WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Oct 18, 2021 – 10:49 am BST

PDB ID   : 7ONS
Title   : PARP1 catalytic domain in complex with isoquinolone-based inhibitor (compound 16)
Authors   : Schimpl, M.; Balazs, A.; Barratt, D.; Bista, M.; Chuba, M.; Degorce, S.L.; Di Fruscia, P.; Embrey, K.; Ghosh, A.; Gill, S.; Gunnarsson, A.; Hande, S.; Hemsley, P.; Heightman, T.D.; Illuzzi, G.; Lane, J.; Larner, C.; Leo, E.; Madin, A.; Martin, S.; McWilliams, L.; Orme, J.; Pachl, F.; Packer, M.; Pike, A.; Staniszewska, A.D.; Talbot, V.; Underwood, E.; Varnes, G.J.; Zhang, A.; Zheng, X.; Johannes, J.W.
Deposited on : 2021-05-25
Resolution : 1.97 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*
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 ①) were used in the production of this report:

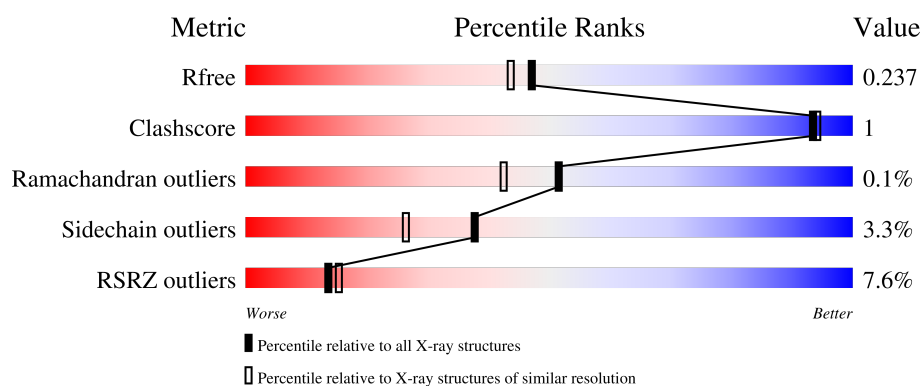|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 1.97 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	Percentile Ranks	Value
Rfree		0.237
Clashscore		1
Ramachandran outliers		0.1%
Sidechain outliers		3.3%
RSRZ outliers		7.6%

Worse .... *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	11647 (2.00-1.96)
Clashscore	141614	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)
RSRZ outliers	127900	11410 (2.00-1.96)

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	352	4% / 92% / 7% ·
1	B	352	11% / 93% / 6% ·

WORLDWIDE
wwPDB
PROTEIN DATA BANK

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 5954 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 Poly [ADP-ribose] polymerase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	350	2747	1749	463	524	11	0	1	0
1	B	348	Total	C	N	O	S	0	0	0
			2710	1727	456	516	11			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	660	GLY	-	expression tag	UNP P09874
A	661	SER	-	expression tag	UNP P09874
A	762	ALA	VAL	engineered mutation	UNP P09874
B	660	GLY	-	expression tag	UNP P09874
B	661	SER	-	expression tag	UNP P09874
B	762	ALA	VAL	engineered mutation	UNP P09874

- Molecule 2 is 7-[[4-(1,5-dimethylimidazol-2-yl)piperazin-1-yl]methyl]-3-ethyl-1 {H}-quinolin-2-one (three-letter code: VKT) (formula: $C_{21}H_{27}N_5O$).

VKT
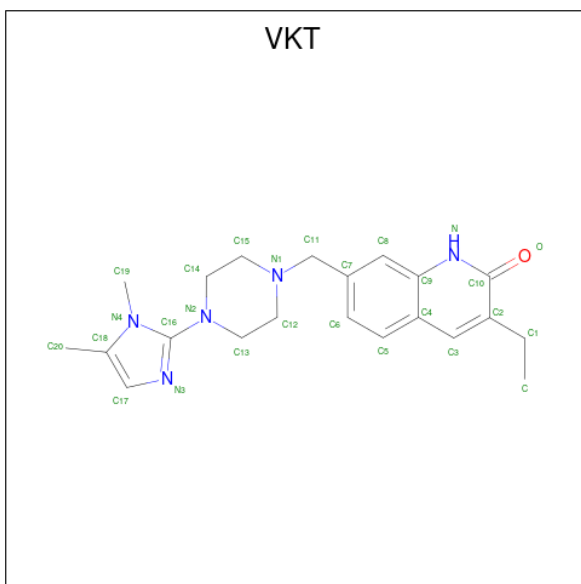
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total 27	C 21	N 5	O 1	0	0
2	B	1	Total 27	C 21	N 5	O 1	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

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

- Molecule 4 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	250	Total 250	O 250	0	0
4	B	168	Total 168	O 168	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: Poly [ADP-ribose] polymerase 1
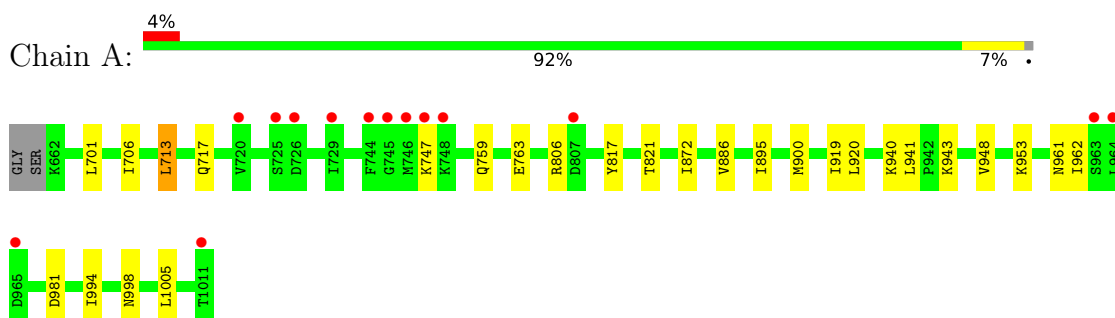
• Molecule 1: Poly [ADP-ribose] polymerase 1
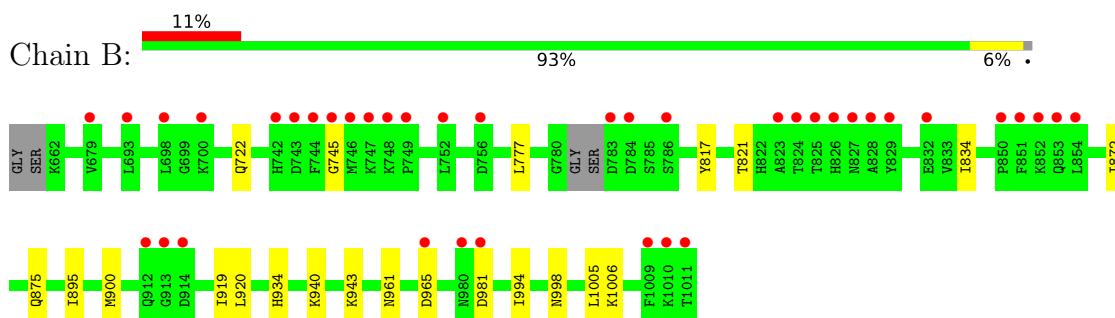
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	48.49Å 93.01Å 163.26Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	80.82 – 1.97 80.82 – 1.97	Depositor EDS
% Data completeness (in resolution range)	98.1 (80.82-1.97) 98.1 (80.82-1.97)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.81 (at 1.97Å)	Xtriage
Refinement program	BUSTER	Depositor
R, R_{free}	0.195 , 0.223 0.208 , 0.237	Depositor DCC
R_{free} test set	2584 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å^2)	17.0	Xtriage
Anisotropy	0.548	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< |L| > = 0.46, < L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	5954	wwPDB-VP
Average B, all atoms (Å^2)	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.43% 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 ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: SO4, VKT

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	0.47	0/2800	0.62	0/3783
1	B	0.47	0/2760	0.62	0/3728
All	All	0.47	0/5560	0.62	0/7511

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	2747	0	2767	10	0
1	B	2710	0	2722	7	0
2	A	27	0	0	0	0
2	B	27	0	0	0	0
3	A	15	0	0	0	0
3	B	10	0	0	0	0
4	A	250	0	0	0	0
4	B	168	0	0	0	0
All	All	5954	0	5489	15	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 1.

All (15) 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:953:LYS:HE2	1:B:722:GLN:HE22	1.42	0.83
1:B:872:ILE:HG21	1:B:920:LEU:HD11	1.88	0.55
1:A:872:ILE:HG21	1:A:920:LEU:HD11	1.89	0.54
1:B:821:THR:HB	1:B:900:MET:HA	1.94	0.49
1:A:895:ILE:HD11	1:A:994:ILE:HG22	1.95	0.48
1:A:821:THR:HB	1:A:900:MET:HA	1.94	0.48
1:B:895:ILE:HD11	1:B:994:ILE:HG22	1.97	0.47
1:A:886:VAL:HG11	1:B:934:HIS:ND1	2.35	0.42
1:A:713:LEU:HD12	1:A:713:LEU:HA	1.88	0.42
1:B:919:ILE:HG21	1:B:1005:LEU:HD11	2.01	0.42
1:A:701:LEU:HD22	1:A:706:ILE:HD11	2.01	0.41
1:A:759:GLN:O	1:A:763:GLU:HG3	2.20	0.41
1:A:941:LEU:HD21	1:A:948:VAL:HG23	2.03	0.41
1:B:834:ILE:HD11	1:B:1006:LYS:HB2	2.03	0.41
1:A:919:ILE:HG21	1:A:1005:LEU:HD11	2.02	0.41

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	349/352 (99%)	346 (99%)	3 (1%)	0	100	100
1	B	344/352 (98%)	341 (99%)	2 (1%)	1 (0%)	41	29
All	All	693/704 (98%)	687 (99%)	5 (1%)	1 (0%)	51	42

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	745	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	304/308 (99%)	293 (96%)	11 (4%)	35	23
1	B	297/308 (96%)	288 (97%)	9 (3%)	41	29
All	All	601/616 (98%)	581 (97%)	20 (3%)	38	26

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	713	LEU
1	A	717	GLN
1	A	747	LYS
1	A	806	ARG
1	A	817	TYR
1	A	940	LYS
1	A	943	LYS
1	A	961	ASN
1	A	962	ILE
1	A	981	ASP
1	A	998	ASN
1	B	777	LEU
1	B	817	TYR
1	B	875	GLN
1	B	940	LYS
1	B	943	LYS
1	B	961	ASN
1	B	965	ASP
1	B	981	ASP
1	B	998	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	906	ASN
1	A	998	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	722	GLN
1	B	906	ASN
1	B	998	ASN

5.3.3　RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5　Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6　Ligand geometry ⓘ

7 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	SO4	B	1102	-	4,4,4	0.30	0	6,6,6	0.34	0
3	SO4	A	1104	-	4,4,4	0.21	0	6,6,6	0.11	0
2	VKT	A	1101	-	26,30,30	1.19	4 (15%)	33,43,43	2.07	8 (24%)
3	SO4	B	1103	-	4,4,4	0.22	0	6,6,6	0.26	0
3	SO4	A	1102	-	4,4,4	0.30	0	6,6,6	0.32	0
3	SO4	A	1103	-	4,4,4	0.27	0	6,6,6	0.15	0
2	VKT	B	1101	-	26,30,30	1.11	3 (11%)	33,43,43	2.02	6 (18%)

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
2	VKT	B	1101	-	-	0/6/20/20	0/4/4/4
2	VKT	A	1101	-	-	0/6/20/20	0/4/4/4

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1101	VKT	C9-N	2.81	1.40	1.35
2	A	1101	VKT	C10-C2	2.60	1.47	1.41
2	B	1101	VKT	C9-N	2.50	1.39	1.35
2	B	1101	VKT	C10-C2	2.41	1.46	1.41
2	A	1101	VKT	C16-N3	-2.29	1.31	1.35
2	B	1101	VKT	C16-N3	-2.25	1.31	1.35
2	A	1101	VKT	C10-N	2.02	1.36	1.33

All (14) bond angle outliers are listed below:

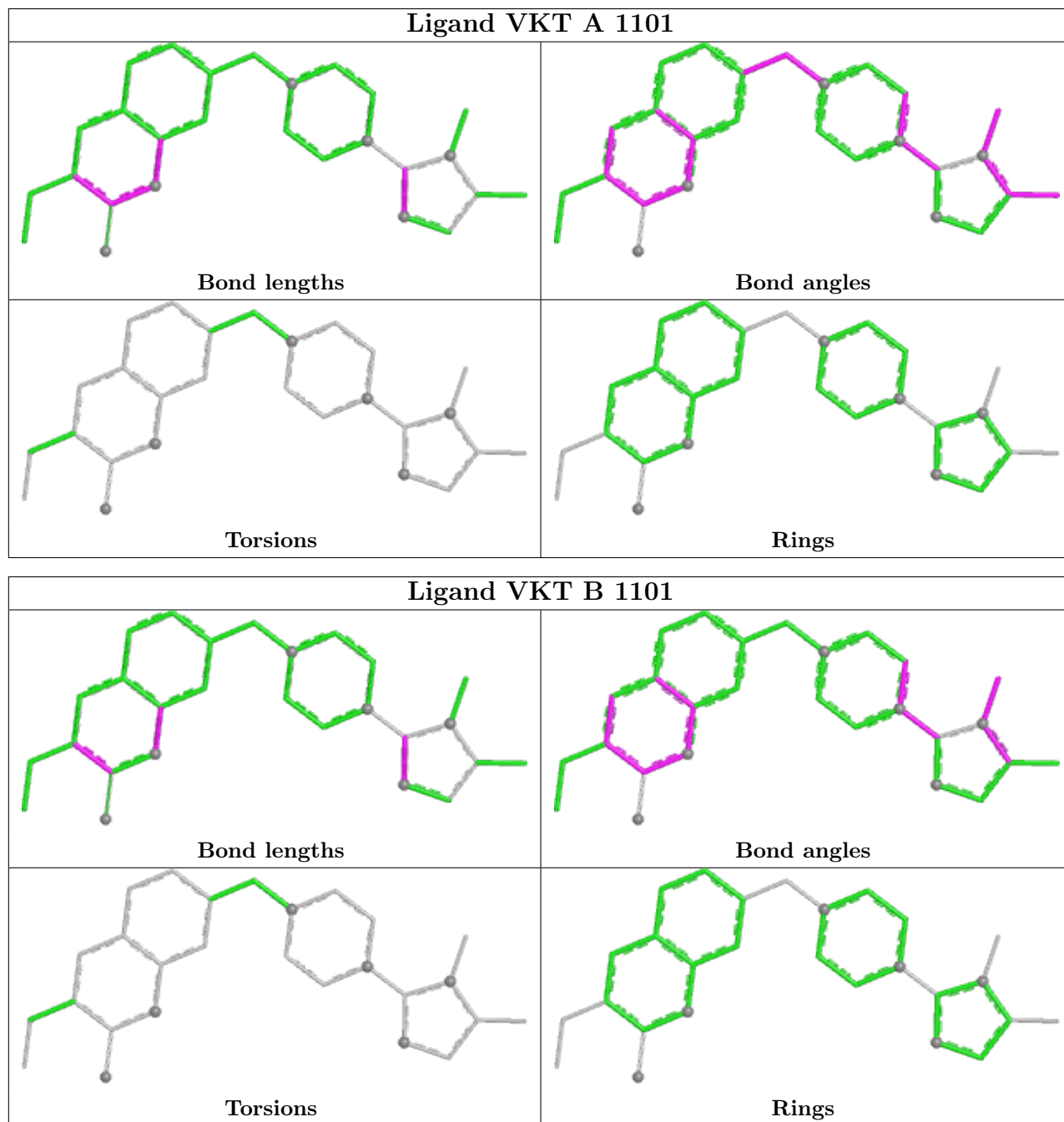| Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	1101	VKT	C2-C10-N	-6.27	116.06	125.25
2	A	1101	VKT	C2-C10-N	-6.26	116.08	125.25
2	B	1101	VKT	C10-N-C9	5.95	124.99	116.83
2	A	1101	VKT	C10-N-C9	5.84	124.85	116.83
2	A	1101	VKT	C3-C2-C10	4.65	120.62	115.46
2	B	1101	VKT	C3-C2-C10	4.57	120.54	115.46
2	A	1101	VKT	C19-N4-C18	3.35	127.38	124.09
2	B	1101	VKT	C19-N4-C18	3.21	127.24	124.09
2	A	1101	VKT	C13-N2-C16	-2.82	115.99	122.03
2	B	1101	VKT	C4-C9-N	-2.70	118.71	122.41
2	A	1101	VKT	C4-C9-N	-2.66	118.76	122.41
2	A	1101	VKT	C20-C18-N4	2.25	125.32	122.44
2	A	1101	VKT	C7-C11-N1	2.23	117.46	113.12
2	B	1101	VKT	C13-N2-C16	-2.08	117.58	122.03

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will

also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand VKT A 1101

Bond lengths

Bond angles

Torsions

Rings

Ligand VKT B 1101

Bond lengths

Bond angles

Torsions

Rings

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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, 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	350/352 (99%)	0.08	14 (4%)	38	40	8, 20, 51, 72	0
1	B	348/352 (98%)	0.52	39 (11%)	5	6	8, 32, 73, 109	0
All	All	698/704 (99%)	0.29	53 (7%)	13	15	8, 25, 63, 109	0

All (53) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1011	THR	6.5
1	B	1011	THR	6.4
1	B	826	HIS	4.7
1	B	743	ASP	4.4
1	B	744	PHE	4.4
1	B	784	ASP	4.4
1	A	964	LEU	4.4
1	A	725	SER	4.1
1	B	825	THR	4.0
1	B	746	MET	3.9
1	A	726	ASP	3.8
1	A	965	ASP	3.8
1	B	980	ASN	3.8
1	B	693	LEU	3.8
1	A	807	ASP	3.7
1	B	783	ASP	3.6
1	A	744	PHE	3.6
1	B	827	ASN	3.6
1	B	823	ALA	3.6
1	B	829	TYR	3.4
1	B	742	HIS	3.3
1	B	854	LEU	3.2
1	B	700	LYS	3.2
1	B	786	SER	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	747	LYS	3.0
1	B	913	GLY	2.9
1	B	853	GLN	2.8
1	A	746	MET	2.8
1	B	1009	PHE	2.8
1	B	679	VAL	2.8
1	A	745	GLY	2.7
1	B	981	ASP	2.7
1	B	832	GLU	2.7
1	A	748	LYS	2.5
1	B	965	ASP	2.4
1	B	851	PHE	2.4
1	B	748	LYS	2.4
1	B	752	LEU	2.3
1	A	747	LYS	2.3
1	B	912	GLN	2.3
1	A	720	VAL	2.3
1	B	1010	LYS	2.3
1	B	745	GLY	2.3
1	B	698	LEU	2.3
1	B	852	LYS	2.3
1	B	914	ASP	2.2
1	A	729	ILE	2.2
1	B	850	PRO	2.2
1	B	824	THR	2.2
1	B	756	ASP	2.2
1	B	828	ALA	2.1
1	B	749	PRO	2.1
1	A	963	SER	2.0

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

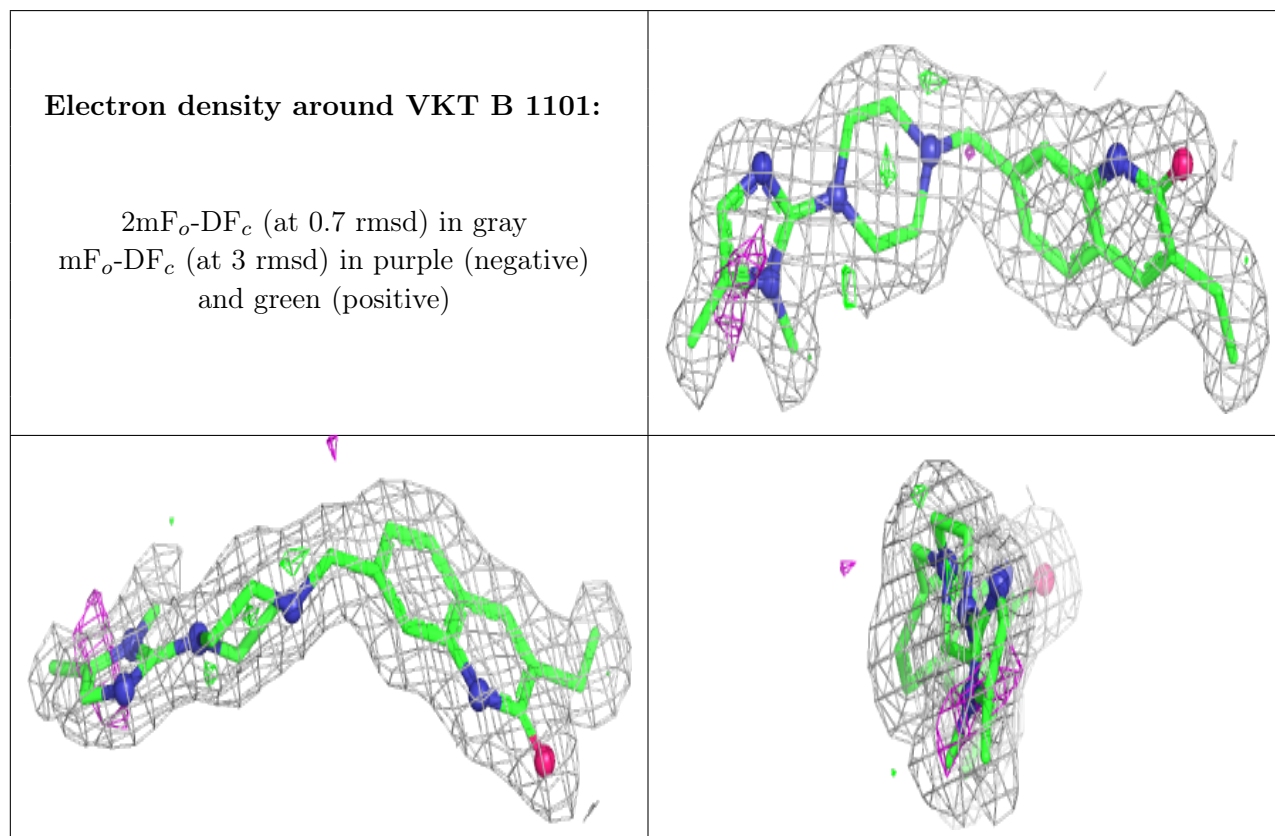There are no non-standard protein/DNA/RNA residues in this entry.

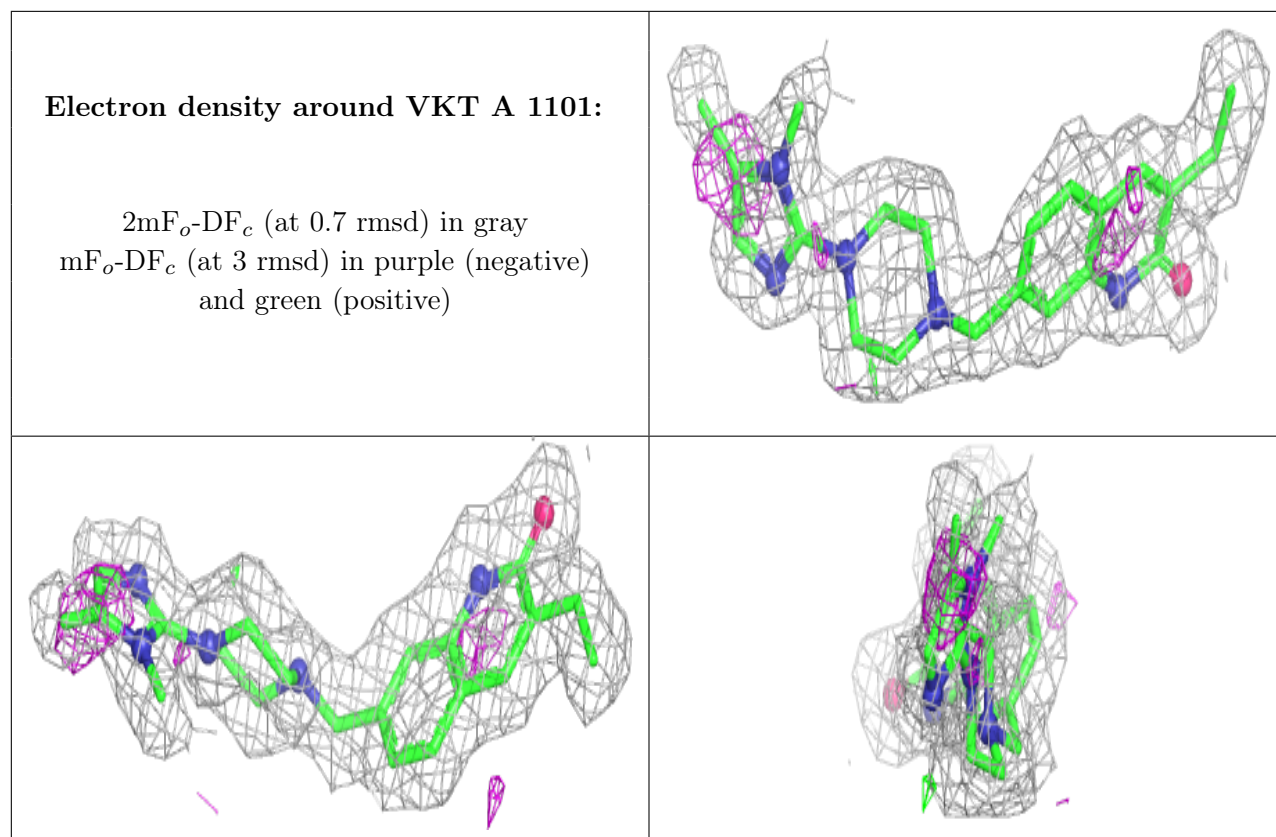## 6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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	SO4	A	1103	5/5	0.85	0.16	58,59,60,61	0
2	VKT	B	1101	27/27	0.94	0.11	10,14,21,22	0
2	VKT	A	1101	27/27	0.94	0.12	3,10,24,25	0
3	SO4	B	1102	5/5	0.95	0.12	26,28,35,37	0
3	SO4	A	1104	5/5	0.96	0.12	44,44,47,48	0
3	SO4	B	1103	5/5	0.96	0.16	48,48,49,49	0
3	SO4	A	1102	5/5	0.99	0.09	10,20,23,23	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around VKT B 1101:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

Electron density around VKT A 1101:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

6.5 Other polymers ⓘ

There are no such residues in this entry.