



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 02:56 PM UTC

PDB ID : 2H2J / pdb_00002h2j
Title : Structure of Rubisco LSMT bound to Sinefungin and Monomethyllysine
Authors : Couture, J.F.; Hauk, G.; Trievel, R.C.
Deposited on : 2006-05-18
Resolution : 2.45 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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Buster-report	:	wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

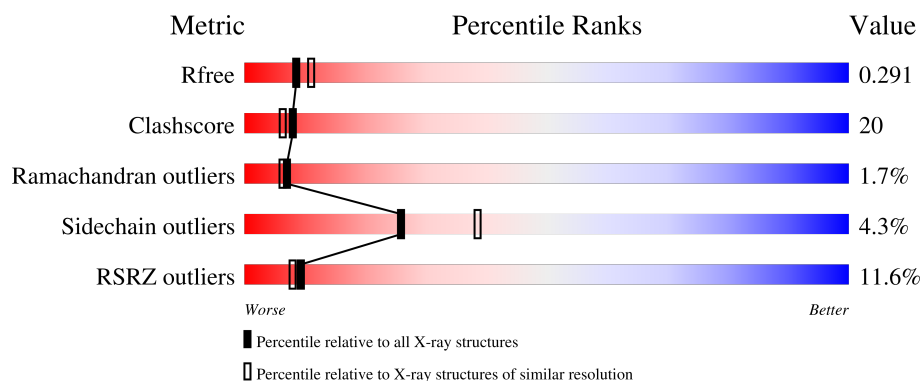
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.45 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}	180053	1190 (2.46-2.46)
Clashscore	190562	1229 (2.46-2.46)
Ramachandran outliers	187476	1218 (2.46-2.46)
Sidechain outliers	187428	1218 (2.46-2.46)
RSRZ outliers	180081	1190 (2.46-2.46)

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 $\leq 5\%$. 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	440	20% 62% 29% . .
1	B	440	9% 67% 30% .
1	C	440	6% 64% 30% 5%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 11028 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.

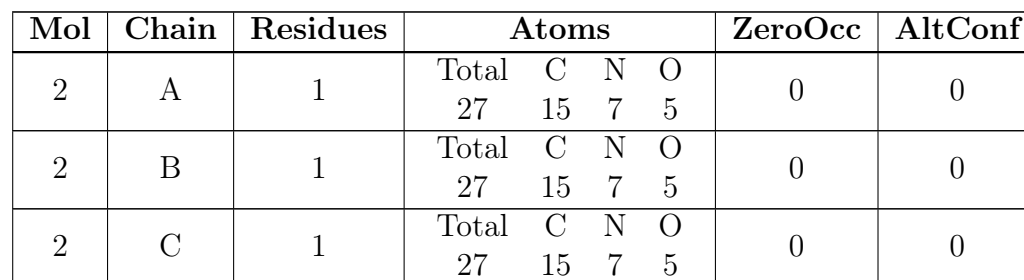
- Molecule 1 is a protein called Ribulose-1,5 bisphosphate carboxylase/oxygenase large subunit N-methyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	423	Total	C	N	O	S	0	0	0
			3406	2184	561	654	7			
1	B	439	Total	C	N	O	S	0	0	0
			3532	2265	583	677	7			
1	C	438	Total	C	N	O	S	0	0	0
			3526	2262	582	675	7			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	483	GLU	-	cloning artifact	UNP Q43088
A	484	ASN	-	cloning artifact	UNP Q43088
A	485	LEU	-	cloning artifact	UNP Q43088
A	486	TYR	-	cloning artifact	UNP Q43088
A	487	PHE	-	cloning artifact	UNP Q43088
A	488	GLN	-	cloning artifact	UNP Q43088
B	483	GLU	-	cloning artifact	UNP Q43088
B	484	ASN	-	cloning artifact	UNP Q43088
B	485	LEU	-	cloning artifact	UNP Q43088
B	486	TYR	-	cloning artifact	UNP Q43088
B	487	PHE	-	cloning artifact	UNP Q43088
B	488	GLN	-	cloning artifact	UNP Q43088
C	483	GLU	-	cloning artifact	UNP Q43088
C	484	ASN	-	cloning artifact	UNP Q43088
C	485	LEU	-	cloning artifact	UNP Q43088
C	486	TYR	-	cloning artifact	UNP Q43088
C	487	PHE	-	cloning artifact	UNP Q43088
C	488	GLN	-	cloning artifact	UNP Q43088

- Molecule 2 is SINEFUNGIN (CCD ID: SFG) (formula: $C_{15}H_{23}N_7O_5$).



- MLZ
-
- The chemical structure of MLZ (Methyl 2-amino-3-oxopentanoate) is shown. It features a central carbon chain with a methyl group (CM) attached to a nitrogen atom (N) labeled NZ. The chain continues through a carbon atom (CE) to another carbon atom (CD), then to a carbon atom (CG), and finally to a carbon atom (CB). The carbon atom (CB) is bonded to a nitrogen atom (N) labeled H₂N. The carbon atom (CB) is also bonded to a carbon atom (CA(S)) which is double-bonded to an oxygen atom (O). The carbon atom (CA(S)) is bonded to a carbon atom (C) which is double-bonded to an oxygen atom (O) and single-bonded to a hydroxyl group (OH) labeled OXT. The carbon atom (C) is also bonded to a methyl group (CM).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			11	7	2	2		

Continued on next page...

Continued from previous page...

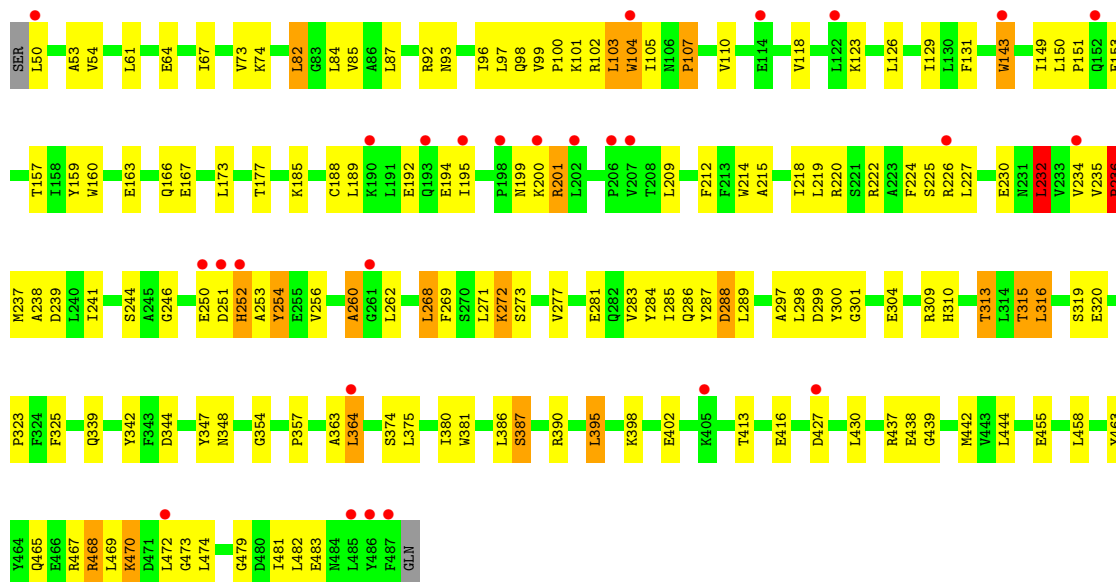
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	B	1	Total	C	N	O	0	0
			11	7	2	2		
3	C	1	Total	C	N	O	0	0
			11	7	2	2		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	138	Total	O	0	0
			138	138		
4	B	138	Total	O	0	0
			138	138		
4	C	174	Total	O	0	0
			174	174		



- Molecule 1: Ribulose-1,5 biphosphate carboxylase/oxygenase large subunit N-methyltransferase



4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α , β , γ	132.37Å 156.50Å 267.59Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	27.79 – 2.45 27.79 – 2.45	Depositor EDS
% Data completeness (in resolution range)	94.8 (27.79-2.45) 94.8 (27.79-2.45)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.05	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.03 (at 2.44Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.253 , 0.292 0.253 , 0.291	Depositor DCC
R_{free} test set	4883 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å ²)	58.6	Xtriage
Anisotropy	0.322	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 49.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	11028	wwPDB-VP
Average B, all atoms (Å ²)	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.89% 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: SFG, MLZ

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	0.46	0/3475	0.87	8/4713 (0.2%)
1	B	0.48	0/3607	0.92	14/4894 (0.3%)
1	C	0.47	0/3601	0.92	12/4886 (0.2%)
All	All	0.47	0/10683	0.91	34/14493 (0.2%)

There are no bond length outliers.

The worst 5 of 34 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	235	VAL	CA-C-N	7.86	129.67	119.84
1	C	235	VAL	C-N-CA	7.86	129.67	119.84
1	C	253	ALA	N-CA-C	-6.76	105.02	113.20
1	B	268	LEU	N-CA-C	6.73	119.90	109.07
1	B	235	VAL	CA-C-N	6.53	128.00	119.84

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	3406	0	3375	138	0
1	B	3532	0	3493	129	0
1	C	3526	0	3488	157	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	27	0	22	1	0
2	B	27	0	22	4	0
2	C	27	0	22	4	0
3	A	11	0	15	0	0
3	B	11	0	15	2	0
3	C	11	0	15	4	0
4	A	138	0	0	9	0
4	B	138	0	0	13	0
4	C	174	0	0	23	0
All	All	11028	0	10467	413	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 20.

The worst 5 of 413 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:B:468:ARG:HB3	1:B:468:ARG:HH11	1.19	1.06
1:C:177:THR:HB	4:C:940:HOH:O	1.65	0.94
1:A:439:GLY:HA2	1:A:442:MET:HE3	1.51	0.92
1:C:315:THR:HG22	4:C:904:HOH:O	1.68	0.92
1:B:286:GLN:HE21	1:B:288:ASP:H	0.96	0.91

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	417/440 (95%)	377 (90%)	33 (8%)	7 (2%)	7 6
1	B	437/440 (99%)	404 (92%)	26 (6%)	7 (2%)	7 7

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	436/440 (99%)	405 (93%)	23 (5%)	8 (2%)	6	5
All	All	1290/1320 (98%)	1186 (92%)	82 (6%)	22 (2%)	7	6

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	251	ASP
1	A	364	LEU
1	A	387	SER
1	B	364	LEU
1	B	387	SER

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	373/386 (97%)	356 (95%)	17 (5%)	24	36
1	B	385/386 (100%)	371 (96%)	14 (4%)	31	45
1	C	384/386 (100%)	366 (95%)	18 (5%)	23	35
All	All	1142/1158 (99%)	1093 (96%)	49 (4%)	26	38

5 of 49 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	427	ASP
1	C	232	LEU
1	B	444	LEU
1	C	107	PRO
1	C	268	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	310	HIS
1	B	465	GLN
1	C	252	HIS
1	B	446	GLN
1	B	484	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

6 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	MLZ	A	900	-	9,10,10	0.75	0	8,11,11	0.41	0
3	MLZ	B	901	-	9,10,10	0.74	0	8,11,11	0.38	0
3	MLZ	C	902	-	9,10,10	0.75	0	8,11,11	0.37	0
2	SFG	B	801	-	28,29,29	1.28	2 (7%)	34,42,42	1.79	6 (17%)
2	SFG	A	800	-	28,29,29	1.09	2 (7%)	34,42,42	1.18	3 (8%)
2	SFG	C	802	-	28,29,29	1.01	2 (7%)	34,42,42	1.42	5 (14%)

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	MLZ	A	900	-	-	2/10/10/10	-
3	MLZ	B	901	-	-	1/10/10/10	-
3	MLZ	C	902	-	-	1/10/10/10	-
2	SFG	B	801	-	-	7/17/33/33	0/3/3/3
2	SFG	A	800	-	-	5/17/33/33	0/3/3/3
2	SFG	C	802	-	-	6/17/33/33	0/3/3/3

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	801	SFG	CG-CD	3.26	1.57	1.53
2	B	801	SFG	C4-N3	2.98	1.40	1.34
2	C	802	SFG	CG-CD	2.53	1.56	1.53
2	C	802	SFG	C4-N3	2.26	1.38	1.34
2	A	800	SFG	C5'-CD	2.15	1.56	1.53

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	801	SFG	O4'-C4'-C5'	6.82	121.19	109.41
2	C	802	SFG	O4'-C4'-C5'	4.29	116.82	109.41
2	B	801	SFG	O4'-C1'-N9	3.57	114.94	108.09
2	B	801	SFG	C3'-C2'-C1'	3.47	108.03	101.46
2	C	802	SFG	C3'-C2'-C1'	3.17	107.45	101.46

There are no chirality outliers.

5 of 22 torsion outliers are listed below:

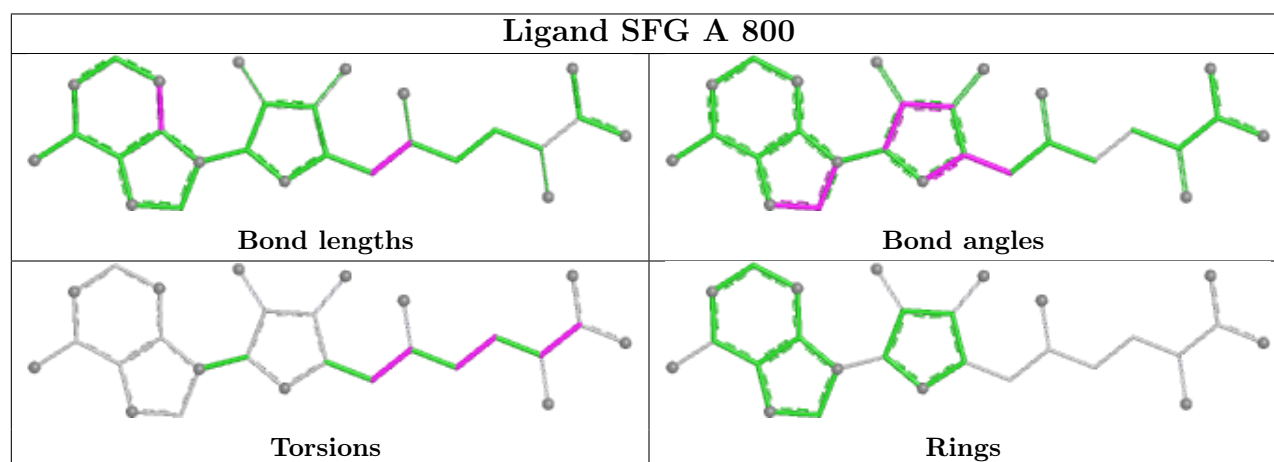
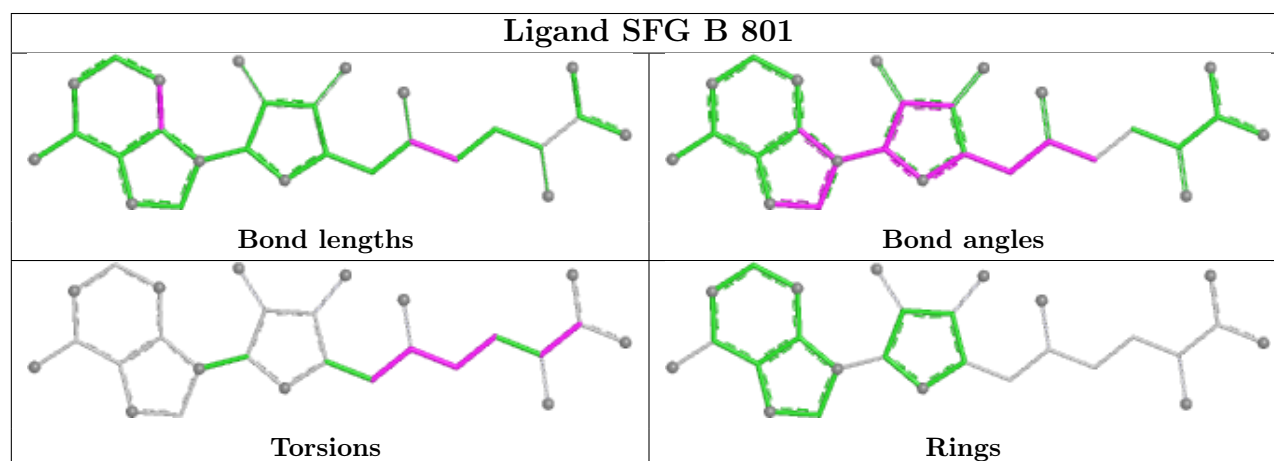
Mol	Chain	Res	Type	Atoms
2	A	800	SFG	CA-CB-CG-CD
2	A	800	SFG	C4'-C5'-CD-CG
2	A	800	SFG	C4'-C5'-CD-NE
2	B	801	SFG	CA-CB-CG-CD
2	B	801	SFG	C4'-C5'-CD-CG

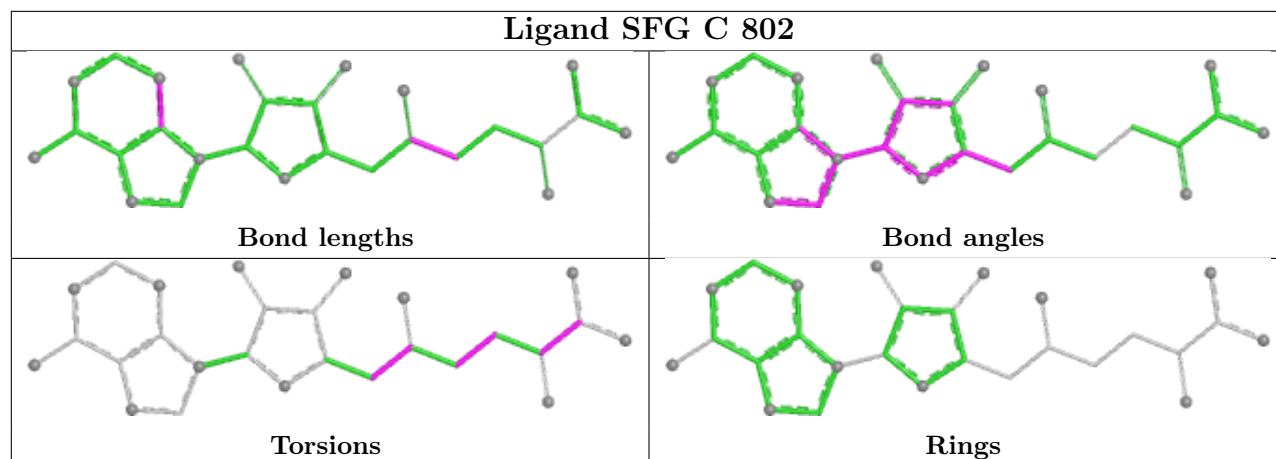
There are no ring outliers.

5 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	901	MLZ	2	0
3	C	902	MLZ	4	0
2	B	801	SFG	4	0
2	A	800	SFG	1	0
2	C	802	SFG	4	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 [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	423/440 (96%)	1.02	86 (20%) 3 2	48, 71, 110, 133	0
1	B	439/440 (99%)	0.53	38 (8%) 16 14	40, 65, 102, 137	0
1	C	438/440 (99%)	0.49	27 (6%) 26 24	43, 64, 98, 118	0
All	All	1300/1320 (98%)	0.68	151 (11%) 9 8	40, 66, 106, 137	0

The worst 5 of 151 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	481	ILE	6.0
1	B	365	GLY	5.9
1	B	482	LEU	5.7
1	C	487	PHE	5.7
1	A	252	HIS	5.3

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 oligosaccharides in this entry.

6.4 Ligands [i](#)

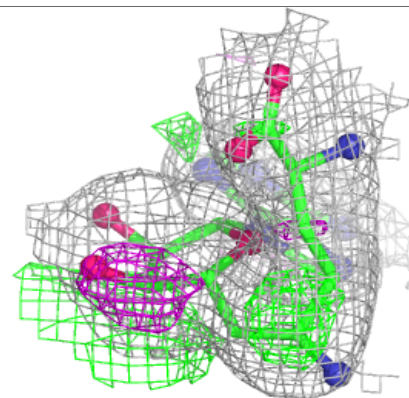
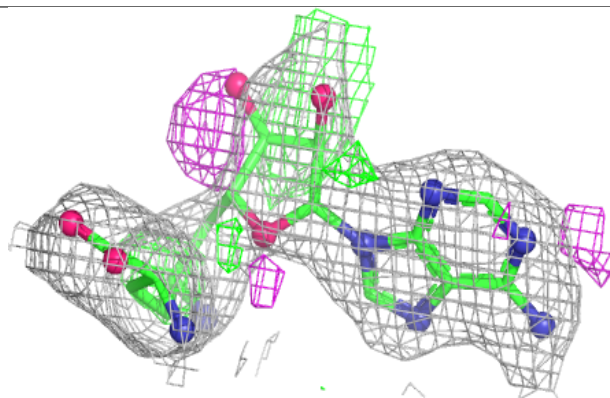
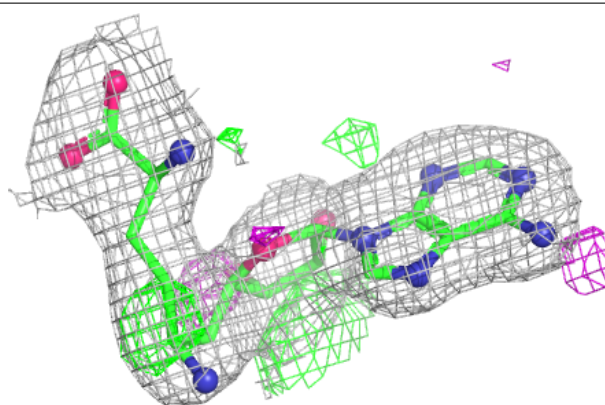
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q < 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
2	SFG	A	800	27/27	0.85	0.14	52,63,69,76	0
2	SFG	C	802	27/27	0.87	0.14	50,56,61,67	0
3	MLZ	A	900	11/11	0.88	0.21	82,87,94,95	0
3	MLZ	B	901	11/11	0.90	0.14	49,58,66,66	0
2	SFG	B	801	27/27	0.92	0.10	38,43,50,53	0
3	MLZ	C	902	11/11	0.93	0.11	50,53,56,57	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.

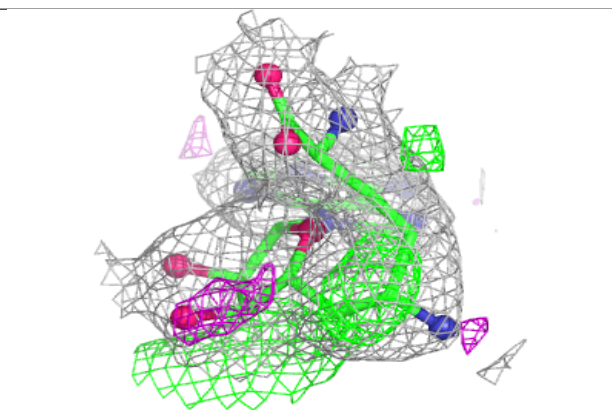
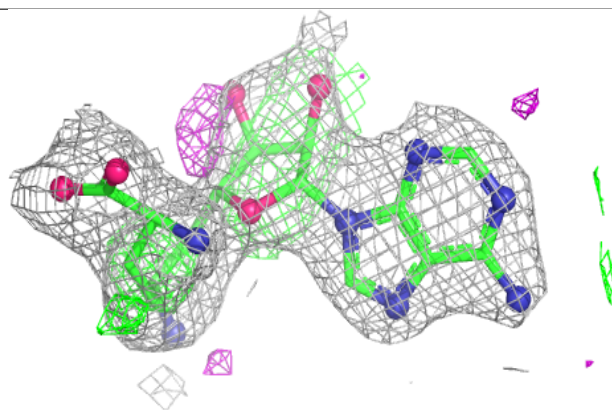
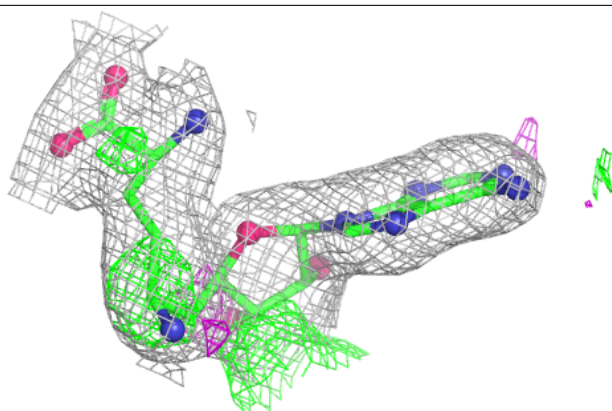
Electron density around SFG A 800:

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

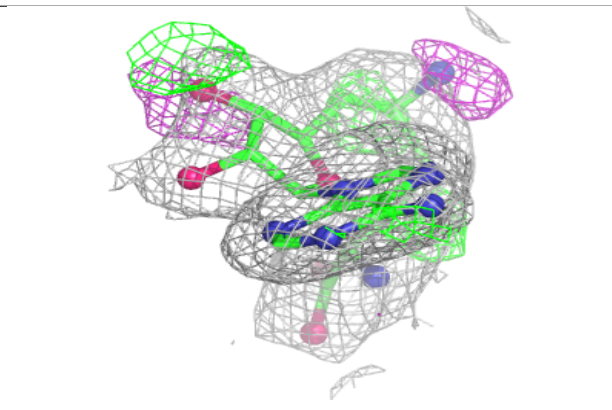
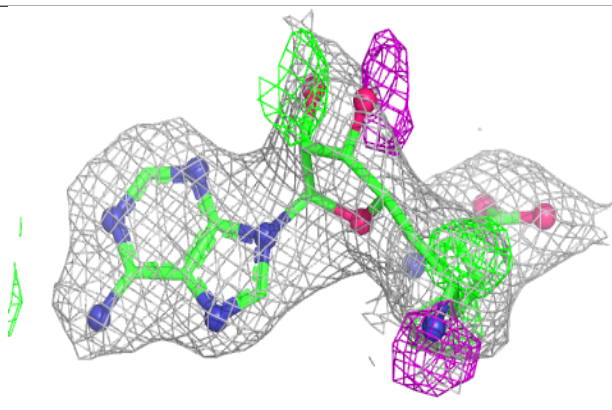
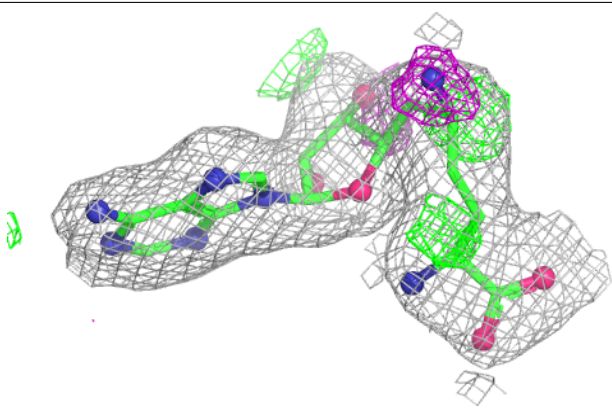


Electron density around SFG C 802:

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

**Electron density around SFG B 801:**

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



6.5 Other polymers [i](#)

There are no such residues in this entry.