



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Nov 15, 2018 – 03:35 PM EST

PDB ID : 5M0R
EMDB ID: : EMD-4139
Title : Cryo-EM reconstruction of the maedi-visna virus (MVV) strand transfer complex
Authors : Pye, V.E.; Ballandras-Colas, A.; Maskell, D.; Locke, J.; Kotecha, A.; Costa, A.; Cherepanov, P.
Deposited on : 2016-10-05
Resolution : 8.20 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

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

MolProbity : 4.02b-467
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20031633

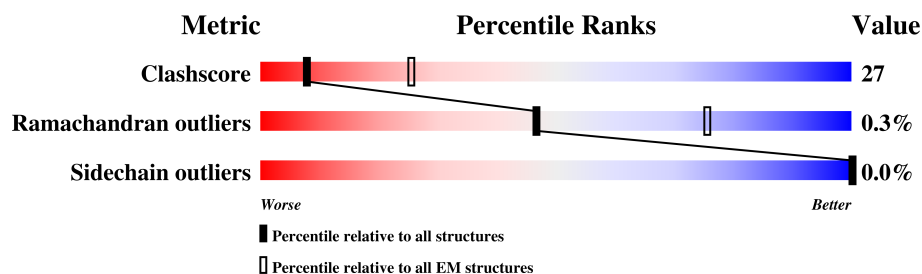
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 8.20 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.













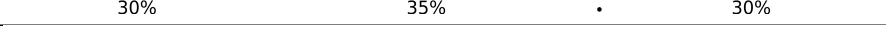
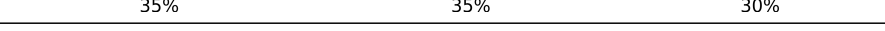
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	136327	1886
Ramachandran outliers	132723	1663
Sidechain outliers	132532	1531

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	281	51% 44% . .
1	B	281	68% 22% 9%
1	C	281	70% 23% 7%
1	D	281	58% 16% . 23%
1	E	281	72% 19% 9%
1	F	281	72% 23% 5%
1	G	281	58% 15% 27%
1	H	281	60% 17% 23%
1	I	281	47% 47% . . .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	J	281	
1	K	281	
1	L	281	
1	M	281	
1	N	281	
1	O	281	
1	P	281	
2	Q	21	
2	S	21	
3	R	50	
3	T	50	
4	U	23	
4	V	23	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 34774 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 integrase.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	276	Total	C	N	O	S	0	0
			2216	1411	387	409	9		
1	B	256	Total	C	N	O	S	0	0
			2075	1321	363	383	8		
1	C	261	Total	C	N	O	S	0	0
			2123	1353	371	391	8		
1	D	216	Total	C	N	O	S	0	0
			1729	1110	297	316	6		
1	E	257	Total	C	N	O	S	0	0
			2090	1333	365	384	8		
1	F	266	Total	C	N	O	S	0	0
			2140	1362	377	393	8		
1	G	205	Total	C	N	O	S	0	0
			1657	1051	284	314	8		
1	H	217	Total	C	N	O	S	0	0
			1761	1129	308	318	6		
1	I	276	Total	C	N	O	S	0	0
			2216	1411	387	409	9		
1	J	256	Total	C	N	O	S	0	0
			2075	1321	363	383	8		
1	K	261	Total	C	N	O	S	0	0
			2123	1353	371	391	8		
1	L	216	Total	C	N	O	S	0	0
			1729	1110	297	316	6		
1	M	257	Total	C	N	O	S	0	0
			2090	1333	365	384	8		
1	N	266	Total	C	N	O	S	0	0
			2140	1362	377	393	8		
1	O	205	Total	C	N	O	S	0	0
			1657	1051	284	314	8		
1	P	217	Total	C	N	O	S	0	0
			1761	1129	308	318	6		

- Molecule 2 is a DNA chain called vDNA, non-transferred strand.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	Q	21	Total	C	N	O	P	0	0
			431	203	79	128	21		
2	S	21	Total	C	N	O	P	0	0
			431	203	79	128	21		

- Molecule 3 is a DNA chain called vDNA-tDNA, transferred strand, joined to a model tDNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	R	41	Total	C	N	O	P	0	0
			834	396	153	244	41		
3	T	41	Total	C	N	O	P	0	0
			834	396	153	244	41		

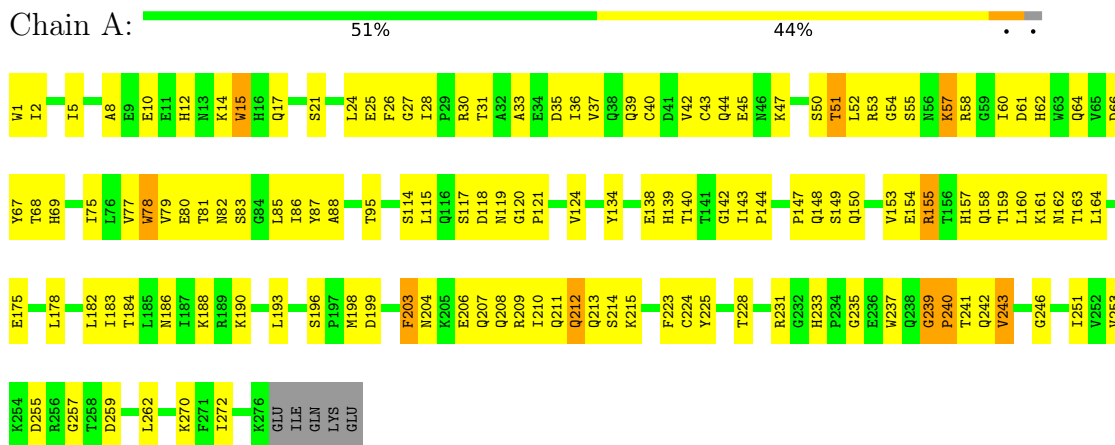
- Molecule 4 is a DNA chain called tDNA.

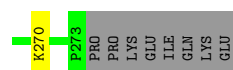
Mol	Chain	Residues	Atoms					AltConf	Trace
4	U	16	Total	C	N	O	P	0	0
			331	159	63	94	15		
4	V	16	Total	C	N	O	P	0	0
			331	159	63	94	15		

3 Residue-property plots [i](#)

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.

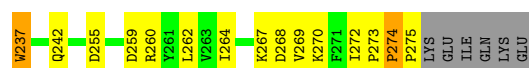
- Molecule 1: integrase





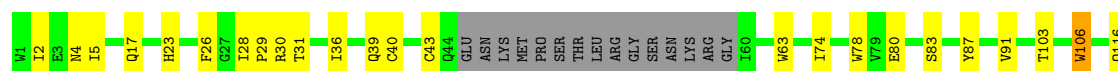
- Molecule 1: integrase

Chain D: 58% 16% 23%



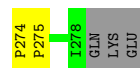
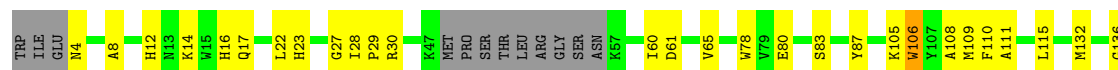
- Molecule 1: integrase

Chain E: 72% 19% 9%



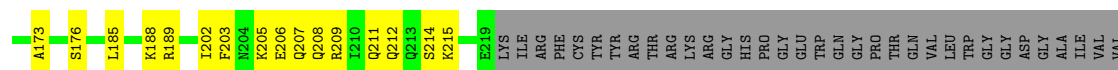
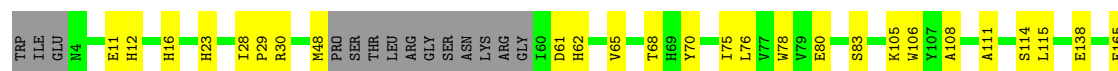
- Molecule 1: integrase

Chain F: 72% 23% 5%



- Molecule 1: integrase

Chain G: 58% 15% 27%



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

- Molecule 1: integrase

Chain H:  60% 17% 23%

TRP
ILE
GLU
ASN
ILE
PRO
LEU
ALA
GLU
GLU
HIS
HIS
LYS
TRP
HIS
GLN
ASP
ALA
VAL
SER
LEU
HIS
LEU
PHE
GLY
ILE
PRO
ARG
THR
ALA
ALA
GLU
ASP
ILE
VAL
GLN
GLN
CYS
ASP
VAL
CYS
GLN
ASN
LYS
MET
PRO
SER
THR
LEU
ARG
GLY
SER
ASN
LYS
ARG
GLY
I60

Q64
T68
I75
V79
E80
S83
E89
G93
E98
K105
W106
A108
M109
F110
L115
N119
T140
T141
G142
W145
N146
P147
Q148
S149
Q150
L152
T156
L164
A173
S176
A177
L178
T181
T184
L185
R189
I202

F203
N204
Q207
Q208
R209
T210
Q211
Q212
F223
Y226
R227
T228
G235
Q238
V252
V263
P275
K276
GLU
ILE
GLN
LYS
GLU

- Molecule 1: integrase

Chain I:  47% 47% . . .

V1
I2
I5
A8
E9
E10
E11
H12
N13
N13
K14
W15
H16
Q17
S21
F26
G27
I28
P29
R30
A33
I36
V37
Q38
Q39
C40
P41
V42
C43
K47
W48
P49
S50
T51
R53
G54
S55
N56
K57
R58
D61
H62
W63
Q64
V65
D66
Y67
T68
H69
I74
I75

L76
V77
W78
V79
E80
T81
H82
S83
G84
L85
H86
Y87
A88
E94
T95
F99
A108
A111
P112
K113
S114
L115
Q116
S117
D118
N119
Q120
P121
A122
F123
V124
A125
E126
E138
H139
L140
T141
G142
I143
P144
W145
N146
P147
Q148
S149
Q150
V153
E154
R155
T156
H157
Q158
T159
L160

K161
N162
T163
E175
L176
L182
N183
T184
I86
Y87
N185
N186
I187
K188
R189
K190
L193
S196
P197
M198
D199
F203
N204
K205
E206
Q207
Q208
R209
Q212
Q213
S214
K215
S216
K217
I221
R222
F223
C224
Y226
R227
T228
R231
G232
H233
P234
G235
E236
Q237
Q238
G239
P240
T241

Q242
V243
G246
I251
V252
V253
K254
D255
R256
Q257
T258
D259
L262
V263
I264
V269
K270
I271
I272
P273
K276
GLU
ILE
GLN
LYS
GLU

- Molecule 1: integrase

Chain J:  69% 21% 9%

TRP
ILE
E3
H23
I28
P29
A33
I36
C40
Q44
GLU
ASN
LYS
MET
PRO
SER
THR
LEU
ARG
GLY
ILE
ASN
LYS
ARG
GLY
I60
D61
H62
W63
Q64
T68
H69
I75
W78
V79
E80
S83
G84
L85
I86
A89
E89
R90
V91
K92
G93
M104
K105

W106
T107
A108
S114
L115
Y134
L135
E138
T159
L160
K161
T163
A173
S176
A177
L178
A179
G180
T181
T184
I187
K190
T195
M198
D199
I200
Q208
R209
I210
Q211
Q212
K215
S216
K217
Q218
R222
W245
G246
G247
A250
I251
V252
V253

L262
P273
PRO
PRO
LYS
GLU
GLN
LYS
GLU

- Molecule 1: integrase

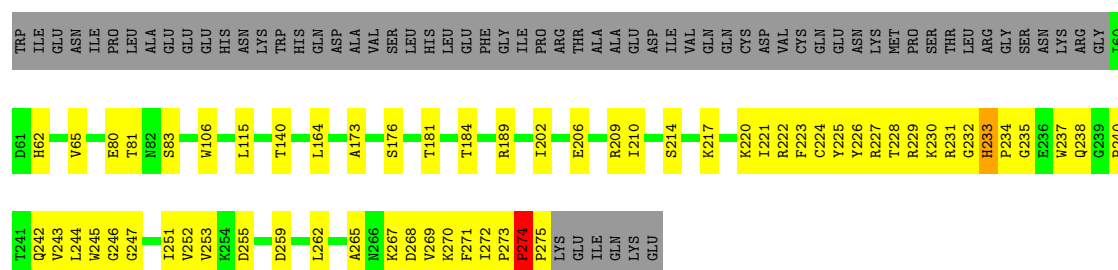
Chain K:  71% 22% 7%

V1
I2
E3
N4
I5
P6
L7
A8
E9
S21
L22
H23
F26
G27
I28
P29
R30
T31
A32
I36
C40
K47
MET
PRO
SER
THR
LEU
ARG
GLY
ILE
ASN
LYS
ARG
GLY
I60
D61
H62
W63
Q64
V65
W78
V79
E80
S83
W106
S114
L115
T140
S149



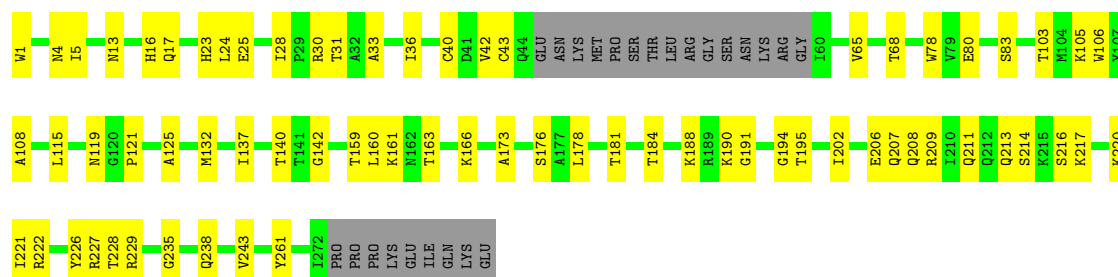
- Molecule 1: integrase

Chain L: 55% 21% 23%



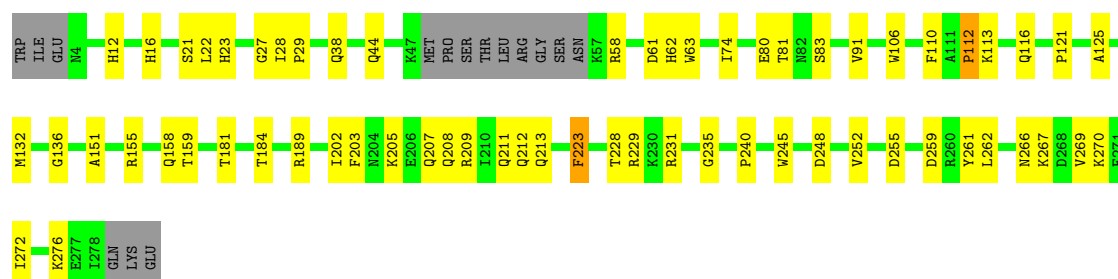
- Molecule 1: integrase

Chain M: 67% 25% 9%



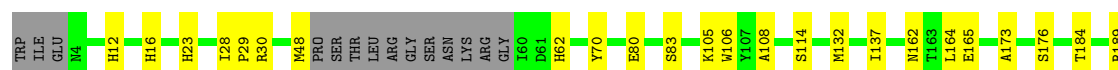
- Molecule 1: integrase

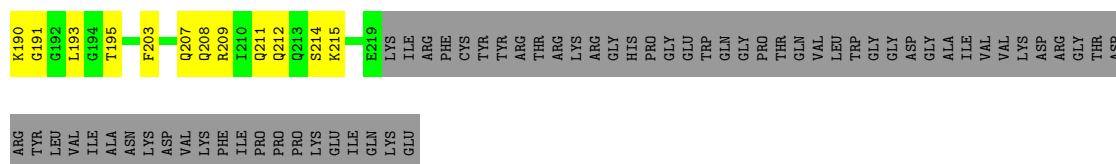
Chain N: 72% 22% 5%



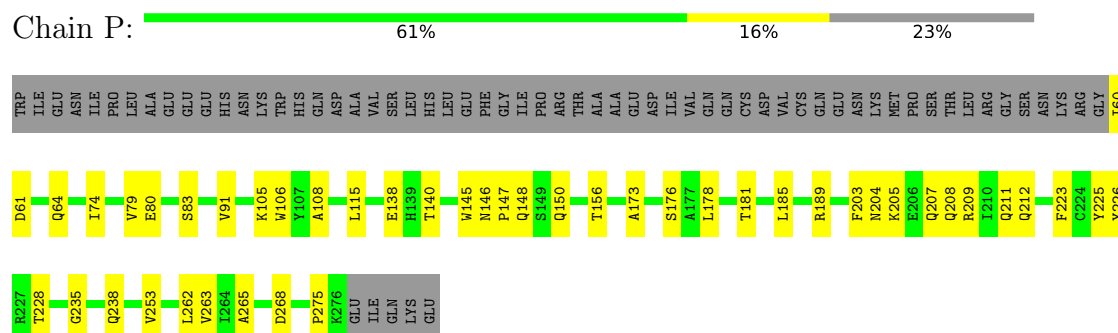
- Molecule 1: integrase

Chain O: 60% 13% 27%

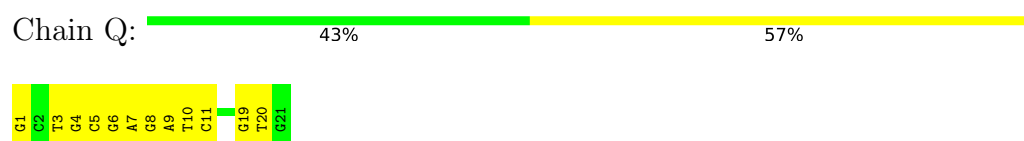




• Molecule 1: integrase



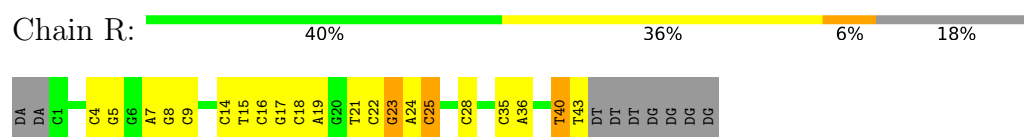
• Molecule 2: vDNA, non-transferred strand



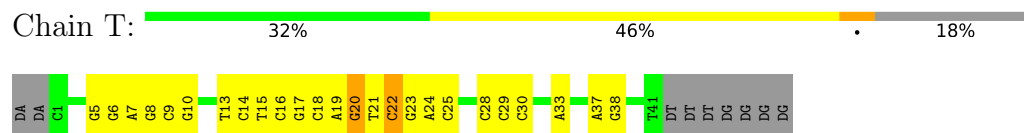
• Molecule 2: vDNA, non-transferred strand



• Molecule 3: vDNA-tDNA, transferred strand, joined to a model tDNA

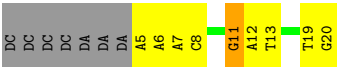


• Molecule 3: vDNA-tDNA, transferred strand, joined to a model tDNA

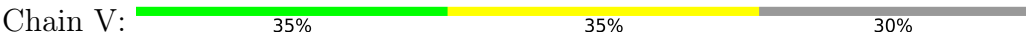


• Molecule 4: tDNA





● Molecule 4: tDNA



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C2	Depositor
Number of particles used	37021	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI POLARA 300	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	1.47	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

5 Model quality ⓘ

5.1 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 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	0.83	4/2269 (0.2%)	0.85	7/3071 (0.2%)
1	B	0.58	1/2124 (0.0%)	0.73	2/2874 (0.1%)
1	C	0.48	1/2174 (0.0%)	0.65	0/2942
1	D	0.57	3/1772 (0.2%)	0.79	4/2401 (0.2%)
1	E	0.53	1/2140 (0.0%)	0.67	0/2896
1	F	0.62	2/2191 (0.1%)	0.72	0/2966
1	G	0.43	1/1693 (0.1%)	0.60	0/2290
1	H	0.47	1/1805 (0.1%)	0.66	0/2441
1	I	0.89	6/2271 (0.3%)	1.22	12/3077 (0.4%)
1	J	0.59	2/2124 (0.1%)	0.74	1/2874 (0.0%)
1	K	0.47	1/2174 (0.0%)	0.65	1/2942 (0.0%)
1	L	0.66	3/1772 (0.2%)	0.68	1/2401 (0.0%)
1	M	0.54	0/2140	0.67	0/2896
1	N	0.61	0/2191	0.71	1/2966 (0.0%)
1	O	0.43	1/1693 (0.1%)	0.60	0/2290
1	P	0.49	1/1805 (0.1%)	0.68	0/2441
2	Q	0.52	0/482	0.85	0/742
2	S	0.75	3/482 (0.6%)	1.07	4/742 (0.5%)
3	R	1.38	3/934 (0.3%)	0.96	3/1437 (0.2%)
3	T	0.66	2/934 (0.2%)	0.98	4/1437 (0.3%)
4	U	0.69	2/372 (0.5%)	0.92	0/574
4	V	0.48	0/372	0.88	0/574
All	All	0.63	38/35914 (0.1%)	0.77	40/49274 (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	A	0	1
1	I	0	3
All	All	0	4

The worst 5 of 38 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	R	40	DT	O3'-P	-38.12	1.15	1.61
1	I	238	GLN	C-N	17.98	1.65	1.33
1	A	239	GLY	C-N	16.57	1.65	1.34
1	L	221	ILE	C-N	14.47	1.67	1.34
1	I	146	ASN	C-N	10.20	1.53	1.34

The worst 5 of 40 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	I	238	GLN	O-C-N	30.56	175.16	123.20
1	I	238	GLN	CA-C-N	-29.87	56.45	116.20
1	I	238	GLN	C-N-CA	-16.82	86.99	122.30
1	D	274	PRO	C-N-CD	-15.44	86.63	120.60
2	S	13	DG	OP1-P-O3'	-10.65	81.76	105.20

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	51	THR	Mainchain
1	I	142	GLY	Peptide
1	I	225	TYR	Mainchain
1	I	51	THR	Mainchain

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	2216	0	2173	407	0
1	B	2075	0	2046	51	0
1	C	2123	0	2091	85	0
1	D	1729	0	1705	152	0
1	E	2090	0	2061	100	0
1	F	2140	0	2106	51	0
1	G	1657	0	1626	57	0
1	H	1761	0	1764	101	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	I	2216	0	2175	484	0
1	J	2075	0	2046	116	0
1	K	2123	0	2095	77	0
1	L	1729	0	1702	160	0
1	M	2090	0	2059	96	0
1	N	2140	0	2105	107	0
1	O	1657	0	1626	60	0
1	P	1761	0	1764	82	0
2	Q	431	0	236	58	0
2	S	431	0	236	93	0
3	R	834	0	461	127	0
3	T	834	0	460	150	0
4	U	331	0	183	12	0
4	V	331	0	182	36	0
All	All	34774	0	32902	1839	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 27.

The worst 5 of 1839 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:A:212:GLN:HB2	1:D:217:LYS:CD	1.21	1.64
1:A:14:LYS:HB3	1:J:134:TYR:CE1	1.27	1.61
1:I:145:TRP:CB	1:N:229:ARG:HD2	1.28	1.61
1:C:270:LYS:HD3	1:H:145:TRP:CE2	1.31	1.59
1:I:119:ASN:HB2	4:V:19:DT:C4'	1.25	1.59

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	270/281 (96%)	253 (94%)	16 (6%)	1 (0%)	36	77
1	B	252/281 (90%)	242 (96%)	8 (3%)	2 (1%)	21	65
1	C	257/281 (92%)	249 (97%)	8 (3%)	0	100	100
1	D	214/281 (76%)	203 (95%)	8 (4%)	3 (1%)	12	52
1	E	253/281 (90%)	244 (96%)	8 (3%)	1 (0%)	36	77
1	F	262/281 (93%)	248 (95%)	12 (5%)	2 (1%)	21	65
1	G	201/281 (72%)	197 (98%)	4 (2%)	0	100	100
1	H	215/281 (76%)	211 (98%)	4 (2%)	0	100	100
1	I	274/281 (98%)	255 (93%)	18 (7%)	1 (0%)	36	77
1	J	252/281 (90%)	237 (94%)	15 (6%)	0	100	100
1	K	257/281 (92%)	253 (98%)	4 (2%)	0	100	100
1	L	214/281 (76%)	203 (95%)	9 (4%)	2 (1%)	19	61
1	M	253/281 (90%)	244 (96%)	9 (4%)	0	100	100
1	N	262/281 (93%)	252 (96%)	9 (3%)	1 (0%)	36	77
1	O	201/281 (72%)	194 (96%)	7 (4%)	0	100	100
1	P	215/281 (76%)	209 (97%)	6 (3%)	0	100	100
All	All	3852/4496 (86%)	3694 (96%)	145 (4%)	13 (0%)	47	81

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	231	ARG
1	D	233	HIS
1	I	240	PRO
1	L	233	HIS
1	D	234	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	237/246 (96%)	236 (100%)	1 (0%)	92	95
1	B	223/246 (91%)	223 (100%)	0	100	100
1	C	228/246 (93%)	228 (100%)	0	100	100
1	D	183/246 (74%)	183 (100%)	0	100	100
1	E	224/246 (91%)	224 (100%)	0	100	100
1	F	228/246 (93%)	228 (100%)	0	100	100
1	G	181/246 (74%)	181 (100%)	0	100	100
1	H	189/246 (77%)	189 (100%)	0	100	100
1	I	237/246 (96%)	237 (100%)	0	100	100
1	J	223/246 (91%)	223 (100%)	0	100	100
1	K	228/246 (93%)	228 (100%)	0	100	100
1	L	183/246 (74%)	183 (100%)	0	100	100
1	M	224/246 (91%)	224 (100%)	0	100	100
1	N	228/246 (93%)	228 (100%)	0	100	100
1	O	181/246 (74%)	181 (100%)	0	100	100
1	P	189/246 (77%)	189 (100%)	0	100	100
All	All	3386/3936 (86%)	3385 (100%)	1 (0%)	100	100

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	212	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	3
1	I	2
1	L	1
3	R	1

The worst 5 of 7 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	44:GLN	C	45:GLU	N	2.19
1	A	224:CYS	C	225:TYR	N	2.02
1	L	221:ILE	C	222:ARG	N	1.67
1	A	239:GLY	C	240:PRO	N	1.65
1	I	238:GLN	C	239:GLY	N	1.65