Full wwPDB X-ray Structure Validation Report (i)

Jan 31, 2016 – 08:50 PM GMT

PDB ID	:	1M9C
Title	:	X-ray crystal structure of Cyclophilin A/HIV-1 CA N-terminal domain (1-146) M-type Complex.
Authors	:	Howard, B.R.; Vajdos, F.F.; Li, S.; Sundquist, W.I.; Hill, C.P.
Deposited on	:	2002-07-28
Resolution	:	2.00 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

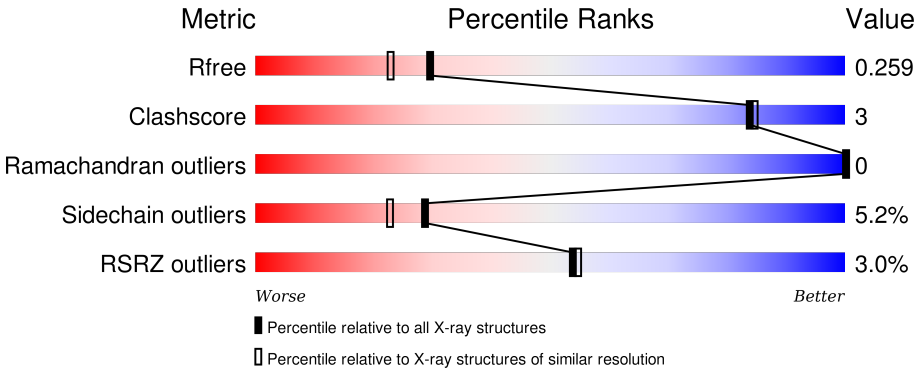
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	6249 (2.00-2.00)
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)
RSRZ outliers	91569	6262 (2.00-2.00)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower 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% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	165	90% 10%
1	B	165	85% 11%
2	C	146	3% 74% 23%
2	D	146	8% 73% 17% 7%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4856 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, 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 Cyclophilin A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	165	Total	C	N	O	S	0	0	0
			1266	802	218	237	9			
1	B	163	Total	C	N	O	S	0	0	0
			1251	792	216	235	8			

- Molecule 2 is a protein called HIV-1 Capsid.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	146	Total	C	N	O	S	15	0	0
			1136	718	200	210	8			
2	D	136	Total	C	N	O	S	10	0	0
			1059	670	186	196	7			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	120	HIS	ASN	see remark 999	UNP Q72497
D	120	HIS	ASN	see remark 999	UNP Q72497

- Molecule 3 is water.

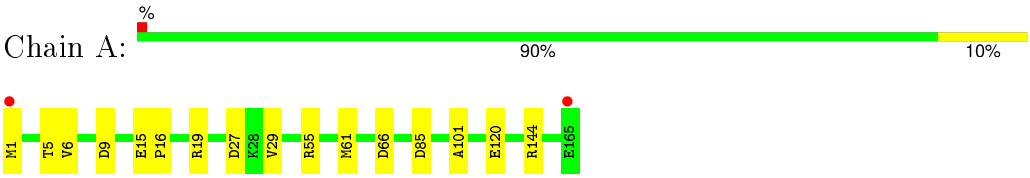
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	59	Total	O	0	0
			59	59		
3	B	30	Total	O	0	0
			30	30		
3	C	40	Total	O	0	0
			40	40		
3	D	15	Total	O	0	0
			15	15		

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 errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). 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.

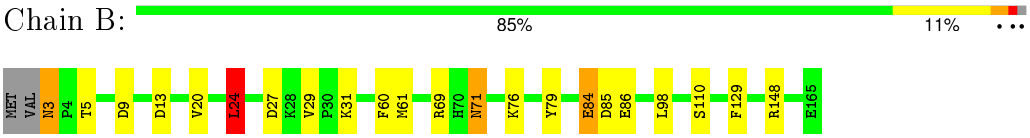
- Molecule 1: Cyclophilin A

Chain A:



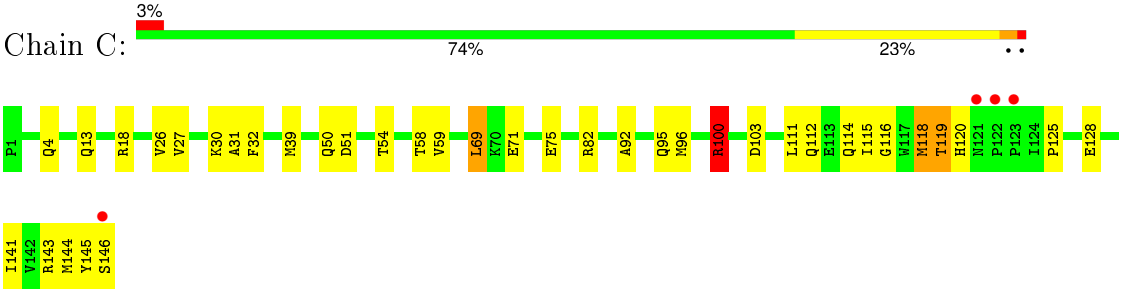
- Molecule 1: Cyclophilin A

Chain B:



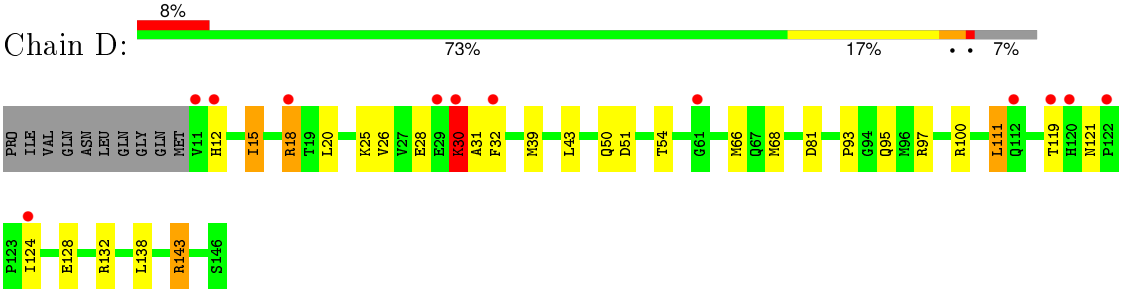
- Molecule 2: HIV-1 Capsid

Chain C:



- Molecule 2: HIV-1 Capsid

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	38.45Å 113.23Å 67.01Å 90.00° 100.53° 90.00°	Depositor
Resolution (Å)	37.80 – 2.00 32.75 – 2.00	Depositor EDS
% Data completeness (in resolution range)	83.0 (37.80-2.00) 83.0 (32.75-2.00)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	4.45 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5	Depositor
R, R_{free}	0.195 , 0.259 0.202 , 0.259	Depositor DCC
R_{free} test set	3125 reflections (10.99%)	DCC
Wilson B-factor (Å^2)	26.5	Xtriage
Anisotropy	0.742	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 36.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Outliers	0 of 31562 reflections	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4856	wwPDB-VP
Average B, all atoms (Å^2)	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.12% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

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	#$\|Z\|>5$	RMSZ	#$\|Z\|>5$
1	A	0.93	2/1294 (0.2%)	1.44	10/1733 (0.6%)
1	B	0.81	1/1279 (0.1%)	1.37	13/1713 (0.8%)
2	C	0.92	3/1164 (0.3%)	1.55	14/1583 (0.9%)
2	D	0.99	3/1086 (0.3%)	1.62	14/1478 (0.9%)
All	All	0.91	9/4823 (0.2%)	1.50	51/6507 (0.8%)

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
2	D	0	1

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	28	GLU	CA-CB	-14.65	1.21	1.53
2	D	31	ALA	C-N	14.07	1.66	1.34
2	D	30	LYS	C-N	12.58	1.62	1.34
2	C	112	GLN	CA-CB	11.46	1.79	1.53
2	C	4	GLN	CA-CB	-8.40	1.35	1.53
1	A	6	VAL	CB-CG2	-5.30	1.41	1.52
1	B	79	TYR	CD2-CE2	-5.18	1.31	1.39
1	A	101	ALA	CA-CB	-5.09	1.41	1.52
2	C	118	MET	SD-CE	-5.03	1.49	1.77

All (51) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	30	LYS	O-C-N	-25.68	81.61	122.70
2	C	100	ARG	NE-CZ-NH1	14.62	127.61	120.30
2	D	100	ARG	NE-CZ-NH1	13.43	127.02	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	100	ARG	NE-CZ-NH2	-11.64	114.48	120.30
2	D	100	ARG	NE-CZ-NH2	-10.18	115.21	120.30
1	A	144	ARG	NE-CZ-NH1	9.83	125.22	120.30
2	D	30	LYS	CA-C-N	-9.70	95.87	117.20
1	B	9	ASP	CB-CG-OD2	9.22	126.60	118.30
1	A	66	ASP	CB-CG-OD2	9.15	126.54	118.30
2	C	51	ASP	CB-CG-OD2	8.66	126.09	118.30
2	C	112	GLN	N-CA-CB	-8.17	95.90	110.60
1	B	85	ASP	CB-CG-OD2	7.95	125.45	118.30
2	D	28	GLU	N-CA-CB	7.68	124.43	110.60
1	A	9	ASP	CB-CG-OD2	7.63	125.17	118.30
1	A	19	ARG	NE-CZ-NH2	7.62	124.11	120.30
2	C	82	ARG	NE-CZ-NH2	-7.59	116.50	120.30
2	C	103	ASP	CB-CG-OD2	-7.20	111.82	118.30
1	B	27	ASP	CB-CG-OD1	7.13	124.72	118.30
1	B	69	ARG	NE-CZ-NH2	-6.96	116.82	120.30
1	B	13	ASP	CB-CG-OD2	6.92	124.53	118.30
2	D	81	ASP	CB-CG-OD2	6.73	124.36	118.30
1	A	120	GLU	OE1-CD-OE2	6.72	131.36	123.30
2	D	66	MET	CG-SD-CE	-6.45	89.89	100.20
1	A	55	ARG	NE-CZ-NH2	-6.38	117.11	120.30
2	C	118	MET	CG-SD-CE	-6.17	90.33	100.20
1	A	85	ASP	CB-CG-OD2	6.03	123.72	118.30
2	C	18	ARG	NE-CZ-NH2	-5.96	117.32	120.30
2	D	18	ARG	NE-CZ-NH1	5.87	123.23	120.30
1	A	120	GLU	CG-CD-OE2	-5.83	106.65	118.30
2	C	103	ASP	CB-CG-OD1	5.82	123.54	118.30
2	D	68	MET	CG-SD-CE	-5.81	90.90	100.20
2	D	51	ASP	CB-CG-OD1	5.80	123.52	118.30
2	D	132	ARG	NE-CZ-NH1	5.79	123.19	120.30
2	C	13	GLN	N-CA-CB	-5.46	100.77	110.60
1	B	86	GLU	OE1-CD-OE2	-5.43	116.78	123.30
1	A	27	ASP	CB-CG-OD1	5.36	123.13	118.30
1	A	5	THR	OG1-CB-CG2	-5.32	97.77	110.00
2	C	69	LEU	CB-CG-CD2	5.30	120.00	111.00
1	B	148	ARG	NE-CZ-NH1	5.28	122.94	120.30
2	C	143	ARG	NE-CZ-NH2	-5.23	117.68	120.30
1	B	3	ASN	N-CA-C	-5.20	96.97	111.00
1	B	20	VAL	CB-CA-C	-5.18	101.56	111.40
2	D	143	ARG	NE-CZ-NH1	5.17	122.89	120.30
2	C	96	MET	CB-CG-SD	-5.16	96.93	112.40
1	B	24	LEU	CB-CG-CD1	5.14	119.75	111.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	100	ARG	CD-NE-CZ	5.14	130.80	123.60
1	B	5	THR	OG1-CB-CG2	-5.12	98.22	110.00
2	D	18	ARG	CG-CD-NE	5.09	122.48	111.80
2	D	97	ARG	NE-CZ-NH1	5.08	122.84	120.30
1	B	85	ASP	CB-CG-OD1	-5.05	113.76	118.30
1	B	3	ASN	N-CA-CB	-5.04	101.53	110.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	30	LYS	Mainchain

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	1266	0	1237	1	0
1	B	1251	0	1216	8	0
2	C	1136	0	1135	15	0
2	D	1059	0	1052	6	0
3	A	59	0	0	0	0
3	B	30	0	0	0	0
3	C	40	0	0	1	0
3	D	15	0	0	0	0
All	All	4856	0	4640	28	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.

All (28) 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:31:LYS:NZ	1:B:84:GLU:OE2	2.01	0.93
2:D:30:LYS:O	2:D:32:PHE:N	2.03	0.91

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:26:VAL:HG21	2:D:39:MET:HG2	1.73	0.70
2:C:100:ARG:HD3	3:C:155:HOH:O	1.95	0.66
2:D:12:HIS:CE1	2:D:111:LEU:HD11	2.31	0.66
2:C:32:PHE:O	2:C:145:TYR:HB2	2.04	0.58
1:B:3:ASN:HB3	1:B:24:LEU:O	2.03	0.58
2:C:50:GLN:HE22	2:C:111:LEU:HD22	1.69	0.57
2:C:54:THR:O	2:C:58:THR:HG23	2.04	0.57
2:C:69:LEU:HB2	2:C:141:ILE:HD11	1.88	0.56
1:B:98:LEU:HG	1:B:129:PHE:CZ	2.45	0.51
2:C:27:VAL:HG11	2:C:59:VAL:HG13	1.94	0.49
2:C:26:VAL:HG21	2:C:39:MET:HG2	1.95	0.49
2:C:115:ILE:O	2:C:119:THR:HB	2.13	0.48
2:C:114:GLN:O	2:C:118:MET:HE3	2.13	0.47
1:B:60:PHE:CD2	2:C:92:ALA:HB2	2.51	0.46
1:B:71:ASN:HD22	1:B:71:ASN:C	2.19	0.46
1:B:84:GLU:H	1:B:84:GLU:CD	2.20	0.45
1:B:76:LYS:O	1:B:110:SER:HB3	2.16	0.45
2:C:125:PRO:HB2	2:C:128:GLU:HB3	1.97	0.45
2:C:71:GLU:O	2:C:75:GLU:HG3	2.18	0.44
1:A:15:GLU:HA	1:A:16:PRO:HD2	1.82	0.43
2:D:39:MET:O	2:D:43:LEU:HG	2.18	0.43
2:D:15:ILE:HD11	2:D:54:THR:HB	2.02	0.42
2:D:138:LEU:HD23	2:D:138:LEU:HA	1.85	0.41
2:C:116:GLY:O	2:C:120:HIS:HB2	2.20	0.41
2:C:30:LYS:O	2:C:31:ALA:C	2.57	0.41
1:B:60:PHE:CE2	2:C:92:ALA:HB2	2.56	0.41

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 X-ray entries followed by that with respect to entries of similar resolution.

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	163/165 (99%)	156 (96%)	7 (4%)	0	100	100
1	B	161/165 (98%)	153 (95%)	8 (5%)	0	100	100
2	C	144/146 (99%)	139 (96%)	5 (4%)	0	100	100
2	D	134/146 (92%)	131 (98%)	3 (2%)	0	100	100
All	All	602/622 (97%)	579 (96%)	23 (4%)	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 X-ray entries followed by that with respect to entries of similar resolution.

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	133/133 (100%)	130 (98%)	3 (2%)	58	60
1	B	131/133 (98%)	126 (96%)	5 (4%)	40	36
2	C	123/123 (100%)	118 (96%)	5 (4%)	37	32
2	D	114/123 (93%)	101 (89%)	13 (11%)	7	4
All	All	501/512 (98%)	475 (95%)	26 (5%)	29	23

All (26) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1	MET
1	A	29	VAL
1	A	61	MET
1	B	24	LEU
1	B	29	VAL
1	B	61	MET
1	B	71	ASN
1	B	84	GLU
2	C	95	GLN
2	C	100	ARG
2	C	119	THR
2	C	144	MET
2	C	146	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	D	15	ILE
2	D	18	ARG
2	D	20	LEU
2	D	25	LYS
2	D	50	GLN
2	D	93	PRO
2	D	95	GLN
2	D	111	LEU
2	D	119	THR
2	D	121	ASN
2	D	124	ILE
2	D	128	GLU
2	D	143	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	71	ASN
2	C	50	GLN
2	C	74	ASN
2	D	13	GLN
2	D	121	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.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	165/165 (100%)	-0.30	2 (1%) 81 81	16, 24, 37, 58	0
1	B	163/165 (98%)	-0.26	0 100 100	20, 31, 46, 55	0
2	C	146/146 (100%)	-0.04	4 (2%) 58 58	18, 30, 49, 69	3 (2%)
2	D	135/146 (92%)	0.27	12 (8%) 12 13	23, 38, 65, 81	1 (0%)
All	All	609/622 (97%)	-0.10	18 (2%) 54 55	16, 30, 51, 81	4 (0%)

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	146	SER	5.2
2	D	32	PHE	4.4
2	D	11	VAL	3.5
2	C	121	ASN	3.4
2	D	122	PRO	3.2
2	D	12	HIS	2.7
2	D	119	THR	2.6
2	D	61	GLY	2.6
2	D	120	HIS	2.6
2	C	123	PRO	2.4
2	D	112	GLN	2.3
2	D	30	LYS	2.3
2	D	18	ARG	2.1
2	D	29	GLU	2.1
1	A	165	GLU	2.1
2	D	124	ILE	2.1
1	A	1	MET	2.1
2	C	122	PRO	2.0

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.