Full wwPDB/EMDataBank EM Map/Model Validation Report (i)

Oct 29, 2019 – 03:56 PM EDT

PDB ID	:	6SEE
EMDB ID:	:	EMD-10153
Title	:	Class2A : CENP-A nucleosome in complex with CENP-C central region
Authors	:	Ali-Ahmad, A.; Bilokapic, S.; Schafer, I.B.; Halic, M.; Sekulic, N.
Deposited on	:	2019-07-29
Resolution	:	4.20 Å(reported)

This is a Full wwPDB/EMDataBank EM Map/Model Validation 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.

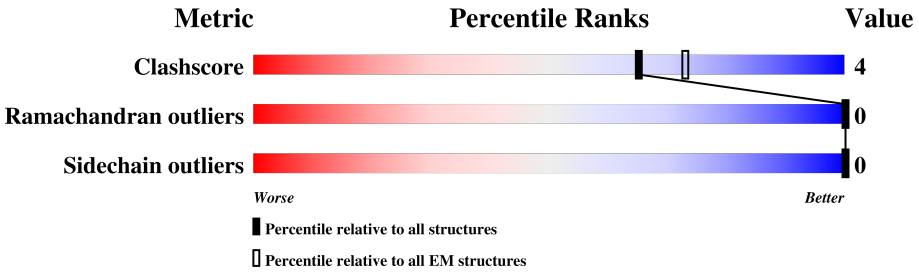
MolProbity	:	4.02b-467
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.4

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 4.20 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 <=5%

Mol	Chain	Length	Quality of chain
1	A	140	60% 10% 30%
1	E	140	64% 6% 30%
2	B	103	61% 17% 22%
2	F	103	68% 8% 24%
3	C	130	62% 12% 27%
3	G	130	71% 12% 18%
4	D	126	64% 9% 27%
4	H	126	61% 10% 29%
5	I	145	88% 11% .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	J	145	
7	V	124	

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 11949 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 Histone H3-like centromeric protein A.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	98	Total	C	N	O	S	0	0
			812	524	155	132	1		
1	E	98	Total	C	N	O	S	0	0
			812	524	155	132	1		

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	80	Total	C	N	O	S	0	0
			637	401	125	110	1		
2	F	78	Total	C	N	O	S	0	0
			618	391	120	106	1		

- Molecule 3 is a protein called Histone H2A type 2-A.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	95	Total	C	N	O	S	0	0
			731	458	143	129	1		
3	G	107	Total	C	N	O	S	0	0
			822	518	161	142	1		

- Molecule 4 is a protein called Histone H2B type 1-C/E/F/G/I.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	92	Total	C	N	O	S	0	0
			719	452	129	136	2		
4	H	90	Total	C	N	O	S	0	0
			699	440	123	134	2		

- Molecule 5 is a DNA chain called DNA (145-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
5	I	145	Total	C	N	O	P	0	0
			2952	1404	537	867	144		

- Molecule 6 is a DNA chain called DNA (145-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
6	J	145	Total	C	N	O	P	0	0
			2987	1416	558	869	144		

- Molecule 7 is a protein called Centromere protein C.

Mol	Chain	Residues	Atoms				AltConf	Trace
7	V	18	Total	C	N	O	0	0
			160	100	34	26		

There are 12 discrepancies between the modelled and reference sequences:

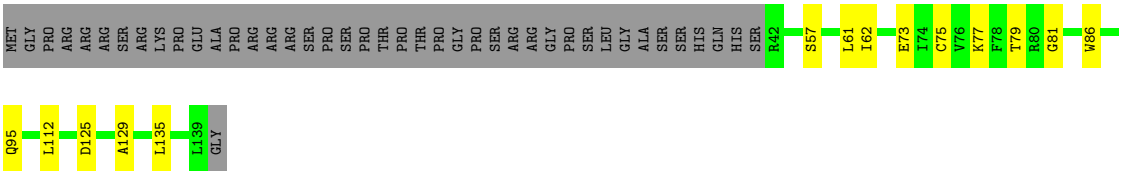
Chain	Residue	Modelled	Actual	Comment	Reference
V	414	GLY	-	expression tag	UNP Q03188
V	415	PRO	-	expression tag	UNP Q03188
V	416	LEU	-	expression tag	UNP Q03188
V	417	GLY	-	expression tag	UNP Q03188
V	418	SER	-	expression tag	UNP Q03188
V	419	PRO	-	expression tag	UNP Q03188
V	420	GLU	-	expression tag	UNP Q03188
V	421	PHE	-	expression tag	UNP Q03188
V	422	GLY	-	expression tag	UNP Q03188
V	423	ARG	-	expression tag	UNP Q03188
V	424	ALA	-	expression tag	UNP Q03188
V	425	THR	-	expression tag	UNP Q03188

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: Histone H3-like centromeric protein A

Chain A:

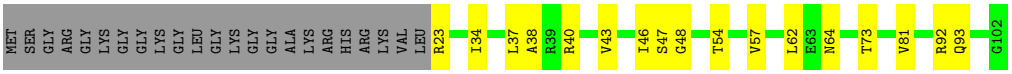


• Molecule 1: Histone H3-like centromeric protein A

Chain E:

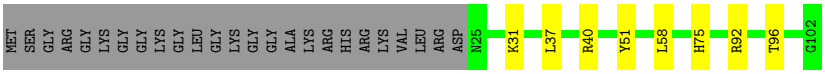
• Molecule 2: Histone H4

Chain B:



• Molecule 2: Histone H4

Chain F:

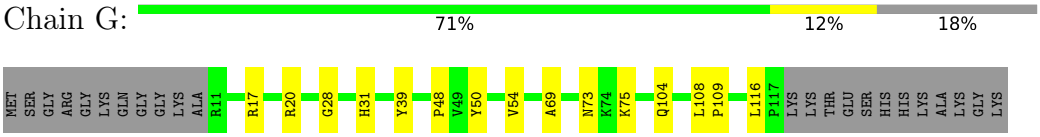


• Molecule 3: Histone H2A type 2-A

Chain C:

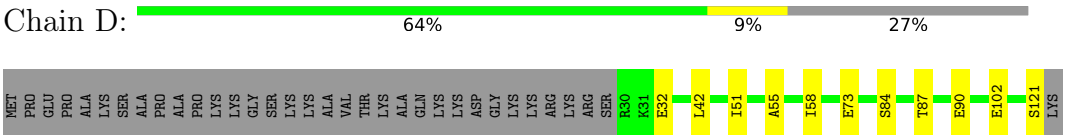
• Molecule 3: Histone H2A type 2-A

Chain G:



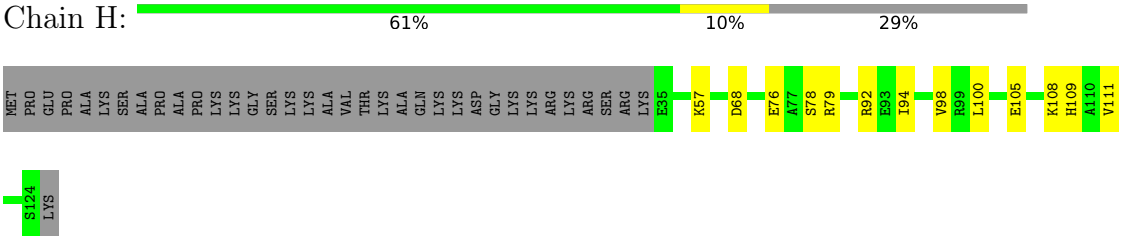
• Molecule 4: Histone H2B type 1-C/E/F/G/I

Chain D:



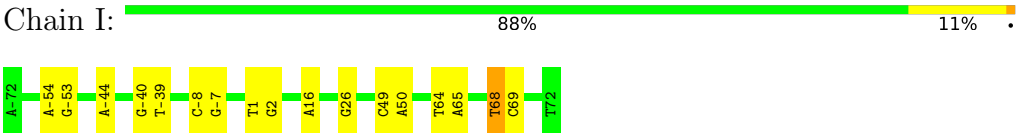
• Molecule 4: Histone H2B type 1-C/E/F/G/I

Chain H:



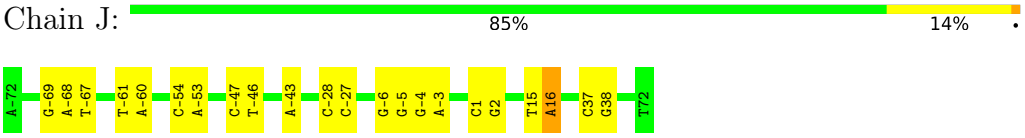
• Molecule 5: DNA (145-MER)

Chain I:



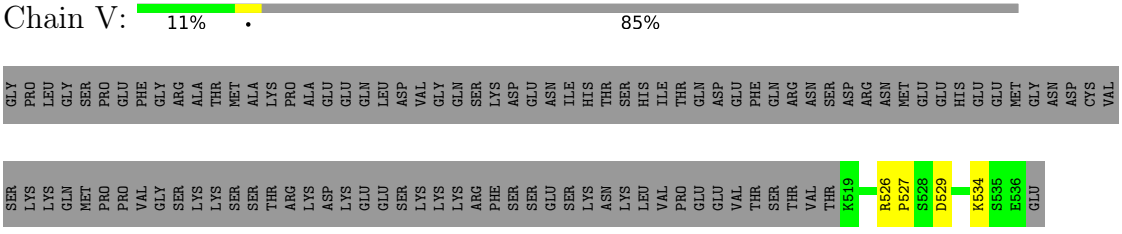
• Molecule 6: DNA (145-MER)

Chain J:



• Molecule 7: Centromere protein C

Chain V:



4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	11000	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 ($e^-/Å^2$)	100	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	0.34	0/827	0.57	0/1111
1	E	0.33	0/827	0.51	0/1111
2	B	0.35	0/644	0.56	0/862
2	F	0.33	0/625	0.52	0/837
3	C	0.33	0/740	0.51	0/997
3	G	0.31	0/832	0.59	1/1121 (0.1%)
4	D	0.32	0/730	0.51	0/982
4	H	0.34	0/710	0.51	1/957 (0.1%)
5	I	0.92	0/3308	1.04	2/5099 (0.0%)
6	J	0.91	0/3354	1.02	1/5180 (0.0%)
7	V	0.32	0/164	0.69	0/219
All	All	0.70	0/12761	0.85	5/18476 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	G	116	LEU	CA-CB-CG	6.28	129.74	115.30
5	I	26	DG	O4'-C1'-N9	-5.88	103.88	108.00
5	I	68	DT	O4'-C1'-N1	5.78	112.04	108.00
4	H	68	ASP	CB-CG-OD2	5.22	123.00	118.30
6	J	16	DA	O4'-C1'-N9	5.05	111.54	108.00

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	812	0	854	11	0
1	E	812	0	854	8	0
2	B	637	0	676	14	0
2	F	618	0	659	8	0
3	C	731	0	769	14	0
3	G	822	0	880	10	0
4	D	719	0	738	9	0
4	H	699	0	712	9	0
5	I	2952	0	1629	9	0
6	J	2987	0	1630	12	0
7	V	160	0	163	3	0
All	All	11949	0	9564	80	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 4.

All (80) 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:E:130:ARG:HH11	1:E:139:LEU:HA	1.56	0.69
3:G:50:TYR:HB3	4:H:94:ILE:HD13	1.82	0.62
2:B:73:THR:HG21	2:B:81:VAL:HG12	1.83	0.61
2:B:64:ASN:OD1	2:B:93:GLN:NE2	2.35	0.60
2:B:23:ARG:NH1	5:I:16:DA:OP1	2.35	0.60
3:G:104:GLN:OE1	4:H:57:LYS:NZ	2.35	0.59
3:C:89:ASN:O	7:V:526:ARG:NH1	2.36	0.58
3:C:100:VAL:HG12	2:F:96:THR:HB	1.85	0.58
1:A:62:ILE:O	1:A:95:GLN:NE2	2.38	0.57
1:A:135:LEU:HD11	3:C:95:LYS:HG3	1.86	0.56
3:C:20:ARG:HE	4:D:121:SER:HB3	1.71	0.56
4:D:55:ALA:HA	4:D:58:ILE:HD12	1.88	0.56
3:G:54:VAL:HG21	4:H:98:VAL:HG21	1.87	0.55
1:A:79:THR:HG23	1:A:81:GLY:H	1.72	0.55
3:G:28:GLY:HA3	5:I:-44:DA:H3'	1.89	0.55
1:E:61:LEU:HD11	2:F:40:ARG:HE	1.73	0.54
6:J:-61:DT:H2'	6:J:-60:DA:C8	2.43	0.54
3:C:81:ARG:HB2	1:E:58:THR:HG21	1.89	0.54
4:H:105:GLU:OE1	4:H:109:HIS:NE2	2.41	0.54
3:C:32:ARG:NH2	4:D:32:GLU:OE2	2.42	0.53
1:A:61:LEU:HD12	2:B:37:LEU:HD23	1.90	0.52
3:G:39:TYR:HB3	4:H:78:SER:HB3	1.90	0.52
1:A:75:CYS:HB2	1:A:86:TRP:HZ2	1.75	0.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:I:-8:DC:H2”	5:I:-7:DG:C8	2.45	0.52
2:F:75:HIS:O	4:H:92:ARG:NH2	2.37	0.51
1:E:62:ILE:O	1:E:95:GLN:NE2	2.42	0.51
3:G:69:ALA:O	3:G:73:ASN:ND2	2.35	0.51
1:E:61:LEU:HD21	2:F:40:ARG:HH21	1.77	0.50
2:B:92:ARG:NE	4:D:73:GLU:OE1	2.45	0.50
2:F:92:ARG:HG3	4:H:100:LEU:HD21	1.92	0.50
1:E:102:LEU:HD11	2:F:58:LEU:HD13	1.93	0.49
3:C:16:SER:HB2	6:J:-43:DA:H4’	1.94	0.49
5:I:68:DT:H2”	5:I:69:DC:C5	2.48	0.49
3:C:78:ILE:HB	4:D:51:ILE:HD13	1.95	0.49
1:A:125:ASP:OD1	1:E:115:HIS:NE2	2.40	0.48
1:A:62:ILE:HD11	2:B:37:LEU:HD11	1.96	0.47
1:A:57:SER:OG	2:B:40:ARG:NH2	2.47	0.47
2:B:47:SER:OG	2:B:48:GLY:N	2.49	0.46
3:C:88:ARG:HA	3:C:94:ASN:HD21	1.81	0.46
2:B:34:ILE:HA	2:B:37:LEU:HD12	1.98	0.46
2:F:31:LYS:HG3	2:F:51:TYR:CZ	2.52	0.44
3:G:108:LEU:HD12	3:G:109:PRO:HD2	1.99	0.44
3:C:41:GLU:HB2	4:D:84:SER:HB2	1.99	0.44
5:I:1:DT:H2”	5:I:2:DG:C8	2.53	0.44
6:J:15:DT:H2”	6:J:16:DA:C8	2.53	0.44
6:J:-69:DG:H2”	6:J:-68:DA:C8	2.52	0.44
6:J:-4:DG:H2”	6:J:-3:DA:C8	2.52	0.44
1:A:77:LYS:HE3	1:A:77:LYS:HB3	1.78	0.44
7:V:529:ASP:H	7:V:534:LYS:NZ	2.16	0.44
3:G:17:ARG:HA	3:G:20:ARG:HG2	1.99	0.43
4:H:108:LYS:HA	4:H:111:VAL:HG12	1.99	0.43
3:C:63:LEU:HD13	4:D:42:LEU:HB2	2.00	0.43
4:D:102:GLU:H	4:D:102:GLU:HG3	1.65	0.43
6:J:-6:DG:H2”	6:J:-5:DG:C8	2.53	0.43
5:I:49:DC:H2’	5:I:50:DA:C8	2.54	0.43
1:E:61:LEU:HD12	2:F:37:LEU:HD23	2.01	0.42
5:I:-54:DA:H2”	5:I:-53:DG:C8	2.54	0.42
6:J:-54:DC:H2”	6:J:-53:DA:C8	2.54	0.42
6:J:-68:DA:H2”	6:J:-67:DT:H71	2.00	0.42
2:B:92:ARG:HA	2:B:92:ARG:HD2	1.84	0.42
3:C:88:ARG:HB3	3:C:108:LEU:HD11	2.00	0.42
2:B:62:LEU:HD12	2:B:62:LEU:HA	1.81	0.42
4:H:76:GLU:OE1	4:H:79:ARG:NH1	2.53	0.42
3:C:87:ILE:HD13	3:C:97:LEU:HD22	2.01	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:73:GLU:O	1:A:77:LYS:HB2	2.19	0.41
2:B:38:ALA:HB1	2:B:43:VAL:HB	2.02	0.41
5:I:64:DT:H2"	5:I:65:DA:H8	1.85	0.41
4:D:87:THR:OG1	4:D:90:GLU:OE1	2.26	0.41
2:B:54:THR:HA	2:B:57:VAL:HG22	2.03	0.41
6:J:-28:DC:H2"	6:J:-27:DC:C6	2.56	0.41
7:V:527:PRO:HB3	7:V:534:LYS:HD3	2.03	0.41
3:G:73:ASN:O	3:G:75:LYS:NZ	2.47	0.41
3:C:20:ARG:HA	3:C:20:ARG:HD2	1.90	0.41
3:G:31:HIS:HD2	3:G:48:PRO:HG3	1.86	0.41
6:J:1:DC:H2"	6:J:2:DG:C8	2.55	0.41
2:B:46:ILE:HD12	2:B:46:ILE:HA	1.89	0.40
1:A:112:LEU:HD21	1:A:129:ALA:HB2	2.03	0.40
5:I:-40:DG:H2"	5:I:-39:DT:H71	2.04	0.40
6:J:-47:DC:H2"	6:J:-46:DT:C5	2.56	0.40
6:J:37:DC:H2"	6:J:38:DG:C8	2.56	0.40

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	96/140 (69%)	92 (96%)	4 (4%)	0	100	100
1	E	96/140 (69%)	94 (98%)	2 (2%)	0	100	100
2	B	78/103 (76%)	75 (96%)	3 (4%)	0	100	100
2	F	76/103 (74%)	73 (96%)	3 (4%)	0	100	100
3	C	93/130 (72%)	88 (95%)	5 (5%)	0	100	100
3	G	105/130 (81%)	103 (98%)	2 (2%)	0	100	100
4	D	90/126 (71%)	89 (99%)	1 (1%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	H	88/126 (70%)	87 (99%)	1 (1%)	0	100	100
7	V	16/124 (13%)	15 (94%)	1 (6%)	0	100	100
All	All	738/1122 (66%)	716 (97%)	22 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	83/118 (70%)	83 (100%)	0	100	100
1	E	83/118 (70%)	83 (100%)	0	100	100
2	B	65/79 (82%)	65 (100%)	0	100	100
2	F	63/79 (80%)	63 (100%)	0	100	100
3	C	74/99 (75%)	74 (100%)	0	100	100
3	G	83/99 (84%)	83 (100%)	0	100	100
4	D	79/106 (74%)	79 (100%)	0	100	100
4	H	77/106 (73%)	77 (100%)	0	100	100
7	V	18/115 (16%)	18 (100%)	0	100	100
All	All	625/919 (68%)	625 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	64	ASN
2	B	93	GLN
3	C	94	ASN
3	G	31	HIS
3	G	112	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	H	49	HIS
4	H	63	ASN

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)

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)

There are no chain breaks in this entry.