



wwPDB EM Map/Model Validation Report ⓘ

Apr 10, 2016 – 02:05 PM BST

PDB ID : 3JC7
EMDB ID: : EMD-6536
Title : Structure of the eukaryotic replicative CMG helicase and pumpjack motion
Authors : Li, H.; Bai, L.; Yuan, Z.; Sun, J.; Georgescu, R.E.; Liu, J.; O' Donnell, M.E.
Deposited on : 2015-11-24
Resolution : 4.80 Å(reported)
Based on PDB ID : 2Q9Q

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could stem from errors in the original structure(s) used in the fitting.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
<http://wwpdb.org/validation/2016/EMValidationReportHelp>

MolProbity : 4.02b-467
Mogul : unknown
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : trunk27241

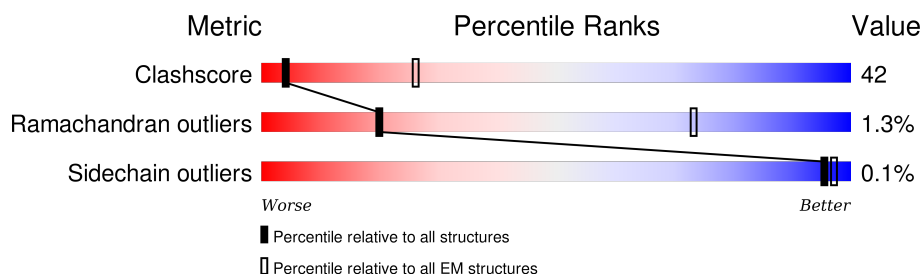
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.80 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)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	2	868	36% 29% . 34%
2	3	971	28% 31% . 41%
3	4	933	32% 33% . 33%
4	5	775	39% 42% . 16%
5	6	1017	32% 32% . 34%
6	7	845	36% 40% . 22%
7	c	650	84% . 15%
8	D	294	38% 36% . 25%
9	B	213	41% 44% 15%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
10	A	208	<div><div></div><div>39%</div><div>59%</div><div></div></div>
11	C	194	<div><div></div><div>41%</div><div>40%</div><div>18%</div><div></div></div>

2 Entry composition

There are 12 unique types of molecules in this entry. The entry contains 40298 atoms, of which 0 are hydrogens and 0 are deuteriums.

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

- Molecule 1 is a protein called DNA replication licensing factor MCM2.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	2	576	Total	C	N	O	S	0	0
			4531	2859	809	847	16		

- Molecule 2 is a protein called DNA replication licensing factor MCM3.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	3	577	Total	C	N	O	S	0	0
			4521	2857	800	851	13		

- Molecule 3 is a protein called DNA replication licensing factor MCM4.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	4	621	Total	C	N	O	S	0	0
			4911	3092	845	947	27		

- Molecule 4 is a protein called Minichromosome maintenance protein 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	5	654	Total	C	N	O	S	0	0
			5172	3250	897	1001	24		

- Molecule 5 is a protein called DNA replication licensing factor MCM6.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	6	668	Total	C	N	O	S	0	0
			5204	3290	913	978	23		

- Molecule 6 is a protein called DNA replication licensing factor MCM7.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	7	657	Total	C	N	O	S	0	0
			5176	3266	900	982	28		

- Molecule 7 is a protein called Cell division control protein 45.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	c	553	Total	C	N	O	S	0	0
			4470	2852	759	846	13		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
c	22	ALA	HIS	CONFLICT	UNP Q08032
c	155	GLU	GLN	CONFLICT	UNP Q08032
c	551	THR	TRP	CONFLICT	UNP Q08032

- Molecule 8 is a protein called DNA replication complex GINS protein SLD5.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	D	221	Total	C	N	O	S	0	0
			1820	1159	300	348	13		

- Molecule 9 is a protein called DNA replication complex GINS protein PSF2.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	B	181	Total	C	N	O	S	0	0
			1513	978	261	270	4		

- Molecule 10 is a protein called DNA replication complex GINS protein PSF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	A	208	Total	C	N	O	S	0	0
			1691	1062	287	332	10		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	161	ALA	VAL	CONFLICT	UNP Q12488
A	192	GLN	ARG	CONFLICT	UNP Q12488
A	207	LEU	LYS	CONFLICT	UNP Q12488

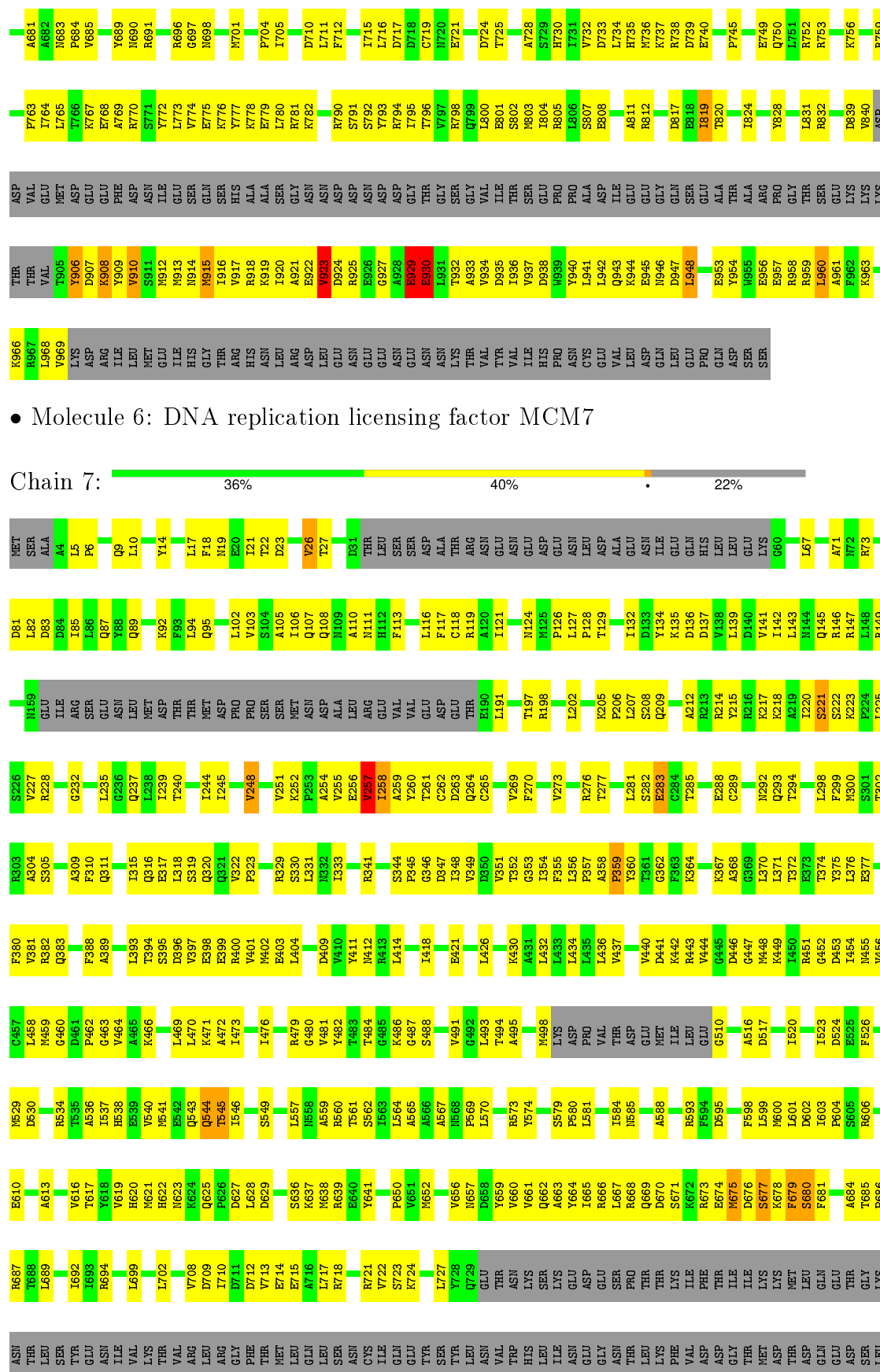
- Molecule 11 is a protein called DNA replication complex GINS protein PSF3.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	C	159	Total	C	N	O	S	0	0
			1288	843	207	232	6		

- Molecule 12 is ZINC ION (three-letter code: ZN) (formula: Zn).


Mol	Chain	Residues	Atoms		AltConf
12	7	1	Total	Zn	0
			1	1	





VAL
SER
THR
PRO
GLY
LYS
LEU
ALA
PRO
GLN
THR
THR
ALA
SER
ALA
ASN
VAL
SER
SER
ALA
GLN
ASP
SER
GLY
ILE
ASP
LEU
GLN
ASP
ALA

- Molecule 7: Cell division control protein 45

Chain c:  84% 15%

MET
THR
TYR
GLY
I5
L27
E97
I98
E102
THR
VAL
ILE
ASP
THR
THR
ASP
GLU
LYS
SER
GLY
Q114
L152
L165
ASP
GLU
GLU
SER
GLY
ASN
ASP
ASP
GLU
LEU
SER
GLY
THR
ASP
GLY
ASN
GLY
GLU
ASP
ASN
GLY
ASP
S462
S591
LEU
ASP
THR
ALA
ASP
ASP
HIS
T597
I601
GLU

L637
L650

- Molecule 8: DNA replication complex GINS protein SLD5

Chain D:  38% 36% 25%

MET
ASP
ILE
GLU
ASP
ASP
ASP
LEU
ASP
ASP
GLU
GLU
LEU
LEU
LYS
GLU
THR
THR
ALA
VAL
ASP
SER
LVS
THR
THR
GLN
GLY
SER
SER
SER
SER
THR
HIS
ARG
ASP
ALA
ASN
THR
VAL
GLY
SER
SER
SER
SER
SER
LEU
LEU
ASP
ASN
ASP
ASP
LVS
THR
GLN
TYR
V54
Q57
Q58
D59
E70

R71
C72
P74
P73
L77
P78
Y79
L83
R86
Y91
Q94
I98
I101
M110
SER
SER
ASN
ALA
ASN
PRO
PRO
MET
PRO
ASN
E121
L124
M129
E132
L133
R135
L136
K137
F138
R141
S142
Y143
I144
R145
C146
R147
L148
S149
K150
K153
F154
I230

R159
Q160
L161
M162
E163
D164
E165
N166
S167
L168
I169
S170
L171
T172
D173
L174
L175
S176
K177
D178
E179
I180
K181
Y182
S187
L188
I189
W190
L191
K192
L193
K200
M202
P203
L206
D211
T212
E213
G214
S215
V216
N217
M218
I219
D220
E221
D223
W224
N225
K226
F227
V228
F229
I230

R231
V232
M233
G234
G238
LYS
TRP
ASN
GLU
ASP
PRO
LEU
LEU
GLN
E248
E249
E250
P254
C255
V256
T257
V258
T259
I260
K264
E265
E266
V267
E268
L269
T270
I274
V275
V276
R277
V279
E280
V281
I282
R283
L286
V291
K292
L293
ILE

- Molecule 9: DNA replication complex GINS protein PSF2

Chain B:  41% 44% 15%

MET
SER
L3
L7
Q8
F11
S12
P13
E14
E15
L16
Q17
V20
E21
P24
I25
K26
I27
F28
P29
R30
T31
T32
THR
ARG
GLN
LYS
ILE
ARG
GLY
ASP
ASP
ARG
GLY
THR
GLY
ASN
HIS
THR
ARG
H50
Q51
L52
I53
T54
T55
D56
D57
K58
A59
L60
I61
R62
R63
V64
A65

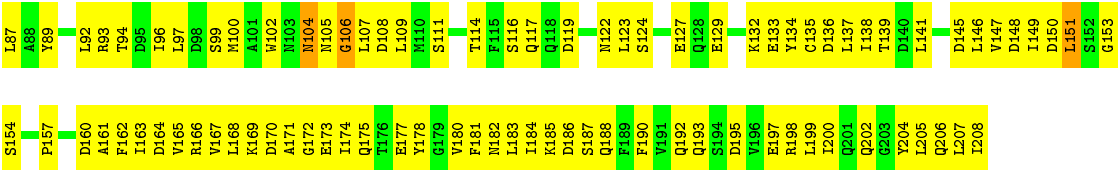
S68
T69
E70
V71
V72
L73
W74
I75
L79
K80
Q81
Q82
S83
K84
C85
S86
I87
V88
T94
L98
I102
F112
L115
P116
W117
M118
W119
L120
V121
L122
F127
A130
F134
P137
I138
H139
R142
L148
R152
K155
V156
L157
L160

L163
H164
E165
S166
H167
L168
Q169
L173
S174
L175
L176
E177
I178
M179
E180
L181
F184
I185
T186
E187
I188
M189
D190
K191
L192
R193
E194
I195
H196
T197
L200
THR
ALA
GLY
THR
GLU
ASN
ASP
GLU
GLU
PHE
ASN
ILE

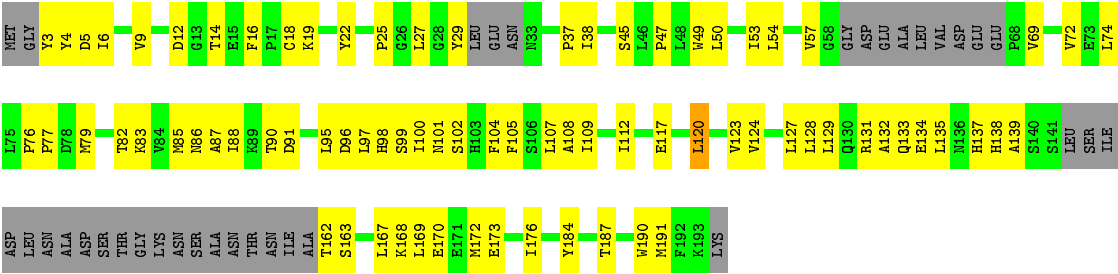
- Molecule 10: DNA replication complex GINS protein PSF1

Chain A:  39% 59%

H1
L5
G6
A13
K14
A15
L16
K17
Q18
L19
R22
S23
N24
Q25
D26
V27
N28
L29
P30
M31
Y32
I36
I37
L40
L41
V44
L47
L54
Q57
Q58
M62
D65
K66
V67
F73
V74
T75
L76
L77
C78
M79
E80
R81
R82
K83
R84
C85
L86



● Molecule 11: DNA replication complex GINS protein PSF3



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of images	86822	Depositor
Resolution determination method	FSC 0.143	Depositor
CTF correction method	CTFFIND4	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	50	Depositor
Minimum defocus (nm)	1500	Depositor
Maximum defocus (nm)	3500	Depositor
Magnification	29000	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z > 2$	RMSZ	$\# Z > 2$
1	2	0.41	0/4604	0.63	2/6215 (0.0%)
10	A	0.39	0/1713	0.66	0/2309
11	C	0.46	0/1320	0.64	1/1784 (0.1%)
2	3	0.50	1/4597 (0.0%)	0.67	1/6232 (0.0%)
3	4	0.41	0/4981	0.65	1/6734 (0.0%)
4	5	0.51	0/5243	0.83	12/7075 (0.2%)
5	6	0.48	1/5282 (0.0%)	0.79	13/7129 (0.2%)
6	7	0.42	0/5256	0.65	1/7099 (0.0%)
7	c	0.39	0/4548	0.63	2/6152 (0.0%)
8	D	0.43	0/1853	0.64	0/2500
9	B	0.44	0/1545	0.69	0/2092
All	All	0.44	2/40942 (0.0%)	0.69	33/55321 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	2	0	1
10	A	0	4
2	3	0	3
3	4	0	5
4	5	0	2
5	6	0	6
6	7	0	6
7	c	0	1
8	D	0	2
All	All	0	30

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	3	232	PRO	N-CD	16.15	1.70	1.47
5	6	929	GLU	C-O	7.77	1.38	1.23

The worst 5 of 33 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	6	954	TYR	CB-CG-CD2	-16.47	111.12	121.00
4	5	732	THR	CA-CB-CG2	-14.07	92.70	112.40
4	5	732	THR	OG1-CB-CG2	13.49	141.04	110.00
4	5	742	ARG	CA-CB-CG	10.90	137.38	113.40
5	6	954	TYR	CB-CG-CD1	9.46	126.68	121.00

There are no chirality outliers.

5 of 30 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	2	803	PHE	Peptide
2	3	163	ALA	Peptide
2	3	165	ALA	Peptide
2	3	428	LEU	Peptide
3	4	202	LYS	Peptide

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	2	4531	0	4575	356	0
2	3	4521	0	4581	520	0
3	4	4911	0	4947	440	0
4	5	5172	0	5224	650	0
5	6	5204	0	5166	449	0
6	7	5176	0	5245	502	0
7	c	4470	0	4491	0	0
8	D	1820	0	1823	200	0
9	B	1513	0	1558	148	0
10	A	1691	0	1686	312	0
11	C	1288	0	1298	132	0
12	7	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	40298	0	40594	3102	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 42.

The worst 5 of 3102 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)
2:3:502:ILE:HG21	6:7:244:ILE:CG1	1.24	1.67
2:3:484:VAL:HG21	6:7:486:LYS:CB	1.16	1.62
4:5:197:PHE:CE2	4:5:329:LYS:HG2	1.19	1.60
1:2:684:ARG:HB3	1:2:685:ASP:CB	1.11	1.58
2:3:502:ILE:CG2	6:7:244:ILE:HG12	1.12	1.55

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	2	564/868 (65%)	512 (91%)	45 (8%)	7 (1%)	16	62
2	3	563/971 (58%)	515 (92%)	41 (7%)	7 (1%)	16	62
3	4	609/933 (65%)	534 (88%)	65 (11%)	10 (2%)	12	57
4	5	638/775 (82%)	571 (90%)	49 (8%)	18 (3%)	6	45
5	6	656/1017 (64%)	594 (90%)	52 (8%)	10 (2%)	13	58
6	7	647/845 (77%)	572 (88%)	66 (10%)	9 (1%)	14	58
7	c	543/650 (84%)	496 (91%)	45 (8%)	2 (0%)	39	80
8	D	215/294 (73%)	197 (92%)	16 (7%)	2 (1%)	21	67
9	B	177/213 (83%)	163 (92%)	14 (8%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
10	A	206/208 (99%)	181 (88%)	24 (12%)	1 (0%)	34	77
11	C	151/194 (78%)	144 (95%)	7 (5%)	0	100	100
All	All	4969/6968 (71%)	4479 (90%)	424 (8%)	66 (1%)	20	60

5 of 66 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	2	291	SER
3	4	450	GLN
3	4	609	VAL
3	4	736	ILE
4	5	340	SER

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 PDB entries followed by that with respect to all EM entries.

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	2	495/770 (64%)	495 (100%)	0	100	100
2	3	502/835 (60%)	502 (100%)	0	100	100
3	4	554/848 (65%)	554 (100%)	0	100	100
4	5	589/688 (86%)	589 (100%)	0	100	100
5	6	548/886 (62%)	548 (100%)	0	100	100
6	7	578/753 (77%)	578 (100%)	0	100	100
7	c	498/585 (85%)	496 (100%)	2 (0%)	93	96
8	D	213/279 (76%)	212 (100%)	1 (0%)	92	96
9	B	171/198 (86%)	170 (99%)	1 (1%)	90	95
10	A	192/192 (100%)	191 (100%)	1 (0%)	92	96
11	C	144/173 (83%)	144 (100%)	0	100	100
All	All	4484/6207 (72%)	4479 (100%)	5 (0%)	95	97

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

Mol	Chain	Res	Type
7	c	27	LEU
7	c	152	LEU
8	D	168	LEU
9	B	175	LEU
10	A	151	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 64 such sidechains are listed below:

Mol	Chain	Res	Type
4	5	259	GLN
4	5	731	GLN
9	B	179	ASN
4	5	411	ASN
4	5	560	HIS

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](#)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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.