



wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 06:24 am BST

PDB ID : 5SXE
Title : Crystal Structure of PI3Kalpha in complex with fragments 19 and 28
Authors : Gabelli, S.B.; Vogelstein, B.; Miller, M.S.; Amzel, L.M.
Deposited on : 2016-08-09
Resolution : 3.51 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) 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 : 2.11
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) : 2.11

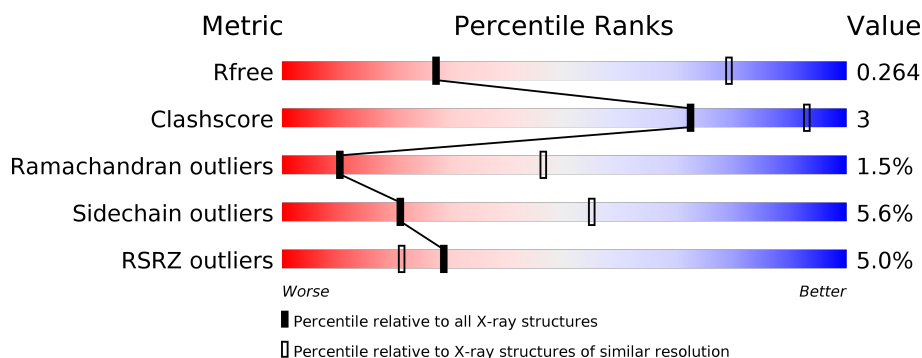
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.51 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	1161 (3.60-3.44)
Clashscore	141614	1244 (3.60-3.44)
Ramachandran outliers	138981	1206 (3.60-3.44)
Sidechain outliers	138945	1207 (3.60-3.44)
RSRZ outliers	127900	1080 (3.60-3.44)

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 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 $\leq 5\%$. 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	1096	<div> <div>4%</div> <div>80%</div> <div>14%</div> <div>• •</div> </div>
2	B	279	<div> <div>8%</div> <div>73%</div> <div>7%</div> <div>•</div> <div>19%</div> </div>

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
3	71G	A	1101	-	-	-	X

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 10518 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 Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic sub-unit alpha isoform.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	1048	Total	C	N	O	P	S	0	0	0
			8567	5468	1471	1556	2	70			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-27	MET	-	initiating methionine	UNP P42336
A	-26	SER	-	expression tag	UNP P42336
A	-25	TYR	-	expression tag	UNP P42336
A	-24	TYR	-	expression tag	UNP P42336
A	-23	HIS	-	expression tag	UNP P42336
A	-22	HIS	-	expression tag	UNP P42336
A	-21	HIS	-	expression tag	UNP P42336
A	-20	HIS	-	expression tag	UNP P42336
A	-19	HIS	-	expression tag	UNP P42336
A	-18	HIS	-	expression tag	UNP P42336
A	-17	ASP	-	expression tag	UNP P42336
A	-16	TYR	-	expression tag	UNP P42336
A	-15	ASP	-	expression tag	UNP P42336
A	-14	ILE	-	expression tag	UNP P42336
A	-13	PRO	-	expression tag	UNP P42336
A	-12	THR	-	expression tag	UNP P42336
A	-11	THR	-	expression tag	UNP P42336
A	-10	GLU	-	expression tag	UNP P42336
A	-9	ASN	-	expression tag	UNP P42336
A	-8	LEU	-	expression tag	UNP P42336
A	-7	TYR	-	expression tag	UNP P42336
A	-6	PHE	-	expression tag	UNP P42336
A	-5	GLN	-	expression tag	UNP P42336
A	-4	GLY	-	expression tag	UNP P42336
A	-3	ALA	-	expression tag	UNP P42336
A	-2	MET	-	expression tag	UNP P42336

Continued on next page...

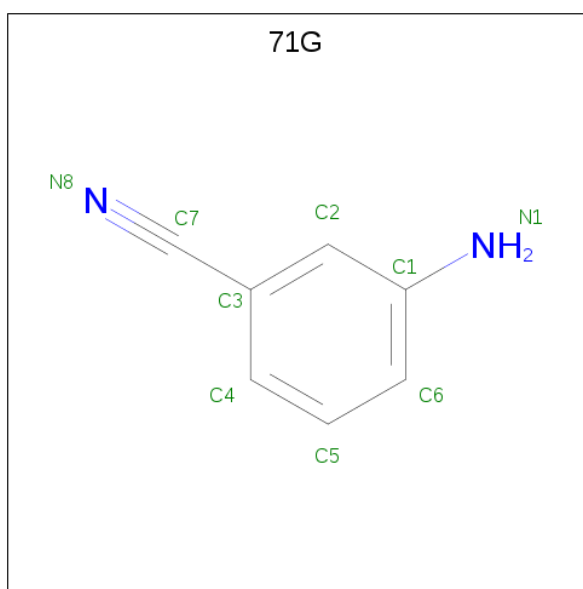
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP P42336
A	0	SER	-	expression tag	UNP P42336

- Molecule 2 is a protein called Phosphatidylinositol 3-kinase regulatory subunit alpha.

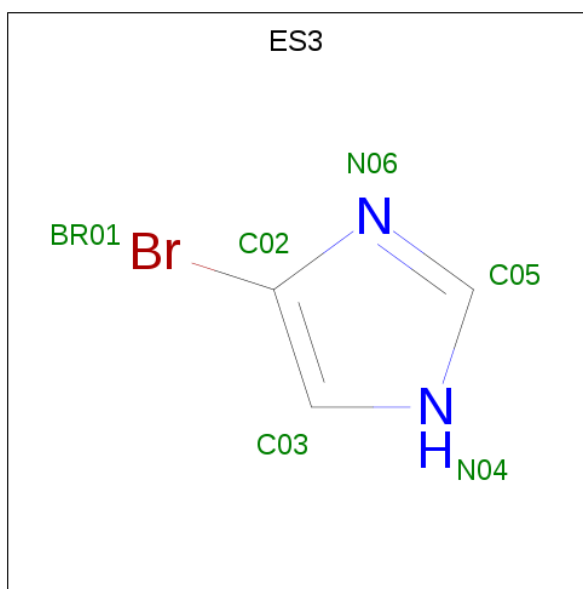
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	225	Total	C	N	O	S	0	0	0
			1936	1215	342	373	6			

- Molecule 3 is 3-aminobenzonitrile (three-letter code: 71G) (formula: $C_7H_6N_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	N	0	0
			9	7	2		

- Molecule 4 is 4-bromo-1H-imidazole (three-letter code: ES3) (formula: $C_3H_3BrN_2$).

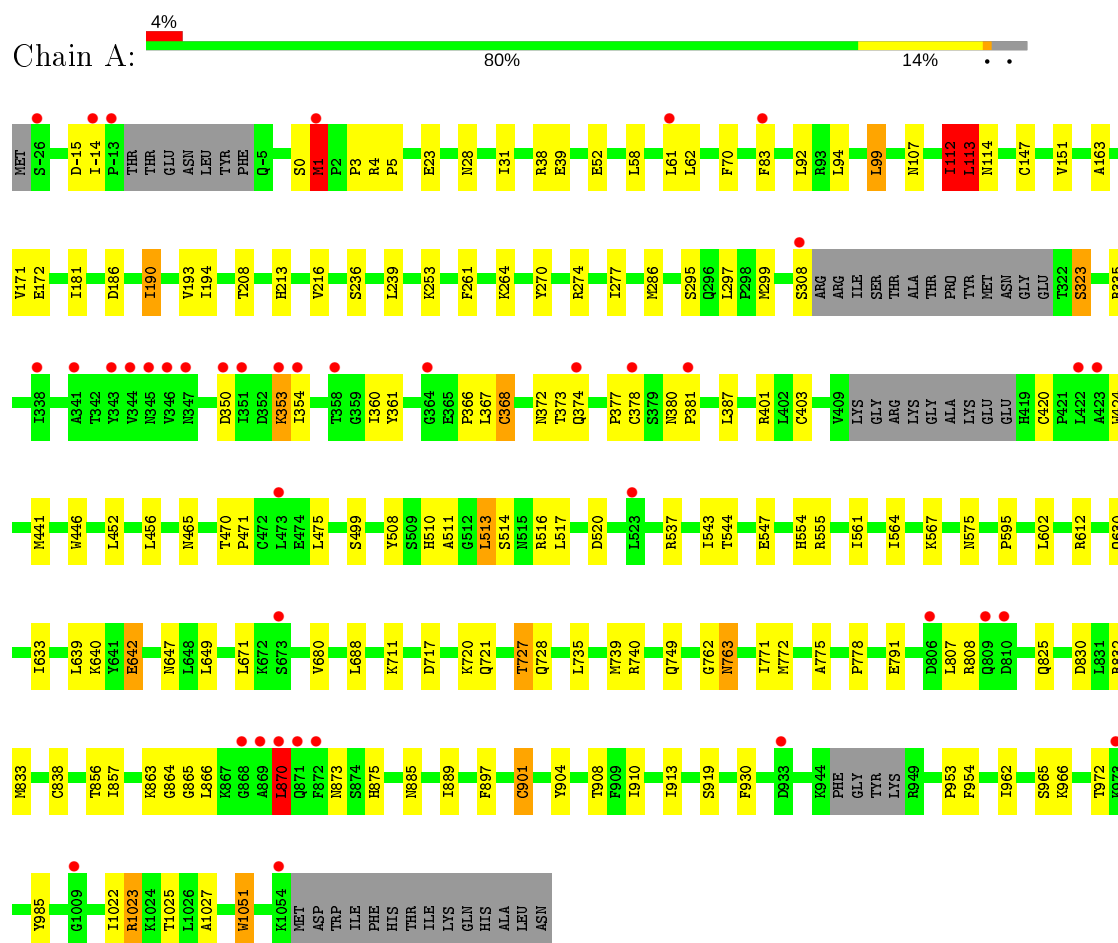


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	Br	C	N		
4	B	1	6	1	3	2	0	0

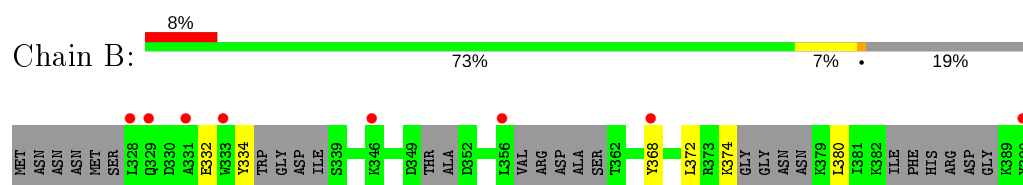
3 Residue-property plots [i](#)

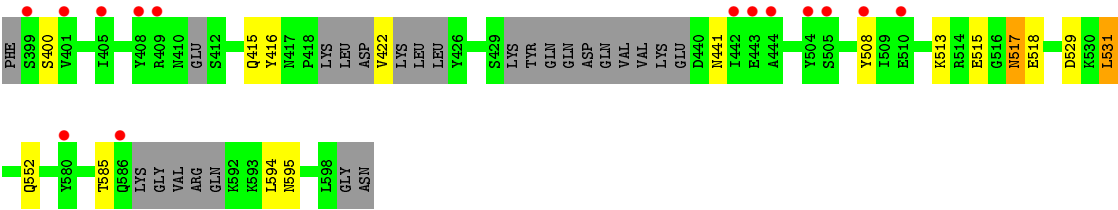
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit alpha isoform



- Molecule 2: Phosphatidylinositol 3-kinase regulatory subunit alpha





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	113.83Å 116.26Å 148.98Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	91.65 – 3.51 48.90 – 3.51	Depositor EDS
% Data completeness (in resolution range)	99.6 (91.65-3.51) 99.7 (48.90-3.51)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.05 (at 3.48Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.191 , 0.271 0.193 , 0.264	Depositor DCC
R_{free} test set	1306 reflections (5.16%)	wwPDB-VP
Wilson B-factor (Å ²)	118.7	Xtriage
Anisotropy	0.046	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 101.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.029 for k,h,-l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	10518	wwPDB-VP
Average B, all atoms (Å ²)	139.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.07% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: 71G, ES3, SEP

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	0.55	0/8743	0.78	5/11815 (0.0%)
2	B	0.51	0/1957	0.70	0/2603
All	All	0.54	0/10700	0.76	5/14418 (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	A	0	3

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	A	870	LEU	CA-CB-CG	6.67	130.65	115.30
1	A	113	LEU	CB-CG-CD1	5.61	120.54	111.00
1	A	513	LEU	CA-CB-CG	5.47	127.87	115.30
1	A	1	MET	CA-CB-CG	5.09	121.95	113.30
1	A	1023	ARG	NE-CZ-NH1	5.03	122.82	120.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	113	LEU	Peptide
1	A	323	SER	Peptide
1	A	511	ALA	Peptide

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	8567	0	8543	69	0
2	B	1936	0	1913	8	0
3	A	9	0	0	0	0
4	B	6	0	3	0	0
All	All	10518	0	10459	72	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.

The worst 5 of 72 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:962:ILE:O	1:A:965:SER:O	1.99	0.80
1:A:910:ILE:O	1:A:1025:THR:HG21	1.99	0.62
1:A:190:ILE:HD11	1:A:213:HIS:HA	1.80	0.61
1:A:261:PHE:HA	1:A:270:TYR:CE2	2.37	0.59
1:A:31:ILE:HD11	2:B:531:LEU:HD13	1.86	0.57

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	1036/1096 (94%)	918 (89%)	101 (10%)	17 (2%)	9 45
2	B	200/279 (72%)	185 (92%)	13 (6%)	2 (1%)	15 54

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	1236/1375 (90%)	1103 (89%)	114 (9%)	19 (2%)	10	46

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	3	PRO
1	A	378	CYS
1	A	514	SER
1	A	763	ASN
1	A	107	ASN

5.3.2 Protein sidechains ⓘ

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	956/997 (96%)	904 (95%)	52 (5%)	22	56
2	B	214/259 (83%)	201 (94%)	13 (6%)	18	52
All	All	1170/1256 (93%)	1105 (94%)	65 (6%)	21	55

5 of 65 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	441	MET
1	A	516	ARG
2	B	531	LEU
1	A	452	LEU
1	A	499	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	634	GLN
1	A	749	GLN
1	A	1042	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	575	ASN
2	B	517	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

2 non-standard protein/DNA/RNA residues 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
1	SEP	A	790	1	8,9,10	0.67	0	8,12,14	2.06	3 (37%)
1	SEP	A	7	1	8,9,10	0.79	0	8,12,14	1.28	1 (12%)

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
1	SEP	A	790	1	-	4/5/8/10	-
1	SEP	A	7	1	-	5/5/8/10	-

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	790	SEP	OG-CB-CA	4.22	112.25	108.14
1	A	7	SEP	P-OG-CB	2.94	126.39	118.30
1	A	790	SEP	P-OG-CB	2.34	124.75	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	790	SEP	O2P-P-OG	-2.21	100.84	106.73

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	790	SEP	N-CA-CB-OG
1	A	790	SEP	CB-OG-P-O2P
1	A	790	SEP	CB-OG-P-O3P
1	A	7	SEP	N-CA-CB-OG
1	A	7	SEP	CA-CB-OG-P

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

2 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	$\# Z > 2$
4	ES3	B	701	-	3,6,6	3.60	1 (33%)	2,7,7	1.97	1 (50%)
3	71G	A	1101	-	9,9,9	2.52	2 (22%)	11,11,11	0.68	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
4	ES3	B	701	-	-	-	0/1/1/1
3	71G	A	1101	-	-	0/2/2/2	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1101	71G	C3-C7	-6.65	1.29	1.44
4	B	701	ES3	BR01-C02	6.23	2.00	1.90
3	A	1101	71G	C1-N1	3.08	1.49	1.38

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	701	ES3	C03-N04-C05	2.46	109.61	105.78

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	1046/1096 (95%)	0.10	40 (3%) 40 30	73, 123, 193, 254	0
2	B	225/279 (80%)	0.58	23 (10%) 6 6	111, 189, 231, 275	0
All	All	1271/1375 (92%)	0.19	63 (4%) 28 21	73, 130, 212, 275	0

The worst 5 of 63 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	409	ARG	5.0
1	A	378	CYS	4.8
1	A	346	VAL	4.5
2	B	328	LEU	4.2
2	B	390	TYR	4.2

6.2 Non-standard residues in protein, DNA, RNA chains [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
1	SEP	A	7	10/11	0.64	0.29	163,190,249,258	0
1	SEP	A	790	10/11	0.94	0.13	94,118,199,203	0

6.3 Carbohydrates [i](#)

There are no carbohydrates 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(\AA^2)	Q<0.9
3	71G	A	1101	9/9	0.57	1.35	146,183,195,196	0
4	ES3	B	701	6/6	0.65	0.28	191,195,207,240	0

6.5 Other polymers [i](#)

There are no such residues in this entry.