wwPDB EM Validation Summary Report

Dec 13, 2022 – 12:06 AM EST

PDB ID	:	1LD4
Title	:	Placement of the Structural Proteins in Sindbis Virus
Authors	:	Zhang, W.; Mukhopadhyay, S.; Pletnev, S.V.; Baker, T.S.; Kuhn, R.J.; Rossmann, M.G.
Deposited on	:	2002-04-08
Resolution	:	11.40 Å(reported)
Based on initial models	:	1SVB, 1YSA, 1I9W

This is a wwPDB EM 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
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

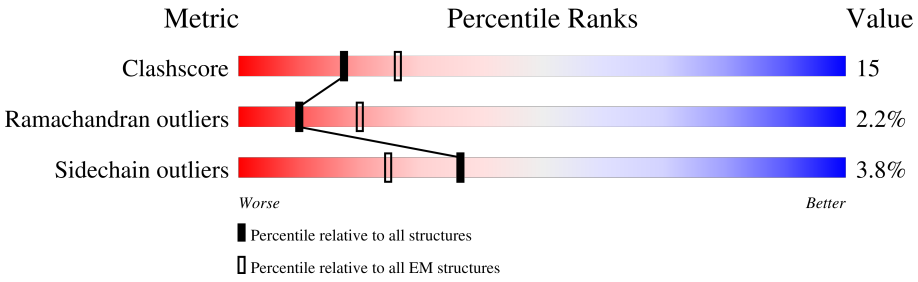
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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 11.40 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	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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	264	47% 9% · 43%
1	B	264	48% 8% · 43%
1	C	264	48% 8% · 43%
1	D	264	48% 8% · 43%
2	E	57	49% 51%
2	F	57	49% 51%
2	G	57	49% 51%
2	H	57	49% 51%
2	I	57	49% 51%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	J	57	49% 51%
2	K	57	49% 51%
2	L	57	49% 51%
3	M	439	54% 28% 16%
3	N	439	55% 28% 16%
3	O	439	55% 28% 16%
3	P	439	55% 27% 16%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 15680 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 Coat protein C.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	151	Total	C	N	O	S	0	0
			1162	731	207	219	5		
1	B	151	Total	C	N	O	S	0	0
			1162	731	207	219	5		
1	C	151	Total	C	N	O	S	0	0
			1162	731	207	219	5		
1	D	151	Total	C	N	O	S	0	0
			1162	731	207	219	5		

- Molecule 2 is a protein called GENERAL CONTROL PROTEIN GCN4.

Mol	Chain	Residues	Atoms		AltConf	Trace
2	E	28	Total	C	0	28
			28	28		
2	F	28	Total	C	0	28
			28	28		
2	G	28	Total	C	0	28
			28	28		
2	H	28	Total	C	0	28
			28	28		
2	I	28	Total	C	0	28
			28	28		
2	J	28	Total	C	0	28
			28	28		
2	K	28	Total	C	0	28
			28	28		
2	L	28	Total	C	0	28
			28	28		

- Molecule 3 is a protein called Spike glycoprotein E1.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	M	369	Total	C	N	O	S	0	15
			2694	1709	446	519	20		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
3	N	369	Total	C	N	O	S	0	15
			2694	1709	446	519	20		
3	O	369	Total	C	N	O	S	0	15
			2694	1709	446	519	20		
3	P	369	Total	C	N	O	S	0	15
			2694	1709	446	519	20		

- Molecule 4 is UNKNOWN ATOM OR ION (three-letter code: UNX) (formula: X).

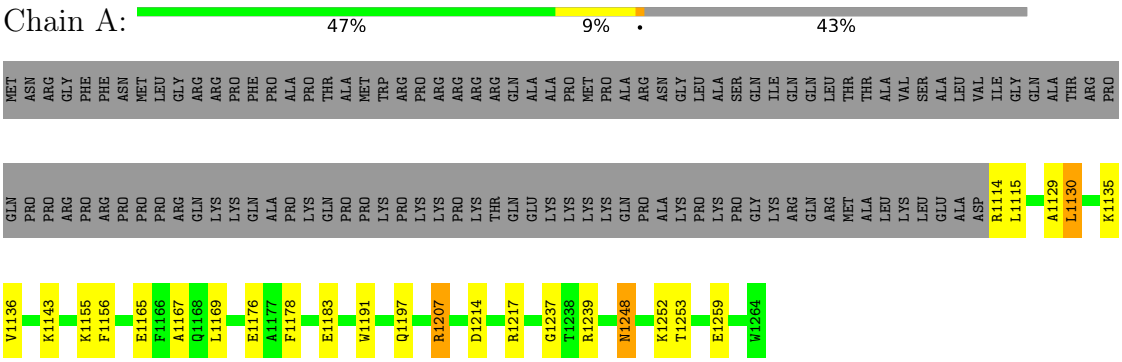
Mol	Chain	Residues	Atoms		AltConf
4	F	2	Total	X	0
			2	2	
4	H	1	Total	X	0
			1	1	
4	J	1	Total	X	0
			1	1	
4	L	1	Total	X	0
			1	1	
4	M	7	Total	X	0
			7	7	
4	N	6	Total	X	0
			6	6	
4	O	8	Total	X	0
			8	8	
4	P	6	Total	X	0
			6	6	

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

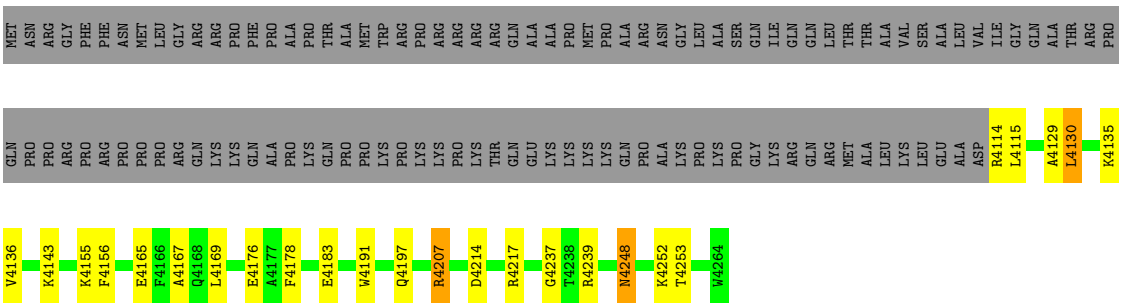
- Molecule 1: Coat protein C

Chain A:



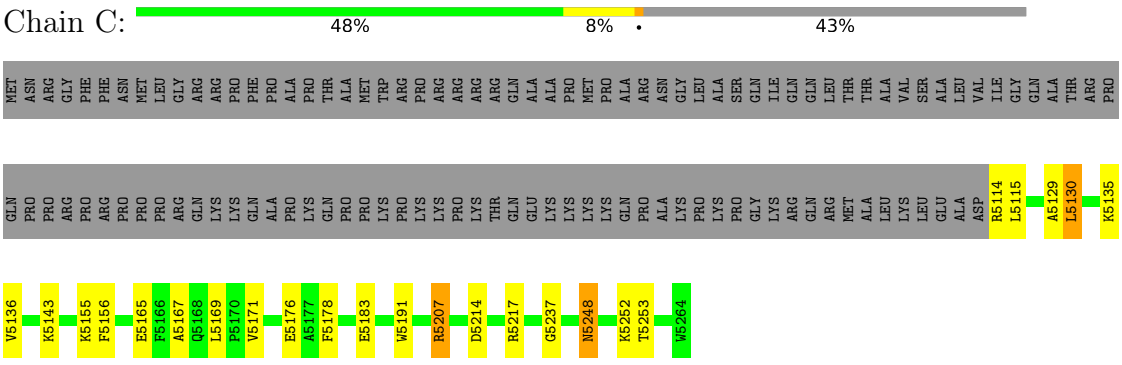
- Molecule 1: Coat protein C

Chain B:



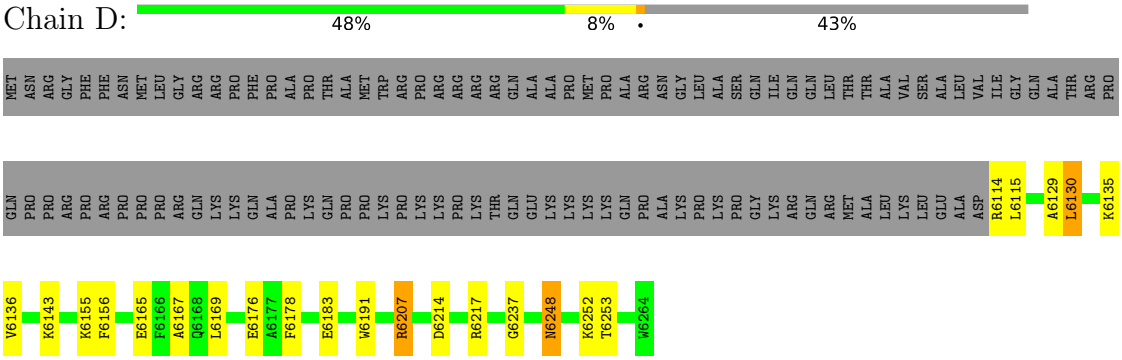
- Molecule 1: Coat protein C

Chain C:



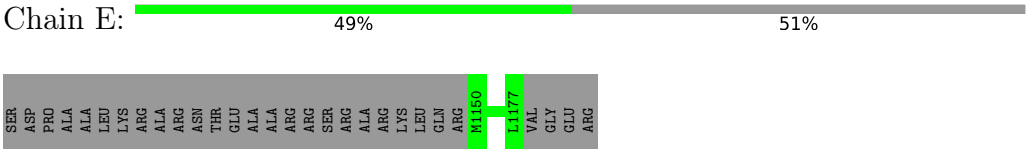
• Molecule 1: Coat protein C

Chain D:



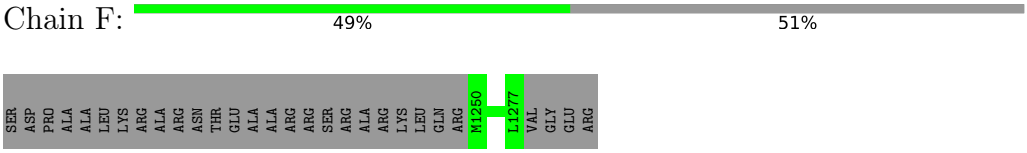
• Molecule 2: GENERAL CONTROL PROTEIN GCN4

Chain E:



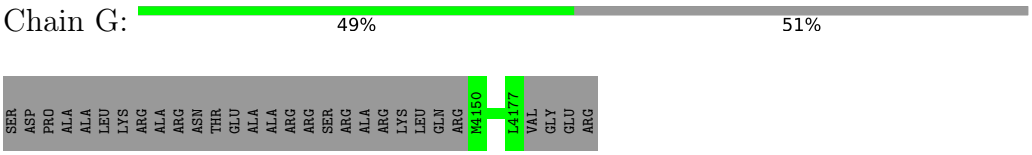
• Molecule 2: GENERAL CONTROL PROTEIN GCN4

Chain F:



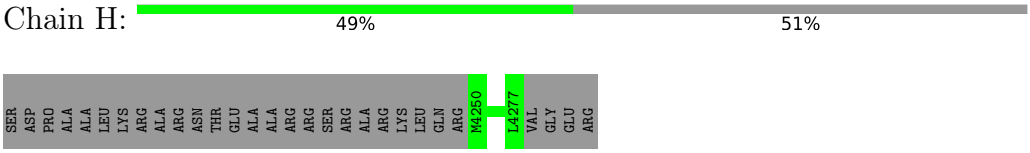
• Molecule 2: GENERAL CONTROL PROTEIN GCN4

Chain G:



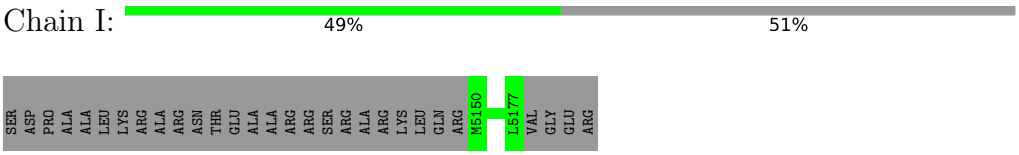
• Molecule 2: GENERAL CONTROL PROTEIN GCN4

Chain H:



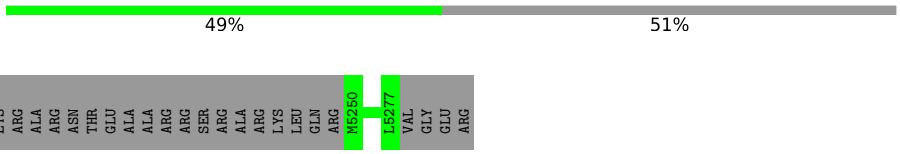
• Molecule 2: GENERAL CONTROL PROTEIN GCN4

Chain I:



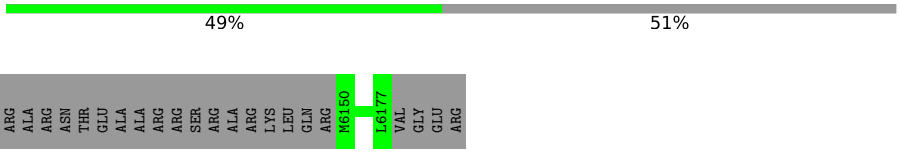
• Molecule 2: GENERAL CONTROL PROTEIN GCN4

Chain J:



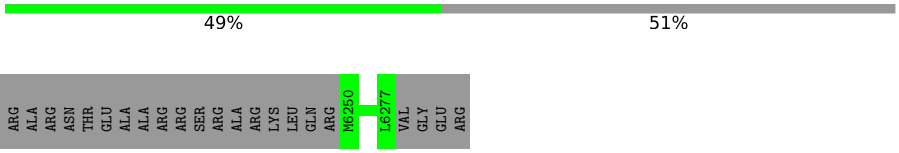
• Molecule 2: GENERAL CONTROL PROTEIN GCN4

Chain K:



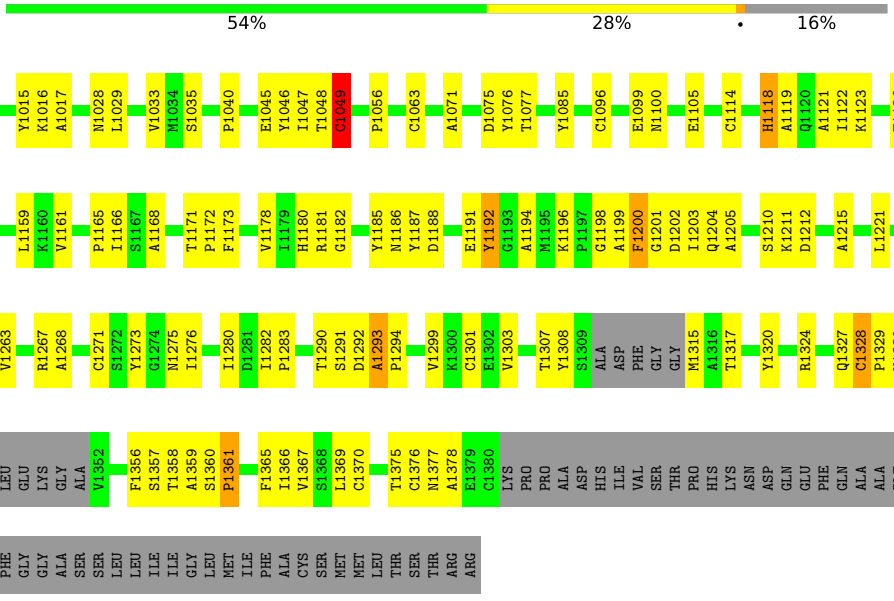
• Molecule 2: GENERAL CONTROL PROTEIN GCN4

Chain L:



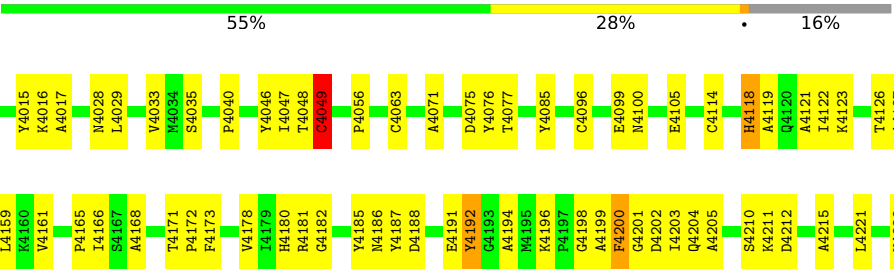
• Molecule 3: Spike glycoprotein E1

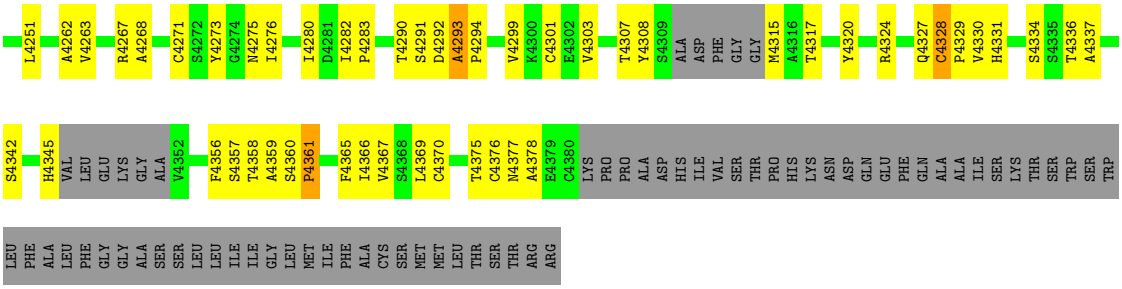
Chain M:



• Molecule 3: Spike glycoprotein E1

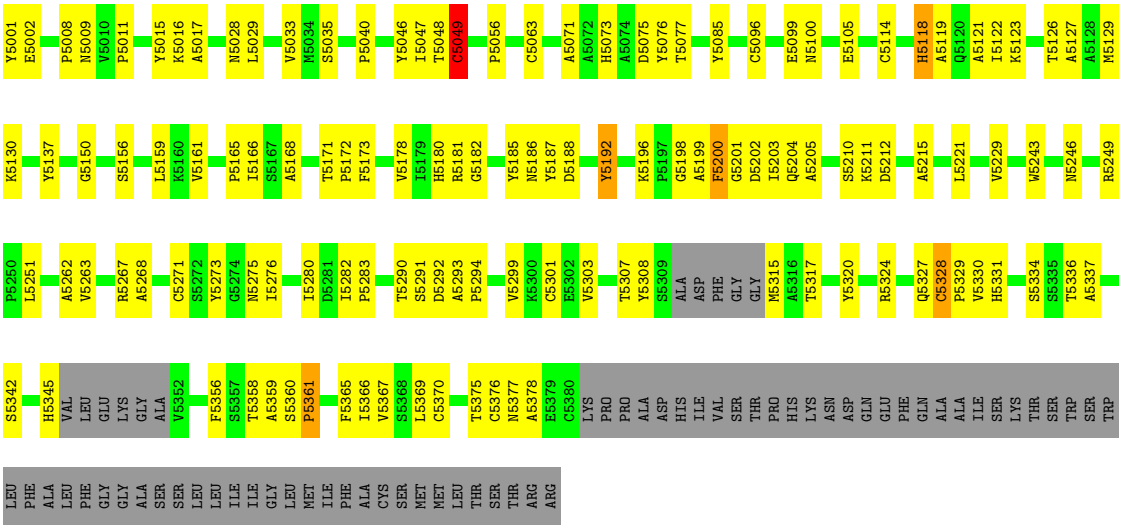
Chain N:





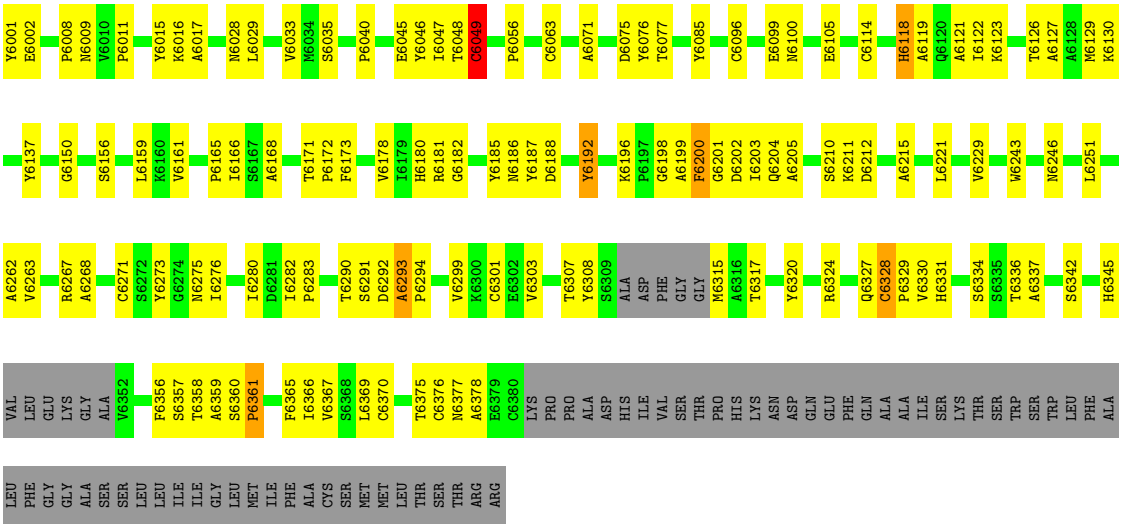
• Molecule 3: Spike glycoprotein E1

Chain O:



• Molecule 3: Spike glycoprotein E1

Chain P:



4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	1.00Å 1.00Å 1.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	(Not available) – 11.40	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-11.40)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	unknown	Depositor
R, R_{free}	(Not available) , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	15680	wwPDB-VP
Average B, all atoms (Å^2)	61.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

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

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	#$\|Z\|>5$	RMSZ	#$\|Z\|>5$
1	A	0.52	0/1190	0.81	1/1607 (0.1%)
1	B	0.52	0/1190	0.81	1/1607 (0.1%)
1	C	0.52	0/1190	0.81	1/1607 (0.1%)
1	D	0.52	0/1190	0.81	1/1607 (0.1%)
3	M	0.34	0/2743	0.54	0/3740
3	N	0.34	0/2743	0.54	0/3740
3	O	0.34	0/2743	0.54	0/3740
3	P	0.34	0/2743	0.54	0/3740
All	All	0.41	0/15732	0.63	4/21388 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	6130	LEU	CA-CB-CG	5.11	127.05	115.30
1	A	1130	LEU	CA-CB-CG	5.10	127.04	115.30
1	B	4130	LEU	CA-CB-CG	5.09	127.02	115.30
1	C	5130	LEU	CA-CB-CG	5.09	127.00	115.30

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	1162	0	1131	16	0
1	B	1162	0	1131	11	0
1	C	1162	0	1131	14	0
1	D	1162	0	1131	9	0
2	E	28	0	0	0	0
2	F	28	0	0	0	0
2	G	28	0	0	0	0
2	H	28	0	0	0	0
2	I	28	0	0	0	0
2	J	28	0	0	0	0
2	K	28	0	0	0	0
2	L	28	0	0	0	0
3	M	2694	0	2605	139	0
3	N	2694	0	2605	135	0
3	O	2694	0	2605	91	0
3	P	2694	0	2605	95	0
4	F	2	0	0	0	0
4	H	1	0	0	0	0
4	J	1	0	0	0	0
4	L	1	0	0	0	0
4	M	7	0	0	0	0
4	N	6	0	0	0	0
4	O	8	0	0	0	0
4	P	6	0	0	0	0
All	All	15680	0	14944	464	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 15.

The worst 5 of 464 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)
3:M:1194:ALA:CB	3:N:4151:VAL:HG12	1.36	1.54
3:M:1151:VAL:HG12	3:N:4194:ALA:CB	1.36	1.48
3:M:1151:VAL:CG1	3:N:4194:ALA:CB	2.05	1.34
3:M:1194:ALA:CB	3:N:4151:VAL:CG1	2.05	1.31
3:O:5330:VAL:HG22	3:O:5369:LEU:CA	1.62	1.29

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	149/264 (56%)	144 (97%)	5 (3%)	0	100	100
1	B	149/264 (56%)	144 (97%)	5 (3%)	0	100	100
1	C	149/264 (56%)	144 (97%)	5 (3%)	0	100	100
1	D	149/264 (56%)	144 (97%)	5 (3%)	0	100	100
3	M	348/439 (79%)	270 (78%)	67 (19%)	11 (3%)	4	26
3	N	348/439 (79%)	270 (78%)	67 (19%)	11 (3%)	4	26
3	O	348/439 (79%)	270 (78%)	67 (19%)	11 (3%)	4	26
3	P	348/439 (79%)	270 (78%)	67 (19%)	11 (3%)	4	26
All	All	1988/2812 (71%)	1656 (83%)	288 (14%)	44 (2%)	10	35

5 of 44 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	M	1126	THR
3	M	1361	PRO
3	N	4126	THR
3	N	4361	PRO
3	O	5126	THR

5.3.2 Protein sidechains (i)

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	122/218 (56%)	115 (94%)	7 (6%)	20	45

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	B	122/218 (56%)	115 (94%)	7 (6%)	20 45
1	C	122/218 (56%)	115 (94%)	7 (6%)	20 45
1	D	122/218 (56%)	115 (94%)	7 (6%)	20 45
3	M	299/370 (81%)	290 (97%)	9 (3%)	41 63
3	N	299/370 (81%)	290 (97%)	9 (3%)	41 63
3	O	299/370 (81%)	290 (97%)	9 (3%)	41 63
3	P	299/370 (81%)	290 (97%)	9 (3%)	41 63
All	All	1684/2352 (72%)	1620 (96%)	64 (4%)	36 57

5 of 64 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	P	6118	HIS
3	P	6192	TYR
1	D	6191	TRP
1	D	6183	GLU
3	P	6200	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
3	O	5345	HIS
3	O	5355	HIS
3	P	6331	HIS
1	D	6190	ASN
1	D	6128	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 32 ligands modelled in this entry, 32 are unknown - 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.