



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Jul 24, 2017 – 08:46 PM EDT

PDB ID : 5VYA
EMDB ID: : EMD-8746
Title : S. cerevisiae Hsp104:casein complex, Extended Conformation
Authors : Gates, S.N.; Yokom, A.L.; Lin, J.-B.; Jackrel, M.E.; Rizo, A.N.; Kendsersky, N.M.; Buell, C.E.; Sweeny, E.A.; Chuang, E.; Torrente, M.P.; Mack, K.L.; Su, M.; Shorter, J.; Southworth, D.R.
Deposited on : unknown
Resolution : 4.00 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
Mogul : 1.7.2 (RC1), CSD as538be (2017)
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-20029824

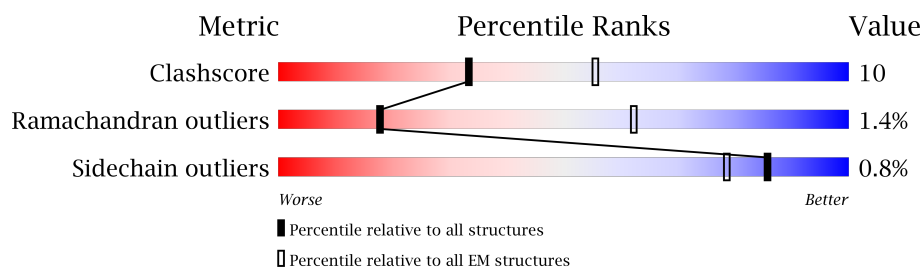
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.00 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	908	
1	B	908	
1	C	908	
1	D	908	
1	E	908	
1	F	908	
2	P	28	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	AGS	B	1002	-	-	X	-
3	AGS	D	1002	-	-	X	-
3	AGS	E	1001	-	-	X	-

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 27691 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.

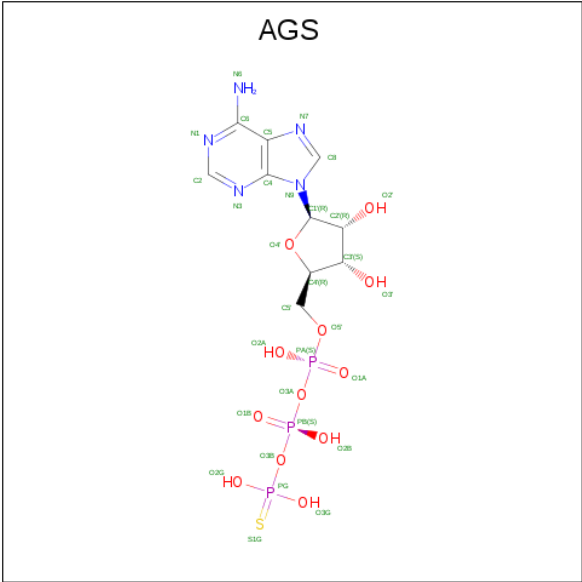
- Molecule 1 is a protein called Heat shock protein 104.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	579	Total	C	N	O	S	0	0
			4530	2850	794	868	18		
1	B	579	Total	C	N	O	S	0	0
			4530	2850	794	868	18		
1	C	579	Total	C	N	O	S	0	0
			4530	2850	794	868	18		
1	D	579	Total	C	N	O	S	0	0
			4530	2850	794	868	18		
1	E	579	Total	C	N	O	S	0	0
			4530	2850	794	868	18		
1	F	579	Total	C	N	O	S	0	0
			4530	2850	794	868	18		

- Molecule 2 is a protein called Alpha-S1-casein.

Mol	Chain	Residues	Atoms				AltConf	Trace
2	P	28	Total	C	N	O	0	0
			139	83	28	28		

- Molecule 3 is PHOSPHOTHIOPHOSPHORIC ACID-ADENYLATE ESTER (three-letter code: AGS) (formula: C₁₀H₁₆N₅O₁₂P₃S).

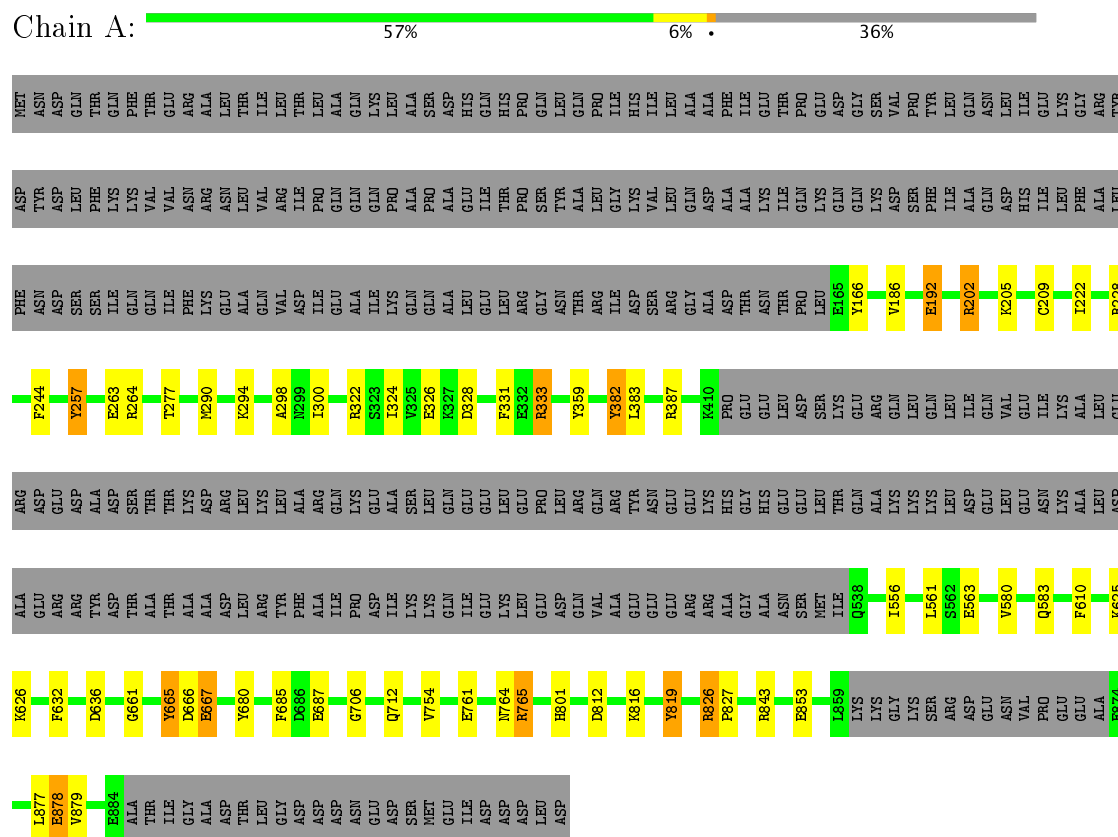


Mol	Chain	Residues	Atoms						AltConf
3	A	1	Total	C	N	O	P	S	0
			62	20	10	24	6	2	
3	A	1	Total	C	N	O	P	S	0
			62	20	10	24	6	2	
3	B	1	Total	C	N	O	P	S	0
			62	20	10	24	6	2	
3	B	1	Total	C	N	O	P	S	0
			62	20	10	24	6	2	
3	C	1	Total	C	N	O	P	S	0
			62	20	10	24	6	2	
3	C	1	Total	C	N	O	P	S	0
			62	20	10	24	6	2	
3	D	1	Total	C	N	O	P	S	0
			62	20	10	24	6	2	
3	D	1	Total	C	N	O	P	S	0
			62	20	10	24	6	2	
3	E	1	Total	C	N	O	P	S	0
			62	20	10	24	6	2	
3	E	1	Total	C	N	O	P	S	0
			62	20	10	24	6	2	
3	F	1	Total	C	N	O	P	S	0
			62	20	10	24	6	2	
3	F	1	Total	C	N	O	P	S	0
			62	20	10	24	6	2	

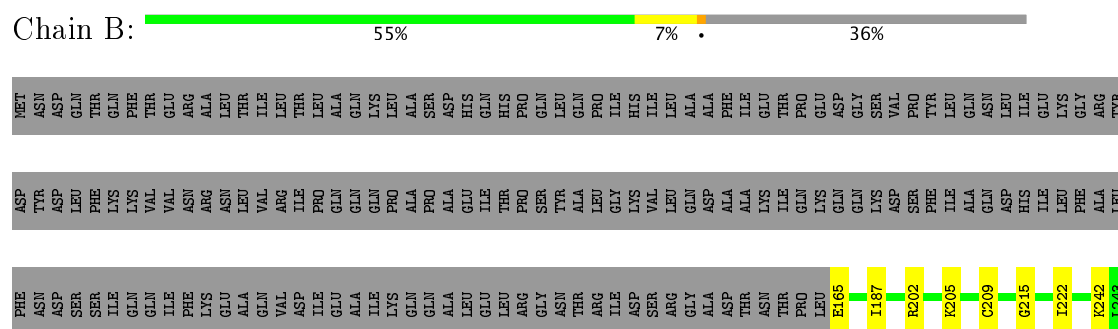
3 Residue-property plots

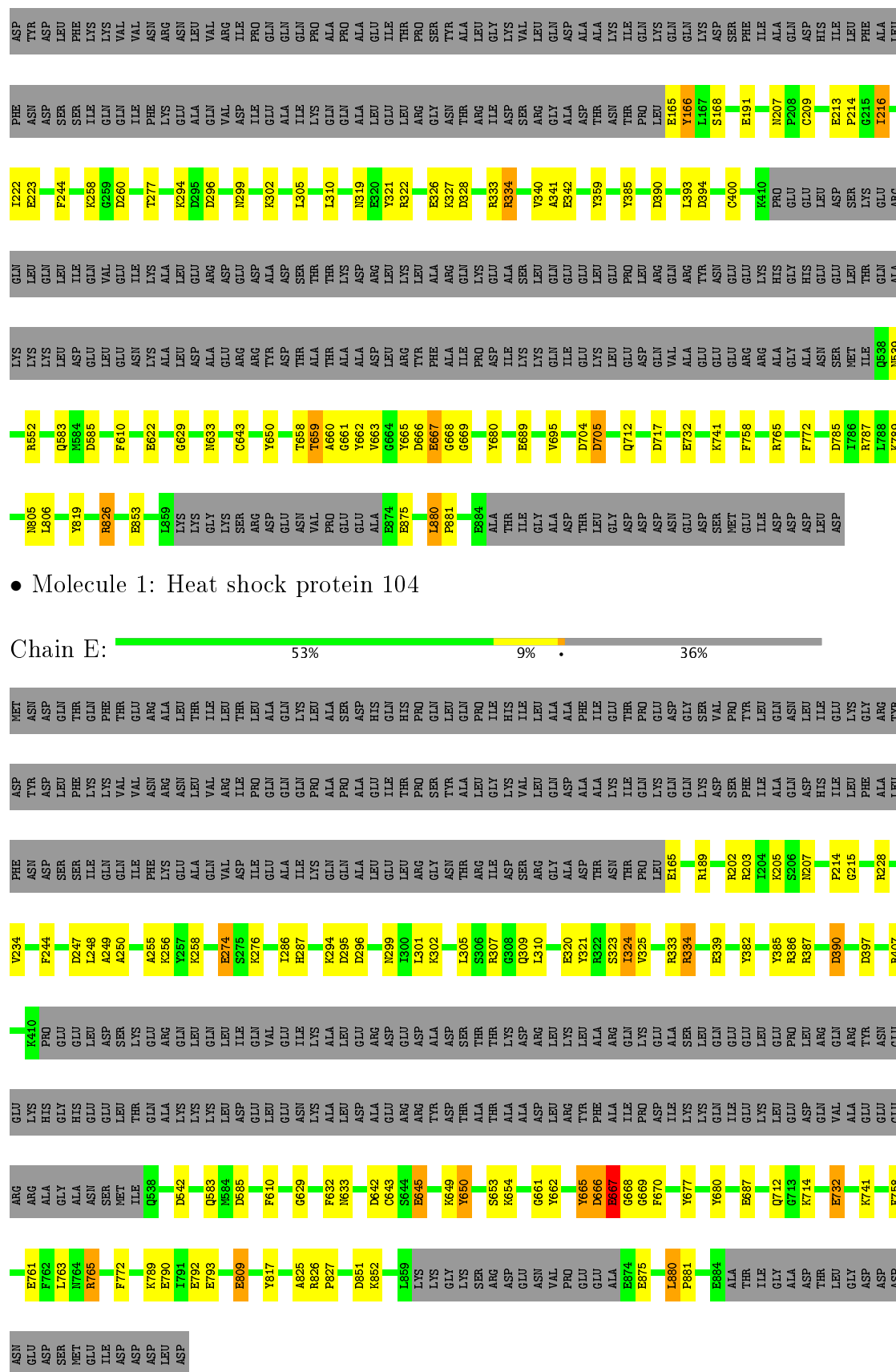
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according 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: Heat shock protein 104



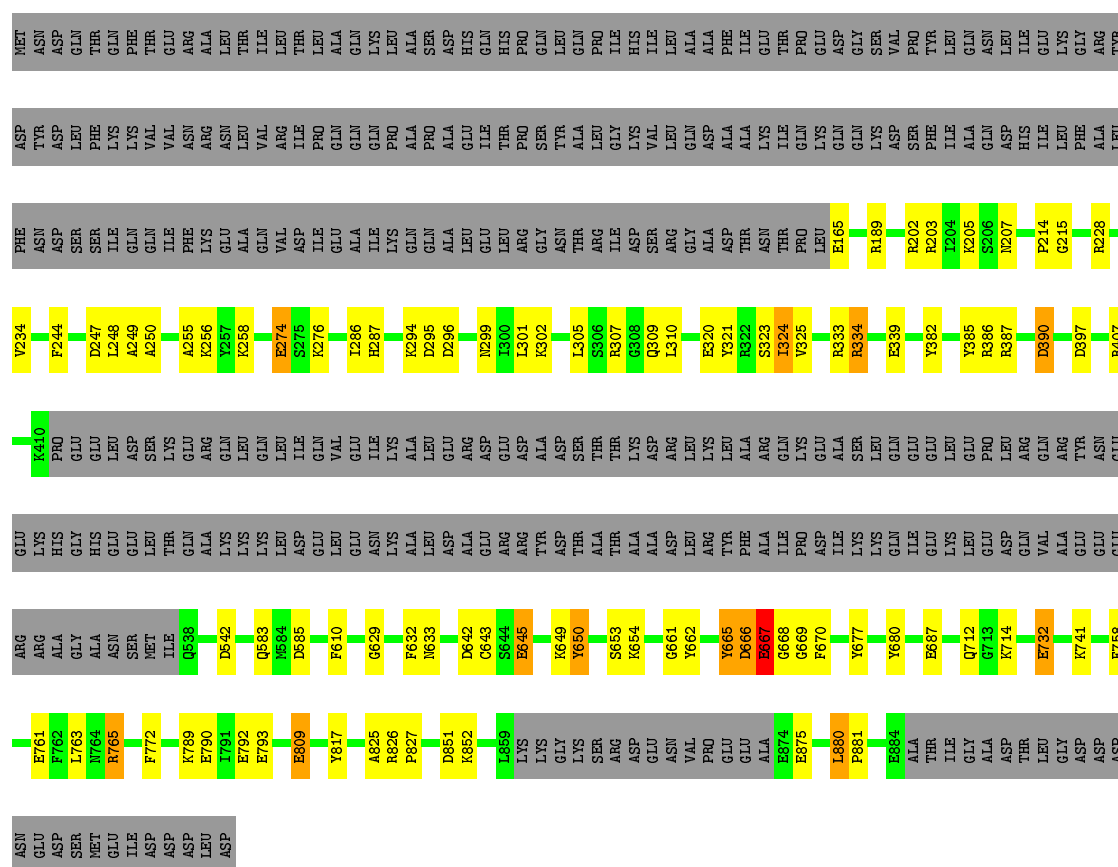
- Molecule 1: Heat shock protein 104



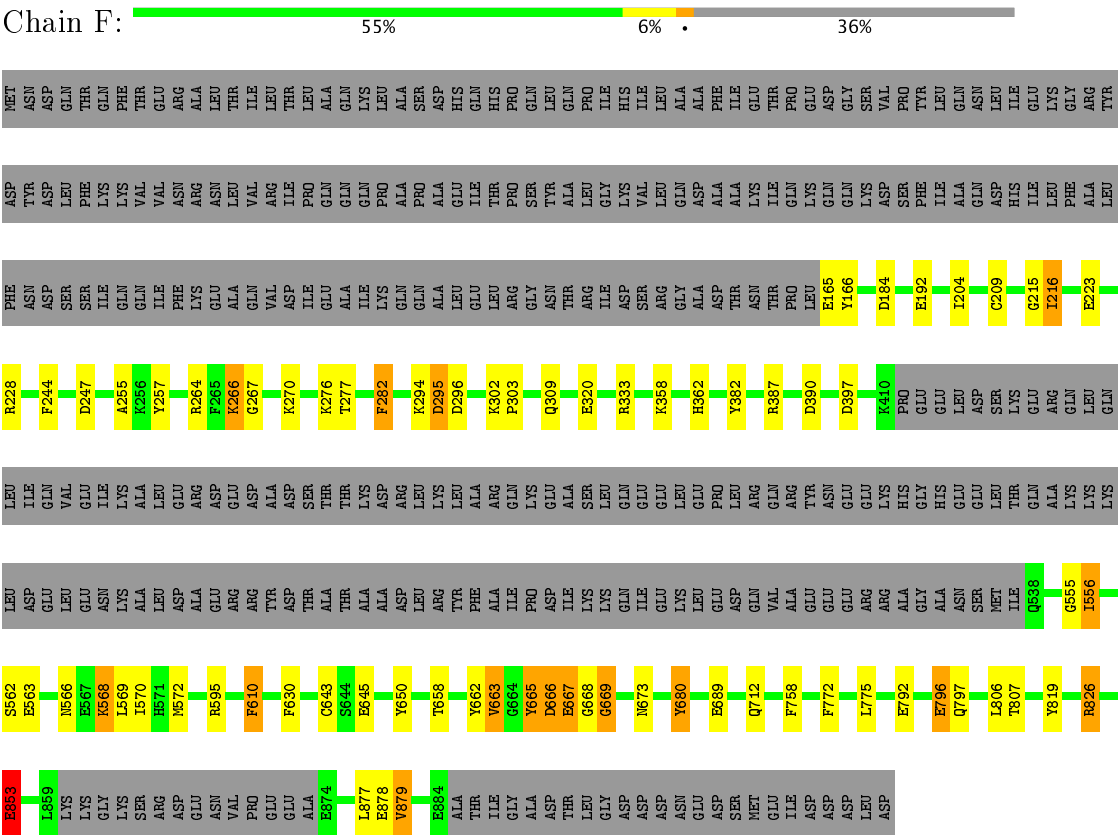


• Molecule 1: Heat shock protein 104

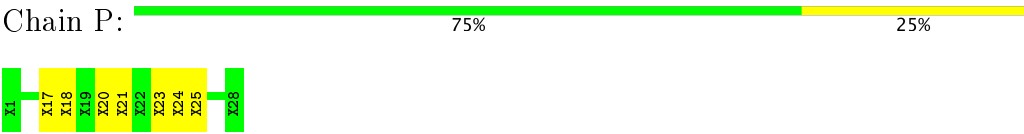
Chain E:



• Molecule 1: Heat shock protein 104



• Molecule 2: Alpha-S1-casein



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	146463	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	0.8	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	50000	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: AGS

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	1.12	13/4586 (0.3%)	0.94	12/6171 (0.2%)
1	B	1.20	17/4586 (0.4%)	0.92	5/6171 (0.1%)
1	C	1.26	20/4586 (0.4%)	0.95	8/6171 (0.1%)
1	D	1.24	18/4586 (0.4%)	0.93	5/6171 (0.1%)
1	E	1.24	29/4586 (0.6%)	0.95	7/6171 (0.1%)
1	F	1.18	21/4586 (0.5%)	0.93	8/6171 (0.1%)
All	All	1.21	118/27516 (0.4%)	0.94	45/37026 (0.1%)

The worst 5 of 118 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	643	CYS	CB-SG	-8.99	1.67	1.82
1	D	610	PHE	CB-CG	-8.63	1.36	1.51
1	F	673	ASN	CB-CG	-8.55	1.31	1.51
1	F	282	PHE	CB-CG	-8.44	1.37	1.51
1	E	643	CYS	CB-SG	-8.34	1.68	1.82

The worst 5 of 45 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	595	ARG	NE-CZ-NH1	8.40	124.50	120.30
1	E	228	ARG	NE-CZ-NH2	-8.20	116.20	120.30
1	C	595	ARG	NE-CZ-NH2	-7.89	116.36	120.30
1	E	321	TYR	CB-CG-CD1	-7.82	116.31	121.00
1	F	166	TYR	CB-CG-CD2	-7.74	116.36	121.00

There are no chirality outliers.

There are no planarity outliers.

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	A	4530	0	4648	71	0
1	B	4530	0	4645	144	0
1	C	4530	0	4645	125	0
1	D	4530	0	4644	133	0
1	E	4530	0	4645	127	0
1	F	4530	0	4643	80	0
2	P	139	0	39	19	0
3	A	62	0	24	7	0
3	B	62	0	24	13	0
3	C	62	0	24	10	0
3	D	62	0	24	14	0
3	E	62	0	24	20	0
3	F	62	0	24	9	0
All	All	27691	0	28053	578	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.

The worst 5 of 578 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:B:662:TYR:HD2	1:C:650:TYR:CE1	1.13	1.60
1:B:662:TYR:CD2	1:C:650:TYR:CE1	1.86	1.56
1:D:665:TYR:CE1	1:D:712:GLN:CG	1.87	1.52
1:E:662:TYR:CD2	1:F:650:TYR:HE1	1.21	1.52
1:D:665:TYR:CE1	1:D:712:GLN:HG2	0.98	1.50

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 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 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	573/908 (63%)	530 (92%)	31 (5%)	12 (2%)	8	48
1	B	573/908 (63%)	544 (95%)	23 (4%)	6 (1%)	18	61
1	C	573/908 (63%)	548 (96%)	21 (4%)	4 (1%)	25	68
1	D	573/908 (63%)	541 (94%)	27 (5%)	5 (1%)	20	63
1	E	573/908 (63%)	542 (95%)	26 (4%)	5 (1%)	20	63
1	F	573/908 (63%)	529 (92%)	29 (5%)	15 (3%)	6	43
All	All	3438/5448 (63%)	3234 (94%)	157 (5%)	47 (1%)	18	55

5 of 47 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	290	MET
1	B	295	ASP
1	B	666	ASP
1	C	295	ASP
1	C	705	ASP

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	497/783 (64%)	494 (99%)	3 (1%)	89	95
1	B	497/783 (64%)	492 (99%)	5 (1%)	80	90
1	C	497/783 (64%)	492 (99%)	5 (1%)	80	90
1	D	497/783 (64%)	493 (99%)	4 (1%)	85	93
1	E	497/783 (64%)	493 (99%)	4 (1%)	85	93
1	F	497/783 (64%)	494 (99%)	3 (1%)	89	95
All	All	2982/4698 (64%)	2958 (99%)	24 (1%)	86	93

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	659	THR
1	D	216	ILE
1	F	568	LYS
1	C	665	TYR
1	C	765	ARG

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

Mol	Chain	Res	Type
1	D	712	GLN
1	E	712	GLN

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](#)

12 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or 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 length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	AGS	A	1001	-	26,33,33	2.02	3 (11%)	22,52,52	2.10	4 (18%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	AGS	A	1002	-	26,33,33	2.12	4 (15%)	22,52,52	2.18	5 (22%)
3	AGS	B	1001	-	26,33,33	1.98	5 (19%)	22,52,52	2.37	5 (22%)
3	AGS	B	1002	-	26,33,33	1.95	6 (23%)	22,52,52	2.53	6 (27%)
3	AGS	C	1001	-	26,33,33	1.86	5 (19%)	22,52,52	2.48	5 (22%)
3	AGS	C	1002	-	26,33,33	1.93	6 (23%)	22,52,52	2.52	6 (27%)
3	AGS	D	1001	-	26,33,33	1.91	3 (11%)	22,52,52	2.28	7 (31%)
3	AGS	D	1002	1	26,33,33	1.88	6 (23%)	22,52,52	2.63	4 (18%)
3	AGS	E	1001	-	26,33,33	2.04	5 (19%)	22,52,52	2.28	7 (31%)
3	AGS	E	1002	-	26,33,33	1.87	5 (19%)	22,52,52	2.63	4 (18%)
3	AGS	F	1001	1	26,33,33	1.93	4 (15%)	22,52,52	2.39	5 (22%)
3	AGS	F	1002	-	26,33,33	1.93	4 (15%)	22,52,52	2.48	6 (27%)

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	AGS	A	1001	-	-	0/17/38/38	0/3/3/3
3	AGS	A	1002	-	-	0/17/38/38	0/3/3/3
3	AGS	B	1001	-	-	0/17/38/38	0/3/3/3
3	AGS	B	1002	-	-	0/17/38/38	0/3/3/3
3	AGS	C	1001	-	-	0/17/38/38	0/3/3/3
3	AGS	C	1002	-	-	0/17/38/38	0/3/3/3
3	AGS	D	1001	-	-	0/17/38/38	0/3/3/3
3	AGS	D	1002	1	-	0/17/38/38	0/3/3/3
3	AGS	E	1001	-	-	0/17/38/38	0/3/3/3
3	AGS	E	1002	-	-	0/17/38/38	0/3/3/3
3	AGS	F	1001	1	-	0/17/38/38	0/3/3/3
3	AGS	F	1002	-	-	0/17/38/38	0/3/3/3

The worst 5 of 56 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	1001	AGS	C2'-C1'	-2.84	1.49	1.53
3	D	1002	AGS	C2'-C1'	-2.67	1.49	1.53
3	E	1002	AGS	C2'-C1'	-2.67	1.49	1.53
3	C	1001	AGS	C2'-C1'	-2.60	1.49	1.53
3	B	1002	AGS	C2'-C1'	-2.51	1.49	1.53

The worst 5 of 64 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
3	E	1002	AGS	PB-O3B-PG	-8.92	103.52	132.35
3	D	1002	AGS	PB-O3B-PG	-8.91	103.56	132.35
3	C	1001	AGS	PB-O3B-PG	-8.73	104.12	132.35
3	C	1002	AGS	PB-O3B-PG	-8.27	105.63	132.35
3	B	1002	AGS	PB-O3B-PG	-8.26	105.66	132.35

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

12 monomers are involved in 73 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1001	AGS	5	0
3	A	1002	AGS	2	0
3	B	1001	AGS	4	0
3	B	1002	AGS	9	0
3	C	1001	AGS	2	0
3	C	1002	AGS	8	0
3	D	1001	AGS	5	0
3	D	1002	AGS	9	0
3	E	1001	AGS	16	0
3	E	1002	AGS	4	0
3	F	1001	AGS	4	0
3	F	1002	AGS	5	0

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.