wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 08:48 PM GMT

PDB ID	:	1M1A
Title	:	LIGAND BINDING ALTERS THE STRUCTURE AND DYNAMICS OF NUCLEOSOMAL DNA
Authors	:	Suto, R.K.; Edayathumangalam, R.S.; White, C.L.; Melander, C.; Gottesfeld, J.M.; Dervan, P.B.; Luger, K.
Deposited on	:	2002-06-18
Resolution	:	2.65 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 (i)) were used in the production of this report:

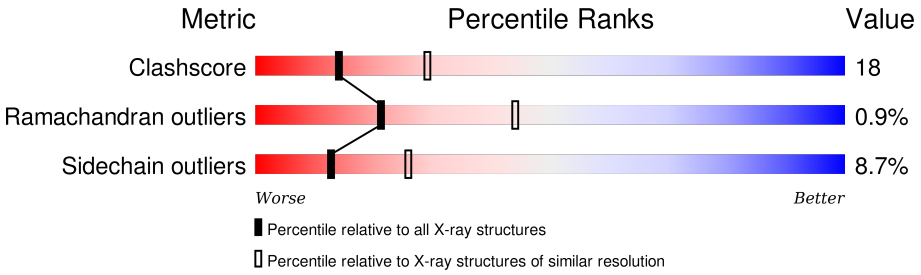
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
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 2.65 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))
Clashscore	102246	3524 (2.70-2.62)
Ramachandran outliers	100387	3469 (2.70-2.62)
Sidechain outliers	100360	3469 (2.70-2.62)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	I	146	38% 59% ·
1	J	146	40% 60%
2	A	135	44% 27% · 27%
2	E	135	55% 17% · 27%
3	B	102	50% 25% · 22%
3	F	102	68% 18% 5% · 9%
4	C	129	60% 16% 5% 19%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	G	129	55% 20% 7% 18%
5	D	125	52% 19% 26%
5	H	125	49% 26% 26%

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
7	PYB	J	1903	-	-	X	-
7	PYB	J	1904	-	-	X	-
7	PYB	J	1908	-	-	X	-
7	DIB	J	1911	-	-	X	-

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 12378 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 DNA chain called Palindromic 146 Base Pair DNA Fragment.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	I	146	Total	C	N	O	P	0	0	0
			2990	1430	541	874	145			
1	J	146	Total	C	N	O	P	0	0	0
			2990	1430	541	874	145			

- Molecule 2 is a protein called Histone H3.2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	99	Total	C	N	O	S	0	0	0
			817	515	158	141	3			
2	E	98	Total	C	N	O	S	0	0	0
			808	509	156	140	3			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	486	SER	ARG	SEE REMARK 999	UNP P02302
E	686	SER	ARG	SEE REMARK 999	UNP P02302

- Molecule 3 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	80	Total	C	N	O	S	0	0	0
			638	401	125	111	1			
3	F	93	Total	C	N	O	S	0	0	0
			737	463	149	124	1			

- Molecule 4 is a protein called Histone H2A.1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	C	105	Total	C	N	O	0	0	0
			809	510	158	141			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	G	106	Total	C	N	O	0	0	0
			818	516	160	142			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	899	ARG	GLY	SEE REMARK 999	UNP P06897
G	1099	ARG	GLY	SEE REMARK 999	UNP P06897

- Molecule 5 is a protein called Histone H2B.1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	D	93	Total	C	N	O	S	0	0	0
			726	457	130	137	2			
5	H	93	Total	C	N	O	S	0	0	0
			726	457	130	137	2			

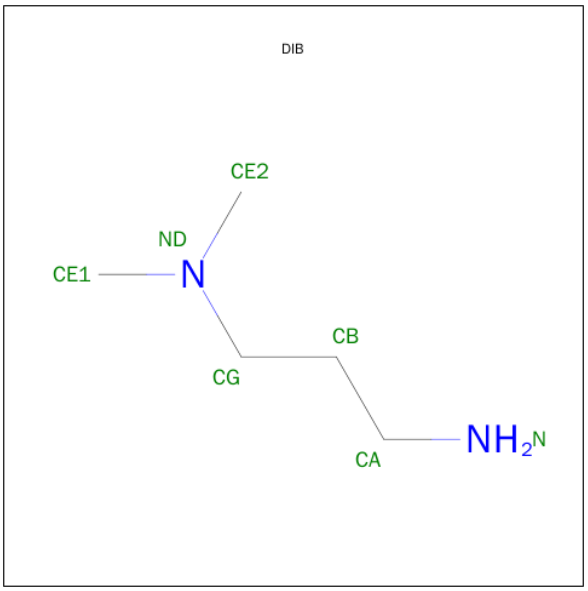
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	1229	THR	SER	SEE REMARK 999	UNP P02281
H	1429	THR	SER	SEE REMARK 999	UNP P02281

- Molecule 6 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	J	4	Total	Mn	0	0
			4	4		
6	I	5	Total	Mn	0	0
			5	5		
6	E	1	Total	Mn	0	0
			1	1		

- Molecule 7 is 4-AMINO-(1-METHYLIMIDAZOLE)-2-CARBOXYLIC ACID (three-letter code: DIB, PYB, IMT, BAL, ABU) (formula: $C_5H_{14}N_2$, $C_6H_8N_2O_2$, $C_5H_7N_3O_2$, $C_3H_7NO_2$, $C_4H_9NO_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	J	11	Total	C	N	O	0	0
			89	58	21	10		

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	24	Total	O	0	0
			24	24		
8	B	17	Total	O	0	0
			17	17		
8	C	30	Total	O	0	0
			30	30		
8	D	8	Total	O	0	0
			8	8		
8	E	35	Total	O	0	0
			35	35		
8	F	40	Total	O	0	0
			40	40		
8	G	16	Total	O	0	0
			16	16		
8	H	11	Total	O	0	0
			11	11		
8	I	11	Total	O	0	0
			11	11		
8	J	28	Total	O	0	0
			28	28		

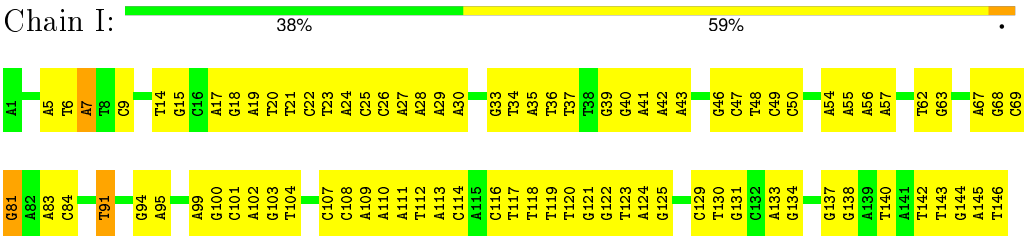
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

Note EDS was not executed.

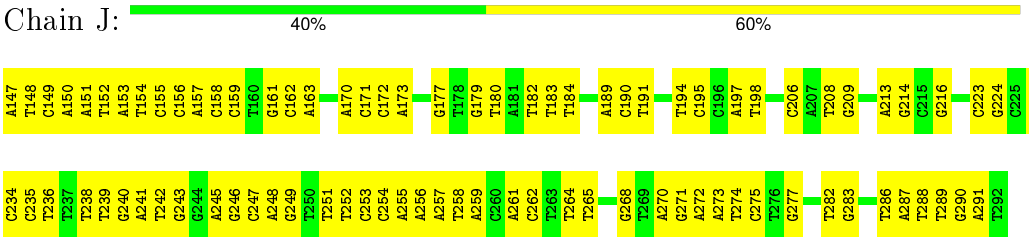
- Molecule 1: Palindromic 146 Base Pair DNA Fragment

Chain I:



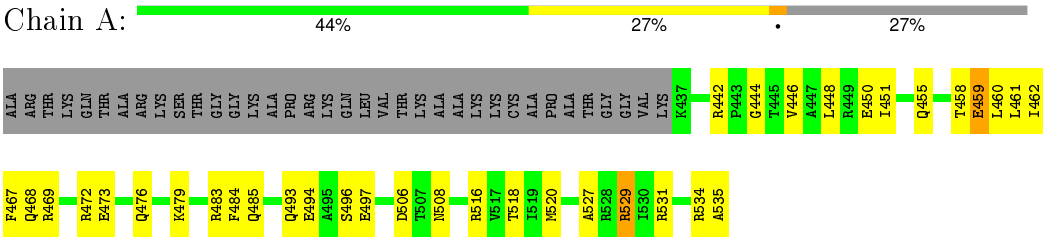
- Molecule 1: Palindromic 146 Base Pair DNA Fragment

Chain J:



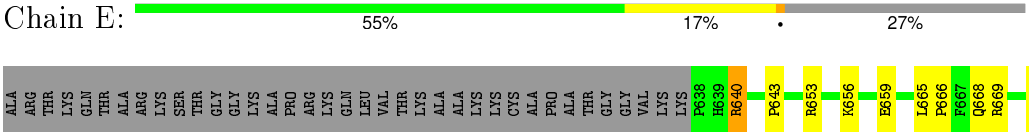
- Molecule 2: Histone H3.2

Chain A:



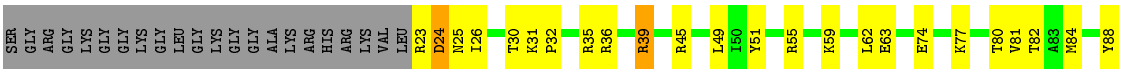
- Molecule 2: Histone H3.2

Chain E:



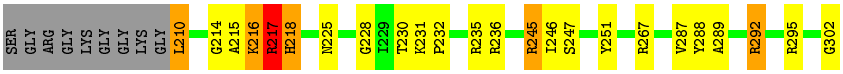
• Molecule 3: Histone H4

Chain B:



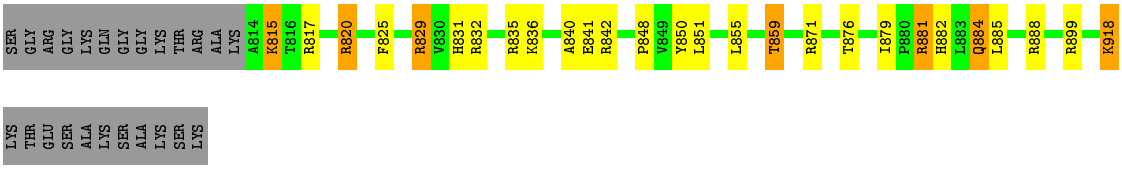
• Molecule 3: Histone H4

Chain F:



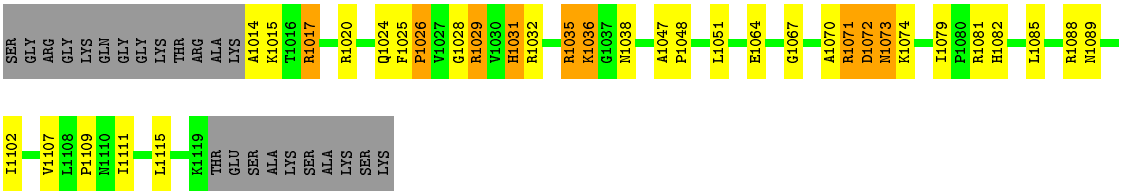
• Molecule 4: Histone H2A.1

Chain C:



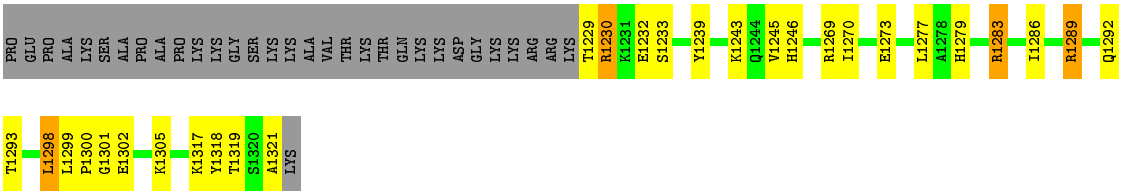
• Molecule 4: Histone H2A.1

Chain G:



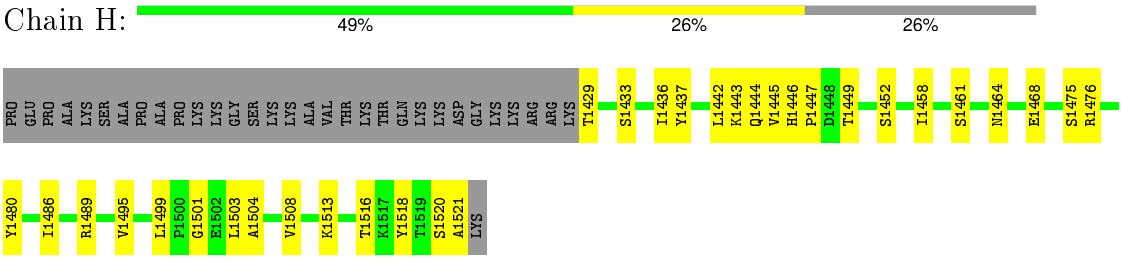
• Molecule 5: Histone H2B.1

Chain D:



• Molecule 5: Histone H2B.1

Chain H:


49%
26%
26%
PRO GLU PRO ALA LYS SER ALA PRO ALA PRO LYS LYS GLY SER LYS LYS ALA VAL THR LYS THR GLN LYS LYS ASP GLY LYS LYS ARG ARG LYS T1429 S1433 I1436 Y1437 L1442 K1443 Q1444 V1445 H1446 P1447 D1448 T1449 S1452 I1458 S1461 N1464 E1468 S1475 R1476
Y1480 I1486 R1489 V1495 L1499 P1500 G1501 E1502 L1503 A1504 V1508 K1513 T1516 K1517 Y1518 T1519 S1520 A1521 LYS

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	106.72Å 109.20Å 177.35Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	60.00 – 2.65	Depositor
% Data completeness (in resolution range)	96.9 (60.00-2.65)	Depositor
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.223 , 0.267	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	12378	wwPDB-VP
Average B, all atoms (Å^2)	88.0	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: MN, DIB, IMT, BAL, PYB, ABU

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	I	0.38	0/3354	0.72	1/5175 (0.0%)
1	J	0.38	0/3354	0.71	0/5175
2	A	0.54	0/829	0.74	0/1111
2	E	0.64	0/820	0.84	0/1099
3	B	0.54	0/645	0.77	0/862
3	F	0.62	0/745	0.93	2/992 (0.2%)
4	C	0.58	0/819	0.81	0/1106
4	G	0.49	0/828	0.73	0/1117
5	D	0.61	0/737	0.76	0/993
5	H	0.56	0/737	0.71	0/993
All	All	0.48	0/12868	0.75	3/18623 (0.0%)

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	I	0	3
1	J	0	1
All	All	0	4

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	F	235	ARG	NE-CZ-NH1	-6.09	117.25	120.30
1	I	81	DG	C5'-C4'-C3'	-5.32	104.53	114.10
3	F	236	ARG	NE-CZ-NH1	-5.12	117.74	120.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	I	7	DA	Sidechain
1	I	78	DG	Sidechain
1	I	91	DT	Sidechain
1	J	213	DA	Sidechain

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	I	2990	0	1651	108	0
1	J	2990	0	1651	114	0
2	A	817	0	858	39	0
2	E	808	0	846	31	0
3	B	638	0	676	27	0
3	F	737	0	793	24	0
4	C	809	0	864	28	0
4	G	818	0	877	38	0
5	D	726	0	747	25	0
5	H	726	0	747	24	0
6	E	1	0	0	0	0
6	I	5	0	0	0	0
6	J	4	0	0	0	0
7	J	89	0	68	26	0
8	A	24	0	0	2	0
8	B	17	0	0	1	0
8	C	30	0	0	1	0
8	D	8	0	0	0	0
8	E	35	0	0	4	0
8	F	40	0	0	5	0
8	G	16	0	0	0	0
8	H	11	0	0	0	0
8	I	11	0	0	2	0
8	J	28	0	0	2	0
All	All	12378	0	9778	398	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 18.

The worst 5 of 398 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:F:295:ARG:HD2	8:F:156:HOH:O	1.36	1.19
4:C:918:LYS:HD2	4:C:918:LYS:H	1.17	1.08
1:J:247:DC:H2"	1:J:248:DA:H5"	1.37	1.03
5:H:1445:VAL:HG12	5:H:1446:HIS:HD2	1.29	0.96
3:F:287:VAL:HG11	3:F:302:GLY:HA3	1.48	0.96

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
2	A	97/135 (72%)	95 (98%)	2 (2%)	0	100 100
2	E	96/135 (71%)	92 (96%)	3 (3%)	1 (1%)	19 41
3	B	78/102 (76%)	73 (94%)	5 (6%)	0	100 100
3	F	91/102 (89%)	86 (94%)	3 (3%)	2 (2%)	8 19
4	C	103/129 (80%)	98 (95%)	5 (5%)	0	100 100
4	G	104/129 (81%)	99 (95%)	3 (3%)	2 (2%)	10 22
5	D	91/125 (73%)	89 (98%)	1 (1%)	1 (1%)	17 38
5	H	91/125 (73%)	85 (93%)	5 (6%)	1 (1%)	17 38
All	All	751/982 (76%)	717 (96%)	27 (4%)	7 (1%)	21 44

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	D	1301	GLY
5	H	1501	GLY

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	G	1072	ASP
2	E	681	ASP
3	F	217	ARG

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
2	A	86/111 (78%)	83 (96%)	3 (4%)	43 71
2	E	85/111 (77%)	80 (94%)	5 (6%)	24 48
3	B	65/78 (83%)	56 (86%)	9 (14%)	4 9
3	F	74/78 (95%)	68 (92%)	6 (8%)	15 30
4	C	83/100 (83%)	73 (88%)	10 (12%)	6 13
4	G	84/100 (84%)	75 (89%)	9 (11%)	8 17
5	D	79/105 (75%)	74 (94%)	5 (6%)	22 44
5	H	79/105 (75%)	71 (90%)	8 (10%)	9 19
All	All	635/788 (81%)	580 (91%)	55 (9%)	13 26

5 of 55 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	D	1283	ARG
2	E	690	MET
5	H	1476	ARG
5	D	1289	ARG
2	E	640	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
5	D	1281	ASN
5	D	1292	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
5	H	1446	HIS
4	C	884	GLN
4	G	1073	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)

Of 21 ligands modelled in this entry, 10 are monoatomic - leaving 11 for Mogul analysis.

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	$\#\|Z\| > 2$
7	IMT	J	1901	7	6,8,10	1.26	1 (16%)	5,10,14	2.47	3 (60%)
7	IMT	J	1902	7	6,9,10	1.80	2 (33%)	5,12,14	1.72	2 (40%)
7	PYB	J	1903	7	8,9,10	1.68	2 (25%)	4,12,14	3.50	2 (50%)
7	PYB	J	1904	7	8,9,10	1.64	2 (25%)	4,12,14	3.55	2 (50%)
7	ABU	J	1905	7	5,5,6	1.57	1 (20%)	4,4,6	0.41	0
7	PYB	J	1906	7	8,9,10	1.27	1 (12%)	4,12,14	3.57	2 (50%)
7	PYB	J	1907	7	8,9,10	1.68	1 (12%)	4,12,14	3.75	2 (50%)
7	PYB	J	1908	7	8,9,10	1.58	2 (25%)	4,12,14	3.76	2 (50%)
7	PYB	J	1909	7	8,9,10	1.56	2 (25%)	4,12,14	3.71	2 (50%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
7	BAL	J	1910	7	3,4,5	0.90	0	0,3,5	0.00	-
7	DIB	J	1911	7	6,6,6	0.80	0	6,6,6	0.49	0

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
7	IMT	J	1901	7	-	0/0/2/4	0/1/1/1
7	IMT	J	1902	7	-	0/0/2/4	0/1/1/1
7	PYB	J	1903	7	-	0/0/2/4	0/1/1/1
7	PYB	J	1904	7	-	0/0/2/4	0/1/1/1
7	ABU	J	1905	7	-	0/3/3/4	0/0/0/0
7	PYB	J	1906	7	-	0/0/2/4	0/1/1/1
7	PYB	J	1907	7	-	0/0/2/4	0/1/1/1
7	PYB	J	1908	7	-	0/0/2/4	0/1/1/1
7	PYB	J	1909	7	-	0/0/2/4	0/1/1/1
7	BAL	J	1910	7	-	0/1/2/3	0/0/0/0
7	DIB	J	1911	7	-	0/4/4/4	0/0/0/0

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	J	1905	ABU	OE2-CD	-3.51	1.23	1.42
7	J	1904	PYB	CB1-CA	2.01	1.41	1.39
7	J	1903	PYB	CB1-CA	2.05	1.41	1.39
7	J	1908	PYB	CB1-CA	2.21	1.41	1.39
7	J	1909	PYB	CB1-CA	2.29	1.41	1.39

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	J	1908	PYB	O-C-CG1	-4.78	116.30	124.44
7	J	1907	PYB	O-C-CG1	-4.38	116.97	124.44
7	J	1903	PYB	O-C-CG1	-4.16	117.35	124.44
7	J	1906	PYB	O-C-CG1	-4.04	117.55	124.44
7	J	1909	PYB	O-C-CG1	-3.93	117.74	124.44

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

10 monomers are involved in 26 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	J	1901	IMT	1	0
7	J	1902	IMT	1	0
7	J	1903	PYB	5	0
7	J	1904	PYB	6	0
7	J	1905	ABU	2	0
7	J	1906	PYB	2	0
7	J	1907	PYB	2	0
7	J	1908	PYB	4	0
7	J	1909	PYB	3	0
7	J	1911	DIB	6	0

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

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates

EDS was not executed - this section will therefore be empty.

6.4 Ligands

EDS was not executed - this section will therefore be empty.

6.5 Other polymers

EDS was not executed - this section will therefore be empty.