wwPDB X-ray Structure Validation Summary Report (i)

Sep 23, 2023 – 09:07 PM EDT

PDB ID	:	5SXB
Title	:	Crystal Structure of PI3Kalpha in complex with fragment 23
Authors	:	Gabelli, S.B.; Vogelstein, B.; Miller, M.S.; Amzel, L.M.
Deposited on	:	2016-08-09
Resolution	:	3.30 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:

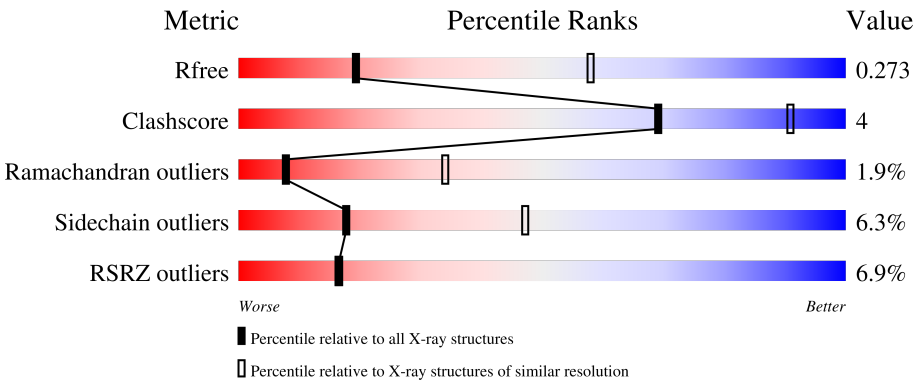
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 3.30 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	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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	1096	4% 80% 15% • •
2	B	279	16% 64% 10% • 24%

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	SOA	B	701	-	-	-	X

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 10504 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 subunit alpha isoform.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	P	S			
1	A	1058	8656	5526	1482	1576	2	70	0	0	0

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...

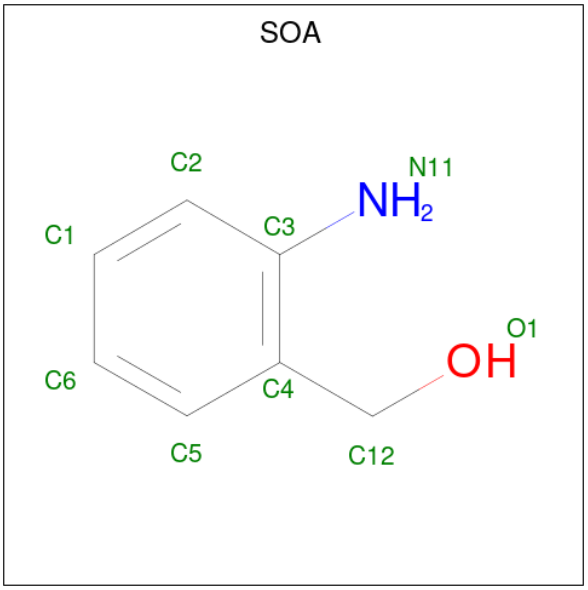
Continued from previous page...

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
			Total	C	N	O	S			
2	B	212	1839	1156	327	350	6	0	0	0

- Molecule 3 is ISATOIC ANHYDRIDE (three-letter code: SOA) (formula: C_7H_9NO).



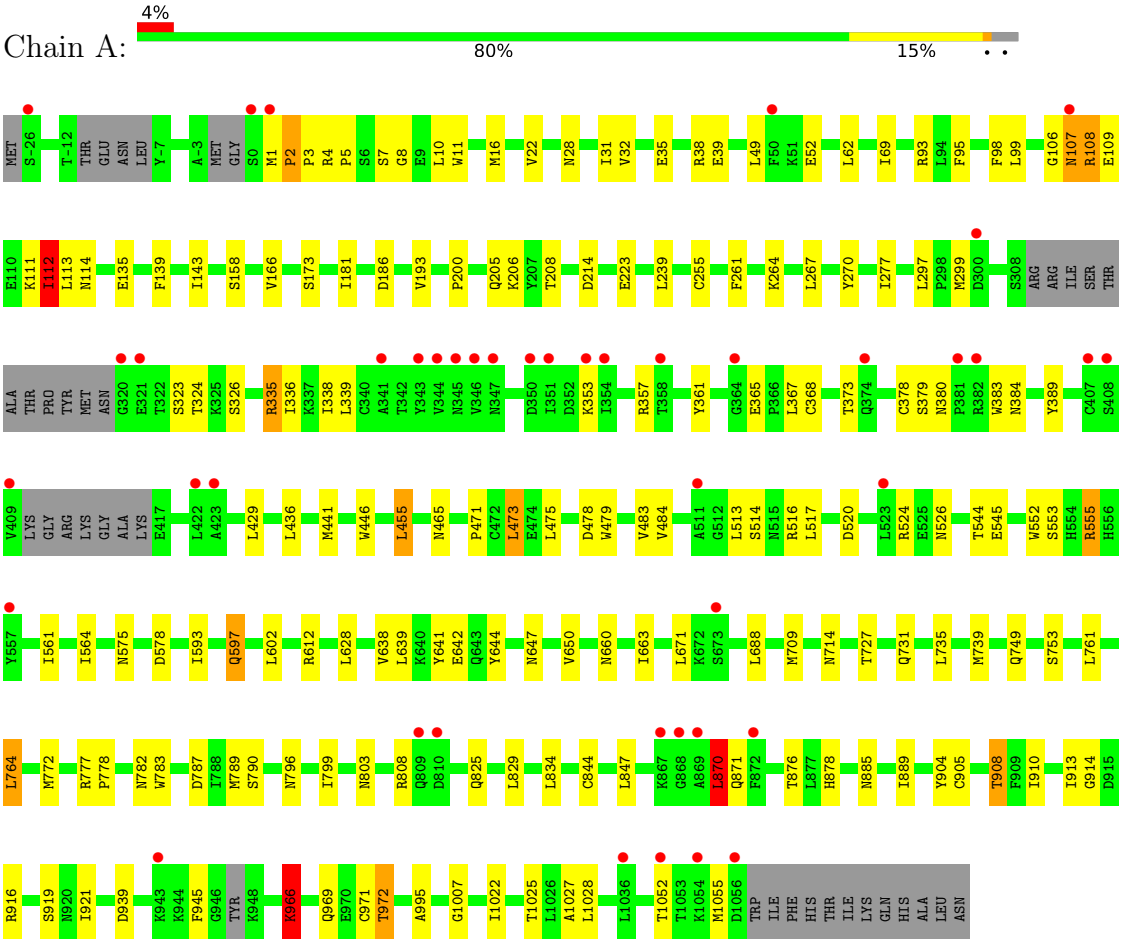
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	B	1	9	7	1	1	0	0

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.

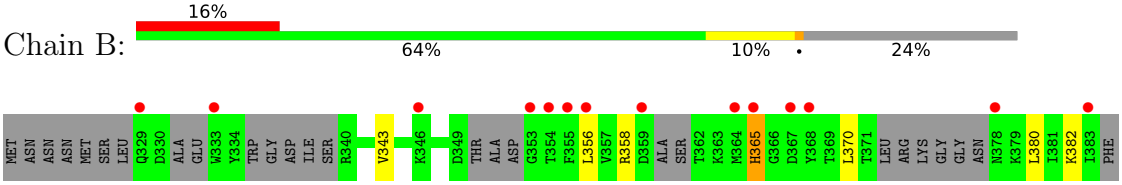
• Molecule 1: Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit alpha isoform

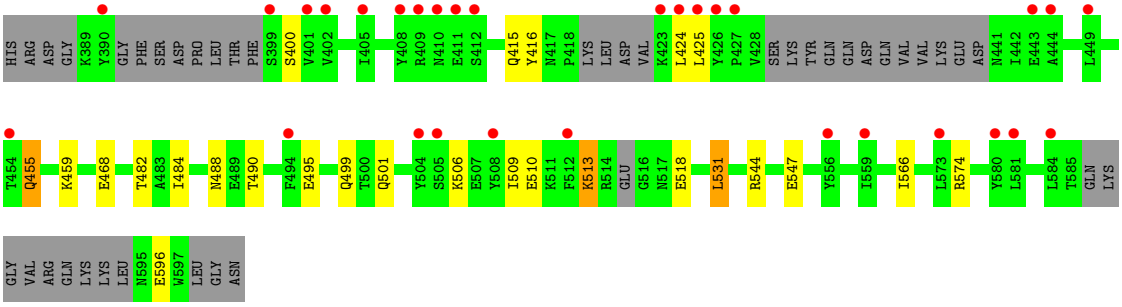
Chain A:



• Molecule 2: Phosphatidylinositol 3-kinase regulatory subunit alpha

Chain B:




HIS
ARG
ASP
GLY
K389
Y390
GLY
PHE
SER
ASP
PRO
LEU
THR
PHE
S399
S400
V401
V402
I405
Y408
R409
N410
E411
S412
Q415
Y416
N417
P418
LYS
LEU
ASP
VAL
K423
L424
L425
Y426
P427
V428
SER
LYS
TYR
GLN
GLN
ASP
GLN
VAL
VAL
LYS
GLU
ASP
N441
I442
E443
A444
L449
T454
Q455
K459
E468
T482
A483
I484
N488
E489
T490
F494
E495
Q499
T500
Q501
Y504
S505
K506
E507
Y508
I509
E510
K511
F512
K513
R514
GLU
G516
N517
E518
L531
R544
E547
Y556
I559
I566
L573
R574
Y580
L581
L584
T585
GLN
LYS
GLY
VAL
ARG
GLN
LYS
LYS
LEU
N595
E596
W597
LEU
GLY
ASN

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	114.46Å 116.77Å 150.10Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	92.16 – 3.30 49.14 – 3.30	Depositor EDS
% Data completeness (in resolution range)	99.8 (92.16-3.30) 99.6 (49.14-3.30)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.42 (at 3.33Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.204 , 0.278 0.203 , 0.273	Depositor DCC
R_{free} test set	1579 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å^2)	100.0	Xtriage
Anisotropy	0.139	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 79.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.012 for k,h,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	10504	wwPDB-VP
Average B, all atoms (Å^2)	121.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.14% 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: SOA, 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	0.54	0/8834	0.77	2/11935 (0.0%)
2	B	0.48	0/1860	0.67	0/2474
All	All	0.53	0/10694	0.75	2/14409 (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	2

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	A	870	LEU	CA-CB-CG	6.16	129.46	115.30
1	A	764	LEU	CA-CB-CG	5.61	128.20	115.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	112	ILE	Peptide
1	A	113	LEU	Peptide

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	8656	0	8617	67	0
2	B	1839	0	1816	14	0
3	B	9	0	9	0	0
All	All	10504	0	10442	77	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 4.

The worst 5 of 77 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:709:MET:HE1	1:A:847:LEU:HD21	1.66	0.76
1:A:561:ILE:O	1:A:564:ILE:HG22	1.96	0.65
1:A:28:ASN:ND2	1:A:62:LEU:HD21	2.14	0.63
1:A:642:GLU:HG2	1:A:647:ASN:CG	2.20	0.61
1:A:98:PHE:CE1	2:B:490:THR:HG23	2.35	0.61

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	1044/1096 (95%)	927 (89%)	96 (9%)	21 (2%)	7 32
2	B	188/279 (67%)	179 (95%)	7 (4%)	2 (1%)	14 45
All	All	1232/1375 (90%)	1106 (90%)	103 (8%)	23 (2%)	8 34

5 of 23 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2	PRO
1	A	107	ASN
1	A	323	SER
1	A	514	SER
1	A	38	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	
1	A	965/997 (97%)	905 (94%)	60 (6%)	18	47
2	B	202/259 (78%)	188 (93%)	14 (7%)	15	44
All	All	1167/1256 (93%)	1093 (94%)	74 (6%)	18	47

5 of 74 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	972	THR
2	B	531	LEU
1	A	1055	MET
2	B	455	GLN
1	A	339	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	795	GLN
1	A	917	HIS
1	A	885	ASN
1	A	1047	HIS
1	A	643	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)

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	0.60	0	8,12,14	1.52	1 (12%)
1	SEP	A	7	1	8,9,10	0.69	0	8,12,14	1.57	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	-	1/5/8/10	-
1	SEP	A	7	1	-	2/5/8/10	-

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	A	7	SEP	OG-CB-CA	-2.71	105.51	108.14
1	A	790	SEP	OG-CB-CA	2.28	110.36	108.14

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	7	SEP	N-CA-CB-OG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
1	A	790	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 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	$\#\|Z\| > 2$
3	SOA	B	701	-	9,9,9	1.26	1 (11%)	11,11,11	0.53	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
3	SOA	B	701	-	-	0/2/2/2	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	701	SOA	C3-N11	3.49	1.49	1.37

There are no bond angle outliers.

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	1056/1096 (96%)	0.22	43 (4%)	37	35	52, 103, 180, 220	0
2	B	212/279 (75%)	0.98	44 (20%)	1	1	91, 178, 226, 259	0
All	All	1268/1375 (92%)	0.34	87 (6%)	16	16	52, 111, 198, 259	0

The worst 5 of 87 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	409	ARG	5.8
2	B	405	ILE	5.6
1	A	381	PRO	5.4
1	A	1056	ASP	4.9
2	B	410	ASN	4.8

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	0.80	0.24	145,156,183,193	0
1	SEP	A	790	10/11	0.95	0.14	75,94,166,169	0

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
3	SOA	B	701	9/9	0.77	0.71	141,156,162,165	0

6.5 Other polymers (i)

There are no such residues in this entry.