wwPDB X-ray Structure Validation Summary Report (i)

Oct 30, 2023 – 06:11 PM JST

PDB ID	:	4XGP
Title	:	Crystal Structure of E112A/H234A Mutant of Stationary Phase Survival Protein (SurE) from Salmonella typhimurium co-crystallized and soaked with AMP.
Authors	:	Mathiharan, Y.K.; Murthy, M.R.N.
Deposited on	:	2015-01-01
Resolution	:	1.90 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:

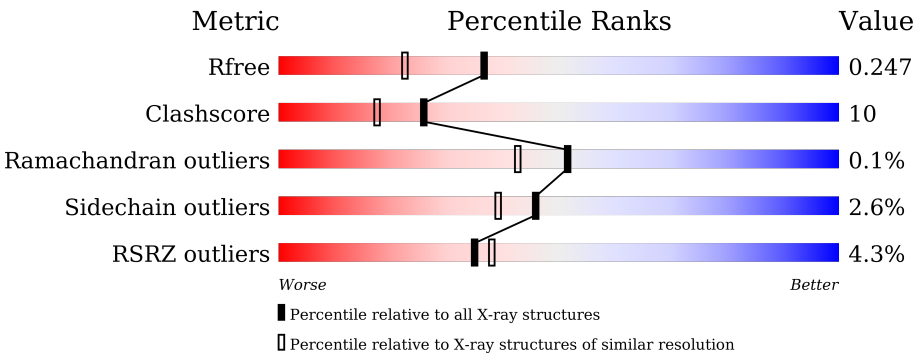
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	267	3% 79% 18% ••
1	B	267	2% 81% 14% • •
1	C	267	3% 80% 15% • •
1	D	267	9% 78% 15% • 5%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 8770 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 5'/3'-nucleotidase SurE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	264	2022	1263	360	389	10	0	10	0
1	B	257	1915	1200	329	377	9	1	6	0
1	C	255	1903	1186	339	371	7	0	2	0
1	D	254	1900	1191	331	371	7	0	7	0

There are 64 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-13	MET	-	expression tag	UNP P66881
A	-12	ARG	-	expression tag	UNP P66881
A	-11	GLY	-	expression tag	UNP P66881
A	-10	SER	-	expression tag	UNP P66881
A	-9	HIS	-	expression tag	UNP P66881
A	-8	HIS	-	expression tag	UNP P66881
A	-7	HIS	-	expression tag	UNP P66881
A	-6	HIS	-	expression tag	UNP P66881
A	-5	HIS	-	expression tag	UNP P66881
A	-4	HIS	-	expression tag	UNP P66881
A	-3	GLY	-	expression tag	UNP P66881
A	-2	MET	-	expression tag	UNP P66881
A	-1	ALA	-	expression tag	UNP P66881
A	0	SER	-	expression tag	UNP P66881
A	112	ALA	GLU	engineered mutation	UNP P66881
A	234	ALA	HIS	engineered mutation	UNP P66881
B	-13	MET	-	expression tag	UNP P66881
B	-12	ARG	-	expression tag	UNP P66881
B	-11	GLY	-	expression tag	UNP P66881
B	-10	SER	-	expression tag	UNP P66881
B	-9	HIS	-	expression tag	UNP P66881

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-8	HIS	-	expression tag	UNP P66881
B	-7	HIS	-	expression tag	UNP P66881
B	-6	HIS	-	expression tag	UNP P66881
B	-5	HIS	-	expression tag	UNP P66881
B	-4	HIS	-	expression tag	UNP P66881
B	-3	GLY	-	expression tag	UNP P66881
B	-2	MET	-	expression tag	UNP P66881
B	-1	ALA	-	expression tag	UNP P66881
B	0	SER	-	expression tag	UNP P66881
B	112	ALA	GLU	engineered mutation	UNP P66881
B	234	ALA	HIS	engineered mutation	UNP P66881
C	-13	MET	-	expression tag	UNP P66881
C	-12	ARG	-	expression tag	UNP P66881
C	-11	GLY	-	expression tag	UNP P66881
C	-10	SER	-	expression tag	UNP P66881
C	-9	HIS	-	expression tag	UNP P66881
C	-8	HIS	-	expression tag	UNP P66881
C	-7	HIS	-	expression tag	UNP P66881
C	-6	HIS	-	expression tag	UNP P66881
C	-5	HIS	-	expression tag	UNP P66881
C	-4	HIS	-	expression tag	UNP P66881
C	-3	GLY	-	expression tag	UNP P66881
C	-2	MET	-	expression tag	UNP P66881
C	-1	ALA	-	expression tag	UNP P66881
C	0	SER	-	expression tag	UNP P66881
C	112	ALA	GLU	engineered mutation	UNP P66881
C	234	ALA	HIS	engineered mutation	UNP P66881
D	-13	MET	-	expression tag	UNP P66881
D	-12	ARG	-	expression tag	UNP P66881
D	-11	GLY	-	expression tag	UNP P66881
D	-10	SER	-	expression tag	UNP P66881
D	-9	HIS	-	expression tag	UNP P66881
D	-8	HIS	-	expression tag	UNP P66881
D	-7	HIS	-	expression tag	UNP P66881
D	-6	HIS	-	expression tag	UNP P66881
D	-5	HIS	-	expression tag	UNP P66881
D	-4	HIS	-	expression tag	UNP P66881
D	-3	GLY	-	expression tag	UNP P66881
D	-2	MET	-	expression tag	UNP P66881
D	-1	ALA	-	expression tag	UNP P66881
D	0	SER	-	expression tag	UNP P66881
D	112	ALA	GLU	engineered mutation	UNP P66881

Continued on next page...

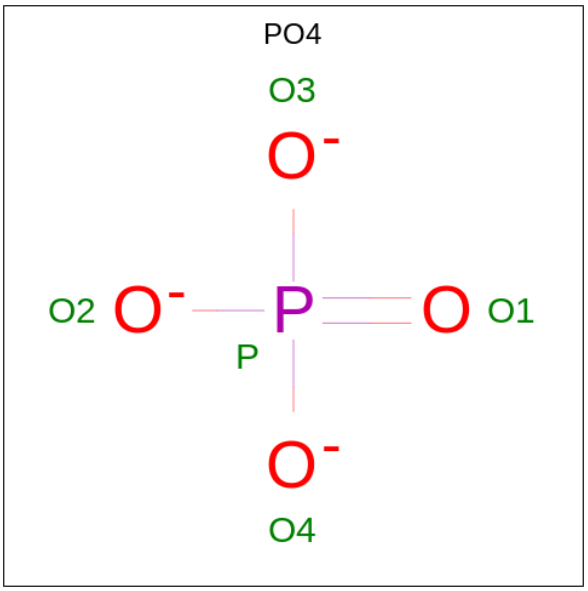
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	234	ALA	HIS	engineered mutation	UNP P66881

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

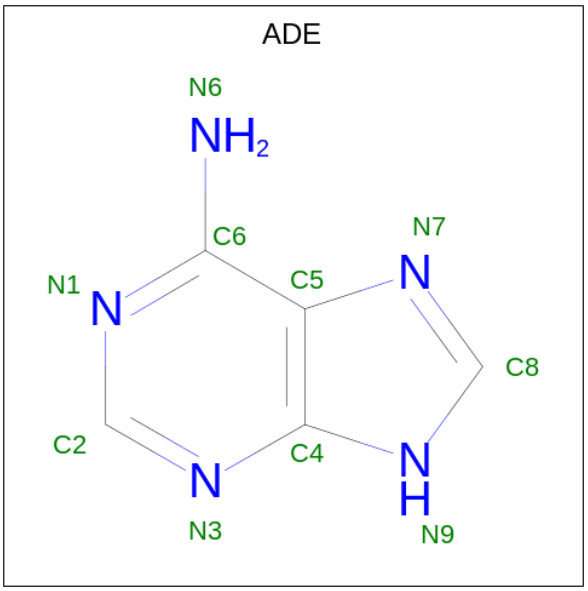
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Mg 1 1	0	0
2	B	1	Total Mg 1 1	0	0
2	C	1	Total Mg 1 1	0	0
2	D	1	Total Mg 1 1	0	0

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



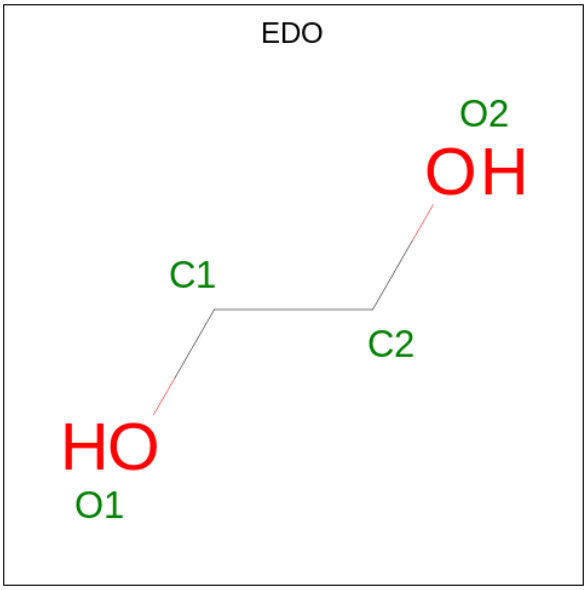
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O P 5 4 1	0	0

- Molecule 4 is ADENINE (three-letter code: ADE) (formula: $C_5H_5N_5$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	N		
4	A	1	10	5	5	0	0

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



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

Continued on next page...

Continued from previous page...

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

- Molecule 6 is water.

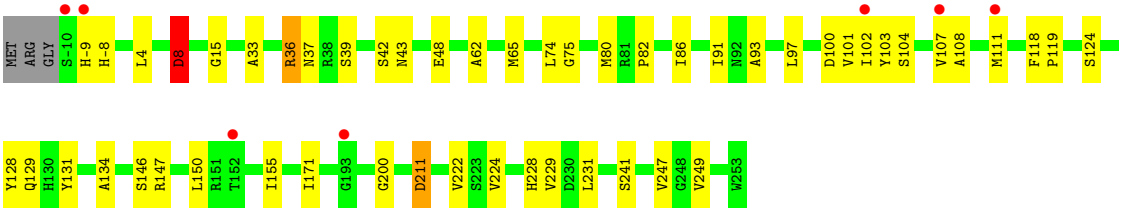
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	A	305	305	305	0	0
			Total	O		
6	B	280	280	280	0	0
			Total	O		
6	C	219	219	219	0	0
			Total	O		
6	D	187	187	187	0	0

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.

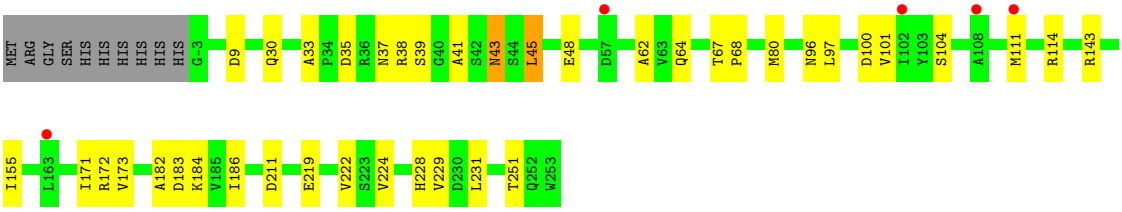
- Molecule 1: 5'/3'-nucleotidase SurE

Chain A:



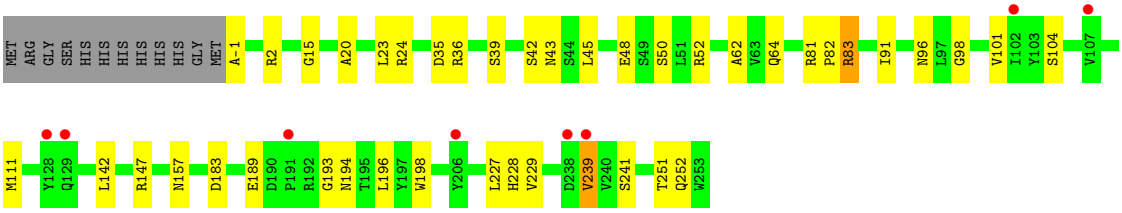
- Molecule 1: 5'/3'-nucleotidase SurE

Chain B:



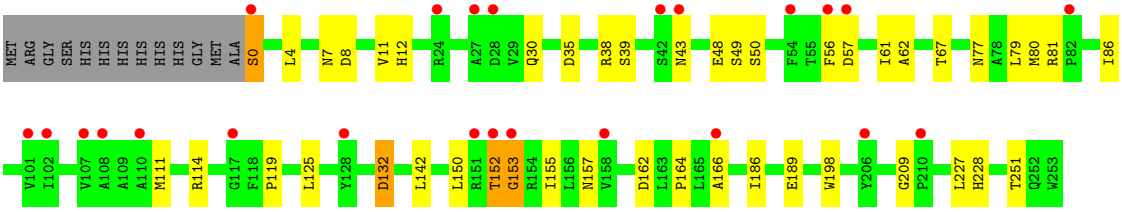
- Molecule 1: 5'/3'-nucleotidase SurE

Chain C:



- Molecule 1: 5'/3'-nucleotidase SurE

Chain D:


MET ARG GLY SER HIS HIS HIS HIS HIS GLY MET ALA S0 L4 N7 D8 V11 H12 R24 A27 D28 V29 Q30 D35 R38 S39 S42 N43 E48 S49 S50 F54 T55 F56 D57 I61 A62 T67 N77 A78 L79 M80 R81 P82 I86
V101 I102 V107 A108 A109 A110 M111 R114 G117 F118 P119 L125 Y128 D132 L142 L150 R151 T152 G153 R154 I155 L156 N157 V158 D162 L163 P164 L165 A166 I186 E189 W198 Y206 G209 P210 L227 H228 T251 Q252 W253

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	162.52Å 96.23Å 93.94Å 90.00° 98.41° 90.00°	Depositor
Resolution (Å)	32.87 – 1.90 31.06 – 1.90	Depositor EDS
% Data completeness (in resolution range)	98.8 (32.87-1.90) 98.9 (31.06-1.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	3.19 (at 1.91Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.189 , 0.240 0.197 , 0.247	Depositor DCC
R_{free} test set	5587 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	31.0	Xtriage
Anisotropy	0.033	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 53.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	8770	wwPDB-VP
Average B, all atoms (Å^2)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.99% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: ADE, EDO, PO4, MG

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	0.95	0/2089	0.97	2/2855 (0.1%)
1	B	0.85	0/1973	0.89	2/2700 (0.1%)
1	C	0.70	0/1949	0.80	0/2666
1	D	0.67	0/1960	0.79	0/2684
All	All	0.80	0/7971	0.87	4/10905 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	8	ASP	CB-CG-OD2	-5.77	113.11	118.30
1	B	45	LEU	CA-CB-CG	5.30	127.49	115.30
1	B	183	ASP	CB-CG-OD1	5.28	123.05	118.30
1	A	8	ASP	N-CA-C	5.18	125.00	111.00

There are no chirality outliers.

There are no planarity outliers.

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	2022	0	1979	58	0
1	B	1915	0	1871	35	0
1	C	1903	0	1867	36	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	1900	0	1848	49	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	5	0	0	0	0
4	A	10	0	4	3	0
5	B	4	0	6	1	0
5	C	12	0	18	0	0
5	D	4	0	6	0	0
6	A	305	0	0	9	0
6	B	280	0	0	13	0
6	C	219	0	0	5	0
6	D	187	0	0	15	0
All	All	8770	0	7599	161	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 161 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:C:252:GLN:HA	6:C:401:HOH:O	1.46	1.16
1:B:143:ARG:HD3	6:B:609:HOH:O	1.60	1.01
1:A:111[B]:MET:HE2	1:A:229:VAL:HG12	1.46	0.96
1:A:111[B]:MET:HE2	1:A:229:VAL:CG1	2.02	0.88
1:A:36:ARG:HG3	6:D:530:HOH:O	1.74	0.86

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	272/267 (102%)	263 (97%)	9 (3%)	0	100	100
1	B	261/267 (98%)	250 (96%)	11 (4%)	0	100	100
1	C	255/267 (96%)	249 (98%)	6 (2%)	0	100	100
1	D	259/267 (97%)	244 (94%)	14 (5%)	1 (0%)	34	24
All	All	1047/1068 (98%)	1006 (96%)	40 (4%)	1 (0%)	51	42

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	153	GLY

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	219/213 (103%)	212 (97%)	7 (3%)	39	30
1	B	205/213 (96%)	202 (98%)	3 (2%)	65	62
1	C	204/213 (96%)	199 (98%)	5 (2%)	47	41
1	D	202/213 (95%)	196 (97%)	6 (3%)	41	33
All	All	830/852 (97%)	809 (98%)	21 (2%)	46	41

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	241	SER
1	D	39	SER
1	D	228	HIS
1	D	132	ASP
1	D	4	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	12	HIS
1	D	96	ASN
1	D	228	HIS
1	D	180	HIS
1	C	228	HIS

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)

Of 11 ligands modelled in this entry, 4 are monoatomic - leaving 7 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	$\#\|Z\| > 2$
5	EDO	B	302	-	3,3,3	0.45	0	2,2,2	0.33	0
4	ADE	A	303	-	9,11,11	1.51	2 (22%)	7,15,15	2.44	4 (57%)
3	PO4	A	302	2	4,4,4	0.91	0	6,6,6	0.43	0
5	EDO	D	302	-	3,3,3	0.53	0	2,2,2	0.38	0
5	EDO	C	304	-	3,3,3	0.61	0	2,2,2	0.24	0
5	EDO	C	303	-	3,3,3	0.59	0	2,2,2	0.41	0
5	EDO	C	302	-	3,3,3	0.32	0	2,2,2	0.67	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
5	EDO	B	302	-	-	0/1/1/1	-
4	ADE	A	303	-	-	-	0/2/2/2
5	EDO	D	302	-	-	1/1/1/1	-
5	EDO	C	304	-	-	0/1/1/1	-
5	EDO	C	303	-	-	1/1/1/1	-
5	EDO	C	302	-	-	1/1/1/1	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	303	ADE	C5-C4	2.84	1.48	1.40
4	A	303	ADE	C2-N3	2.31	1.35	1.32

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	303	ADE	N3-C2-N1	-3.81	122.72	128.68
4	A	303	ADE	C2-N3-C4	3.34	121.27	113.45
4	A	303	ADE	C4-C5-N7	-2.59	106.70	109.40
4	A	303	ADE	C2-N1-C6	2.26	122.62	118.75

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	C	303	EDO	O1-C1-C2-O2
5	D	302	EDO	O1-C1-C2-O2
5	C	302	EDO	O1-C1-C2-O2

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	302	EDO	1	0
4	A	303	ADE	3	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 (i)

6.1 Protein, DNA and RNA chains (i)

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, 95th 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	264/267 (98%)	-0.17	7 (2%)	54	57	19, 30, 52, 73	1 (0%)
1	B	257/267 (96%)	-0.12	5 (1%)	66	69	21, 34, 56, 75	0
1	C	255/267 (95%)	0.07	8 (3%)	49	51	26, 41, 64, 75	0
1	D	254/267 (95%)	0.51	24 (9%)	8	9	30, 48, 77, 92	0
All	All	1030/1068 (96%)	0.07	44 (4%)	35	38	19, 38, 66, 92	1 (0%)

The worst 5 of 44 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	152	THR	5.6
1	D	128	TYR	5.0
1	D	107	VAL	3.7
1	C	206	TYR	3.7
1	D	153	GLY	3.3

6.2 Non-standard residues in protein, DNA, RNA chains (i)

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, 95th 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
5	EDO	C	303	4/4	0.83	0.18	59,61,66,67	0
5	EDO	D	302	4/4	0.89	0.13	69,69,69,70	0
5	EDO	C	302	4/4	0.91	0.12	54,56,57,61	0
4	ADE	A	303	10/10	0.91	0.12	56,69,74,75	0
5	EDO	B	302	4/4	0.91	0.21	61,64,67,69	0
5	EDO	C	304	4/4	0.93	0.20	49,59,62,64	0
2	MG	D	301	1/1	0.94	0.06	43,43,43,43	0
3	PO4	A	302	5/5	0.95	0.12	25,35,38,40	4
2	MG	C	301	1/1	0.96	0.06	40,40,40,40	0
2	MG	B	301	1/1	0.98	0.04	32,32,32,32	0
2	MG	A	301	1/1	0.99	0.03	28,28,28,28	0

6.5 Other polymers (i)

There are no such residues in this entry.