



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 05:51 PM GMT

PDB ID : 4JPS  
Title : Co-crystal Structures of the Lipid Kinase PI3K alpha with Pan and Isoform Selective Inhibitors  
Authors : Knapp, M.S.; Elling, R.A.  
Deposited on : 2013-03-19  
Resolution : 2.20 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

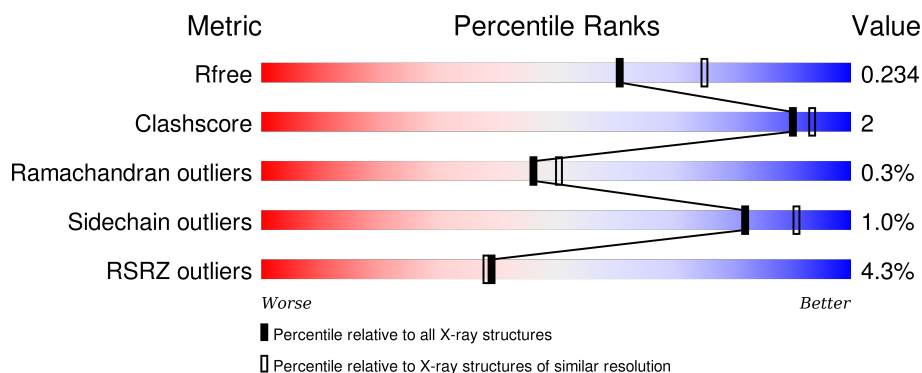
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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))
$R_{free}$	91344	3774 (2.20-2.20)
Clashscore	102246	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)
RSRZ outliers	91569	3781 (2.20-2.20)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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  $\leq 5\%$ . 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	1074	<div> <div></div> <div> <div></div> <div>86%</div> <div>6%</div> <div>7%</div> </div> </div>
2	B	293	<div> <div>13%</div> <div> <div></div> <div>70%</div> <div>6%</div> <div>23%</div> </div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 10533 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
1	A	999	Total	C	N	O	S	0	6	0
			8111	5179	1393	1474	65			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	232	LYS	MET	ENGINEERED MUTATION	UNP P42336
A	233	LYS	LEU	ENGINEERED MUTATION	UNP P42336
A	1069	HIS	-	EXPRESSION TAG	UNP P42336
A	1070	HIS	-	EXPRESSION TAG	UNP P42336
A	1071	HIS	-	EXPRESSION TAG	UNP P42336
A	1072	HIS	-	EXPRESSION TAG	UNP P42336
A	1073	HIS	-	EXPRESSION TAG	UNP P42336
A	1074	HIS	-	EXPRESSION TAG	UNP P42336

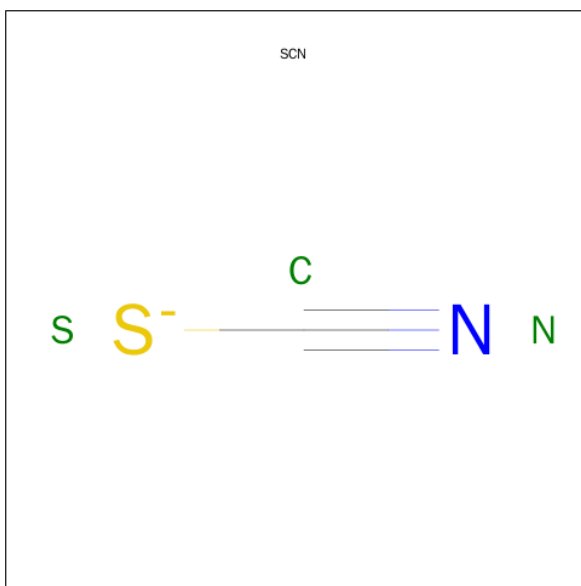
- Molecule 2 is a protein called Phosphatidylinositol 3-kinase regulatory subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	227	Total	C	N	O	S	0	1	0
			1883	1175	337	367	4			

There are 6 discrepancies between the modelled and reference sequences:

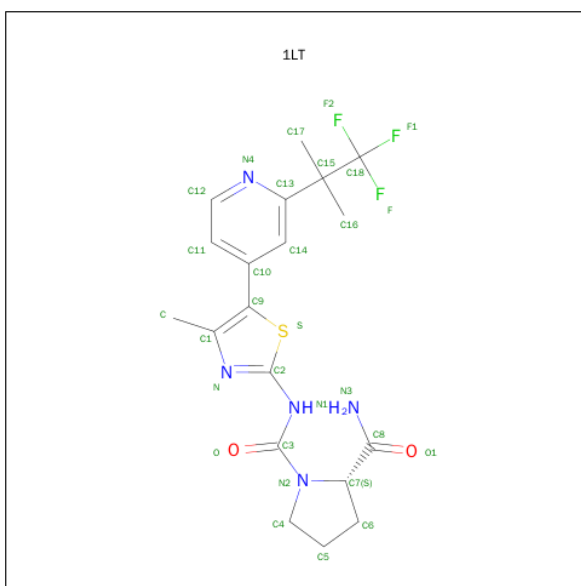
Chain	Residue	Modelled	Actual	Comment	Reference
B	301	MET	-	EXPRESSION TAG	UNP P27986
B	302	HIS	-	EXPRESSION TAG	UNP P27986
B	303	ASN	-	EXPRESSION TAG	UNP P27986
B	304	LEU	-	EXPRESSION TAG	UNP P27986
B	305	GLN	-	EXPRESSION TAG	UNP P27986
B	306	TYR	-	EXPRESSION TAG	UNP P27986

- Molecule 3 is THIOCYANATE ION (three-letter code: SCN) (formula: CNS).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	S	0	0
			3	1	1	1		

- Molecule 4 is (2S)-N 1 -{4-METHYL-5-[2-(1,1,1-TRIFLUORO-2-METHYLPROPAN-2-YL)PYRIDIN-4-YL]-1,3-THIAZOL-2-YL}PYRROLIDINE-1,2-DICARBOXAMIDE (three-letter code: 1LT) (formula: C<sub>19</sub>H<sub>22</sub>F<sub>3</sub>N<sub>5</sub>O<sub>2</sub>S).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
4	A	1	Total	C	F	N	O	S	0	0
			30	19	3	5	2	1		

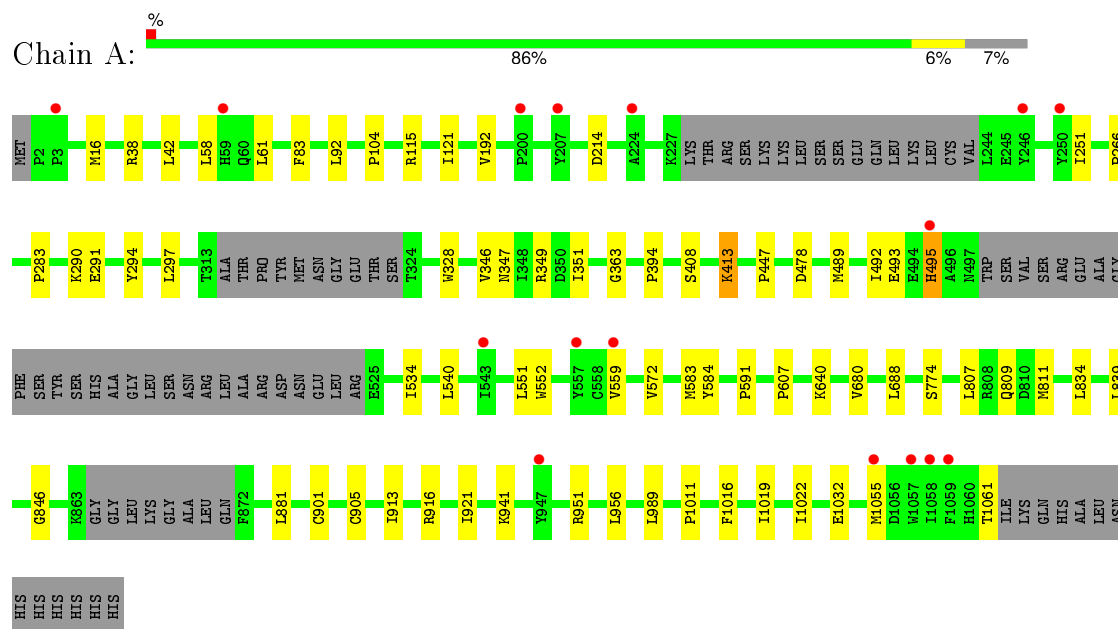
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	439	Total 439	O 439	0	0
5	B	67	Total 67	O 67	0	0

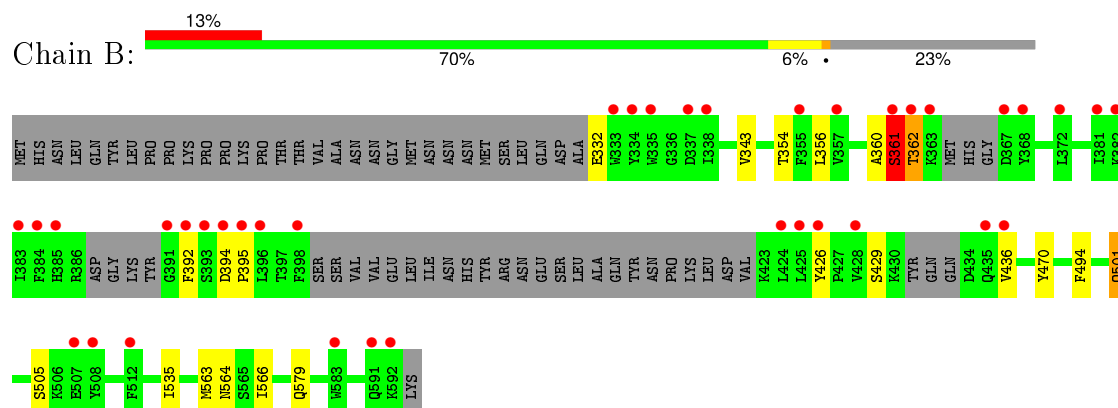
### 3 Residue-property plots

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.

- 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, $\alpha$ , $\beta$ , $\gamma$	104.86Å 106.19Å 133.67Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.20 74.62 – 2.00	Depositor EDS
% Data completeness (in resolution range)	96.6 (50.00-2.20) 97.0 (74.62-2.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.05	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.61 (at 2.00Å)	Xtriage
Refinement program	BUSTER 2.11.4	Depositor
R, $R_{free}$	0.205 , 0.228 0.208 , 0.234	Depositor DCC
$R_{free}$ test set	3663 reflections (5.23%)	DCC
Wilson B-factor (Å <sup>2</sup> )	41.4	Xtriage
Anisotropy	0.184	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 48.6	EDS
Estimated twinning fraction	0.014 for k,h,l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 98189 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	10533	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.56% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: 1LT, SCN

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	0.43	0/8299	0.61	0/11236
2	B	0.40	0/1909	0.60	1/2560 (0.0%)
All	All	0.43	0/10208	0.61	1/13796 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	361	SER	C-N-CA	6.20	137.21	121.70

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	8111	0	8001	35	0
2	B	1883	0	1796	11	0
3	A	3	0	0	0	0
4	A	30	0	22	1	0
5	A	439	0	0	1	0
5	B	67	0	0	1	0
All	All	10533	0	9819	46	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 2.

All (46) 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)
2:B:436:VAL:HB	2:B:579:GLN:HG2	1.68	0.75
1:A:121:ILE:HG12	1:A:688:LEU:HB3	1.71	0.71
2:B:394:ASP:HB3	2:B:395:PRO:HD3	1.76	0.68
1:A:807:LEU:HD12	1:A:846:GLY:HA3	1.81	0.62
2:B:332:GLU:N	2:B:429:SER:HG	1.99	0.60
1:A:42:LEU:HD21	1:A:92:LEU:HD11	1.85	0.58
1:A:58:LEU:HB3	1:A:61:LEU:HD12	1.87	0.57
1:A:251:ILE:HG23	1:A:290:LYS:HG2	1.88	0.54
1:A:294:TYR:HA	1:A:297:LEU:HD12	1.89	0.54
1:A:916:ARG:HB3	1:A:921:ILE:HD11	1.90	0.53
2:B:563:MET:HA	2:B:566:ILE:HG22	1.89	0.53
1:A:16:MET:HB2	1:A:38:ARG:HG3	1.92	0.52
1:A:492:ILE:HG21	1:A:584:TYR:HD2	1.75	0.52
2:B:494:PHE:HB3	2:B:535:ILE:HG12	1.94	0.50
1:A:351:ILE:HG21	1:A:408:SER:HB2	1.95	0.48
1:A:941:LYS:HD2	1:A:951:ARG:HD3	1.96	0.48
1:A:809[B]:GLN:HE21	1:A:1011:PRO:HD2	1.79	0.47
1:A:811:MET:HE3	1:A:839:LEU:HD12	1.98	0.46
2:B:360:ALA:C	2:B:362:THR:H	2.19	0.46
1:A:83:PHE:HE1	1:A:104:PRO:HB3	1.80	0.46
1:A:1019:ILE:O	1:A:1022:ILE:HG22	2.15	0.46
1:A:552:TRP:HZ3	1:A:583:MET:HE2	1.79	0.46
1:A:413:LYS:HD2	1:A:413:LYS:HA	1.77	0.46
2:B:354:THR:HA	2:B:426:TYR:O	2.16	0.46
1:A:774:SER:HB3	5:A:1555:HOH:O	2.15	0.45
1:A:347:ASN:OD1	1:A:349:ARG:HG2	2.16	0.45
1:A:328:TRP:HA	1:A:394:PRO:HB3	1.99	0.45
1:A:905[B]:CYS:SG	1:A:956:LEU:HD11	2.56	0.45
1:A:534:ILE:HG21	1:A:551:LEU:HD11	1.99	0.45
1:A:192:VAL:HG12	1:A:283:PRO:HG2	1.99	0.45
1:A:572:VAL:HG21	1:A:583:MET:HG2	2.00	0.44
1:A:346:VAL:HB	1:A:351:ILE:HD13	2.00	0.43
1:A:447:PRO:HB3	2:B:470:TYR:OH	2.19	0.43
2:B:343:VAL:HG22	2:B:356:LEU:HD11	2.01	0.43
2:B:501:GLN:O	2:B:505:SER:HB2	2.19	0.42
1:A:640:LYS:HE2	1:A:680:VAL:HG11	2.00	0.42
1:A:989:LEU:HD22	1:A:1032:GLU:HG2	2.01	0.42

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:540:LEU:HD21	1:A:1016:PHE:CD1	2.55	0.42
2:B:564:ASN:HA	5:B:665:HOH:O	2.20	0.41
1:A:363:GLY:N	1:A:607:PRO:HG3	2.35	0.41
1:A:881:LEU:HA	1:A:881:LEU:HD12	1.95	0.41
4:A:1102:1LT:S	4:A:1102:1LT:O	2.78	0.41
1:A:214:ASP:HA	1:A:266:PRO:HB3	2.00	0.41
1:A:559:VAL:HG13	1:A:591:PRO:HD3	2.03	0.41
1:A:291:GLU:HA	1:A:294:TYR:HB2	2.01	0.41
1:A:489:MET:O	1:A:493:GLU:HG3	2.22	0.40

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	995/1074 (93%)	967 (97%)	26 (3%)	2 (0%)	52	59
2	B	218/293 (74%)	212 (97%)	4 (2%)	2 (1%)	21	19
All	All	1213/1367 (89%)	1179 (97%)	30 (2%)	4 (0%)	46	50

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	362	THR
1	A	495	HIS
2	B	361	SER
1	A	1055	MET

### 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	894/980 (91%)	886 (99%)	8 (1%)	84	92
2	B	197/272 (72%)	194 (98%)	3 (2%)	72	84
All	All	1091/1252 (87%)	1080 (99%)	11 (1%)	82	91

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	115	ARG
1	A	413	LYS
1	A	478	ASP
1	A	495	HIS
1	A	834	LEU
1	A	901	CYS
1	A	913	ILE
1	A	1061	THR
2	B	361	SER
2	B	392	PHE
2	B	501	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	701	HIS
1	A	796	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

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

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	$\# Z  > 2$
3	SCN	A	1101	-	2,2,2	1.87	1 (50%)	1,1,1	0.54	0
4	1LT	A	1102	-	27,32,32	0.63	1 (3%)	25,49,49	1.36	2 (8%)

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	SCN	A	1101	-	-	0/0/0/0	0/0/0/0
4	1LT	A	1102	-	-	0/27/41/41	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1101	SCN	C-S	2.62	1.80	1.63
4	A	1102	1LT	C3-N1	2.72	1.40	1.36

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1102	1LT	C2-N1-C3	-4.26	124.42	130.00
4	A	1102	1LT	C10-C9-C1	4.84	134.55	127.03

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1102	1LT	1	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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	999/1074 (93%)	-0.01	16 (1%) 74 73	27, 47, 76, 112	0
2	B	227/293 (77%)	0.69	37 (16%) 2 2	35, 63, 111, 138	0
All	All	1226/1367 (89%)	0.12	53 (4%) 39 38	27, 49, 89, 138	0

All (53) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	361	SER	6.2
2	B	385	HIS	6.1
1	A	1055	MET	5.9
2	B	384	PHE	5.8
2	B	383	ILE	5.1
1	A	246	TYR	5.1
2	B	394	ASP	4.5
2	B	362	THR	4.5
2	B	338	ILE	4.5
2	B	428	VAL	4.1
2	B	391	GLY	4.0
2	B	424	LEU	4.0
2	B	396	LEU	4.0
2	B	333	TRP	3.9
1	A	557	TYR	3.9
1	A	559	VAL	3.8
2	B	368	TYR	3.7
2	B	335	TRP	3.7
2	B	398	PHE	3.6
2	B	337	ASP	3.5
1	A	1059	PHE	3.5
2	B	583	TRP	3.5
2	B	367	ASP	3.4
2	B	508	TYR	3.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	436	VAL	3.4
2	B	355	PHE	3.3
2	B	507	GLU	3.3
2	B	393	SER	3.3
1	A	224	ALA	3.3
2	B	591	GLN	3.1
2	B	381	ILE	3.0
2	B	512	PHE	3.0
1	A	200	PRO	3.0
2	B	334	TYR	2.8
2	B	357	VAL	2.7
1	A	1057	TRP	2.7
2	B	395	PRO	2.7
2	B	372	LEU	2.7
1	A	207	TYR	2.7
2	B	425	LEU	2.6
2	B	592	LYS	2.6
1	A	250	TYR	2.6
1	A	947	TYR	2.5
2	B	382	LYS	2.4
1	A	3	PRO	2.4
1	A	495	HIS	2.4
1	A	1058	ILE	2.3
2	B	435	GLN	2.2
2	B	363	LYS	2.2
2	B	392	PHE	2.0
1	A	543	ILE	2.0
1	A	59	HIS	2.0
2	B	426	TYR	2.0

## 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 ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	1LT	A	1102	30/30	0.97	0.11	-0.90	27,31,41,43	0
3	SCN	A	1101	3/3	0.91	0.16	-	71,71,72,73	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.