



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 9, 2020 – 02:31 AM BST

PDB ID : 2FPC
Title : Structure of Strictosidine Synthase, the Biosynthetic Entry to the Monoterpenoid Indole Alkaloid Family
Authors : Panjikar, S.
Deposited on : 2006-01-16
Resolution : 3.00 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 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 : 2.13.1
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) : 2.13.1

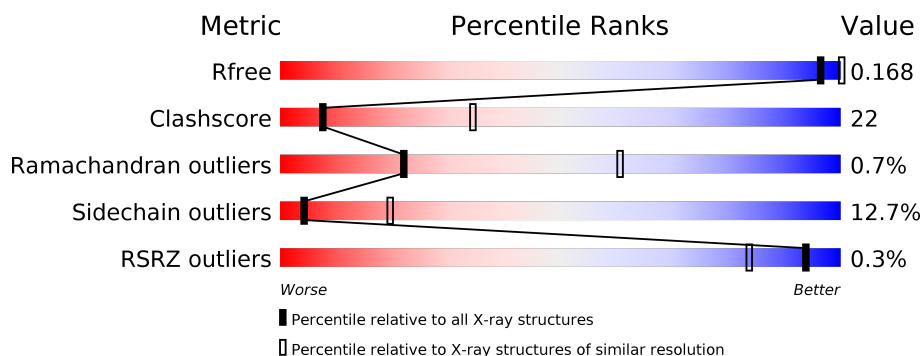
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 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	322	
1	B	322	

2 Entry composition [i](#)

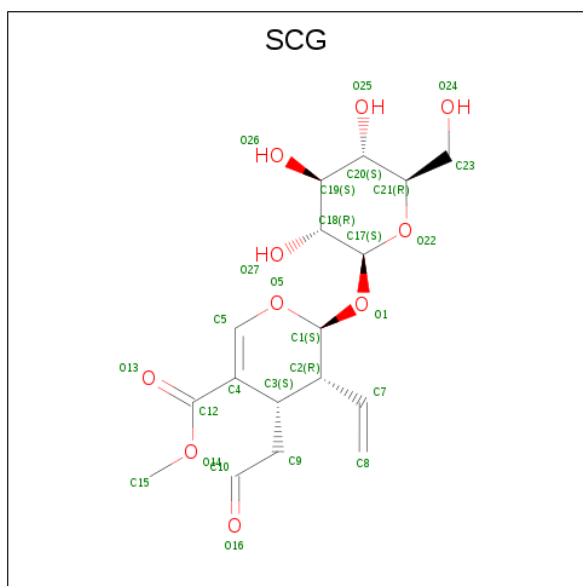
There are 3 unique types of molecules in this entry. The entry contains 4920 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Strictosidine synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	302	Total	C	N	O	S	0	0	0
			2382	1529	388	460	5			
1	B	302	Total	C	N	O	S	0	0	0
			2382	1529	388	460	5			

- Molecule 2 is Secologanin (three-letter code: SCG) (formula: $C_{17}H_{24}O_{10}$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			27	17	10		
2	B	1	Total	C	O	0	0
			27	17	10		


- Molecule 3 is water.

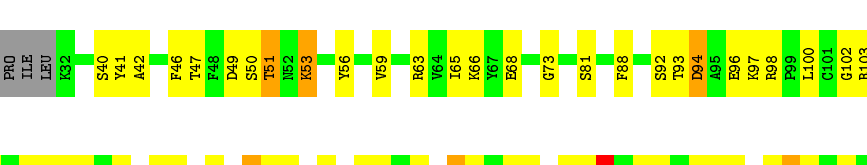
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	55	Total 55	O 55	0	0
3	B	47	Total 47	O 47	0	0

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.

- Chain A: 49% 38% 7% 6%

Category	Amino Acid	Percentage
SER	SER	49%
	LEU	38%
	ALA	7%
	LEU	6%
	SER	49%
	SER	38%
	PRO	7%
	ILE	6%
	LEU	38%
	K32	49%
E37	E37	49%
	A38	38%
	P39	7%
	S40	6%
	Y41	49%
	A42	38%
	S45	7%
	F46	6%
	T47	49%
	F48	38%
D49	D49	49%
	S50	38%
	T51	7%
	N52	6%
	K53	49%
	G54	38%
	F55	7%
	Y56	6%
	T57	49%
	S58	38%
V59	V59	49%
	Q60	38%
	D61	7%
	K66	6%
	Y67	49%
	E68	38%
	N71	7%
	S72	6%
	F73	49%
	F74	38%
S81	S81	49%
	P82	38%
	Y83	7%
	W84	6%
	N85	49%
	C89	38%
	S92	7%
	I93	6%
	D94	49%
	A95	38%
E96	E96	49%
	K97	38%
	R98	7%
	P99	6%
	R103	49%
	T104	38%
	Y105	7%
	D106	6%
	I107	49%
	S108	38%
Y109	Y109	49%
	Q112	38%
	N113	7%
	M114	6%
	P115	49%
	L116	38%
	D120	7%
	G121	6%
	Y122	49%
	I123	38%
H124	H124	49%
	L125	38%
	S126	7%
	E131	6%
	A135	49%
	L138	38%
	S141	7%
	P146	6%
	F147	49%
	K148	38%
W149	W149	49%
	L150	38%
	Y151	7%
	A152	6%
	V155	49%
	D156	38%
	Q157	7%
	R158	6%
	T165	49%
	D166	38%
L170	L170	49%
	Y171	38%
	D172	7%
	D173	6%
	R174	49%
	M180	38%
	K185	7%
	T186	6%
	D187	49%
	R188	38%
P189	P189	49%
	K191	38%
	Y192	7%
	D193	6%
	P194	49%
	K197	38%
	T200	7%
	L201	6%
	K204	49%
	E205	38%
L206	L206	49%
	H207	38%
	V208	7%
	P209	6%
	G210	49%
	G211	38%
	A212	7%
	G213	6%
	E214	49%
	V214	38%
D217	D217	49%
	G218	38%
	V221	7%
	L222	6%
	V223	49%
	A224	38%
	E225	7%
	F226	6%
	L227	49%
	S228	38%
H229	H229	49%
	Q230	38%
	Y234	7%
	W235	6%
	L236	49%
	K240	38%
	K241	7%
	G242	6%

- Chain B: 



4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α , β , γ	148.61Å 148.61Å 121.07Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 3.00 19.97 – 3.00	Depositor EDS
% Data completeness (in resolution range)	100.0 (20.00-3.00) 100.0 (19.97-3.00)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.26 (at 2.98Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.186 , 0.240 0.170 , 0.168	Depositor DCC
R_{free} test set	773 reflections (3.88%)	wwPDB-VP
Wilson B-factor (Å ²)	55.7	Xtriage
Anisotropy	0.734	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 61.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.024 for h,-h-k,-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	4920	wwPDB-VP
Average B, all atoms (Å ²)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.15% 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: SCG

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	0.95	1/2447 (0.0%)	1.01	9/3329 (0.3%)
1	B	0.83	1/2447 (0.0%)	0.97	8/3329 (0.2%)
All	All	0.90	2/4894 (0.0%)	0.99	17/6658 (0.3%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	180	MET	SD-CE	7.74	2.21	1.77
1	B	59	VAL	CB-CG1	-5.01	1.42	1.52

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	63	ARG	NE-CZ-NH1	-6.69	116.95	120.30
1	A	120	ASP	CB-CG-OD2	6.64	124.27	118.30
1	A	106	ASP	CB-CG-OD2	6.42	124.08	118.30
1	B	94	ASP	CB-CG-OD2	6.34	124.01	118.30
1	A	288	ASP	CB-CG-OD2	6.14	123.83	118.30

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	2382	0	2287	115	0
1	B	2382	0	2287	91	0
2	A	27	0	15	1	0
2	B	27	0	15	4	0
3	A	55	0	0	4	0
3	B	47	0	0	4	0
All	All	4920	0	4604	207	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 22.

The worst 5 of 207 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:180:MET:SD	1:A:180:MET:CE	2.21	1.28
1:B:229:HIS:HD2	1:B:252:ASN:H	1.13	0.94
1:A:240:LYS:O	1:A:243:THR:HB	1.69	0.93
1:B:122:TYR:OH	1:B:173:ASP:HB2	1.69	0.91
1:A:288:ASP:HB3	1:A:294:LEU:HD11	1.51	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	300/322 (93%)	280 (93%)	18 (6%)	2 (1%)	22	60
1	B	300/322 (93%)	280 (93%)	18 (6%)	2 (1%)	22	60
All	All	600/644 (93%)	560 (93%)	36 (6%)	4 (1%)	22	60

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	173	ASP
1	A	210	GLY
1	A	173	ASP
1	B	210	GLY

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	260/278 (94%)	233 (90%)	27 (10%)	7	27
1	B	260/278 (94%)	221 (85%)	39 (15%)	3	14
All	All	520/556 (94%)	454 (87%)	66 (13%)	4	19

5 of 66 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	53	LYS
1	B	107	ILE
1	B	248	VAL
1	B	65	ILE
1	B	93	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	252	ASN
1	A	292	ASN
1	B	310	GLN
1	A	263	HIS
1	A	275	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 ⓘ

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$
2	SCG	A	1001	-	27,28,28	1.54	3 (11%)	34,39,39	1.68	7 (20%)
2	SCG	B	2001	-	27,28,28	2.07	3 (11%)	34,39,39	2.27	9 (26%)

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	SCG	A	1001	-	-	6/17/54/54	0/2/2/2
2	SCG	B	2001	-	-	4/17/54/54	0/2/2/2

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	2001	SCG	O14-C12	7.17	1.49	1.33
2	B	2001	SCG	O5-C5	6.17	1.46	1.35
2	A	1001	SCG	O5-C5	5.19	1.45	1.35
2	A	1001	SCG	O14-C12	4.10	1.42	1.33
2	A	1001	SCG	C2-C3	-2.89	1.51	1.54

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	2001	SCG	O5-C1-C2	-5.81	103.70	112.86
2	B	2001	SCG	C15-O14-C12	5.44	126.14	115.86
2	B	2001	SCG	O13-C12-C4	-5.04	118.02	124.34
2	B	2001	SCG	O14-C12-C4	4.65	118.69	112.64
2	A	1001	SCG	O5-C1-C2	-4.64	105.53	112.86

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

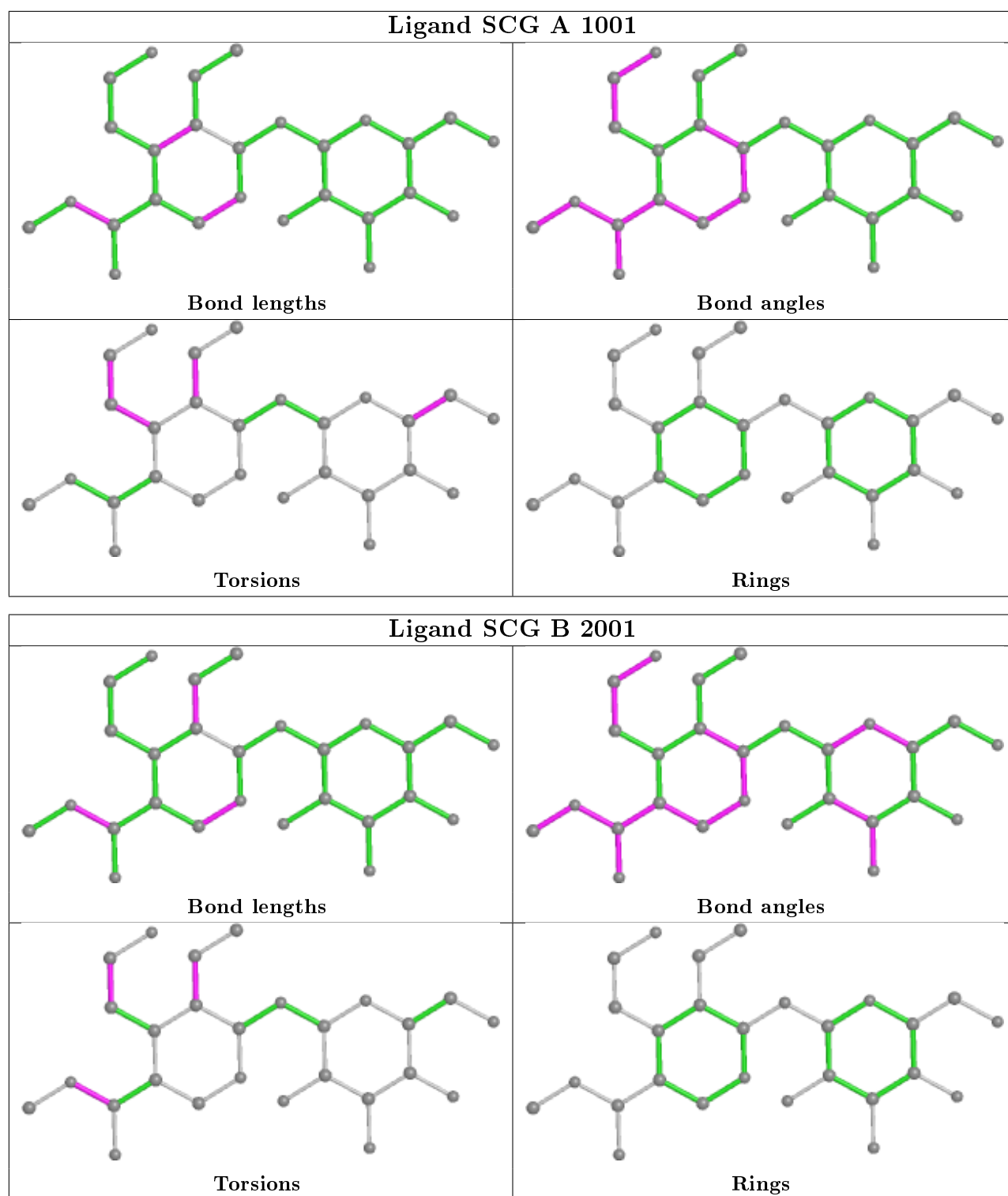
Mol	Chain	Res	Type	Atoms
2	A	1001	SCG	C1-C2-C7-C8
2	B	2001	SCG	C4-C12-O14-C15
2	B	2001	SCG	O13-C12-O14-C15
2	A	1001	SCG	C20-C21-C23-O24
2	A	1001	SCG	O22-C21-C23-O24

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1001	SCG	1	0
2	B	2001	SCG	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 ⓘ

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	302/322 (93%)	-0.76	0	100 100	20, 26, 31, 44	0
1	B	302/322 (93%)	-0.55	2 (0%)	87 69	20, 26, 31, 39	0
All	All	604/644 (93%)	-0.66	2 (0%)	94 84	20, 26, 32, 44	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	275	ASN	4.4
1	B	112	GLN	2.2

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](#)

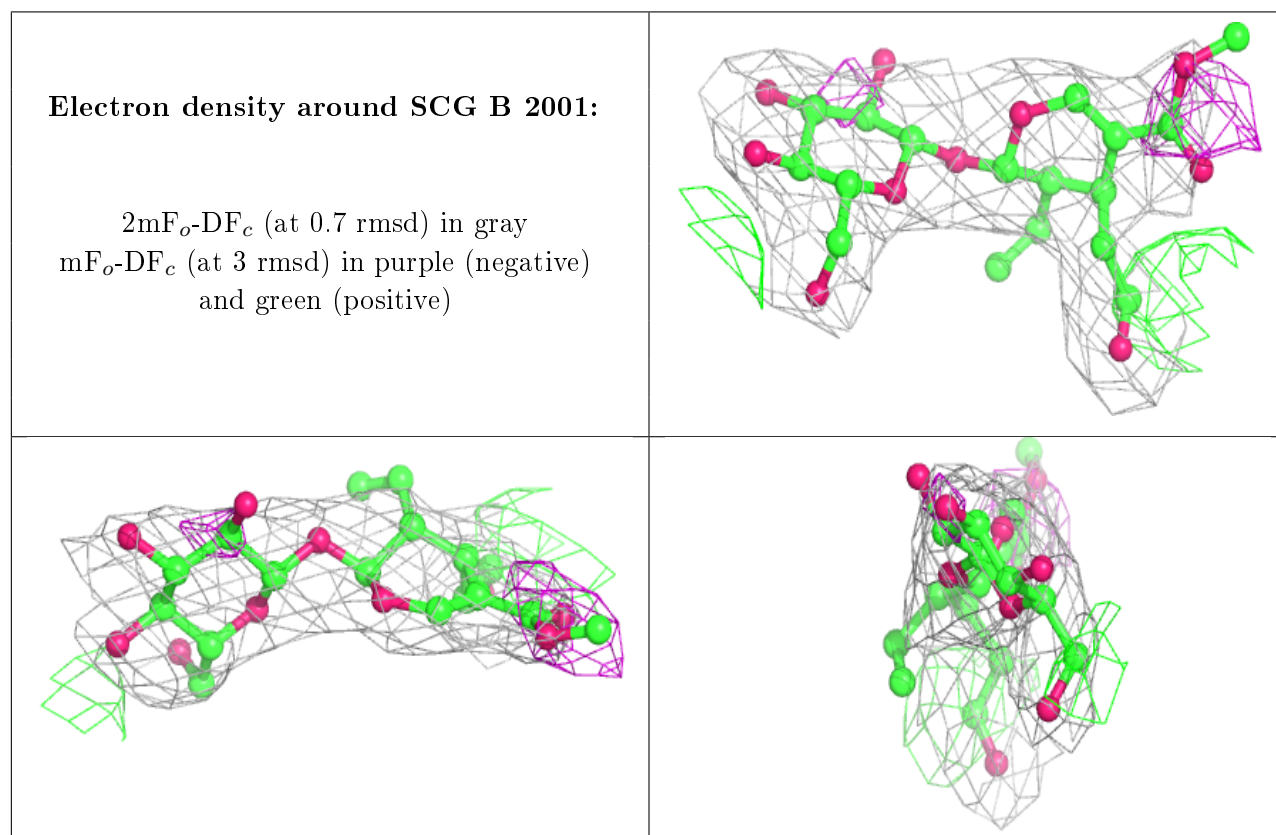
There are no monosaccharides in this entry.

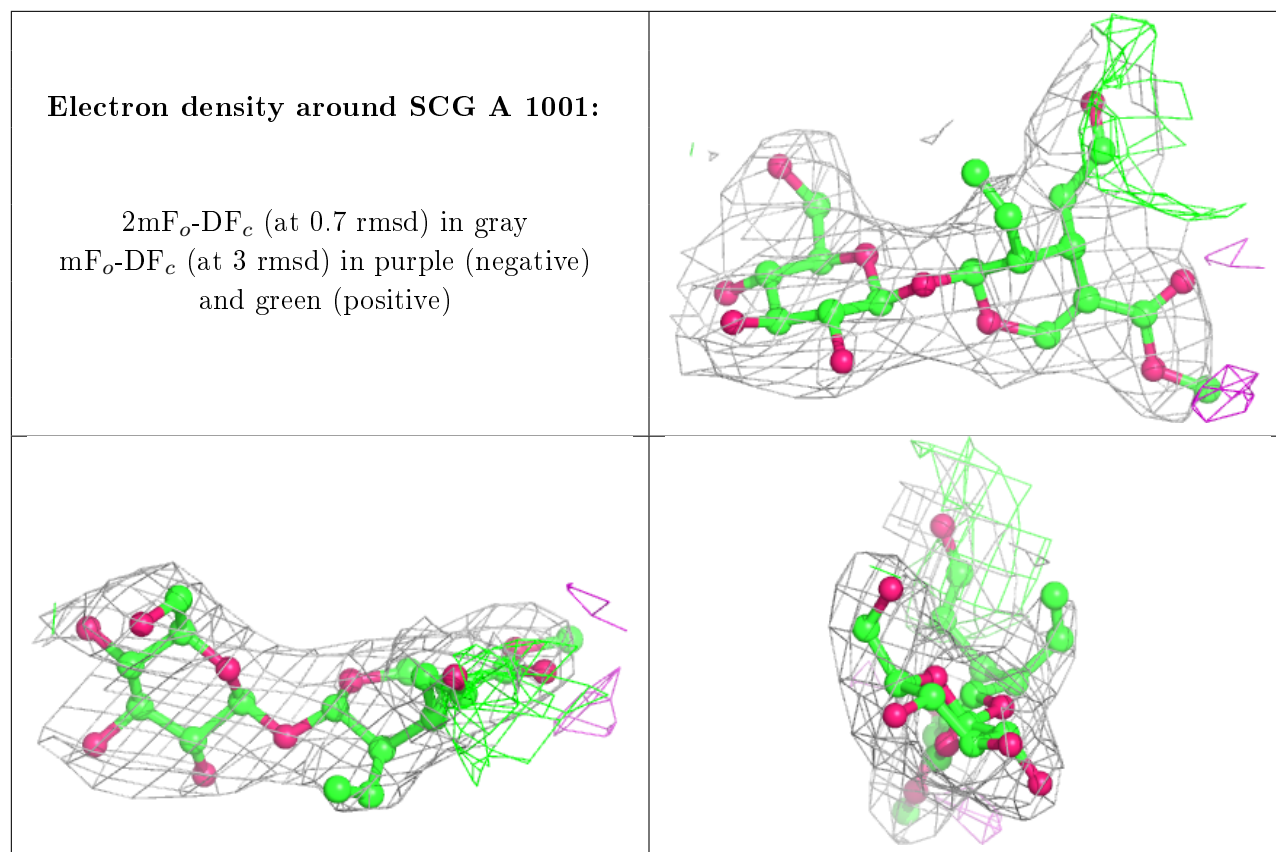
6.4 Ligands [i](#)

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(Å ²)	Q<0.9
2	SCG	B	2001	27/27	0.83	0.27	56,63,71,72	0
2	SCG	A	1001	27/27	0.92	0.16	48,54,63,67	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.