



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 7, 2026 – 05:01 AM UTC

PDB ID : 1GM8 / pdb_00001gm8
Title : Crystal structures of penicillin acylase enzyme-substrate complexes: Structural insights into the catalytic mechanism
Authors : McVey, C.E.; Walsh, M.A.; Dodson, G.G.; Wilson, K.S.; Brannigan, J.A.
Deposited on : 2001-09-11
Resolution : 2.00 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 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

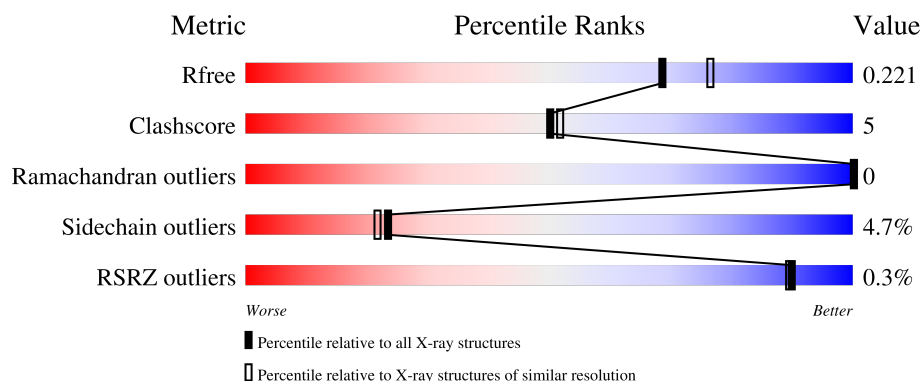
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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}	180053	10052 (2.00-2.00)
Clashscore	190562	11152 (2.00-2.00)
Ramachandran outliers	187476	11031 (2.00-2.00)
Sidechain outliers	187428	11029 (2.00-2.00)
RSRZ outliers	180081	10067 (2.00-2.00)

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 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	209	 73% 22% . .
2	B	557	 70% 25% .

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 6754 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 PENICILLIN G ACYLASE ALPHA SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	207	Total	C	N	O	S	0	0	0
			1662	1061	279	314	8			

- Molecule 2 is a protein called PENICILLIN G ACYLASE BETA SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	557	Total	C	N	O	S	0	1	0
			4423	2810	769	834	10			

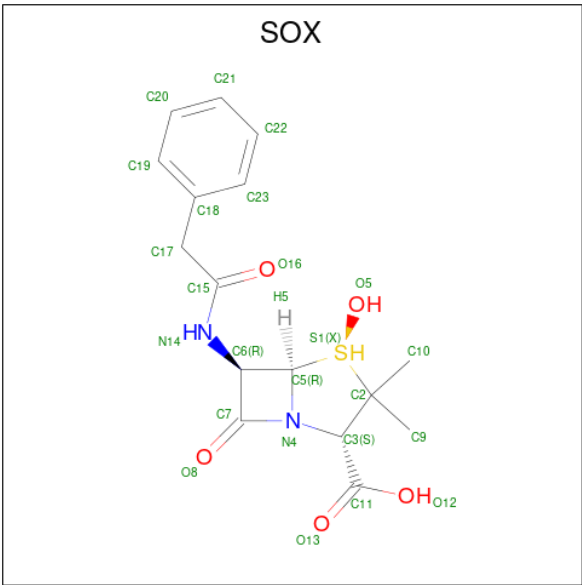
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	241	ALA	ASN	engineered mutation	UNP P06875

- Molecule 3 is CALCIUM ION (CCD ID: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Ca	0	0
			1	1		

- Molecule 4 is N-[(2S,4S,6R)-2-(DIHYDROXYMETHYL)-4-HYDROXY-3,3-DIMETHYL-7-OXO-4LAMBDA 4 -THIA-1-AZABICYCLO[3.2.0]HEPT-6-YL]-2-PHENYLAC ETAMIDE (CCD ID: SOX) (formula: C₁₆H₂₀N₂O₅S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	B	1	Total	C	N	O	S	0	0
			24	16	2	5	1		

- Molecule 5 is water.

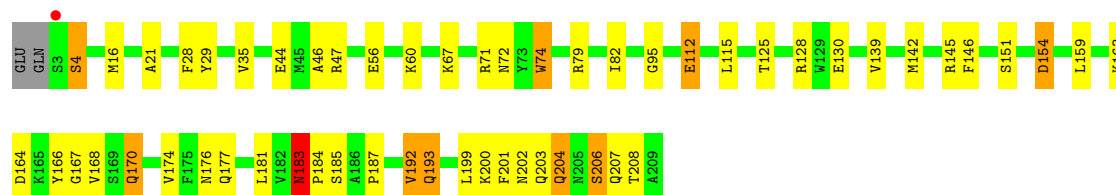
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	182	Total	O	3	3
			182	182		
5	B	462	Total	O	4	4
			462	462		

3 Residue-property plots

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.

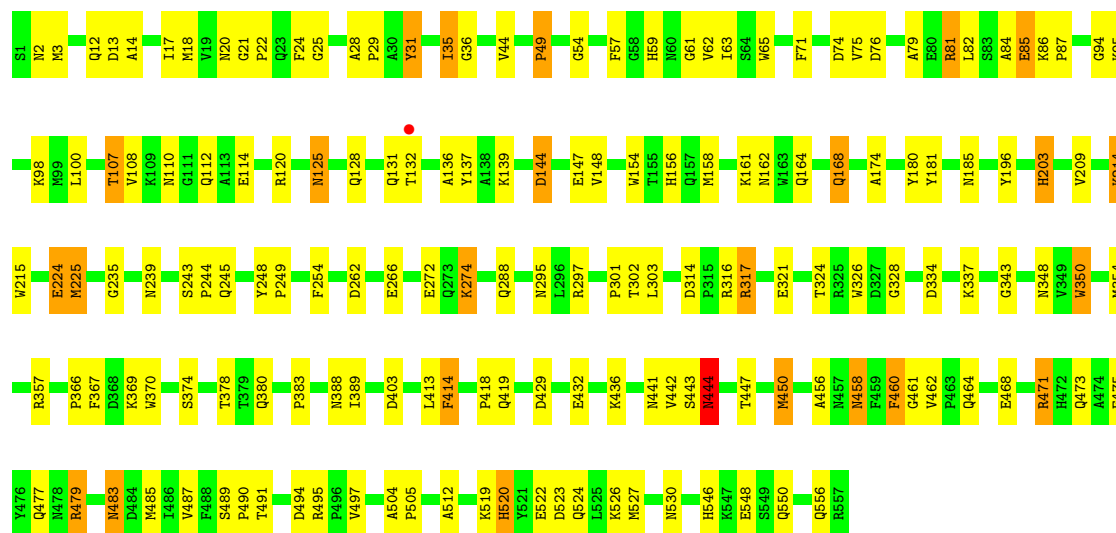
• Molecule 1: PENICILLIN G ACYLASE ALPHA SUBUNIT

Chain A: 



• Molecule 2: PENICILLIN G ACYLASE BETA SUBUNIT

Chain B: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	51.20Å 131.20Å 63.90Å 90.00° 105.80° 90.00°	Depositor
Resolution (Å)	15.00 – 2.00 15.00 – 2.00	Depositor EDS
% Data completeness (in resolution range)	93.4 (15.00-2.00) 93.1 (15.00-2.00)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.63 (at 2.00Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.162 , 0.234 0.154 , 0.221	Depositor DCC
R_{free} test set	1540 reflections (3.02%)	wwPDB-VP
Wilson B-factor (Å ²)	13.6	Xtriage
Anisotropy	0.425	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.41 , 63.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	6754	wwPDB-VP
Average B, all atoms (Å ²)	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.94% 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: SOX, CA

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	1.11	2/1704 (0.1%)	2.03	57/2312 (2.5%)
2	B	1.11	1/4549 (0.0%)	2.04	137/6203 (2.2%)
All	All	1.11	3/6253 (0.0%)	2.04	194/8515 (2.3%)

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	3
2	B	0	5
All	All	0	8

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	130	GLU	N-CA	6.08	1.50	1.46
2	B	57	PHE	CA-CB	5.14	1.61	1.53
1	A	167	GLY	N-CA	5.01	1.50	1.45

All (194) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	81	ARG	CD-NE-CZ	20.42	152.99	124.40
2	B	471	ARG	CD-NE-CZ	19.55	151.76	124.40
2	B	471	ARG	NE-CZ-NH2	-13.95	106.65	119.20
2	B	75	VAL	CA-C-O	-11.99	107.65	120.36
1	A	176	ASN	OD1-CG-ND2	-11.49	111.11	122.60
2	B	471	ARG	NH1-CZ-NH2	9.88	132.14	119.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	460	PHE	CA-CB-CG	9.62	123.42	113.80
1	A	154	ASP	CA-C-O	-9.41	110.58	120.55
2	B	530	ASN	OD1-CG-ND2	-9.02	113.58	122.60
2	B	316	ARG	CA-C-O	8.63	129.88	120.82
2	B	224	GLU	CB-CG-CD	8.49	127.04	112.60
2	B	136	ALA	CA-C-O	-8.21	111.83	120.70
2	B	443	SER	CA-C-N	-7.97	110.17	122.60
2	B	443	SER	C-N-CA	-7.97	110.17	122.60
2	B	144	ASP	CA-C-O	-7.74	112.44	120.80
2	B	477	GLN	OE1-CD-NE2	-7.74	114.86	122.60
1	A	28	PHE	CA-CB-CG	-7.61	106.19	113.80
1	A	177	GLN	OE1-CD-NE2	-7.57	115.03	122.60
1	A	145	ARG	NE-CZ-NH2	-7.54	112.41	119.20
2	B	464	GLN	CA-C-O	7.25	127.24	119.34
2	B	462	VAL	N-CA-C	7.10	116.43	108.05
2	B	348	ASN	CA-CB-CG	7.01	119.61	112.60
2	B	137	TYR	CA-C-O	-7.00	112.80	120.43
2	B	74	ASP	CA-CB-CG	6.94	119.54	112.60
1	A	125	THR	CA-C-N	6.92	126.84	119.85
1	A	125	THR	C-N-CA	6.92	126.84	119.85
1	A	145	ARG	NH1-CZ-NH2	6.87	128.23	119.30
2	B	168	GLN	OE1-CD-NE2	6.86	129.46	122.60
1	A	176	ASN	CB-CG-ND2	6.75	126.53	116.40
2	B	479	ARG	NE-CZ-NH1	-6.75	114.75	121.50
2	B	161	LYS	CB-CG-CD	6.73	126.78	111.30
2	B	29	PRO	CA-C-N	6.71	130.37	120.90
2	B	29	PRO	C-N-CA	6.71	130.37	120.90
2	B	239	ASN	CA-C-O	-6.66	112.35	120.54
2	B	367	PHE	CA-C-O	-6.66	112.31	119.97
1	A	183	ASN	CA-CB-CG	6.65	119.25	112.60
1	A	192	VAL	CB-CA-C	-6.65	103.33	112.04
2	B	350	TRP	CA-C-O	6.60	127.75	120.82
1	A	184	PRO	N-CA-CB	6.54	110.02	103.48
2	B	128	GLN	CB-CG-CD	6.54	123.71	112.60
2	B	441	ASN	CA-C-O	-6.52	114.77	122.13
2	B	243	SER	CA-C-O	6.48	126.30	120.02
2	B	203	HIS	CA-CB-CG	6.47	120.27	113.80
2	B	249	PRO	O-C-N	-6.45	115.88	123.10
1	A	168	VAL	N-CA-C	6.40	116.51	110.30
1	A	145	ARG	CD-NE-CZ	6.36	133.31	124.40
2	B	461	GLY	N-CA-C	-6.36	106.35	115.64
1	A	16	MET	CA-C-N	6.35	126.85	119.99

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	16	MET	C-N-CA	6.35	126.85	119.99
1	A	60	LYS	N-CA-C	6.35	119.01	111.71
2	B	244	PRO	CB-CA-C	-6.34	102.85	112.11
2	B	546	HIS	CA-CB-CG	-6.32	107.48	113.80
1	A	170	GLN	OE1-CD-NE2	-6.31	116.29	122.60
1	A	200	LYS	CA-C-O	-6.29	114.04	120.71
2	B	473	GLN	CG-CD-NE2	6.28	125.82	116.40
1	A	176	ASN	CA-CB-CG	-6.26	106.34	112.60
2	B	14	ALA	CA-C-O	-6.24	114.24	121.10
1	A	154	ASP	O-C-N	6.21	128.70	122.12
2	B	473	GLN	OE1-CD-NE2	-6.20	116.40	122.60
2	B	21	GLY	CA-C-N	6.18	126.39	119.90
2	B	21	GLY	C-N-CA	6.18	126.39	119.90
2	B	357	ARG	NE-CZ-NH2	6.17	124.76	119.20
2	B	254	PHE	O-C-N	-6.16	115.01	122.22
2	B	487	VAL	CA-C-O	-6.15	113.93	120.39
2	B	374	SER	CA-C-O	6.14	126.49	119.18
2	B	334	ASP	CA-CB-CG	6.14	118.74	112.60
2	B	512	ALA	CA-C-N	6.13	126.31	119.32
2	B	512	ALA	C-N-CA	6.13	126.31	119.32
2	B	161	LYS	CA-CB-CG	6.09	126.28	114.10
1	A	47	ARG	CD-NE-CZ	6.08	132.92	124.40
2	B	288	GLN	OE1-CD-NE2	-6.07	116.53	122.60
2	B	162	ASN	O-C-N	-6.07	116.89	123.26
2	B	107	THR	CA-C-O	-6.03	114.17	121.05
1	A	146	PHE	N-CA-C	6.03	120.47	113.18
1	A	192	VAL	O-C-N	-6.02	115.99	121.89
2	B	489	SER	CA-C-N	5.99	125.93	119.76
2	B	489	SER	C-N-CA	5.99	125.93	119.76
2	B	20	ASN	CA-CB-CG	5.92	118.52	112.60
2	B	369	LYS	CA-C-O	-5.91	112.33	119.49
2	B	295	ASN	CA-CB-CG	-5.88	106.72	112.60
2	B	456	ALA	CA-C-N	5.88	130.47	122.19
2	B	456	ALA	C-N-CA	5.88	130.47	122.19
2	B	272	GLU	CA-CB-CG	-5.87	102.35	114.10
1	A	46	ALA	N-CA-C	-5.87	104.96	111.36
1	A	151	SER	N-CA-C	-5.84	103.50	111.56
2	B	137	TYR	CA-C-N	-5.83	114.47	122.93
2	B	137	TYR	C-N-CA	-5.83	114.47	122.93
2	B	76	ASP	OD1-CG-OD2	-5.82	108.93	122.90
2	B	61	GLY	N-CA-C	-5.81	108.17	115.08
2	B	248	TYR	CA-C-N	5.80	125.99	119.90

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	248	TYR	C-N-CA	5.80	125.99	119.90
1	A	145	ARG	CA-C-O	-5.78	112.45	119.61
2	B	483	ASN	O-C-N	5.77	130.31	123.27
1	A	187	PRO	N-CA-CB	5.73	108.27	103.17
2	B	429	ASP	CB-CA-C	-5.73	101.28	110.79
2	B	380	GLN	N-CA-C	-5.70	106.28	113.18
2	B	196	TYR	CA-CB-CG	-5.70	103.64	113.90
2	B	456	ALA	O-C-N	-5.70	114.32	122.41
2	B	443	SER	CA-C-O	-5.69	111.94	119.10
2	B	181	TYR	N-CA-CB	-5.68	100.71	111.00
2	B	49	PRO	N-CA-CB	5.66	108.36	103.27
1	A	167	GLY	CA-C-O	-5.64	118.06	122.52
2	B	137	TYR	O-C-N	5.62	129.62	123.27
2	B	491	THR	N-CA-C	-5.61	106.43	113.28
2	B	28	ALA	CB-CA-C	-5.61	101.82	110.62
1	A	142	MET	N-CA-C	5.59	117.82	111.11
2	B	148	VAL	CA-C-O	-5.58	114.93	120.85
2	B	497	VAL	CA-CB-CG2	5.58	119.88	110.40
2	B	245	GLN	OE1-CD-NE2	5.57	128.16	122.60
2	B	108	VAL	CA-C-O	-5.56	114.66	120.27
2	B	243	SER	O-C-N	-5.56	115.88	121.27
2	B	520	HIS	CA-CB-CG	-5.55	108.25	113.80
2	B	54	GLY	CA-C-N	-5.54	113.58	121.50
2	B	54	GLY	C-N-CA	-5.54	113.58	121.50
2	B	110	ASN	CB-CA-C	-5.54	103.72	111.80
2	B	2	ASN	CA-C-O	5.53	127.99	121.58
1	A	79	ARG	CA-C-O	-5.53	114.69	120.55
2	B	185	ASN	CA-CB-CG	-5.52	107.08	112.60
1	A	204	GLN	CA-CB-CG	5.52	125.14	114.10
2	B	366	PRO	CB-CA-C	-5.51	98.22	112.00
1	A	130	GLU	CA-C-N	5.51	126.72	119.84
1	A	130	GLU	C-N-CA	5.51	126.72	119.84
2	B	225	MET	CA-CB-CG	-5.49	103.11	114.10
2	B	447	THR	CB-CA-C	5.49	117.34	109.11
2	B	475	GLU	CA-CB-CG	-5.48	103.14	114.10
2	B	125	ASN	CA-CB-CG	5.47	118.07	112.60
2	B	235	GLY	N-CA-C	-5.46	107.79	115.32
2	B	418	PRO	N-CA-CB	5.44	108.16	103.27
2	B	174	ALA	N-CA-CB	-5.43	101.94	110.46
2	B	297	ARG	CD-NE-CZ	-5.43	116.80	124.40
1	A	115	LEU	CA-C-O	-5.43	114.93	119.76
2	B	314	ASP	CA-C-N	5.42	124.94	119.19

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	314	ASP	C-N-CA	5.42	124.94	119.19
2	B	321	GLU	CA-CB-CG	-5.41	103.29	114.10
2	B	18	MET	CA-C-O	5.40	126.19	120.36
2	B	429	ASP	CA-C-O	5.39	126.27	120.55
2	B	490	PRO	N-CA-CB	5.39	108.11	103.31
2	B	57	PHE	CA-CB-CG	-5.38	108.42	113.80
1	A	35	VAL	CA-C-N	5.34	127.44	120.28
1	A	35	VAL	C-N-CA	5.34	127.44	120.28
2	B	316	ARG	O-C-N	-5.34	116.57	122.07
2	B	442	VAL	O-C-N	-5.34	116.66	121.89
1	A	21	ALA	CA-C-O	-5.33	115.75	121.45
1	A	125	THR	O-C-N	-5.32	116.42	121.57
2	B	530	ASN	CB-CG-ND2	5.29	124.34	116.40
1	A	206	SER	CA-CB-OG	-5.27	100.56	111.10
1	A	28	PHE	CA-C-N	5.27	127.29	120.44
1	A	28	PHE	C-N-CA	5.27	127.29	120.44
1	A	128	ARG	CA-CB-CG	-5.25	103.60	114.10
2	B	483	ASN	CA-C-O	-5.25	114.62	120.66
2	B	85	GLU	O-C-N	-5.25	116.08	122.22
2	B	94	GLY	N-CA-C	5.23	122.33	115.47
1	A	174	VAL	N-CA-CB	5.22	118.41	110.58
2	B	156	HIS	CA-CB-CG	-5.21	108.59	113.80
2	B	266	GLU	CA-C-O	-5.20	114.22	120.10
2	B	162	ASN	CA-CB-CG	-5.18	107.42	112.60
2	B	62	VAL	O-C-N	5.18	125.97	121.98
1	A	154	ASP	CA-CB-CG	-5.17	107.43	112.60
2	B	84	ALA	CA-C-O	-5.16	114.27	120.10
2	B	302	THR	O-C-N	5.16	127.59	122.12
2	B	328	GLY	N-CA-C	-5.15	108.21	115.32
2	B	164	GLN	CG-CD-NE2	5.14	124.11	116.40
2	B	301	PRO	N-CA-CB	5.14	108.75	103.15
2	B	25	GLY	N-CA-C	-5.12	105.00	112.17
1	A	115	LEU	N-CA-CB	-5.12	101.66	110.01
2	B	450	MET	N-CA-CB	-5.12	102.50	110.29
2	B	79	ALA	CA-C-O	-5.12	115.45	120.98
2	B	31	TYR	O-C-N	5.12	127.55	122.12
2	B	168	GLN	CB-CA-C	-5.11	102.32	110.79
2	B	324	THR	CA-C-O	5.10	125.86	120.10
1	A	4	SER	O-C-N	-5.09	116.26	122.22
1	A	82	ILE	CA-C-N	5.09	127.62	120.28
1	A	82	ILE	C-N-CA	5.09	127.62	120.28
2	B	262	ASP	CA-CB-CG	-5.09	107.51	112.60

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	82	LEU	CA-C-N	-5.09	113.78	122.33
2	B	82	LEU	C-N-CA	-5.09	113.78	122.33
1	A	125	THR	CA-C-O	5.09	126.46	120.41
2	B	87	PRO	N-CA-CB	5.08	107.84	103.27
2	B	168	GLN	CB-CG-CD	-5.06	104.00	112.60
2	B	475	GLU	N-CA-C	5.05	116.88	107.99
2	B	76	ASP	CB-CG-OD2	5.04	130.00	118.40
1	A	74	TRP	CA-C-N	5.04	125.06	119.32
1	A	74	TRP	C-N-CA	5.04	125.06	119.32
2	B	17	ILE	CA-C-N	5.04	130.09	122.99
2	B	17	ILE	C-N-CA	5.04	130.09	122.99
2	B	414	PHE	CA-CB-CG	-5.04	108.76	113.80
1	A	74	TRP	O-C-N	-5.03	116.74	121.32
2	B	524	GLN	OE1-CD-NE2	5.03	127.63	122.60
1	A	67	LYS	N-CA-CB	5.03	117.51	110.12
2	B	156	HIS	O-C-N	-5.03	116.08	122.27
1	A	202	ASN	OD1-CG-ND2	5.01	127.61	122.60
1	A	183	ASN	CA-C-N	5.01	124.67	119.56
1	A	183	ASN	C-N-CA	5.01	124.67	119.56
2	B	485	MET	CA-C-O	5.00	126.60	120.60

There are no chirality outliers.

All (8) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	203	GLN	Mainchain
1	A	206	SER	Mainchain
1	A	44	GLU	Mainchain
2	B	144	ASP	Mainchain
2	B	403	ASP	Mainchain
2	B	444	ASN	Mainchain
2	B	495[B]	ARG	Mainchain
2	B	85	GLU	Mainchain

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	1662	0	1609	20	1
2	B	4423	0	4254	52	1
3	B	1	0	0	0	0
4	B	24	0	20	1	0
5	A	182	0	0	3	0
5	B	462	0	0	4	0
All	All	6754	0	5883	64	1

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 (64) 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)
2:B:214:LYS:H	2:B:214:LYS:HE2	1.18	1.06
2:B:214:LYS:H	2:B:214:LYS:CE	1.91	0.83
1:A:181:LEU:HD21	1:A:201:PHE:HB2	1.63	0.81
2:B:479:ARG:HH21	2:B:483:ASN:HD22	1.31	0.79
2:B:317:ARG:HG3	2:B:317:ARG:HH11	1.48	0.77
2:B:214:LYS:HE2	2:B:214:LYS:N	1.97	0.76
1:A:164:ASP:HB3	5:A:2137:HOH:O	1.88	0.72
2:B:388:ASN:HD22	2:B:389:ILE:H	1.38	0.71
2:B:520:HIS:HE1	2:B:548:GLU:OE2	1.74	0.70
2:B:120:ARG:HH11	2:B:125:ASN:ND2	1.90	0.69
1:A:199:LEU:HD11	2:B:225:MET:HE3	1.76	0.66
2:B:444:ASN:C	2:B:444:ASN:HD22	2.04	0.66
1:A:71:ARG:NH2	5:A:2055:HOH:O	2.27	0.64
2:B:203:HIS:HD2	5:B:2173:HOH:O	1.81	0.62
2:B:520:HIS:HD2	2:B:523:ASP:OD2	1.82	0.61
1:A:199:LEU:HD21	2:B:225:MET:HE1	1.83	0.61
2:B:458:ASN:C	2:B:458:ASN:HD22	2.09	0.60
1:A:56:GLU:HG2	2:B:107:THR:HB	1.85	0.59
2:B:44:VAL:HG11	2:B:158:MET:HB3	1.83	0.59
2:B:471:ARG:NH2	2:B:522:GLU:OE1	2.31	0.58
1:A:183:ASN:C	1:A:183:ASN:HD22	2.15	0.55
2:B:458:ASN:HD21	2:B:460:PHE:HB2	1.71	0.55
2:B:504:ALA:HA	2:B:505:PRO:C	2.31	0.55
1:A:183:ASN:HD22	1:A:185:SER:H	1.54	0.55
2:B:388:ASN:ND2	2:B:389:ILE:H	2.03	0.54
5:A:2129:HOH:O	2:B:203:HIS:HE1	1.90	0.54
2:B:378:THR:CG2	2:B:450:MET:HE2	2.38	0.53
2:B:214:LYS:HG2	2:B:215:TRP:CD1	2.43	0.53

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:468:GLU:CD	2:B:468:GLU:H	2.16	0.53
1:A:72:ASN:HD21	2:B:139:LYS:NZ	2.06	0.53
2:B:378:THR:HG22	2:B:383:PRO:HG3	1.91	0.52
1:A:183:ASN:ND2	1:A:185:SER:H	2.08	0.51
1:A:112:GLU:H	1:A:112:GLU:CD	2.17	0.51
1:A:139:VAL:HG22	2:B:147:GLU:HB3	1.94	0.49
2:B:479:ARG:HH21	2:B:483:ASN:ND2	2.06	0.49
2:B:3:MET:HE3	2:B:180:TYR:HB2	1.95	0.49
1:A:181:LEU:HD21	1:A:201:PHE:CB	2.38	0.47
2:B:388:ASN:HD22	2:B:389:ILE:N	2.08	0.47
1:A:193:GLN:H	1:A:193:GLN:NE2	2.12	0.47
1:A:166:TYR:HB3	1:A:170:GLN:HG2	1.97	0.46
1:A:29:TYR:HA	1:A:95:GLY:O	2.15	0.45
2:B:35:ILE:HG13	2:B:36:GLY:N	2.31	0.45
2:B:274:LYS:HE2	5:B:2226:HOH:O	2.16	0.45
2:B:12:GLN:O	2:B:13:ASP:HB2	2.17	0.45
2:B:71:PHE:CZ	4:B:1559:SOX:H5	2.52	0.44
2:B:378:THR:CG2	2:B:383:PRO:HG3	2.48	0.44
1:A:159:LEU:HD11	1:A:163:LYS:HE3	2.01	0.43
2:B:378:THR:HG23	2:B:450:MET:HE2	2.00	0.43
2:B:22:PRO:HB2	2:B:24:PHE:CE2	2.53	0.42
2:B:112:GLN:NE2	5:B:2095:HOH:O	2.46	0.42
2:B:444:ASN:C	2:B:444:ASN:ND2	2.72	0.42
2:B:59:HIS:HB2	2:B:63:ILE:O	2.20	0.41
2:B:303:LEU:HD21	2:B:350:TRP:CE2	2.55	0.41
2:B:414:PHE:CD1	2:B:419:GLN:HG2	2.55	0.41
2:B:350:TRP:CZ2	2:B:354:MET:HG3	2.56	0.41
2:B:526:LYS:HD3	2:B:526:LYS:HA	1.86	0.41
2:B:527:MET:HE2	5:B:2433:HOH:O	2.21	0.41
1:A:154:ASP:HB3	2:B:370:TRP:CH2	2.56	0.41
1:A:207:GLN:HA	2:B:215:TRP:CZ2	2.56	0.41
2:B:326:TRP:CZ3	2:B:343:GLY:HA3	2.56	0.41
2:B:414:PHE:CG	2:B:419:GLN:HG2	2.56	0.41
1:A:74:TRP:CZ2	2:B:100:LEU:HG	2.57	0.40
2:B:31:TYR:CE2	2:B:49:PRO:HB3	2.56	0.40
2:B:65:TRP:HA	2:B:180:TYR:O	2.22	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:208:THR:OG1	2:B:494:ASP:OD1[2_645]	2.15	0.05

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	205/209 (98%)	199 (97%)	6 (3%)	0	100	100
2	B	556/557 (100%)	545 (98%)	11 (2%)	0	100	100
All	All	761/766 (99%)	744 (98%)	17 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	178/180 (99%)	172 (97%)	6 (3%)	32	33
2	B	460/459 (100%)	436 (95%)	24 (5%)	21	18
All	All	638/639 (100%)	608 (95%)	30 (5%)	23	22

All (30) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	4	SER
1	A	112	GLU
1	A	183	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	192	VAL
1	A	193	GLN
1	A	204	GLN
2	B	35	ILE
2	B	81	ARG
2	B	86	LYS
2	B	95	LYS
2	B	98	LYS
2	B	114	GLU
2	B	131	GLN
2	B	132	THR
2	B	154	TRP
2	B	168	GLN
2	B	209	VAL
2	B	214	LYS
2	B	224	GLU
2	B	274	LYS
2	B	317	ARG
2	B	337	LYS
2	B	413	LEU
2	B	432	GLU
2	B	436	LYS
2	B	444	ASN
2	B	458	ASN
2	B	519	LYS
2	B	550	GLN
2	B	556	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (23) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	22	ASN
1	A	72	ASN
1	A	120	ASN
1	A	144	ASN
1	A	183	ASN
1	A	193	GLN
1	A	204	GLN
2	B	93	ASN
2	B	112	GLN
2	B	125	ASN
2	B	134	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	203	HIS
2	B	233	GLN
2	B	239	ASN
2	B	245	GLN
2	B	318	GLN
2	B	348	ASN
2	B	388	ASN
2	B	409	GLN
2	B	444	ASN
2	B	458	ASN
2	B	483	ASN
2	B	520	HIS

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](#)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 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	SOX	B	1559	-	21,26,26	0.91	0	28,40,40	1.52	3 (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	SOX	B	1559	-	-	2/12/48/48	0/3/3/3

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	1559	SOX	O16-C15-N14	4.22	130.10	122.95
4	B	1559	SOX	C2-C3-N4	-3.35	103.04	105.89
4	B	1559	SOX	O16-C15-C17	-2.62	116.32	121.99

There are no chirality outliers.

All (2) torsion outliers are listed below:

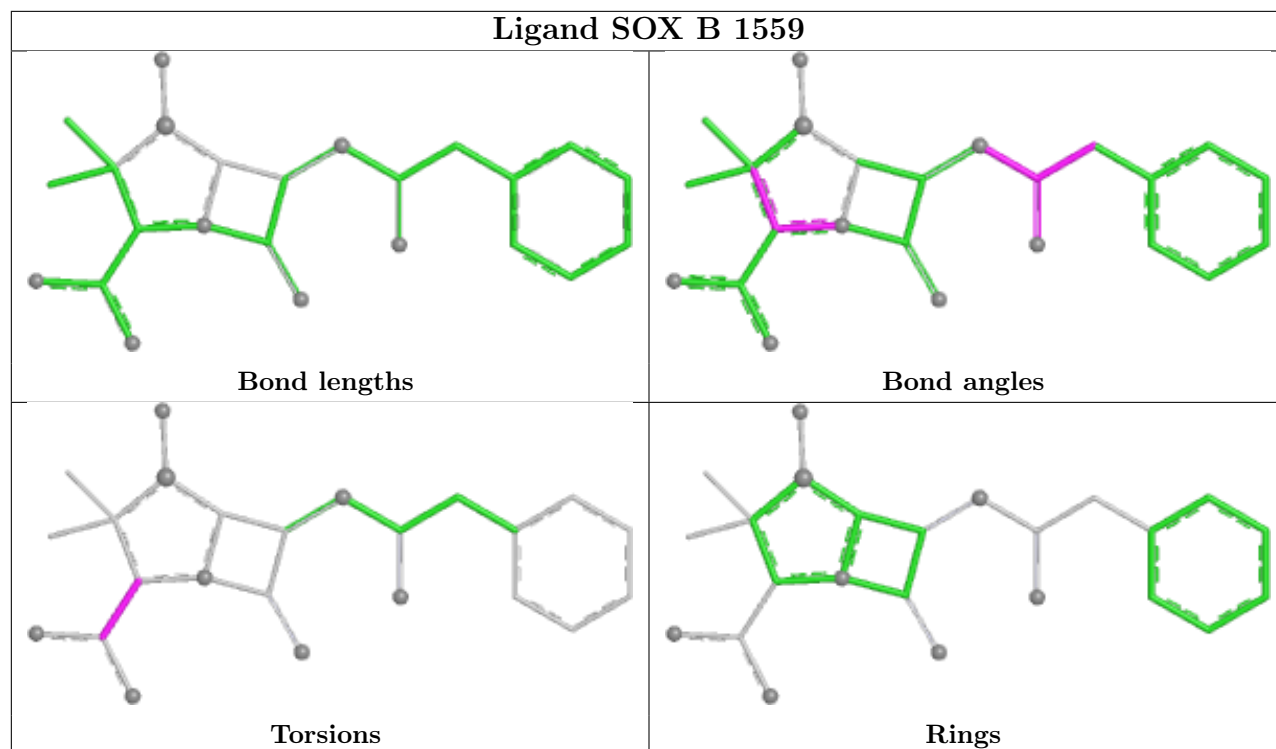
Mol	Chain	Res	Type	Atoms
4	B	1559	SOX	O13-C11-C3-C2
4	B	1559	SOX	O12-C11-C3-C2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	1559	SOX	1	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	207/209 (99%)	-0.80	1 (0%) 87 87	9, 16, 27, 46	0
2	B	557/557 (100%)	-0.82	1 (0%) 91 90	7, 15, 32, 72	1 (0%)
All	All	764/766 (99%)	-0.81	2 (0%) 90 89	7, 15, 31, 72	1 (0%)

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	3	SER	2.4
2	B	132	THR	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](#)

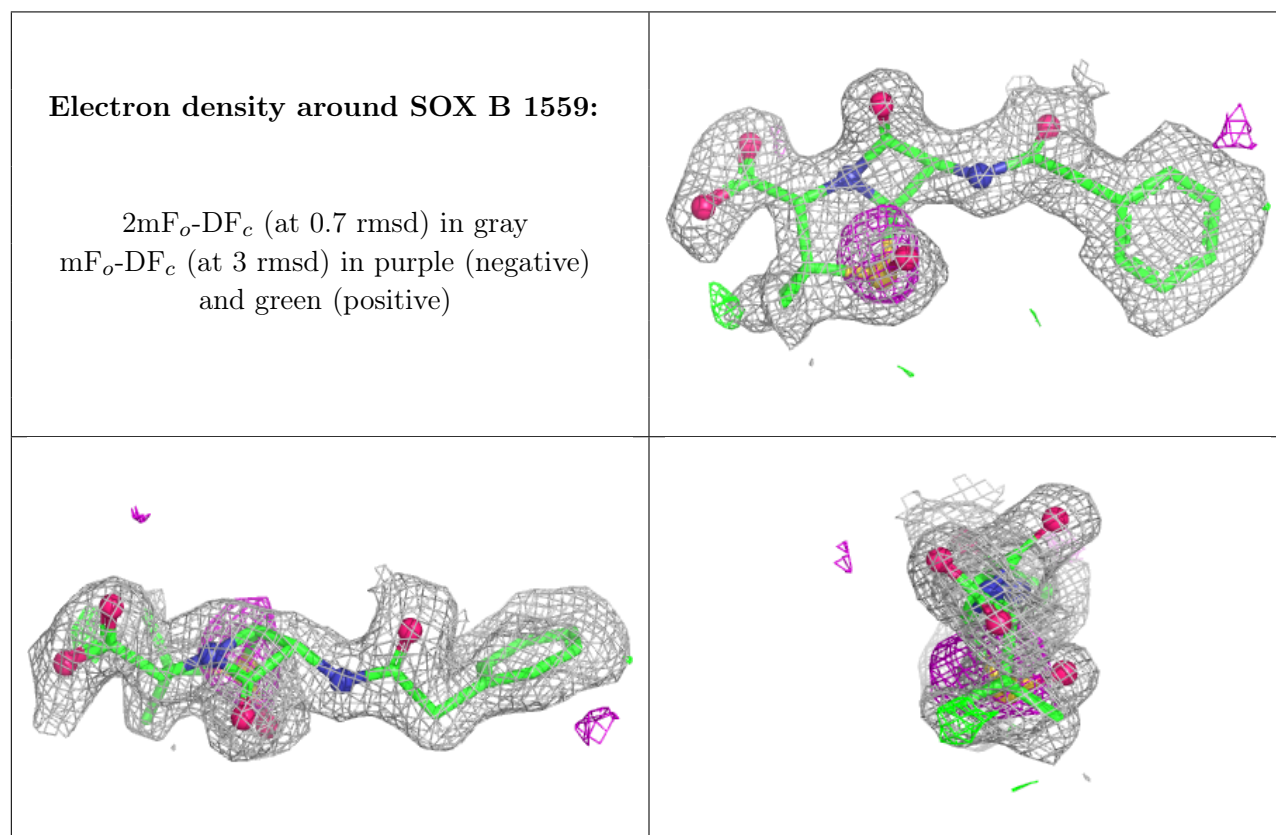
There are no oligosaccharides in this entry.

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(Å ²)	Q<0.9
4	SOX	B	1559	24/24	0.90	0.09	17,26,35,38	0
3	CA	B	1558	1/1	1.00	0.04	14,14,14,14	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.