wwPDB X-ray Structure Validation Summary Report (i)

Feb 15, 2024 – 03:43 AM EST

PDB ID	:	3O62
Title	:	Nucleosome core particle modified with a cisplatin 1,3-cis-{Pt(NH3)2}2+-d(GpTpG) intrastrand cross-link
Authors	:	Lippard, S.J.; Todd, R.C.
Deposited on	:	2010-07-28
Resolution	:	3.22 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
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:

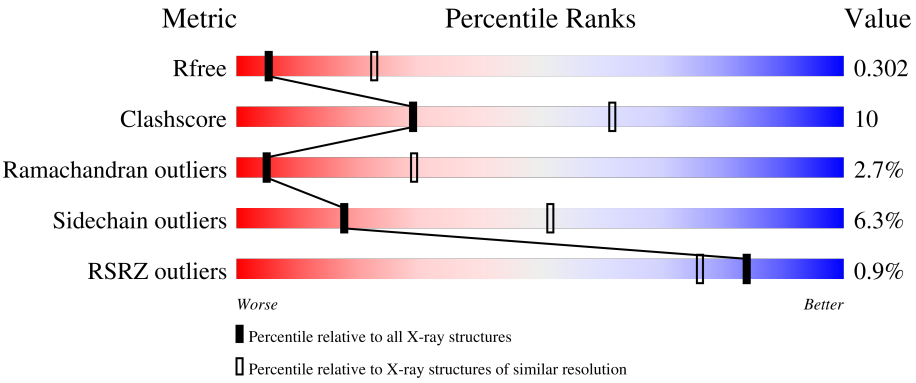
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 3.22 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}	130704	1335 (3.24-3.20)
Clashscore	141614	1460 (3.24-3.20)
Ramachandran outliers	138981	1437 (3.24-3.20)
Sidechain outliers	138945	1436 (3.24-3.20)
RSRZ outliers	127900	1291 (3.24-3.20)

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 of the lower 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% 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	135	55% 13% • 27%
1	E	135	49% 21% • 27%
2	B	102	2% 60% 15% • 22%
2	F	102	67% 13% • 18%
3	C	128	65% 16% • 18%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	128	2% 62% 20% 17%
4	D	122	49% 25% 22%
4	H	122	2% 57% 16% 5% 22%
5	I	146	2% 42% 38% 19%
6	J	146	% 40% 42% 17%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 12032 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 Histone H3.2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	98	Total	C	N	O	S	0	0	0
			808	509	156	140	3			
1	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	102	ALA	GLY	conflict	UNP P84233
E	102	ALA	GLY	conflict	UNP P84233

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	80	Total	C	N	O	S	0	0	0
			638	401	125	111	1			
2	F	84	Total	C	N	O	S	0	0	0
			678	428	135	114	1			

- Molecule 3 is a protein called Histone H2A type 1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	105	Total	C	N	O	0	0	0
			809	510	158	141			
3	G	106	Total	C	N	O	0	0	0
			818	516	160	142			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	99	ARG	GLY	conflict	UNP P06897
C	123	SER	ALA	conflict	UNP P06897

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	?	-	ALA	deletion	UNP P06897
G	99	ARG	GLY	conflict	UNP P06897
G	123	SER	ALA	conflict	UNP P06897
G	?	-	ALA	deletion	UNP P06897

- Molecule 4 is a protein called Histone H2B 1.1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	95	Total	C	N	O	S	0	0	0
			745	469	134	140	2			
4	H	95	Total	C	N	O	S	0	0	0
			745	469	134	140	2			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	29	THR	SER	conflict	UNP P02281
H	29	THR	SER	conflict	UNP P02281

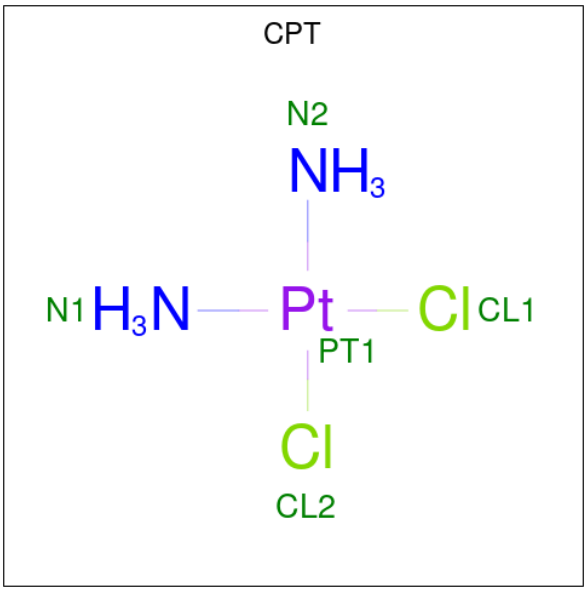
- Molecule 5 is a DNA chain called DNA (146-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	I	146	Total	C	N	O	P	0	0	0
			2968	1421	538	864	145			

- Molecule 6 is a DNA chain called DNA (146-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	J	146	Total	C	N	O	P	0	0	0
			3012	1439	544	884	145			

- Molecule 7 is Cisplatin (three-letter code: CPT) (formula: $Cl_2H_6N_2Pt$).



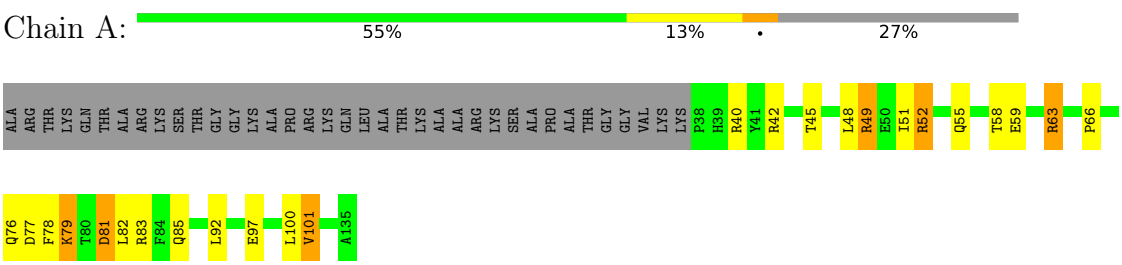
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	N	Pt		
7	I	1	3	2	1	0	0

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 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.

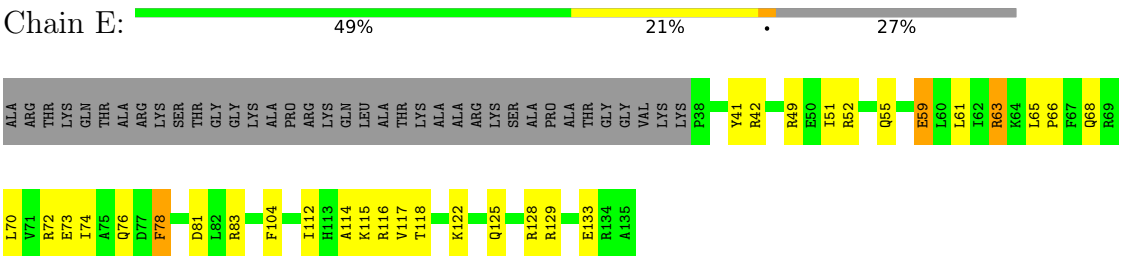
• Molecule 1: Histone H3.2

Chain A:



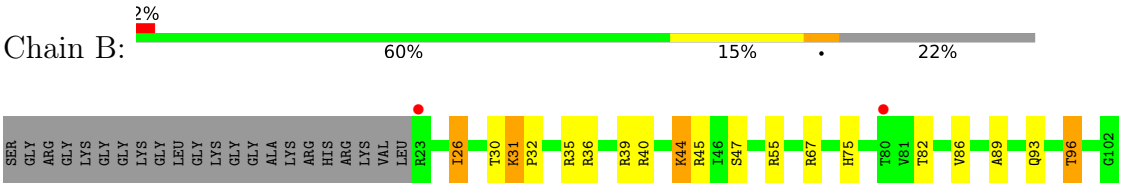
• Molecule 1: Histone H3.2

Chain E:



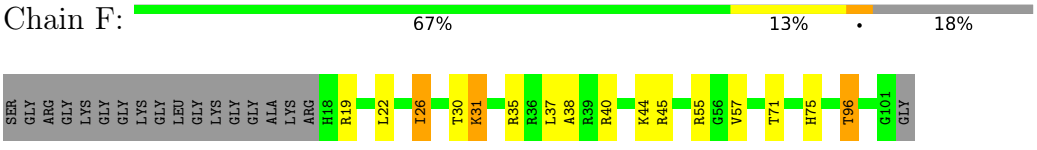
• Molecule 2: Histone H4

Chain B:



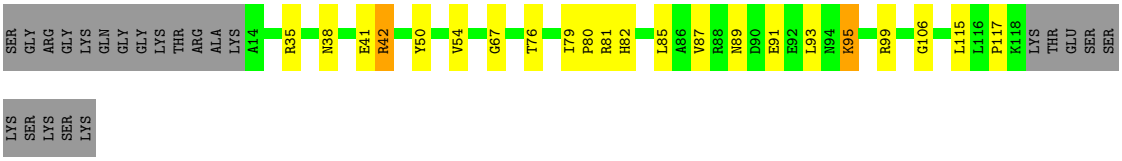
• Molecule 2: Histone H4

Chain F:



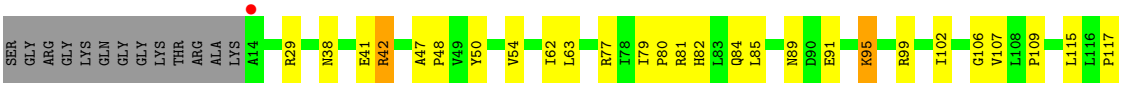
• Molecule 3: Histone H2A type 1

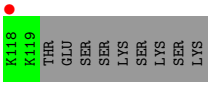
Chain C:



• Molecule 3: Histone H2A type 1

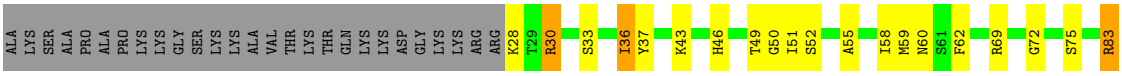
Chain G:

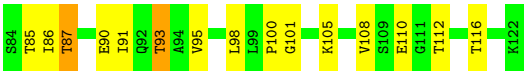




• Molecule 4: Histone H2B 1.1

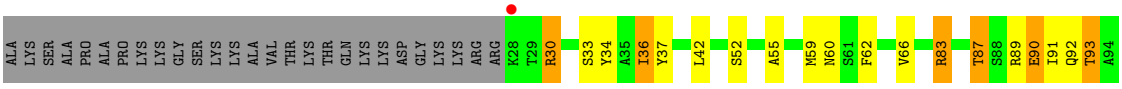
Chain D:

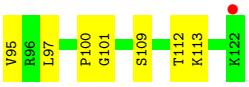




• Molecule 4: Histone H2B 1.1

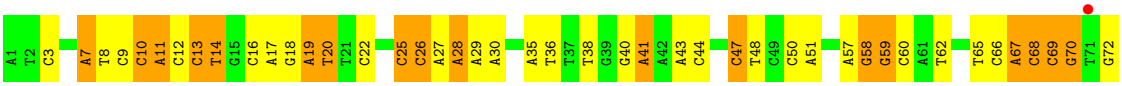
Chain H:

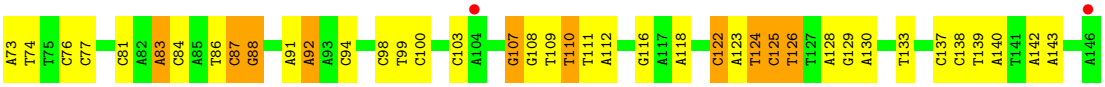




• Molecule 5: DNA (146-MER)

Chain I:

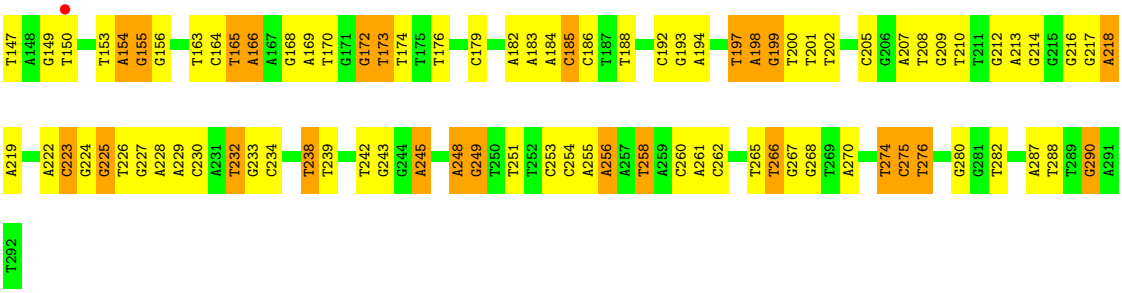




• Molecule 6: DNA (146-MER)

Chain J:


%
40%
42%
17%


T147 A148 G149 T150 T153 A154 G155 G156 T163 C164 T165 A166 A167 G168 A169 T170 G171 G172 T173 T174 T175 T176 C179 A182 A183 A184 C185 C186 T187 T188 C192 G193 A194 T197 A198 G199 T200 T201 T202 C205 G206 A207 T208 G209 T210 T211 G212 A213 G214 G215 G216 G217 A218
A219 A222 C223 G224 G225 T226 G227 A228 A229 C230 A231 T232 G233 C234 T238 T239 T242 G243 G244 A245 A248 G249 T250 T251 T252 C253 C254 A255 A256 A257 T258 A259 C260 A261 C262 T265 T266 G267 G268 T269 A270 T274 C275 T276 G280 G281 T282 A287 T288 T289 G290 A291
T292

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	106.09Å 109.55Å 177.18Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.03 – 3.22 42.03 – 3.22	Depositor EDS
% Data completeness (in resolution range)	68.8 (42.03-3.22) 68.9 (42.03-3.22)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.67 (at 3.19Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.249 , 0.306 0.246 , 0.302	Depositor DCC
R_{free} test set	1209 reflections (5.13%)	wwPDB-VP
Wilson B-factor (Å^2)	71.6	Xtriage
Anisotropy	0.131	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.21 , 79.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.44$, $< L^2 > = 0.27$	Xtriage
Estimated twinning fraction	0.039 for k,h,-l	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	12032	wwPDB-VP
Average B, all atoms (Å^2)	158.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.48% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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

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	0.50	0/820	0.73	0/1099
1	E	0.55	0/820	0.77	0/1099
2	B	0.52	0/645	0.75	0/862
2	F	0.52	0/686	0.75	0/918
3	C	0.49	0/819	0.67	0/1106
3	G	0.45	0/828	0.65	0/1117
4	D	0.56	0/756	0.72	0/1015
4	H	0.51	0/756	0.68	0/1015
5	I	1.04	1/3328 (0.0%)	1.78	80/5127 (1.6%)
6	J	1.03	0/3380	1.78	79/5223 (1.5%)
All	All	0.83	1/12838 (0.0%)	1.41	159/18581 (0.9%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	I	72	DG	N7-C5	6.38	1.43	1.39

The worst 5 of 159 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	I	26	DC	P-O3'-C3'	11.03	132.94	119.70
5	I	138	DC	P-O3'-C3'	8.39	129.77	119.70
5	I	125	DC	O4'-C1'-C2'	-8.38	99.19	105.90
6	J	147	DT	O4'-C1'-N1	8.37	113.86	108.00
6	J	172	DG	O4'-C1'-C2'	-8.27	99.28	105.90

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	808	0	846	28	0
1	E	808	0	846	29	0
2	B	638	0	676	19	0
2	F	678	0	726	16	0
3	C	809	0	864	19	0
3	G	818	0	877	23	0
4	D	745	0	773	24	0
4	H	745	0	773	19	0
5	I	2968	0	1646	49	0
6	J	3012	0	1656	51	0
7	I	3	0	0	2	0
All	All	12032	0	9683	209	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 10.

The worst 5 of 209 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:49:ARG:HG3	1:A:49:ARG:HH11	1.26	0.98
5:I:28:DA:H2"	5:I:29:DA:H5"	1.47	0.93
6:J:255:DA:H2"	6:J:256:DA:C8	2.11	0.85
4:H:91:ILE:O	4:H:95:VAL:HG23	1.80	0.81
4:D:33:SER:HB2	4:D:60:ASN:HD21	1.51	0.76

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	96/135 (71%)	86 (90%)	7 (7%)	3 (3%)	4 26
1	E	96/135 (71%)	81 (84%)	11 (12%)	4 (4%)	3 19
2	B	78/102 (76%)	71 (91%)	6 (8%)	1 (1%)	12 46
2	F	82/102 (80%)	74 (90%)	6 (7%)	2 (2%)	6 32
3	C	103/128 (80%)	91 (88%)	11 (11%)	1 (1%)	15 52
3	G	104/128 (81%)	92 (88%)	11 (11%)	1 (1%)	15 52
4	D	93/122 (76%)	78 (84%)	12 (13%)	3 (3%)	4 25
4	H	93/122 (76%)	70 (75%)	18 (19%)	5 (5%)	2 13
All	All	745/974 (76%)	643 (86%)	82 (11%)	20 (3%)	5 29

5 of 20 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	117	PRO
3	G	117	PRO
1	A	81	ASP
1	A	101	VAL
4	D	51	ILE

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	85/110 (77%)	78 (92%)	7 (8%)	11 40
1	E	85/110 (77%)	82 (96%)	3 (4%)	36 68
2	B	65/78 (83%)	61 (94%)	4 (6%)	18 52
2	F	70/78 (90%)	66 (94%)	4 (6%)	20 55
3	C	83/101 (82%)	79 (95%)	4 (5%)	25 60
3	G	84/101 (83%)	80 (95%)	4 (5%)	25 60

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
4	D	81/102 (79%)	73 (90%)	8 (10%)	8 30
4	H	81/102 (79%)	75 (93%)	6 (7%)	13 45
All	All	634/782 (81%)	594 (94%)	40 (6%)	18 52

5 of 40 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	F	44	LYS
4	H	83	ARG
2	F	96	THR
3	G	95	LYS
4	H	90	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
3	G	38	ASN
4	H	60	ASN
4	H	64	ASN
4	D	60	ASN
4	D	64	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

1 ligand is modelled in this entry.

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	CPT	I	147	5	0,2,4	-	-	-		

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.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	I	147	CPT	2	0

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.

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	98/135 (72%)	-0.30	0 100 100	48, 120, 210, 276	0
1	E	98/135 (72%)	-0.49	0 100 100	31, 67, 151, 197	0
2	B	80/102 (78%)	-0.09	2 (2%) 57 44	52, 109, 167, 199	0
2	F	84/102 (82%)	-0.50	0 100 100	28, 59, 118, 199	0
3	C	105/128 (82%)	-0.57	0 100 100	27, 66, 143, 210	0
3	G	106/128 (82%)	-0.22	2 (1%) 66 54	56, 112, 192, 246	0
4	D	95/122 (77%)	-0.54	0 100 100	31, 75, 156, 181	0
4	H	95/122 (77%)	-0.46	2 (2%) 63 50	39, 103, 195, 212	0
5	I	146/146 (100%)	-0.21	3 (2%) 63 50	99, 219, 280, 358	0
6	J	146/146 (100%)	-0.18	1 (0%) 87 82	100, 221, 297, 341	0
All	All	1053/1266 (83%)	-0.34	10 (0%) 84 76	27, 113, 260, 358	0

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	I	146	DA	6.8
4	H	28	LYS	3.2
2	B	23	ARG	3.0
6	J	150	DT	3.0
5	I	71	DT	2.8

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 (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
7	CPT	I	147	3/5	0.91	0.11	247,247,252,311	0

6.5 Other polymers (i)

There are no such residues in this entry.