wwPDB NMR Structure Validation Summary Report (i)

Feb 26, 2024 – 10:17 AM EST

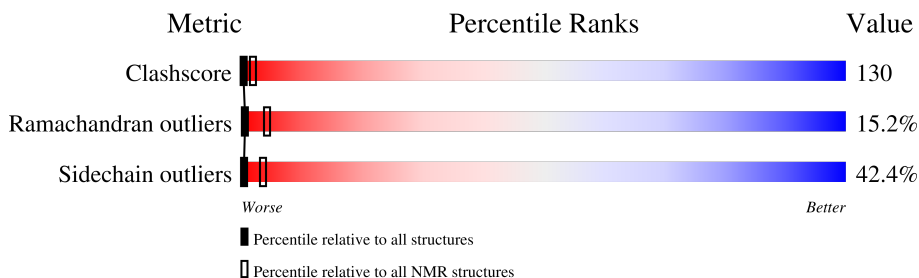|		
PDB ID	:	1RF8
Title	:	Solution structure of the yeast translation initiation factor eIF4E in complex with m7GDP and eIF4GI residues 393 to 490
Authors	:	Gross, J.D.; Moerke, N.J.; von der Haar, T.; Lugovskoy, A.A.; Sachs, A.B.; McCarthy, J.E.G.; Wagner, G.
Deposited on	:	2003-11-07

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment was not calculated.

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	Percentile Ranks	Value
Clashscore		130
Ramachandran outliers		15.2%
Sidechain outliers		42.4%

Worse   *Better*

∎ Percentile relative to all structures

▯ Percentile relative to all NMR structures

Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and 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%

Mol	Chain	Length	Quality of chain
1	A	213	• 41% 37% • 17%
2	B	100	8% 37% 28% • 25%

2   Ensemble composition and analysis ⓘ

This entry contains 11 models. Model 2 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative.

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	**Residue range (total)**	**Backbone RMSD (Å)**	**Medoid model**
1	A:22-A:50, A:56-A:188, A:193-A:202, A:210-A:213, B:234-B:308 (251)	1.33	2

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 2 clusters. No single-model clusters were found.

Cluster number	Models
1	1, 4, 5, 9, 10, 11
2	2, 3, 6, 7, 8

3 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 5073 atoms, of which 2506 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Eukaryotic translation initiation factor 4E.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	213	3371	1089	1663	285	329	5	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	120	CYS	ARG	conflict	UNP P07260
A	132	CYS	ARG	conflict	UNP P07260
A	169	CYS	SER	conflict	UNP P07260
A	200	CYS	SER	conflict	UNP P07260

- Molecule 2 is a protein called Eukaryotic initiation factor 4F subunit p150.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	B	100	1550	498	768	123	160	1	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	215	GLY	-	cloning artifact	UNP P39935
B	216	SER	-	cloning artifact	UNP P39935

- Molecule 3 is S-[(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl] methanesulfonothioate (three-letter code: MTN) (formula: $C_{10}H_{18}NO_3S_2$).

Mol	Chain	Residues	Atoms					
3	A	1	Total	C	H	N	O	S
			27	9	15	1	1	1
3	A	1	Total	C	H	N	O	S
			27	9	15	1	1	1
3	A	1	Total	C	H	N	O	S
			27	9	15	1	1	1
3	A	1	Total	C	H	N	O	S
			27	9	15	1	1	1

- Molecule 4 is 7N-METHYL-8-HYDROGUANOSINE-5'-DIPHOSPHATE (three-letter code: M7G) (formula: $C_{11}H_{18}N_5O_{11}P_2$).
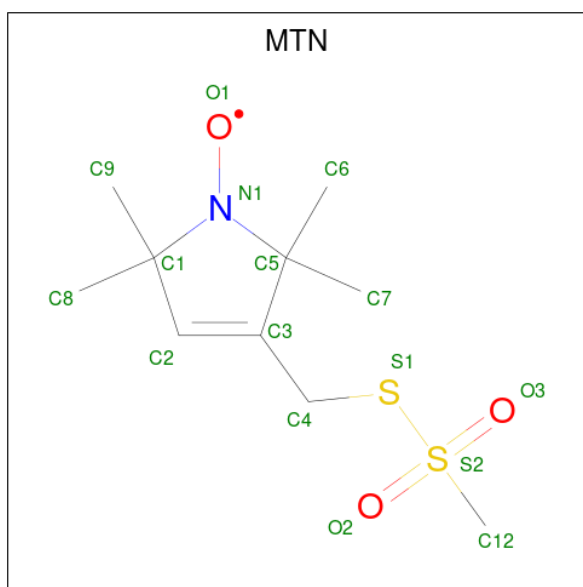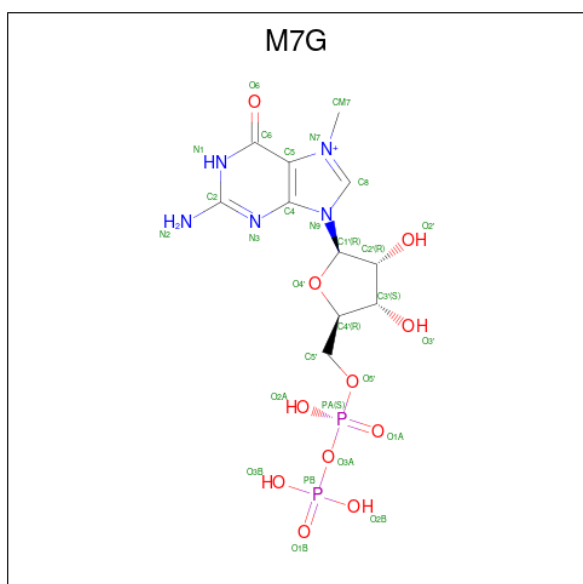
Mol	Chain	Residues	Atoms					
			Total	C	H	N	O	P
4	A	1	44	11	15	5	11	2

4 Residue-property plots ⓘ

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-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 outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

- Molecule 1: Eukaryotic translation initiation factor 4E

Chain A:

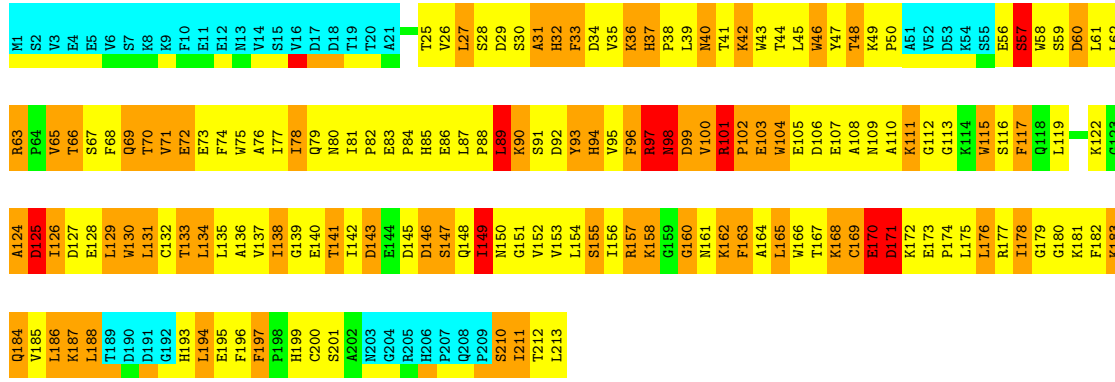- Molecule 2: Eukaryotic initiation factor 4F subunit p150

Chain B:
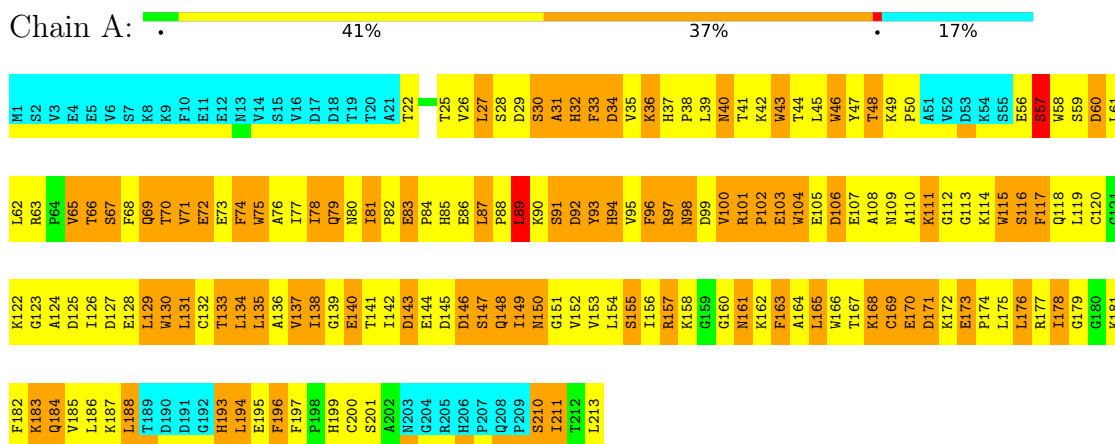
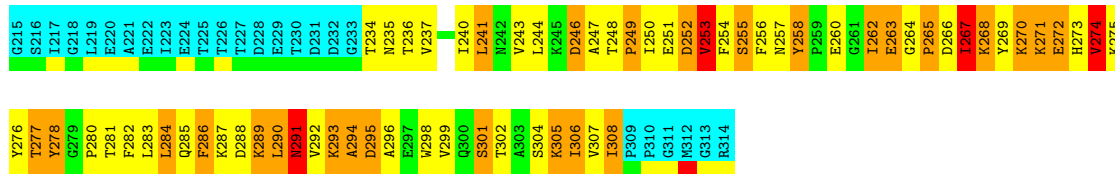## 4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 2. Colouring as in section 4.1 above.

- Molecule 1: Eukaryotic translation initiation factor 4E

Chain A:

M1 S2 V3 E4 E5 V6 S7 K8 K9 F10 E11 E12 N13 V14 S15 V16 D17 D18 T19 T20 A21 T25 V26 L27 S28 D29 S30 A31 H32 F33 D34 V35 K36 H37 P38 L39 N40 T41 K42 W43 T44 L45 W46 Y47 T48 K49 P50 A51 V52 D53 K54 S55 E56 S57 W58 S59 D60 L61 L62

R63 P64 V65 T66 S67 F68 Q69 T70 V71 E72 E73 F74 W75 A76 I77 I78 Q79 N80 I81 P82 E83 P84 H85 E86 L87 P88 L89 K90 S91 D92 Y93 H94 V95 F96 R97 N98 D99 V100 R101 P102 E103 W104 E105 D106 E107 A108 N109 A110 K111 G112 G113 K114 W115 S116 F117 Q118 L119 K122 G123

A124 D125 I126 D127 E128 L129 W130 L131 C132 T133 L134 L135 A136 V137 I138 G139 E140 T141 I142 D143 E144 D145 D146 S147 Q148 I149 N150 G151 V152 V153 L154 S155 I156 R157 K158 G159 G160 N161 K162 F163 A164 L165 W166 T167 K168 C169 E170 D171 K172 E173 P174 L175 L176 R177 I178 G179 G180 K181 F182 K183

Q184 V185 L186 K187 L188 T189 D190 D191 G192 H193 L194 E195 F196 F197 P198 H199 C200 S201 A202 N203 G204 R205 H206 P207 Q208 P209 S210 I211 T212 L213

• Molecule 2: Eukaryotic initiation factor 4F subunit p150

Chain B:　　10%　　　35%　　　26%　　•　　25%

G215 S216 I217 G218 L219 E220 A221 E222 I223 E224 T225 T226 T227 D228 E229 T230 D231 D232 G233 T234 N235 T236 V237 I240 L241 N242 V243 L244 K245 D246 A247 T248 P249 I250 E251 D252 V253 F254 S255 F256 N257 Y258 P259 E260 G261 I262 E263 G264 P265 D266 I267 K268 Y269 K270 K271 E272 H273 V274 K275

Y276 T277 Y278 G279 P280 T281 F282 L283 L284 Q285 F286 K287 D288 K289 L290 R291 V292 K293 A294 D295 A296 E297 W298 V299 Q300 S301 T302 A303 S304 K305 I306 V307 I308 P309 P310 G311 M312 G313 R314

5 Refinement protocol and experimental data overview ⓘ

Of the 50 calculated structures, 11 were deposited, based on the following criterion: *structures with the least restraint violations.*

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
X-PLOR	structure solution	3.851
X-PLOR	refinement	3.851

No chemical shift data was provided.

6 Model quality ⓘ

6.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: MTN, M7G

There are no covalent bond-length or bond-angle outliers.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

6.2 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 each 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 averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	1428	1403	1397	438±19
2	B	605	608	606	188±21
3	A	48	60	60	2±2
4	A	29	15	16	16±4
All	All	23210	22946	22869	5990

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 130.

5 of 2804 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models Worst	Total
2:B:250:ILE:HD11	2:B:256:PHE:CE1	1.14	1.77	8	8
1:A:135:LEU:HD21	2:B:283:LEU:HD13	1.13	1.15	8	3
1:A:35:VAL:CG1	2:B:250:ILE:HD12	1.07	1.79	7	7
2:B:250:ILE:HD11	2:B:256:PHE:CZ	1.07	1.84	7	11
1:A:149:ILE:HD11	1:A:167:THR:HG22	1.07	1.10	11	5

6.3 Torsion angles ⓘ

6.3.1 Protein backbone ⓘ

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 NMR 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	175/213 (82%)	110±4 (63±2%)	39±4 (22±2%)	25±3 (14±2%)	0	4
2	B	75/100 (75%)	50±2 (67±3%)	12±2 (16±3%)	13±3 (17±4%)	0	3
All	All	2750/3443 (80%)	1767 (64%)	564 (21%)	419 (15%)	0	4

5 of 93 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	89	LEU	11
1	A	100	VAL	11
1	A	169	CYS	11
2	B	252	ASP	11
2	B	253	VAL	11

6.3.2 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 NMR 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	158/191 (83%)	87±4 (55±3%)	71±4 (45±3%)	0	2
2	B	68/87 (78%)	43±2 (63±3%)	25±2 (37±3%)	1	7
All	All	2486/3058 (81%)	1432 (58%)	1054 (42%)	0	3

5 of 194 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	33	PHE	11
1	A	48	THR	11

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	57	SER	11
1	A	59	SER	11
1	A	69	GLN	11

6.3.3 RNA ⓘ

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.6 Ligand geometry ⓘ

5 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length is the number of standard deviations the observed value is removed from the expected value. A bond length with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
3	MTN	A	321	1	9,12,16	1.10±0.01	1±0 (11±0%)
3	MTN	A	323	1	9,12,16	1.10±0.01	1±0 (11±0%)
3	MTN	A	322	1	9,12,16	1.10±0.01	1±0 (11±0%)
3	MTN	A	320	1	9,12,16	1.09±0.01	1±0 (11±0%)
4	M7G	A	325	-	26,31,31	1.04±0.02	2±0 (7±0%)

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of angles that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond angle is the number of standard

deviations the observed value is removed from the expected value. A bond angle with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

| Mol | Type | Chain | Res | Link | Bond angles | | |
					Counts	RMSZ	#Z>2
3	MTN	A	321	1	9,20,27	1.15±0.01	2±0 (22±0%)
3	MTN	A	323	1	9,20,27	1.14±0.01	2±0 (22±0%)
3	MTN	A	322	1	9,20,27	1.14±0.02	2±0 (22±0%)
3	MTN	A	320	1	9,20,27	1.15±0.01	2±0 (22±0%)
4	M7G	A	325	-	28,49,49	1.06±0.03	2±0 (7±0%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	MTN	A	321	1	-	0±0,0,25,29	0±0,1,1,1
4	M7G	A	325	-	-	0±0,12,32,32	0±0,3,3,3
3	MTN	A	320	1	-	0±0,0,25,29	0±0,1,1,1
3	MTN	A	322	1	-	0±0,0,25,29	0±0,1,1,1
3	MTN	A	323	1	-	0±0,0,25,29	0±0,1,1,1

5 of 6 unique bond outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) | Models | |
								Worst	Total
3	A	323	MTN	O1-N1	3.05	1.27	1.43	2	11
3	A	321	MTN	O1-N1	3.04	1.27	1.43	10	11
3	A	320	MTN	O1-N1	3.03	1.27	1.43	10	11
3	A	322	MTN	O1-N1	3.03	1.27	1.43	1	11
4	A	325	M7G	O4'-C1'	2.84	1.45	1.41	11	11

5 of 10 unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

| Mol | Chain | Res | Type | Atoms | Z | Observed(^{o}) | Ideal(^{o}) | Models | |
								Worst	Total
4	A	325	M7G	C3'-C2'-C1'	3.75	106.63	100.98	10	11
3	A	321	MTN	C9-C1-N1	2.46	113.14	110.00	11	11
3	A	320	MTN	C9-C1-N1	2.39	113.06	110.00	2	11
3	A	322	MTN	C9-C1-N1	2.38	113.05	110.00	3	11

Continued from previous page...

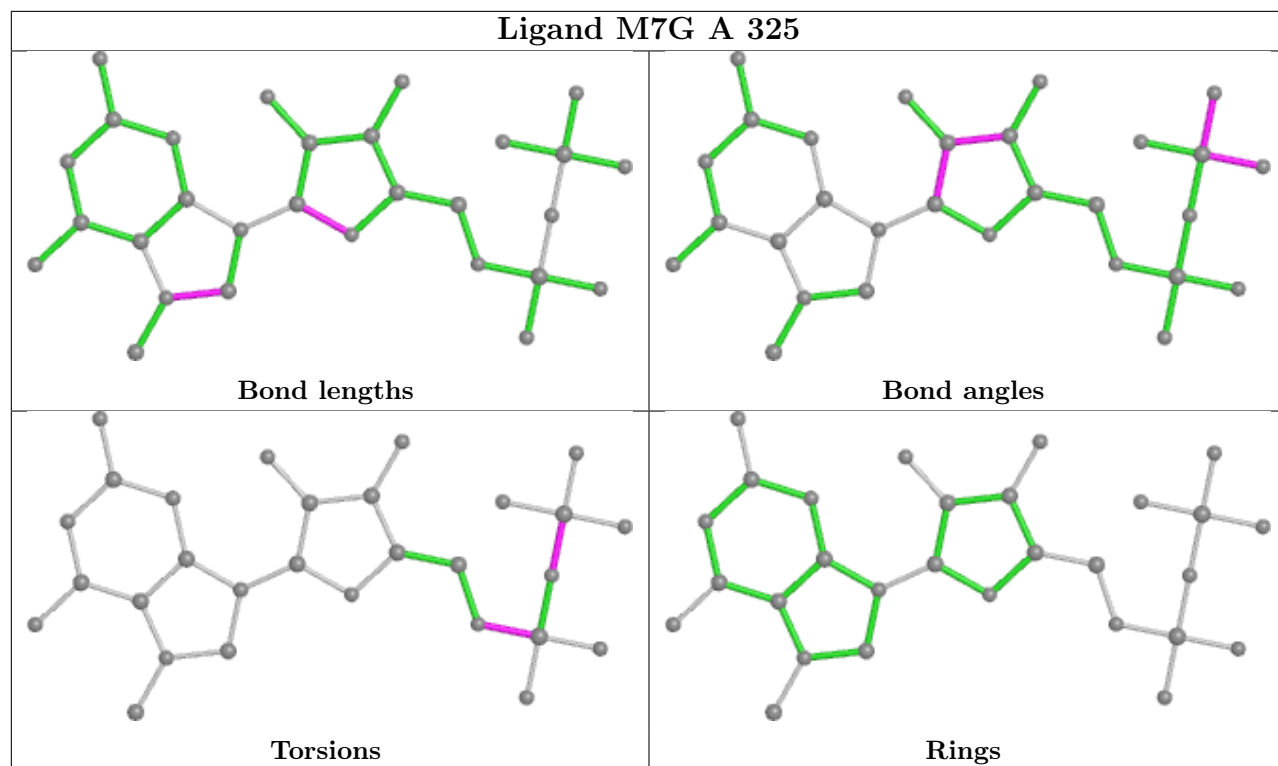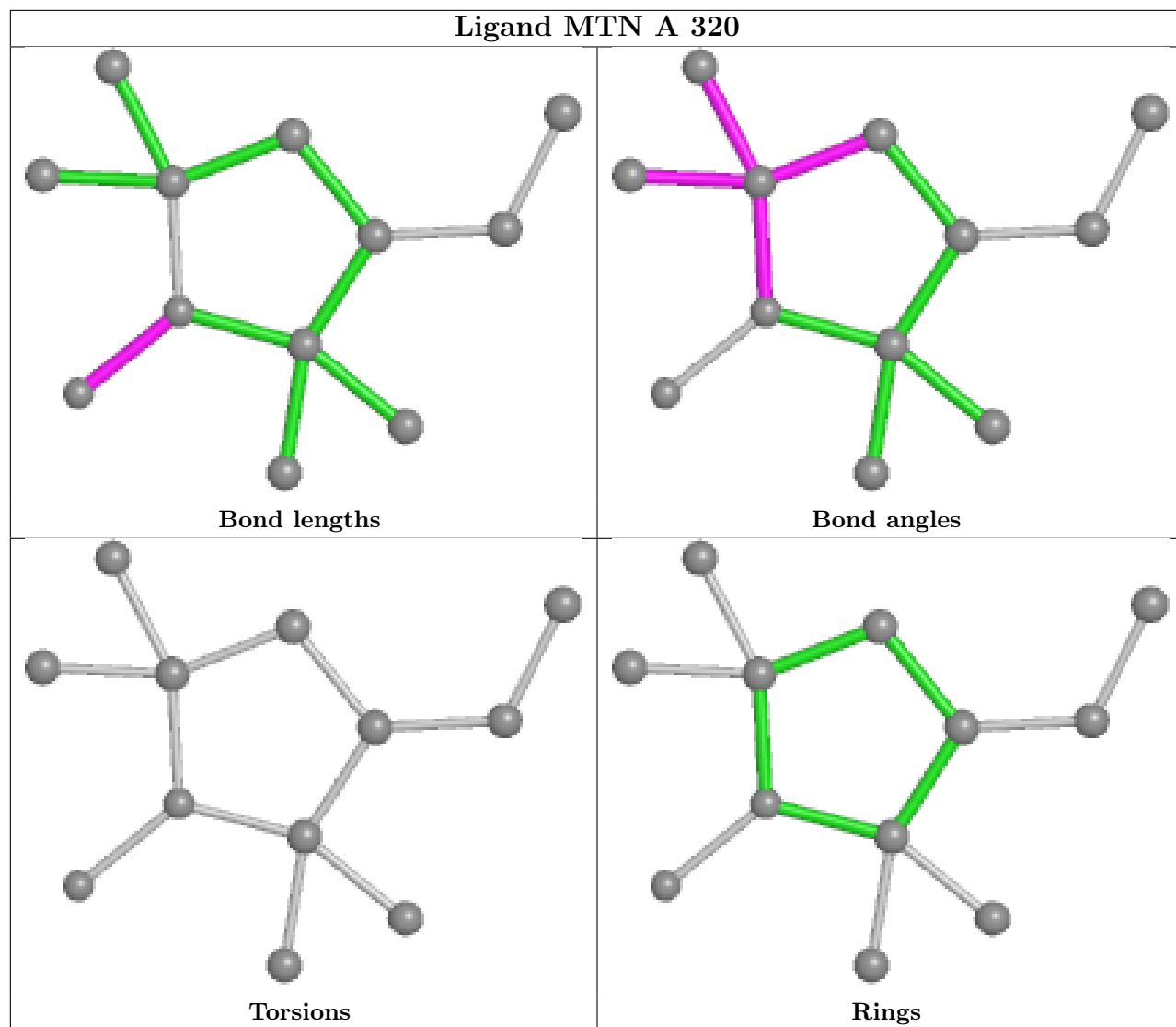| Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
3	A	323	MTN	C9-C1-N1	2.38	113.05	110.00	11	11

There are no chirality outliers.
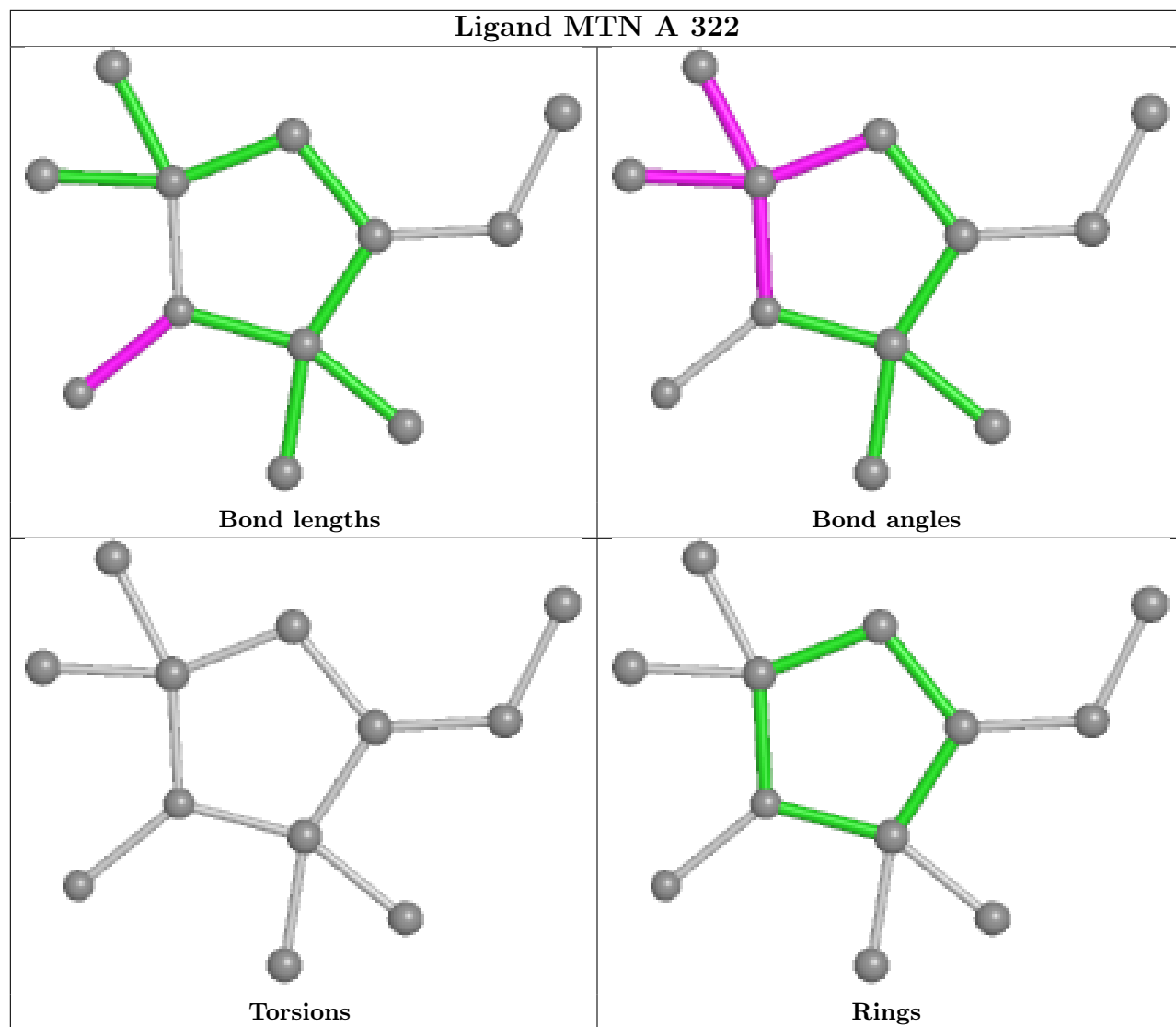
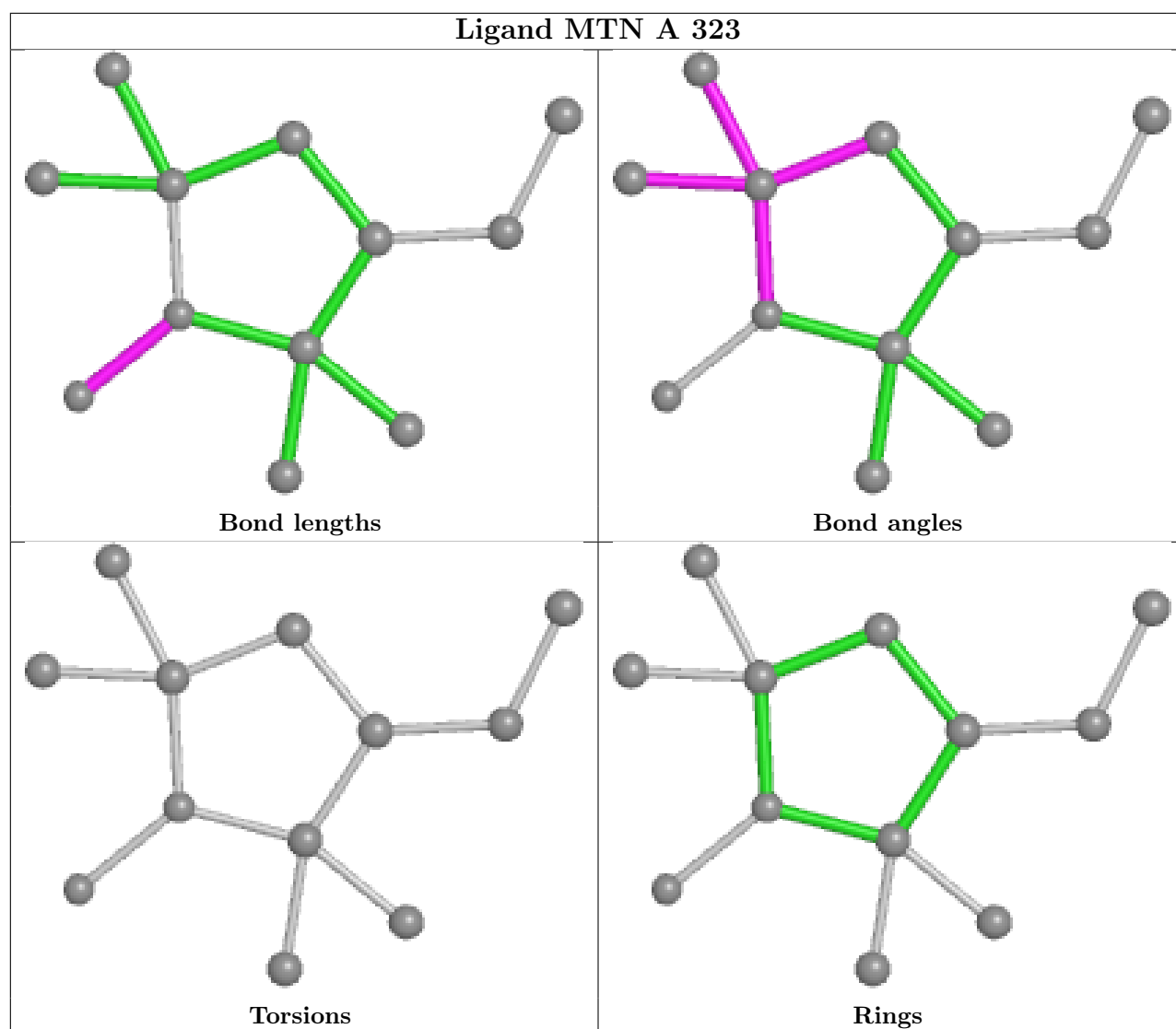There are no torsion outliers.

There are no ring outliers.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand MTN A 321

Bond lengths

Bond angles

Torsions

Rings

Ligand M7G A 325

Bond lengths

Bond angles

Torsions

Rings

Ligand MTN A 320

Bond lengths

Bond angles

Torsions

Rings

Ligand MTN A 322

Bond lengths

Bond angles

Torsions

Rings

Ligand MTN A 323

Bond lengths

Bond angles

Torsions

Rings
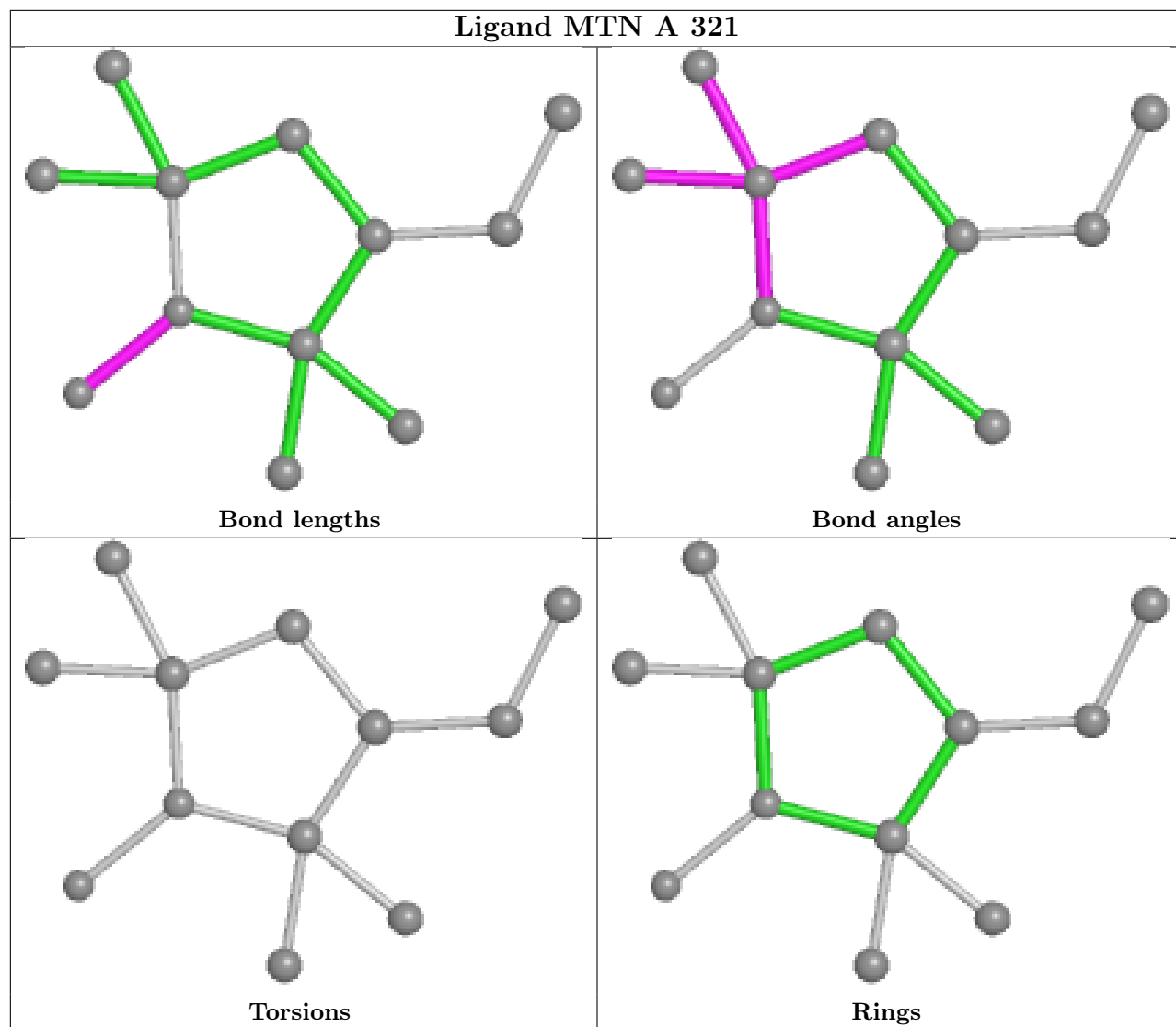
6.7 Other polymers ⓘ

There are no such molecules in this entry.

6.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

7 Chemical shift validation ⓘ

No chemical shift data were provided