# wwPDB EM Validation Summary Report ⓘ

Nov 11, 2024 – 03:45 PM JST

| | | |
|---|---|---|
| PDB ID | : | 7EGE |
| EMDB ID | : | EMD-31114 |
| Title | : | TFIID in canonical conformation |
| Authors | : | Chen, X.; Wu, Z.; Li, J.; Zhao, D.; Xu, Y. |
| Deposited on | : | 2021-03-24 |
| Resolution | : | 9.00 Å(reported) |

---

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

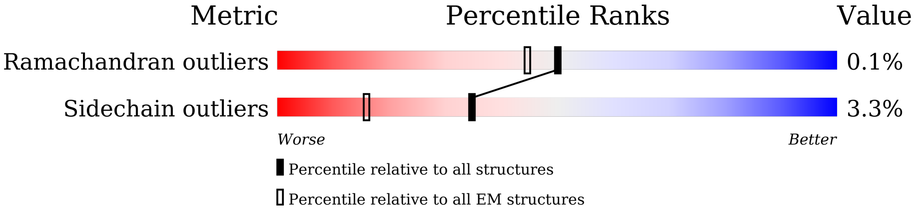
| | | |
|---|---|---|
| EMDB validation analysis | : | 0.0.1.dev113 |
| MolProbity | : | 4.02b-467 |
| Percentile statistics | : | 20231227.v01 (using entries in the PDB archive December 27th 2023) |
| MapQ | : | 1.9.13 |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | 2.39 |

# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
*ELECTRON MICROSCOPY*

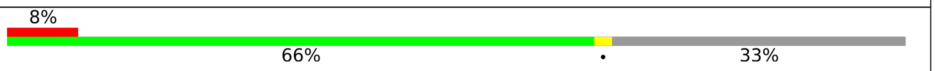
The reported resolution of this entry is 9.00 Å.

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.



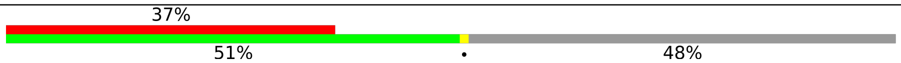
| Metric | Whole archive (#Entries) | EM structures (#Entries) |
|---|---|---|
| Ramachandran outliers | 207382 | 16835 |
| Sidechain outliers | 206894 | 16415 |

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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 EM map (all-atom inclusion < 40%). The numeric value is given above the bar.

| Mol | Chain | Length | Quality of chain |
|---|---|---|---|
| 1 | A | 1872 |  |
| 2 | B | 1199 |  |
| 3 | D | 1085 |  |
| 3 | d | 1085 |  |
| 4 | E | 800 |  |
| 4 | e | 800 |  |
| 5 | F | 677 |  |
| 5 | f | 677 |  |
| 6 | G | 349 |  |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Length | Quality of chain |
|---|---|---|---|
| 7 | H | 310 |  |
| 8 | I | 264 |  |
| 8 | i | 264 |  |
| 9 | J | 218 |  |
| 9 | j | 218 |  |
| 10 | L | 161 |  |
| 10 | l | 161 |  |
| 11 | P | 339 |  |
| 12 | c | 929 |  |
| 13 | k | 211 |  |
| 14 | m | 124 |  |

## 2 Entry composition ⓘ

There are 14 unique types of molecules in this entry. The entry contains 41435 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 Transcription initiation factor TFIID subunit 1.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | |
| 1 | A | 549 | 4505 | 2879 | 781 | 818 | 27 | 0 | 0 |

- Molecule 2 is a protein called Transcription initiation factor TFIID subunit 2.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | |
| 2 | B | 963 | 7796 | 5011 | 1315 | 1412 | 58 | 0 | 0 |

- Molecule 3 is a protein called Transcription initiation factor TFIID subunit 4.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | |
| 3 | D | 154 | 1266 | 789 | 231 | 243 | 3 | 0 | 0 |
| | | | Total | C | N | O | S | | |
| 3 | d | 158 | 1307 | 814 | 238 | 252 | 3 | 0 | 0 |

- Molecule 4 is a protein called Transcription initiation factor TFIID subunit 5.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | |
| 4 | E | 547 | 4376 | 2774 | 759 | 822 | 21 | 0 | 0 |
| | | | Total | C | N | O | S | | |
| 4 | e | 539 | 4327 | 2746 | 748 | 814 | 19 | 0 | 0 |

- Molecule 5 is a protein called Transcription initiation factor TFIID subunit 6.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | |
| 5 | F | 412 | 3149 | 1997 | 551 | 583 | 18 | 0 | 0 |
| | | | Total | C | N | O | S | | |
| 5 | f | 403 | 3081 | 1954 | 533 | 576 | 18 | 0 | 0 |

- Molecule 6 is a protein called Transcription initiation factor TFIID subunit 7.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 6 | G | 144 | Total | C | N | O | S | 0 | 0 |
| | | | 1171 | 742 | 215 | 210 | 4 | | |

- Molecule 7 is a protein called Transcription initiation factor TFIID subunit 8.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 7 | H | 209 | Total | C | N | O | S | 0 | 0 |
| | | | 1622 | 1026 | 281 | 310 | 5 | | |

- Molecule 8 is a protein called Transcription initiation factor TFIID subunit 9.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 8 | I | 120 | Total | C | N | O | S | 0 | 0 |
| | | | 959 | 610 | 166 | 177 | 6 | | |
| 8 | i | 121 | Total | C | N | O | S | 0 | 0 |
| | | | 967 | 615 | 167 | 178 | 7 | | |

- Molecule 9 is a protein called Transcription initiation factor TFIID subunit 10.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 9 | J | 90 | Total | C | N | O | S | 0 | 0 |
| | | | 720 | 466 | 115 | 135 | 4 | | |
| 9 | j | 95 | Total | C | N | O | S | 0 | 0 |
| | | | 759 | 488 | 124 | 143 | 4 | | |

- Molecule 10 is a protein called Transcription initiation factor TFIID subunit 12.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 10 | L | 76 | Total | C | N | O | S | 0 | 0 |
| | | | 622 | 388 | 109 | 122 | 3 | | |
| 10 | l | 107 | Total | C | N | O | S | 0 | 0 |
| | | | 876 | 547 | 158 | 166 | 5 | | |

- Molecule 11 is a protein called TATA-box-binding protein.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 11 | P | 177 | Total | C | N | O | S | 0 | 0 |
| | | | 1412 | 918 | 249 | 238 | 7 | | |

- Molecule 12 is a protein called Transcription initiation factor TFIID subunit 3.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | |
| 12 | c | 127 | 1011 | 638 | 174 | 193 | 6 | 0 | 0 |

- Molecule 13 is a protein called Transcription initiation factor TFIID subunit 11.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | |
| 13 | k | 98 | 785 | 499 | 142 | 139 | 5 | 0 | 0 |

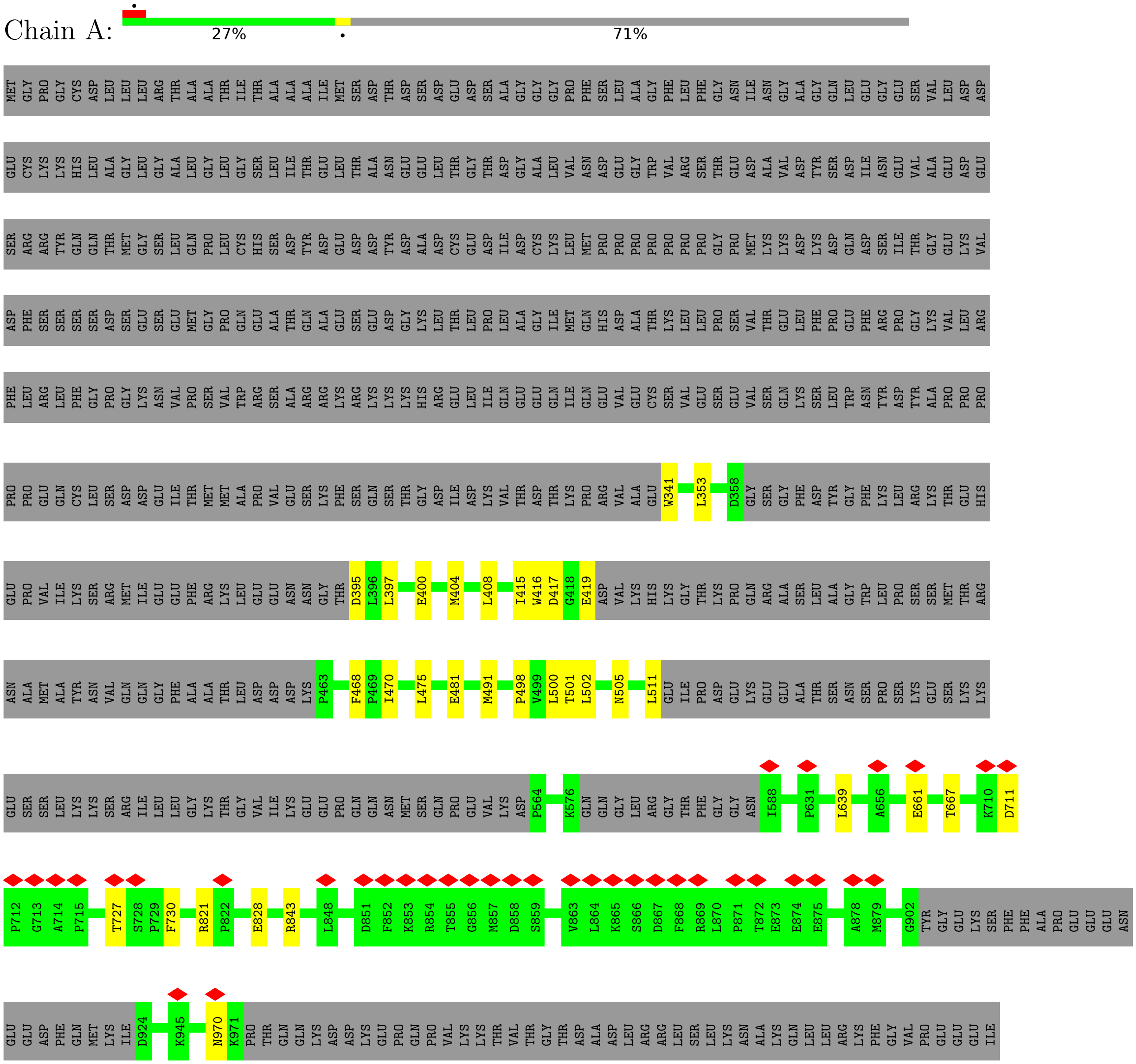
- Molecule 14 is a protein called Transcription initiation factor TFIID subunit 13.

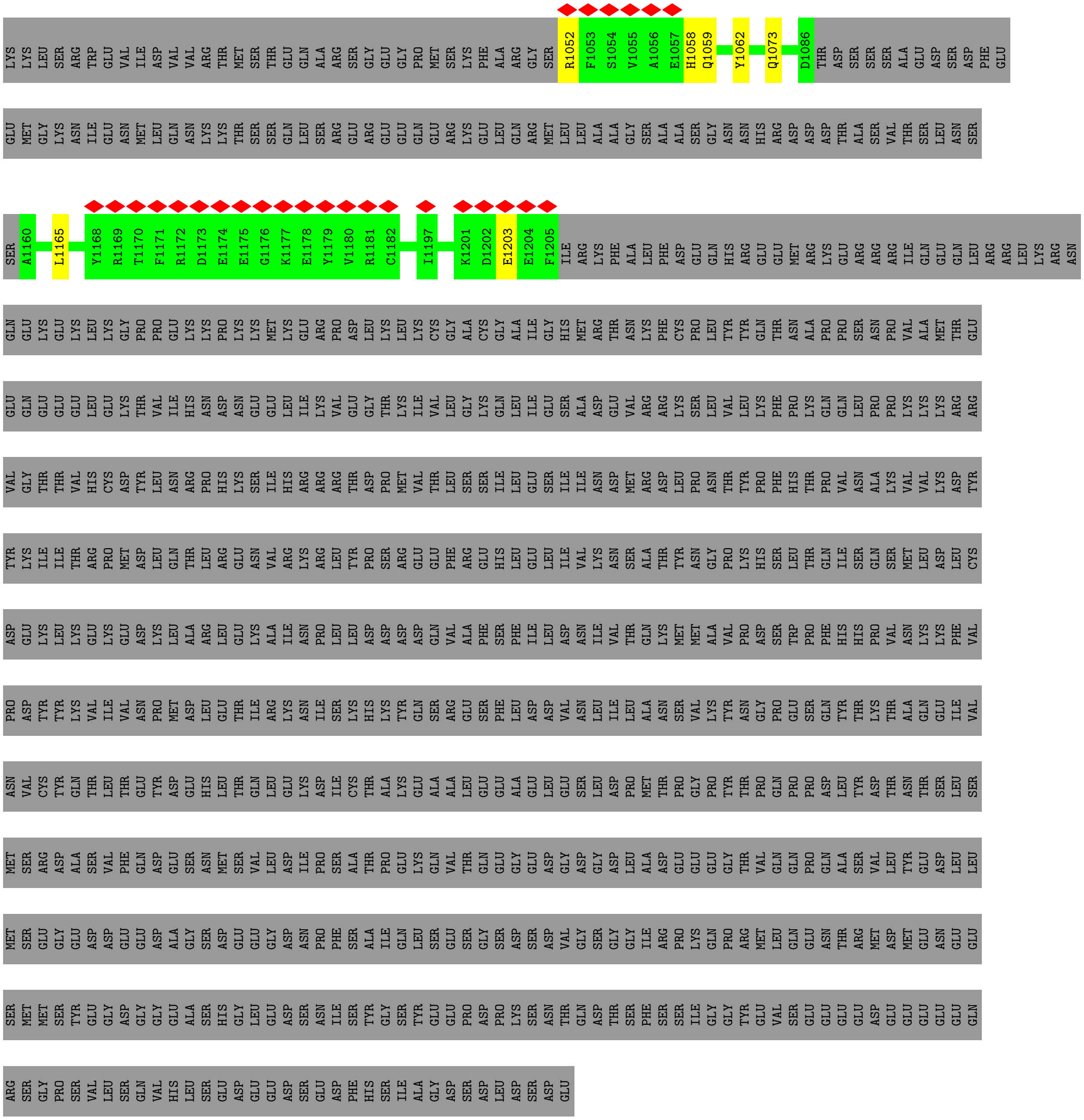
| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | |
| 14 | m | 87 | 724 | 456 | 131 | 131 | 6 | 0 | 0 |

## 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 and atom inclusion in map 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 diamond above a residue indicates a poor fit to the EM map for this residue (all-atom inclusion < 40%). 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.

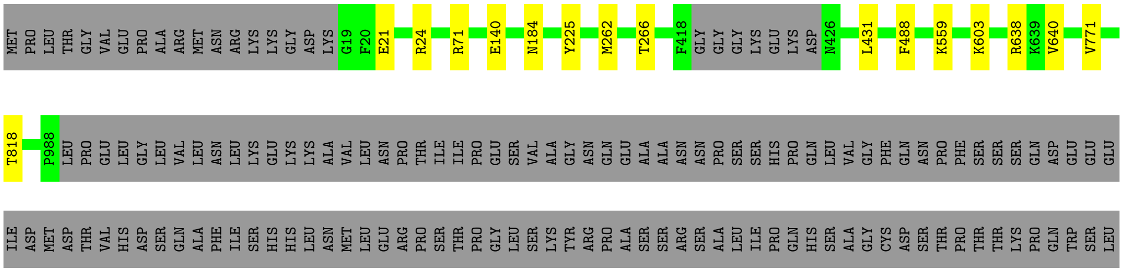
- Molecule 1: Transcription initiation factor TFIID subunit 1

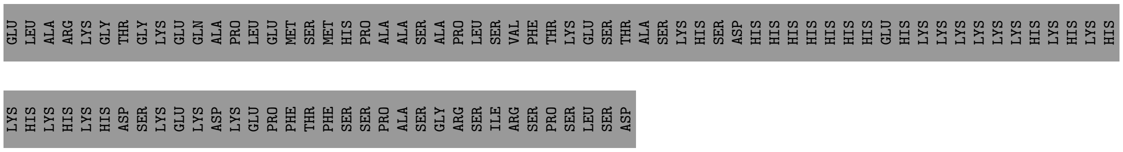




- Molecule 2: Transcription initiation factor TFIID subunit 2

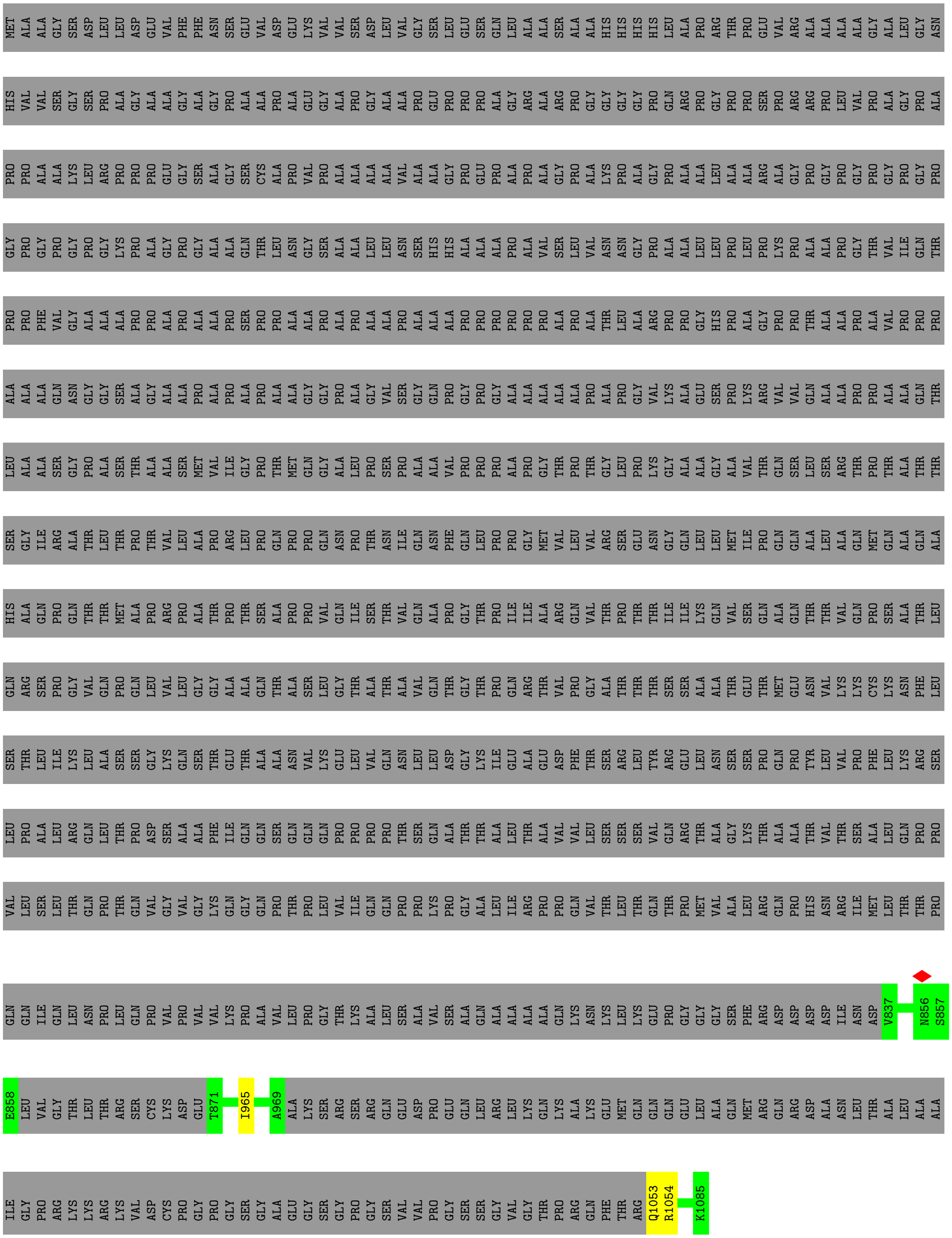
Chain B:





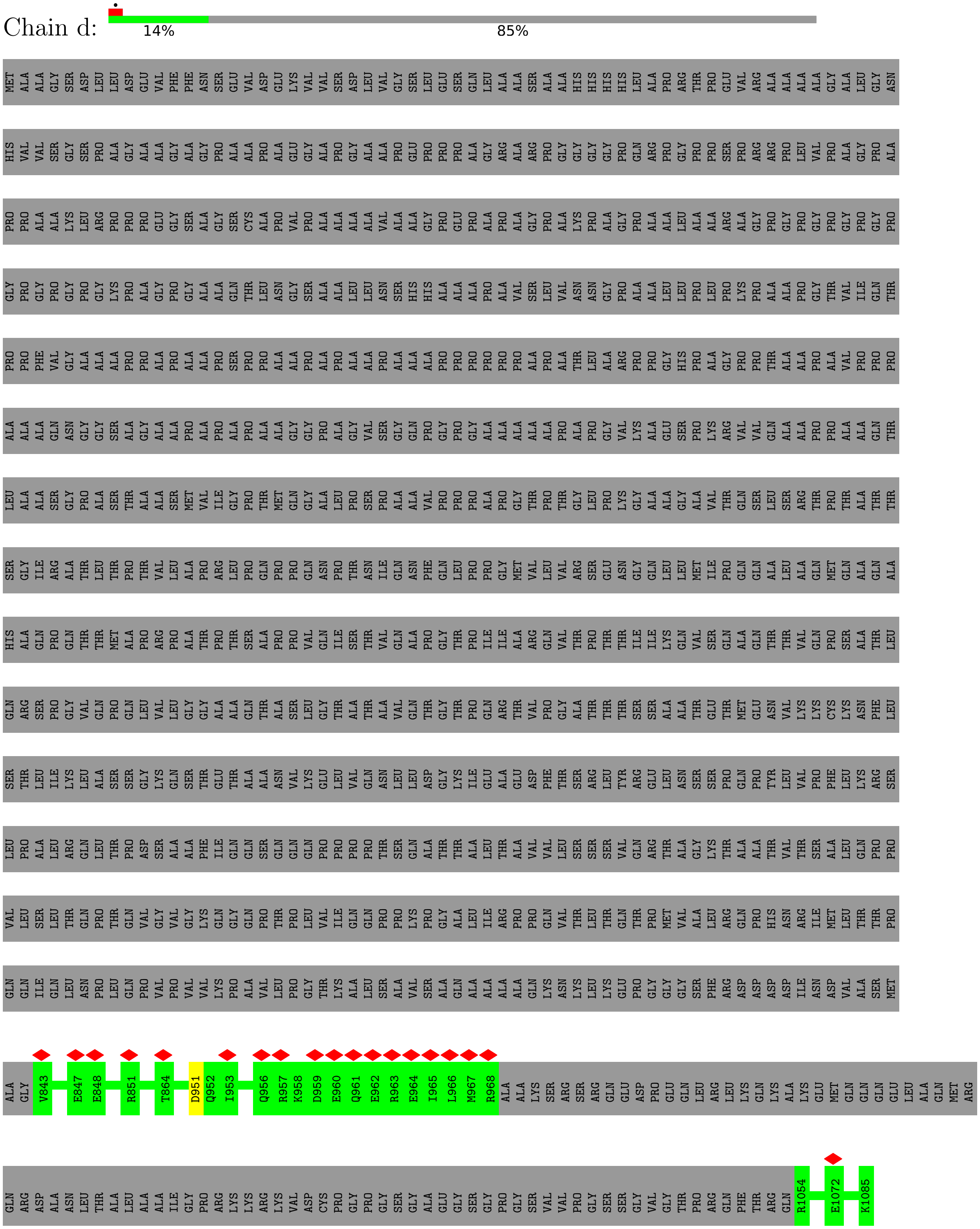
- Molecule 3: Transcription initiation factor TFIID subunit 4

Chain D:



- Molecule 3: Transcription initiation factor TFIID subunit 4

Chain d:



MET ALA ALA GLY SER ASP LEU LEU ASP GLU VAL PHE PHE ASN SER GLU VAL ASP GLU LYS VAL VAL SER ASP LEU VAL GLY SER LEU GLU SER GLN LEU ALA ALA SER ALA ALA HIS HIS HIS HIS LEU ALA PRO ARG THR PRO GLU VAL ARG ALA ALA ALA ALA GLY ALA LEU GLY ASN

HIS VAL VAL SER GLY SER PRO ALA GLY ALA ALA GLY ALA GLY PRO ALA ALA PRO ALA GLU GLY ALA PRO GLY ALA ALA PRO GLU PRO PRO ALA GLY ARG ALA ARG PRO GLY GLY GLY GLY PRO GLN ARG PRO GLY PRO PRO SER PRO ARG ARG PRO LEU VAL PRO ALA GLY PRO ALA

PRO PRO ALA ALA LYS LEU ARG PRO PRO PRO GLU GLY SER ALA GLY SER CYS ALA PRO VAL PRO ALA ALA ALA ALA VAL ALA ALA GLY PRO GLU PRO ALA PRO ALA GLY PRO ALA LYS PRO ALA GLY PRO ALA ALA LEU ALA ALA ARG ALA GLY PRO GLY PRO GLY PRO GLY PRO GLY PRO

GLY PRO GLY PRO GLY PRO GLY LYS PRO ALA GLY PRO GLY ALA ALA GLN THR LEU ASN GLY SER ALA ALA LEU LEU ASN SER HIS HIS ALA ALA ALA PRO ALA VAL SER LEU VAL ASN ASN GLY PRO ALA ALA LEU LEU PRO LEU PRO LYS PRO ALA ALA PRO GLY THR VAL ILE GLN THR

PRO PRO PHE VAL GLY ALA ALA ALA PRO PRO ALA PRO ALA ALA PRO SER PRO PRO ALA ALA PRO ALA PRO ALA ALA PRO ALA ALA ALA PRO PRO PRO PRO PRO PRO ALA PRO ALA THR LEU ALA ARG PRO PRO GLY HIS PRO ALA GLY PRO PRO THR ALA ALA PRO ALA VAL PRO PRO PRO

ALA ALA ALA GLN ASN GLY GLY SER ALA GLY ALA ALA PRO ALA PRO ALA PRO ALA ALA GLY GLY PRO ALA GLY VAL SER GLY GLN PRO GLY PRO GLY ALA ALA ALA ALA ALA PRO ALA PRO GLY VAL LYS ALA GLU SER PRO LYS ARG VAL VAL GLN ALA ALA PRO PRO ALA ALA GLN THR

LEU ALA ALA SER GLY PRO ALA SER THR ALA ALA SER MET VAL ILE GLY PRO THR MET GLN GLY ALA LEU PRO SER PRO ALA ALA VAL PRO PRO PRO ALA PRO GLY THR PRO THR GLY LEU PRO LYS GLY ALA ALA GLY ALA VAL THR GLN SER LEU SER ARG THR PRO THR ALA THR THR

SER GLY ILE ARG ALA THR LEU THR PRO THR VAL LEU ALA PRO ARG LEU PRO GLN PRO PRO GLN ASN PRO THR ASN ILE GLN ASN PHE GLN LEU PRO PRO GLY MET VAL LEU VAL ARG SER GLU ASN GLY GLN LEU LEU MET ILE PRO GLN GLN ALA LEU ALA GLN MET GLN ALA GLN ALA

HIS ALA GLN PRO GLN THR THR MET ALA PRO ARG PRO ALA THR PRO THR SER ALA PRO PRO VAL GLN ILE SER THR VAL GLN ALA PRO GLY THR PRO ILE ILE ALA ARG GLN VAL THR PRO THR THR ILE ILE LYS GLN VAL SER GLN ALA GLN THR THR VAL GLN PRO SER ALA THR LEU

GLN ARG SER PRO GLY VAL GLN PRO GLN LEU VAL LEU GLY GLY ALA ALA GLN THR ALA SER LEU GLY THR ALA THR ALA VAL GLN THR GLY THR PRO GLN ARG THR VAL PRO GLY ALA THR THR THR SER SER ALA ALA THR GLU THR MET GLU ASN VAL LYS LYS CYS LYS ASN PHE LEU

SER THR LEU ILE LYS LEU ALA SER SER GLY LYS GLN SER THR GLU THR ALA ALA ASN VAL LYS GLU LEU VAL GLN ASN LEU LEU ASP GLY LYS ILE GLU ALA GLU ASP PHE THR SER ARG LEU TYR ARG GLU LEU ASN SER SER PRO GLN PRO TYR LEU VAL PRO PHE LEU LYS ARG SER

LEU PRO ALA LEU ARG GLN LEU THR PRO ASP SER ALA ALA PHE ILE GLN GLN SER GLN GLN GLN PRO PRO PRO THR SER GLN ALA THR THR ALA LEU THR ALA VAL VAL LEU SER SER SER VAL GLN ARG THR ALA GLY LYS THR ALA ALA THR VAL THR SER ALA LEU GLN PRO PRO

VAL LEU SER LEU THR GLN PRO THR GLN VAL GLY VAL GLY LYS GLN GLY GLN PRO THR PRO LEU VAL ILE GLN GLN PRO PRO LYS PRO GLY ALA LEU ILE ARG PRO PRO GLN VAL THR LEU THR GLN THR PRO MET VAL ALA LEU ARG GLN PRO HIS ASN ARG ILE MET LEU THR THR PRO

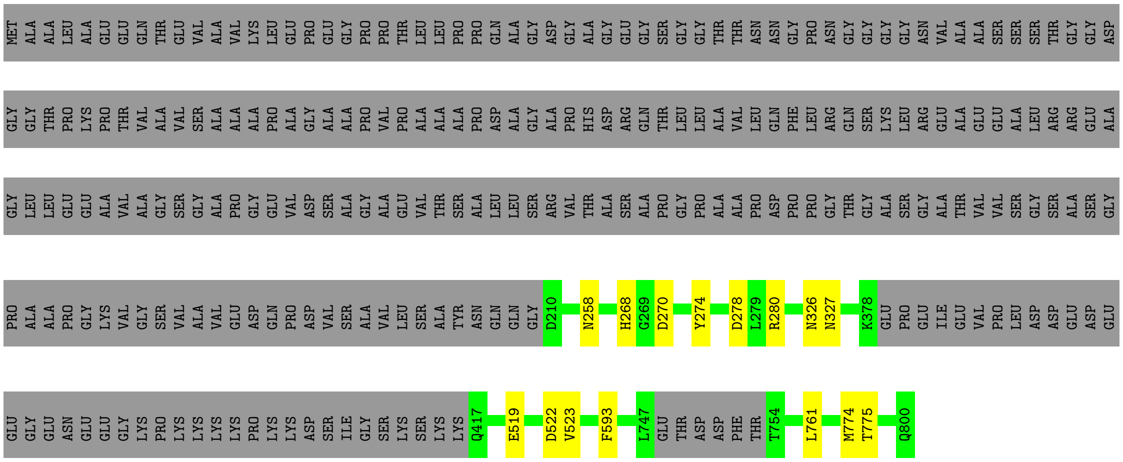
GLN GLN ILE GLN LEU ASN PRO LEU GLN PRO VAL PRO VAL VAL LYS PRO ALA VAL LEU PRO GLY THR LYS ALA LEU SER ALA VAL SER ALA GLN ALA ALA ALA ALA GLN LYS ASN LYS LEU LYS GLU PRO GLY GLY GLY SER PHE ARG ASP ASP ASP ASP ILE ASN ASP VAL ALA SER MET

ALA GLY V843 E847 E848 R851 T864 D951 Q952 I953 Q956 R957 K958 D959 E960 Q961 E962 R963 E964 I965 L966 M967 R968 ALA ALA LYS SER ARG SER ARG GLN GLU ASP PRO GLU GLN LEU ARG LEU LYS GLN LYS ALA LYS GLU MET GLN GLN GLN GLU LEU ALA GLN MET ARG

GLN ARG ASP ALA ASN LEU THR ALA LEU ALA ALA ILE GLY PRO ARG LYS LYS ARG LYS VAL ASP CYS PRO GLY PRO GLY SER GLY ALA GLU GLY SER GLY PRO GLY SER VAL VAL PRO GLY SER SER GLY VAL GLY THR PRO ARG GLN PHE THR ARG GLN R1054 E1072 K1085

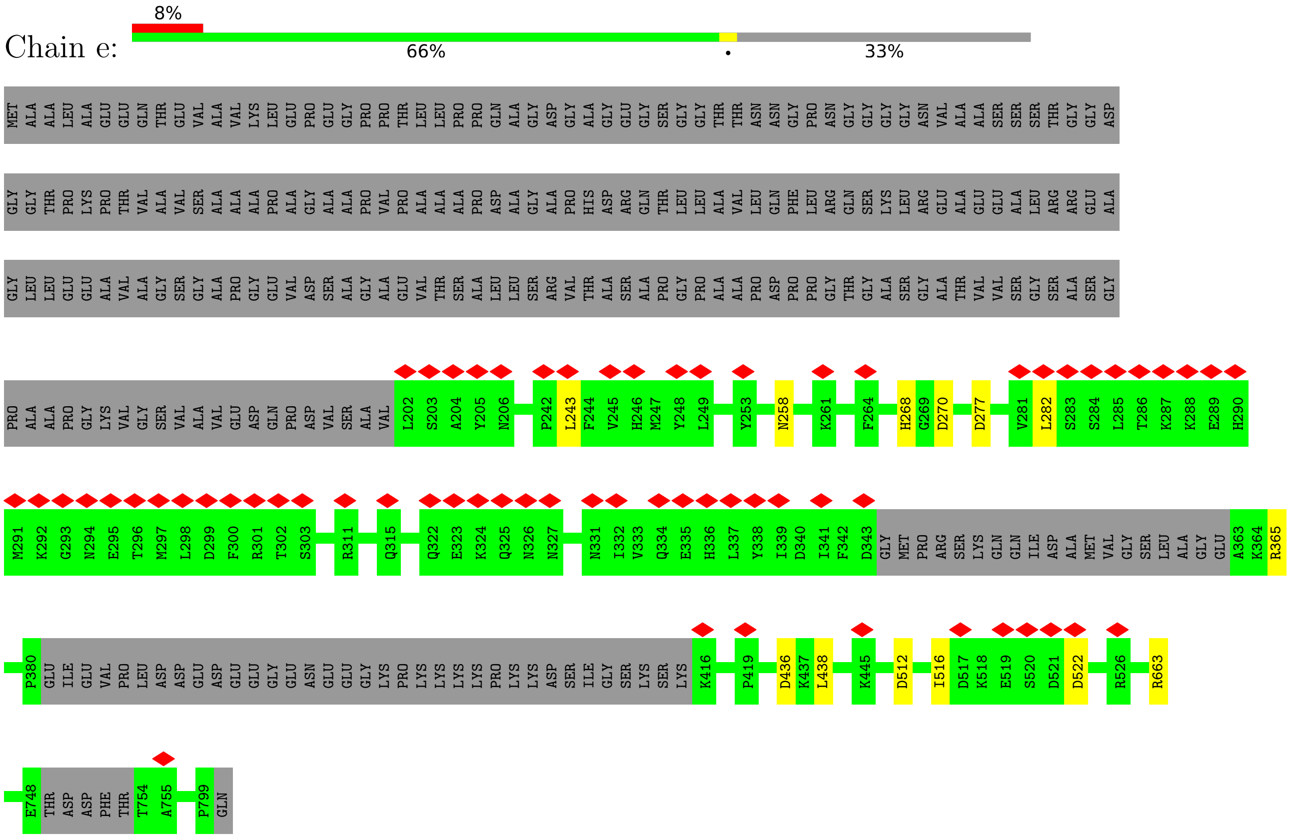
- Molecule 4: Transcription initiation factor TFIID subunit 5

Chain E:



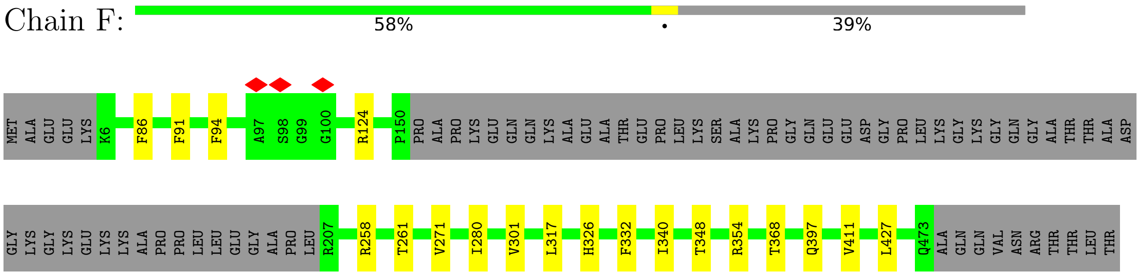
- Molecule 4: Transcription initiation factor TFIID subunit 5

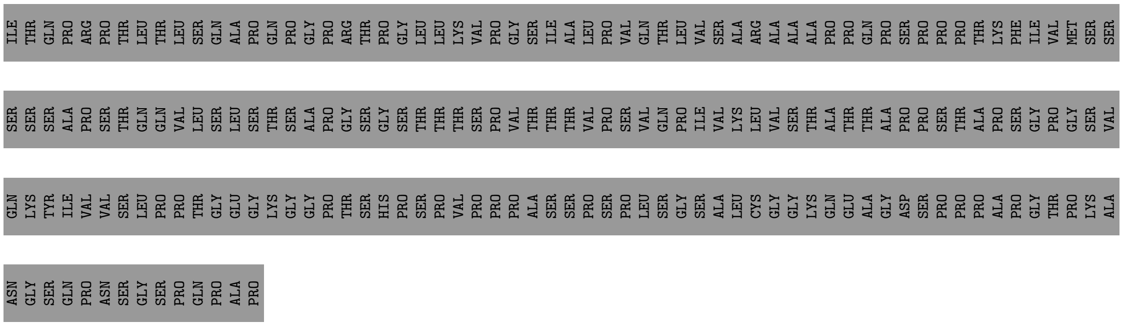
Chain e:



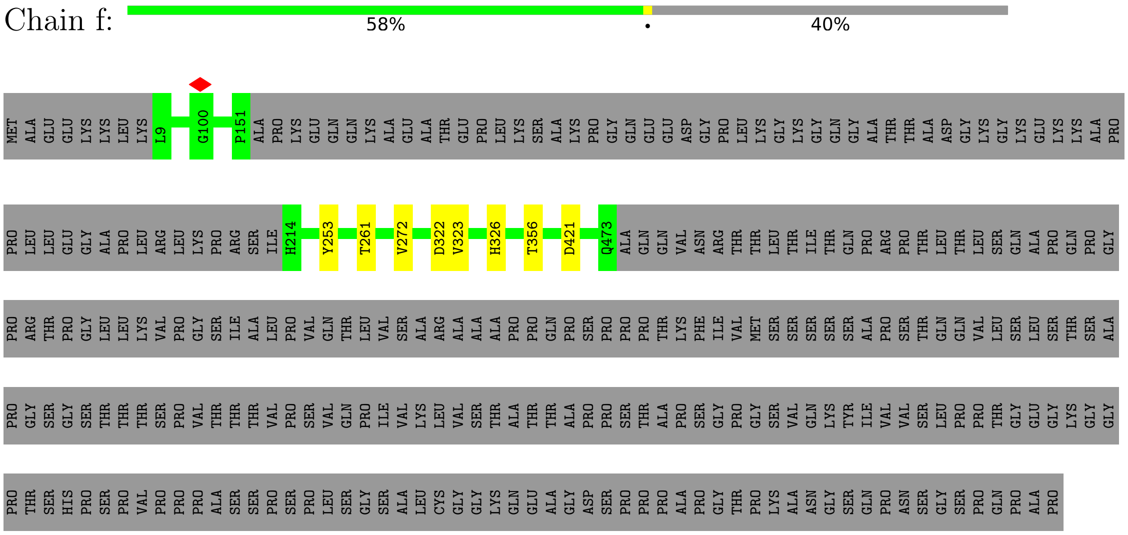
- Molecule 5: Transcription initiation factor TFIID subunit 6

Chain F:

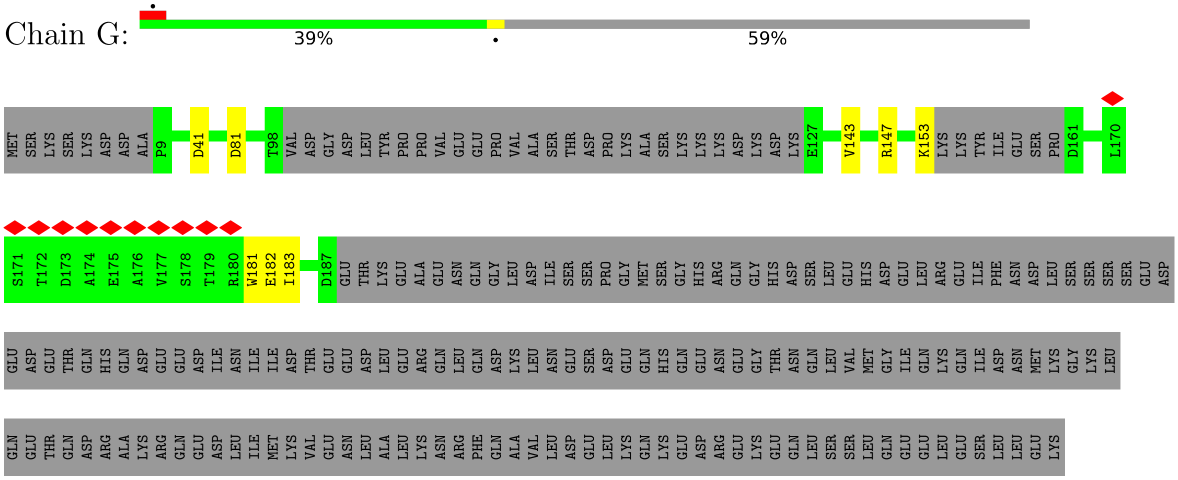




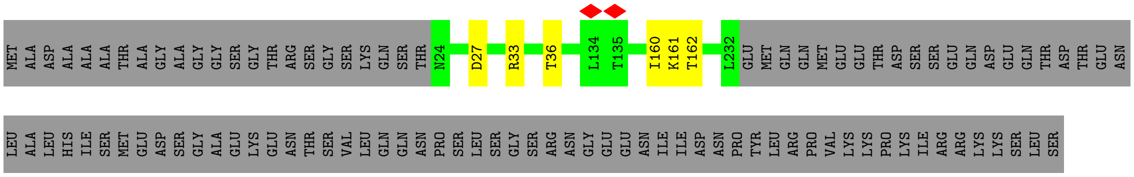
- Molecule 5: Transcription initiation factor TFIID subunit 6



- Molecule 6: Transcription initiation factor TFIID subunit 7

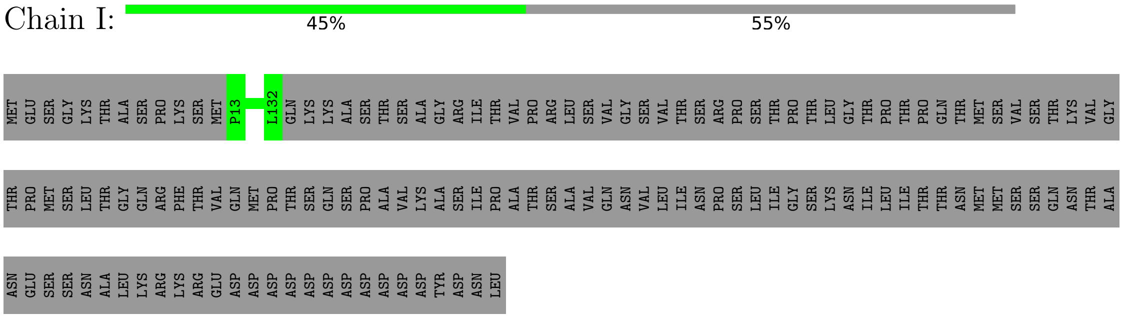


- Molecule 7: Transcription initiation factor TFIID subunit 8



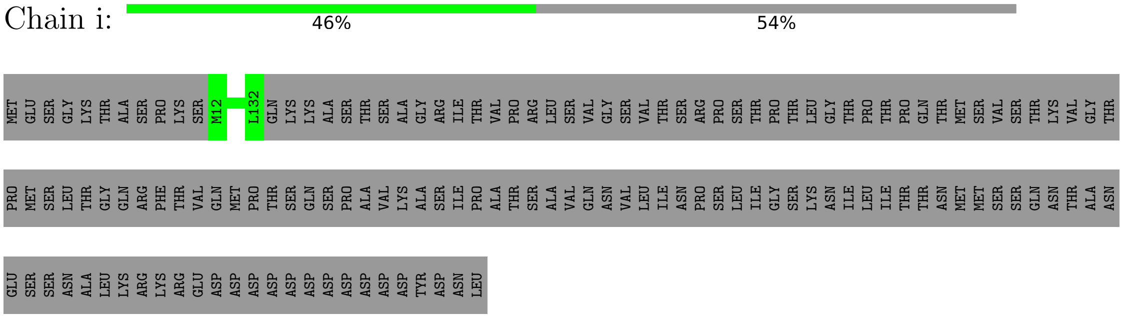
- Molecule 8: Transcription initiation factor TFIID subunit 9

Chain I:



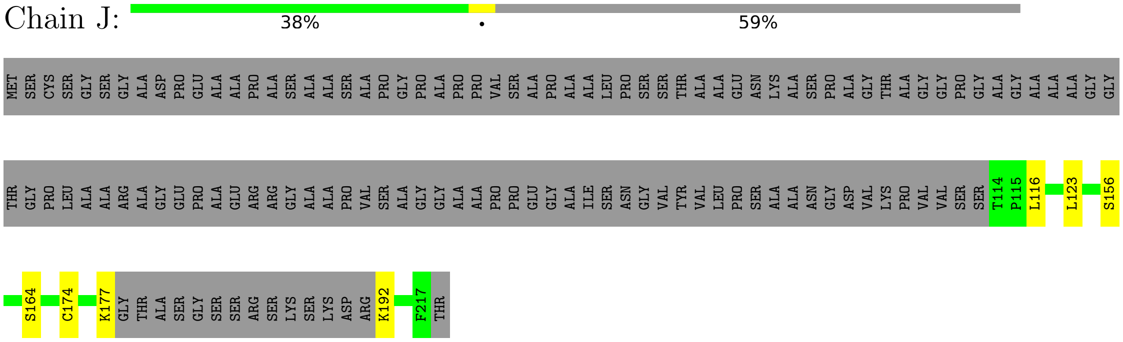
- Molecule 8: Transcription initiation factor TFIID subunit 9

Chain i:



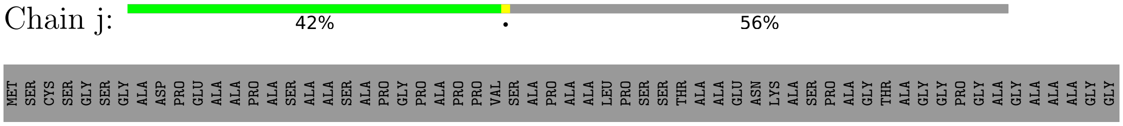
- Molecule 9: Transcription initiation factor TFIID subunit 10

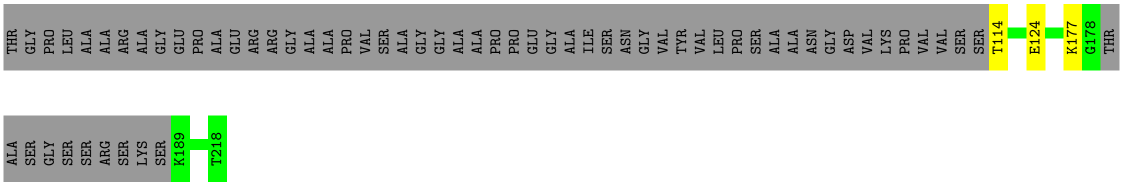
Chain J:



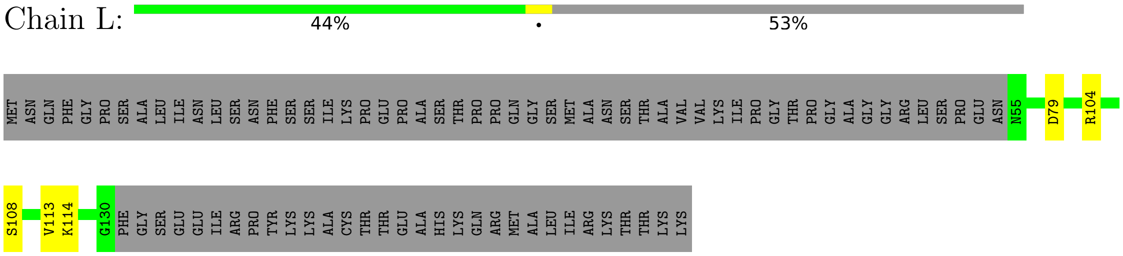
- Molecule 9: Transcription initiation factor TFIID subunit 10

Chain j:

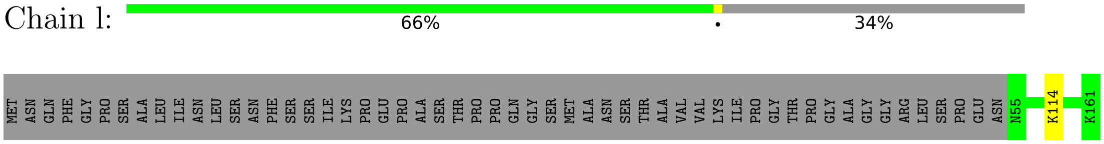




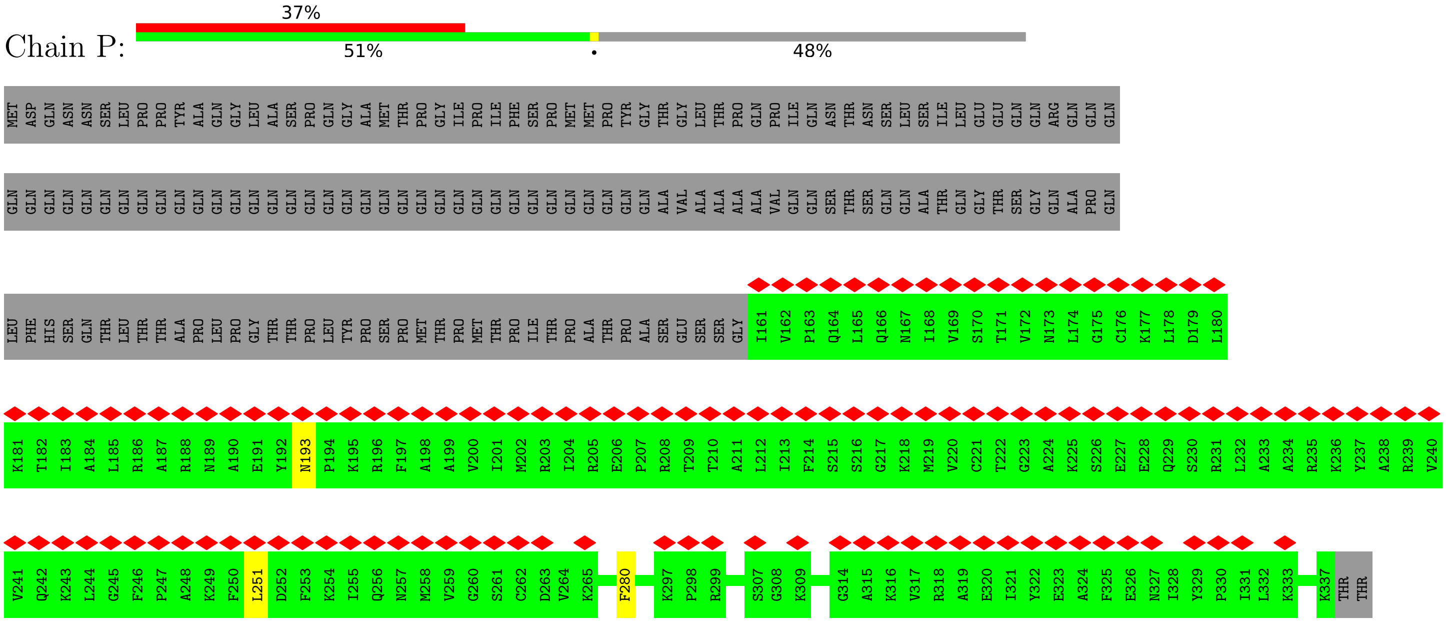
- Molecule 10: Transcription initiation factor TFIID subunit 12



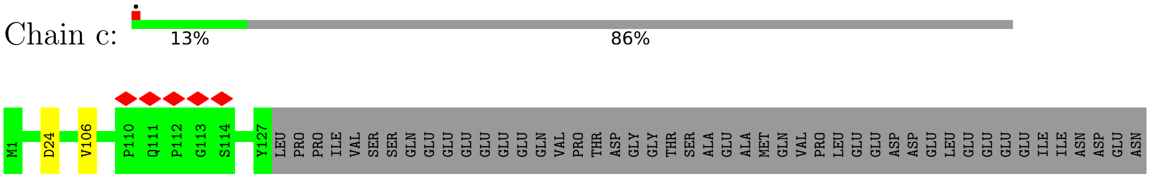
- Molecule 10: Transcription initiation factor TFIID subunit 12

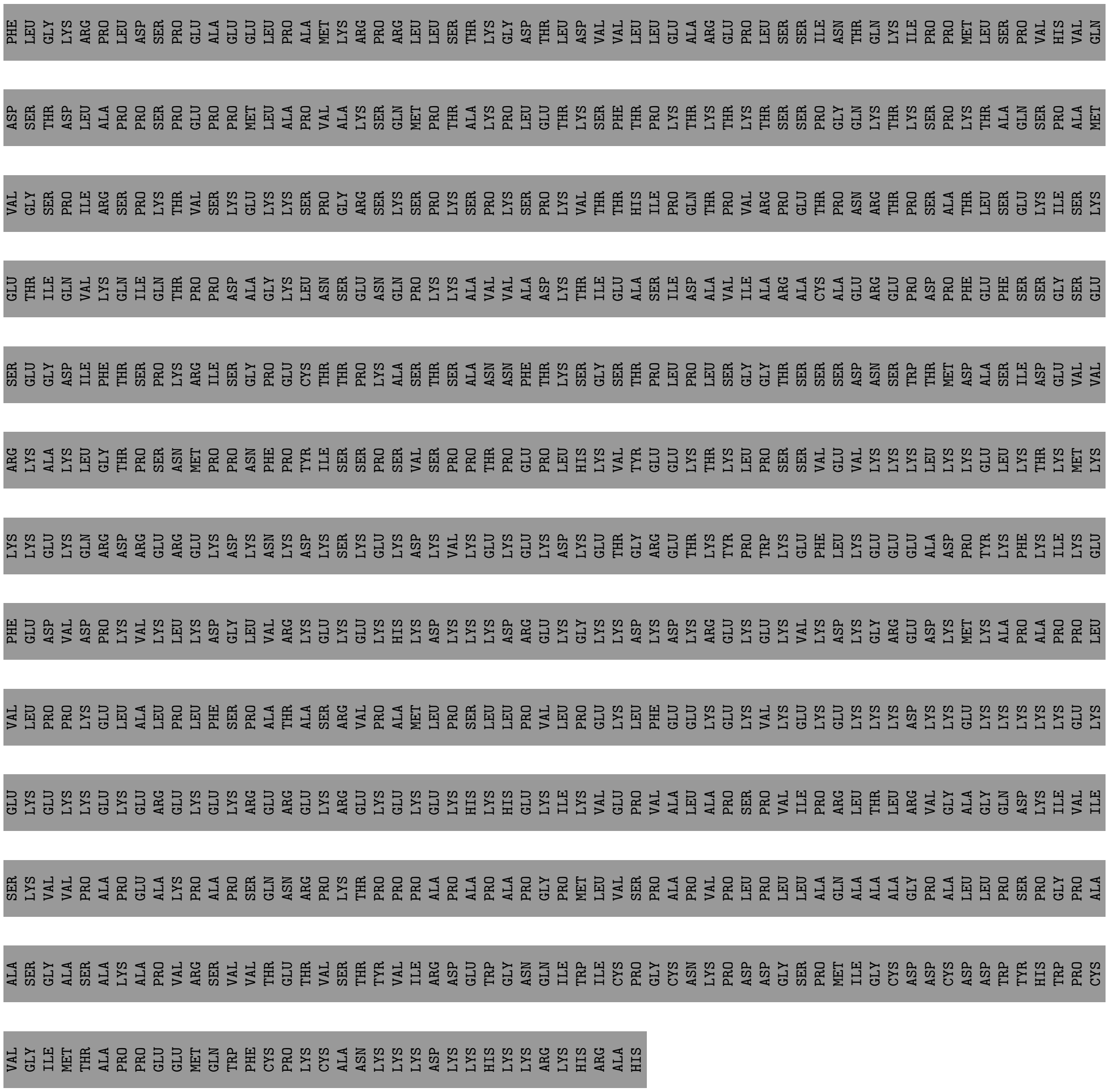


- Molecule 11: TATA-box-binding protein

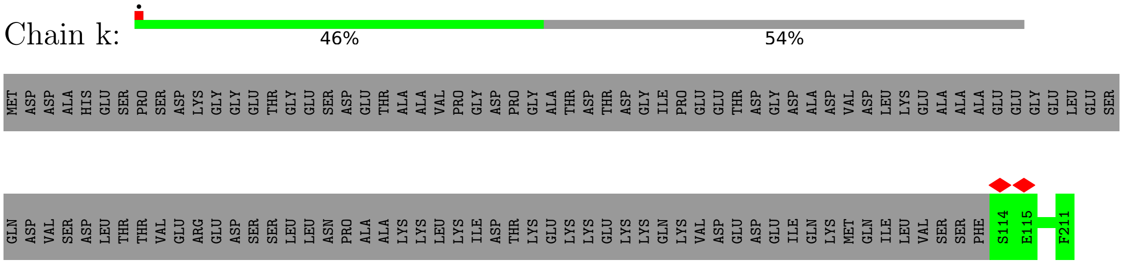


- Molecule 12: Transcription initiation factor TFIID subunit 3

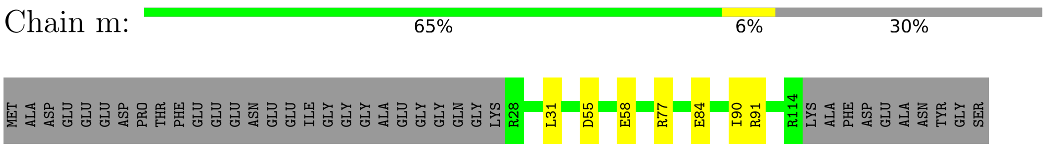




- Molecule 13: Transcription initiation factor TFIID subunit 11



- Molecule 14: Transcription initiation factor TFIID subunit 13



# 4 Experimental information ⓘ

| **Property** | **Value** | **Source** |
|---|---|---|
| EM reconstruction method | SINGLE PARTICLE | Depositor |
| Imposed symmetry | POINT, Not provided | |
| Number of particles used | 66920 | Depositor |
| Resolution determination method | FSC 0.143 CUT-OFF | Depositor |
| CTF correction method | PHASE FLIPPING AND AMPLITUDE CORRECTION | Depositor |
| Microscope | FEI TITAN KRIOS | Depositor |
| Voltage (kV) | 300 | Depositor |
| Electron dose $(e^-/\text{Å}^2)$ | 50 | Depositor |
| Minimum defocus (nm) | Not provided | |
| Maximum defocus (nm) | Not provided | |
| Magnification | Not provided | |
| Image detector | GATAN K2 SUMMIT (4k x 4k) | Depositor |
| Maximum map value | 0.043 | Depositor |
| Minimum map value | -0.014 | Depositor |
| Average map value | 0.000 | Depositor |
| Map value standard deviation | 0.002 | Depositor |
| Recommended contour level | 0.0062 | Depositor |
| Map size (Å) | 432.0, 432.0, 432.0 | wwPDB |
| Map dimensions | 320, 320, 320 | wwPDB |
| Map angles (°) | 90.0, 90.0, 90.0 | wwPDB |
| Pixel spacing (Å) | 1.35, 1.35, 1.35 | 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 | $\#\|Z\| > 5$ | RMSZ | $\#\|Z\| > 5$ |
| 1 | A | 0.50 | 0/4621 | 0.67 | 0/6240 |
| 2 | B | 0.46 | 0/7993 | 0.60 | 0/10836 |
| 3 | D | 0.37 | 0/1279 | 0.52 | 0/1714 |
| 3 | d | 0.39 | 0/1321 | 0.53 | 0/1772 |
| 4 | E | 0.39 | 0/4482 | 0.59 | 0/6069 |
| 4 | e | 0.43 | 0/4433 | 0.59 | 0/6004 |
| 5 | F | 0.48 | 0/3207 | 0.69 | 0/4354 |
| 5 | f | 0.40 | 0/3140 | 0.63 | 0/4268 |
| 6 | G | 0.51 | 0/1190 | 0.62 | 0/1601 |
| 7 | H | 0.49 | 0/1662 | 0.66 | 0/2272 |
| 8 | I | 0.28 | 0/981 | 0.48 | 0/1332 |
| 8 | i | 0.29 | 0/989 | 0.46 | 0/1343 |
| 9 | J | 0.57 | 0/736 | 0.69 | 0/998 |
| 9 | j | 0.53 | 0/775 | 0.63 | 0/1049 |
| 10 | L | 0.49 | 0/630 | 0.71 | 0/852 |
| 10 | l | 0.43 | 0/888 | 0.55 | 0/1194 |
| 11 | P | 0.46 | 0/1438 | 0.57 | 0/1935 |
| 12 | c | 0.39 | 0/1035 | 0.54 | 0/1406 |
| 13 | k | 0.30 | 0/799 | 0.47 | 0/1070 |
| 14 | m | 0.59 | 0/733 | 0.64 | 0/977 |
| All | All | 0.45 | 0/42332 | 0.61 | 0/57286 |

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts

Due to software issues we are unable to calculate clashes - this section is therefore empty.

## 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 | A | 533/1872 (28%) | 513 (96%) | 19 (4%) | 1 (0%) | 44 | 78 |
| 2 | B | 959/1199 (80%) | 911 (95%) | 48 (5%) | 0 | 100 | 100 |
| 3 | D | 148/1085 (14%) | 143 (97%) | 5 (3%) | 0 | 100 | 100 |
| 3 | d | 154/1085 (14%) | 149 (97%) | 5 (3%) | 0 | 100 | 100 |
| 4 | E | 541/800 (68%) | 516 (95%) | 24 (4%) | 1 (0%) | 44 | 78 |
| 4 | e | 531/800 (66%) | 482 (91%) | 48 (9%) | 1 (0%) | 44 | 78 |
| 5 | F | 408/677 (60%) | 391 (96%) | 16 (4%) | 1 (0%) | 44 | 78 |
| 5 | f | 399/677 (59%) | 380 (95%) | 19 (5%) | 0 | 100 | 100 |
| 6 | G | 138/349 (40%) | 135 (98%) | 3 (2%) | 0 | 100 | 100 |
| 7 | H | 207/310 (67%) | 194 (94%) | 13 (6%) | 0 | 100 | 100 |
| 8 | I | 118/264 (45%) | 115 (98%) | 3 (2%) | 0 | 100 | 100 |
| 8 | i | 119/264 (45%) | 115 (97%) | 4 (3%) | 0 | 100 | 100 |
| 9 | J | 86/218 (39%) | 82 (95%) | 4 (5%) | 0 | 100 | 100 |
| 9 | j | 91/218 (42%) | 85 (93%) | 4 (4%) | 2 (2%) | 5 | 29 |
| 10 | L | 74/161 (46%) | 72 (97%) | 2 (3%) | 0 | 100 | 100 |
| 10 | l | 105/161 (65%) | 100 (95%) | 5 (5%) | 0 | 100 | 100 |
| 11 | P | 175/339 (52%) | 165 (94%) | 9 (5%) | 1 (1%) | 22 | 60 |
| 12 | c | 125/929 (14%) | 116 (93%) | 9 (7%) | 0 | 100 | 100 |
| 13 | k | 96/211 (46%) | 91 (95%) | 5 (5%) | 0 | 100 | 100 |
| 14 | m | 85/124 (68%) | 82 (96%) | 3 (4%) | 0 | 100 | 100 |
| All | All | 5092/11743 (43%) | 4837 (95%) | 248 (5%) | 7 (0%) | 50 | 83 |

5 of 7 Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 4 | E | 523 | VAL |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type |
|---|---|---|---|
| 4 | e | 522 | ASP |
| 9 | j | 124 | GLU |
| 5 | F | 411 | VAL |
| 9 | j | 177 | LYS |

### 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 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 | A | 487/1665 (29%) | 449 (92%) | 38 (8%) | 10 | 29 |
| 2 | B | 876/1083 (81%) | 860 (98%) | 16 (2%) | 54 | 71 |
| 3 | D | 139/815 (17%) | 136 (98%) | 3 (2%) | 47 | 65 |
| 3 | d | 146/815 (18%) | 145 (99%) | 1 (1%) | 81 | 87 |
| 4 | E | 480/657 (73%) | 466 (97%) | 14 (3%) | 37 | 56 |
| 4 | e | 475/657 (72%) | 463 (98%) | 12 (2%) | 42 | 61 |
| 5 | F | 329/574 (57%) | 311 (94%) | 18 (6%) | 18 | 39 |
| 5 | f | 322/574 (56%) | 314 (98%) | 8 (2%) | 42 | 61 |
| 6 | G | 132/322 (41%) | 124 (94%) | 8 (6%) | 15 | 37 |
| 7 | H | 178/270 (66%) | 172 (97%) | 6 (3%) | 32 | 51 |
| 8 | I | 106/235 (45%) | 106 (100%) | 0 | 100 | 100 |
| 8 | i | 107/235 (46%) | 107 (100%) | 0 | 100 | 100 |
| 9 | J | 79/154 (51%) | 72 (91%) | 7 (9%) | 8 | 25 |
| 9 | j | 83/154 (54%) | 82 (99%) | 1 (1%) | 67 | 78 |
| 10 | L | 71/141 (50%) | 66 (93%) | 5 (7%) | 12 | 32 |
| 10 | l | 98/141 (70%) | 97 (99%) | 1 (1%) | 73 | 82 |
| 11 | P | 153/293 (52%) | 151 (99%) | 2 (1%) | 65 | 77 |
| 12 | c | 113/833 (14%) | 111 (98%) | 2 (2%) | 54 | 71 |
| 13 | k | 87/182 (48%) | 87 (100%) | 0 | 100 | 100 |
| 14 | m | 80/106 (76%) | 73 (91%) | 7 (9%) | 8 | 25 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Analysed | Rotameric | Outliers | Percentiles | |
|---|---|---|---|---|---|---|
| All | All | 4541/9906 (46%) | 4392 (97%) | 149 (3%) | 35 | 52 |

5 of 149 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 11 | P | 193 | ASN |
| 14 | m | 55 | ASP |
| 3 | d | 951 | ASP |
| 4 | e | 516 | ILE |
| 2 | B | 559 | LYS |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 71 such sidechains are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 3 | d | 1069 | ASN |
| 4 | e | 294 | ASN |
| 9 | j | 173 | HIS |
| 2 | B | 916 | ASN |
| 2 | B | 908 | 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 oligosaccharides 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.

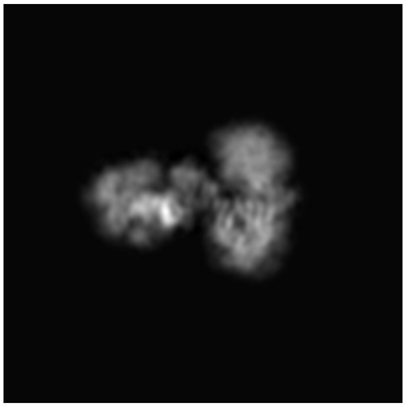
# 6 Map visualisation ⓘ

This section contains visualisations of the EMDB entry EMD-31114. These allow visual inspection of the internal detail of the map and identification of artifacts.

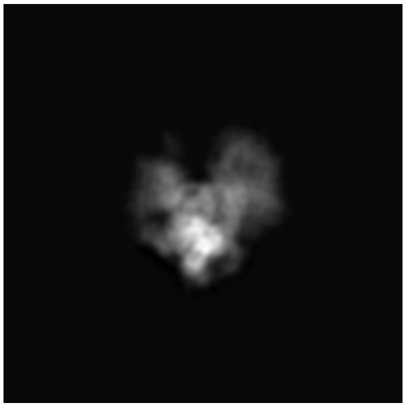
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

## 6.1 Orthogonal projections ⓘ

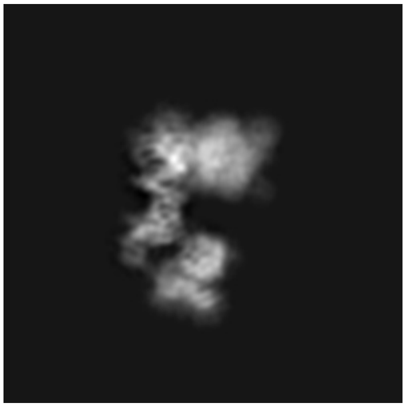
### 6.1.1 Primary map



X



Y

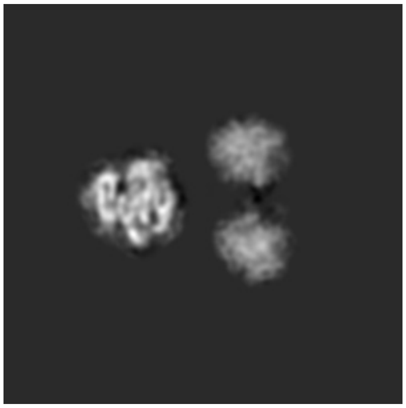


Z

The images above show the map projected in three orthogonal directions.

## 6.2 Central slices ⓘ

### 6.2.1 Primary map



X Index: 160

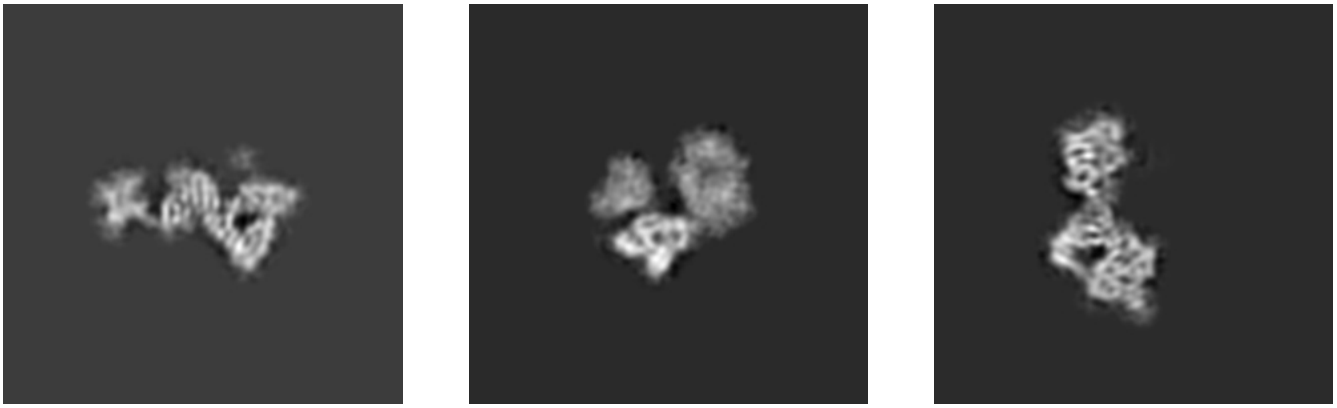
Y Index: 160

Z Index: 160

The images above show central slices of the map in three orthogonal directions.

## 6.3 Largest variance slices

### 6.3.1 Primary map

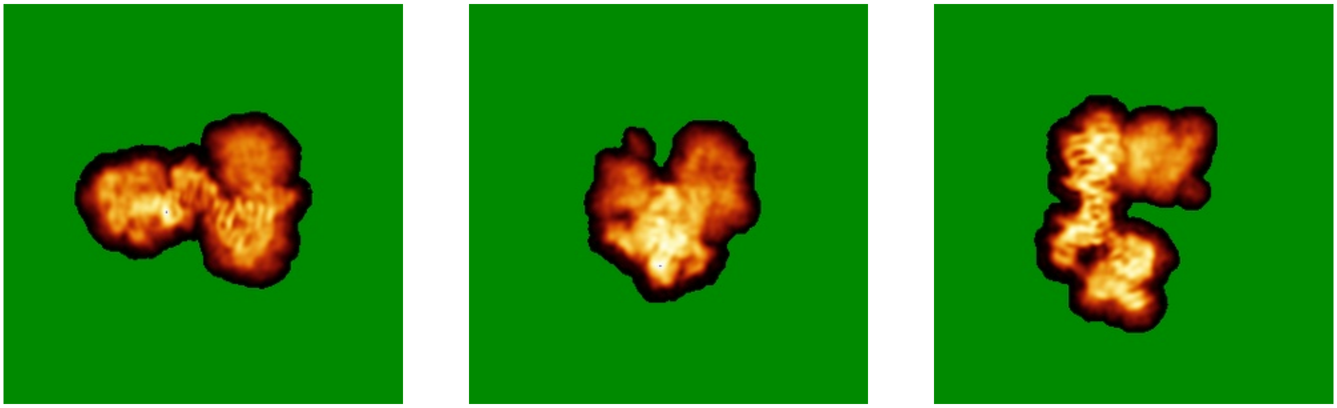


X Index: 131 Y Index: 205 Z Index: 155

The images above show the largest variance slices of the map in three orthogonal directions.

## 6.4 Orthogonal standard-deviation projections (False-color)

### 6.4.1 Primary map

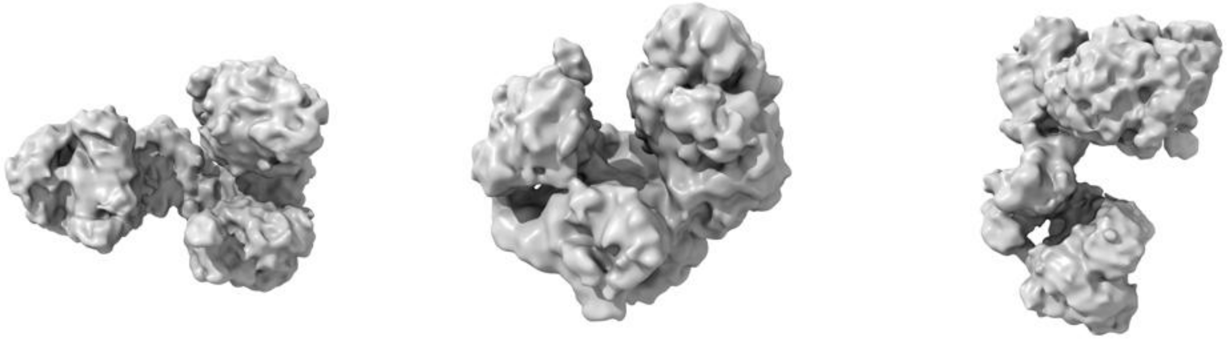


X Y Z

The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

## 6.5 Orthogonal surface views (i)

### 6.5.1 Primary map



X Y Z

The images above show the 3D surface view of the map at the recommended contour level 0.0062. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

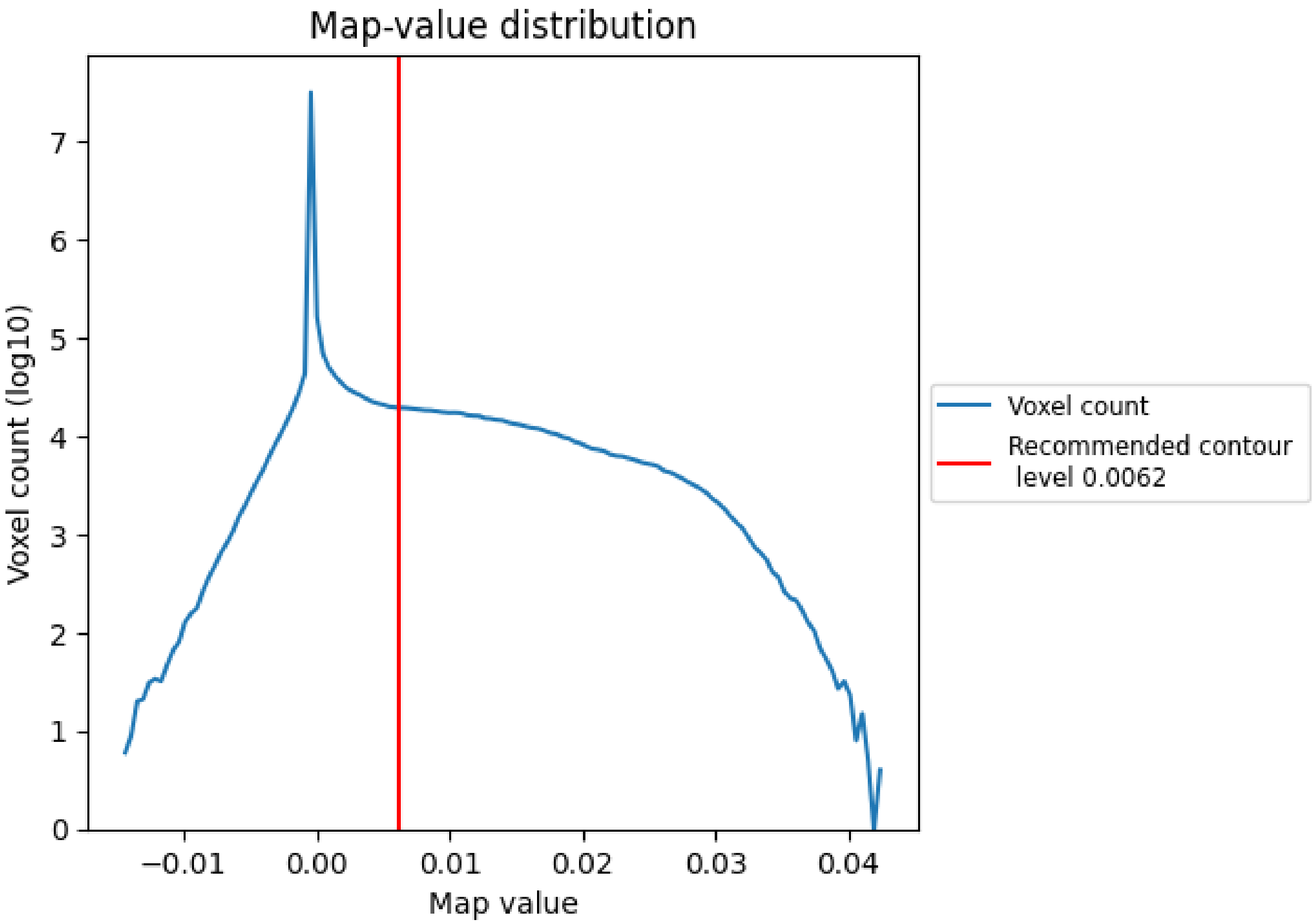
## 6.6 Mask visualisation (i)

This section was not generated. No masks/segmentation were deposited.

# 7 Map analysis ⓘ

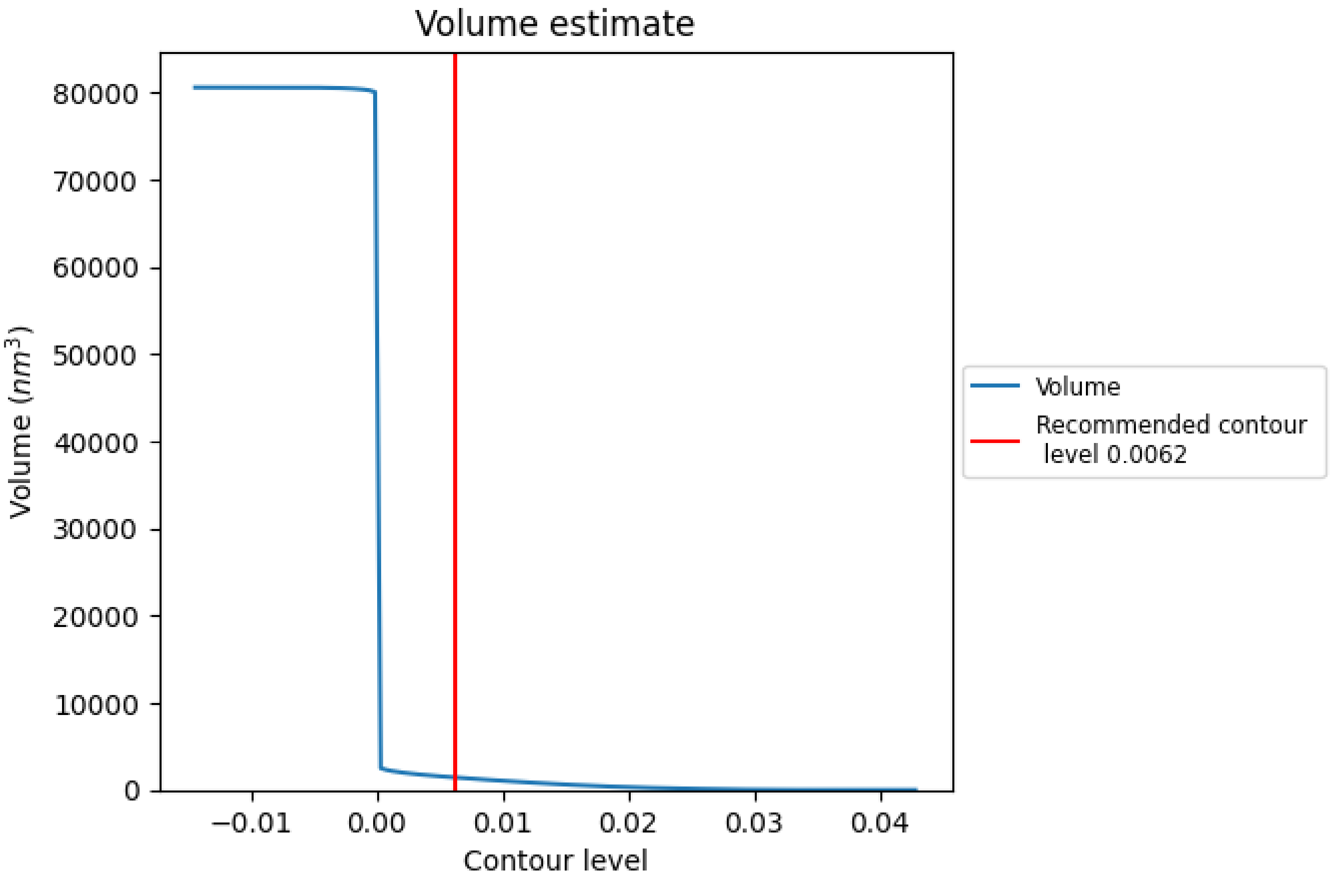
This section contains the results of statistical analysis of the map.

## 7.1 Map-value distribution ⓘ



The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

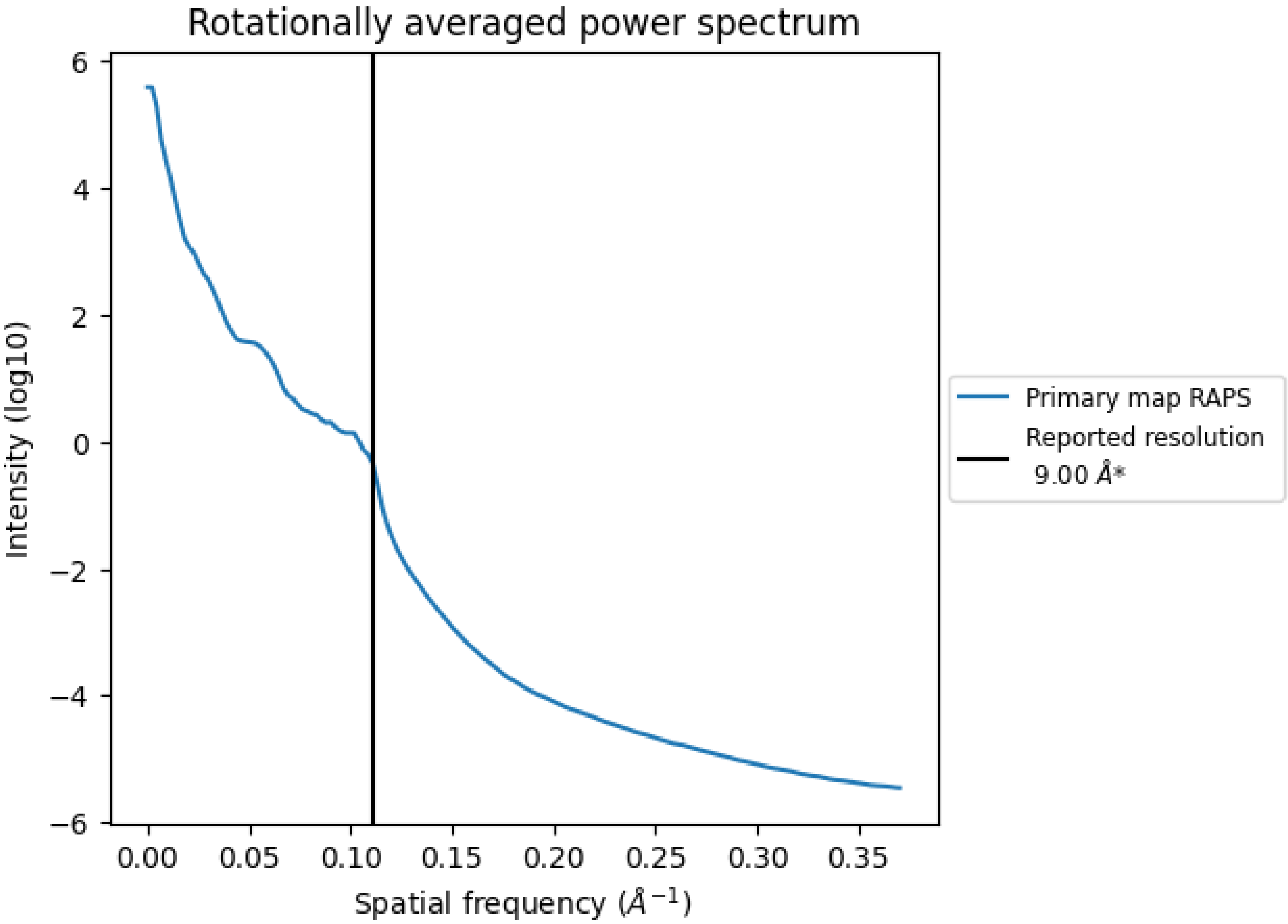
## 7.2 Volume estimate ⓘ



The volume at the recommended contour level is 1436 $nm^3$; this corresponds to an approximate mass of 1297 kDa.

The volume estimate graph shows how the enclosed volume varies with the contour level. The recommended contour level is shown as a vertical line and the intersection between the line and the curve gives the volume of the enclosed surface at the given level.

### 7.3 Rotationally averaged power spectrum ⓘ



*Reported resolution corresponds to spatial frequency of 0.111 Å$^{-1}$

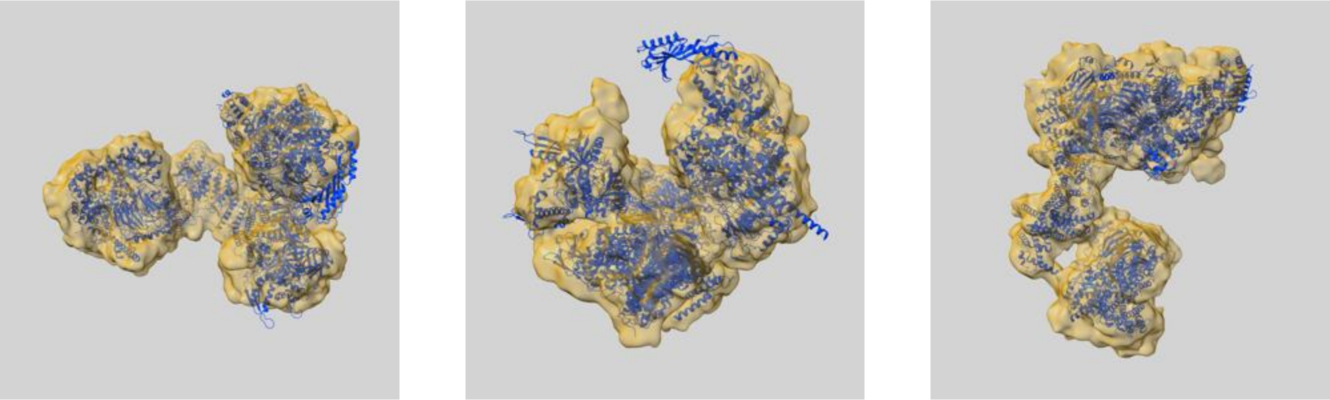
# 8 Fourier-Shell correlation ⓘ

This section was not generated. No FSC curve or half-maps provided.

# 9 Map-model fit ⓘ

This section contains information regarding the fit between EMDB map EMD-31114 and PDB model 7EGE. Per-residue inclusion information can be found in section 3 on page 7.

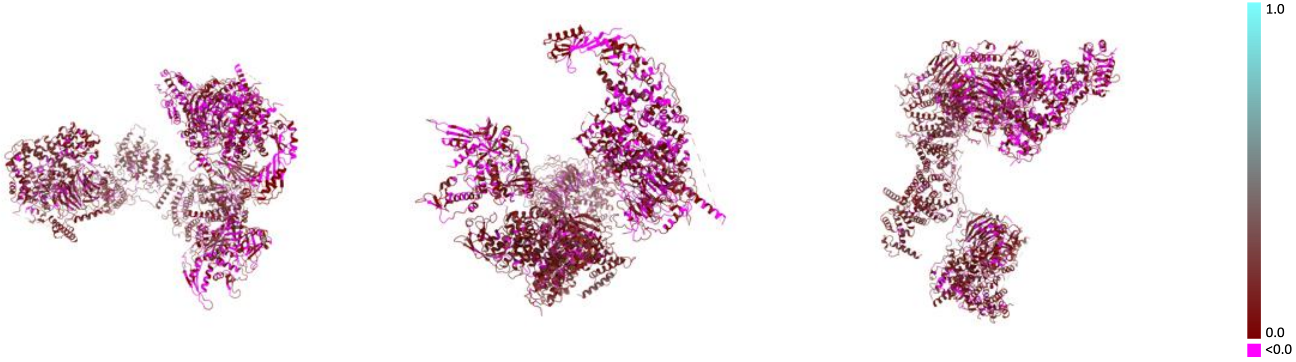
## 9.1 Map-model overlay ⓘ



X Y Z

The images above show the 3D surface view of the map at the recommended contour level 0.0062 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

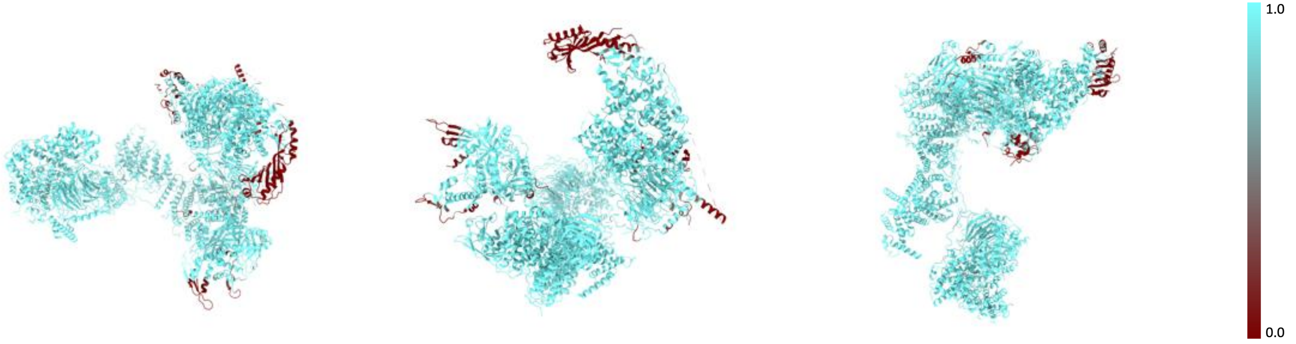
## 9.2 Q-score mapped to coordinate model ⓘ



X Y Z

The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions at which atoms can be resolved. Low Q-score values may therefore be expected for many entries.

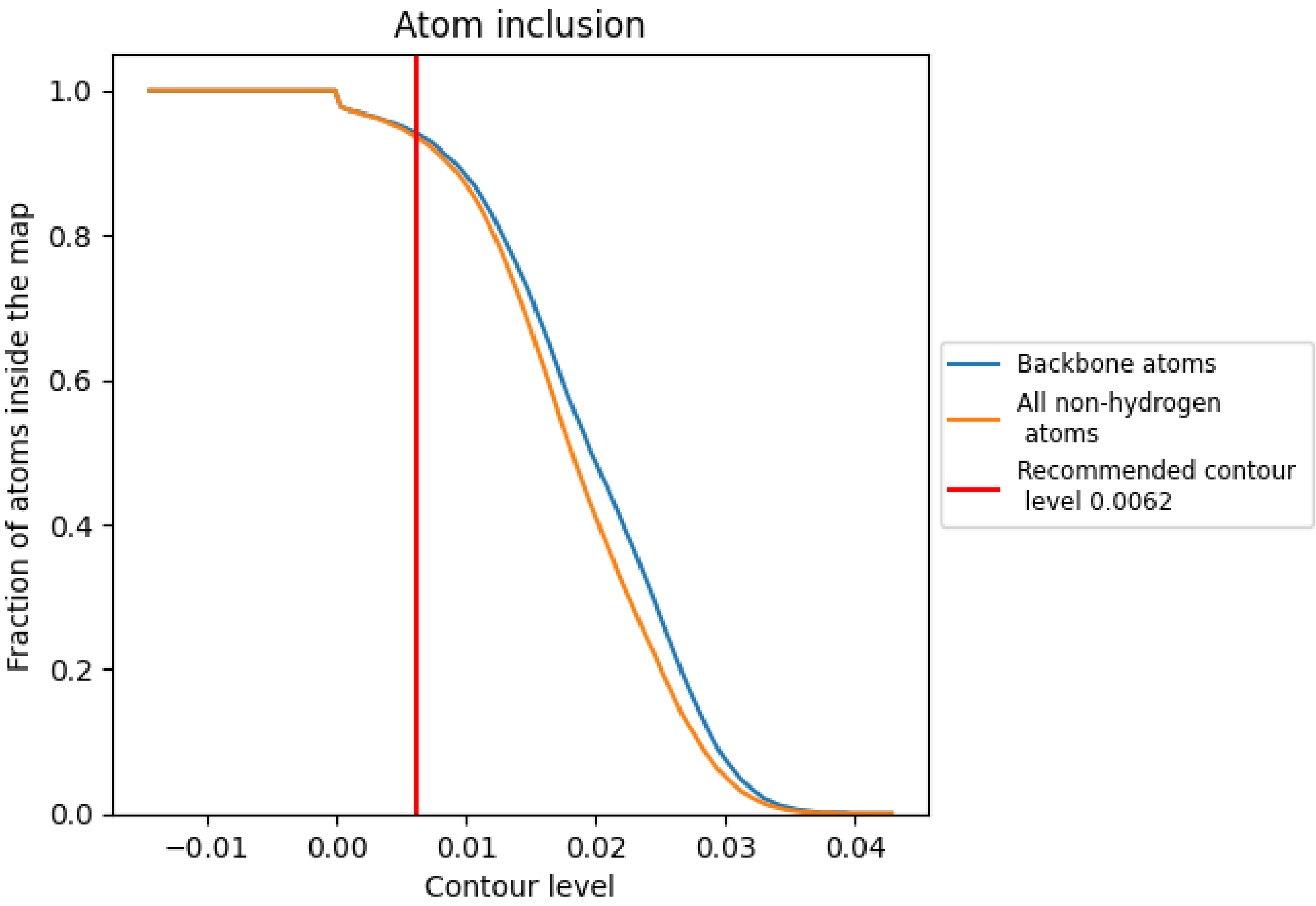
## 9.3 Atom inclusion mapped to coordinate model ⓘ



X Y Z

The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.0062).

## 9.4 Atom inclusion ⓘ



At the recommended contour level, 94% of all backbone atoms, 94% of all non-hydrogen atoms, are inside the map.

## 9.5 Map-model fit summary

The table lists the average atom inclusion at the recommended contour level (0.0062) and Q-score for the entire model and for each chain.

| Chain | Atom inclusion | Q-score |
|---|---|---|
| All | 0.9350 | 0.0780 |
| A | 0.8710 | 0.0530 |
| B | 0.9940 | 0.1120 |
| D | 0.9880 | 0.1000 |
| E | 0.9970 | 0.0960 |
| F | 0.9910 | 0.1180 |
| G | 0.9310 | 0.0470 |
| H | 0.9890 | 0.1270 |
| I | 1.0000 | 0.1280 |
| J | 1.0000 | 0.0770 |
| L | 1.0000 | 0.1190 |
| P | 0.2660 | -0.0120 |
| c | 0.9630 | 0.0470 |
| d | 0.8670 | 0.0310 |
| e | 0.8610 | 0.0280 |
| f | 0.9880 | 0.1120 |
| i | 0.9990 | 0.0490 |
| j | 1.0000 | 0.0410 |
| k | 0.9790 | 0.0440 |
| l | 0.9880 | 0.0310 |
| m | 0.9960 | 0.0160 |