



Full wwPDB X-ray Structure Validation Report ⓘ

Dec 1, 2016 – 11:47 AM EST

PDB ID : 1GM9
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 : 2001-09-12
Resolution : 1.80 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

<http://wwpdb.org/validation/2016/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 : 1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20028320
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20028320

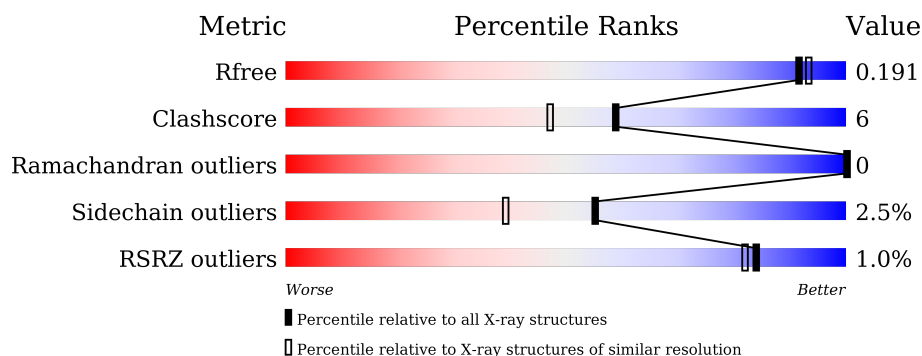
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION



The reported resolution of this entry is 1.80 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}	91344	4533 (1.80-1.80)
Clashscore	102246	5383 (1.80-1.80)
Ramachandran outliers	100387	5320 (1.80-1.80)
Sidechain outliers	100360	5319 (1.80-1.80)
RSRZ outliers	91569	4547 (1.80-1.80)

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. 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	
2	B	557	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	EDO	A	1210	-	-	X	X
3	EDO	B	1559	-	-	-	X
3	EDO	B	1560	-	-	-	X
3	EDO	B	1563	-	-	-	X
3	EDO	B	1564	-	-	-	X
3	EDO	B	1565	-	-	-	X
5	SOX	B	1569	-	-	-	X

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 6882 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.

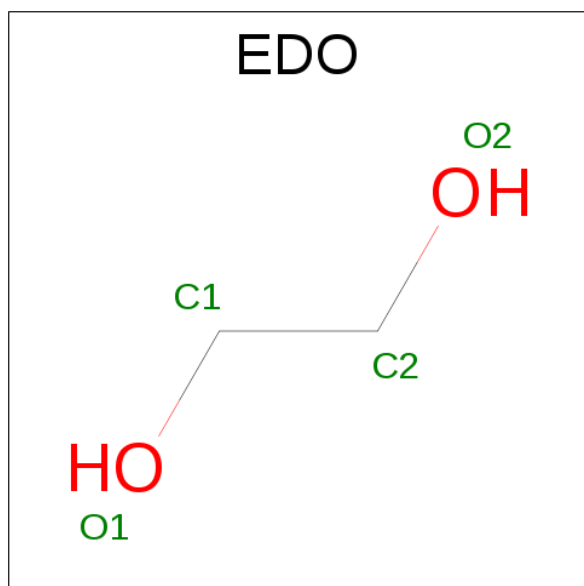
- Molecule 1 is a protein called PENICILLIN G ACYLASE ALPHA SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	207	Total	C	N	O	S	8	2	0
			1678	1072	280	317	9			

- 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	3	2	0
			4434	2815	772	837	10			

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			4	2	2		
3	B	1	Total	C	O	0	0
			4	2	2		

Continued on next page...

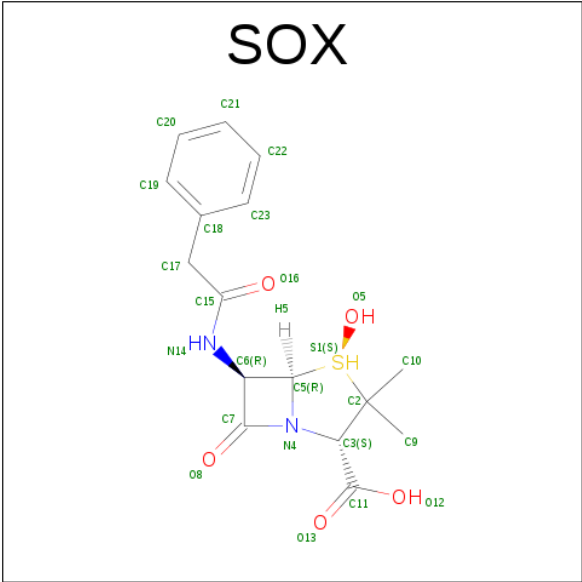
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

- Molecule 5 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-PHENYLACETAMIDE
(three-letter code: SOX) (formula: C₁₆H₂₀N₂O₅S).



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

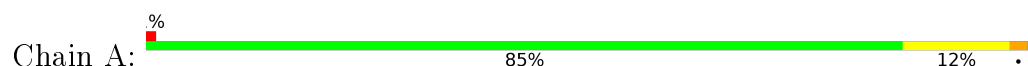
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	192	Total	O	3	3
			192	192		
6	B	509	Total	O	2	2
			509	509		

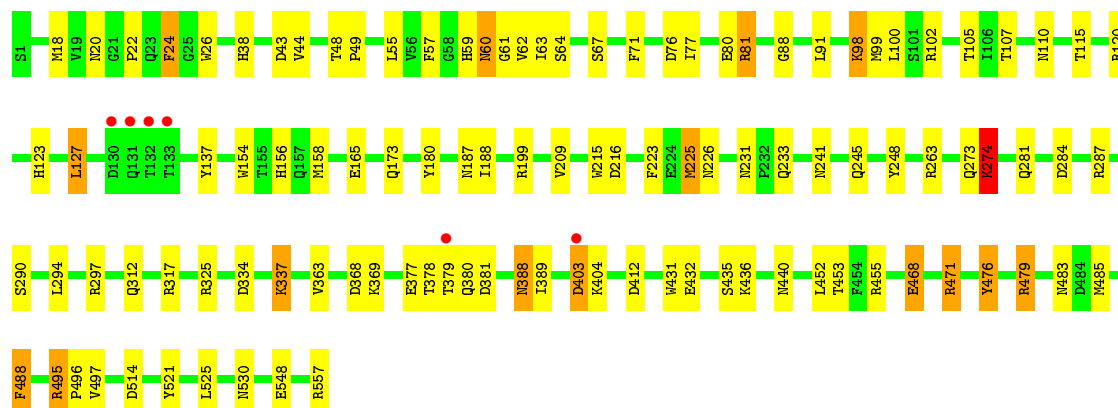
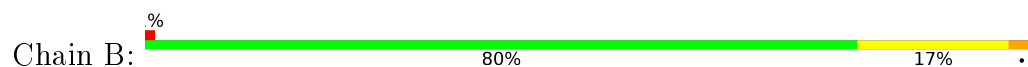
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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



• Molecule 2: PENICILLIN G ACYLASE BETA SUBUNIT



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	51.20 Å 131.70 Å 64.00 Å 90.00° 105.90° 90.00°	Depositor
Resolution (Å)	30.00 – 1.80 29.31 – 1.76	Depositor EDS
% Data completeness (in resolution range)	96.5 (30.00-1.80) 96.3 (29.31-1.76)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.00 (at 1.76 Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.153 , 0.190 0.155 , 0.191	Depositor DCC
R_{free} test set	2195 reflections (3.08%)	DCC
Wilson B-factor (Å ²)	13.1	Xtriage
Anisotropy	0.143	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 57.2	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	0.96	EDS
Total number of atoms	6882	wwPDB-VP
Average B, all atoms (Å ²)	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.90% 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: SOX, CA, SME, EDO

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	0.97	2/1711 (0.1%)	1.62	22/2319 (0.9%)
2	B	1.04	3/4560 (0.1%)	1.59	67/6218 (1.1%)
All	All	1.02	5/6271 (0.1%)	1.60	89/8537 (1.0%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	432	GLU	CG-CD	38.74	2.10	1.51
1	A	60	LYS	CD-CE	16.98	1.93	1.51
2	B	290	SER	CA-CB	7.74	1.64	1.52
2	B	165	GLU	CD-OE2	7.49	1.33	1.25
1	A	204	GLN	CG-CD	7.08	1.67	1.51

All (89) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	455	ARG	NE-CZ-NH1	17.20	128.90	120.30
2	B	455	ARG	NE-CZ-NH2	-15.02	112.79	120.30
1	A	48	ARG	NE-CZ-NH2	-14.16	113.22	120.30
2	B	337	LYS	CA-CB-CG	11.99	139.77	113.40
2	B	381	ASP	CB-CG-OD2	-11.42	108.03	118.30
1	A	70	ARG	NE-CZ-NH2	-11.01	114.79	120.30
2	B	137	TYR	CB-CG-CD2	-10.98	114.41	121.00
1	A	61	ASP	CB-CG-OD2	-10.92	108.47	118.30
1	A	39	ARG	NE-CZ-NH2	-10.29	115.15	120.30
2	B	495[A]	ARG	NE-CZ-NH2	9.99	125.30	120.30
2	B	495[B]	ARG	NE-CZ-NH2	9.99	125.30	120.30
2	B	297	ARG	NE-CZ-NH1	9.94	125.27	120.30
1	A	23	ASP	CB-CG-OD1	9.78	127.10	118.30
1	A	71	ARG	NE-CZ-NH1	-9.61	115.50	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	105	ASP	CB-CG-OD1	9.56	126.90	118.30
1	A	71	ARG	CD-NE-CZ	9.54	136.96	123.60
2	B	137	TYR	CB-CG-CD1	9.32	126.59	121.00
2	B	199	ARG	NE-CZ-NH1	8.88	124.74	120.30
2	B	557	ARG	NE-CZ-NH1	-8.56	116.02	120.30
2	B	479	ARG	NE-CZ-NH2	-8.51	116.04	120.30
1	A	48	ARG	NE-CZ-NH1	8.37	124.48	120.30
2	B	81	ARG	NE-CZ-NH2	8.28	124.44	120.30
2	B	378	THR	C-N-CA	8.13	142.03	121.70
2	B	64	SER	N-CA-CB	8.04	122.56	110.50
1	A	60	LYS	CD-CE-NZ	7.92	129.92	111.70
2	B	471	ARG	CD-NE-CZ	7.88	134.63	123.60
1	A	70	ARG	NH1-CZ-NH2	7.75	127.92	119.40
1	A	96	TYR	CB-CG-CD2	-7.54	116.48	121.00
2	B	81	ARG	CD-NE-CZ	7.51	134.12	123.60
2	B	102	ARG	NE-CZ-NH2	-7.42	116.59	120.30
2	B	403	ASP	CB-CG-OD1	-7.28	111.75	118.30
2	B	488	PHE	CB-CG-CD1	-7.09	115.84	120.80
2	B	455	ARG	CD-NE-CZ	6.91	133.27	123.60
2	B	199	ARG	NE-CZ-NH2	-6.90	116.85	120.30
2	B	290	SER	CB-CA-C	-6.81	97.16	110.10
2	B	263	ARG	NE-CZ-NH1	-6.76	116.92	120.30
2	B	76	ASP	CB-CG-OD1	6.64	124.27	118.30
2	B	317	ARG	NE-CZ-NH1	6.61	123.61	120.30
2	B	403	ASP	OD1-CG-OD2	6.55	135.75	123.30
2	B	403	ASP	CB-CG-OD2	-6.55	112.40	118.30
2	B	471	ARG	NE-CZ-NH1	6.52	123.56	120.30
2	B	223	PHE	CB-CG-CD2	6.46	125.32	120.80
2	B	521	TYR	CB-CG-CD1	6.42	124.85	121.00
2	B	412	ASP	CB-CG-OD1	6.24	123.92	118.30
2	B	297	ARG	NE-CZ-NH2	-6.20	117.20	120.30
1	A	61	ASP	OD1-CG-OD2	6.12	134.93	123.30
2	B	290	SER	CA-CB-OG	-6.11	94.71	111.20
1	A	204	GLN	CG-CD-OE1	-6.09	109.43	121.60
2	B	24	PHE	CB-CG-CD1	-6.08	116.54	120.80
2	B	312	GLN	CB-CG-CD	6.07	127.39	111.60
1	A	70	ARG	NE-CZ-NH1	-6.04	117.28	120.30
2	B	379	THR	O-C-N	-6.04	113.03	122.70
2	B	165	GLU	CG-CD-OE1	5.97	130.23	118.30
2	B	325	ARG	NE-CZ-NH2	-5.86	117.37	120.30
2	B	274	LYS	CA-CB-CG	5.84	126.24	113.40
2	B	325	ARG	NE-CZ-NH1	5.83	123.21	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	334	ASP	CB-CG-OD1	5.80	123.52	118.30
2	B	377	GLU	OE1-CD-OE2	5.79	130.25	123.30
2	B	432	GLU	CB-CG-CD	-5.73	98.72	114.20
2	B	476	TYR	CB-CG-CD1	-5.63	117.62	121.00
2	B	317	ARG	CD-NE-CZ	5.63	131.48	123.60
2	B	180	TYR	CB-CG-CD1	-5.62	117.63	121.00
2	B	514	ASP	CB-CG-OD1	5.61	123.35	118.30
2	B	412	ASP	CB-CG-OD2	-5.61	113.25	118.30
2	B	468	GLU	CB-CG-CD	-5.61	99.06	114.20
1	A	145	ARG	NE-CZ-NH2	-5.57	117.52	120.30
2	B	471	ARG	NE-CZ-NH2	-5.56	117.52	120.30
2	B	245	GLN	N-CA-CB	-5.52	100.66	110.60
2	B	525	LEU	CA-CB-CG	5.46	127.85	115.30
2	B	431	TRP	CH2-CZ2-CE2	5.44	122.84	117.40
1	A	11	ARG	CD-NE-CZ	5.44	131.21	123.60
2	B	225	MET	CG-SD-CE	-5.40	91.56	100.20
2	B	380	GLN	N-CA-CB	5.39	120.30	110.60
1	A	79	ARG	NE-CZ-NH1	5.37	122.98	120.30
2	B	284	ASP	CB-CG-OD1	5.37	123.13	118.30
1	A	56	GLU	CG-CD-OE2	-5.34	107.62	118.30
2	B	248	TYR	CB-CG-CD1	5.29	124.18	121.00
2	B	273	GLN	CA-CB-CG	5.29	125.04	113.40
2	B	215	TRP	CA-CB-CG	-5.29	103.65	113.70
2	B	548	GLU	OE1-CD-OE2	5.21	129.56	123.30
1	A	14	TYR	CB-CG-CD1	5.20	124.12	121.00
1	A	71	ARG	NE-CZ-NH2	5.18	122.89	120.30
2	B	120	ARG	NE-CZ-NH1	5.16	122.88	120.30
2	B	100	LEU	CA-CB-CG	5.08	126.99	115.30
2	B	287	ARG	NE-CZ-NH1	-5.07	117.77	120.30
2	B	127	LEU	CB-CA-C	-5.03	100.64	110.20
1	A	14	TYR	CB-CG-CD2	-5.02	117.98	121.00
2	B	530	ASN	CB-CG-OD1	5.02	131.63	121.60
2	B	110	ASN	O-C-N	-5.00	114.70	123.20

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	1678	0	1627	11	0
2	B	4434	0	4259	61	0
3	A	4	0	6	6	0
3	B	40	0	59	5	0
4	B	1	0	0	0	0
5	B	24	0	18	4	0
6	A	192	0	0	1	0
6	B	509	0	0	4	0
All	All	6882	0	5969	70	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 6.

All (70) 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)
3:A:1210:EDO:H21	2:B:77:ILE:HD11	1.44	0.99
3:A:1210:EDO:H21	2:B:77:ILE:CD1	1.98	0.93
2:B:59:HIS:HD2	2:B:61:GLY:H	1.24	0.81
2:B:479:ARG:HH21	2:B:483:ASN:HD22	1.25	0.80
3:A:1210:EDO:C2	2:B:77:ILE:HD11	2.12	0.80
1:A:39:ARG:HH11	1:A:42:GLN:HE22	1.30	0.78
2:B:80:GLU:OE2	2:B:123:HIS:HD2	1.69	0.76
2:B:488:PHE:CE1	2:B:497:VAL:HG22	2.22	0.75
1:A:194:GLU:OE2	2:B:233:GLN:HG2	1.87	0.74
2:B:388:ASN:HD22	2:B:389:ILE:H	1.39	0.70
1:A:18:HIS:ND1	2:B:38:HIS:HE1	1.90	0.69
2:B:187:ASN:HA	2:B:231:ASN:HD21	1.59	0.67
2:B:452:LEU:HD12	2:B:453:THR:H	1.59	0.67
2:B:123:HIS:HE1	2:B:216:ASP:OD1	1.81	0.64
2:B:71:PHE:CE2	5:B:1569:SOX:H5	2.33	0.64
2:B:26:TRP:CD2	2:B:452:LEU:HD11	2.33	0.63
2:B:38:HIS:HD2	2:B:43:ASP:OD1	1.82	0.63
2:B:452:LEU:HD12	2:B:453:THR:N	2.14	0.63
1:A:181[A]:LEU:HD22	3:B:1564:EDO:H12	1.81	0.61
2:B:81:ARG:HH21	2:B:91:LEU:HD22	1.66	0.60
2:B:468:GLU:H	2:B:468:GLU:CD	2.06	0.59
2:B:60:ASN:ND2	2:B:63:ILE:H	2.02	0.58
2:B:88:GLY:O	2:B:98:LYS:HE3	2.04	0.58

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:44:VAL:HG11	2:B:158:MET:HB3	1.87	0.56
2:B:59:HIS:CD2	2:B:61:GLY:H	2.15	0.56
2:B:71:PHE:HE2	5:B:1569:SOX:H5	1.70	0.56
2:B:479:ARG:HH21	2:B:483:ASN:ND2	2.02	0.55
2:B:98:LYS:HD3	2:B:99:MET:O	2.06	0.55
2:B:105:THR:OG1	2:B:115:THR:HG22	2.07	0.55
2:B:226:ASN:O	3:B:1566:EDO:H22	2.07	0.54
2:B:60:ASN:HD21	2:B:63:ILE:H	1.55	0.54
2:B:71:PHE:HZ	5:B:1569:SOX:HS	1.56	0.53
1:A:183:ASN:HD22	1:A:183:ASN:C	2.11	0.53
3:B:1563:EDO:H22	6:B:2229:HOH:O	2.07	0.53
2:B:488:PHE:HE1	2:B:497:VAL:HG22	1.74	0.52
2:B:26:TRP:CE3	2:B:452:LEU:HD11	2.46	0.51
2:B:60:ASN:C	2:B:60:ASN:HD22	2.13	0.51
3:A:1210:EDO:O2	6:A:2192:HOH:O	2.00	0.50
1:A:183:ASN:HD22	1:A:185:SER:H	1.59	0.50
2:B:337:LYS:HG3	6:B:2307:HOH:O	2.11	0.50
2:B:60:ASN:HD22	2:B:62:VAL:H	1.60	0.49
2:B:369:LYS:HE2	6:B:2328:HOH:O	2.12	0.49
2:B:388:ASN:ND2	2:B:389:ILE:H	2.09	0.49
2:B:488:PHE:CD1	2:B:497:VAL:HG22	2.49	0.47
2:B:188:ILE:H	2:B:231:ASN:ND2	2.12	0.47
1:A:56:GLU:HG2	2:B:107:THR:HB	1.95	0.47
2:B:81:ARG:NH2	2:B:91:LEU:HD22	2.30	0.46
3:A:1210:EDO:H12	2:B:127:LEU:HD13	1.97	0.46
1:A:183:ASN:ND2	1:A:185:SER:H	2.14	0.45
2:B:241:ASN:ND2	5:B:1569:SOX:H6	2.32	0.45
2:B:48:THR:HB	2:B:55:LEU:HA	1.98	0.45
2:B:225:MET:O	3:B:1566:EDO:H12	2.18	0.44
2:B:48:THR:HA	2:B:49:PRO:HD3	1.88	0.44
2:B:22:PRO:HB2	2:B:24:PHE:CE2	2.53	0.44
1:A:130:GLU:HB2	1:A:131:PRO:HD2	2.00	0.44
2:B:452:LEU:HB2	2:B:476:TYR:HA	2.01	0.42
2:B:20:ASN:HD21	2:B:67:SER:H	1.67	0.42
2:B:98:LYS:HE2	2:B:99:MET:O	2.19	0.42
2:B:18:MET:O	2:B:485:MET:HA	2.19	0.42
2:B:495[A]:ARG:HA	2:B:496:PRO:HD3	1.81	0.42
2:B:274:LYS:HE3	2:B:281:GLN:HB3	2.02	0.42
2:B:363:VAL:O	2:B:368:ASP:HB3	2.20	0.41
2:B:80:GLU:OE2	2:B:123:HIS:CD2	2.61	0.41
2:B:59:HIS:HE1	6:B:2480:HOH:O	2.04	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:150:THR:HG23	3:A:1210:EDO:H22	2.03	0.40
2:B:294:LEU:H	3:B:1559:EDO:C2	2.34	0.40
2:B:435:SER:HB2	2:B:440:ASN:HD22	1.86	0.40
1:A:89:ASP:OD1	2:B:156:HIS:HE1	2.05	0.40
2:B:22:PRO:HG3	2:B:57:PHE:CZ	2.56	0.40
2:B:471:ARG:HD2	2:B:471:ARG:HA	1.89	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	206/209 (99%)	202 (98%)	4 (2%)	0	100	100
2	B	557/557 (100%)	545 (98%)	12 (2%)	0	100	100
All	All	763/766 (100%)	747 (98%)	16 (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	179/179 (100%)	173 (97%)	6 (3%)	44	26
2	B	462/460 (100%)	452 (98%)	10 (2%)	60	45

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	641/639 (100%)	625 (98%)	16 (2%)	55 39

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	4	SER
1	A	31	TYR
1	A	90	MET
1	A	183	ASN
1	A	200	LYS
1	A	204	GLN
2	B	60	ASN
2	B	98	LYS
2	B	154	TRP
2	B	173	GLN
2	B	209	VAL
2	B	274	LYS
2	B	388	ASN
2	B	403	ASP
2	B	404	LYS
2	B	436	LYS

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

Mol	Chain	Res	Type
1	A	42	GLN
1	A	108	ASN
1	A	170	GLN
1	A	183	ASN
2	B	20	ASN
2	B	38	HIS
2	B	59	HIS
2	B	60	ASN
2	B	93	ASN
2	B	110	ASN
2	B	123	HIS
2	B	187	ASN
2	B	231	ASN
2	B	245	GLN
2	B	318	GLN
2	B	388	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	401	GLN
2	B	420	GLN
2	B	440	ASN
2	B	464	GLN
2	B	472	HIS
2	B	473	GLN
2	B	483	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue 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	$\# Z > 2$
1	SME	A	16	1	6,8,9	0.92	0	6,9,11	2.09	3 (50%)

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
1	SME	A	16	1	-	0/5/7/9	0/0/0/0

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	16	SME	O-C-CA	-2.37	119.37	125.72

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	16	SME	CE-S-CG	2.26	103.07	97.59
1	A	16	SME	OE-S-CG	3.56	114.59	105.57

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 13 ligands modelled in this entry, 1 is monoatomic - leaving 12 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	$\# Z > 2$
3	EDO	A	1210	-	3,3,3	0.68	0	2,2,2	1.38	0
3	EDO	B	1558	-	3,3,3	0.59	0	2,2,2	0.31	0
3	EDO	B	1559	-	3,3,3	1.03	0	2,2,2	0.84	0
3	EDO	B	1560	-	3,3,3	0.56	0	2,2,2	0.05	0
3	EDO	B	1561	-	3,3,3	0.54	0	2,2,2	0.11	0
3	EDO	B	1562	-	3,3,3	0.47	0	2,2,2	0.45	0
3	EDO	B	1563	-	3,3,3	0.73	0	2,2,2	0.78	0
3	EDO	B	1564	-	3,3,3	0.59	0	2,2,2	0.20	0
3	EDO	B	1565	-	3,3,3	0.64	0	2,2,2	1.13	0
3	EDO	B	1566	-	3,3,3	0.69	0	2,2,2	2.11	1 (50%)
3	EDO	B	1567	-	3,3,3	0.56	0	2,2,2	0.28	0
5	SOX	B	1569	-	18,26,26	1.05	1 (5%)	26,40,40	3.59	6 (23%)

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	EDO	A	1210	-	-	0/1/1/1	0/0/0/0
3	EDO	B	1558	-	-	0/1/1/1	0/0/0/0
3	EDO	B	1559	-	-	0/1/1/1	0/0/0/0
3	EDO	B	1560	-	-	0/1/1/1	0/0/0/0
3	EDO	B	1561	-	-	0/1/1/1	0/0/0/0
3	EDO	B	1562	-	-	0/1/1/1	0/0/0/0
3	EDO	B	1563	-	-	0/1/1/1	0/0/0/0
3	EDO	B	1564	-	-	0/1/1/1	0/0/0/0
3	EDO	B	1565	-	-	0/1/1/1	0/0/0/0
3	EDO	B	1566	-	-	0/1/1/1	0/0/0/0
3	EDO	B	1567	-	-	0/1/1/1	0/0/0/0
5	SOX	B	1569	-	-	0/8/48/48	0/3/3/3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	1569	SOX	O8-C7	-2.06	1.18	1.22

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	1569	SOX	O8-C7-C6	-7.36	114.24	136.55
5	B	1569	SOX	C6-N14-C15	-4.84	112.03	121.82
5	B	1569	SOX	C7-C6-N14	-4.08	102.95	115.49
5	B	1569	SOX	C10-C2-C3	-2.55	109.55	112.75
3	B	1566	EDO	O1-C1-C2	2.40	128.68	112.23
5	B	1569	SOX	C3-N4-C7	5.01	137.93	125.89
5	B	1569	SOX	O8-C7-N4	13.74	151.69	131.62

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 15 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1210	EDO	6	0
3	B	1559	EDO	1	0
3	B	1563	EDO	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	1564	EDO	1	0
3	B	1566	EDO	2	0
5	B	1569	SOX	4	0

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	206/209 (98%)	-0.71	2 (0%) 84 82	9, 14, 25, 47	3 (1%)
2	B	557/557 (100%)	-0.61	6 (1%) 82 80	6, 14, 30, 65	1 (0%)
All	All	763/766 (99%)	-0.63	8 (1%) 84 82	6, 14, 28, 65	4 (0%)

All (8) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	132	THR	6.5
2	B	131	GLN	4.9
2	B	133	THR	3.7
1	A	3	SER	3.1
2	B	379	THR	2.7
2	B	130	ASP	2.4
1	A	4	SER	2.3
2	B	403	ASP	2.0

6.2 Non-standard residues in protein, DNA, RNA chains [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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å ²)	Q<0.9
1	SME	A	16	9/10	0.98	0.06	-	10,13,14,17	0

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(\AA^2)	Q<0.9
5	SOX	B	1569	24/24	0.85	0.17	8.45	10,36,46,54	0
3	EDO	B	1560	4/4	0.95	0.11	8.41	15,17,21,26	0
3	EDO	B	1565	4/4	0.92	0.18	5.65	20,24,26,28	0
3	EDO	B	1563	4/4	0.93	0.11	3.86	18,21,26,26	0
3	EDO	B	1559	4/4	0.94	0.12	3.45	18,22,24,26	0
3	EDO	A	1210	4/4	0.93	0.10	2.21	29,31,34,35	0
3	EDO	B	1564	4/4	0.93	0.12	2.16	15,21,23,24	0
3	EDO	B	1562	4/4	0.98	0.09	0.89	12,14,15,17	0
3	EDO	B	1566	4/4	0.93	0.10	0.67	19,22,28,30	0
3	EDO	B	1567	4/4	0.92	0.07	0.58	31,31,33,42	0
4	CA	B	1568	1/1	1.00	0.06	-0.75	10,10,10,10	0
3	EDO	B	1561	4/4	0.98	0.05	-3.10	14,21,23,25	0
3	EDO	B	1558	4/4	0.98	0.06	-	16,17,19,21	0

6.5 Other polymers [i](#)

There are no such residues in this entry.