



wwPDB/EMDatabank EM Map/Model Validation Summary Report ⓘ

Feb 21, 2018 – 08:57 PM EST

PDB ID : 6B0X
EMDB ID: : EMD-7030
Title : Capsid protein and C-terminal part of scaffolding protein in the Staphylococcus aureus phage 80alpha procapsid
Authors : Kizziah, J.L.; Dearborn, A.D.; Dokland, T.
Deposited on : 2017-09-15
Resolution : 3.72 Å(reported)

This is a wwPDB/EMDatabank EM Map/Model Validation Summary Report
for a publicly released PDB/EMDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/EMValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

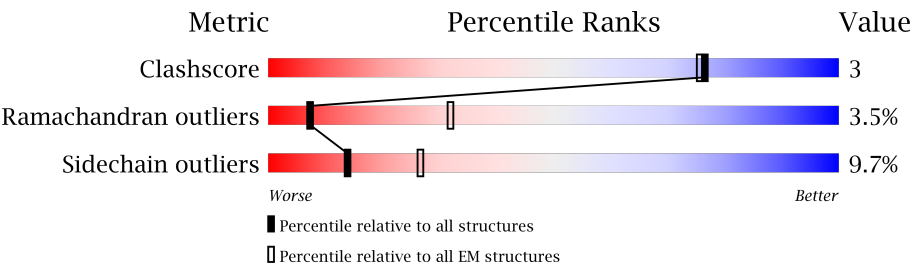
MolProbity : 4.02b-467
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20030736

1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 3.72 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	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026

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	A	324	<div><div>72%</div><div>15%</div><div>•</div><div>12%</div></div>
1	B	324	<div><div>73%</div><div>13%</div><div>•</div><div>12%</div></div>
1	C	324	<div><div>73%</div><div>14%</div><div>•</div><div>12%</div></div>
1	D	324	<div><div>68%</div><div>18%</div><div>•</div><div>12%</div></div>
1	E	324	<div><div>77%</div><div>9%</div><div>•</div><div>12%</div></div>
1	F	324	<div><div>73%</div><div>14%</div><div>•</div><div>12%</div></div>
1	G	324	<div><div>69%</div><div>18%</div><div>•</div><div>12%</div></div>
2	a	206	<div><div>7%</div><div>93%</div></div>
2	b	206	<div><div>7%</div><div>92%</div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	c	206	 7% 93%
2	d	206	 8% 92%
2	e	206	 8% 92%
2	f	206	 7% 92%
2	g	206	 7% 92%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 16791 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 Major head protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	284	Total	C	N	O	S	0	0
			2270	1446	370	443	11		
1	B	284	Total	C	N	O	S	0	0
			2270	1446	370	443	11		
1	C	284	Total	C	N	O	S	0	0
			2270	1446	370	443	11		
1	D	284	Total	C	N	O	S	0	0
			2270	1446	370	443	11		
1	E	284	Total	C	N	O	S	0	0
			2270	1446	370	443	11		
1	F	284	Total	C	N	O	S	0	0
			2270	1446	370	443	11		
1	G	284	Total	C	N	O	S	0	0
			2270	1446	370	443	11		

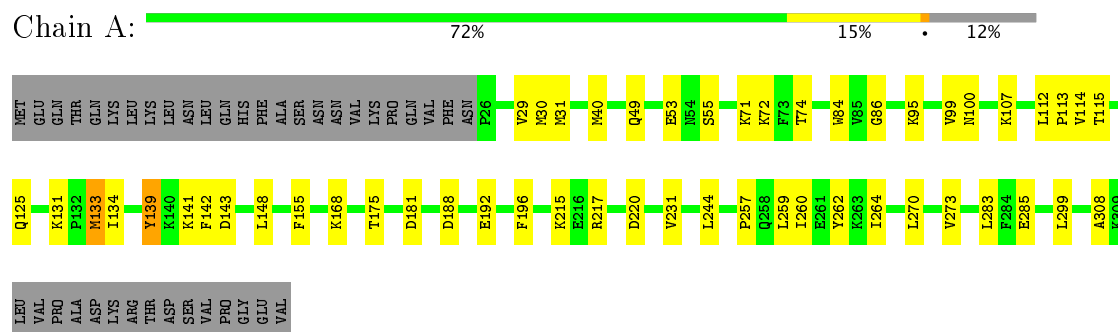
- Molecule 2 is a protein called Scaffold protein.

Mol	Chain	Residues	Atoms				AltConf	Trace
2	a	14	Total	C	N	O	0	0
			116	72	25	19		
2	b	16	Total	C	N	O	0	0
			132	82	28	22		
2	c	15	Total	C	N	O	0	0
			125	77	27	21		
2	d	16	Total	C	N	O	0	0
			132	82	28	22		
2	e	16	Total	C	N	O	0	0
			132	82	28	22		
2	f	16	Total	C	N	O	0	0
			132	82	28	22		
2	g	16	Total	C	N	O	0	0
			132	82	28	22		

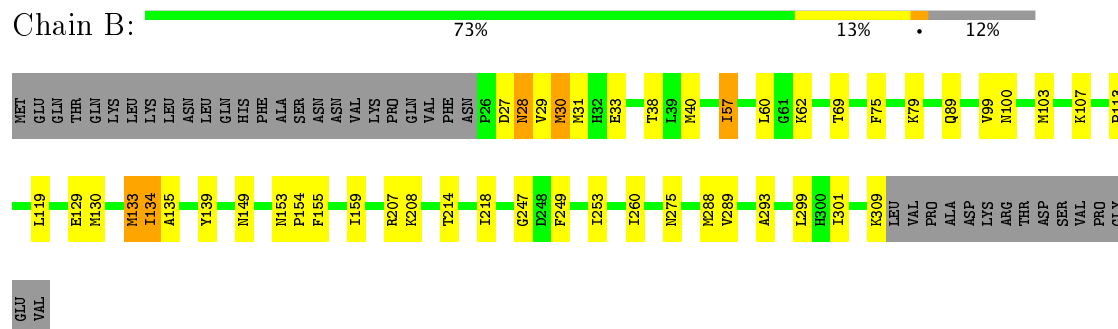
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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. 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.

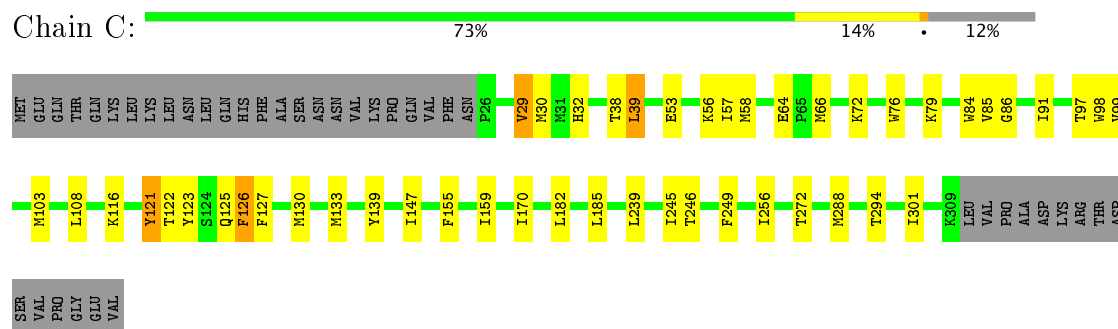
- Molecule 1: Major head protein



- Molecule 1: Major head protein

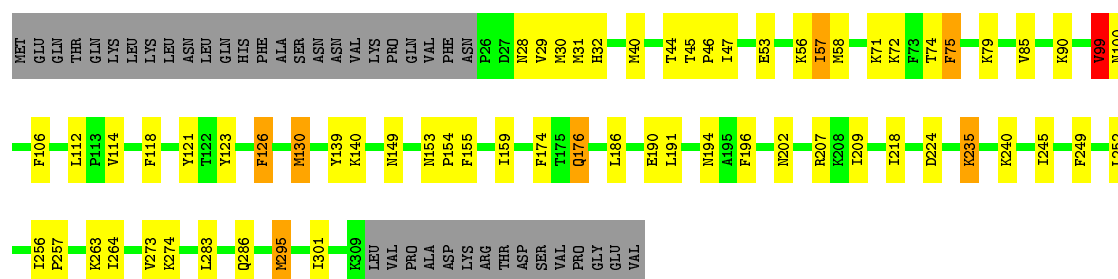


- Molecule 1: Major head protein




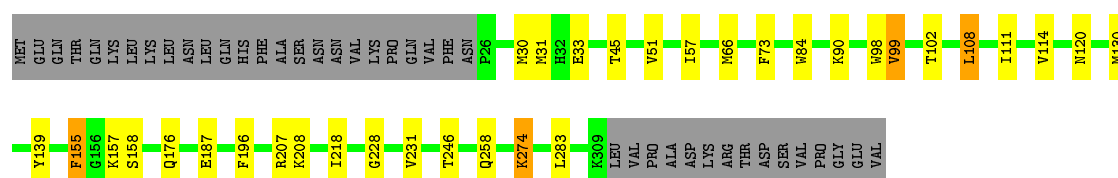
- Molecule 1: Major head protein

Chain D:  68% 18% 12%



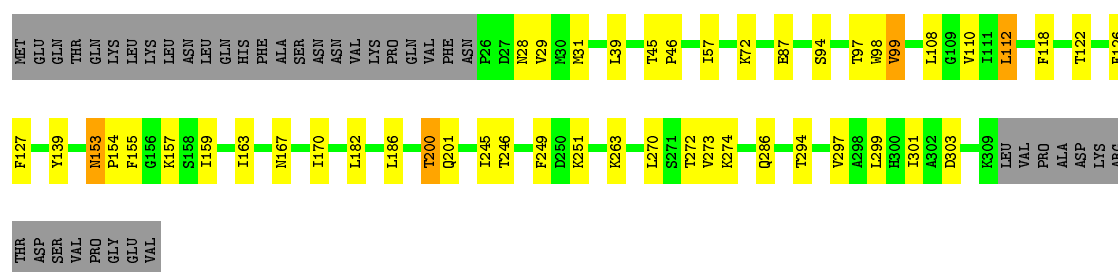
- Molecule 1: Major head protein

Chain E:  77% 9% 12%



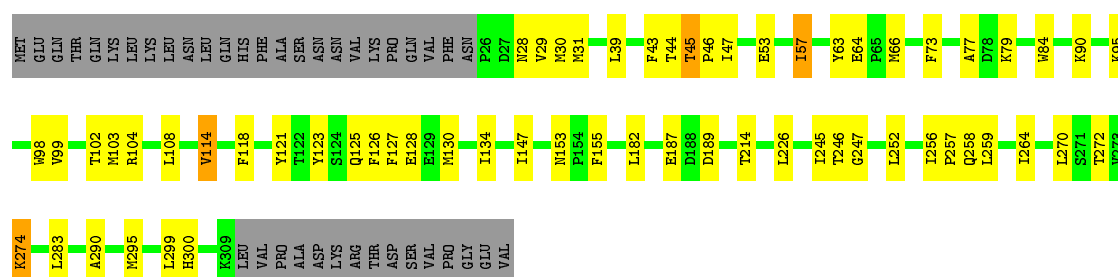
- Molecule 1: Major head protein

Chain F:  73% 14% 12%



- Molecule 1: Major head protein

Chain G:  69% 18% 12%



- Molecule 2: Scaffold protein

Chain a:  7% 93%

GLU
ALA
SER
ARG
LYS
MET
LEU
SER
GLU
ALA
GLU
VAL
ASP
SER
SER
ASP
GLU
VAL
VAL
ASN
LEU
VAL
VAL
THR
THR
ASP
THR
ALA
GLU
GLN
LYS
THR
LYS
SER
ASN
VAL
ASP
GLU
ALA
PHE
SER
ASN
ALA
VAL
LYS
LYS
ALA
VAL
ASN
GLU
ALA
ARG
LYS
THR
GLU
LEU
PRO
SER
GLN
GLY
THR
GLY
ARG

ASP
SER
PHE
ASN
HIS
SER
THR
LYS
LYS
P191
K206

- Molecule 2: Scaffold protein

Chain e: 8% 92%

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

VAL
ALA
HIS
GLU
LYS
LYS
LYS
ALA
ASP
TYR
ALA
LYS
GLU
LYS
LYS
VAL
ALA
SER
VAL
GLU
ALA
ALA
ASP
LEU
VAL
VAL
GLU
THR
ALA
ASP
LYS
LYS
GLY
ALA
LEU
ALA
GLU
GLN
LYS
THR
LYS
ASN
LYS
ASN
ASP
VAL
GLN
LYS
LYS
ALA
ASP
LYS
PHE
GLU
TYR
GLU
ARG
GLU
VAL
GLN
MET
LYS
LYS
GLU
SER
VAL
ARG
SER
LYS
LYS
VAL
GLN
GLN
LYS
ASP
GLY
LEU
ASN
ILE
LEU
MET
GLU
THR
GLY
SER

GLU
ALA
ARG
GLU
MET
LYS
LEU
SER
GLU
GLU
VAL
VAL
ASP
SER
SER
SER
ASP
GLU
GLN
PHE
PHE
ALA
ASP
GLU
VAL
VAL
SER
VAL
ASP
ASN
LEU
VAL
VAL
GLU
THR
ALA
ASP
LYS
LYS
THR
LYS
ASN
LYS
ASN
ASP
VAL
GLN
LYS
LYS
ALA
ASP
PHE
GLU
SER
ASN
ALA
VAL
LYS
LYS
MET
GLU
THR
ILE
THR
PHE
THR

ASP
SER
PHE
ASN
HIS
SER
THR
LYS
LYS
P191
K206

- Molecule 2: Scaffold protein

Chain f: 7% 92%

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

VAL
ALA
HIS
GLU
LYS
LYS
LYS
ALA
ASP
TYR
ALA
LYS
GLU
LYS
LYS
VAL
ALA
SER
VAL
GLU
ALA
ALA
ASP
LEU
VAL
VAL
GLU
THR
ALA
ASP
LYS
LYS
GLY
ALA
LEU
ALA
GLU
GLN
LYS
THR
LYS
ASN
LYS
ASN
ASP
VAL
GLN
LYS
LYS
ALA
ASP
PHE
GLU
TYR
GLU
ARG
GLU
VAL
GLN
MET
LYS
LYS
GLU
SER
VAL
ARG
SER
LYS
LYS
VAL
GLN
GLN
LYS
ASP
GLY
LEU
ASN
ILE
LEU
MET
GLU
THR
GLY
SER

GLU
ALA
ARG
GLU
MET
LYS
LEU
SER
GLU
GLU
VAL
VAL
ASP
SER
SER
SER
ASP
GLU
GLN
PHE
PHE
ALA
ASP
GLU
VAL
VAL
SER
VAL
ASP
ASN
LEU
VAL
VAL
GLU
THR
ALA
ASP
LYS
LYS
THR
LYS
ASN
LYS
ASN
ASP
VAL
GLN
LYS
LYS
ALA
ASP
PHE
GLU
SER
ASN
ALA
VAL
LYS
LYS
MET
GLU
THR
ILE
THR
PHE
THR

ASP
SER
PHE
ASN
HIS
SER
THR
LYS
LYS
P191
Q192
N193
K206

- Molecule 2: Scaffold protein

Chain g: 7% 92%

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

VAL
ALA
HIS
GLU
LYS
LYS
LYS
ALA
ASP
TYR
ALA
LYS
GLU
LYS
LYS
VAL
ALA
SER
VAL
GLU
ALA
ALA
ASP
LEU
VAL
VAL
GLU
THR
ALA
ASP
LYS
LYS
GLY
ALA
LEU
ALA
GLU
GLN
LYS
THR
LYS
ASN
LYS
ASN
ASP
VAL
GLN
LYS
LYS
ALA
ASP
PHE
GLU
TYR
GLU
ARG
GLU
VAL
GLN
MET
LYS
LYS
VAL
VAL
ASN
GLU
PHE
THR

GLU
ALA
ARG
LYS
MET
LYS
LEU
SER
GLU
GLU
VAL
VAL
ASP
SER
SER
SER
ASP
GLU
GLN
PHE
PHE
ALA
ASP
GLU
VAL
VAL
SER
VAL
ASP
ASN
LEU
VAL
VAL
GLU
THR
ALA
ASP
LYS
LYS
THR
LYS
ASN
LYS
ASN
VAL
VAL
GLU
VAL
VAL
ASN
GLU
ALA
ARG
VAL
LYS
LYS
ALA
VAL
VAL
VAL
ASN
GLU
ALA
VAL
VAL
VAL
VAL
ASN
GLU
PHE
THR

ASP
SER
PHE
ASN
HIS
SER
THR
LYS
LYS
P191
R202
K205
K206

4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of particles used	10557	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING ONLY	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	40	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	DIRECT ELECTRON DE-20 (5k x 3k)	Depositor

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 >2	RMSZ	# Z >2
1	A	0.42	0/2312	0.59	0/3116
1	B	0.40	0/2312	0.56	0/3116
1	C	0.39	0/2312	0.58	0/3116
1	D	0.40	0/2312	0.58	0/3116
1	E	0.40	0/2312	0.58	0/3116
1	F	0.40	0/2312	0.57	0/3116
1	G	0.39	0/2312	0.58	0/3116
2	a	0.42	0/115	0.51	0/151
2	b	0.35	0/132	0.60	0/174
2	c	0.38	0/124	0.51	0/163
2	d	0.32	0/132	0.54	0/174
2	e	0.33	0/132	0.53	0/174
2	f	0.36	0/132	0.51	0/174
2	g	0.34	0/132	0.54	0/174
All	All	0.40	0/17083	0.57	0/22996

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	2270	0	2261	5	0
1	B	2270	0	2261	6	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	2270	0	2261	32	0
1	D	2270	0	2261	19	0
1	E	2270	0	2261	6	0
1	F	2270	0	2261	17	0
1	G	2270	0	2261	30	0
2	a	116	0	131	0	0
2	b	132	0	147	0	0
2	c	125	0	139	0	0
2	d	132	0	147	0	0
2	e	132	0	147	0	0
2	f	132	0	147	0	0
2	g	132	0	147	0	0
All	All	16791	0	16832	112	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 3.

The worst 5 of 112 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:G:39:LEU:HD21	1:G:127:PHE:CE2	1.76	1.21
1:C:126:PHE:CD1	1:C:130:MET:HG3	1.78	1.16
1:G:39:LEU:HD21	1:G:127:PHE:HE2	1.13	1.01
1:C:126:PHE:CE1	1:C:130:MET:CG	2.44	1.00
1:C:126:PHE:CD1	1:C:130:MET:CG	2.53	0.91

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	A	282/324 (87%)	235 (83%)	32 (11%)	15 (5%)	2	26
1	B	282/324 (87%)	230 (82%)	39 (14%)	13 (5%)	3	30
1	C	282/324 (87%)	236 (84%)	39 (14%)	7 (2%)	6	44
1	D	282/324 (87%)	236 (84%)	36 (13%)	10 (4%)	4	38
1	E	282/324 (87%)	231 (82%)	44 (16%)	7 (2%)	6	44
1	F	282/324 (87%)	239 (85%)	34 (12%)	9 (3%)	5	40
1	G	282/324 (87%)	228 (81%)	43 (15%)	11 (4%)	3	34
2	a	12/206 (6%)	11 (92%)	1 (8%)	0	100	100
2	b	14/206 (7%)	13 (93%)	0	1 (7%)	1	20
2	c	13/206 (6%)	13 (100%)	0	0	100	100
2	d	14/206 (7%)	13 (93%)	1 (7%)	0	100	100
2	e	14/206 (7%)	12 (86%)	2 (14%)	0	100	100
2	f	14/206 (7%)	14 (100%)	0	0	100	100
2	g	14/206 (7%)	14 (100%)	0	0	100	100
All	All	2069/3710 (56%)	1725 (83%)	271 (13%)	73 (4%)	7	38

5 of 73 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	30	MET
1	A	113	PRO
1	A	114	VAL
1	A	192	GLU
1	B	113	PRO

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	A	249/286 (87%)	217 (87%)	32 (13%)	5	29
1	B	249/286 (87%)	222 (89%)	27 (11%)	7	37

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	249/286 (87%)	228 (92%)	21 (8%)	13	49
1	D	249/286 (87%)	217 (87%)	32 (13%)	5	29
1	E	249/286 (87%)	228 (92%)	21 (8%)	13	49
1	F	249/286 (87%)	234 (94%)	15 (6%)	22	61
1	G	249/286 (87%)	223 (90%)	26 (10%)	8	39
2	a	12/178 (7%)	12 (100%)	0	100	100
2	b	14/178 (8%)	13 (93%)	1 (7%)	17	56
2	c	13/178 (7%)	13 (100%)	0	100	100
2	d	14/178 (8%)	14 (100%)	0	100	100
2	e	14/178 (8%)	14 (100%)	0	100	100
2	f	14/178 (8%)	13 (93%)	1 (7%)	17	56
2	g	14/178 (8%)	12 (86%)	2 (14%)	4	26
All	All	1838/3248 (57%)	1660 (90%)	178 (10%)	14	42

5 of 178 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	31	MET
1	D	186	LEU
1	G	189	ASP
1	D	47	ILE
1	D	85	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	32	HIS
1	C	238	ASN
1	G	202	ASN
1	B	178	ASN
1	B	202	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.