WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB NMR Structure Validation Summary Report (i)

Oct 2, 2021 – 08:15 PM EDT

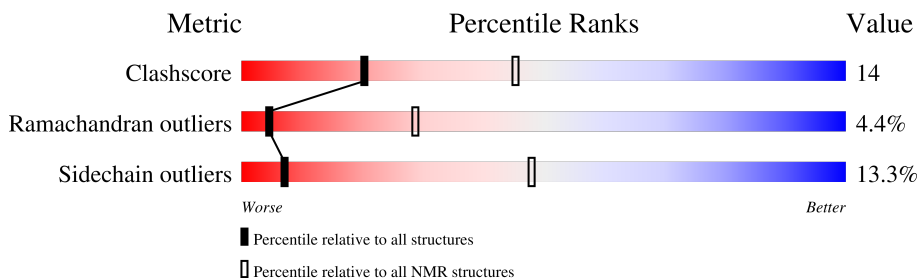|		
PDB ID	:	2KU2
Title	:	Dynamic Regulation of Archaeal Proteasome Gate Opening as Studied by TROSY-NMR
Authors	:	Religa, T.L.; Sprangers, R.; Kay, L.E.
Deposited on	:	2010-02-11

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		14
Ramachandran outliers		4.4%
Sidechain outliers		13.3%

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	237	55% 32% 5% · 6%
1	B	237	59% 27% 7% 6%
1	C	237	61% 27% 6% 6%
1	D	237	60% 26% 7% 6%
1	E	237	55% 32% 7% 6%
1	F	237	59% 30% 5% · 6%
1	G	237	58% 32% · 6%

2 Ensemble composition and analysis (i)

This entry contains 50 models. Model 10 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

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:12-A:233, B:12-B:233, C:12-C:233, D:12-D:233, E:12-E:233, F:12-F:233, G:12-G:233 (1554)	0.04	10

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 9 clusters and 1 single-model cluster was found.

Cluster number	Models
1	5, 6, 9, 10, 15, 16, 17, 19, 21, 26, 37, 39, 45, 47
2	2, 13, 29, 41, 42, 44, 48, 49, 50
3	14, 18, 32, 36, 38
4	7, 8, 25, 34, 43
5	11, 22, 23, 24, 40
6	27, 28, 31, 33
7	1, 3, 4
8	12, 20
9	35, 46
Single-model clusters	30

3 Entry composition ⓘ

There is only 1 type of molecule in this entry. The entry contains 25802 atoms, of which 13027 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Proteasome subunit alpha.

Mol	Chain	Residues	Atoms						Trace
1	A	237	Total	C	H	N	O	S	0
			3686	1153	1861	312	354	6	
1	B	237	Total	C	H	N	O	S	0
			3686	1153	1861	312	354	6	
1	C	237	Total	C	H	N	O	S	0
			3686	1153	1861	312	354	6	
1	D	237	Total	C	H	N	O	S	0
			3686	1153	1861	312	354	6	
1	E	237	Total	C	H	N	O	S	0
			3686	1153	1861	312	354	6	
1	F	237	Total	C	H	N	O	S	0
			3686	1153	1861	312	354	6	
1	G	237	Total	C	H	N	O	S	0
			3686	1153	1861	312	354	6	

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	expression tag	UNP P25156
A	-2	ALA	-	expression tag	UNP P25156
A	-1	MET	-	expression tag	UNP P25156
A	0	GLY	-	expression tag	UNP P25156
A	8	GLY	TYR	engineered mutation	UNP P25156
A	9	GLY	ASP	engineered mutation	UNP P25156
B	-3	GLY	-	expression tag	UNP P25156
B	-2	ALA	-	expression tag	UNP P25156
B	-1	MET	-	expression tag	UNP P25156
B	0	GLY	-	expression tag	UNP P25156
B	8	GLY	TYR	engineered mutation	UNP P25156
B	9	GLY	ASP	engineered mutation	UNP P25156
C	-3	GLY	-	expression tag	UNP P25156
C	-2	ALA	-	expression tag	UNP P25156
C	-1	MET	-	expression tag	UNP P25156
C	0	GLY	-	expression tag	UNP P25156
C	8	GLY	TYR	engineered mutation	UNP P25156
C	9	GLY	ASP	engineered mutation	UNP P25156
D	-3	GLY	-	expression tag	UNP P25156

Continued on next page...
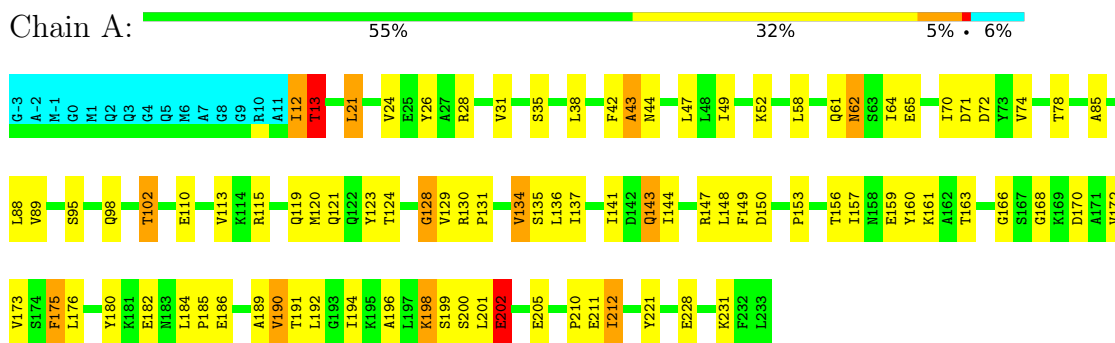
Continued from previous page...

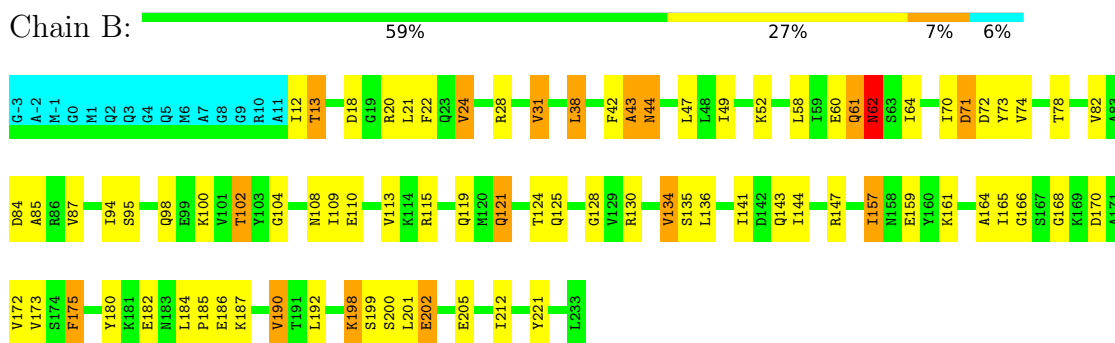| Chain	Residue	Modelled	Actual	Comment	Reference
D	-2	ALA	-	expression tag	UNP P25156
D	-1	MET	-	expression tag	UNP P25156
D	0	GLY	-	expression tag	UNP P25156
D	8	GLY	TYR	engineered mutation	UNP P25156
D	9	GLY	ASP	engineered mutation	UNP P25156
E	-3	GLY	-	expression tag	UNP P25156
E	-2	ALA	-	expression tag	UNP P25156
E	-1	MET	-	expression tag	UNP P25156
E	0	GLY	-	expression tag	UNP P25156
E	8	GLY	TYR	engineered mutation	UNP P25156
E	9	GLY	ASP	engineered mutation	UNP P25156
F	-3	GLY	-	expression tag	UNP P25156
F	-2	ALA	-	expression tag	UNP P25156
F	-1	MET	-	expression tag	UNP P25156
F	0	GLY	-	expression tag	UNP P25156
F	8	GLY	TYR	engineered mutation	UNP P25156
F	9	GLY	ASP	engineered mutation	UNP P25156
G	-3	GLY	-	expression tag	UNP P25156
G	-2	ALA	-	expression tag	UNP P25156
G	-1	MET	-	expression tag	UNP P25156
G	0	GLY	-	expression tag	UNP P25156
G	8	GLY	TYR	engineered mutation	UNP P25156
G	9	GLY	ASP	engineered mutation	UNP P25156

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.
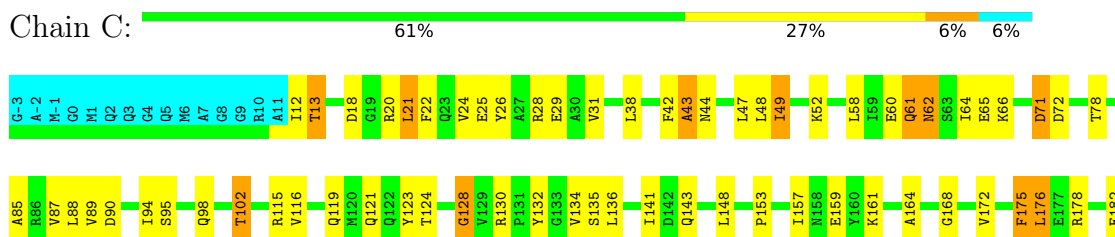
• Molecule 1: Proteasome subunit alpha
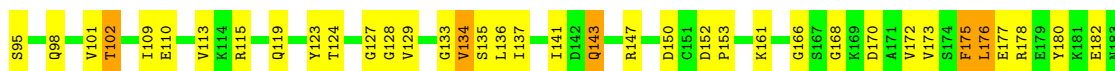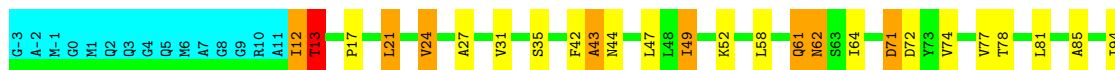
Chain A:

• Molecule 1: Proteasome subunit alpha

Chain B:

• Molecule 1: Proteasome subunit alpha

Chain C:

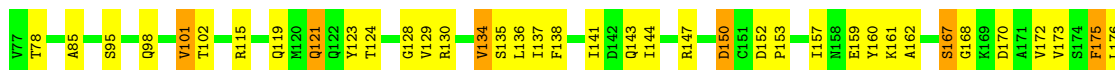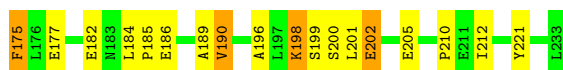E186 V190 K198 S199 S200 L201 E202 E205 P210 E211 I212 Y221 V229 K230 L233

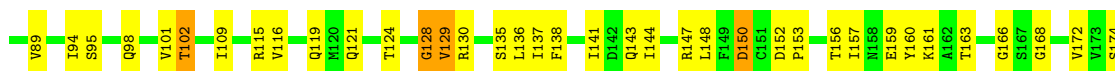● Molecule 1: Proteasome subunit alpha

Chain D:　　　　　　60%　　　　　　　　26%　　　　7%　　6%

G-3 A-2 M-1 G0 M1 Q2 Q3 G4 Q5 M6 A7 G8 G9 R10 A11 I12 T13 P17 L21 V24 A27 V31 S35 F42 A43 N44 L47 L48 I49 K52 L58 Q61 N62 S63 I64 D71 D72 Y73 V74 V77 T78 L81 A85 I94

S95 Q98 V101 T102 I109 E110 V113 K114 R115 Q119 Y123 T124 G127 G128 V129 G133 V134 S135 L136 I137 I141 D142 Q143 R147 D150 C151 D152 P153 K161 G166 S167 G168 K169 D170 A171 V172 V173 S174 F175 L176 E177 R178 E179 Y180 K181 E182 N183

L184 A189 V190 T191 L192 G193 I194 K195 A196 L197 K198 S199 S200 L201 E202 E205 E211 I212 Y221 L233

● Molecule 1: Proteasome subunit alpha

Chain E:　　　　　55%　　　　　　　　32%　　　　7%　　6%

G-3 A-2 M-1 G0 M1 Q2 Q3 G4 Q5 M6 A7 G8 G9 R10 A11 I12 T13 L21 V24 E25 Y26 A27 R28 V31 K32 K33 G34 L38 G39 M40 K41 F42 A43 N44 L47 L48 I49 K52 L58 Q61 M62 S63 I64 L69 I70 D71 D72 Y73 V74 A75 A76

V77 T78 A85 S95 Q98 V101 T102 R115 Q119 M120 Q121 Q122 Y123 T124 G128 V129 R130 V134 S135 L136 I137 F138 I141 D142 Q143 I144 R147 D150 C151 D152 P153 I157 N158 E159 Y160 K161 A162 S167 G168 K169 D170 A171 V172 V173 S174 F175 L176

Y180 K181 E182 N183 L184 P185 E186 K187 E188 A189 V190 T191 L192 G193 I194 K195 A196 L197 K198 S199 S200 L201 E202 E205 E206 L207 K208 A209 P210 E211 Y221 E228 K231 F232 L233

● Molecule 1: Proteasome subunit alpha

Chain F:　　　　　59%　　　　　　　　30%　　　5% • 6%

G-3 A-2 M-1 G0 M1 Q2 Q3 G4 Q5 M6 A7 G8 G9 R10 A11 I12 T13 P17 D18 G19 R20 L21 V24 E25 V31 S35 L38 F42 A43 N44 L47 L48 I49 K52 L58 I59 E60 Q61 N62 S63 I64 I70 D71 D72 Y73 V74 T78 A85

V89 I94 S95 Q98 V101 T102 I109 R115 V116 Q119 M120 Q121 T124 G128 V129 R130 S135 L136 I137 F138 I141 D142 Q143 I144 R147 L148 F149 D150 C151 D152 P153 T156 I157 N158 E159 Y160 K161 A162 T163 G166 S167 G168 V172 V173 S174

F175 L176 E177 E182 N183 L184 P185 E186 A189 V190 A196 L197 K198 S199 S200 L201 E202 E205 P210 E211 I212 Y221 L233

● Molecule 1: Proteasome subunit alpha

Chain G:　　　　　58%　　　　　　　　32%　　　•　　6%

G-3 A-2 M-1 G0 M1 Q2 Q3 G4 Q5 M6 A7 G8 G9 R10 A11 I12 T13 S16 P17 D18 G19 R20 L21 F22 Q23 V24 E25 Y26 A27 R28 V31 F42 A43 N44 L47 L48 I49 K52 Q61 N62 S63 I64 E65 K66 D71 A76 V77 T78 V82 A83 D84

A85 L88 S95 A96 Q97 Q98 T102 V107 K114 R115 V116 A117 D118 Q119 Y123 T124 G128 V129 R130 P131 Y132 G133 V134 S135 L136 I137 F138 I141 D142 Q143 D150 C151 D152 P153 I157 Y160 K161 A162 T163 G168 K169 D170 A171 V172 V173 S174 F175 L176

4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 10. Colouring as in section 4.1 above.
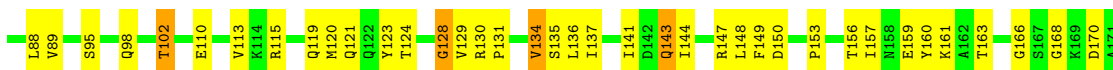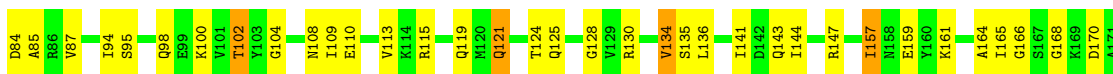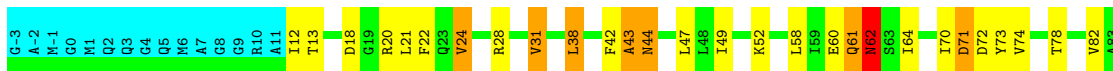
• Molecule 1: Proteasome subunit alpha

Chain A:

• Molecule 1: Proteasome subunit alpha

Chain B:

• Molecule 1: Proteasome subunit alpha

Chain C:

• Molecule 1: Proteasome subunit alpha

Chain D:

• Molecule 1: Proteasome subunit alpha
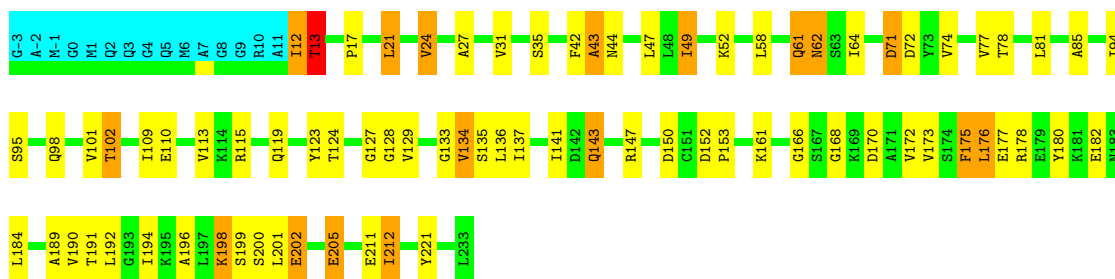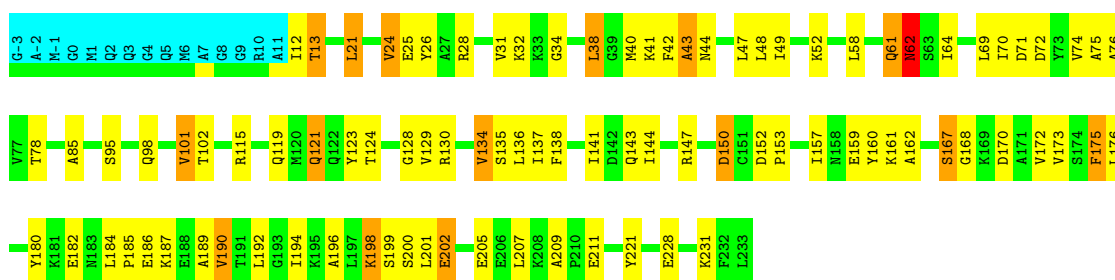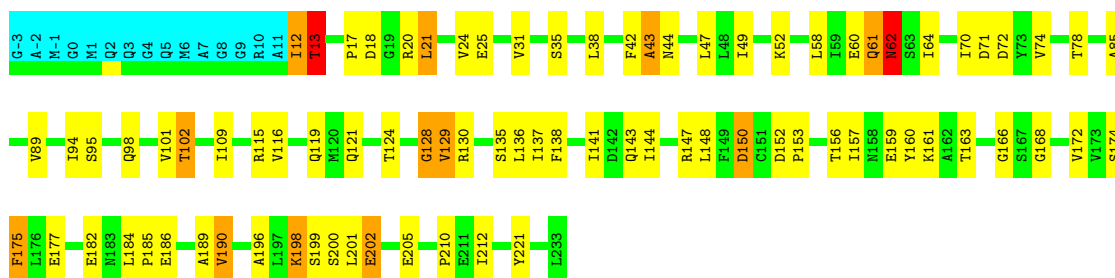
Chain E:

• Molecule 1: Proteasome subunit alpha

Chain F:

• Molecule 1: Proteasome subunit alpha

Chain G:
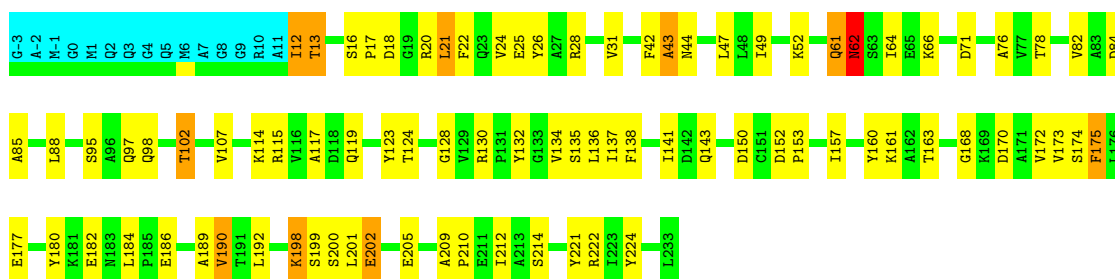
5 Refinement protocol and experimental data overview ⓘ

The models were refined using the following method: *simulated annealing.*

Of the 166 calculated structures, 50 were deposited, based on the following criterion: *structures with the lowest energy.*

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

Software name	Classification	Version
X-PLOR NIH	structure solution	2.23
X-PLOR NIH	refinement	2.23

No chemical shift data was provided.

6 Model quality ⓘ

6.1 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 (average) 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	0.53±0.00	0±0/1751 (0.0± 0.0%)	0.72±0.00	1±0/2359 (0.0± 0.0%)
1	B	0.51±0.00	0±0/1751 (0.0± 0.0%)	0.71±0.00	0±0/2359 (0.0± 0.0%)
1	C	0.53±0.00	0±0/1751 (0.0± 0.0%)	0.70±0.00	1±0/2359 (0.0± 0.0%)
1	D	0.55±0.00	0±0/1751 (0.0± 0.0%)	0.72±0.00	0±0/2359 (0.0± 0.0%)
1	E	0.53±0.00	0±0/1751 (0.0± 0.0%)	0.71±0.00	0±0/2359 (0.0± 0.0%)
1	F	0.54±0.00	0±0/1751 (0.0± 0.0%)	0.69±0.00	0±0/2359 (0.0± 0.0%)
1	G	0.53±0.00	0±0/1751 (0.0± 0.0%)	0.71±0.00	0±0/2359 (0.0± 0.0%)
All	All	0.53	0/612850 (0.0%)	0.71	100/825650 (0.0%)

There are no bond-length outliers.

All 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(°)	Ideal(°)	Models	
								Worst	Total
1	C	128	GLY	N-CA-C	5.36	126.49	113.10	1	50
1	A	128	GLY	N-CA-C	5.07	125.78	113.10	1	50

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	1728	1765	1765	59±1
1	B	1728	1765	1765	50±2

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	C	1728	1765	1765	42±1
1	D	1728	1765	1765	46±1
1	E	1728	1765	1765	56±1
1	F	1728	1765	1765	55±1
1	G	1728	1765	1765	49±1
All	All	604800	617750	617750	16650

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 14.

5 of 466 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models Worst	Total
1:A:12:ILE:HD12	1:A:12:ILE:N	0.77	1.94	4	2
1:F:12:ILE:HD13	1:F:12:ILE:H	0.76	1.39	21	1
1:A:52:LYS:NZ	1:A:62:ASN:HA	0.76	1.96	1	50
1:C:12:ILE:HD13	1:C:12:ILE:H	0.74	1.41	50	2
1:A:12:ILE:HD12	1:A:12:ILE:O	0.73	1.83	6	10

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	221/237 (93%)	186±1 (84±0%)	25±1 (11±0%)	10±1 (5±0%)	4	27
1	B	221/237 (93%)	187±1 (84±0%)	25±1 (11±0%)	10±1 (4±0%)	5	29
1	C	221/237 (93%)	185±1 (84±0%)	28±1 (13±0%)	8±1 (3±0%)	6	36
1	D	221/237 (93%)	189±1 (85±0%)	22±1 (10±0%)	10±1 (4±0%)	4	28
1	E	221/237 (93%)	184±1 (83±0%)	26±1 (12±0%)	12±1 (5±0%)	4	24
1	F	221/237 (93%)	187±1 (85±0%)	23±0 (10±0%)	11±1 (5±0%)	4	26
1	G	221/237 (93%)	188±1 (85±0%)	24±1 (11±0%)	9±0 (4±0%)	5	32
All	All	77350/82950 (93%)	65258 (84%)	8702 (11%)	3390 (4%)	4	29

5 of 77 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	43	ALA	50
1	A	72	ASP	50
1	A	128	GLY	50
1	A	129	VAL	50
1	A	182	GLU	50

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	185/192 (96%)	163±0 (88±0%)	22±0 (12±0%)	8	51
1	B	185/192 (96%)	160±0 (86±0%)	25±0 (14±0%)	7	47
1	C	185/192 (96%)	159±0 (86±0%)	26±0 (14±0%)	6	46
1	D	185/192 (96%)	158±1 (86±0%)	27±1 (14±0%)	6	45
1	E	185/192 (96%)	158±0 (85±0%)	27±0 (15±0%)	6	44
1	F	185/192 (96%)	162±0 (87±0%)	23±0 (13±0%)	8	49
1	G	185/192 (96%)	163±0 (88±0%)	22±0 (12±0%)	9	52
All	All	64750/67200 (96%)	56115 (87%)	8635 (13%)	7	48

5 of 178 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	13	THR	50
1	A	21	LEU	50
1	A	38	LEU	50
1	A	44	ASN	50
1	A	47	LEU	50

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 ⓘ

There are no ligands in this entry.

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 (i)

No chemical shift data were provided