wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 08:31 PM GMT

PDB ID	:	1KPG
Title	:	Crystal Structure of mycolic acid cyclopropane synthase CmaA1 complexed with SAH and CTAB
Authors	:	Huang, C.-C.; Smith, C.V.; Jacobs Jr., W.R.; Glickman, M.S.; Sacchettini, J.C.; TB Structural Genomics Consortium (TBSGC)
Deposited on	:	2001-12-30
Resolution	:	2.00 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:

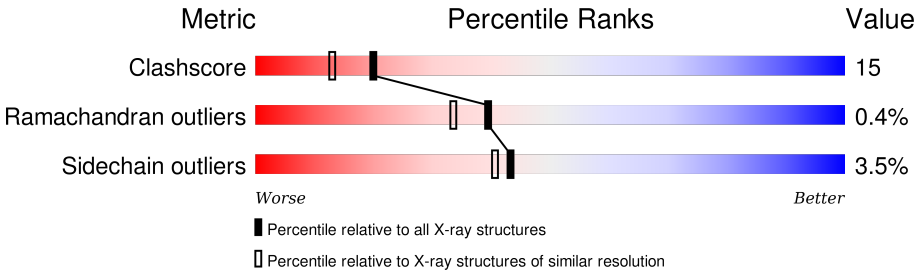
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
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 2.00 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))
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	287	75% 23%
1	B	287	72% 25%
1	C	287	69% 28%
1	D	287	70% 27%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 9906 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 CYCLOPROPANE-FATTY-ACYL-PHOSPHOLIPID SYNTHASE 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	285	Total	C	N	O	S	Se	0	0	0
			2269	1452	388	415	5	9			
1	B	284	Total	C	N	O	S	Se	0	0	0
			2261	1448	387	412	5	9			
1	C	285	Total	C	N	O	S	Se	0	0	0
			2269	1452	388	415	5	9			
1	D	283	Total	C	N	O	S	Se	0	0	0
			2252	1443	386	409	5	9			

There are 40 discrepancies between the modelled and reference sequences:

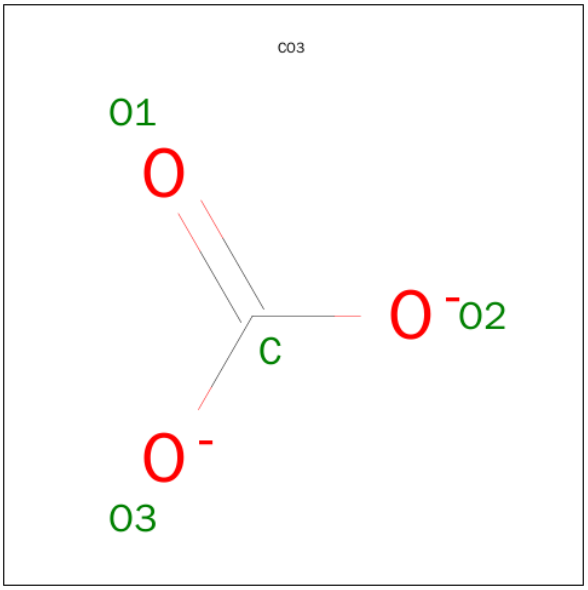
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	CLONING ARTIFACT	UNP Q11195
A	43	MSE	MET	CLONING ARTIFACT	UNP Q11195
A	66	MSE	MET	CLONING ARTIFACT	UNP Q11195
A	79	MSE	MET	CLONING ARTIFACT	UNP Q11195
A	80	MSE	MET	CLONING ARTIFACT	UNP Q11195
A	164	MSE	MET	CLONING ARTIFACT	UNP Q11195
A	184	MSE	MET	CLONING ARTIFACT	UNP Q11195
A	210	MSE	MET	CLONING ARTIFACT	UNP Q11195
A	263	MSE	MET	CLONING ARTIFACT	UNP Q11195
A	272	MSE	MET	CLONING ARTIFACT	UNP Q11195
B	1	MSE	MET	CLONING ARTIFACT	UNP Q11195
B	43	MSE	MET	CLONING ARTIFACT	UNP Q11195
B	66	MSE	MET	CLONING ARTIFACT	UNP Q11195
B	79	MSE	MET	CLONING ARTIFACT	UNP Q11195
B	80	MSE	MET	CLONING ARTIFACT	UNP Q11195
B	164	MSE	MET	CLONING ARTIFACT	UNP Q11195
B	184	MSE	MET	CLONING ARTIFACT	UNP Q11195
B	210	MSE	MET	CLONING ARTIFACT	UNP Q11195
B	263	MSE	MET	CLONING ARTIFACT	UNP Q11195
B	272	MSE	MET	CLONING ARTIFACT	UNP Q11195

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	1	MSE	MET	CLONING ARTIFACT	UNP Q11195
C	43	MSE	MET	CLONING ARTIFACT	UNP Q11195
C	66	MSE	MET	CLONING ARTIFACT	UNP Q11195
C	79	MSE	MET	CLONING ARTIFACT	UNP Q11195
C	80	MSE	MET	CLONING ARTIFACT	UNP Q11195
C	164	MSE	MET	CLONING ARTIFACT	UNP Q11195
C	184	MSE	MET	CLONING ARTIFACT	UNP Q11195
C	210	MSE	MET	CLONING ARTIFACT	UNP Q11195
C	263	MSE	MET	CLONING ARTIFACT	UNP Q11195
C	272	MSE	MET	CLONING ARTIFACT	UNP Q11195
D	1	MSE	MET	CLONING ARTIFACT	UNP Q11195
D	43	MSE	MET	CLONING ARTIFACT	UNP Q11195
D	66	MSE	MET	CLONING ARTIFACT	UNP Q11195
D	79	MSE	MET	CLONING ARTIFACT	UNP Q11195
D	80	MSE	MET	CLONING ARTIFACT	UNP Q11195
D	164	MSE	MET	CLONING ARTIFACT	UNP Q11195
D	184	MSE	MET	CLONING ARTIFACT	UNP Q11195
D	210	MSE	MET	CLONING ARTIFACT	UNP Q11195
D	263	MSE	MET	CLONING ARTIFACT	UNP Q11195
D	272	MSE	MET	CLONING ARTIFACT	UNP Q11195

- Molecule 2 is CARBONATE ION (three-letter code: CO3) (formula: CO_3).



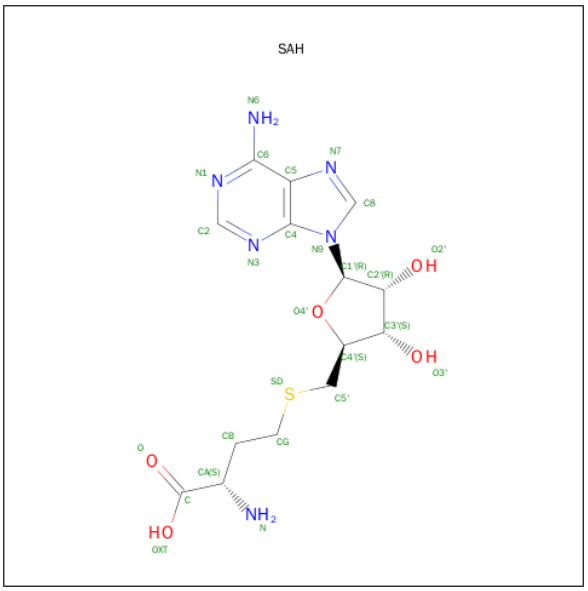
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 4 1 3	0	0

Continued on next page...

Continued from previous page...

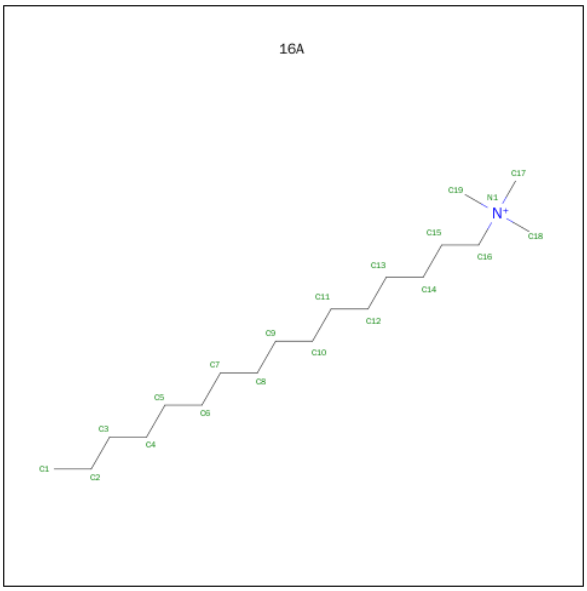
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total C O 4 1 3	0	0
2	C	1	Total C O 4 1 3	0	0
2	D	1	Total C O 4 1 3	0	0

- Molecule 3 is S-ADENOSYL-L-HOMOCYSTEINE (three-letter code: SAH) (formula: $C_{14}H_{20}N_6O_5S$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C N O S 26 14 6 5 1	0	0
3	B	1	Total C N O S 26 14 6 5 1	0	0
3	C	1	Total C N O S 26 14 6 5 1	0	0
3	D	1	Total C N O S 26 14 6 5 1	0	0

- Molecule 4 is CETYL-TRIMETHYL-AMMONIUM (three-letter code: 16A) (formula: $C_{19}H_{42}N$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N 20 19 1	0	0
4	B	1	Total C N 20 19 1	0	0
4	C	1	Total C N 20 19 1	0	0
4	D	1	Total C N 20 19 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	185	Total O 185 185	0	0
5	B	140	Total O 140 140	0	0
5	C	180	Total O 180 180	0	0
5	D	150	Total O 150 150	0	0

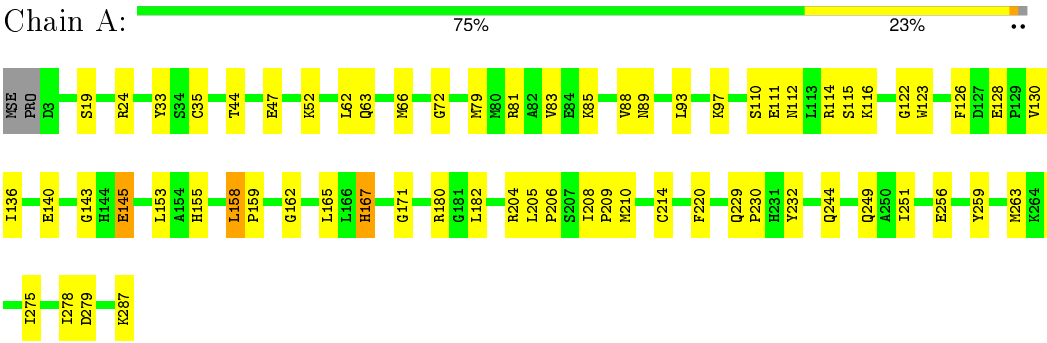
3 Residue-property plots (i)

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.

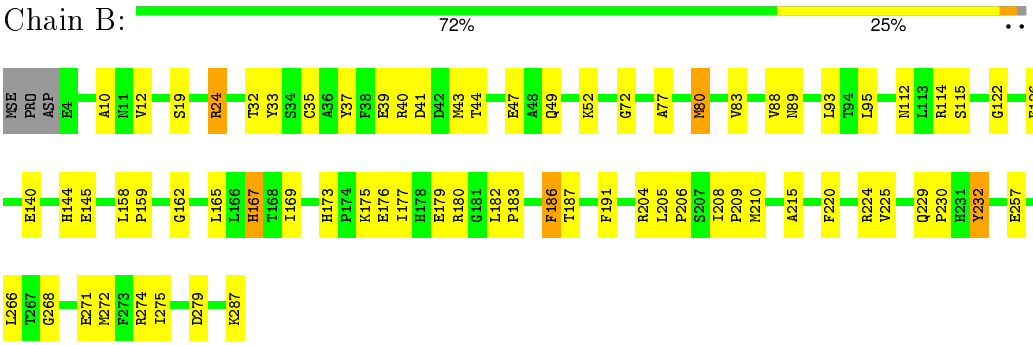
• Molecule 1: CYCLOPROPANE-FATTY-ACYL-PHOSPHOLIPID SYNTHASE 1

Chain A:



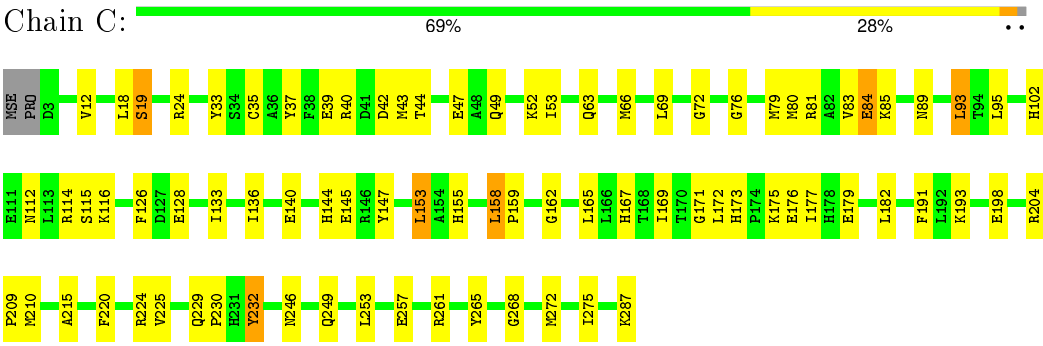
• Molecule 1: CYCLOPROPANE-FATTY-ACYL-PHOSPHOLIPID SYNTHASE 1

Chain B:



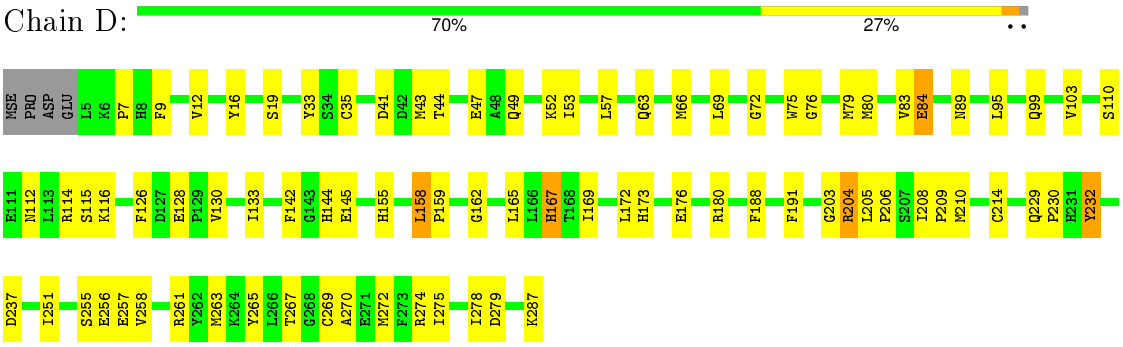
• Molecule 1: CYCLOPROPANE-FATTY-ACYL-PHOSPHOLIPID SYNTHASE 1

Chain C:



• Molecule 1: CYCLOPROPANE-FATTY-ACYL-PHOSPHOLIPID SYNTHASE 1

Chain D:



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 43	Depositor
Cell constants a, b, c, α, β, γ	77.32Å 77.32Å 173.49Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.68 – 2.00	Depositor
% Data completeness (in resolution range)	95.9 (29.68-2.00)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.191 , 0.233	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	9906	wwPDB-VP
Average B, all atoms (Å^2)	19.0	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: CO3, SAH, 16A

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	0.34	0/2314	0.59	0/3118
1	B	0.32	0/2306	0.56	0/3107
1	C	0.33	0/2314	0.59	1/3118 (0.0%)
1	D	0.32	0/2297	0.58	1/3095 (0.0%)
All	All	0.33	0/9231	0.58	2/12438 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	76	GLY	N-CA-C	5.07	125.77	113.10
1	C	76	GLY	N-CA-C	5.04	125.70	113.10

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	2269	0	2226	66	0
1	B	2261	0	2222	71	0
1	C	2269	0	2226	74	0
1	D	2252	0	2216	66	0
2	A	4	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	4	0	0	0	0
2	C	4	0	0	0	0
2	D	4	0	0	0	0
3	A	26	0	19	2	0
3	B	26	0	19	3	0
3	C	26	0	19	1	0
3	D	26	0	19	1	0
4	A	20	0	42	2	0
4	B	20	0	42	5	0
4	C	20	0	42	4	0
4	D	20	0	42	5	0
5	A	185	0	0	5	0
5	B	140	0	0	3	0
5	C	180	0	0	4	0
5	D	150	0	0	0	0
All	All	9906	0	9134	272	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 15.

The worst 5 of 272 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:24:ARG:HB2	1:B:24:ARG:HH11	1.22	1.04
1:C:69:LEU:HD11	1:C:93:LEU:HD11	1.39	1.00
1:B:182:LEU:HD22	1:B:275:ILE:HG23	1.54	0.87
1:D:205:LEU:HB2	4:D:4901:16A:H91	1.60	0.83
1:D:63:GLN:HB2	1:D:66:MSE:HE3	1.63	0.81

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	283/287 (99%)	272 (96%)	10 (4%)	1 (0%)	39	33
1	B	282/287 (98%)	269 (95%)	12 (4%)	1 (0%)	39	33
1	C	283/287 (99%)	273 (96%)	9 (3%)	1 (0%)	39	33
1	D	281/287 (98%)	276 (98%)	4 (1%)	1 (0%)	39	33
All	All	1129/1148 (98%)	1090 (96%)	35 (3%)	4 (0%)	39	33

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	19	SER
1	A	19	SER
1	B	19	SER
1	D	19	SER

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	241/233 (103%)	233 (97%)	8 (3%)	45	43
1	B	240/233 (103%)	231 (96%)	9 (4%)	40	36
1	C	241/233 (103%)	232 (96%)	9 (4%)	41	38
1	D	239/233 (103%)	231 (97%)	8 (3%)	45	43
All	All	961/932 (103%)	927 (96%)	34 (4%)	43	40

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	232	TYR
1	C	93	LEU
1	D	204	ARG
1	C	33	TYR
1	A	232	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such

sidechains are listed below:

Mol	Chain	Res	Type
1	B	144	HIS
1	C	13	GLN
1	D	173	HIS
1	B	246	ASN
1	B	249	GLN

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)

12 ligands are 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	$\#\|Z\| > 2$
3	SAH	A	1900	-	20,28,28	1.05	2 (10%)	19,40,40	1.19	2 (10%)
4	16A	A	1901	-	19,19,19	0.65	0	21,21,21	0.99	0
2	CO3	A	1902	-	0,3,3	0.00	-	0,3,3	0.00	-
3	SAH	B	2900	-	20,28,28	1.17	2 (10%)	19,40,40	1.10	1 (5%)
4	16A	B	2901	-	19,19,19	0.63	0	21,21,21	0.92	0
2	CO3	B	2902	-	0,3,3	0.00	-	0,3,3	0.00	-
3	SAH	C	3900	-	20,28,28	1.22	2 (10%)	19,40,40	1.08	1 (5%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
4	16A	C	3901	-	19,19,19	0.63	0	21,21,21	0.91	0
2	CO3	C	3902	-	0,3,3	0.00	-	0,3,3	0.00	-
3	SAH	D	4900	-	20,28,28	1.20	2 (10%)	19,40,40	1.09	1 (5%)
4	16A	D	4901	-	19,19,19	0.62	0	21,21,21	1.21	2 (9%)
2	CO3	D	4902	-	0,3,3	0.00	-	0,3,3	0.00	-

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
3	SAH	A	1900	-	-	0/7/31/31	0/3/3/3
4	16A	A	1901	-	-	0/17/17/17	0/0/0/0
2	CO3	A	1902	-	-	0/0/0/0	0/0/0/0
3	SAH	B	2900	-	-	0/7/31/31	0/3/3/3
4	16A	B	2901	-	-	0/17/17/17	0/0/0/0
2	CO3	B	2902	-	-	0/0/0/0	0/0/0/0
3	SAH	C	3900	-	-	0/7/31/31	0/3/3/3
4	16A	C	3901	-	-	0/17/17/17	0/0/0/0
2	CO3	C	3902	-	-	0/0/0/0	0/0/0/0
3	SAH	D	4900	-	-	0/7/31/31	0/3/3/3
4	16A	D	4901	-	-	0/17/17/17	0/0/0/0
2	CO3	D	4902	-	-	0/0/0/0	0/0/0/0

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	3900	SAH	C8-N7	-3.33	1.28	1.34
3	D	4900	SAH	C8-N7	-2.85	1.29	1.34
3	B	2900	SAH	C8-N7	-2.53	1.29	1.34
3	A	1900	SAH	C8-N7	-2.48	1.29	1.34
3	A	1900	SAH	O4'-C1'	2.56	1.44	1.41

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	D	4901	16A	C14-C13-C12	-2.03	104.04	114.53
3	A	1900	SAH	O4'-C4'-C5'	2.19	114.83	108.85
4	D	4901	16A	C15-C16-N1	2.82	129.15	116.08
3	C	3900	SAH	C4-C5-N7	3.01	112.25	109.48
3	B	2900	SAH	C4-C5-N7	3.05	112.29	109.48

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 22 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1900	SAH	2	0
4	A	1901	16A	2	0
3	B	2900	SAH	3	0
4	B	2901	16A	5	0
3	C	3900	SAH	1	0
4	C	3901	16A	4	0
3	D	4900	SAH	1	0
4	D	4901	16A	5	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

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.