



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 19, 2016 – 10:04 PM GMT

PDB ID : 5C3I  
Title : Crystal structure of the quaternary complex of histone H3-H4 heterodimer with chaperone ASF1 and the replicative helicase subunit MCM2  
Authors : Wang, H.; Wang, M.; Yang, N.; Xu, R.M.  
Deposited on : 2015-06-17  
Resolution : 3.50 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

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

|                                |   |  |
|--------------------------------|---|--|
| MolProbity                     | : | 4.02b-467  |
| Mogul                          | : | unknown  |
| Xtriage (Phenix)               | : | 1.9-1692   |
| EDS                            | : | rb-20026982  |
| Percentile statistics          | : | 20151230.v01 (using entries in the PDB archive December 30th 2015) |
| Refmac                         | : | 5.8.0135   |
| CCP4                           | : | 6.5.0  |
| Ideal geometry (proteins)      | : | Engh & Huber (2001)  |
| Ideal geometry (DNA, RNA)      | : | Parkinson et al. (1996)  |
| Validation Pipeline (wwPDB-VP) | : | rb-20026982  |

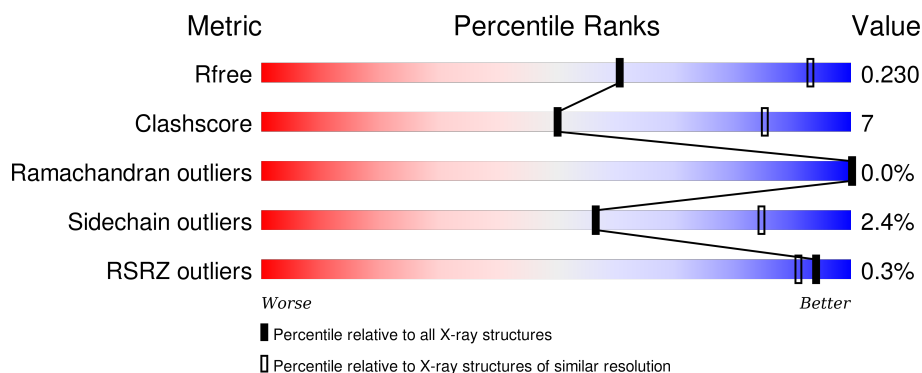
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.50 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) | Similar resolution<br>(#Entries, resolution range(Å)) |
|-----------------------|-----------------------------|---|
| $R_{free}$            | 91344                       | 1051 (3.60-3.40)                                      |
| Clashscore            | 102246                      | 1157 (3.60-3.40)                                      |
| Ramachandran outliers | 100387                      | 1120 (3.60-3.40)                                      |
| Sidechain outliers    | 100360                      | 1121 (3.60-3.40)                                      |
| RSRZ outliers         | 91569                       | 1058 (3.60-3.40)                                      |

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower 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  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

| Mol | Chain | Length | Quality of chain  |
|-----|-------|--------|---|
| 1   | A     | 188    | <div> <div>67%</div> <div>13%</div> <div>•</div> <div>19%</div> </div>              |
| 1   | E     | 188    | <div> <div>69%</div> <div>11%</div> <div>•</div> <div>19%</div> </div>              |
| 1   | I     | 188    | <div> <div>%</div> <div>65%</div> <div>16%</div> <div>•</div> <div>19%</div> </div> |
| 1   | M     | 188    | <div> <div>66%</div> <div>14%</div> <div>•</div> <div>19%</div> </div>              |
| 1   | Q     | 188    | <div> <div>61%</div> <div>20%</div> <div>•</div> <div>18%</div> </div>              |

*Continued on next page...*

Continued from previous page...

| Mol | Chain | Length | Quality of chain |
|-----|-------|--------|------------------|
| 1   | U     | 188    |                  |
| 2   | B     | 136    |                  |
| 2   | F     | 136    |                  |
| 2   | J     | 136    |                  |
| 2   | N     | 136    |                  |
| 2   | R     | 136    |                  |
| 2   | V     | 136    |                  |
| 3   | C     | 103    |                  |
| 3   | G     | 103    |                  |
| 3   | K     | 103    |                  |
| 3   | O     | 103    |                  |
| 3   | S     | 103    |                  |
| 3   | W     | 103    |                  |
| 4   | D     | 93     |                  |
| 4   | H     | 93     |                  |
| 4   | L     | 93     |                  |
| 4   | P     | 93     |                  |
| 4   | T     | 93     |                  |
| 4   | X     | 93     |                  |

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 17655 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Histone chaperone ASF1A.

| Mol | Chain | Residues | Atoms |     |     |     |   | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---|---------|---------|-------|
| 1   | A     | 152      | Total | C   | N   | O   | S | 0       | 0       | 0     |
|     |       |          | 1223  | 786 | 200 | 234 | 3 |         |         |       |
| 1   | E     | 153      | Total | C   | N   | O   | S | 0       | 0       | 0     |
|     |       |          | 1232  | 791 | 201 | 237 | 3 |         |         |       |
| 1   | I     | 153      | Total | C   | N   | O   | S | 0       | 0       | 0     |
|     |       |          | 1232  | 791 | 201 | 237 | 3 |         |         |       |
| 1   | M     | 152      | Total | C   | N   | O   | S | 0       | 0       | 0     |
|     |       |          | 1223  | 786 | 200 | 234 | 3 |         |         |       |
| 1   | Q     | 154      | Total | C   | N   | O   | S | 0       | 0       | 0     |
|     |       |          | 1240  | 795 | 202 | 240 | 3 |         |         |       |
| 1   | U     | 150      | Total | C   | N   | O   | S | 0       | 0       | 0     |
|     |       |          | 1204  | 772 | 197 | 232 | 3 |         |         |       |

There are 78 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment        | Reference  |
|-------|---------|----------|--------|----------------|------------|
| A     | -12     | MET      | -      | expression tag | UNP Q9Y294 |
| A     | -11     | GLY      | -      | expression tag | UNP Q9Y294 |
| A     | -10     | SER      | -      | expression tag | UNP Q9Y294 |
| A     | -9      | HIS      | -      | expression tag | UNP Q9Y294 |
| A     | -8      | HIS      | -      | expression tag | UNP Q9Y294 |
| A     | -7      | HIS      | -      | expression tag | UNP Q9Y294 |
| A     | -6      | HIS      | -      | expression tag | UNP Q9Y294 |
| A     | -5      | HIS      | -      | expression tag | UNP Q9Y294 |
| A     | -4      | HIS      | -      | expression tag | UNP Q9Y294 |
| A     | -3      | SER      | -      | expression tag | UNP Q9Y294 |
| A     | -2      | ASN      | -      | expression tag | UNP Q9Y294 |
| A     | -1      | ASP      | -      | expression tag | UNP Q9Y294 |
| A     | 0       | PRO      | -      | expression tag | UNP Q9Y294 |
| E     | -12     | MET      | -      | expression tag | UNP Q9Y294 |
| E     | -11     | GLY      | -      | expression tag | UNP Q9Y294 |
| E     | -10     | SER      | -      | expression tag | UNP Q9Y294 |
| E     | -9      | HIS      | -      | expression tag | UNP Q9Y294 |

*Continued on next page...*

*Continued from previous page...*

| Chain | Residue | Modelled | Actual | Comment        | Reference  |
|-------|---------|----------|--------|----------------|------------|
| E     | -8      | HIS      | -      | expression tag | UNP Q9Y294 |
| E     | -7      | HIS      | -      | expression tag | UNP Q9Y294 |
| E     | -6      | HIS      | -      | expression tag | UNP Q9Y294 |
| E     | -5      | HIS      | -      | expression tag | UNP Q9Y294 |
| E     | -4      | HIS      | -      | expression tag | UNP Q9Y294 |
| E     | -3      | SER      | -      | expression tag | UNP Q9Y294 |
| E     | -2      | ASN      | -      | expression tag | UNP Q9Y294 |
| E     | -1      | ASP      | -      | expression tag | UNP Q9Y294 |
| E     | 0       | PRO      | -      | expression tag | UNP Q9Y294 |
| I     | -12     | MET      | -      | expression tag | UNP Q9Y294 |
| I     | -11     | GLY      | -      | expression tag | UNP Q9Y294 |
| I     | -10     | SER      | -      | expression tag | UNP Q9Y294 |
| I     | -9      | HIS      | -      | expression tag | UNP Q9Y294 |
| I     | -8      | HIS      | -      | expression tag | UNP Q9Y294 |
| I     | -7      | HIS      | -      | expression tag | UNP Q9Y294 |
| I     | -6      | HIS      | -      | expression tag | UNP Q9Y294 |
| I     | -5      | HIS      | -      | expression tag | UNP Q9Y294 |
| I     | -4      | HIS      | -      | expression tag | UNP Q9Y294 |
| I     | -3      | SER      | -      | expression tag | UNP Q9Y294 |
| I     | -2      | ASN      | -      | expression tag | UNP Q9Y294 |
| I     | -1      | ASP      | -      | expression tag | UNP Q9Y294 |
| I     | 0       | PRO      | -      | expression tag | UNP Q9Y294 |
| M     | -12     | MET      | -      | expression tag | UNP Q9Y294 |
| M     | -11     | GLY      | -      | expression tag | UNP Q9Y294 |
| M     | -10     | SER      | -      | expression tag | UNP Q9Y294 |
| M     | -9      | HIS      | -      | expression tag | UNP Q9Y294 |
| M     | -8      | HIS      | -      | expression tag | UNP Q9Y294 |
| M     | -7      | HIS      | -      | expression tag | UNP Q9Y294 |
| M     | -6      | HIS      | -      | expression tag | UNP Q9Y294 |
| M     | -5      | HIS      | -      | expression tag | UNP Q9Y294 |
| M     | -4      | HIS      | -      | expression tag | UNP Q9Y294 |
| M     | -3      | SER      | -      | expression tag | UNP Q9Y294 |
| M     | -2      | ASN      | -      | expression tag | UNP Q9Y294 |
| M     | -1      | ASP      | -      | expression tag | UNP Q9Y294 |
| M     | 0       | PRO      | -      | expression tag | UNP Q9Y294 |
| Q     | -12     | MET      | -      | expression tag | UNP Q9Y294 |
| Q     | -11     | GLY      | -      | expression tag | UNP Q9Y294 |
| Q     | -10     | SER      | -      | expression tag | UNP Q9Y294 |
| Q     | -9      | HIS      | -      | expression tag | UNP Q9Y294 |
| Q     | -8      | HIS      | -      | expression tag | UNP Q9Y294 |
| Q     | -7      | HIS      | -      | expression tag | UNP Q9Y294 |
| Q     | -6      | HIS      | -      | expression tag | UNP Q9Y294 |

*Continued on next page...*

*Continued from previous page...*

| Chain | Residue | Modelled | Actual | Comment        | Reference  |
|-------|---------|----------|--------|----------------|------------|
| Q     | -5      | HIS      | -      | expression tag | UNP Q9Y294 |
| Q     | -4      | HIS      | -      | expression tag | UNP Q9Y294 |
| Q     | -3      | SER      | -      | expression tag | UNP Q9Y294 |
| Q     | -2      | ASN      | -      | expression tag | UNP Q9Y294 |
| Q     | -1      | ASP      | -      | expression tag | UNP Q9Y294 |
| Q     | 0       | PRO      | -      | expression tag | UNP Q9Y294 |
| U     | -12     | MET      | -      | expression tag | UNP Q9Y294 |
| U     | -11     | GLY      | -      | expression tag | UNP Q9Y294 |
| U     | -10     | SER      | -      | expression tag | UNP Q9Y294 |
| U     | -9      | HIS      | -      | expression tag | UNP Q9Y294 |
| U     | -8      | HIS      | -      | expression tag | UNP Q9Y294 |
| U     | -7      | HIS      | -      | expression tag | UNP Q9Y294 |
| U     | -6      | HIS      | -      | expression tag | UNP Q9Y294 |
| U     | -5      | HIS      | -      | expression tag | UNP Q9Y294 |
| U     | -4      | HIS      | -      | expression tag | UNP Q9Y294 |
| U     | -3      | SER      | -      | expression tag | UNP Q9Y294 |
| U     | -2      | ASN      | -      | expression tag | UNP Q9Y294 |
| U     | -1      | ASP      | -      | expression tag | UNP Q9Y294 |
| U     | 0       | PRO      | -      | expression tag | UNP Q9Y294 |

- Molecule 2 is a protein called Histone H3.1.

| Mol | Chain | Residues | Atoms |     |     |     |   | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---|---------|---------|-------|
| 2   | B     | 87       | Total | C   | N   | O   | S | 0       | 0       | 0     |
|     |       |          | 720   | 454 | 137 | 125 | 4 |         |         |       |
| 2   | F     | 88       | Total | C   | N   | O   | S | 0       | 0       | 0     |
|     |       |          | 725   | 457 | 138 | 126 | 4 |         |         |       |
| 2   | J     | 84       | Total | C   | N   | O   | S | 0       | 0       | 0     |
|     |       |          | 688   | 435 | 130 | 119 | 4 |         |         |       |
| 2   | N     | 76       | Total | C   | N   | O   | S | 0       | 0       | 0     |
|     |       |          | 615   | 390 | 115 | 106 | 4 |         |         |       |
| 2   | R     | 74       | Total | C   | N   | O   | S | 0       | 0       | 0     |
|     |       |          | 599   | 381 | 110 | 104 | 4 |         |         |       |
| 2   | V     | 76       | Total | C   | N   | O   | S | 0       | 0       | 0     |
|     |       |          | 619   | 392 | 115 | 108 | 4 |         |         |       |

- Molecule 3 is a protein called Histone H4.

| Mol | Chain | Residues | Atoms |     |     |     |   | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---|---------|---------|-------|
| 3   | C     | 78       | Total | C   | N   | O   | S | 0       | 0       | 0     |
|     |       |          | 622   | 393 | 120 | 108 | 1 |         |         |       |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Residues | Atoms |     |     |     |   | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---|---------|---------|-------|
| 3   | G     | 77       | Total | C   | N   | O   | S | 0       | 0       | 0     |
|     |       |          | 618   | 391 | 119 | 107 | 1 |         |         |       |
| 3   | K     | 76       | Total | C   | N   | O   | S | 0       | 0       | 0     |
|     |       |          | 606   | 385 | 117 | 103 | 1 |         |         |       |
| 3   | O     | 78       | Total | C   | N   | O   | S | 0       | 0       | 0     |
|     |       |          | 622   | 393 | 120 | 108 | 1 |         |         |       |
| 3   | S     | 78       | Total | C   | N   | O   | S | 0       | 0       | 0     |
|     |       |          | 622   | 393 | 120 | 108 | 1 |         |         |       |
| 3   | W     | 77       | Total | C   | N   | O   | S | 0       | 0       | 0     |
|     |       |          | 618   | 391 | 119 | 107 | 1 |         |         |       |

- Molecule 4 is a protein called DNA replication licensing factor MCM2,MCM2.

| Mol | Chain | Residues | Atoms |     |    |     |   | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|-----|----|-----|---|---------|---------|-------|
| 4   | D     | 63       | Total | C   | N  | O   | S | 0       | 0       | 0     |
|     |       |          | 481   | 284 | 86 | 109 | 2 |         |         |       |
| 4   | H     | 56       | Total | C   | N  | O   | S | 0       | 0       | 0     |
|     |       |          | 446   | 263 | 79 | 102 | 2 |         |         |       |
| 4   | L     | 56       | Total | C   | N  | O   | S | 0       | 0       | 0     |
|     |       |          | 446   | 263 | 79 | 102 | 2 |         |         |       |
| 4   | P     | 48       | Total | C   | N  | O   | S | 0       | 0       | 0     |
|     |       |          | 365   | 219 | 58 | 87  | 1 |         |         |       |
| 4   | T     | 56       | Total | C   | N  | O   | S | 0       | 0       | 0     |
|     |       |          | 443   | 262 | 79 | 100 | 2 |         |         |       |
| 4   | X     | 56       | Total | C   | N  | O   | S | 0       | 0       | 0     |
|     |       |          | 446   | 263 | 79 | 102 | 2 |         |         |       |

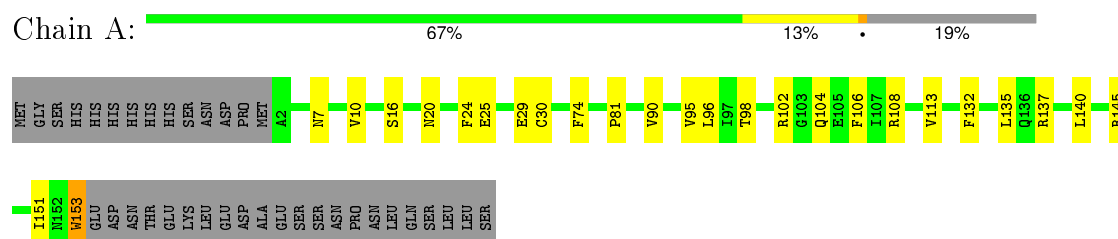
There are 6 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment        | Reference  |
|-------|---------|----------|--------|----------------|------------|
| D     | 62      | SER      | -      | expression tag | UNP P49736 |
| H     | 62      | SER      | -      | expression tag | UNP P49736 |
| L     | 62      | SER      | -      | expression tag | UNP P49736 |
| P     | 62      | SER      | -      | expression tag | UNP P49736 |
| T     | 62      | SER      | -      | expression tag | UNP P49736 |
| X     | 62      | SER      | -      | expression tag | UNP P49736 |

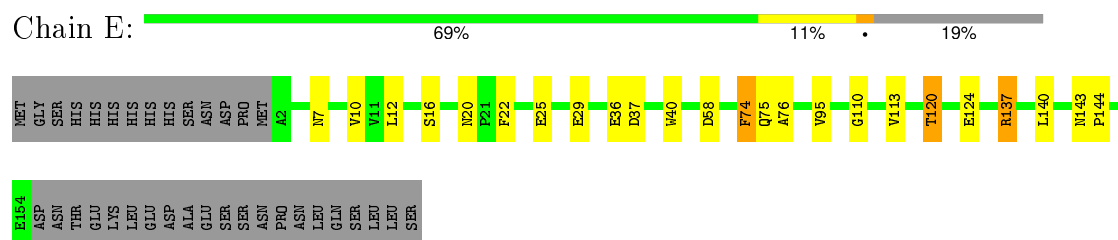
### 3 Residue-property plots [i](#)

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 and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

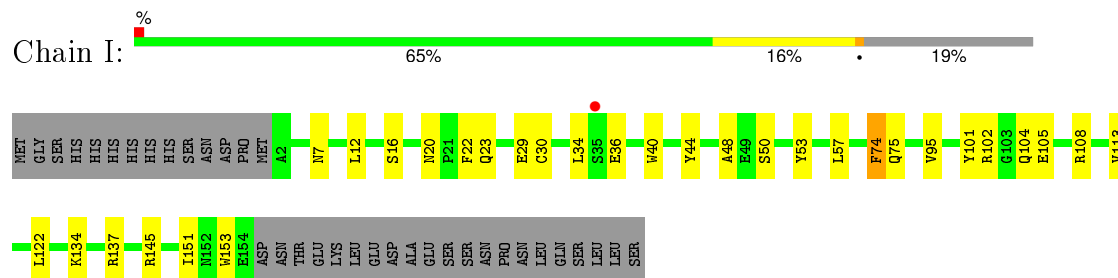
#### • Molecule 1: Histone chaperone ASF1A



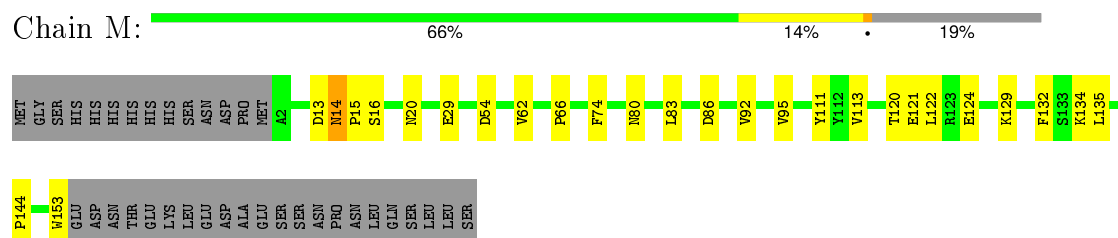
#### • Molecule 1: Histone chaperone ASF1A



#### • Molecule 1: Histone chaperone ASF1A



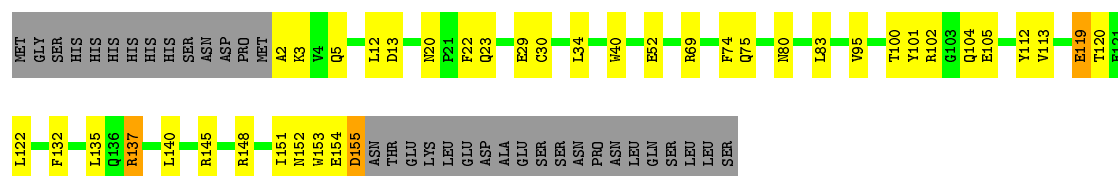
#### • Molecule 1: Histone chaperone ASF1A





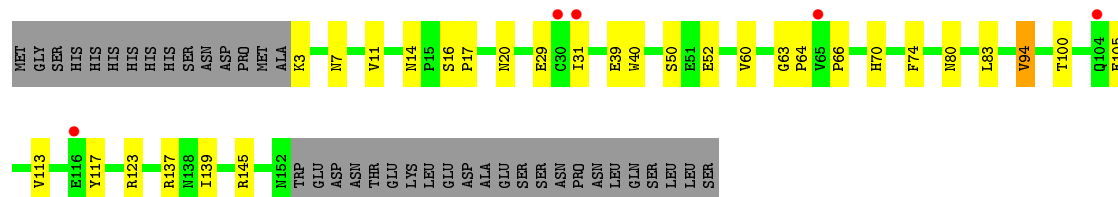
- Molecule 1: Histone chaperone ASF1A

Chain Q: 



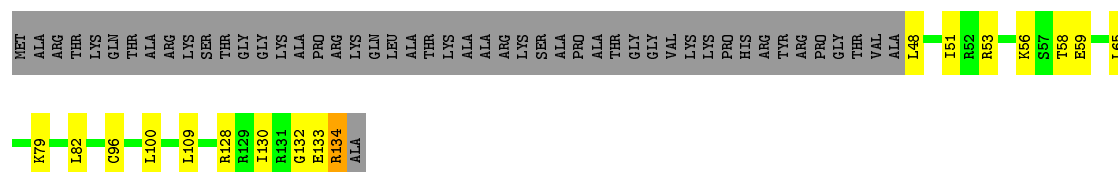
- Molecule 1: Histone chaperone ASF1A

Chain U: 



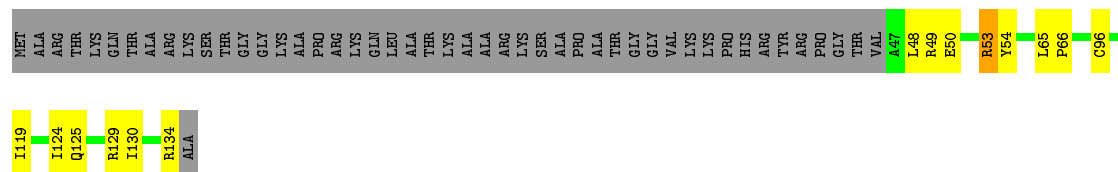
- Molecule 2: Histone H3.1

Chain B: 



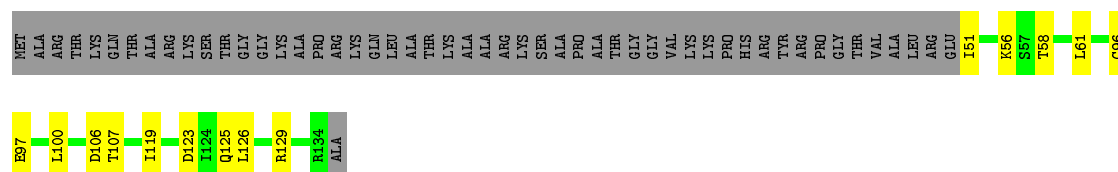
- Molecule 2: Histone H3.1

Chain F: 



- Molecule 2: Histone H3.1

Chain J: 



- Molecule 2: Histone H3.1

Chain N:  44% 12% 44%

MET ALA ARG THR LYS GLN THR ALA ARG LYS SER SER THR GLY LYS ALA PRO ARG LYS GLN LEU THR LYS ALA ALA LYS LYS LYS VAL GLY THR LYS LYS PRO HIS ARG TYR ARG PRO GLY THR VAL ALA LEU ARG GLU ILE ARG ARG TYR GLN LYS SER THR E59

L65 R69 E73 K79 L82 M90 I112 H113 A114 K115 R116 V117 T118 I119 K122 Q125 R129 R131 R134 ALA

- Molecule 2: Histone H3.1

Chain R:  43% 12% 46%

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

L60 L61 L62 R63 R64 L65 Q68 E73 L82 R83 C96 E97 L100 T107 N108 L109 D123 I124 Q125 R128 R129 E133 ARG ALA

- Molecule 2: Histone H3.1

Chain V:  47% 9% 44%

MET ALA ARG THR LYS GLN THR ALA ARG LYS SER SER THR GLY LYS ALA PRO ARG LYS GLN LEU THR LYS ALA ALA LYS LYS LYS VAL GLY THR LYS LYS PRO HIS ARG TYR ARG PRO GLY THR VAL ALA LEU ARG GLU ILE ARG ARG TYR GLN LYS SER THR E59

L60 L61 L62 R63 R64 L65 R69 E73 Q76 K79 T80 D81 L82 E97 L100 I130 R134 ALA

- Molecule 3: Histone H4

Chain C:  64% 12% 24%

MET SER GLY ARG GLY LYS GLY LYS LYS LEU GLY LYS GLY GLY ALA ALA LYS ARG HIS ARG LYS VAL LEU ARG D24 D25 I29 T30 K31 P32 L58 L62 R67 D68 A69 Y72 A76 F100 G101 GLY

- Molecule 3: Histone H4

Chain G:  63% 11% 25%

MET SER GLY ARG GLY LYS GLY LYS LYS LEU GLY LYS GLY GLY ALA ALA LYS ARG HIS ARG LYS VAL LEU ARG D24 D25 I29 I46 I50 E53 V57 F61 L62 V65 T71 T82 R92 G99 F100 GLY GLY

- Molecule 3: Histone H4

Chain K:  70% 26%

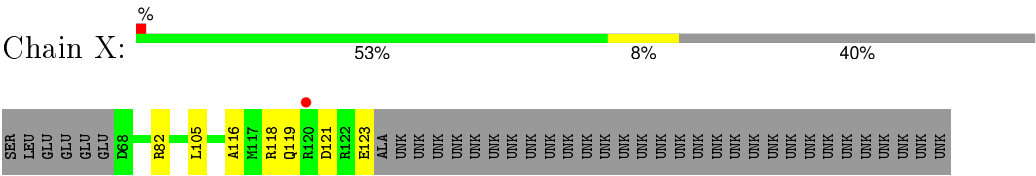
MET SER GLY ARG GLY LYS GLY LYS LYS LEU GLY LYS GLY GLY ALA ALA LYS ARG HIS ARG LYS VAL LEU ARG ASP ASN I26 R40 I46 L68 A76 G101 GLY

- Molecule 3: Histone H4

Chain O:  66% 10% 24%



● Molecule 4: DNA replication licensing factor MCM2,MCM2



## 4 Data and refinement statistics

| Property  | Value   | Source           |
|---|---|------------------|
| Space group   | P 31 2 1  | Depositor        |
| Cell constants<br>a, b, c, $\alpha$ , $\beta$ , $\gamma$                | 147.70Å 147.70Å 261.80Å<br>90.00° 90.00° 120.00°              | Depositor        |
| Resolution (Å)  | 48.99 – 3.50<br>48.98 – 3.48                                  | Depositor<br>EDS |
| % Data completeness<br>(in resolution range)                            | 92.8 (48.99-3.50)<br>92.9 (48.98-3.48)                        | Depositor<br>EDS |
| $R_{merge}$   | 0.16  | Depositor        |
| $R_{sym}$   | (Not available)   | Depositor        |
| $\langle I/\sigma(I) \rangle$ <sup>1</sup>                              | 4.20 (at 3.48Å)   | Xtriage          |
| Refinement program  | REFMAC  | Depositor        |
| R, $R_{free}$   | 0.216 , 0.253<br>0.191 , 0.230                                | Depositor<br>DCC |
| $R_{free}$ test set   | 1972 reflections (5.24%)                                      | DCC              |
| Wilson B-factor (Å <sup>2</sup> )                                       | 46.2  | Xtriage          |
| Anisotropy  | 0.318   | Xtriage          |
| Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> ) | 0.34 , -4.9   | EDS              |
| Estimated twinning fraction   | 0.787 for H, K, L<br>0.213 for -H, -K, L<br>0.155 for -h,-k,l | Xtriage          |
| Reported twinning fraction  | 0.787 for H, K, L<br>0.213 for -H, -K, L                      | Depositor        |
| L-test for twinning <sup>2</sup>  | $\langle  L  \rangle = 0.40$ , $\langle L^2 \rangle = 0.22$   | Xtriage          |
| Outliers  | 0 of 39980 reflections  | Xtriage          |
| $F_o, F_c$ correlation  | 0.87  | EDS              |
| Total number of atoms   | 17655   | wwPDB-VP         |
| Average B, all atoms (Å <sup>2</sup> )                                  | 47.0  | wwPDB-VP         |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.44% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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

| Mol | Chain | Bond lengths |                | Bond angles |         |
|-----|-------|--------------|----------------|-------------|---------|
|     |       | RMSZ         | # Z  >5        | RMSZ        | # Z  >5 |
| 1   | A     | 0.42         | 1/1256 (0.1%)  | 0.52        | 0/1719  |
| 1   | E     | 0.42         | 0/1265         | 0.48        | 0/1731  |
| 1   | I     | 0.42         | 1/1265 (0.1%)  | 0.51        | 0/1731  |
| 1   | M     | 0.42         | 1/1256 (0.1%)  | 0.50        | 0/1719  |
| 1   | Q     | 0.42         | 1/1273 (0.1%)  | 0.52        | 0/1742  |
| 1   | U     | 0.38         | 1/1235 (0.1%)  | 0.47        | 0/1689  |
| 2   | B     | 0.33         | 0/728          | 0.49        | 0/974   |
| 2   | F     | 0.32         | 0/733          | 0.49        | 0/981   |
| 2   | J     | 0.33         | 0/696          | 0.51        | 0/932   |
| 2   | N     | 0.33         | 0/622          | 0.49        | 0/834   |
| 2   | R     | 0.33         | 0/606          | 0.55        | 0/813   |
| 2   | V     | 0.33         | 0/626          | 0.50        | 0/839   |
| 3   | C     | 0.32         | 0/629          | 0.51        | 0/843   |
| 3   | G     | 0.32         | 0/625          | 0.47        | 0/838   |
| 3   | K     | 0.33         | 0/613          | 0.49        | 0/821   |
| 3   | O     | 0.33         | 0/629          | 0.50        | 0/843   |
| 3   | S     | 0.32         | 0/629          | 0.48        | 0/843   |
| 3   | W     | 0.31         | 0/625          | 0.47        | 0/838   |
| 4   | D     | 0.28         | 0/448          | 0.51        | 0/600   |
| 4   | H     | 0.27         | 0/448          | 0.48        | 0/600   |
| 4   | L     | 0.29         | 0/448          | 0.51        | 0/600   |
| 4   | P     | 0.31         | 0/367          | 0.52        | 0/495   |
| 4   | T     | 0.31         | 0/445          | 0.51        | 0/596   |
| 4   | X     | 0.28         | 0/448          | 0.51        | 0/600   |
| All | All   | 0.36         | 5/17915 (0.0%) | 0.50        | 0/24221 |

All (5) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms   | Z    | Observed(Å) | Ideal(Å) |
|-----|-------|-----|------|---------|------|-------------|----------|
| 1   | A     | 153 | TRP  | CD2-CE2 | 5.04 | 1.47        | 1.41     |
| 1   | U     | 40  | TRP  | CD2-CE2 | 5.03 | 1.47        | 1.41     |
| 1   | Q     | 40  | TRP  | CD2-CE2 | 5.03 | 1.47        | 1.41     |
| 1   | M     | 153 | TRP  | CD2-CE2 | 5.02 | 1.47        | 1.41     |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type | Atoms   | Z    | Observed(Å) | Ideal(Å) |
|-----|-------|-----|------|---------|------|-------------|----------|
| 1   | I     | 153 | TRP  | CD2-CE2 | 5.01 | 1.47        | 1.41     |

There are no bond angle outliers.

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     | 1223  | 0        | 1176     | 19      | 0            |
| 1   | E     | 1232  | 0        | 1182     | 16      | 0            |
| 1   | I     | 1232  | 0        | 1182     | 22      | 0            |
| 1   | M     | 1223  | 0        | 1176     | 16      | 0            |
| 1   | Q     | 1240  | 0        | 1186     | 29      | 0            |
| 1   | U     | 1204  | 0        | 1161     | 23      | 0            |
| 2   | B     | 720   | 0        | 758      | 18      | 0            |
| 2   | F     | 725   | 0        | 763      | 11      | 0            |
| 2   | J     | 688   | 0        | 722      | 14      | 0            |
| 2   | N     | 615   | 0        | 645      | 14      | 0            |
| 2   | R     | 599   | 0        | 630      | 13      | 0            |
| 2   | V     | 619   | 0        | 649      | 9       | 0            |
| 3   | C     | 622   | 0        | 660      | 9       | 0            |
| 3   | G     | 618   | 0        | 657      | 11      | 0            |
| 3   | K     | 606   | 0        | 650      | 6       | 0            |
| 3   | O     | 622   | 0        | 660      | 8       | 0            |
| 3   | S     | 622   | 0        | 660      | 14      | 0            |
| 3   | W     | 618   | 0        | 657      | 15      | 0            |
| 4   | D     | 481   | 0        | 412      | 6       | 0            |
| 4   | H     | 446   | 0        | 403      | 7       | 0            |
| 4   | L     | 446   | 0        | 403      | 6       | 0            |
| 4   | P     | 365   | 0        | 322      | 1       | 0            |
| 4   | T     | 443   | 0        | 404      | 11      | 0            |
| 4   | X     | 446   | 0        | 403      | 5       | 0            |
| All | All   | 17655 | 0        | 17521    | 247     | 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 7.

The worst 5 of 247 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) |
|-----------------|------------------|--------------------------|-------------------|
| 3:W:29:ILE:CD1  | 3:W:58:LEU:HD23  | 1.59                     | 1.31              |
| 3:W:29:ILE:HD13 | 3:W:58:LEU:CD2   | 1.75                     | 1.17              |
| 1:Q:112:TYR:HB2 | 1:Q:140:LEU:HD12 | 1.34                     | 1.08              |
| 3:W:29:ILE:HD13 | 3:W:58:LEU:HD23  | 1.08                     | 1.04              |
| 1:I:16:SER:HB2  | 1:I:20:ASN:HD22  | 1.25                     | 1.01              |

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

| Mol | Chain | Analysed      | Favoured   | Allowed | Outliers | Percentiles |     |
|-----|-------|---------------|------------|---------|----------|-------------|-----|
| 1   | A     | 150/188 (80%) | 150 (100%) | 0       | 0        | 100         | 100 |
| 1   | E     | 151/188 (80%) | 149 (99%)  | 2 (1%)  | 0        | 100         | 100 |
| 1   | I     | 151/188 (80%) | 149 (99%)  | 2 (1%)  | 0        | 100         | 100 |
| 1   | M     | 150/188 (80%) | 145 (97%)  | 5 (3%)  | 0        | 100         | 100 |
| 1   | Q     | 152/188 (81%) | 149 (98%)  | 3 (2%)  | 0        | 100         | 100 |
| 1   | U     | 148/188 (79%) | 143 (97%)  | 5 (3%)  | 0        | 100         | 100 |
| 2   | B     | 85/136 (62%)  | 83 (98%)   | 2 (2%)  | 0        | 100         | 100 |
| 2   | F     | 86/136 (63%)  | 84 (98%)   | 2 (2%)  | 0        | 100         | 100 |
| 2   | J     | 82/136 (60%)  | 79 (96%)   | 3 (4%)  | 0        | 100         | 100 |
| 2   | N     | 74/136 (54%)  | 72 (97%)   | 2 (3%)  | 0        | 100         | 100 |
| 2   | R     | 72/136 (53%)  | 70 (97%)   | 2 (3%)  | 0        | 100         | 100 |
| 2   | V     | 74/136 (54%)  | 72 (97%)   | 2 (3%)  | 0        | 100         | 100 |
| 3   | C     | 76/103 (74%)  | 75 (99%)   | 1 (1%)  | 0        | 100         | 100 |

*Continued on next page...*



Continued from previous page...

| Mol | Chain | Analysed        | Favoured   | Allowed | Outliers | Percentiles |     |
|-----|-------|-----------------|------------|---------|----------|-------------|-----|
| 3   | G     | 75/103 (73%)    | 75 (100%)  | 0       | 0        | 100         | 100 |
| 3   | K     | 74/103 (72%)    | 74 (100%)  | 0       | 0        | 100         | 100 |
| 3   | O     | 76/103 (74%)    | 73 (96%)   | 3 (4%)  | 0        | 100         | 100 |
| 3   | S     | 76/103 (74%)    | 74 (97%)   | 2 (3%)  | 0        | 100         | 100 |
| 3   | W     | 75/103 (73%)    | 75 (100%)  | 0       | 0        | 100         | 100 |
| 4   | D     | 54/93 (58%)     | 51 (94%)   | 3 (6%)  | 0        | 100         | 100 |
| 4   | H     | 54/93 (58%)     | 54 (100%)  | 0       | 0        | 100         | 100 |
| 4   | L     | 54/93 (58%)     | 52 (96%)   | 1 (2%)  | 1 (2%)   | 10          | 51  |
| 4   | P     | 46/93 (50%)     | 46 (100%)  | 0       | 0        | 100         | 100 |
| 4   | T     | 54/93 (58%)     | 53 (98%)   | 1 (2%)  | 0        | 100         | 100 |
| 4   | X     | 54/93 (58%)     | 51 (94%)   | 3 (6%)  | 0        | 100         | 100 |
| All | All   | 2143/3120 (69%) | 2098 (98%) | 44 (2%) | 1 (0%)   | 100         | 100 |

All (1) Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 4   | L     | 121 | ASP  |

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

| Mol | Chain | Analysed      | Rotameric | Outliers | Percentiles |    |
|-----|-------|---------------|-----------|----------|-------------|----|
| 1   | A     | 137/171 (80%) | 133 (97%) | 4 (3%)   | 50          | 81 |
| 1   | E     | 138/171 (81%) | 135 (98%) | 3 (2%)   | 60          | 85 |
| 1   | I     | 138/171 (81%) | 136 (99%) | 2 (1%)   | 74          | 91 |
| 1   | M     | 137/171 (80%) | 134 (98%) | 3 (2%)   | 60          | 85 |
| 1   | Q     | 139/171 (81%) | 131 (94%) | 8 (6%)   | 25          | 65 |
| 1   | U     | 136/171 (80%) | 133 (98%) | 3 (2%)   | 60          | 85 |
| 2   | B     | 77/111 (69%)  | 76 (99%)  | 1 (1%)   | 76          | 91 |

Continued on next page...

*Continued from previous page...*

| Mol | Chain | Analysed        | Rotameric  | Outliers | Percentiles |     |
|-----|-------|-----------------|------------|----------|-------------|-----|
| 2   | F     | 77/111 (69%)    | 75 (97%)   | 2 (3%)   | 54          | 83  |
| 2   | J     | 73/111 (66%)    | 72 (99%)   | 1 (1%)   | 74          | 91  |
| 2   | N     | 65/111 (59%)    | 60 (92%)   | 5 (8%)   | 16          | 54  |
| 2   | R     | 64/111 (58%)    | 62 (97%)   | 2 (3%)   | 47          | 81  |
| 2   | V     | 66/111 (60%)    | 66 (100%)  | 0        | 100         | 100 |
| 3   | C     | 64/79 (81%)     | 63 (98%)   | 1 (2%)   | 70          | 89  |
| 3   | G     | 64/79 (81%)     | 63 (98%)   | 1 (2%)   | 70          | 89  |
| 3   | K     | 62/79 (78%)     | 62 (100%)  | 0        | 100         | 100 |
| 3   | O     | 64/79 (81%)     | 64 (100%)  | 0        | 100         | 100 |
| 3   | S     | 64/79 (81%)     | 62 (97%)   | 2 (3%)   | 47          | 81  |
| 3   | W     | 64/79 (81%)     | 64 (100%)  | 0        | 100         | 100 |
| 4   | D     | 44/50 (88%)     | 41 (93%)   | 3 (7%)   | 20          | 60  |
| 4   | H     | 44/50 (88%)     | 43 (98%)   | 1 (2%)   | 58          | 85  |
| 4   | L     | 44/50 (88%)     | 42 (96%)   | 2 (4%)   | 34          | 73  |
| 4   | P     | 35/50 (70%)     | 35 (100%)  | 0        | 100         | 100 |
| 4   | T     | 43/50 (86%)     | 42 (98%)   | 1 (2%)   | 58          | 85  |
| 4   | X     | 44/50 (88%)     | 44 (100%)  | 0        | 100         | 100 |
| All | All   | 1883/2466 (76%) | 1838 (98%) | 45 (2%)  | 57          | 85  |

5 of 45 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 4   | L     | 98  | ASP  |
| 2   | N     | 115 | LYS  |
| 4   | T     | 120 | ARG  |
| 1   | M     | 74  | PHE  |
| 2   | N     | 117 | VAL  |

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1   | I     | 104 | GLN  |
| 1   | M     | 143 | ASN  |
| 1   | U     | 150 | HIS  |
| 1   | M     | 20  | ASN  |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 3   | O     | 25  | ASN  |

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

### 5.5 Carbohydrates [i](#)

There are no 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.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

| Mol | Chain | Analysed      | <RSRZ> | #RSRZ>2      | OWAB(Å <sup>2</sup> ) | Q<0.9 |
|-----|-------|---------------|--------|--------------|-----------------------|-------|
| 1   | A     | 152/188 (80%) | -0.11  | 0 100 100    | 26, 32, 55, 60        | 0     |
| 1   | E     | 153/188 (81%) | 0.14   | 0 100 100    | 31, 48, 84, 97        | 0     |
| 1   | I     | 153/188 (81%) | 0.05   | 1 (0%) 89 82 | 27, 40, 65, 72        | 0     |
| 1   | M     | 152/188 (80%) | -0.04  | 0 100 100    | 27, 37, 52, 60        | 0     |
| 1   | Q     | 154/188 (81%) | -0.03  | 0 100 100    | 24, 31, 46, 60        | 0     |
| 1   | U     | 150/188 (79%) | 0.46   | 5 (3%) 50 41 | 49, 83, 117, 132      | 0     |
| 2   | B     | 87/136 (63%)  | -0.08  | 0 100 100    | 27, 33, 98, 111       | 0     |
| 2   | F     | 88/136 (64%)  | -0.16  | 0 100 100    | 29, 35, 61, 70        | 0     |
| 2   | J     | 84/136 (61%)  | -0.18  | 0 100 100    | 27, 39, 59, 69        | 0     |
| 2   | N     | 76/136 (55%)  | -0.16  | 0 100 100    | 27, 41, 72, 77        | 0     |
| 2   | R     | 74/136 (54%)  | -0.18  | 0 100 100    | 29, 34, 49, 50        | 0     |
| 2   | V     | 76/136 (55%)  | 0.08   | 0 100 100    | 40, 51, 66, 72        | 0     |
| 3   | C     | 78/103 (75%)  | -0.14  | 0 100 100    | 24, 33, 47, 50        | 0     |
| 3   | G     | 77/103 (74%)  | -0.13  | 0 100 100    | 31, 42, 57, 69        | 0     |
| 3   | K     | 76/103 (73%)  | -0.08  | 0 100 100    | 27, 38, 56, 66        | 0     |
| 3   | O     | 78/103 (75%)  | 0.03   | 0 100 100    | 34, 50, 72, 76        | 0     |
| 3   | S     | 78/103 (75%)  | -0.21  | 0 100 100    | 31, 38, 48, 51        | 0     |
| 3   | W     | 77/103 (74%)  | -0.06  | 0 100 100    | 42, 50, 74, 83        | 0     |
| 4   | D     | 56/93 (60%)   | 0.05   | 0 100 100    | 26, 51, 70, 73        | 0     |
| 4   | H     | 56/93 (60%)   | 0.18   | 0 100 100    | 28, 52, 83, 87        | 0     |
| 4   | L     | 56/93 (60%)   | 0.17   | 0 100 100    | 30, 58, 75, 81        | 0     |
| 4   | P     | 48/93 (51%)   | 0.20   | 0 100 100    | 36, 57, 87, 91        | 0     |
| 4   | T     | 56/93 (60%)   | 0.08   | 0 100 100    | 35, 54, 72, 74        | 0     |
| 4   | X     | 56/93 (60%)   | 0.42   | 1 (1%) 71 62 | 41, 70, 93, 97        | 0     |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Analysed        | <RSRZ> | #RSRZ>2      | OWAB(Å <sup>2</sup> ) | Q<0.9 |
|-----|-------|-----------------|--------|--------------|-----------------------|-------|
| All | All   | 2191/3120 (70%) | 0.01   | 7 (0%) 94 91 | 24, 42, 83, 132       | 0     |

The worst 5 of 7 RSRZ outliers are listed below:

| Mol | Chain | Res | Type | RSRZ |
|-----|-------|-----|------|------|
| 1   | I     | 35  | SER  | 2.8  |
| 1   | U     | 30  | CYS  | 2.6  |
| 4   | X     | 120 | ARG  | 2.1  |
| 1   | U     | 31  | ILE  | 2.1  |
| 1   | U     | 65  | VAL  | 2.1  |

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

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

## 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.