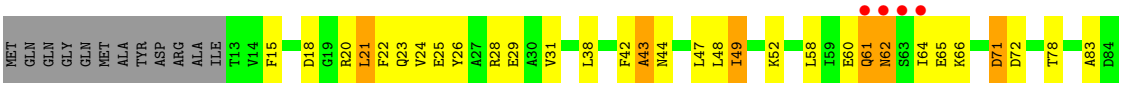

• Molecule 1: PROTEASOME

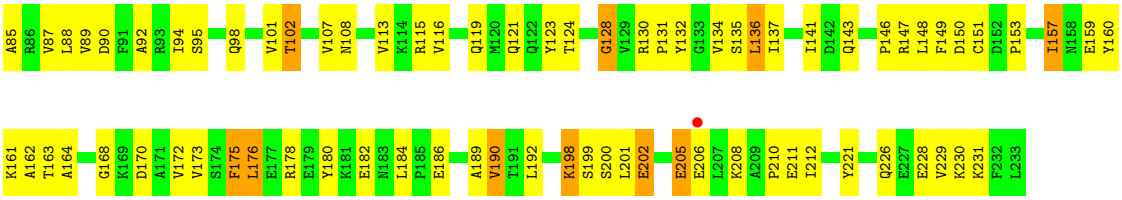
Chain C: 51% 35% 8% 5%




• Molecule 1: PROTEASOME

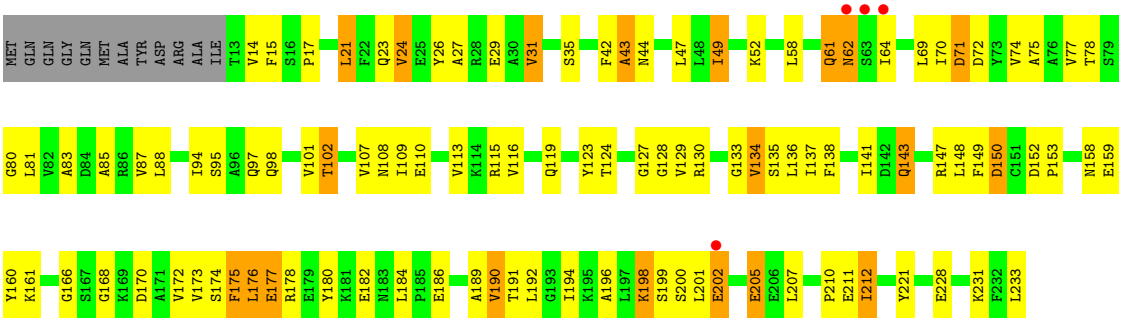
Chain D: 2% 49% 39% 7% 5%





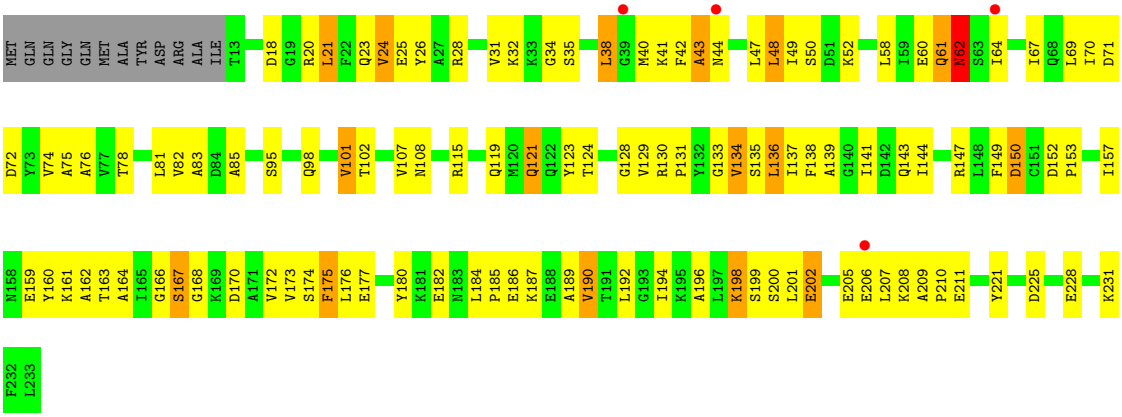

• Molecule 1: PROTEASOME

Chain E:



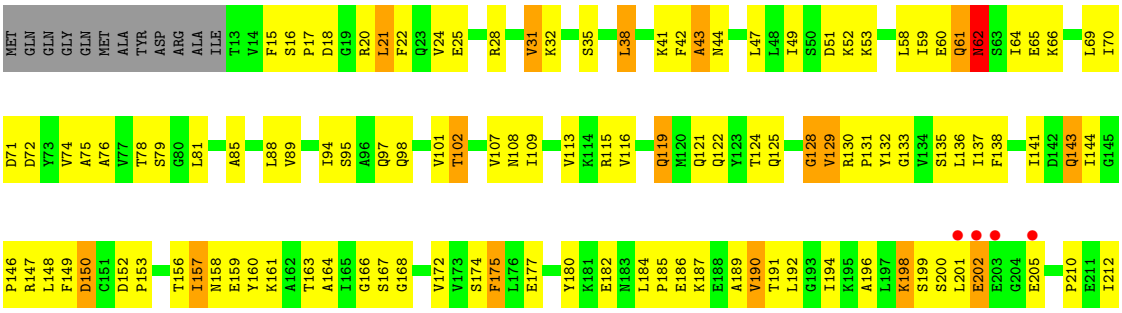

• Molecule 1: PROTEASOME

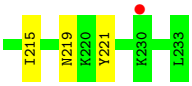
Chain F:




• Molecule 1: PROTEASOME

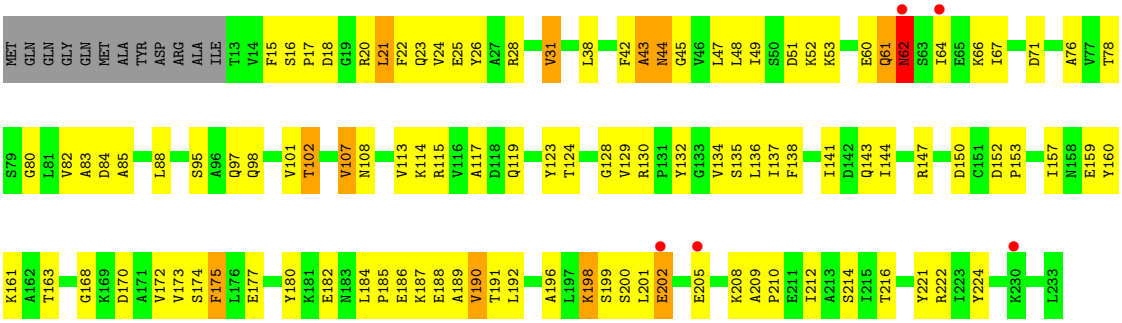
Chain G:





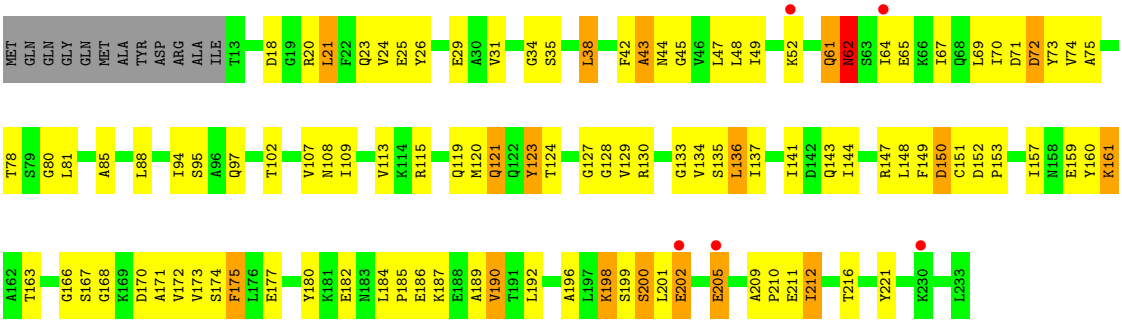

• Molecule 1: PROTEASOME

Chain H:



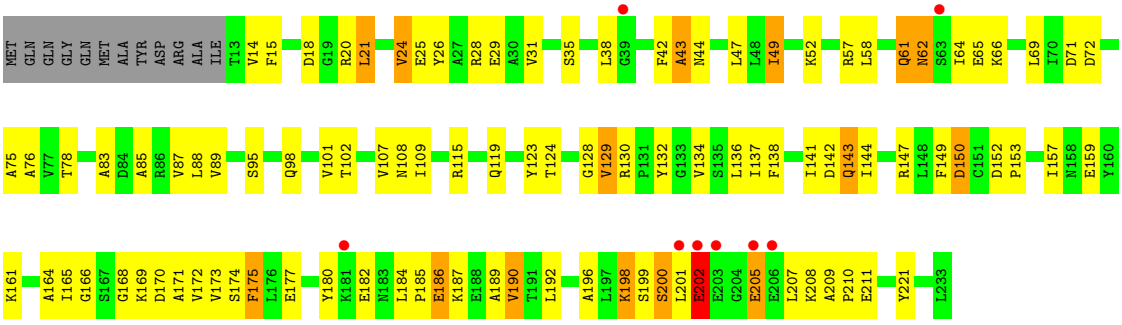

• Molecule 1: PROTEASOME

Chain I:



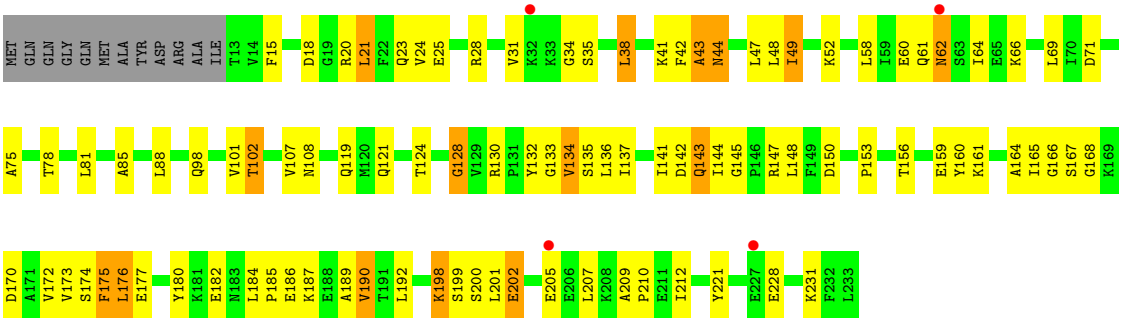

• Molecule 1: PROTEASOME

Chain J:



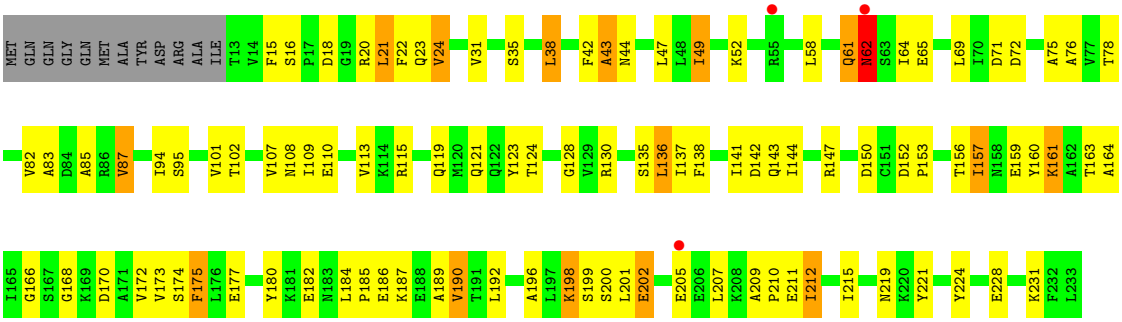

• Molecule 1: PROTEASOME

Chain K:



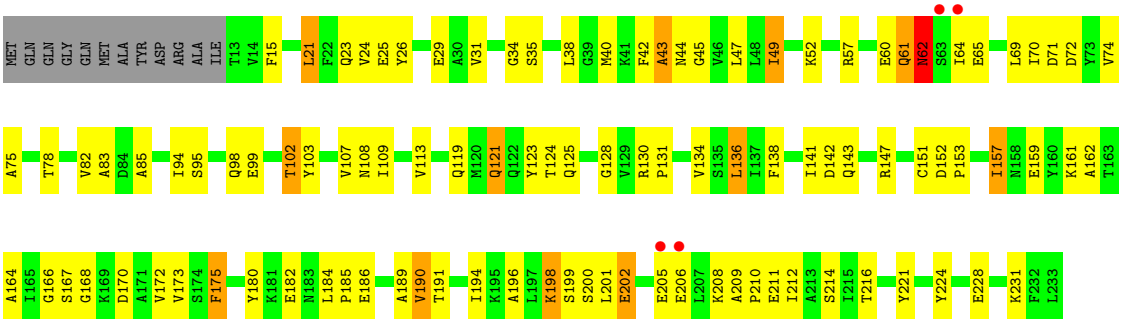

- Molecule 1: PROTEASOME

Chain L:



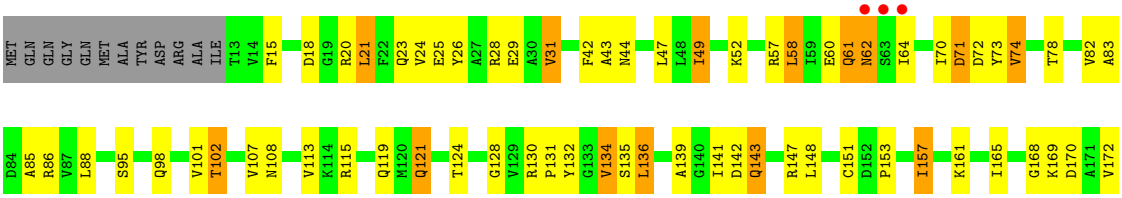

- Molecule 1: PROTEASOME

Chain M:



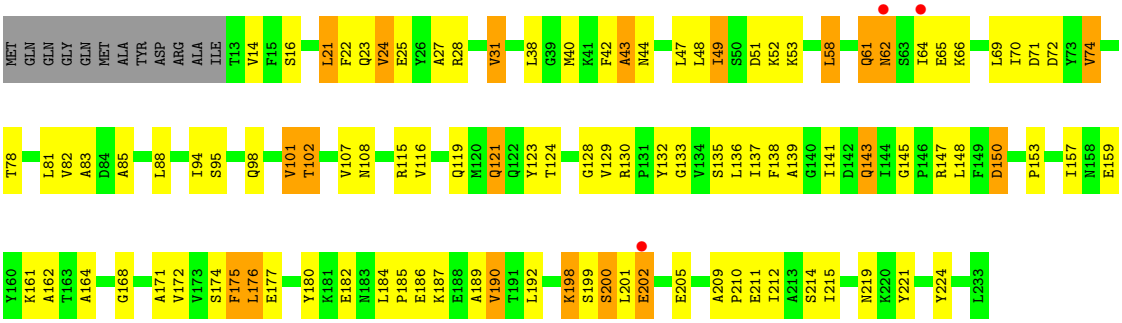

- Molecule 1: PROTEASOME

Chain N:



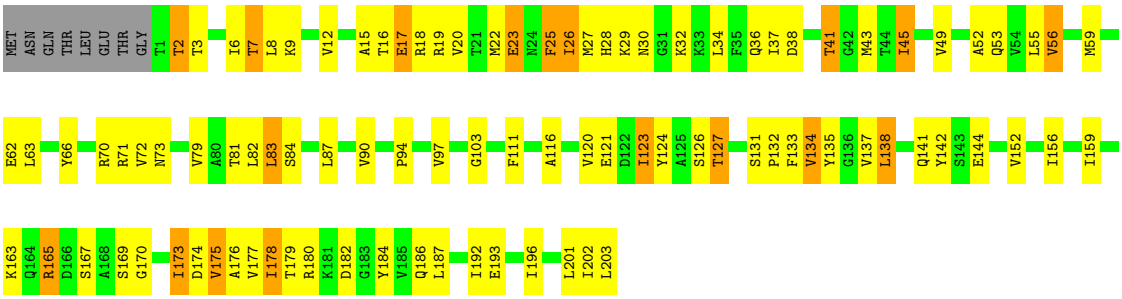
• Molecule 1: PROTEASOME

Chain O:



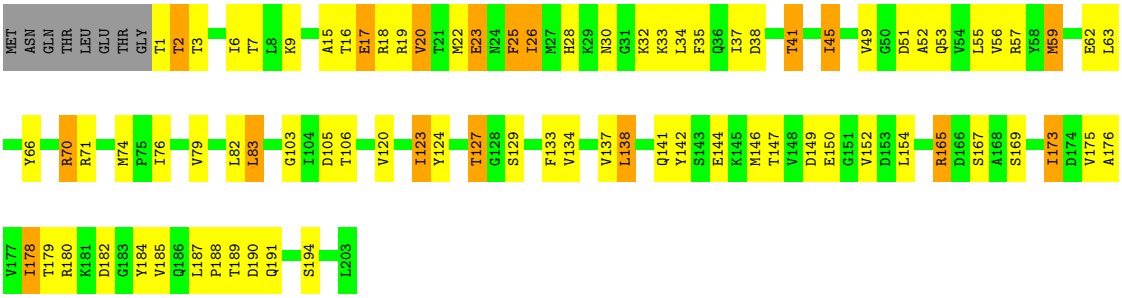

• Molecule 2: PROTEASOME

Chain B:



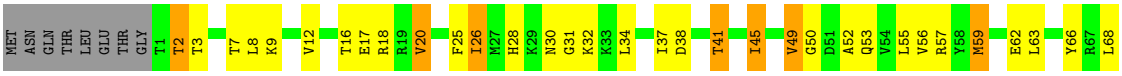

• Molecule 2: PROTEASOME

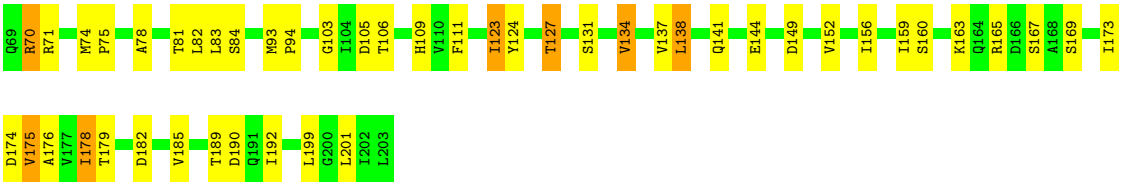
Chain P:




• Molecule 2: PROTEASOME

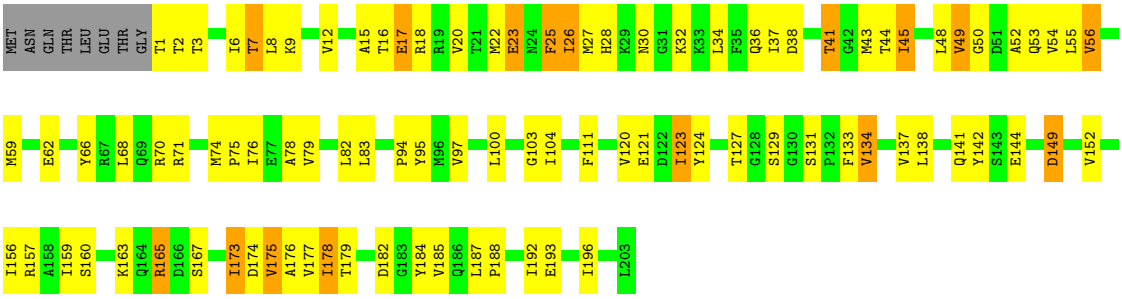
Chain Q:





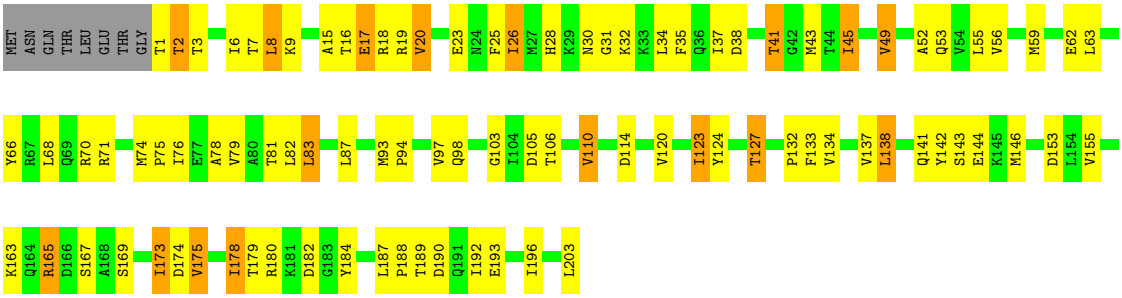

- Molecule 2: PROTEASOME

Chain R:



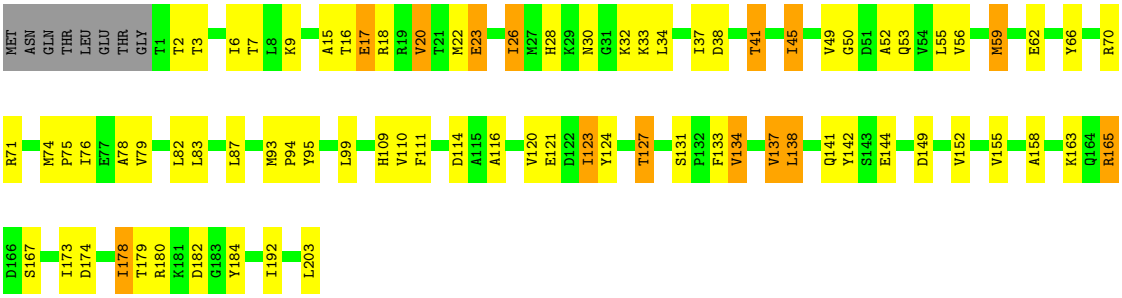

- Molecule 2: PROTEASOME

Chain S:



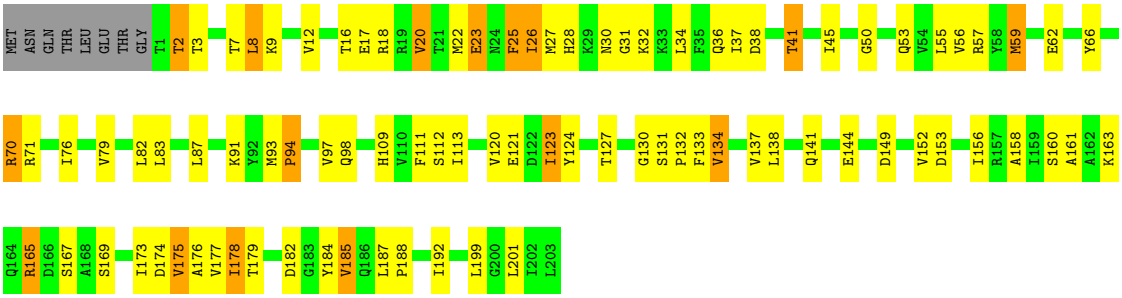

- Molecule 2: PROTEASOME

Chain T:



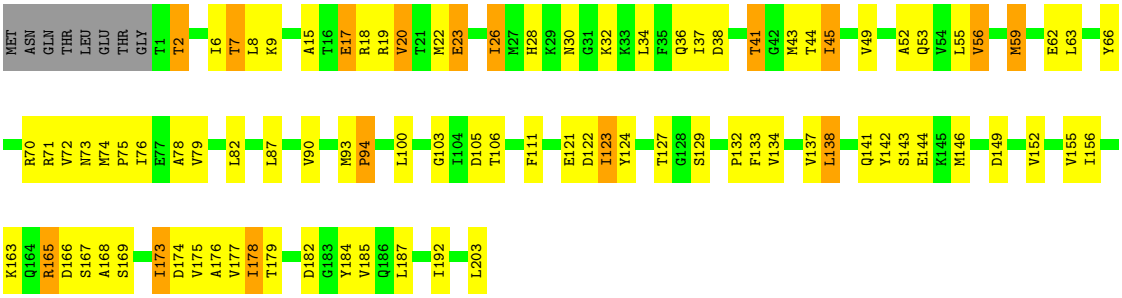

- Molecule 2: PROTEASOME

Chain U:



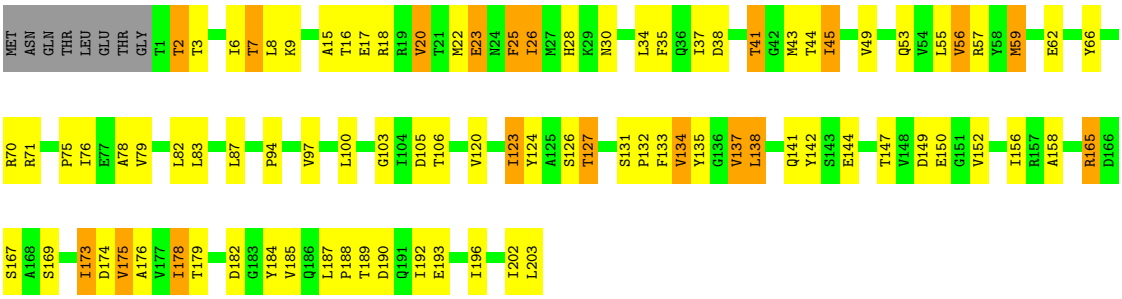

- Molecule 2: PROTEASOME

Chain V:



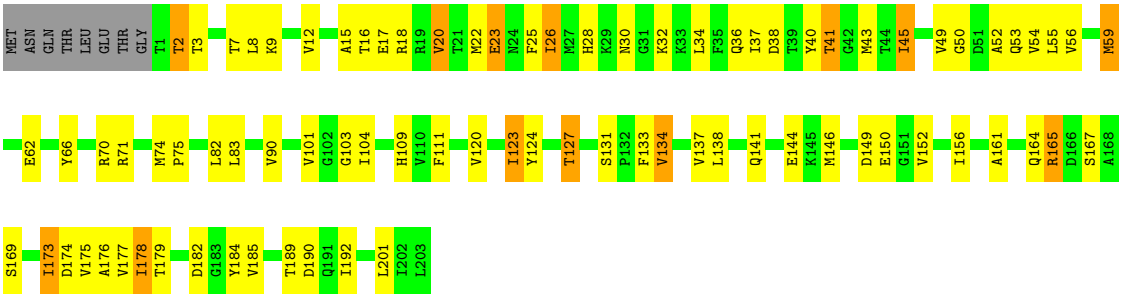

- Molecule 2: PROTEASOME

Chain W:




- Molecule 2: PROTEASOME

Chain X:




- Molecule 2: PROTEASOME

Chain Y: 58% 30% 8% •

MET ASN GLN THR LEU GLU THR GLY T1 T2 T3 T7 L8 K9 V12 A15 T16 E17 R18 R19 V20 T21 M22 E23 N24 F25 I26 M27 H28 K29 N30 G31 K32 K33 L34 F35 Q36 I37 D38 T41 T45 V49 G50 D51 A52 Q53 V54 L55 V56 M59 E62 L63

Y66 R70 R71 I76 V79 L82 L83 L87 P94 F111 V120 E121 D122 I123 Y124 T127 F133 V134 V137 L138 Q141 E144 D149 V152 I156 A161 Q164 R165 D166 S167 A168 S169 I173 D174 V175 A176 V177 I178 T179 R180

K181 D182 G183 Y184 V185 T189 D190 Q191 I192 E193 L203

• Molecule 2: PROTEASOME

Chain Z: 55% 35% 7% •

MET ASN GLN THR LEU GLU THR GLY T1 T2 T3 I6 T7 L8 K9 V12 A15 T16 E17 R18 R19 V20 I26 M27 H28 K29 N30 G31 K32 K33 L34 I37 D38 T41 G42 M43 T44 I45 V49 G50 D51 A52 Q53 V54 L55 V56 M59 E62 L63 Y66

R70 R71 M74 A78 V79 L82 L83 L87 M93 P94 L100 V110 F111 D114 V120 E121 D122 I123 Y124 T127 S131 P132 F133 V134 V137 L138 Q141 Y142 S143 E144 D149 V152 V155 I156 K163 Q164 R165 D166 S167 A168 S169

I173 D174 V175 A176 V177 I178 T179 D182 G183 Y184 V185 Q186 L187 P188 T189 D190 Q191 I192 L203

• Molecule 2: PROTEASOME

Chain 1: 54% 34% 9% •

MET ASN GLN THR LEU GLU THR GLY T1 T2 T3 I6 T7 L8 K9 V12 A15 T16 E17 R18 R19 V20 T21 M22 E23 I26 M27 H28 K29 N30 G31 K32 K33 L34 F35 Q36 I37 D38 T41 T44 I45 V49 G50 D51 A52 Q53 V54 L55 V56 M59 E62

Y66 R70 R71 P75 A78 V79 L82 L83 M93 P94 V97 Q98 L99 L100 G103 I104 F111 V120 E121 D122 I123 Y124 T127 S131 P132 F133 V134 V137 L138 Q141 Y142 S143 E144 D149 V152 D153 L154 V155 I156 R157 A158 K163 Q164

R165 D166 S167 A168 S169 I173 D174 V175 A176 V177 I178 T179 R180 V185 Q186 T189 D190 Q191 I192 L201 I202 L203

• Molecule 2: PROTEASOME

Chain 2: 55% 33% 8% •

MET ASN GLN THR LEU GLU THR GLY T1 T2 T3 I6 T7 L8 K9 V12 T16 E17 R18 R19 M22 E23 N24 F25 I26 N30 G31 K32 K33 L34 F35 Q36 I37 D38 T41 T45 V49 G50 D51 A52 Q53 V54 L55 V56 M59 E62 L63 Y66 R67

L68 Q69 R70 R71 I76 V79 L82 L83 L87 V90 P94 Y95 L99 G103 F111 E121 D122 I123 Y124 T127 S131 P132 F133 V134 V137 L138 Q141 Y142 S143 E144 K145 M146 D149 V152 T156 R157 A158 R165 D166 S167 A168 S169

I173 D174 V175 A176 V177 I178 T179 D182 G183 Y184 L187 P188 T189 D190 Q191 I192 L201 I202 L203

# 4 Data and refinement statistics (i)

| Property | Value | Source |
|---|---|---|
| Space group | P 21 21 21 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 311.90Å 209.70Å 117.20Å<br>90.00° 90.00° 90.00° | Depositor |
| Resolution (Å) | 10.00 – 3.40<br>20.97 – 3.40 | Depositor<br>EDS |
| % Data completeness<br>(in resolution range) | 85.7 (10.00-3.40)<br>85.0 (20.97-3.40) | Depositor<br>EDS |
| $R_{merge}$ | 0.10 | Depositor |
| $R_{sym}$ | (Not available) | Depositor |
| $< I/\sigma(I) >$ [1] | 2.72 (at 3.37Å) | Xtriage |
| Refinement program | X-PLOR 3.1 | Depositor |
| R, $R_{free}$ | 0.221 , (Not available)<br>0.208 , (Not available) | Depositor<br>DCC |
| $R_{free}$ test set | No test flags present. | wwPDB-VP |
| Wilson B-factor (Å$^2$) | 80.4 | Xtriage |
| Anisotropy | 0.268 | Xtriage |
| Bulk solvent $k_{sol}$(e/Å$^3$), $B_{sol}$(Å$^2$) | 0.31 , 131.3 | EDS |
| L-test for twinning[2] | $< \|L\| > = 0.48$, $< L^2 > = 0.31$ | Xtriage |
| Estimated twinning fraction | No twinning to report. | Xtriage |
| $F_o$,$F_c$ correlation | 0.91 | EDS |
| Total number of atoms | 56294 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 47.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.52% 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)

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 | 0/1743 | 0.71 | 1/2348 (0.0%) |
| 1 | C | 0.51 | 0/1743 | 0.71 | 0/2348 |
| 1 | D | 0.52 | 0/1743 | 0.70 | 1/2348 (0.0%) |
| 1 | E | 0.55 | 0/1743 | 0.72 | 0/2348 |
| 1 | F | 0.52 | 0/1743 | 0.70 | 0/2348 |
| 1 | G | 0.54 | 0/1743 | 0.69 | 0/2348 |
| 1 | H | 0.53 | 0/1743 | 0.70 | 0/2348 |
| 1 | I | 0.50 | 0/1743 | 0.70 | 0/2348 |
| 1 | J | 0.51 | 0/1743 | 0.69 | 0/2348 |
| 1 | K | 0.51 | 0/1743 | 0.69 | 1/2348 (0.0%) |
| 1 | L | 0.53 | 0/1743 | 0.70 | 0/2348 |
| 1 | M | 0.51 | 0/1743 | 0.70 | 0/2348 |
| 1 | N | 0.54 | 0/1743 | 0.70 | 0/2348 |
| 1 | O | 0.52 | 0/1743 | 0.71 | 0/2348 |
| 2 | 1 | 0.55 | 0/1577 | 0.74 | 1/2129 (0.0%) |
| 2 | 2 | 0.53 | 0/1577 | 0.75 | 1/2129 (0.0%) |
| 2 | B | 0.56 | 0/1577 | 0.76 | 0/2129 |
| 2 | P | 0.57 | 0/1577 | 0.76 | 0/2129 |
| 2 | Q | 0.54 | 0/1577 | 0.75 | 0/2129 |
| 2 | R | 0.55 | 0/1577 | 0.77 | 2/2129 (0.1%) |
| 2 | S | 0.57 | 0/1577 | 0.77 | 0/2129 |
| 2 | T | 0.55 | 0/1577 | 0.75 | 1/2129 (0.0%) |
| 2 | U | 0.54 | 0/1577 | 0.74 | 0/2129 |
| 2 | V | 0.56 | 0/1577 | 0.77 | 0/2129 |
| 2 | W | 0.57 | 0/1577 | 0.75 | 1/2129 (0.0%) |
| 2 | X | 0.54 | 0/1577 | 0.73 | 0/2129 |
| 2 | Y | 0.52 | 0/1577 | 0.74 | 0/2129 |
| 2 | Z | 0.54 | 0/1577 | 0.75 | 0/2129 |
| All | All | 0.53 | 0/46480 | 0.72 | 9/62678 (0.0%) |

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 2 | R | 95 | TYR | N-CA-C | -5.91 | 95.06 | 111.00 |
| 2 | 1 | 8 | LEU | CA-CB-CG | 5.55 | 128.06 | 115.30 |
| 1 | D | 128 | GLY | N-CA-C | 5.36 | 126.49 | 113.10 |
| 2 | W | 8 | LEU | CA-CB-CG | 5.32 | 127.54 | 115.30 |
| 1 | K | 128 | GLY | N-CA-C | 5.31 | 126.39 | 113.10 |

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 | 1720 | 393 | 1753 | 76 | 0 |
| 1 | C | 1720 | 393 | 1753 | 75 | 0 |
| 1 | D | 1720 | 393 | 1753 | 76 | 1 |
| 1 | E | 1720 | 393 | 1753 | 78 | 0 |
| 1 | F | 1720 | 393 | 1753 | 83 | 0 |
| 1 | G | 1720 | 393 | 1753 | 97 | 0 |
| 1 | H | 1720 | 393 | 1753 | 81 | 0 |
| 1 | I | 1720 | 393 | 1753 | 80 | 0 |
| 1 | J | 1720 | 393 | 1753 | 74 | 0 |
| 1 | K | 1720 | 393 | 1753 | 71 | 0 |
| 1 | L | 1720 | 393 | 1753 | 79 | 0 |
| 1 | M | 1720 | 393 | 1753 | 74 | 0 |
| 1 | N | 1720 | 393 | 1753 | 67 | 1 |
| 1 | O | 1720 | 393 | 1753 | 75 | 0 |
| 2 | 1 | 1558 | 350 | 1609 | 76 | 0 |
| 2 | 2 | 1558 | 350 | 1609 | 62 | 0 |
| 2 | B | 1558 | 350 | 1609 | 82 | 0 |
| 2 | P | 1558 | 350 | 1609 | 64 | 0 |
| 2 | Q | 1558 | 350 | 1609 | 56 | 0 |
| 2 | R | 1558 | 350 | 1609 | 68 | 0 |
| 2 | S | 1558 | 350 | 1609 | 77 | 0 |
| 2 | T | 1558 | 350 | 1609 | 61 | 0 |
| 2 | U | 1558 | 350 | 1609 | 69 | 0 |
| 2 | V | 1558 | 350 | 1609 | 74 | 0 |
| 2 | W | 1558 | 350 | 1609 | 68 | 0 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|---|---|---|---|---|---|---|
| 2 | X | 1558 | 350 | 1609 | 63 | 0 |
| 2 | Y | 1558 | 350 | 1609 | 66 | 0 |
| 2 | Z | 1558 | 350 | 1609 | 69 | 0 |
| All | All | 45892 | 10402 | 47068 | 1857 | 1 |

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 20.

The worst 5 of 1857 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) |
|---|---|---|---|
| 2:B:38:ASP:HB3 | 2:B:41:THR:HG23 | 1.38 | 1.03 |
| 2:T:38:ASP:HB3 | 2:T:41:THR:HG23 | 1.44 | 1.00 |
| 2:Q:45:ILE:HG12 | 2:Q:52:ALA:HB1 | 1.44 | 0.99 |
| 2:Q:38:ASP:HB3 | 2:Q:41:THR:HG23 | 1.46 | 0.98 |
| 2:R:38:ASP:HB3 | 2:R:41:THR:HG23 | 1.45 | 0.98 |

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

| Atom-1 | Atom-2 | Interatomic distance (Å) | Clash overlap (Å) |
|---|---|---|---|
| 1:D:178:ARG:HH22 | 1:N:178:ARG:HH22[4_457] | 1.29 | 0.31 |

## 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 | 219/233 (94%) | 184 (84%) | 26 (12%) | 9 (4%) | 3 18 |
| 1 | C | 219/233 (94%) | 185 (84%) | 25 (11%) | 9 (4%) | 3 18 |
| 1 | D | 219/233 (94%) | 184 (84%) | 28 (13%) | 7 (3%) | 4 22 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Analysed | Favoured | Allowed | Outliers | Percentiles | |
|---|---|---|---|---|---|---|---|
| 1 | E | 219/233 (94%) | 188 (86%) | 22 (10%) | 9 (4%) | 3 | 18 |
| 1 | F | 219/233 (94%) | 183 (84%) | 25 (11%) | 11 (5%) | 2 | 14 |
| 1 | G | 219/233 (94%) | 186 (85%) | 23 (10%) | 10 (5%) | 2 | 15 |
| 1 | H | 219/233 (94%) | 186 (85%) | 25 (11%) | 8 (4%) | 3 | 20 |
| 1 | I | 219/233 (94%) | 188 (86%) | 22 (10%) | 9 (4%) | 3 | 18 |
| 1 | J | 219/233 (94%) | 182 (83%) | 26 (12%) | 11 (5%) | 2 | 14 |
| 1 | K | 219/233 (94%) | 187 (85%) | 26 (12%) | 6 (3%) | 5 | 26 |
| 1 | L | 219/233 (94%) | 186 (85%) | 24 (11%) | 9 (4%) | 3 | 18 |
| 1 | M | 219/233 (94%) | 184 (84%) | 26 (12%) | 9 (4%) | 3 | 18 |
| 1 | N | 219/233 (94%) | 187 (85%) | 23 (10%) | 9 (4%) | 3 | 18 |
| 1 | O | 219/233 (94%) | 188 (86%) | 23 (10%) | 8 (4%) | 3 | 20 |
| 2 | 1 | 201/211 (95%) | 184 (92%) | 15 (8%) | 2 (1%) | 15 | 46 |
| 2 | 2 | 201/211 (95%) | 181 (90%) | 19 (10%) | 1 (0%) | 29 | 61 |
| 2 | B | 201/211 (95%) | 185 (92%) | 14 (7%) | 2 (1%) | 15 | 46 |
| 2 | P | 201/211 (95%) | 183 (91%) | 16 (8%) | 2 (1%) | 15 | 46 |
| 2 | Q | 201/211 (95%) | 180 (90%) | 20 (10%) | 1 (0%) | 29 | 61 |
| 2 | R | 201/211 (95%) | 184 (92%) | 15 (8%) | 2 (1%) | 15 | 46 |
| 2 | S | 201/211 (95%) | 181 (90%) | 16 (8%) | 4 (2%) | 7 | 30 |
| 2 | T | 201/211 (95%) | 176 (88%) | 23 (11%) | 2 (1%) | 15 | 46 |
| 2 | U | 201/211 (95%) | 184 (92%) | 15 (8%) | 2 (1%) | 15 | 46 |
| 2 | V | 201/211 (95%) | 179 (89%) | 20 (10%) | 2 (1%) | 15 | 46 |
| 2 | W | 201/211 (95%) | 186 (92%) | 13 (6%) | 2 (1%) | 15 | 46 |
| 2 | X | 201/211 (95%) | 181 (90%) | 18 (9%) | 2 (1%) | 15 | 46 |
| 2 | Y | 201/211 (95%) | 183 (91%) | 16 (8%) | 2 (1%) | 15 | 46 |
| 2 | Z | 201/211 (95%) | 185 (92%) | 14 (7%) | 2 (1%) | 15 | 46 |
| All | All | 5880/6216 (95%) | 5150 (88%) | 578 (10%) | 152 (3%) | 5 | 26 |

5 of 152 Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | A | 128 | GLY |
| 1 | A | 182 | GLU |
| 1 | A | 200 | SER |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | C | 128 | GLY |
| 1 | C | 200 | 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 | 184/193 (95%) | 163 (89%) | 21 (11%) | 5 | 21 |
| 1 | C | 184/193 (95%) | 160 (87%) | 24 (13%) | 4 | 16 |
| 1 | D | 184/193 (95%) | 159 (86%) | 25 (14%) | 3 | 14 |
| 1 | E | 184/193 (95%) | 159 (86%) | 25 (14%) | 3 | 14 |
| 1 | F | 184/193 (95%) | 158 (86%) | 26 (14%) | 3 | 13 |
| 1 | G | 184/193 (95%) | 162 (88%) | 22 (12%) | 5 | 19 |
| 1 | H | 184/193 (95%) | 164 (89%) | 20 (11%) | 6 | 23 |
| 1 | I | 184/193 (95%) | 162 (88%) | 22 (12%) | 5 | 19 |
| 1 | J | 184/193 (95%) | 162 (88%) | 22 (12%) | 5 | 19 |
| 1 | K | 184/193 (95%) | 160 (87%) | 24 (13%) | 4 | 16 |
| 1 | L | 184/193 (95%) | 159 (86%) | 25 (14%) | 3 | 14 |
| 1 | M | 184/193 (95%) | 162 (88%) | 22 (12%) | 5 | 19 |
| 1 | N | 184/193 (95%) | 160 (87%) | 24 (13%) | 4 | 16 |
| 1 | O | 184/193 (95%) | 157 (85%) | 27 (15%) | 3 | 12 |
| 2 | 1 | 170/177 (96%) | 148 (87%) | 22 (13%) | 4 | 16 |
| 2 | 2 | 170/177 (96%) | 144 (85%) | 26 (15%) | 2 | 11 |
| 2 | B | 170/177 (96%) | 145 (85%) | 25 (15%) | 3 | 12 |
| 2 | P | 170/177 (96%) | 148 (87%) | 22 (13%) | 4 | 16 |
| 2 | Q | 170/177 (96%) | 143 (84%) | 27 (16%) | 2 | 10 |
| 2 | R | 170/177 (96%) | 143 (84%) | 27 (16%) | 2 | 10 |
| 2 | S | 170/177 (96%) | 146 (86%) | 24 (14%) | 3 | 13 |
| 2 | T | 170/177 (96%) | 152 (89%) | 18 (11%) | 6 | 24 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Analysed | Rotameric | Outliers | Percentiles | |
|---|---|---|---|---|---|---|
| 2 | U | 170/177 (96%) | 146 (86%) | 24 (14%) | 3 | 13 |
| 2 | V | 170/177 (96%) | 145 (85%) | 25 (15%) | 3 | 12 |
| 2 | W | 170/177 (96%) | 145 (85%) | 25 (15%) | 3 | 12 |
| 2 | X | 170/177 (96%) | 149 (88%) | 21 (12%) | 4 | 17 |
| 2 | Y | 170/177 (96%) | 144 (85%) | 26 (15%) | 2 | 11 |
| 2 | Z | 170/177 (96%) | 150 (88%) | 20 (12%) | 5 | 19 |
| All | All | 4956/5180 (96%) | 4295 (87%) | 661 (13%) | 4 | 15 |

5 of 661 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 2 | T | 123 | ILE |
| 2 | Y | 41 | THR |
| 2 | U | 20 | VAL |
| 2 | T | 83 | LEU |
| 2 | V | 185 | VAL |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 126 such sidechains are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | L | 44 | ASN |
| 2 | X | 30 | ASN |
| 1 | N | 125 | GLN |
| 2 | X | 28 | HIS |
| 2 | Z | 141 | GLN |

### 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)

There are no ligands in this entry.

## 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 | 221/233 (94%) | -0.15 | 4 (1%) | 68 | 67 | 17, 56, 117, 155 | 0 |
| 1 | C | 221/233 (94%) | -0.20 | 1 (0%) | 91 | 90 | 13, 53, 115, 154 | 0 |
| 1 | D | 221/233 (94%) | -0.25 | 5 (2%) | 60 | 59 | 21, 53, 116, 153 | 0 |
| 1 | E | 221/233 (94%) | -0.22 | 4 (1%) | 68 | 67 | 17, 52, 115, 155 | 0 |
| 1 | F | 221/233 (94%) | -0.13 | 4 (1%) | 68 | 67 | 16, 57, 116, 153 | 0 |
| 1 | G | 221/233 (94%) | -0.10 | 5 (2%) | 60 | 59 | 20, 56, 118, 155 | 0 |
| 1 | H | 221/233 (94%) | -0.14 | 5 (2%) | 60 | 59 | 21, 57, 116, 154 | 0 |
| 1 | I | 221/233 (94%) | -0.14 | 5 (2%) | 60 | 59 | 17, 55, 116, 155 | 0 |
| 1 | J | 221/233 (94%) | -0.14 | 8 (3%) | 42 | 42 | 21, 55, 116, 156 | 0 |
| 1 | K | 221/233 (94%) | -0.25 | 4 (1%) | 68 | 67 | 20, 55, 116, 153 | 0 |
| 1 | L | 221/233 (94%) | -0.24 | 3 (1%) | 75 | 74 | 13, 53, 116, 153 | 0 |
| 1 | M | 221/233 (94%) | -0.22 | 4 (1%) | 68 | 67 | 17, 53, 117, 153 | 0 |
| 1 | N | 221/233 (94%) | -0.24 | 4 (1%) | 68 | 67 | 19, 52, 116, 156 | 0 |
| 1 | O | 221/233 (94%) | -0.24 | 3 (1%) | 75 | 74 | 18, 51, 116, 152 | 0 |
| 2 | 1 | 203/211 (96%) | -0.56 | 0 | 100 | 100 | 9, 33, 69, 85 | 0 |
| 2 | 2 | 203/211 (96%) | -0.61 | 0 | 100 | 100 | 6, 33, 69, 88 | 0 |
| 2 | B | 203/211 (96%) | -0.55 | 0 | 100 | 100 | 11, 34, 69, 90 | 0 |
| 2 | P | 203/211 (96%) | -0.54 | 0 | 100 | 100 | 8, 35, 68, 89 | 0 |
| 2 | Q | 203/211 (96%) | -0.65 | 0 | 100 | 100 | 8, 33, 71, 90 | 0 |
| 2 | R | 203/211 (96%) | -0.69 | 0 | 100 | 100 | 6, 32, 68, 86 | 0 |
| 2 | S | 203/211 (96%) | -0.63 | 0 | 100 | 100 | 7, 33, 68, 91 | 0 |
| 2 | T | 203/211 (96%) | -0.61 | 0 | 100 | 100 | 10, 35, 73, 90 | 0 |
| 2 | U | 203/211 (96%) | -0.59 | 0 | 100 | 100 | 14, 34, 70, 92 | 0 |
| 2 | V | 203/211 (96%) | -0.60 | 0 | 100 | 100 | 11, 32, 69, 89 | 0 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Analysed | <RSRZ> | #RSRZ>2 | | | OWAB(Å$^2$) | Q<0.9 |
|---|---|---|---|---|---|---|---|---|
| 2 | W | 203/211 (96%) | -0.62 | 0 | 100 | 100 | 11, 34, 69, 85 | 0 |
| 2 | X | 203/211 (96%) | -0.66 | 0 | 100 | 100 | 8, 32, 70, 90 | 0 |
| 2 | Y | 203/211 (96%) | -0.66 | 0 | 100 | 100 | 10, 32, 68, 92 | 0 |
| 2 | Z | 203/211 (96%) | -0.71 | 0 | 100 | 100 | 9, 32, 69, 91 | 0 |
| All | All | 5936/6216 (95%) | -0.40 | 59 (0%) | 82 | 81 | 6, 43, 107, 156 | 0 |

The worst 5 of 59 RSRZ outliers are listed below:

| Mol | Chain | Res | Type | RSRZ |
|---|---|---|---|---|
| 1 | G | 205 | GLU | 4.2 |
| 1 | I | 64 | ILE | 4.0 |
| 1 | L | 55 | ARG | 4.0 |
| 1 | I | 205 | GLU | 3.8 |
| 1 | D | 63 | SER | 3.8 |

## 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)

There are no ligands in this entry.

## 6.5 Other polymers (i)

There are no such residues in this entry.