

## wwPDB EM Map/Model Validation Report ⓘ

Apr 10, 2016 – 02:02 PM BST

PDB ID : 4CKD  
EMDB ID: : EMD-2548  
Title : Model of complex between the E.coli enzyme beta-galactosidase and four single chain Fv antibody domains scFv13R4.  
Authors : Vinothkumar, K.R.; McMullan, G.; Henderson, R.  
Deposited on : 2014-01-03  
Resolution : 13.00 Å(reported)  
Based on PDB ID : 1F4A

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](mailto: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

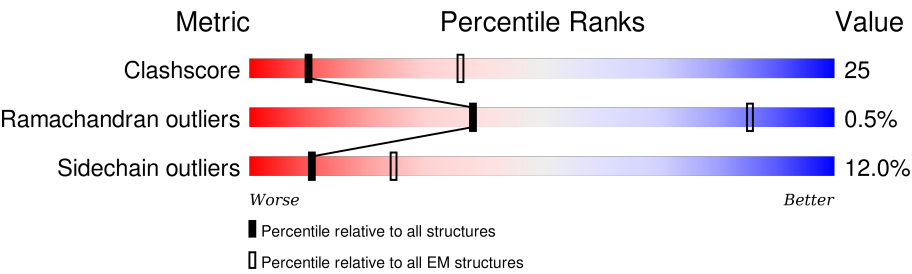


# 1 Overall quality at a glance i

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

The reported resolution of this entry is 13.00 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<br>(#Entries) | EM structures<br>(#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  $\geq 3$ , 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   | A     | 1024   | <div><div>48%</div><div>38%</div><div>11%</div><div>.</div></div> |
| 1   | B     | 1024   | <div><div>49%</div><div>38%</div><div>11%</div><div>.</div></div> |
| 1   | C     | 1024   | <div><div>48%</div><div>39%</div><div>10%</div><div>.</div></div> |
| 1   | D     | 1024   | <div><div>48%</div><div>39%</div><div>11%</div><div>.</div></div> |
| 2   | H     | 114    | <div><div>75%</div><div>25%</div><div>.</div></div>               |
| 2   | I     | 114    | <div><div>73%</div><div>26%</div><div>.</div></div>               |
| 2   | J     | 114    | <div><div>75%</div><div>25%</div><div>.</div></div>               |
| 2   | K     | 114    | <div><div>76%</div><div>23%</div><div>.</div></div>               |
| 3   | L     | 107    | <div><div>83%</div><div>16%</div><div>.</div></div>               |

Continued on next page...



*Continued from previous page...*

| Mol | Chain | Length | Quality of chain   |
|-----|-------|--------|--|
| 3   | M     | 107    |  83%16%• |
| 3   | N     | 107    |  81%18%• |
| 3   | O     | 107    |  81%18%• |



## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 39832 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 BETA-GALACTOSIDASE.

| Mol | Chain | Residues | Atoms |      |      |      |    | AltConf | Trace |
|-----|-------|----------|-------|------|------|------|----|---------|-------|
| 1   | A     | 1021     | Total | C    | N    | O    | S  | 7       | 0     |
|     |       |          | 8238  | 5209 | 1466 | 1525 | 38 |         |       |
| 1   | B     | 1021     | Total | C    | N    | O    | S  | 7       | 0     |
|     |       |          | 8238  | 5209 | 1466 | 1525 | 38 |         |       |
| 1   | C     | 1021     | Total | C    | N    | O    | S  | 7       | 0     |
|     |       |          | 8238  | 5209 | 1466 | 1525 | 38 |         |       |
| 1   | D     | 1021     | Total | C    | N    | O    | S  | 7       | 0     |
|     |       |          | 8238  | 5209 | 1466 | 1525 | 38 |         |       |

- Molecule 2 is a protein called SCFV13R4 ANTIBODY FV HEAVY CHAIN.

| Mol | Chain | Residues | Atoms |     |     |     |   | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---|---------|-------|
| 2   | H     | 114      | Total | C   | N   | O   | S | 0       | 0     |
|     |       |          | 902   | 567 | 144 | 188 | 3 |         |       |
| 2   | I     | 114      | Total | C   | N   | O   | S | 0       | 0     |
|     |       |          | 902   | 567 | 144 | 188 | 3 |         |       |
| 2   | J     | 114      | Total | C   | N   | O   | S | 0       | 0     |
|     |       |          | 902   | 567 | 144 | 188 | 3 |         |       |
| 2   | K     | 114      | Total | C   | N   | O   | S | 0       | 0     |
|     |       |          | 902   | 567 | 144 | 188 | 3 |         |       |

- Molecule 3 is a protein called SCFV13R4 ANTIBODY FV LIGHT CHAIN.

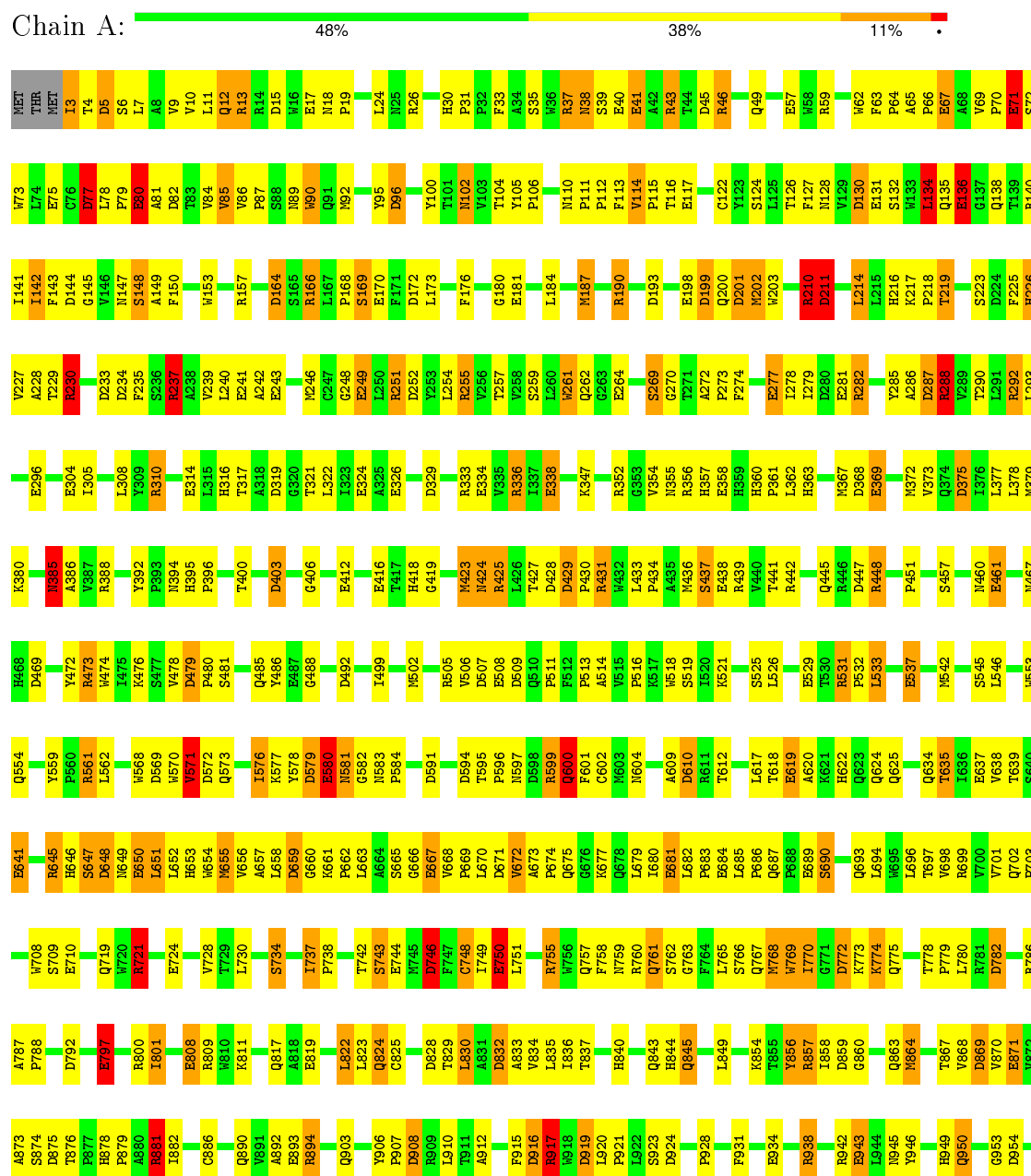
| Mol | Chain | Residues | Atoms |     |     |     |   | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---|---------|-------|
| 3   | L     | 107      | Total | C   | N   | O   | S | 0       | 0     |
|     |       |          | 818   | 511 | 137 | 167 | 3 |         |       |
| 3   | M     | 107      | Total | C   | N   | O   | S | 0       | 0     |
|     |       |          | 818   | 511 | 137 | 167 | 3 |         |       |
| 3   | N     | 107      | Total | C   | N   | O   | S | 0       | 0     |
|     |       |          | 818   | 511 | 137 | 167 | 3 |         |       |
| 3   | O     | 107      | Total | C   | N   | O   | S | 0       | 0     |
|     |       |          | 818   | 511 | 137 | 167 | 3 |         |       |



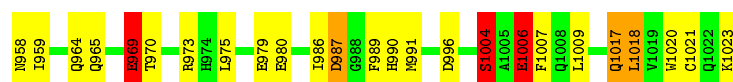
### 3 Residue-property plots

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: BETA-GALACTOSIDASE

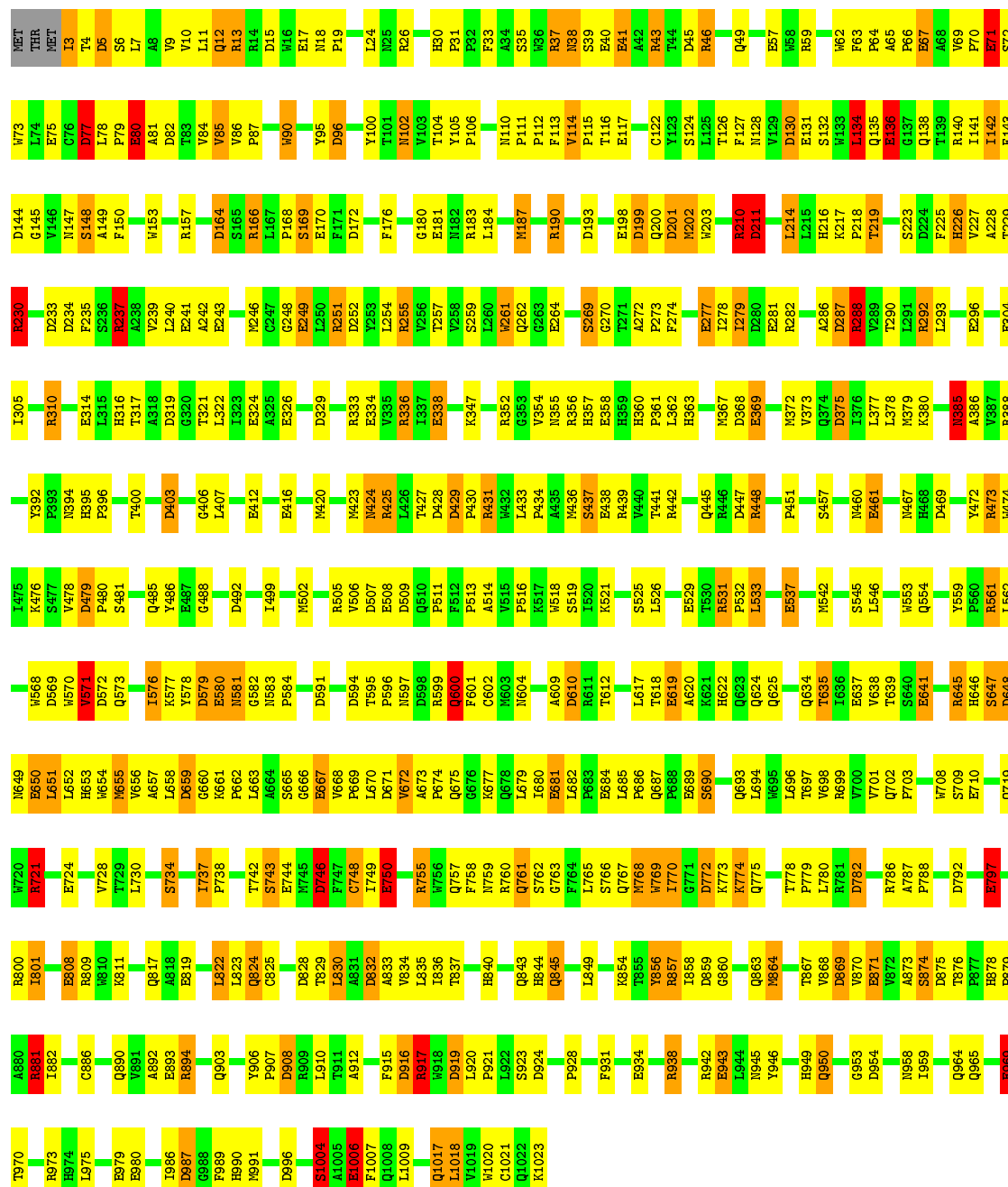






• Molecule 1: BETA-GALACTOSIDASE

Chain B: 49% 38% 11%



• Molecule 1: BETA-GALACTOSIDASE

Chain C: 48% 39% 10%

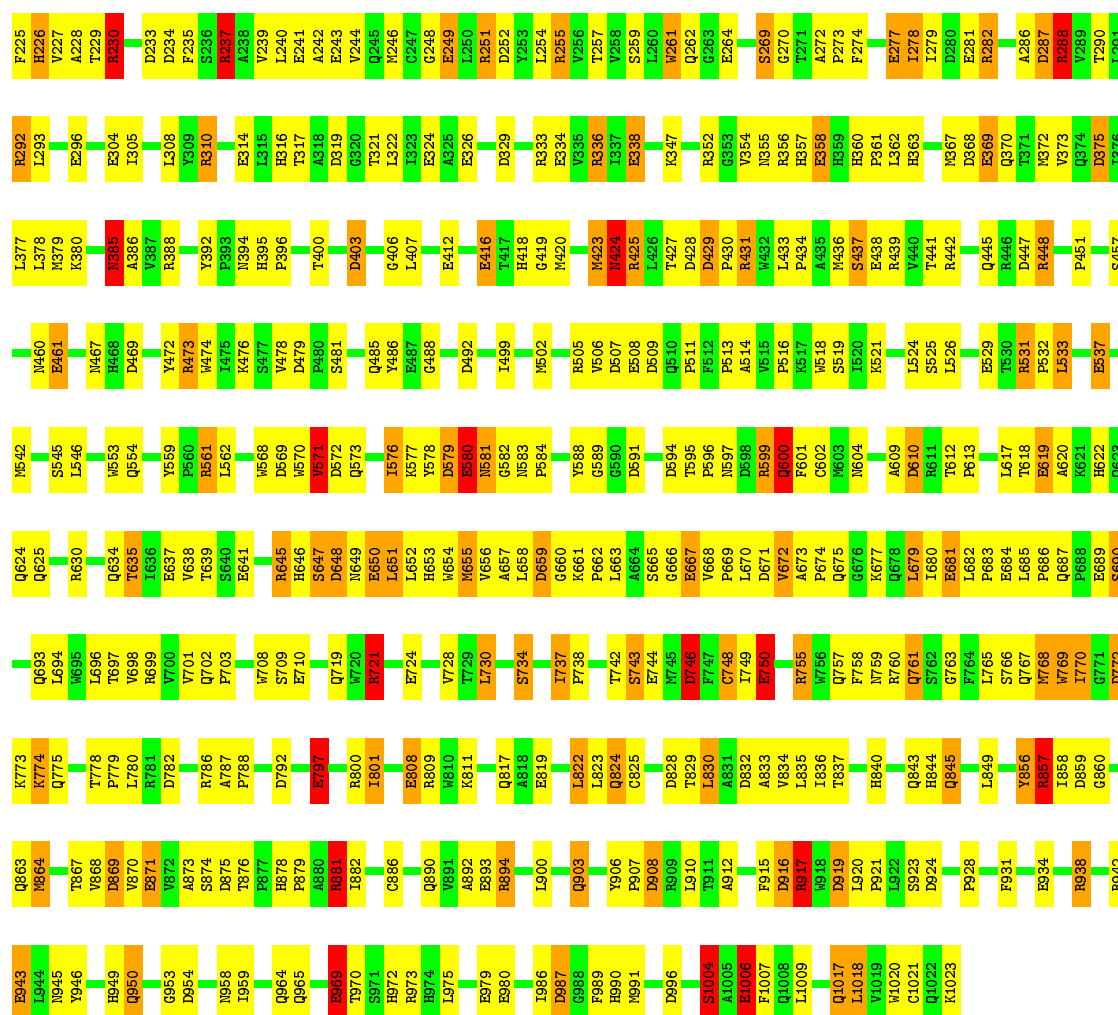




| Response   | Percentage |
|------------|------------|
| Yes        | 48%        |
| No         | 39%        |
| Don't know | 11%        |

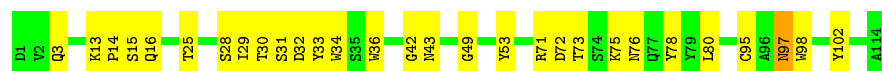






• Molecule 2: SCFV13R4 ANTIBODY FV HEAVY CHAIN

Chain H: 75% 25%



• Molecule 2: SCFV13R4 ANTIBODY FV HEAVY CHAIN

Chain I: 73% 26%




• Molecule 2: SCFV13R4 ANTIBODY FV HEAVY CHAIN

Chain J: 75% 25%




• Molecule 2: SCFV13R4 ANTIBODY FV HEAVY CHAIN



Chain K:  76% 23%




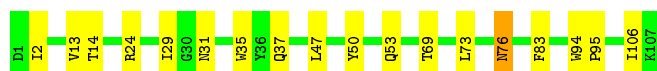
• Molecule 3: SCFV13R4 ANTIBODY FV LIGHT CHAIN

Chain L:  83% 16%




• Molecule 3: SCFV13R4 ANTIBODY FV LIGHT CHAIN

Chain M:  83% 16%




• Molecule 3: SCFV13R4 ANTIBODY FV LIGHT CHAIN

Chain N:  81% 18%



• Molecule 3: SCFV13R4 ANTIBODY FV LIGHT CHAIN

Chain O:  81% 18%





## 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                | DONE INSIDE FREALIGN    | Depositor |
| Microscope                           | OTHER                   | Depositor |
| Voltage (kV)                         | 300                     | Depositor |
| Electron dose ( $e^-/\text{\AA}^2$ ) | 67                      | Depositor |
| Minimum defocus (nm)                 | 2678                    | Depositor |
| Maximum defocus (nm)                 | 4027                    | Depositor |
| Magnification                        | 59000                   | Depositor |
| Image detector                       | FEI FALCON II (4K X 4K) | 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   | A     | 1.08         | 52/8515 (0.6%)   | 1.61        | 174/11615 (1.5%) |
| 1   | B     | 1.08         | 52/8515 (0.6%)   | 1.61        | 175/11615 (1.5%) |
| 1   | C     | 1.08         | 52/8515 (0.6%)   | 1.61        | 173/11615 (1.5%) |
| 1   | D     | 1.08         | 53/8515 (0.6%)   | 1.61        | 180/11615 (1.5%) |
| 2   | H     | 0.31         | 0/925            | 0.65        | 0/1263           |
| 2   | I     | 0.31         | 0/925            | 0.65        | 0/1263           |
| 2   | J     | 0.31         | 0/925            | 0.65        | 0/1263           |
| 2   | K     | 0.31         | 0/925            | 0.65        | 0/1263           |
| 3   | L     | 0.32         | 0/837            | 0.58        | 0/1133           |
| 3   | M     | 0.32         | 0/837            | 0.58        | 0/1133           |
| 3   | N     | 0.32         | 0/837            | 0.58        | 0/1133           |
| 3   | O     | 0.32         | 0/837            | 0.59        | 0/1133           |
| All | All   | 1.00         | 209/41108 (0.5%) | 1.49        | 702/56044 (1.3%) |

The worst 5 of 209 bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms  | Z    | Observed(Å) | Ideal(Å) |
|-----|-------|-----|------|--------|------|-------------|----------|
| 1   | B     | 75  | GLU  | CD-OE2 | 9.50 | 1.36        | 1.25     |
| 1   | A     | 75  | GLU  | CD-OE2 | 9.50 | 1.36        | 1.25     |
| 1   | D     | 75  | GLU  | CD-OE2 | 9.48 | 1.36        | 1.25     |
| 1   | C     | 75  | GLU  | CD-OE2 | 9.41 | 1.36        | 1.25     |
| 1   | B     | 710 | GLU  | CD-OE1 | 7.65 | 1.34        | 1.25     |

The worst 5 of 702 bond angle outliers are listed below:

| Mol | Chain | Res    | Type | Atoms     | Z     | Observed(°) | Ideal(°) |
|-----|-------|--------|------|-----------|-------|-------------|----------|
| 1   | B     | 809[A] | ARG  | NE-CZ-NH1 | 11.67 | 126.14      | 120.30   |
| 1   | B     | 809[B] | ARG  | NE-CZ-NH1 | 11.67 | 126.14      | 120.30   |
| 1   | D     | 809[A] | ARG  | NE-CZ-NH1 | 11.59 | 126.09      | 120.30   |
| 1   | D     | 809[B] | ARG  | NE-CZ-NH1 | 11.59 | 126.09      | 120.30   |
| 1   | A     | 809[A] | ARG  | NE-CZ-NH1 | 11.58 | 126.09      | 120.30   |

There are no chirality outliers.



There are no planarity outliers.

## 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|
| 1   | A     | 8238  | 0        | 7821     | 477     | 0            |
| 1   | B     | 8238  | 0        | 7822     | 458     | 0            |
| 1   | C     | 8238  | 0        | 7821     | 460     | 0            |
| 1   | D     | 8238  | 0        | 7821     | 471     | 0            |
| 2   | H     | 902   | 0        | 838      | 111     | 0            |
| 2   | I     | 902   | 0        | 838      | 110     | 0            |
| 2   | J     | 902   | 0        | 839      | 108     | 0            |
| 2   | K     | 902   | 0        | 839      | 109     | 0            |
| 3   | L     | 818   | 0        | 780      | 12      | 0            |
| 3   | M     | 818   | 0        | 780      | 12      | 0            |
| 3   | N     | 818   | 0        | 780      | 13      | 0            |
| 3   | O     | 818   | 0        | 780      | 13      | 0            |
| All | All   | 39832 | 0        | 37759    | 1937    | 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 25.

The worst 5 of 1937 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) |
|----------------|-----------------|--------------------------|-------------------|
| 1:C:584:PRO:CB | 2:H:30:THR:HG22 | 1.48                     | 1.42              |
| 1:A:584:PRO:CB | 2:J:30:THR:HG22 | 1.50                     | 1.40              |
| 1:B:609:ALA:CB | 2:K:53:TYR:OH   | 1.70                     | 1.40              |
| 1:B:584:PRO:CB | 2:K:30:THR:HG22 | 1.50                     | 1.38              |
| 1:D:584:PRO:CB | 2:I:30:THR:HG22 | 1.50                     | 1.38              |

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   | A     | 1026/1024 (100%) | 954 (93%)  | 66 (6%)  | 6 (1%)   | 30          | 74  |
| 1   | B     | 1026/1024 (100%) | 953 (93%)  | 67 (6%)  | 6 (1%)   | 30          | 74  |
| 1   | C     | 1026/1024 (100%) | 953 (93%)  | 67 (6%)  | 6 (1%)   | 30          | 74  |
| 1   | D     | 1026/1024 (100%) | 953 (93%)  | 67 (6%)  | 6 (1%)   | 30          | 74  |
| 2   | H     | 112/114 (98%)    | 109 (97%)  | 3 (3%)   | 0        | 100         | 100 |
| 2   | I     | 112/114 (98%)    | 109 (97%)  | 3 (3%)   | 0        | 100         | 100 |
| 2   | J     | 112/114 (98%)    | 109 (97%)  | 3 (3%)   | 0        | 100         | 100 |
| 2   | K     | 112/114 (98%)    | 109 (97%)  | 3 (3%)   | 0        | 100         | 100 |
| 3   | L     | 105/107 (98%)    | 102 (97%)  | 3 (3%)   | 0        | 100         | 100 |
| 3   | M     | 105/107 (98%)    | 102 (97%)  | 3 (3%)   | 0        | 100         | 100 |
| 3   | N     | 105/107 (98%)    | 102 (97%)  | 3 (3%)   | 0        | 100         | 100 |
| 3   | O     | 105/107 (98%)    | 102 (97%)  | 3 (3%)   | 0        | 100         | 100 |
| All | All   | 4972/4980 (100%) | 4657 (94%) | 291 (6%) | 24 (0%)  | 38          | 77  |

5 of 24 Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1   | A     | 647 | SER  |
| 1   | B     | 647 | SER  |
| 1   | C     | 647 | SER  |
| 1   | D     | 647 | SER  |
| 1   | A     | 77  | ASP  |

### 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   | A     | 880/876 (100%)   | 754 (86%)  | 126 (14%) | 4           | 25 |
| 1   | B     | 880/876 (100%)   | 754 (86%)  | 126 (14%) | 4           | 25 |
| 1   | C     | 880/876 (100%)   | 754 (86%)  | 126 (14%) | 4           | 25 |
| 1   | D     | 880/876 (100%)   | 754 (86%)  | 126 (14%) | 4           | 25 |
| 2   | H     | 102/102 (100%)   | 101 (99%)  | 1 (1%)    | 82          | 92 |
| 2   | I     | 102/102 (100%)   | 101 (99%)  | 1 (1%)    | 82          | 92 |
| 2   | J     | 102/102 (100%)   | 101 (99%)  | 1 (1%)    | 82          | 92 |
| 2   | K     | 102/102 (100%)   | 101 (99%)  | 1 (1%)    | 82          | 92 |
| 3   | L     | 94/94 (100%)     | 93 (99%)   | 1 (1%)    | 80          | 91 |
| 3   | M     | 94/94 (100%)     | 93 (99%)   | 1 (1%)    | 80          | 91 |
| 3   | N     | 94/94 (100%)     | 93 (99%)   | 1 (1%)    | 80          | 91 |
| 3   | O     | 94/94 (100%)     | 93 (99%)   | 1 (1%)    | 80          | 91 |
| All | All   | 4304/4288 (100%) | 3792 (88%) | 512 (12%) | 11          | 31 |

5 of 512 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1   | B     | 843 | GLN  |
| 1   | C     | 251 | ARG  |
| 1   | D     | 773 | LYS  |
| 1   | B     | 894 | ARG  |
| 1   | C     | 71  | GLU  |

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 115 such sidechains are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1   | C     | 216 | HIS  |
| 1   | C     | 622 | HIS  |
| 2   | K     | 5   | GLN  |
| 1   | C     | 226 | HIS  |
| 1   | C     | 424 | ASN  |

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.



## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

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.