



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 9, 2017 – 10:40 AM EST

PDB ID : 5SWO
Title : Crystal Structure of PI3Kalpha in complex with fragments 4 and 19
Authors : Gabelli, S.B.; Vogelstein, B.; Miller, M.S.; Amzel, L.M.
Deposited on : 2016-08-08
Resolution : 3.50 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20028442
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)	:	rb-20028442

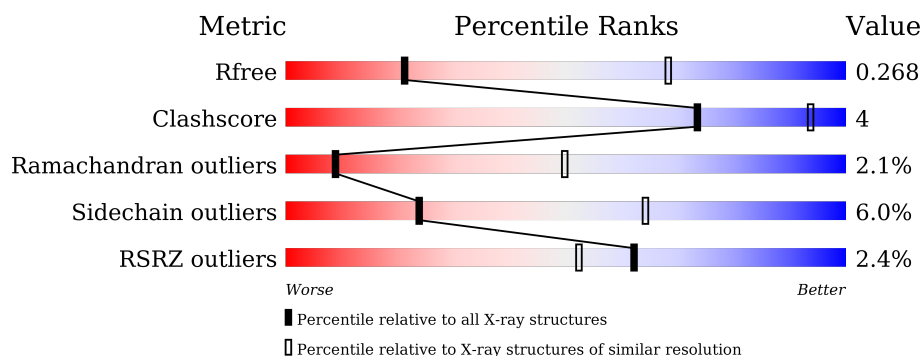
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.50 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	1051 (3.60-3.40)
Clashscore	102246	1157 (3.60-3.40)
Ramachandran outliers	100387	1120 (3.60-3.40)
Sidechain outliers	100360	1121 (3.60-3.40)
RSRZ outliers	91569	1058 (3.60-3.40)

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 $\leq 5\%$. 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 style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 16%, green 78%, grey 6%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> % 78% 16% ... </div> </div>
2	B	279	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 7%, orange 1%, yellow 11%, green 84%, grey 1%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> 7% 84% 11% ... </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
4	70T	A	1102	-	-	-	X
5	2ZV	B	701	-	-	-	X

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 10899 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	1051	Total	C	N	O	P	S	0	0	0
			8593	5489	1471	1563	1	69			

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

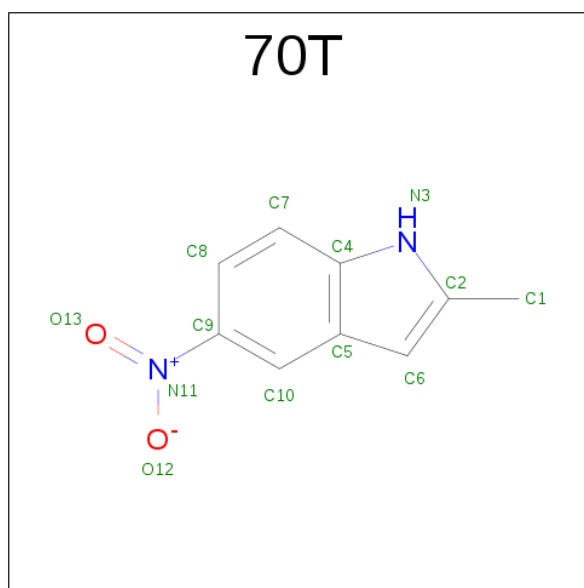
- Molecule 2 is a protein called Phosphatidylinositol 3-kinase regulatory subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	267	Total	C	N	O	S	0	0	0
			2281	1433	407	435	6			

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

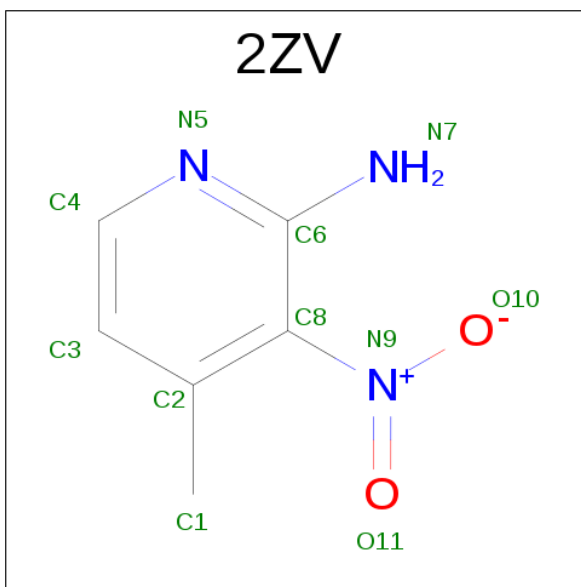
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Cl	0	0
			1	1		

- Molecule 4 is 2-methyl-5-nitro-1H-indole (three-letter code: 70T) (formula: C₉H₈N₂O₂).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			13	9	2	2		

- Molecule 5 is 4-methyl-3-nitropyridin-2-amine (three-letter code: 2ZV) (formula: C₆H₇N₃O₂).

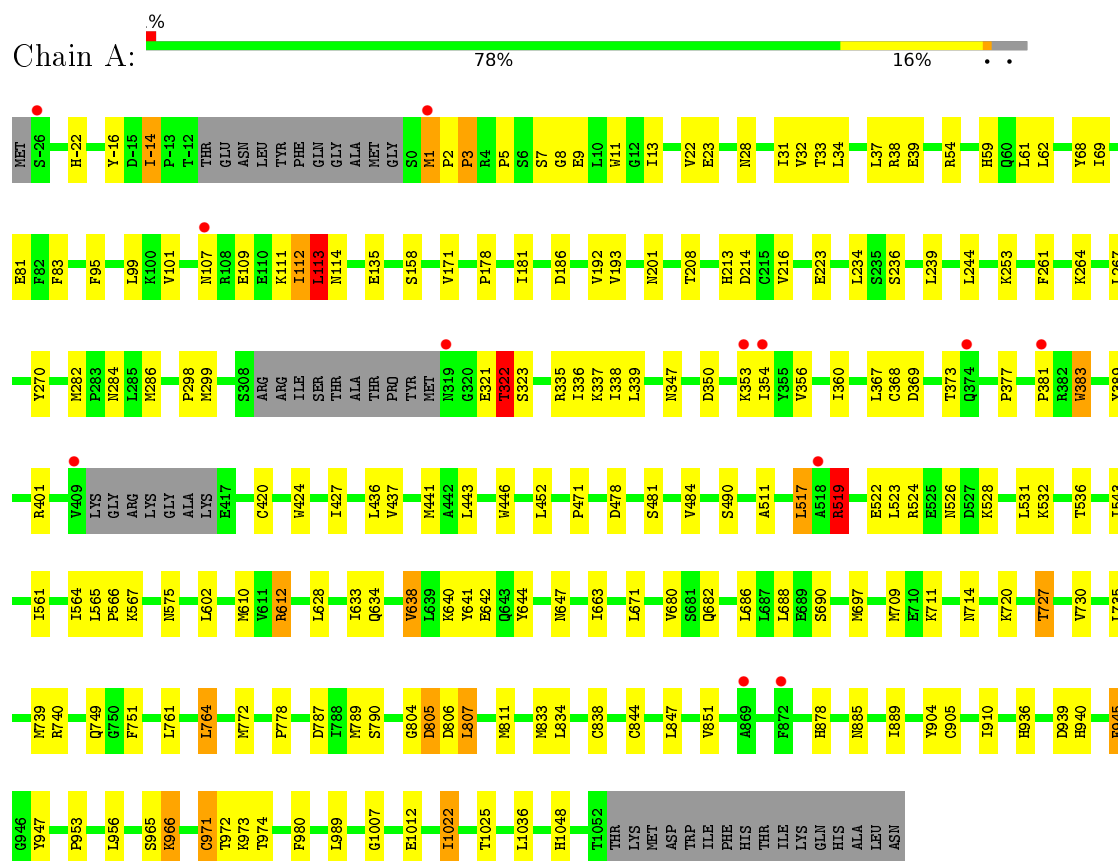


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
5	B	1	11	6	3	2	0	0

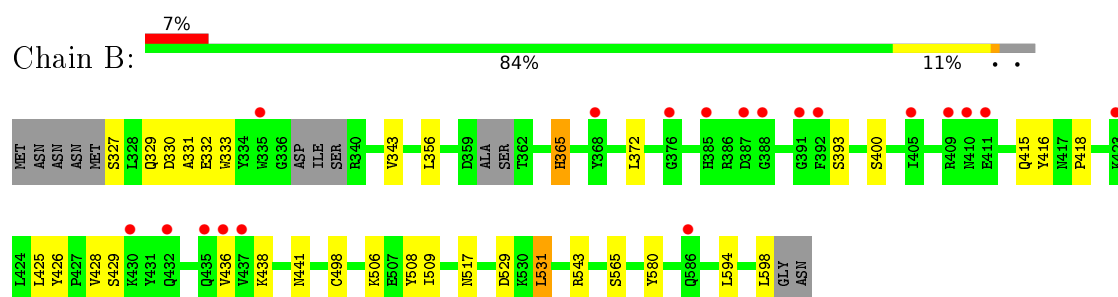
3 Residue-property plots

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, α , β , γ	114.83Å 116.36Å 149.44Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	91.81 – 3.50 48.68 – 3.50	Depositor EDS
% Data completeness (in resolution range)	99.3 (91.81-3.50) 99.5 (48.68-3.50)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.91 (at 3.48Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.189 , 0.274 0.190 , 0.268	Depositor DCC
R_{free} test set	1322 reflections (5.44%)	DCC
Wilson B-factor (Å ²)	99.6	Xtriage
Anisotropy	0.252	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 76.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.022 for k,h,-l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	10899	wwPDB-VP
Average B, all atoms (Å ²)	117.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.06% 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: 70T, 2ZV, CL, 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	0/8783	0.78	3/11874 (0.0%)
2	B	0.49	0/2319	0.71	0/3103
All	All	0.53	0/11102	0.77	3/14977 (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	1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	A	612	ARG	NE-CZ-NH2	-5.68	117.46	120.30
1	A	764	LEU	CA-CB-CG	5.55	128.06	115.30
1	A	113	LEU	CB-CG-CD1	5.05	119.59	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	113	LEU	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	8593	0	8554	82	0
2	B	2281	0	2262	11	0
3	A	1	0	0	0	0
4	A	13	0	0	0	0
5	B	11	0	7	1	0
All	All	10899	0	10823	89	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 89 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)
5:B:701:2ZV:O11	5:B:701:2ZV:N9	1.58	1.37
1:A:31:ILE:HD11	2:B:531:LEU:HD13	1.73	0.68
1:A:543:ILE:HD11	1:A:567:LYS:HD3	1.79	0.64
1:A:519:ARG:HB2	1:A:523:LEU:HA	1.81	0.63
1:A:640:LYS:HE2	1:A:680:VAL:HG11	1.81	0.63

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	1042/1096 (95%)	918 (88%)	103 (10%)	21 (2%)	<div><div>9</div><div>51</div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	261/279 (94%)	228 (87%)	27 (10%)	6 (2%)	8	48
All	All	1303/1375 (95%)	1146 (88%)	130 (10%)	27 (2%)	9	50

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	3	PRO
1	A	945	PHE
1	A	966	LYS
1	A	1	MET
1	A	322	THR

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	960/998 (96%)	902 (94%)	58 (6%)	24	64
2	B	249/259 (96%)	235 (94%)	14 (6%)	26	66
All	All	1209/1257 (96%)	1137 (94%)	72 (6%)	24	64

5 of 72 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	420	CYS
1	A	522	GLU
2	B	441	ASN
1	A	437	VAL
1	A	478	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	575	ASN
1	A	597	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	749	GLN
1	A	526	ASN
1	A	714	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue 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	$\# Z > 2$
1	SEP	A	790	1	7,9,10	0.77	0	8,12,14	1.91	2 (25%)

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	-	0/5/8/10	0/0/0/0

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	790	SEP	O2P-P-OG	-2.21	100.28	106.72
1	A	790	SEP	OG-CB-CA	4.05	111.79	108.26

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
1	A	790	SEP	1	0

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 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$
4	70T	A	1102	-	11,14,14	1.08	2 (18%)	14,20,20	2.30	3 (21%)
5	2ZV	B	701	-	9,11,11	6.31	3 (33%)	10,15,15	1.37	1 (10%)

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	70T	A	1102	-	-	0/4/4/4	0/2/2/2
5	2ZV	B	701	-	-	0/2/4/4	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	701	2ZV	C1-C2	-4.81	1.41	1.51
4	A	1102	70T	C7-C8	2.05	1.40	1.36
4	A	1102	70T	C10-C9	2.16	1.41	1.36
5	B	701	2ZV	C4-N5	3.39	1.42	1.34

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	701	2ZV	O11-N9	17.89	1.58	1.22

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1102	70T	C1-C2-C6	-4.45	116.54	128.96
5	B	701	2ZV	C3-C4-N5	-2.61	120.82	123.88
4	A	1102	70T	O13-N11-C9	3.72	121.34	118.67
4	A	1102	70T	C1-C2-N3	5.12	132.65	119.95

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
5	B	701	2ZV	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 [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	1050/1096 (95%)	-0.12	12 (1%) 82 73	53, 100, 162, 231	0
2	B	267/279 (95%)	0.32	19 (7%) 19 15	97, 162, 206, 228	0
All	All	1317/1375 (95%)	-0.03	31 (2%) 62 52	53, 108, 187, 231	0

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	319	ASN	4.9
2	B	391	GLY	4.5
1	A	872	PHE	4.4
1	A	869	ALA	4.0
1	A	-26	SER	3.7

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. 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, 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	LLDF	B-factors(Å ²)	Q<0.9
1	SEP	A	790	10/11	0.93	0.20	-	82,100,161,174	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. 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, 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	LLDF	B-factors(\AA^2)	Q<0.9
4	70T	A	1102	13/13	0.86	0.50	5.09	109,132,144,144	0
5	2ZV	B	701	11/11	0.66	0.73	3.33	153,178,200,201	0
3	CL	A	1101	1/1	0.96	0.23	-0.80	71,71,71,71	0

6.5 Other polymers [i](#)

There are no such residues in this entry.