



wwPDB EM Map/Model Validation Report ⓘ

Apr 10, 2016 – 02:14 PM BST

PDB ID : 4CR4
EMDB ID: : EMD-2596
Title : Deep classification of a large cryo-EM dataset defines the conformational landscape of the 26S proteasome
Authors : Unverdorben, P.; Beck, F.; Sledz, P.; Schweitzer, A.; Pfeifer, G.; Plitzko, J.M.; Baumeister, W.; Foerster, F.
Deposited on : 2014-02-25
Resolution : 8.80 Å(reported)

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could stem from errors in the original structure(s) used in the fitting.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/EMValidationReportHelp>

MolProbity : 4.02b-467
Mogul : unknown
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : trunk27241

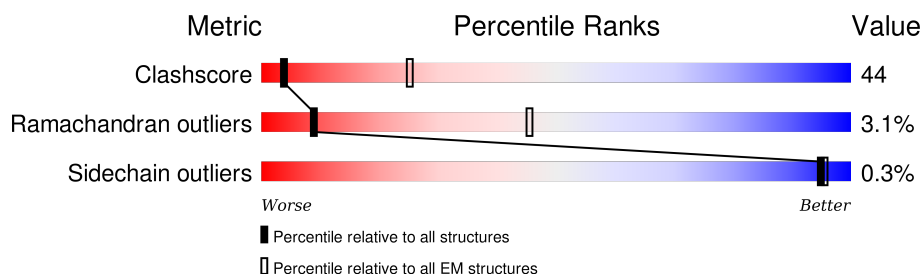
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 8.80 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.







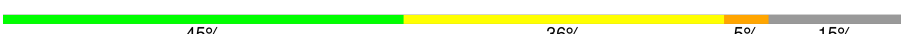
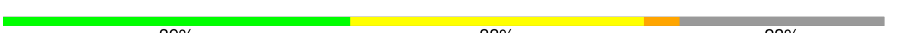
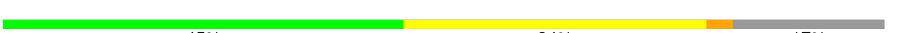



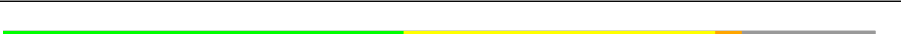



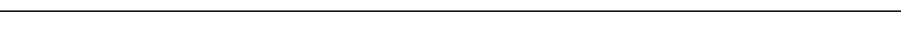
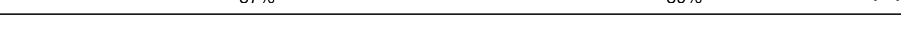

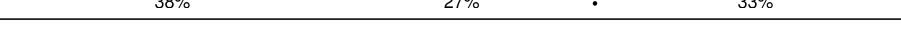



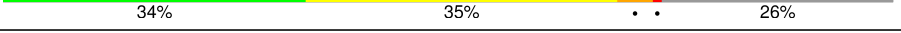


| Metric | Whole archive (#Entries) | EM structures (#Entries) |
|-----------------------|-----------------------------|-----------------------------|
| Clashscore | 114402 | 924 |
| Ramachandran outliers | 111179 | 726 |
| Sidechain outliers | 111093 | 686 |

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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 $\leq 5\%$.

| Mol | Chain | Length | Quality of chain |
|-----|-------|--------|------------------|
| 1 | 1 | 215 | 50% 42% • 5% |
| 2 | 2 | 261 | 54% 28% • 15% |
| 3 | 3 | 205 | 56% 40% • |
| 4 | 4 | 198 | 63% 35% • |
| 5 | 5 | 287 | 44% 27% • 26% |
| 6 | 6 | 241 | 49% 41% • 8% |
| 7 | 7 | 266 | 50% 36% • 12% |
| 8 | A | 252 | 50% 43% • • |
| 9 | B | 250 | 64% 34% • |

Continued on next page...

Continued from previous page...

| Mol | Chain | Length | Quality of chain |
|-----|-------|--------|--|
| 10 | C | 258 |  |
| 11 | D | 254 |  |
| 12 | E | 260 |  |
| 13 | F | 234 |  |
| 14 | G | 288 |  |
| 15 | H | 467 |  |
| 16 | I | 437 |  |
| 17 | J | 405 |  |
| 18 | K | 428 |  |
| 19 | L | 437 |  |
| 20 | M | 434 |  |
| 21 | N | 945 |  |
| 22 | O | 393 |  |
| 23 | P | 445 |  |
| 24 | Q | 434 |  |
| 25 | R | 429 |  |
| 26 | S | 523 |  |
| 27 | T | 274 |  |
| 28 | U | 338 |  |
| 29 | V | 306 |  |
| 30 | W | 268 |  |
| 31 | X | 156 |  |
| 32 | Y | 89 |  |
| 33 | Z | 993 |  |

2 Entry composition

There are 33 unique types of molecules in this entry. The entry contains 80139 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 COMPONENT PRE3.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---|---------|-------|
| 1 | 1 | 205 | Total | C | N | O | S | 0 | 0 |
| | | | 1576 | 996 | 261 | 312 | 7 | | |

- Molecule 2 is a protein called PROTEASOME COMPONENT PUP1.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|---|---------|-------|
| 2 | 2 | 223 | Total | C | N | O | S | 0 | 0 |
| | | | 1692 | 1067 | 294 | 324 | 7 | | |

- Molecule 3 is a protein called PROTEASOME COMPONENT PUP3.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|---|---------|-------|
| 3 | 3 | 204 | Total | C | N | O | S | 0 | 0 |
| | | | 1581 | 1010 | 258 | 305 | 8 | | |

- Molecule 4 is a protein called PROTEASOME COMPONENT C11.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|---|---------|-------|
| 4 | 4 | 198 | Total | C | N | O | S | 0 | 0 |
| | | | 1585 | 1005 | 269 | 305 | 6 | | |

- Molecule 5 is a protein called PROTEASOME COMPONENT PRE2.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|---|---------|-------|
| 5 | 5 | 212 | Total | C | N | O | S | 0 | 0 |
| | | | 1646 | 1045 | 282 | 312 | 7 | | |

There is a discrepancy between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|----------------|------------|
| 5 | 33 | ARG | LYS | SEE REMARK 999 | UNP P30656 |

- Molecule 6 is a protein called PROTEASOME COMPONENT C5.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|---|---------|-------|
| 6 | 6 | 222 | Total | C | N | O | S | 0 | 0 |
| | | | 1757 | 1115 | 303 | 335 | 4 | | |

- Molecule 7 is a protein called PROTEASOME COMPONENT PRE4.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|---|---------|-------|
| 7 | 7 | 233 | Total | C | N | O | S | 0 | 0 |
| | | | 1824 | 1154 | 312 | 351 | 7 | | |

- Molecule 8 is a protein called PROTEASOME COMPONENT C7-ALPHA.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|---|---------|-------|
| 8 | A | 243 | Total | C | N | O | S | 0 | 0 |
| | | | 1921 | 1221 | 322 | 370 | 8 | | |

- Molecule 9 is a protein called PROTEASOME COMPONENT Y7.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|---|---------|-------|
| 9 | B | 250 | Total | C | N | O | S | 0 | 0 |
| | | | 1915 | 1219 | 315 | 377 | 4 | | |

- Molecule 10 is a protein called PROTEASOME COMPONENT Y13.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|---|---------|-------|
| 10 | C | 245 | Total | C | N | O | S | 0 | 0 |
| | | | 1913 | 1207 | 323 | 380 | 3 | | |

- Molecule 11 is a protein called PROTEASOME COMPONENT PRE6.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|---|---------|-------|
| 11 | D | 242 | Total | C | N | O | S | 0 | 0 |
| | | | 1899 | 1186 | 333 | 376 | 4 | | |

- Molecule 12 is a protein called PROTEASOME COMPONENT PUP2.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|---|---------|-------|
| 12 | E | 243 | Total | C | N | O | S | 0 | 0 |
| | | | 1867 | 1165 | 315 | 380 | 7 | | |

- Molecule 13 is a protein called PROTEASOME COMPONENT PRE5.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|---|---------|-------|
| 13 | F | 233 | Total | C | N | O | S | 0 | 0 |
| | | | 1795 | 1129 | 312 | 350 | 4 | | |

- Molecule 14 is a protein called PROTEASOME COMPONENT C1.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|---|---------|-------|
| 14 | G | 245 | Total | C | N | O | S | 0 | 0 |
| | | | 1900 | 1207 | 331 | 358 | 4 | | |

- Molecule 15 is a protein called 26S PROTEASE REGULATORY SUBUNIT 7 HOMOLOG.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|----|---------|-------|
| 15 | H | 359 | Total | C | N | O | S | 0 | 0 |
| | | | 2792 | 1755 | 499 | 523 | 15 | | |

- Molecule 16 is a protein called 26S PROTEASE REGULATORY SUBUNIT 4 HOMOLOG.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|----|---------|-------|
| 16 | I | 362 | Total | C | N | O | S | 0 | 0 |
| | | | 2822 | 1773 | 471 | 563 | 15 | | |

- Molecule 17 is a protein called 26S PROTEASE REGULATORY SUBUNIT 8 HOMOLOG.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|----|---------|-------|
| 17 | J | 373 | Total | C | N | O | S | 0 | 0 |
| | | | 2928 | 1837 | 527 | 547 | 17 | | |

- Molecule 18 is a protein called 26S PROTEASE REGULATORY SUBUNIT 6B HOMOLOG.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|----|---------|-------|
| 18 | K | 381 | Total | C | N | O | S | 0 | 0 |
| | | | 3019 | 1898 | 530 | 581 | 10 | | |

- Molecule 19 is a protein called 26S PROTEASE SUBUNIT RPT4.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|----|---------|-------|
| 19 | L | 361 | Total | C | N | O | S | 0 | 0 |
| | | | 2853 | 1798 | 507 | 536 | 12 | | |

- Molecule 20 is a protein called 26S PROTEASE REGULATORY SUBUNIT 6A.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|----|---------|-------|
| 20 | M | 367 | Total | C | N | O | S | 0 | 0 |
| | | | 2866 | 1799 | 503 | 553 | 11 | | |

- Molecule 21 is a protein called 26S PROTEASOME REGULATORY SUBUNIT RPN2.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|------|------|----|---------|-------|
| 21 | N | 849 | Total | C | N | O | S | 0 | 0 |
| | | | 6562 | 4174 | 1099 | 1261 | 28 | | |

- Molecule 22 is a protein called 26S PROTEASOME REGULATORY SUBUNIT RPN9.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|---|---------|-------|
| 22 | O | 387 | Total | C | N | O | S | 0 | 0 |
| | | | 3182 | 2047 | 520 | 606 | 9 | | |

- Molecule 23 is a protein called 26S PROTEASOME REGULATORY SUBUNIT RPN5.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|---|---------|-------|
| 23 | P | 415 | Total | C | N | O | S | 0 | 0 |
| | | | 3401 | 2166 | 571 | 655 | 9 | | |

- Molecule 24 is a protein called 26S PROTEASOME REGULATORY SUBUNIT RPN6.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|----|---------|-------|
| 24 | Q | 431 | Total | C | N | O | S | 0 | 0 |
| | | | 3471 | 2205 | 574 | 676 | 16 | | |

- Molecule 25 is a protein called 26S PROTEASOME REGULATORY SUBUNIT RPN7.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|----|---------|-------|
| 25 | R | 400 | Total | C | N | O | S | 0 | 0 |
| | | | 3218 | 2051 | 527 | 630 | 10 | | |

- Molecule 26 is a protein called 26S PROTEASOME REGULATORY SUBUNIT RPN3.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|----|---------|-------|
| 26 | S | 353 | Total | C | N | O | S | 0 | 0 |
| | | | 2893 | 1857 | 482 | 541 | 13 | | |

- Molecule 27 is a protein called 26S PROTEASOME REGULATORY SUBUNIT RPN12.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|---|---------|-------|
| 27 | T | 272 | Total | C | N | O | S | 0 | 0 |
| | | | 2235 | 1432 | 355 | 441 | 7 | | |

- Molecule 28 is a protein called 26S PROTEASOME REGULATORY SUBUNIT RPN8.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|---|---------|-------|
| 28 | U | 255 | Total | C | N | O | S | 0 | 0 |
| | | | 2061 | 1312 | 352 | 391 | 6 | | |

- Molecule 29 is a protein called 26S PROTEASOME REGULATORY SUBUNIT RPN11.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|----|---------|-------|
| 29 | V | 247 | Total | C | N | O | S | 0 | 0 |
| | | | 1942 | 1225 | 328 | 376 | 13 | | |

- Molecule 30 is a protein called 26S PROTEASOME REGULATORY SUBUNIT RPN10.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---|---------|-------|
| 30 | W | 197 | Total | C | N | O | S | 0 | 0 |
| | | | 1534 | 962 | 269 | 300 | 3 | | |

- Molecule 31 is a protein called 26S PROTEASOME REGULATORY SUBUNIT RPN13.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---|---------|-------|
| 31 | X | 127 | Total | C | N | O | S | 0 | 0 |
| | | | 1032 | 664 | 169 | 195 | 4 | | |

- Molecule 32 is a protein called 26S PROTEASOME COMPLEX SUBUNIT SEM1.

| Mol | Chain | Residues | Atoms | | | | AltConf | Trace |
|-----|-------|----------|-------|-----|----|----|---------|-------|
| 32 | Y | 19 | Total | C | N | O | 0 | 0 |
| | | | 168 | 101 | 30 | 37 | | |

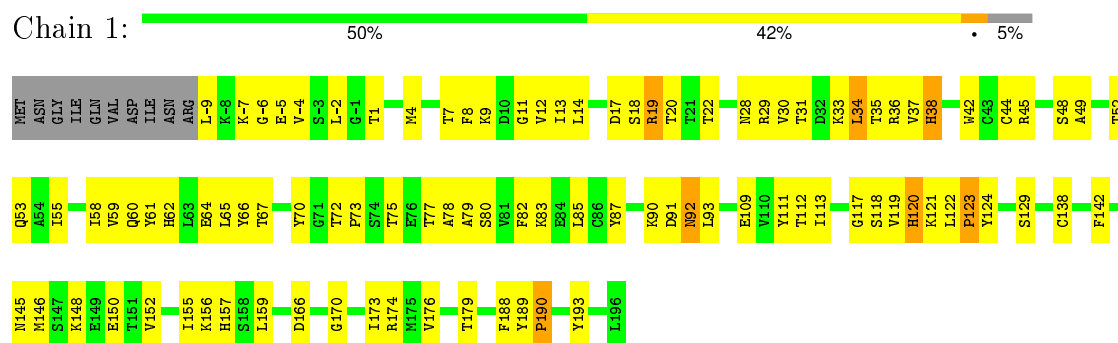
- Molecule 33 is a protein called 26S PROTEASOME REGULATORY SUBUNIT RPN1.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|------|------|----|---------|-------|
| 33 | Z | 813 | Total | C | N | O | S | 0 | 0 |
| | | | 6289 | 3995 | 1029 | 1236 | 29 | | |

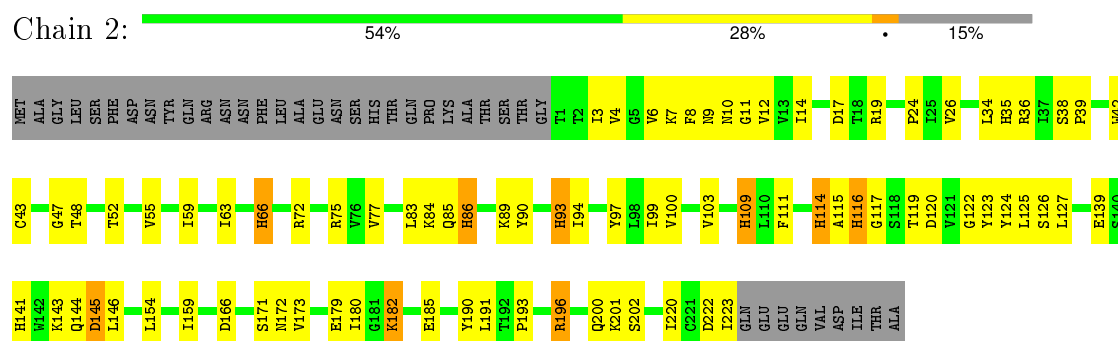
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

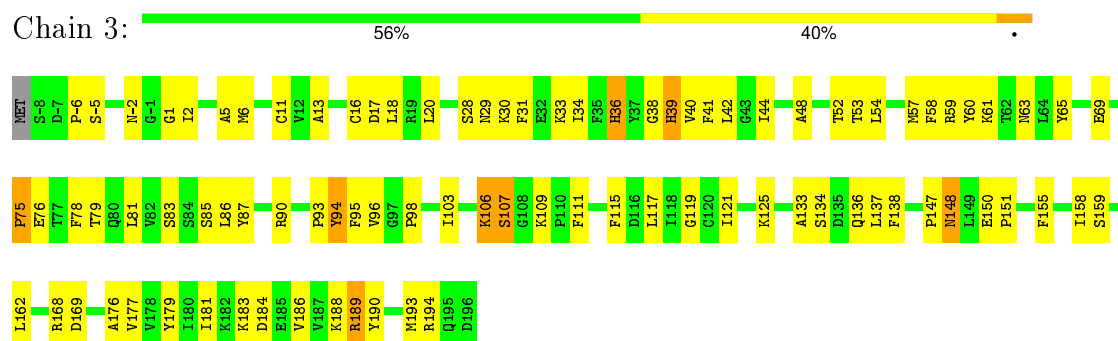
• Molecule 1: PROTEASOME COMPONENT PRE3



• Molecule 2: PROTEASOME COMPONENT PUP1



• Molecule 3: PROTEASOME COMPONENT PUP3



• Molecule 4: PROTEASOME COMPONENT C11

- Molecule 5: PROTEASOME COMPONENT PRE2

- Molecule 6: PROTEASOME COMPONENT C5

- Molecule 7: PROTEASOME COMPONENT PRE4

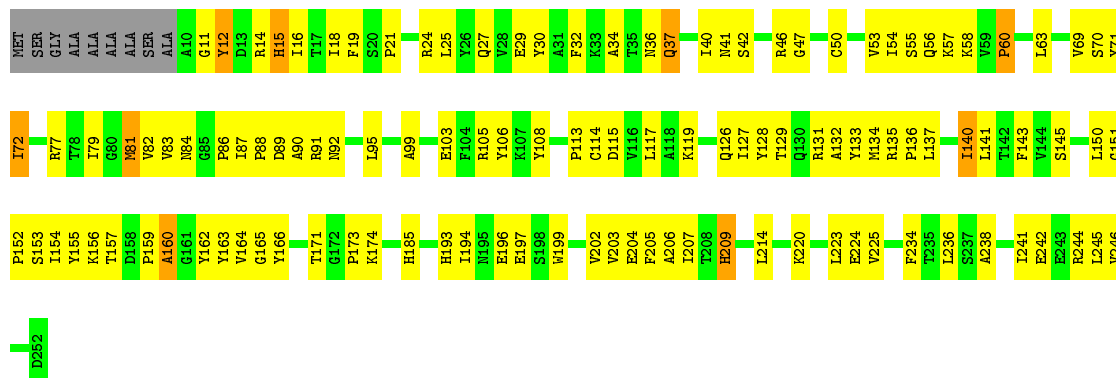
WORLDWIDE
PDB
PROTEIN DATA BANK

 **EMDataBank**
Unified Data Resource for 3DEM



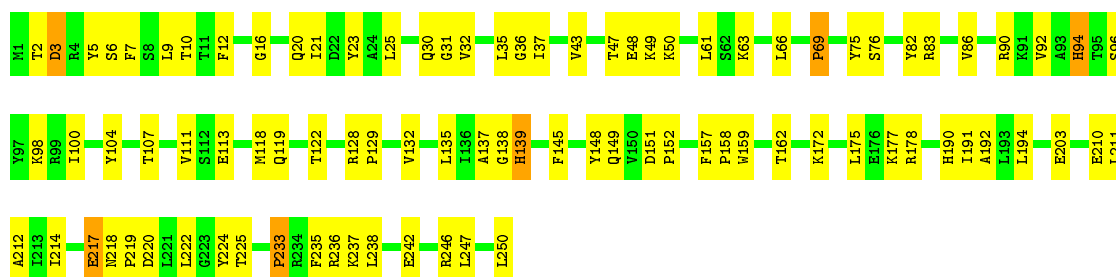
• Molecule 8: PROTEASOME COMPONENT C7-ALPHA

Chain A: 50% 43%



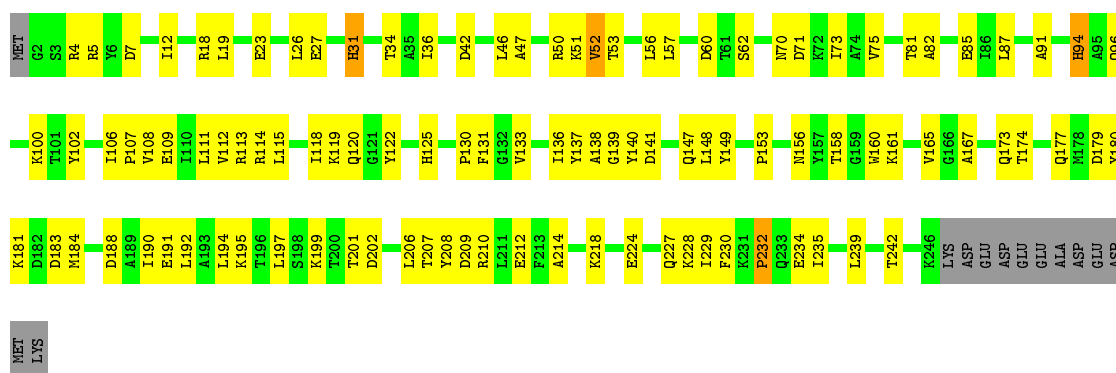
• Molecule 9: PROTEASOME COMPONENT Y7

Chain B: 64% 34%



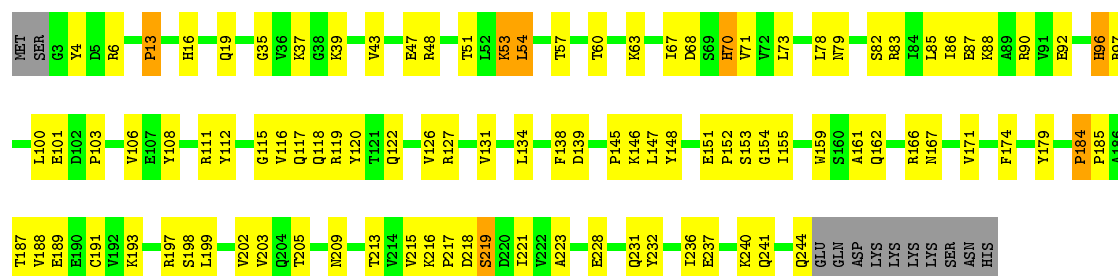
• Molecule 10: PROTEASOME COMPONENT Y13

Chain C: 54% 39% 5%



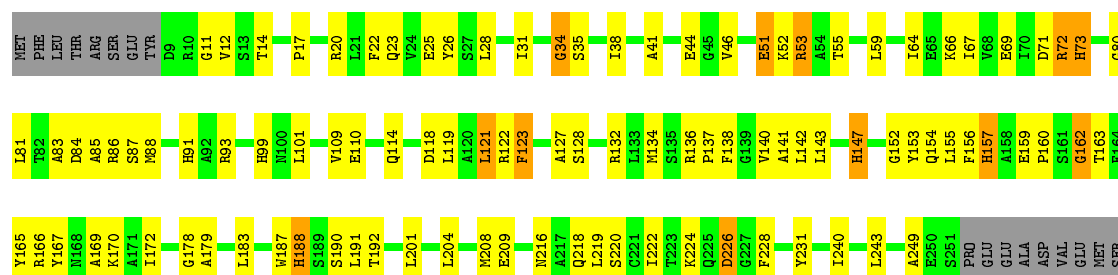
• Molecule 11: PROTEASOME COMPONENT PRE6

Chain D: 56% 37% 5%



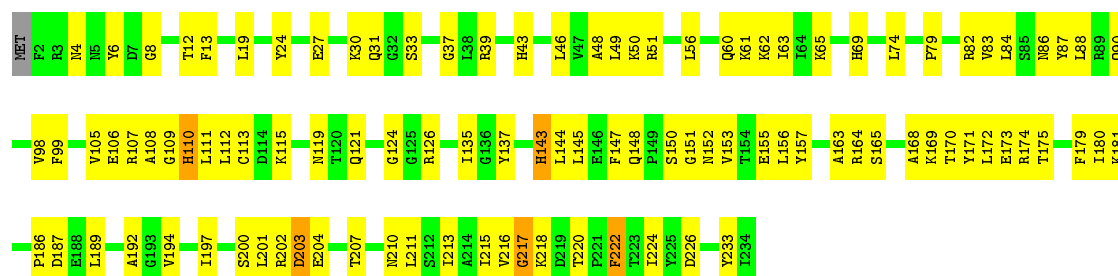
- Molecule 12: PROTEASOME COMPONENT PUP2

Chain E: 55% 34% 5% 7%



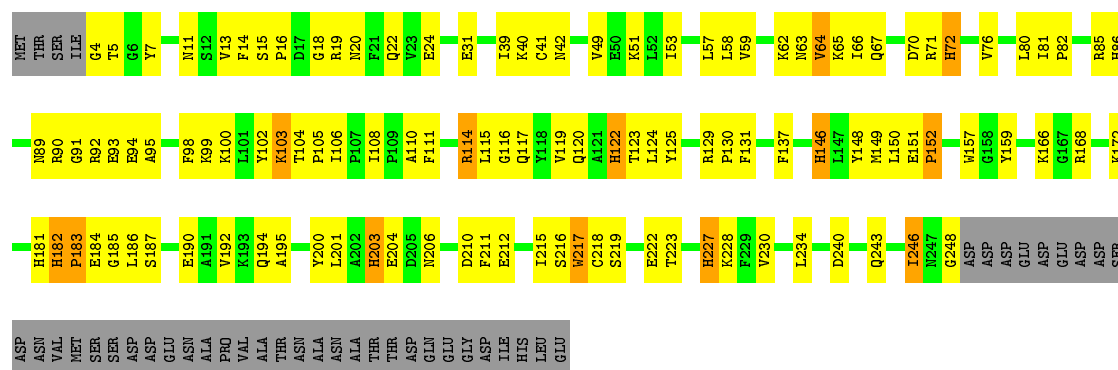
- Molecule 13: PROTEASOME COMPONENT PRE5

Chain F: 56% 42%



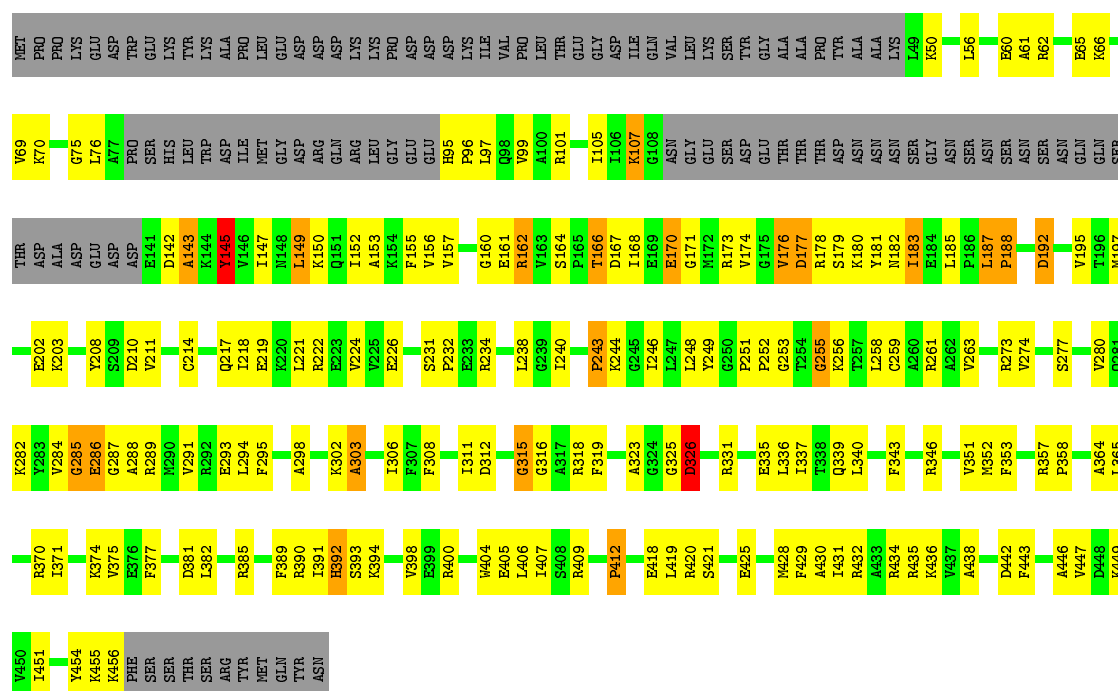
- Molecule 14: PROTEASOME COMPONENT C1

Chain G: 45% 36% 5% 15%



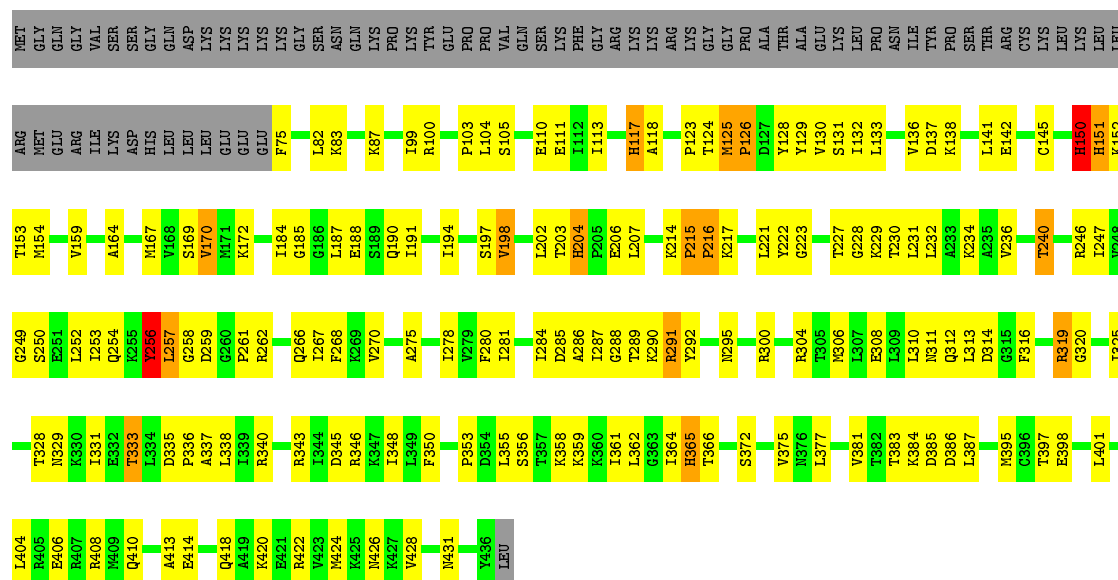
- Molecule 15: 26S PROTEASE REGULATORY SUBUNIT 7 HOMOLOG

Chain H:  39% 33% . 23%



● Molecule 16: 26S PROTEASE REGULATORY SUBUNIT 4 HOMOLOG

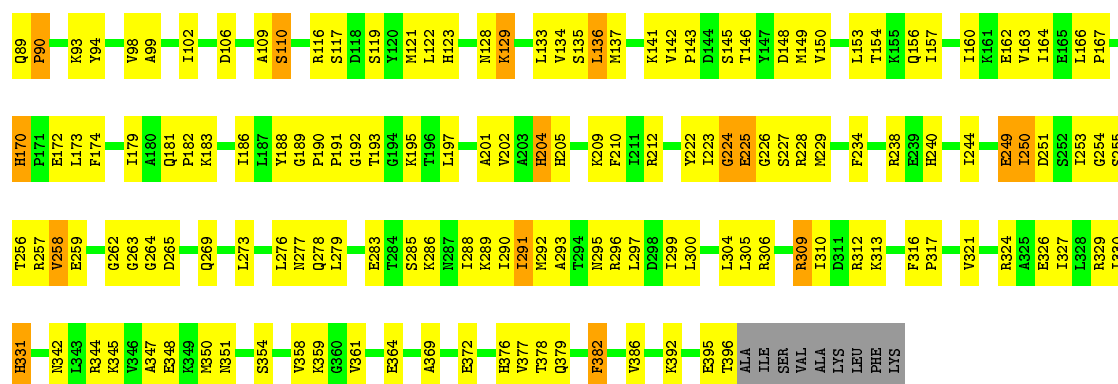
Chain I:  45% 34% • 17%



• Molecule 17: 26S PROTEASE REGULATORY SUBUNIT 8 HOMOLOG

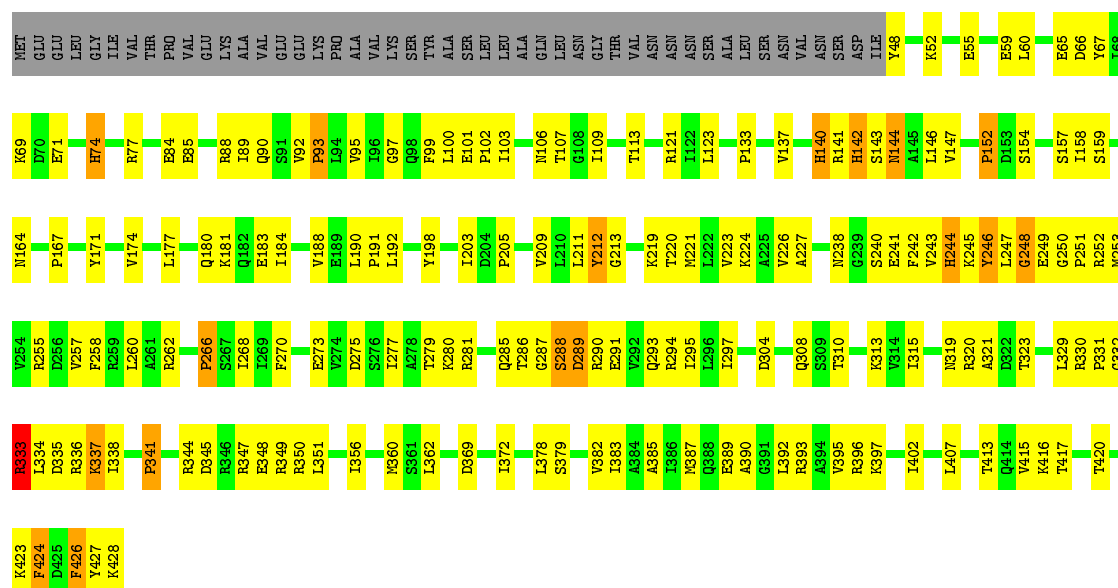
Chain J:  50% 38% 8%





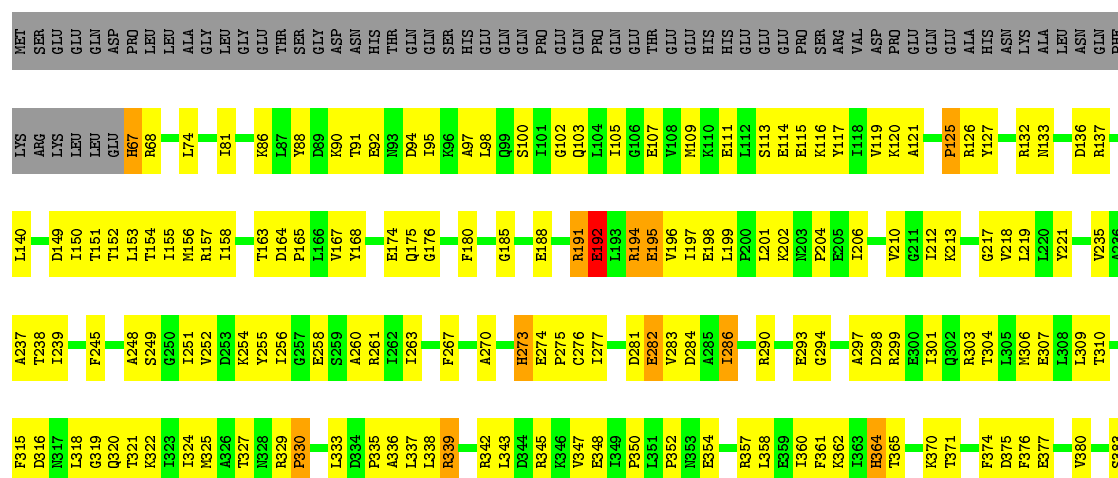
• Molecule 18: 26S PROTEASE REGULATORY SUBUNIT 6B HOMOLOG

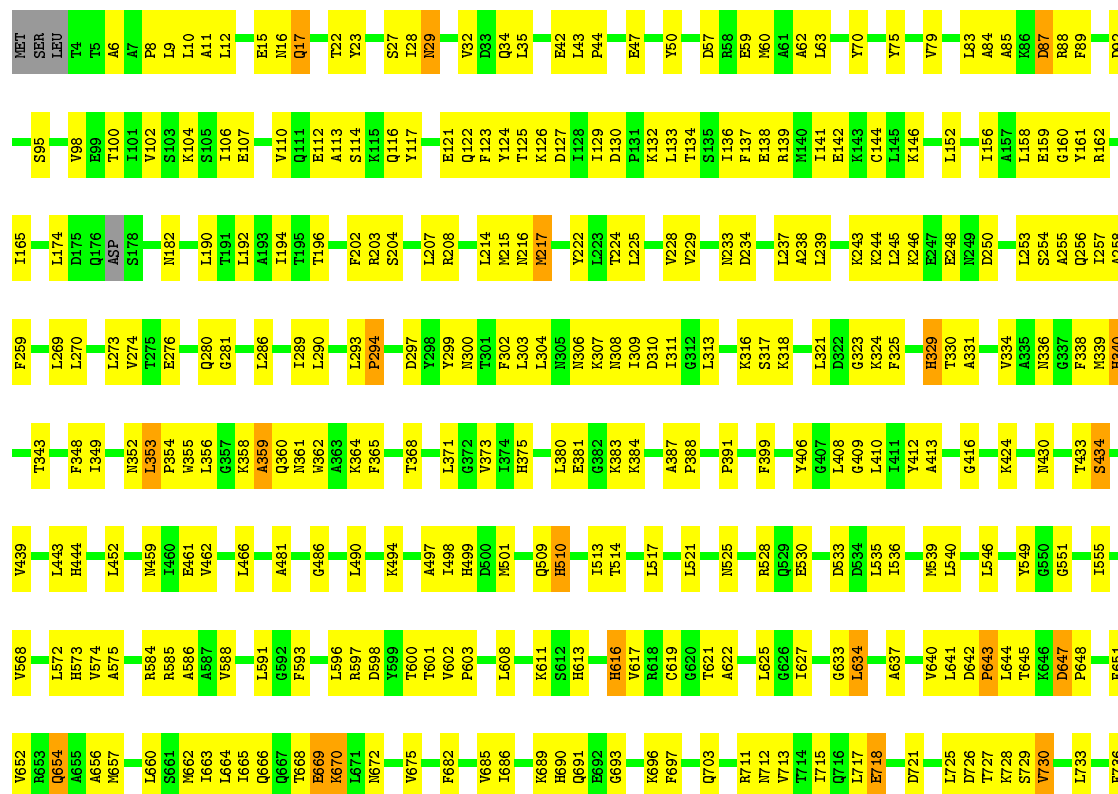
Chain K: 49% 36% 11%



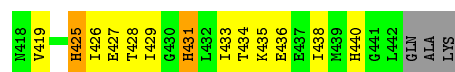
• Molecule 19: 26S PROTEASE SUBUNIT RPT4

Chain L: 42% 38% 17%



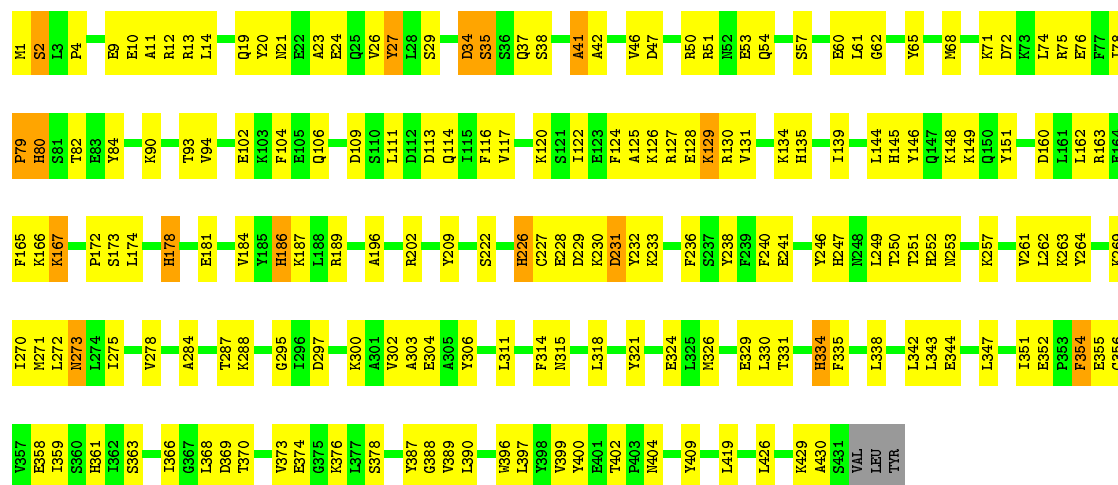






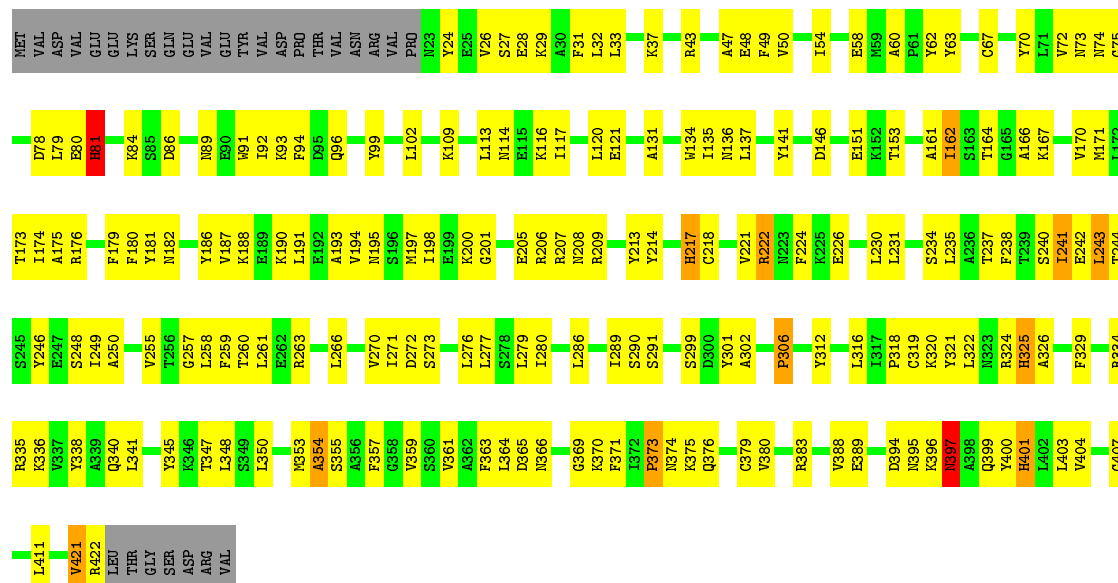
• Molecule 24: 26S PROTEASOME REGULATORY SUBUNIT RPN6

Chain Q: 57% 39%



• Molecule 25: 26S PROTEASOME REGULATORY SUBUNIT RPN7

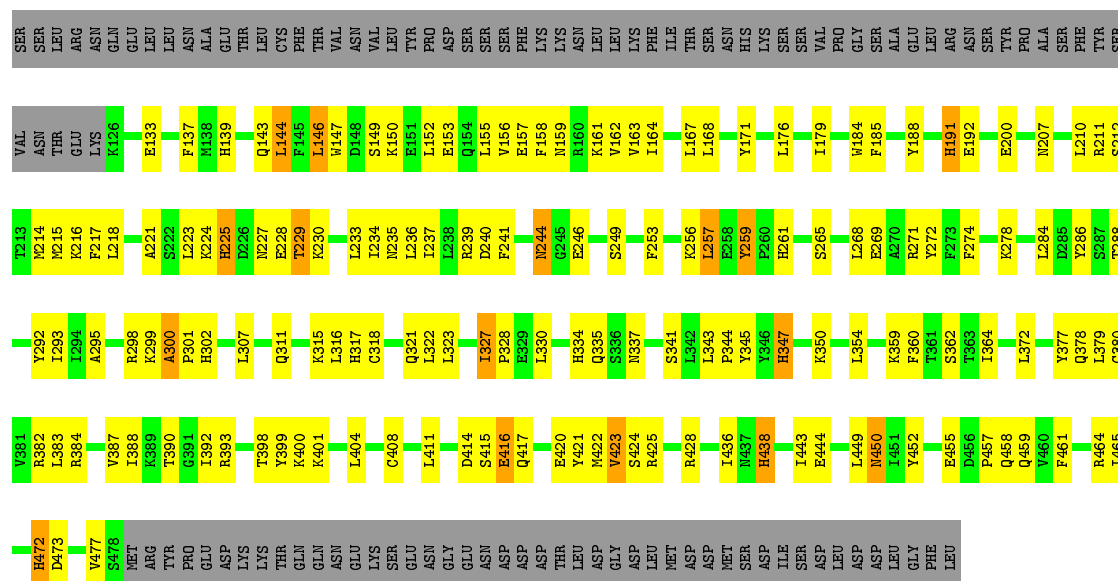
Chain R: 49% 41% 7%



• Molecule 26: 26S PROTEASOME REGULATORY SUBUNIT RPN3

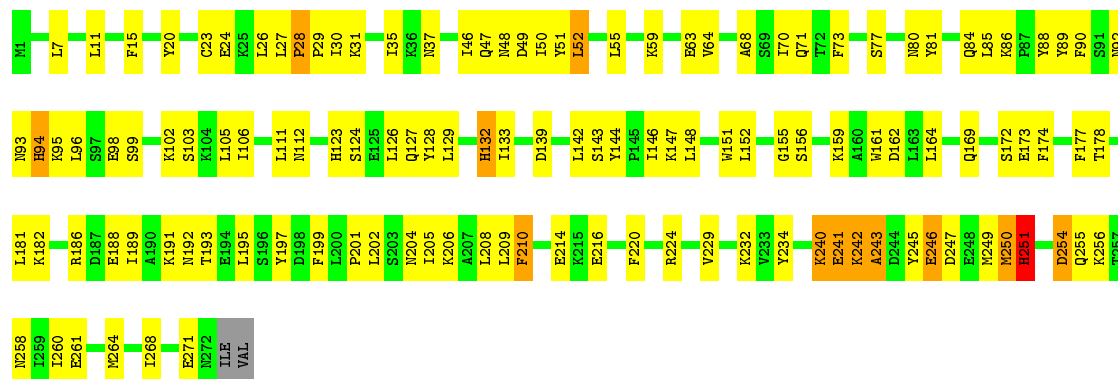
Chain S: 38% 27% 33%





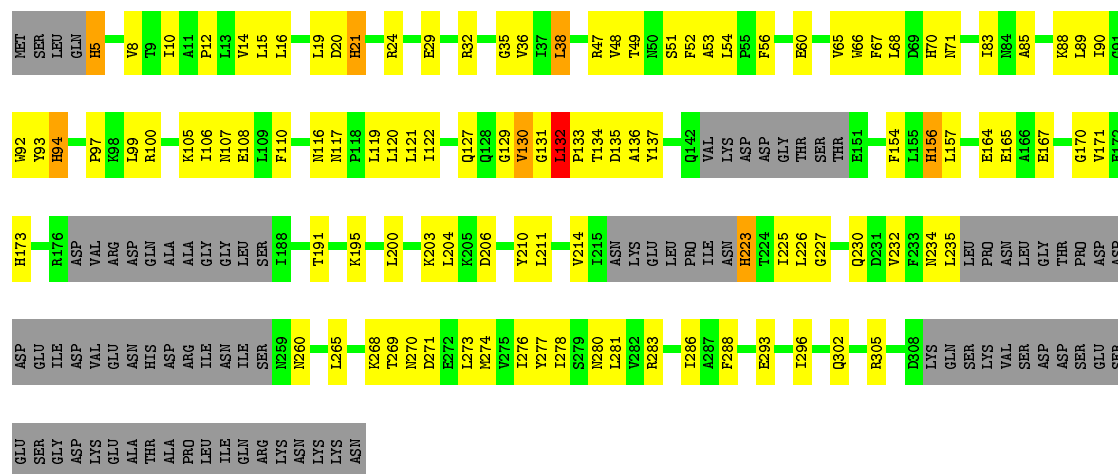
• Molecule 27: 26S PROTEASOME REGULATORY SUBUNIT RPN12

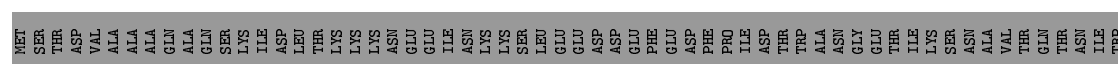
Chain T: 54% 41%

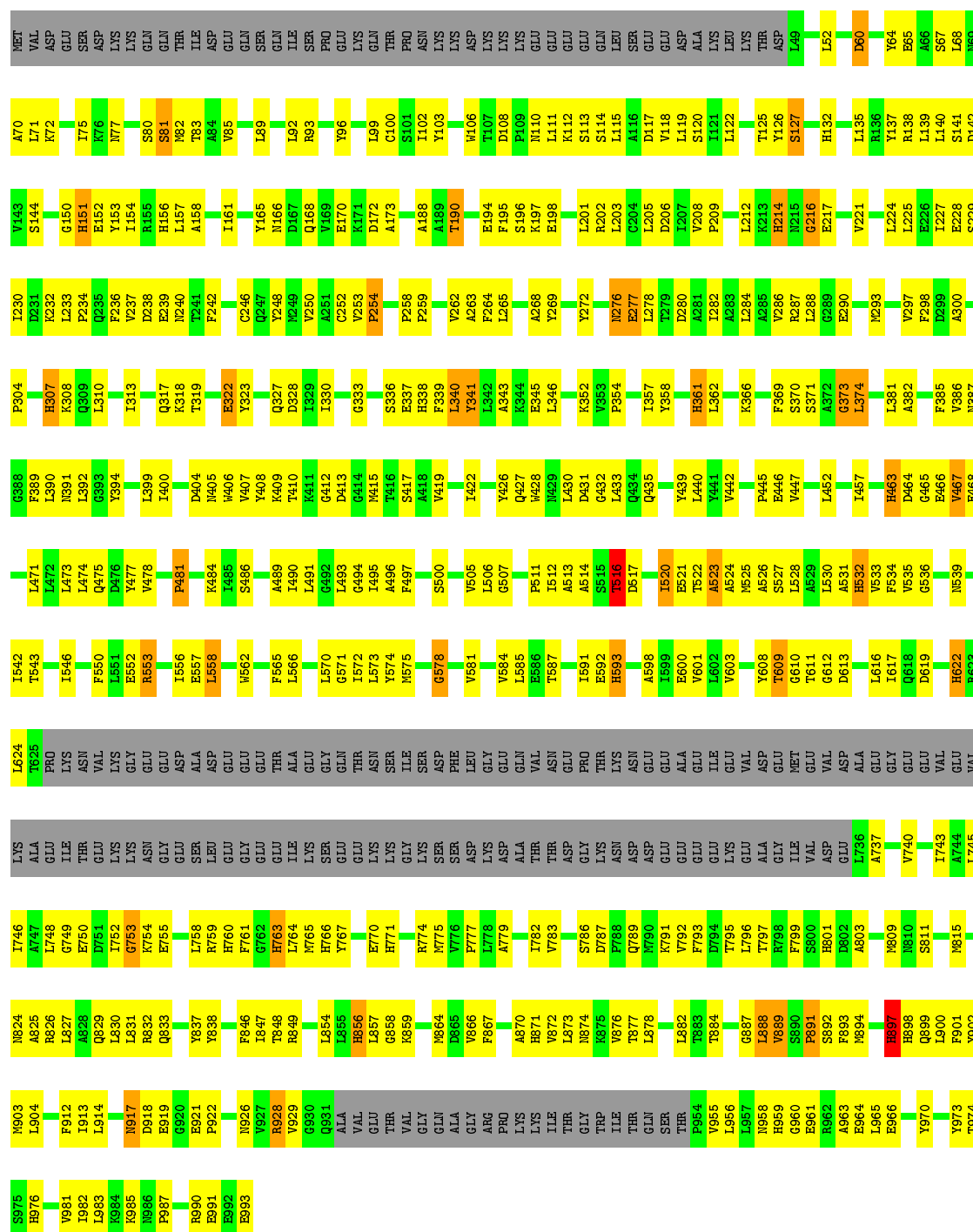


• Molecule 28: 26S PROTEASOME REGULATORY SUBUNIT RPN8

Chain U: 43% 30% 25%







4 Experimental information

| Property | Value | Source |
|--------------------------------------|-----------------------------|-----------|
| Reconstruction method | SINGLE PARTICLE | Depositor |
| Imposed symmetry | POINT, Not provided | Depositor |
| Number of images | Not provided | Depositor |
| Resolution determination method | Not provided | Depositor |
| CTF correction method | MICROGRAPH | Depositor |
| Microscope | OTHER | Depositor |
| Voltage (kV) | 200 | Depositor |
| Electron dose ($e^-/\text{\AA}^2$) | 25 | Depositor |
| Minimum defocus (nm) | 1000 | Depositor |
| Maximum defocus (nm) | 3500 | Depositor |
| Magnification | Not provided | Depositor |
| Image detector | TVIPS TEMCAM-F816 (8K X 8K) | Depositor |

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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 >2 | RMSZ | # Z >2 |
| 1 | 1 | 1.19 | 7/1605 (0.4%) | 1.21 | 4/2171 (0.2%) |
| 10 | C | 1.16 | 8/1943 (0.4%) | 1.11 | 1/2629 (0.0%) |
| 11 | D | 1.13 | 6/1928 (0.3%) | 1.18 | 3/2610 (0.1%) |
| 12 | E | 1.22 | 13/1892 (0.7%) | 1.22 | 2/2549 (0.1%) |
| 13 | F | 1.22 | 9/1823 (0.5%) | 1.20 | 1/2463 (0.0%) |
| 14 | G | 1.29 | 16/1940 (0.8%) | 1.16 | 2/2619 (0.1%) |
| 15 | H | 1.19 | 8/2830 (0.3%) | 1.35 | 17/3805 (0.4%) |
| 16 | I | 1.20 | 12/2859 (0.4%) | 1.23 | 10/3853 (0.3%) |
| 17 | J | 1.21 | 12/2964 (0.4%) | 1.20 | 6/3981 (0.2%) |
| 18 | K | 1.13 | 10/3061 (0.3%) | 1.21 | 10/4129 (0.2%) |
| 19 | L | 1.15 | 8/2895 (0.3%) | 1.20 | 11/3892 (0.3%) |
| 2 | 2 | 1.21 | 11/1723 (0.6%) | 1.21 | 3/2337 (0.1%) |
| 20 | M | 1.18 | 14/2903 (0.5%) | 1.18 | 5/3909 (0.1%) |
| 21 | N | 1.14 | 27/6670 (0.4%) | 1.15 | 6/9023 (0.1%) |
| 22 | O | 1.05 | 9/3243 (0.3%) | 1.09 | 6/4374 (0.1%) |
| 23 | P | 1.13 | 14/3452 (0.4%) | 1.16 | 8/4657 (0.2%) |
| 24 | Q | 1.08 | 12/3527 (0.3%) | 1.11 | 5/4748 (0.1%) |
| 25 | R | 1.09 | 9/3272 (0.3%) | 1.09 | 3/4412 (0.1%) |
| 26 | S | 1.22 | 14/2945 (0.5%) | 1.12 | 5/3976 (0.1%) |
| 27 | T | 1.08 | 7/2279 (0.3%) | 1.10 | 1/3077 (0.0%) |
| 28 | U | 1.09 | 12/2087 (0.6%) | 1.14 | 2/2811 (0.1%) |
| 29 | V | 1.32 | 13/1969 (0.7%) | 1.29 | 8/2652 (0.3%) |
| 3 | 3 | 1.13 | 6/1611 (0.4%) | 1.18 | 3/2174 (0.1%) |
| 30 | W | 1.02 | 8/1557 (0.5%) | 1.13 | 6/2111 (0.3%) |
| 31 | X | 1.03 | 4/1058 (0.4%) | 1.32 | 6/1432 (0.4%) |
| 32 | Y | 1.13 | 0/169 | 0.93 | 0/223 |
| 33 | Z | 1.13 | 32/6403 (0.5%) | 1.14 | 7/8686 (0.1%) |
| 4 | 4 | 1.14 | 8/1613 (0.5%) | 1.15 | 0/2173 |
| 5 | 5 | 1.19 | 8/1683 (0.5%) | 1.23 | 4/2277 (0.2%) |
| 6 | 6 | 1.17 | 11/1795 (0.6%) | 1.22 | 3/2420 (0.1%) |
| 7 | 7 | 1.09 | 3/1855 (0.2%) | 1.19 | 5/2514 (0.2%) |
| 8 | A | 1.16 | 11/1959 (0.6%) | 1.16 | 5/2652 (0.2%) |
| 9 | B | 1.16 | 5/1952 (0.3%) | 1.21 | 3/2642 (0.1%) |
| All | All | 1.15 | 347/81465 (0.4%) | 1.17 | 161/109981 (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 | 1 | 0 | 1 |
| 14 | G | 0 | 1 |
| 15 | H | 0 | 7 |
| 16 | I | 0 | 1 |
| 17 | J | 0 | 4 |
| 18 | K | 0 | 2 |
| 19 | L | 0 | 4 |
| 2 | 2 | 0 | 1 |
| 20 | M | 0 | 2 |
| 21 | N | 0 | 1 |
| 22 | O | 0 | 2 |
| 23 | P | 0 | 2 |
| 24 | Q | 0 | 1 |
| 25 | R | 0 | 2 |
| 26 | S | 0 | 1 |
| 27 | T | 0 | 1 |
| 28 | U | 0 | 1 |
| 29 | V | 0 | 1 |
| 30 | W | 0 | 6 |
| 31 | X | 0 | 4 |
| 33 | Z | 0 | 7 |
| 9 | B | 0 | 1 |
| All | All | 0 | 53 |

The worst 5 of 347 bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|-----|-------|-----|------|-------|--------|-------------|----------|
| 15 | H | 285 | GLY | C-N | 27.78 | 1.98 | 1.34 |
| 26 | S | 257 | LEU | CA-C | 22.13 | 2.10 | 1.52 |
| 29 | V | 290 | ASN | C-N | 21.67 | 1.83 | 1.34 |
| 26 | S | 257 | LEU | N-CA | -20.09 | 1.06 | 1.46 |
| 20 | M | 257 | GLY | C-N | 15.26 | 1.69 | 1.34 |

The worst 5 of 161 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(°) | Ideal(°) |
|-----|-------|-----|------|--------|--------|-------------|----------|
| 15 | H | 187 | LEU | C-N-CD | -30.74 | 52.98 | 120.60 |
| 31 | X | 27 | ILE | C-N-CD | -23.01 | 69.98 | 120.60 |

Continued on next page...

Continued from previous page...

| Mol | Chain | Res | Type | Atoms | Z | Observed(°) | Ideal(°) |
|-----|-------|-----|------|-----------|--------|-------------|----------|
| 29 | V | 291 | ASN | O-C-N | 16.75 | 149.51 | 122.70 |
| 15 | H | 285 | GLY | C-N-CA | -13.58 | 87.74 | 121.70 |
| 15 | H | 162 | ARG | NE-CZ-NH2 | 12.91 | 126.76 | 120.30 |

There are no chirality outliers.

5 of 53 planarity outliers are listed below:

| Mol | Chain | Res | Type | Group |
|-----|-------|-----|------|---------|
| 1 | 1 | -2 | LEU | Peptide |
| 2 | 2 | 202 | SER | Peptide |
| 9 | B | 9 | LEU | Peptide |
| 14 | G | 246 | ILE | Peptide |
| 15 | H | 101 | ARG | 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 | 1 | 1576 | 0 | 1554 | 146 | 0 |
| 2 | 2 | 1692 | 0 | 1698 | 94 | 0 |
| 3 | 3 | 1581 | 0 | 1573 | 103 | 0 |
| 4 | 4 | 1585 | 0 | 1590 | 82 | 0 |
| 5 | 5 | 1646 | 0 | 1595 | 137 | 0 |
| 6 | 6 | 1757 | 0 | 1711 | 129 | 0 |
| 7 | 7 | 1824 | 0 | 1832 | 150 | 0 |
| 8 | A | 1921 | 0 | 1910 | 220 | 0 |
| 9 | B | 1915 | 0 | 1929 | 136 | 0 |
| 10 | C | 1913 | 0 | 1914 | 148 | 0 |
| 11 | D | 1899 | 0 | 1908 | 131 | 0 |
| 12 | E | 1867 | 0 | 1841 | 164 | 0 |
| 13 | F | 1795 | 0 | 1796 | 204 | 0 |
| 14 | G | 1900 | 0 | 1889 | 216 | 0 |
| 15 | H | 2792 | 0 | 2875 | 265 | 0 |
| 16 | I | 2822 | 0 | 2869 | 277 | 0 |
| 17 | J | 2928 | 0 | 3056 | 275 | 0 |
| 18 | K | 3019 | 0 | 3082 | 327 | 0 |
| 19 | L | 2853 | 0 | 2925 | 329 | 0 |

Continued on next page...

Continued from previous page...

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|
| 20 | M | 2866 | 0 | 2937 | 301 | 0 |
| 21 | N | 6562 | 0 | 6624 | 555 | 0 |
| 22 | O | 3182 | 0 | 3207 | 543 | 0 |
| 23 | P | 3401 | 0 | 3483 | 346 | 0 |
| 24 | Q | 3471 | 0 | 3491 | 306 | 0 |
| 25 | R | 3218 | 0 | 3216 | 385 | 0 |
| 26 | S | 2893 | 0 | 2937 | 262 | 0 |
| 27 | T | 2235 | 0 | 2207 | 202 | 0 |
| 28 | U | 2061 | 0 | 2114 | 267 | 0 |
| 29 | V | 1942 | 0 | 1948 | 258 | 0 |
| 30 | W | 1534 | 0 | 1542 | 141 | 0 |
| 31 | X | 1032 | 0 | 1015 | 249 | 0 |
| 32 | Y | 168 | 0 | 153 | 6 | 0 |
| 33 | Z | 6289 | 0 | 6232 | 673 | 0 |
| All | All | 80139 | 0 | 80653 | 7035 | 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 44.

The worst 5 of 7035 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) |
|-------------------|-------------------|--------------------------|-------------------|
| 23:P:427:GLU:HA | 29:V:230:TYR:CE1 | 1.27 | 1.67 |
| 22:O:185:PHE:CE2 | 22:O:223:LEU:HD13 | 1.30 | 1.66 |
| 22:O:310:PHE:HZ | 22:O:341:ILE:CG2 | 1.04 | 1.62 |
| 12:E:12:VAL:HG12 | 12:E:123:PHE:CE2 | 1.37 | 1.60 |
| 33:Z:233:LEU:HD22 | 33:Z:268:ALA:CB | 1.32 | 1.59 |

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 PDB entries followed by that with respect to all EM entries.

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 | 1 | 203/215 (94%) | 189 (93%) | 12 (6%) | 2 (1%) | 19 | 65 |
| 2 | 2 | 221/261 (85%) | 209 (95%) | 9 (4%) | 3 (1%) | 14 | 58 |
| 3 | 3 | 202/205 (98%) | 179 (89%) | 14 (7%) | 9 (4%) | 3 | 33 |
| 4 | 4 | 196/198 (99%) | 178 (91%) | 15 (8%) | 3 (2%) | 13 | 57 |
| 5 | 5 | 210/287 (73%) | 199 (95%) | 8 (4%) | 3 (1%) | 14 | 58 |
| 6 | 6 | 220/241 (91%) | 207 (94%) | 9 (4%) | 4 (2%) | 11 | 53 |
| 7 | 7 | 231/266 (87%) | 208 (90%) | 18 (8%) | 5 (2%) | 8 | 49 |
| 8 | A | 241/252 (96%) | 222 (92%) | 16 (7%) | 3 (1%) | 16 | 61 |
| 9 | B | 248/250 (99%) | 235 (95%) | 11 (4%) | 2 (1%) | 24 | 69 |
| 10 | C | 243/258 (94%) | 229 (94%) | 13 (5%) | 1 (0%) | 39 | 80 |
| 11 | D | 240/254 (94%) | 222 (92%) | 15 (6%) | 3 (1%) | 15 | 60 |
| 12 | E | 241/260 (93%) | 224 (93%) | 11 (5%) | 6 (2%) | 7 | 46 |
| 13 | F | 231/234 (99%) | 216 (94%) | 13 (6%) | 2 (1%) | 21 | 67 |
| 14 | G | 243/288 (84%) | 224 (92%) | 17 (7%) | 2 (1%) | 24 | 69 |
| 15 | H | 351/467 (75%) | 297 (85%) | 37 (10%) | 17 (5%) | 3 | 32 |
| 16 | I | 358/437 (82%) | 322 (90%) | 28 (8%) | 8 (2%) | 8 | 49 |
| 17 | J | 371/405 (92%) | 336 (91%) | 24 (6%) | 11 (3%) | 5 | 42 |
| 18 | K | 377/428 (88%) | 343 (91%) | 23 (6%) | 11 (3%) | 6 | 43 |
| 19 | L | 357/437 (82%) | 320 (90%) | 28 (8%) | 9 (2%) | 7 | 46 |
| 20 | M | 363/434 (84%) | 331 (91%) | 22 (6%) | 10 (3%) | 6 | 44 |
| 21 | N | 843/945 (89%) | 772 (92%) | 45 (5%) | 26 (3%) | 5 | 42 |
| 22 | O | 385/393 (98%) | 332 (86%) | 26 (7%) | 27 (7%) | 1 | 22 |
| 23 | P | 413/445 (93%) | 377 (91%) | 19 (5%) | 17 (4%) | 3 | 35 |
| 24 | Q | 429/434 (99%) | 398 (93%) | 20 (5%) | 11 (3%) | 7 | 45 |
| 25 | R | 398/429 (93%) | 351 (88%) | 36 (9%) | 11 (3%) | 6 | 44 |
| 26 | S | 351/523 (67%) | 318 (91%) | 19 (5%) | 14 (4%) | 4 | 35 |
| 27 | T | 270/274 (98%) | 233 (86%) | 25 (9%) | 12 (4%) | 3 | 33 |
| 28 | U | 245/338 (72%) | 237 (97%) | 5 (2%) | 3 (1%) | 16 | 61 |
| 29 | V | 239/306 (78%) | 214 (90%) | 15 (6%) | 10 (4%) | 3 | 34 |
| 30 | W | 195/268 (73%) | 171 (88%) | 12 (6%) | 12 (6%) | 2 | 26 |
| 31 | X | 125/156 (80%) | 83 (66%) | 18 (14%) | 24 (19%) | 0 | 3 |
| 32 | Y | 17/89 (19%) | 17 (100%) | 0 | 0 | 100 | 100 |

Continued on next page...

Continued from previous page...

| Mol | Chain | Analysed | Favoured | Allowed | Outliers | Percentiles | |
|-----|-------|-------------------|------------|----------|----------|-------------|----|
| 33 | Z | 807/993 (81%) | 714 (88%) | 59 (7%) | 34 (4%) | 3 | 34 |
| All | All | 10064/11670 (86%) | 9107 (90%) | 642 (6%) | 315 (3%) | 9 | 42 |

5 of 315 Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 2 | 2 | 196 | ARG |
| 3 | 3 | 194 | ARG |
| 5 | 5 | 31 | VAL |
| 7 | 7 | 209 | MET |
| 10 | C | 224 | GLU |

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 PDB entries followed by that with respect to all EM entries.

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 | 1 | 169/178 (95%) | 169 (100%) | 0 | 100 | 100 |
| 2 | 2 | 182/214 (85%) | 181 (100%) | 1 (0%) | 92 | 96 |
| 3 | 3 | 172/173 (99%) | 172 (100%) | 0 | 100 | 100 |
| 4 | 4 | 175/175 (100%) | 175 (100%) | 0 | 100 | 100 |
| 5 | 5 | 169/235 (72%) | 169 (100%) | 0 | 100 | 100 |
| 6 | 6 | 185/201 (92%) | 185 (100%) | 0 | 100 | 100 |
| 7 | 7 | 199/224 (89%) | 199 (100%) | 0 | 100 | 100 |
| 8 | A | 207/210 (99%) | 207 (100%) | 0 | 100 | 100 |
| 9 | B | 209/209 (100%) | 208 (100%) | 1 (0%) | 92 | 96 |
| 10 | C | 204/216 (94%) | 204 (100%) | 0 | 100 | 100 |
| 11 | D | 214/226 (95%) | 214 (100%) | 0 | 100 | 100 |
| 12 | E | 199/215 (93%) | 199 (100%) | 0 | 100 | 100 |
| 13 | F | 192/193 (100%) | 191 (100%) | 1 (0%) | 92 | 96 |
| 14 | G | 201/239 (84%) | 200 (100%) | 1 (0%) | 92 | 96 |
| 15 | H | 303/399 (76%) | 300 (99%) | 3 (1%) | 82 | 92 |

Continued on next page...

Continued from previous page...

| Mol | Chain | Analysed | Rotameric | Outliers | Percentiles | |
|-----|-------|------------------|-------------|----------|-------------|-----|
| 16 | I | 319/385 (83%) | 316 (99%) | 3 (1%) | 84 | 93 |
| 17 | J | 325/352 (92%) | 325 (100%) | 0 | 100 | 100 |
| 18 | K | 334/374 (89%) | 334 (100%) | 0 | 100 | 100 |
| 19 | L | 308/377 (82%) | 306 (99%) | 2 (1%) | 90 | 95 |
| 20 | M | 315/375 (84%) | 314 (100%) | 1 (0%) | 94 | 96 |
| 21 | N | 713/797 (90%) | 711 (100%) | 2 (0%) | 94 | 96 |
| 22 | O | 363/368 (99%) | 363 (100%) | 0 | 100 | 100 |
| 23 | P | 388/415 (94%) | 385 (99%) | 3 (1%) | 86 | 94 |
| 24 | Q | 388/391 (99%) | 388 (100%) | 0 | 100 | 100 |
| 25 | R | 351/379 (93%) | 350 (100%) | 1 (0%) | 94 | 96 |
| 26 | S | 330/489 (68%) | 329 (100%) | 1 (0%) | 94 | 96 |
| 27 | T | 254/256 (99%) | 254 (100%) | 0 | 100 | 100 |
| 28 | U | 234/308 (76%) | 234 (100%) | 0 | 100 | 100 |
| 29 | V | 217/268 (81%) | 217 (100%) | 0 | 100 | 100 |
| 30 | W | 171/230 (74%) | 169 (99%) | 2 (1%) | 78 | 90 |
| 31 | X | 116/144 (81%) | 116 (100%) | 0 | 100 | 100 |
| 32 | Y | 18/81 (22%) | 18 (100%) | 0 | 100 | 100 |
| 33 | Z | 692/850 (81%) | 691 (100%) | 1 (0%) | 95 | 97 |
| All | All | 8816/10146 (87%) | 8793 (100%) | 23 (0%) | 95 | 96 |

5 of 23 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 19 | L | 192 | GLU |
| 20 | M | 433 | TYR |
| 30 | W | 22 | PRO |
| 19 | L | 195 | GLU |
| 21 | N | 740 | TRP |

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 166 such sidechains are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 19 | L | 320 | GLN |
| 22 | O | 75 | GLN |
| 33 | Z | 214 | HIS |

Continued on next page...

Continued from previous page...

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 20 | M | 311 | GLN |
| 21 | N | 430 | ASN |

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 carbohydrates 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.