



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 31, 2016 – 09:58 PM GMT

PDB ID : 1RI5  
Title : Structure and mechanism of mRNA cap (guanine N-7) methyltransferase  
Authors : Fabrega, C.; Hausmann, S.; Shen, V.; Shuman, S.; Lima, C.D.; Burley, S.K.;  
New York SGX Research Center for Structural Genomics (NYSGXRC)  
Deposited on : 2003-11-16  
Resolution : 2.10 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
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) : trunk26865

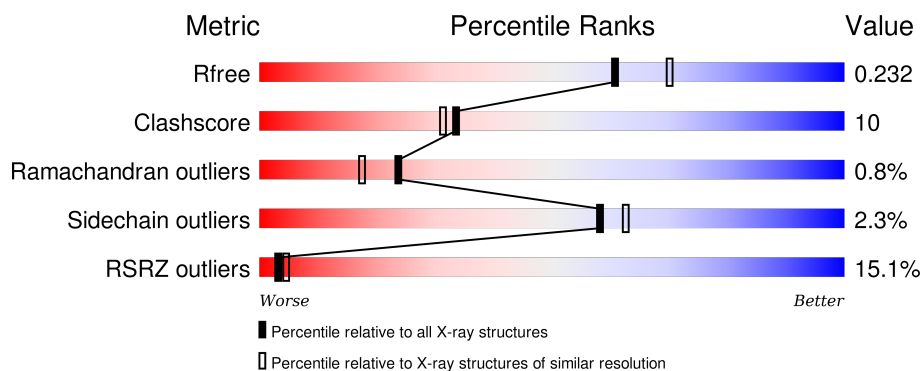
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	3939 (2.10-2.10)
Clashscore	102246	4460 (2.10-2.10)
Ramachandran outliers	100387	4413 (2.10-2.10)
Sidechain outliers	100360	4414 (2.10-2.10)
RSRZ outliers	91569	3948 (2.10-2.10)

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	298	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2220 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 mRNA CAPPING ENZYME.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	252	Total	C	N	O	S	0	0	0
			2061	1304	367	378	12			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	159	Total	O	0	0
			159	159		

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.

Chain A:

13% 62% 22% 15%

0.00 0.05 0.10 0.15 0.20 0.25

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300

MET ASP SER SER SER PRO LEU LYS THR PHE ARG LYS ASP GLN ALA MET GLU GLY LYS LYS LYS GLU GLU ILE ARG ARG ARG GLY ARG ARG GLU SER SER ARG ARG GLN ARG S41 I44 N45 I46 R47 K54 I58 R59 T62 S67 I80 E83 R90

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	63.77Å 63.77Å 112.78Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.73 – 2.10 19.73 – 2.05	Depositor EDS
% Data completeness (in resolution range)	94.7 (19.73-2.10) 94.4 (19.73-2.05)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	12.83 (at 2.06Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.230 , 0.263 0.231 , 0.232	Depositor DCC
$R_{free}$ test set	771 reflections (5.07%)	DCC
Wilson B-factor (Å <sup>2</sup> )	31.6	Xtriage
Anisotropy	0.262	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.41 , 57.2	EDS
Estimated twinning fraction	0.044 for -h,-k,l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 16194 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	2220	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.65% 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	0.45	0/2097	0.47	0/2806

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 [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2061	0	2053	42	0
2	A	159	0	0	7	0
All	All	2220	0	2053	42	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 10.

All (42) 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:A:54:LYS:HD2	1:A:141:PHE:HE1	1.53	0.72
1:A:177:ARG:HH22	1:A:207:GLU:HA	1.55	0.71
1:A:216:LEU:HB3	1:A:219:SER:HB3	1.74	0.67
1:A:246:GLU:HB3	1:A:288:VAL:HB	1.82	0.62

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:266:LYS:HA	2:A:454:HOH:O	2.03	0.58
1:A:213:ARG:NH1	1:A:222:ASN:HD21	2.03	0.56
1:A:190:SER:HB3	1:A:195:LYS:HA	1.89	0.55
1:A:133:GLU:HB3	1:A:165:ARG:HG3	1.89	0.54
1:A:263:GLU:HG2	1:A:264:LEU:H	1.71	0.54
1:A:182:GLU:O	1:A:186:GLN:HG3	2.09	0.53
1:A:132:LYS:HG3	2:A:444:HOH:O	2.10	0.52
1:A:202:GLU:O	1:A:204:VAL:HG23	2.11	0.51
1:A:265:SER:OG	1:A:270:LEU:HD23	2.11	0.49
1:A:177:ARG:CZ	1:A:181:LEU:HD11	2.43	0.49
1:A:44:ILE:HG23	1:A:45:ASN:N	2.28	0.49
1:A:128:MET:O	1:A:163:HIS:HE1	1.97	0.48
1:A:80:LEU:HD12	1:A:84:ARG:HH21	1.78	0.48
1:A:247:ARG:HA	1:A:286:VAL:O	2.14	0.48
1:A:251:ILE:O	1:A:255:GLU:HG3	2.14	0.47
1:A:119:ARG:HH11	1:A:130:LEU:HD23	1.80	0.47
1:A:190:SER:CB	1:A:195:LYS:HA	2.44	0.46
1:A:175:PRO:HG3	1:A:284:TYR:CE2	2.50	0.46
1:A:199:GLU:HB2	1:A:209:VAL:HG12	1.96	0.46
1:A:44:ILE:HG23	1:A:45:ASN:H	1.80	0.46
1:A:67:SER:OG	1:A:134:PHE:HA	2.16	0.46
1:A:263:GLU:HG2	1:A:264:LEU:N	2.30	0.45
1:A:199:GLU:HG3	2:A:327:HOH:O	2.16	0.45
1:A:206:MET:O	1:A:209:VAL:HG22	2.17	0.45
1:A:140:GLN:HE22	1:A:173:THR:HG23	1.82	0.45
1:A:47:ARG:HB2	2:A:443:HOH:O	2.16	0.44
1:A:224:ILE:HG12	2:A:347:HOH:O	2.17	0.44
1:A:58:ILE:O	1:A:62:THR:HG23	2.16	0.44
1:A:210:ARG:HD2	2:A:442:HOH:O	2.17	0.44
1:A:203:ASP:O	1:A:205:PRO:HD3	2.18	0.44
1:A:190:SER:HB2	1:A:194:TYR:O	2.18	0.43
1:A:107:ALA:HA	1:A:110:MET:HG2	2.00	0.43
1:A:266:LYS:C	1:A:268:MET:H	2.22	0.43
1:A:111:LYS:HG3	2:A:367:HOH:O	2.19	0.43
1:A:87:ILE:O	1:A:114:PHE:HB3	2.19	0.42
1:A:58:ILE:HG23	1:A:62:THR:HG21	2.02	0.42
1:A:140:GLN:HE22	1:A:173:THR:CG2	2.34	0.40
1:A:220:VAL:HB	1:A:223:CYS:HB2	2.02	0.40

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	250/298 (84%)	234 (94%)	14 (6%)	2 (1%)	24 17

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	142	SER
1	A	267	LYS

### 5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	222/264 (84%)	217 (98%)	5 (2%)	58 62

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	59	ARG
1	A	80	LEU
1	A	83	GLU
1	A	91	TYR
1	A	263	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:



Mol	Chain	Res	Type
1	A	140	GLN
1	A	261	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	252/298 (84%)	0.77	38 (15%) 3 4	22, 36, 85, 96	0

All (38) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	200	LYS	9.8
1	A	269	GLY	7.7
1	A	201	MET	7.1
1	A	268	MET	6.3
1	A	270	LEU	6.2
1	A	199	GLU	5.7
1	A	203	ASP	5.6
1	A	266	LYS	5.6
1	A	202	GLU	5.5
1	A	264	LEU	5.1
1	A	259	ARG	4.4
1	A	263	GLU	4.3
1	A	261	ASN	3.7
1	A	272	CYS	3.7
1	A	239	ARG	3.6
1	A	131	GLY	3.2
1	A	204	VAL	3.2
1	A	111	LYS	3.0
1	A	265	SER	3.0
1	A	198	LEU	2.9
1	A	180	ILE	2.9
1	A	218	ASP	2.8
1	A	292	LEU	2.6
1	A	224	ILE	2.6
1	A	267	LYS	2.5
1	A	207	GLU	2.5
1	A	90	TYR	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	276	GLU	2.4
1	A	212	TYR	2.4
1	A	160	ILE	2.4
1	A	238	LYS	2.3
1	A	113	ARG	2.2
1	A	237	PHE	2.2
1	A	213	ARG	2.2
1	A	41	SER	2.1
1	A	126	ARG	2.1
1	A	137	ILE	2.0
1	A	235	ASP	2.0

## 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.