



Full wwPDB X-ray Structure Validation Report ⓘ

May 15, 2020 – 02:28 am BST

PDB ID : 5QSM
Title : PanDDA analysis group deposition – Crystal Structure of human STAG1 in complex with Z57261895
Authors : Newman, J.A.; Katis, V.L.; Gavard, A.E.; von Delft, F.; Arrowsmith, C.H.; Edwards, A.; Bountra, C.; Gileadi, O.
Deposited on : 2019-05-25
Resolution : 2.74 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 ⓘ](#)) 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
buster-report : 1.1.7 (2018)
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

2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 7047 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.

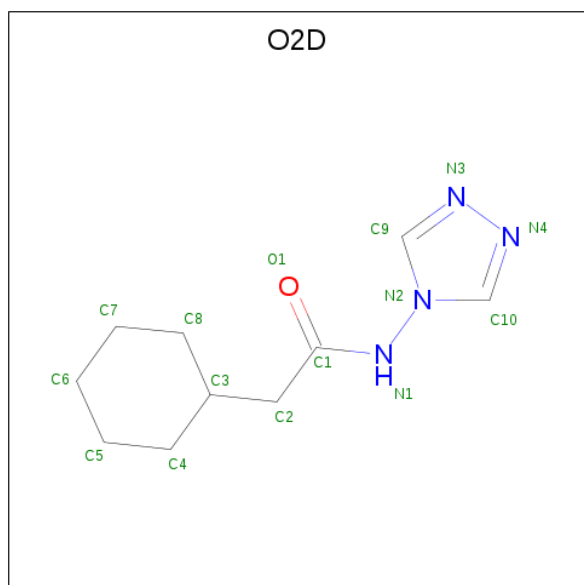
- Molecule 1 is a protein called Cohesin subunit SA-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	435	Total 3483	C 2220	N 578	O 663	S 22	0	0	0
1	B	425	Total 3409	C 2176	N 563	O 648	S 22	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	457	SER	-	expression tag	UNP Q8WVM7
A	458	MET	-	expression tag	UNP Q8WVM7
B	457	SER	-	expression tag	UNP Q8WVM7
B	458	MET	-	expression tag	UNP Q8WVM7

- Molecule 2 is 2-cyclohexyl-N-(4H-1,2,4-triazol-4-yl)acetamide (three-letter code: O2D) (formula: C₁₀H₁₆N₄O) (labeled as "Ligand of Interest" by author).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	15	10	4	1	0	0

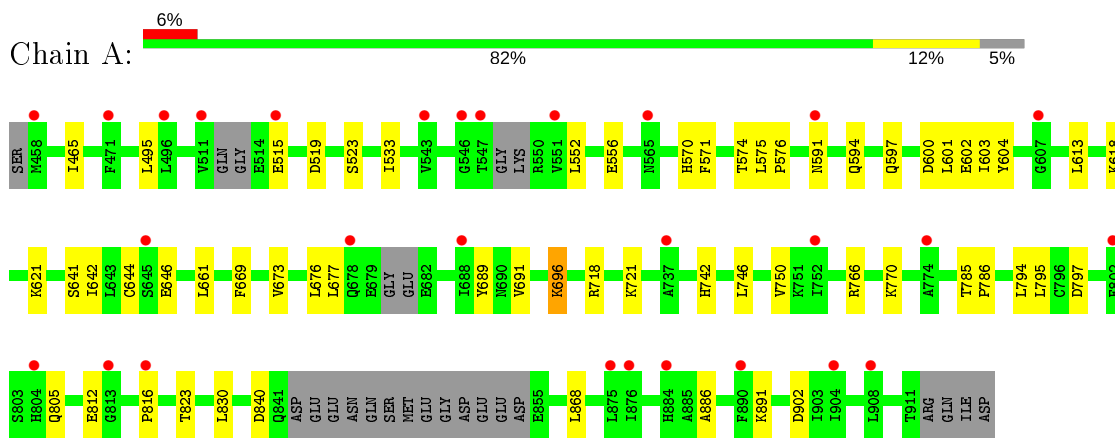
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	102	102	102	0	0
3	B	38	38	38	0	0

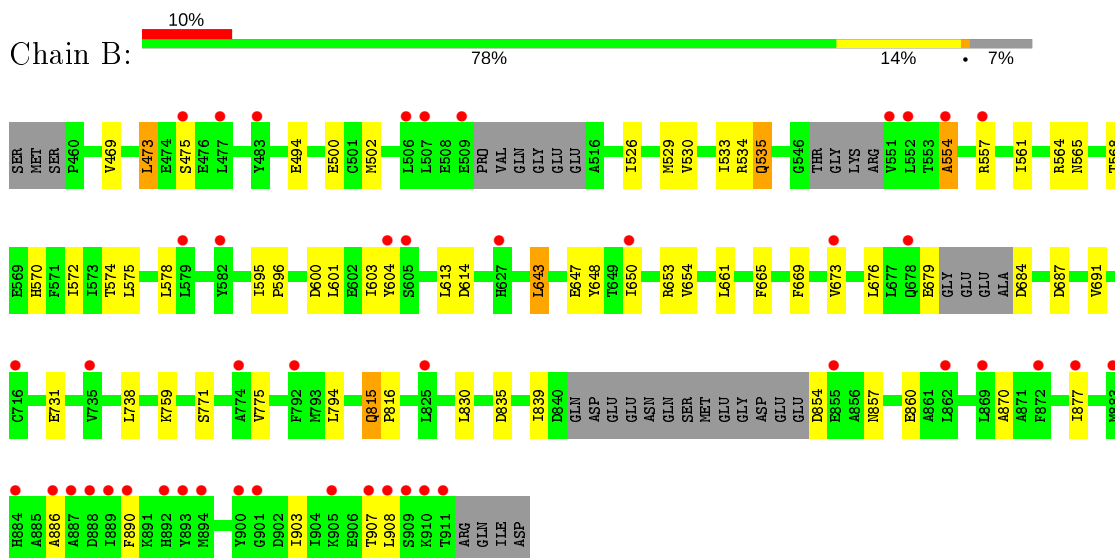
3 Residue-property plots [i](#)

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: Cohesin subunit SA-1



- Molecule 1: Cohesin subunit SA-1



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	155.14Å 168.17Å 46.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	114.03 – 2.74 114.03 – 2.74	Depositor EDS
% Data completeness (in resolution range)	99.5 (114.03-2.74) 99.5 (114.03-2.74)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.69 (at 2.73Å)	Xtrriage
Refinement program	REFMAC 5.8.0189	Depositor
R, R_{free}	0.260 , 0.299 0.290 , 0.322	Depositor DCC
R_{free} test set	1675 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å ²)	55.1	Xtrriage
Anisotropy	0.059	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 64.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	7047	wwPDB-VP
Average B, all atoms (Å ²)	94.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.91% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: O2D

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	0.63	0/3544	0.79	0/4791
1	B	0.62	0/3469	0.74	0/4687
All	All	0.62	0/7013	0.77	0/9478

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 [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	3483	0	3458	27	0
1	B	3409	0	3386	36	0
2	A	15	0	0	5	0
3	A	102	0	0	0	0
3	B	38	0	0	0	0
All	All	7047	0	6844	63	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 5.

All (63) 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:689:TYR:CD1	2:A:1001:O2D:C2	2.37	1.08
1:B:604:TYR:CE1	1:B:613:LEU:HB2	1.93	1.01
1:A:689:TYR:HD1	2:A:1001:O2D:C2	1.72	1.00
1:A:696:LYS:HE2	2:A:1001:O2D:C6	1.92	0.98
1:B:604:TYR:CD1	1:B:613:LEU:HB2	2.06	0.90
1:B:614:ASP:OD1	1:B:653:ARG:HD3	1.81	0.80
1:B:469:VAL:O	1:B:473:LEU:HD23	1.89	0.72
1:B:604:TYR:HB3	1:B:650:ILE:HG21	1.69	0.71
1:A:689:TYR:CE1	2:A:1001:O2D:C2	2.74	0.71
1:A:677:LEU:HD12	1:A:718:ARG:NH2	2.06	0.70
1:A:696:LYS:CE	2:A:1001:O2D:C6	2.74	0.65
1:A:766:ARG:NH1	1:A:816:PRO:O	2.30	0.65
1:A:676:LEU:HD22	1:A:691:VAL:HG21	1.79	0.63
1:B:604:TYR:CE1	1:B:613:LEU:CB	2.78	0.58
1:B:601:LEU:HB2	1:B:648:TYR:CE2	2.39	0.58
1:B:564:ARG:HD3	1:B:600:ASP:HB2	1.87	0.56
1:B:604:TYR:CB	1:B:650:ILE:HG21	2.36	0.55
1:B:877:ILE:HD11	1:B:907:THR:HA	1.89	0.53
1:A:552:LEU:HD22	1:A:556:GLU:HB3	1.89	0.53
1:B:643:LEU:HB3	1:B:654:VAL:HG11	1.91	0.53
1:A:742:HIS:CG	1:A:794:LEU:HD22	2.45	0.51
1:A:600:ASP:OD1	1:A:602:GLU:HG2	2.11	0.51
1:B:568:THR:HG22	1:B:572:ILE:CD1	2.40	0.51
1:B:561:ILE:O	1:B:565:ASN:ND2	2.44	0.50
1:B:502:MET:SD	1:B:529:MET:HG3	2.52	0.50
1:A:604:TYR:CD1	1:A:613:LEU:HB2	2.47	0.49
1:A:677:LEU:CD1	1:A:718:ARG:NH2	2.75	0.49
1:B:568:THR:HG22	1:B:572:ILE:HD11	1.95	0.49
1:B:676:LEU:HD22	1:B:691:VAL:HG21	1.95	0.49
1:B:771:SER:O	1:B:775:VAL:HG23	2.13	0.49
1:B:533:ILE:O	1:B:534:ARG:C	2.51	0.49
1:B:469:VAL:O	1:B:473:LEU:CD2	2.58	0.48
1:B:554:ALA:HA	1:B:557:ARG:HB3	1.93	0.48
1:B:570:HIS:CE1	1:B:574:THR:HG21	2.48	0.48
1:B:830:LEU:CD1	1:B:886:ALA:HA	2.43	0.48
1:B:857:ASN:O	1:B:860:GLU:HB3	2.14	0.48
1:A:533:ILE:HD11	1:A:571:PHE:CE2	2.49	0.48
1:A:746:LEU:O	1:A:750:VAL:HG23	2.14	0.48
1:B:684:ASP:HB3	1:B:687:ASP:OD2	2.14	0.47
1:B:661:LEU:HD11	1:B:665:PHE:CZ	2.50	0.47
1:A:902:ASP:N	1:A:902:ASP:OD1	2.47	0.47
1:A:677:LEU:HD12	1:A:718:ARG:CZ	2.45	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:465:ILE:HG13	1:A:495:LEU:HD11	1.98	0.46
1:A:570:HIS:CE1	1:A:574:THR:HG21	2.51	0.46
1:B:526:ILE:O	1:B:530:VAL:HG23	2.16	0.46
1:B:534:ARG:O	1:B:535:GLN:C	2.55	0.45
1:B:815:GLN:N	1:B:816:PRO:HD2	2.32	0.45
1:A:830:LEU:HD11	1:A:886:ALA:HA	1.99	0.45
1:B:603:ILE:HG22	1:B:603:ILE:O	2.17	0.45
1:B:575:LEU:HA	1:B:578:LEU:HD12	1.99	0.44
1:B:870:ALA:HB2	1:B:903:ILE:HD12	2.00	0.44
1:A:600:ASP:HB3	1:A:603:ILE:HD12	2.00	0.43
1:A:594:GLN:O	1:A:597:GLN:HG2	2.18	0.42
1:A:785:THR:N	1:A:786:PRO:CD	2.82	0.42
1:A:621:LYS:HG3	1:A:661:LEU:HD13	2.01	0.42
1:A:669:PHE:O	1:A:673:VAL:HG23	2.20	0.42
1:A:601:LEU:HD12	1:A:646:GLU:HB2	2.02	0.41
1:B:669:PHE:O	1:B:673:VAL:HG23	2.19	0.41
1:A:575:LEU:N	1:A:576:PRO:CD	2.84	0.41
1:B:738:LEU:HB3	1:B:794:LEU:CD1	2.51	0.40
1:B:890:PHE:CD1	1:B:908:LEU:HD21	2.56	0.40
1:B:643:LEU:HA	1:B:643:LEU:HD12	1.80	0.40
1:B:595:ILE:HB	1:B:596:PRO:HD3	2.02	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	425/459 (93%)	406 (96%)	17 (4%)	2 (0%)	29	48
1	B	415/459 (90%)	385 (93%)	29 (7%)	1 (0%)	47	69
All	All	840/918 (92%)	791 (94%)	46 (6%)	3 (0%)	34	55

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	515	GLU
1	B	554	ALA
1	A	644	CYS

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	389/415 (94%)	372 (96%)	17 (4%)	28	47
1	B	381/415 (92%)	367 (96%)	14 (4%)	34	54
All	All	770/830 (93%)	739 (96%)	31 (4%)	31	52

All (31) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	519	ASP
1	A	523	SER
1	A	591	ASN
1	A	618	LYS
1	A	641	SER
1	A	642	ILE
1	A	696	LYS
1	A	721	LYS
1	A	770	LYS
1	A	795	LEU
1	A	797	ASP
1	A	805	GLN
1	A	812	GLU
1	A	823	THR
1	A	840	ASP
1	A	868	LEU
1	A	891	LYS
1	B	473	LEU
1	B	475	SER
1	B	494	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	500	GLU
1	B	535	GLN
1	B	643	LEU
1	B	647	GLU
1	B	679	GLU
1	B	731	GLU
1	B	759	LYS
1	B	815	GLN
1	B	835	ASP
1	B	839	ILE
1	B	854	ASP

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

Mol	Chain	Res	Type
1	A	480	HIS
1	A	565	ASN
1	A	703	ASN
1	A	790	GLN
1	B	565	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 [i](#)

There are no carbohydrates 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
2	O2D	A	1001	-	10,16,16	1.37	1 (10%)	10,20,20	1.92	3 (30%)

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	O2D	A	1001	-	-	0/6/16/16	0/2/2/2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1001	O2D	O1-C1	-2.40	1.18	1.23

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1001	O2D	C3-C2-C1	-3.79	104.26	112.87
2	A	1001	O2D	C8-C3-C2	-2.79	108.28	111.52
2	A	1001	O2D	C5-C4-C3	-2.72	107.01	112.15

There are no chirality outliers.

There are no torsion outliers.

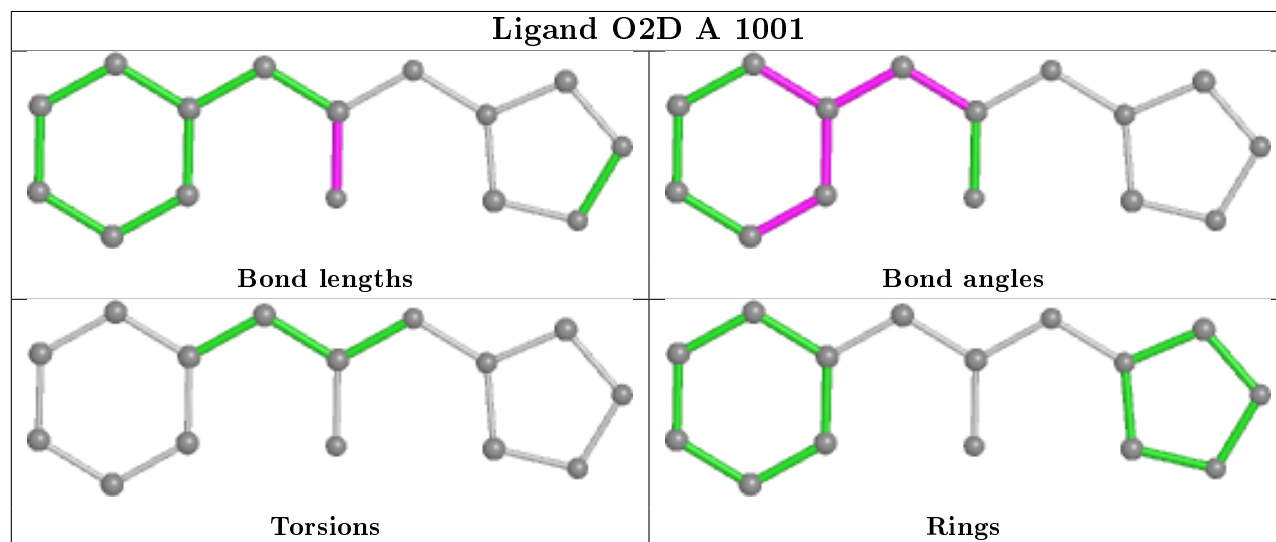
There are no ring outliers.

1 monomer is involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1001	O2D	5	0

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 than 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.



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, 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	435/459 (94%)	0.85	28 (6%) 19 22	56, 82, 122, 143	0
1	B	425/459 (92%)	1.05	46 (10%) 5 5	62, 101, 143, 173	0
All	All	860/918 (93%)	0.94	74 (8%) 10 12	56, 92, 137, 173	0

All (74) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	893	TYR	7.1
1	B	884	HIS	6.8
1	A	547	THR	5.9
1	B	892	HIS	5.4
1	B	889	ILE	5.3
1	B	908	LEU	5.1
1	A	813	GLY	4.7
1	B	872	PHE	4.7
1	B	477	LEU	4.5
1	B	678	GLN	4.5
1	B	869	LEU	4.3
1	A	904	ILE	3.6
1	B	907	THR	3.5
1	B	552	LEU	3.4
1	A	908	LEU	3.3
1	B	551	VAL	3.1
1	B	909	SER	3.0
1	B	910	LYS	3.0
1	B	900	TYR	3.0
1	B	579	LEU	2.9
1	A	546	GLY	2.8
1	A	458	MET	2.8
1	B	483	TYR	2.7
1	B	627	HIS	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	886	ALA	2.7
1	B	894	MET	2.7
1	A	511	VAL	2.7
1	B	716	CYS	2.7
1	B	888	ASP	2.7
1	A	890	PHE	2.6
1	A	515	GLU	2.5
1	A	875	LEU	2.5
1	A	804	HIS	2.5
1	B	650	ILE	2.5
1	B	887	ALA	2.5
1	B	911	THR	2.5
1	B	901	GLY	2.4
1	B	475	SER	2.4
1	B	605	SER	2.4
1	A	802	PHE	2.3
1	B	509	GLU	2.3
1	A	737	ALA	2.3
1	B	506	LEU	2.3
1	A	876	ILE	2.3
1	B	877	ILE	2.3
1	B	792	PHE	2.2
1	A	678	GLN	2.2
1	B	582	TYR	2.2
1	A	645	SER	2.2
1	A	551	VAL	2.2
1	A	774	ALA	2.2
1	A	688	ILE	2.2
1	A	752	ILE	2.2
1	B	673	VAL	2.2
1	B	883	MET	2.1
1	B	890	PHE	2.1
1	A	543	VAL	2.1
1	B	507	LEU	2.1
1	A	607	GLY	2.1
1	B	905	LYS	2.1
1	B	862	LEU	2.1
1	A	816	PRO	2.1
1	B	825	LEU	2.1
1	B	774	ALA	2.1
1	A	591	ASN	2.1
1	A	496	LEU	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	735	VAL	2.1
1	B	855	GLU	2.1
1	A	565	ASN	2.0
1	B	557	ARG	2.0
1	B	604	TYR	2.0
1	B	554	ALA	2.0
1	A	471	PHE	2.0
1	A	884	HIS	2.0

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

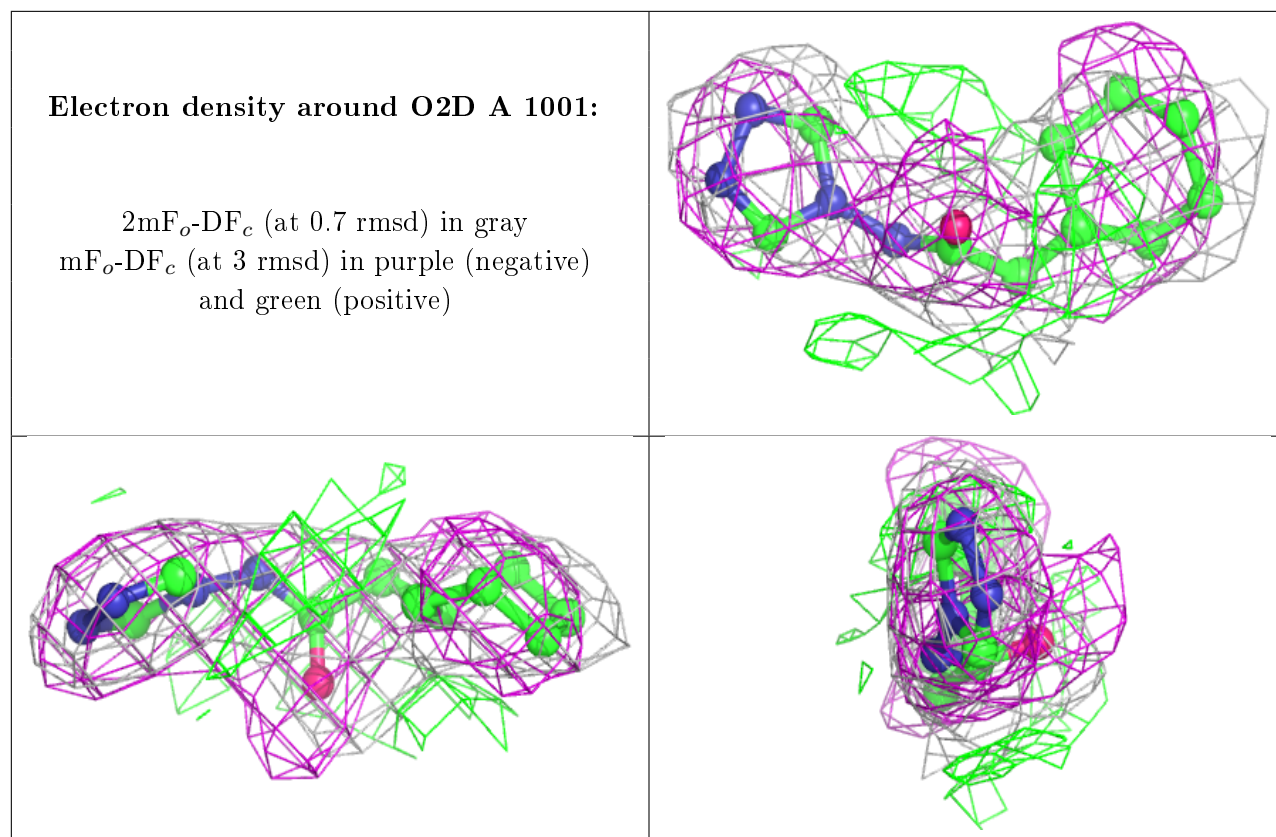
There are no carbohydrates in this entry.

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
2	O2D	A	1001	15/15	0.90	0.38	20,20,20,20	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.



6.5 Other polymers [i](#)

There are no such residues in this entry.