



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 12, 2026 – 11:57 AM UTC

PDB ID : 8RPX / pdb_00008rpx
Title : NhoI restriction endonuclease in complex with quadruply methylated DNA target
Authors : Rafalski, D.; Krakowska, K.; Gilski, M.; Bochtler, M.
Deposited on : 2024-01-17
Resolution : 1.81 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
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

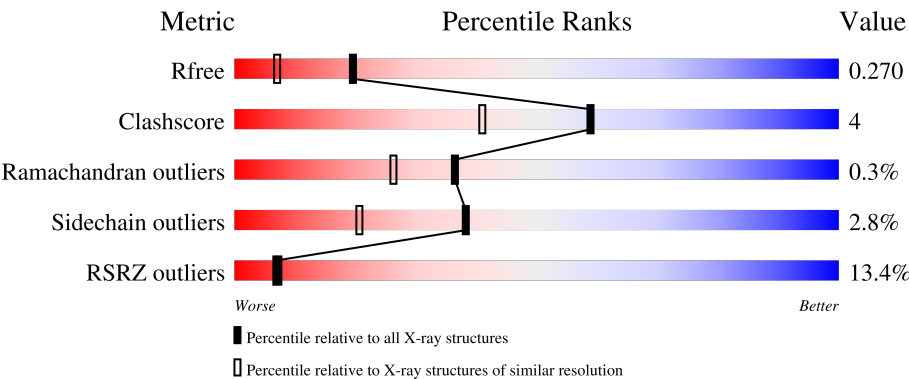
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.81 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.



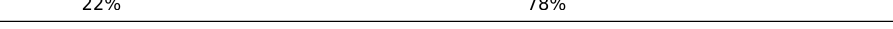
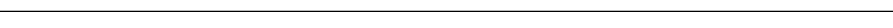
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	1112 (1.82-1.82)
Clashscore	190562	1148 (1.82-1.82)
Ramachandran outliers	187476	1140 (1.82-1.82)
Sidechain outliers	187428	1140 (1.82-1.82)
RSRZ outliers	180081	1112 (1.82-1.82)

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	171	4% 86%6%8%
1	B	171	% 85%6%9%
1	C	171	47% 80%12%8%
1	D	171	8% 91%7%..
1	E	171	5% 85%6%8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	10	 70% 30%
2	H	10	 30% 20% 80%
2	J	10	 90% 10%
3	G	9	 56% 44%
3	I	9	 22% 78%
3	K	9	 56% 44%

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 8612 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 Restriction endonuclease (NhoI).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	158	Total	C	N	O	S	0	4	0
			1271	792	244	234	1			
1	B	156	Total	C	N	O	S	0	1	0
			1238	766	233	238	1			
1	C	157	Total	C	N	O	S	0	1	0
			1178	737	226	214	1			
1	D	168	Total	C	N	O	S	0	1	0
			1299	809	241	247	2			
1	E	157	Total	C	N	O	S	0	1	0
			1241	772	237	231	1			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP I4EG67
A	0	SER	-	expression tag	UNP I4EG67
B	-1	GLY	-	expression tag	UNP I4EG67
B	0	SER	-	expression tag	UNP I4EG67
C	-1	GLY	-	expression tag	UNP I4EG67
C	0	SER	-	expression tag	UNP I4EG67
D	-1	GLY	-	expression tag	UNP I4EG67
D	0	SER	-	expression tag	UNP I4EG67
E	-1	GLY	-	expression tag	UNP I4EG67
E	0	SER	-	expression tag	UNP I4EG67

- Molecule 2 is a DNA chain called DNA (5'-D(*CP*TP*GP*(5CM)P*AP*GP*(5CM)P*TP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	9	Total	C	N	O	P	0	9	0
			361	177	61	107	16			

Continued on next page...

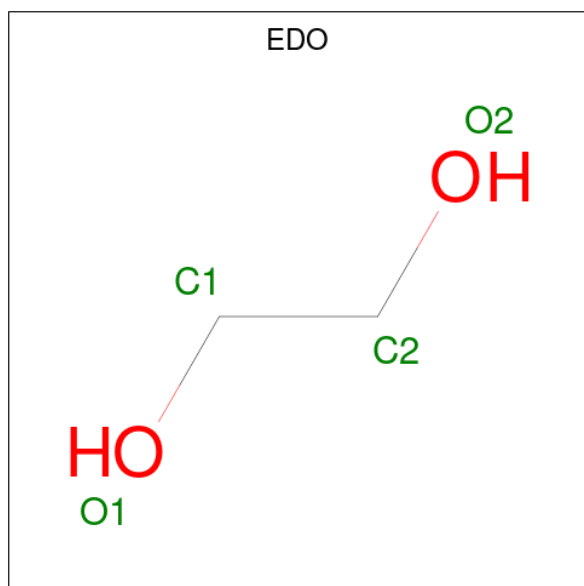
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	9	Total	C	N	O	P	0	9	0
			360	176	62	106	16			
2	J	9	Total	C	N	O	P	0	9	0
			180	88	31	53	8			

- Molecule 3 is a DNA chain called DNA(5'-D(*GP*AP*GP*(5CM)P*TP*GP*(5CM)P*AP*GP)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	9	Total	C	N	O	P	0	9	0
			374	180	76	102	16			
3	I	9	Total	C	N	O	P	0	9	0
			374	180	76	102	16			
3	K	9	Total	C	N	O	P	0	9	0
			187	90	38	51	8			

- Molecule 4 is 1,2-ETHANEDIOL (CCD ID: EDO) (formula: C₂H₆O₂).



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

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	D	1	Total C O 4 2 2	0	0
4	D	1	Total C O 4 2 2	0	0
4	D	1	Total C O 4 2 2	0	0
4	H	1	Total C O 4 2 2	0	0

- Molecule 5 is CALCIUM ION (CCD ID: CA) (formula: Ca) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total Ca 1 1	0	0
5	B	1	Total Ca 1 1	0	0
5	C	1	Total Ca 1 1	0	0
5	D	1	Total Ca 1 1	0	0
5	E	1	Total Ca 1 1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	98	Total O 98 98	0	0
6	B	111	Total O 111 111	0	0
6	C	44	Total O 44 44	0	0
6	D	78	Total O 78 78	0	0
6	E	80	Total O 80 80	0	0
6	F	21	Total O 21 21	0	0
6	G	25	Total O 25 25	0	0
6	H	14	Total O 14 14	0	0

Continued on next page...

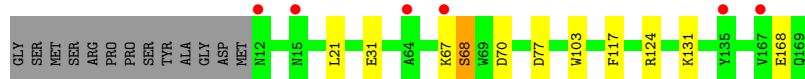
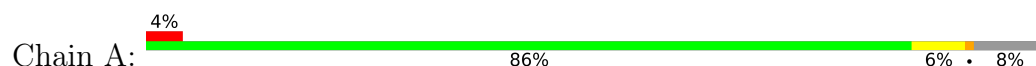
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	I	13	Total 13	O 13	0	0
6	J	14	Total 14	O 14	0	0
6	K	14	Total 14	O 14	0	0

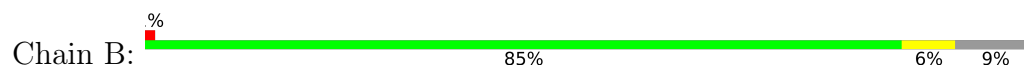
3 Residue-property plots [i](#)

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.

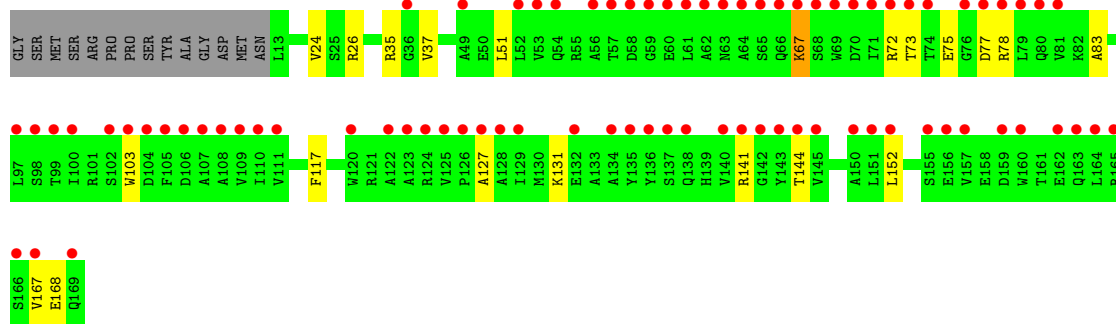
- Molecule 1: Restriction endonuclease (NhoI)



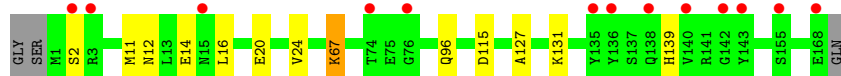
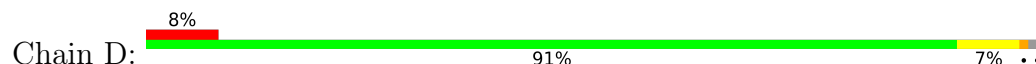
- Molecule 1: Restriction endonuclease (NhoI)



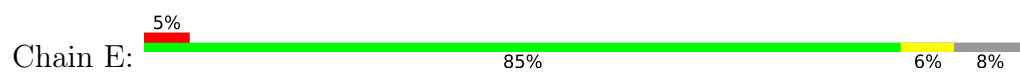
- Molecule 1: Restriction endonuclease (NhoI)



- Molecule 1: Restriction endonuclease (NhoI)



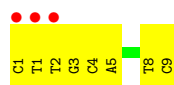
- Molecule 1: Restriction endonuclease (NhoI)



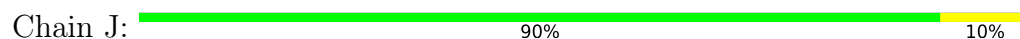
- Molecule 2: DNA (5'-D(*CP*TP*GP*(5CM)P*AP*GP*(5CM)P*TP*C)-3')



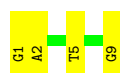
- Molecule 2: DNA (5'-D(*CP*TP*GP*(5CM)P*AP*GP*(5CM)P*TP*C)-3')



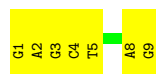
- Molecule 2: DNA (5'-D(*CP*TP*GP*(5CM)P*AP*GP*(5CM)P*TP*C)-3')



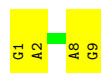
- Molecule 3: DNA(5'-D(*GP*AP*GP*(5CM)P*TP*GP*(5CM)P*AP*GP)-3')



- Molecule 3: DNA(5'-D(*GP*AP*GP*(5CM)P*TP*GP*(5CM)P*AP*GP)-3')



- Molecule 3: DNA(5'-D(*GP*AP*GP*(5CM)P*TP*GP*(5CM)P*AP*GP)-3')



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	114.64Å 167.06Å 56.51Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.81 – 1.81 46.81 – 1.81	Depositor EDS
% Data completeness (in resolution range)	99.8 (46.81-1.81) 99.9 (46.81-1.81)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.60 (at 1.81Å)	Xtriage
Refinement program	PHENIX 1.13_2998	Depositor
R, R_{free}	0.205 , 0.243 (Not available) , 0.270	Depositor DCC
R_{free} test set	4983 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	36.6	Xtriage
Anisotropy	0.411	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 57.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	8612	wwPDB-VP
Average B, all atoms (Å ²)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.33% 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: EDO, CA, 5CM

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	0.27	0/1298	0.46	0/1760
1	B	0.33	0/1257	0.49	0/1704
1	C	0.25	0/1199	0.51	0/1631
1	D	0.31	0/1325	0.49	0/1803
1	E	0.27	0/1261	0.44	0/1710
2	F	0.45	0/307	0.74	0/456
2	H	0.26	0/308	0.58	0/462
2	J	0.57	0/154	0.84	0/231
3	G	0.36	0/328	0.64	0/496
3	I	0.29	0/328	0.49	0/496
3	K	0.56	0/164	0.76	0/248
All	All	0.31	0/7929	0.52	0/10997

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1271	0	1242	7	0
1	B	1238	0	1207	9	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	1178	0	1126	12	0
1	D	1299	0	1235	17	0
1	E	1241	0	1212	5	0
2	F	361	0	202	8	0
2	H	360	0	200	15	0
2	J	180	0	101	1	0
3	G	374	0	202	7	0
3	I	374	0	203	15	0
3	K	187	0	101	0	20
4	A	4	0	6	0	0
4	B	8	0	12	0	0
4	D	16	0	24	3	0
4	H	4	0	6	0	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
5	C	1	0	0	0	0
5	D	1	0	0	0	0
5	E	1	0	0	0	0
6	A	98	0	0	1	0
6	B	111	0	0	2	0
6	C	44	0	0	1	0
6	D	78	0	0	0	0
6	E	80	0	0	1	0
6	F	21	0	0	0	0
6	G	25	0	0	1	0
6	H	14	0	0	0	0
6	I	13	0	0	0	0
6	J	14	0	0	0	0
6	K	14	0	0	0	0
All	All	8612	0	7079	63	20

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 4.

All (63) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:67:LYS:H	4:D:201:EDO:H21	1.41	0.83
2:F:9[B]:DC:N3	3:G:1[B]:DG:N1	2.29	0.76
1:C:141[A]:ARG:HH21	1:D:139:HIS:HA	1.51	0.76
2:F:9[B]:DC:N4	3:G:1[B]:DG:O6	2.22	0.72

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:70:ASP:OD1	6:B:301:HOH:O	2.10	0.70
1:C:73:THR:HG23	1:C:75:GLU:H	1.57	0.68
1:E:70:ASP:OD1	6:E:301:HOH:O	2.11	0.68
2:H:1[B]:DC:N4	3:I:9[B]:DG:O6	2.27	0.68
2:F:9[B]:DC:H5	6:G:118:HOH:O	1.77	0.67
1:A:70:ASP:OD1	6:A:301:HOH:O	2.14	0.65
2:H:8[B]:DT:N3	3:I:2[B]:DA:N1	2.47	0.62
2:H:2[A]:DT:N3	3:I:8[A]:DA:N1	2.50	0.60
1:D:139:HIS:NE2	3:I:4[B]:5CM:OP2	2.28	0.59
2:H:9[B]:DC:N4	3:I:1[B]:DG:O6	2.36	0.59
1:D:96:GLN:OE1	2:H:8[B]:DT:O4	2.23	0.57
1:C:83:ALA:O	6:C:501:HOH:O	2.18	0.56
1:D:67:LYS:HD2	4:D:201:EDO:H12	1.88	0.56
2:H:9[B]:DC:N3	3:I:1[B]:DG:N1	2.41	0.54
3:G:9[A]:DG:H2''	3:I:9[A]:DG:H2'	1.90	0.53
1:A:21:LEU:HD11	1:B:31:GLU:HG3	1.92	0.51
2:F:5[B]:DA:N1	3:G:5[B]:DT:N3	2.58	0.51
1:C:72:ARG:HA	1:C:77:ASP:O	2.11	0.51
1:D:67:LYS:N	4:D:201:EDO:H21	2.19	0.51
1:E:75:GLU:OE1	1:E:75:GLU:N	2.33	0.50
2:F:8[B]:DT:N3	3:G:2[B]:DA:N1	2.59	0.50
2:H:1[A]:DC:N3	3:I:9[A]:DG:N1	2.49	0.50
1:D:20:GLU:O	1:D:24:VAL:HG22	2.12	0.50
1:A:103:TRP:HB3	1:A:131:LYS:HE2	1.93	0.49
1:E:21:LEU:O	1:E:24:VAL:HG12	2.13	0.49
1:B:92:ASN:CG	3:I:2[B]:DA:H4'	2.37	0.49
1:B:31:GLU:OE2	1:B:34:ARG:NH2	2.35	0.49
1:B:92:ASN:CG	2:H:2[A]:DT:H4'	2.37	0.49
1:A:31:GLU:HG3	1:B:21:LEU:HD11	1.95	0.48
1:C:35:ARG:NH2	1:D:16:LEU:O	2.42	0.48
1:D:139:HIS:NE2	2:H:4[A]:5CM:OP2	2.29	0.48
1:A:77:ASP:OD2	1:A:124:ARG:NH2	2.45	0.47
1:C:103:TRP:HB3	1:C:131:LYS:HE2	1.96	0.47
2:F:9[B]:DC:H2''	2:H:9[B]:DC:H2'	1.95	0.47
2:H:5[B]:DA:N1	3:I:5[B]:DT:N3	2.63	0.47
1:C:26:ARG:HG2	1:D:115:ASP:O	2.15	0.47
2:H:2[A]:DT:H2'	2:H:3[A]:DG:O4'	2.16	0.46
1:D:67:LYS:NZ	3:I:3[A]:DG:OP1	2.49	0.46
1:E:103:TRP:HB3	1:E:131:LYS:HE2	1.98	0.45
1:D:67:LYS:NZ	2:H:3[B]:DG:OP1	2.49	0.45
2:F:5[A]:DA:N1	3:G:5[A]:DT:N3	2.66	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:I:2[B]:DA:H2'	3:I:3[B]:DG:O4'	2.17	0.44
1:C:127:ALA:O	1:C:131:LYS:HG3	2.17	0.44
1:C:141[A]:ARG:NH2	1:D:139:HIS:HA	2.27	0.43
1:E:96:GLN:OE1	2:J:8[A]:DT:O4	2.35	0.43
1:D:127:ALA:O	1:D:131:LYS:HG3	2.19	0.43
1:A:21:LEU:HD11	1:B:31:GLU:CG	2.49	0.42
1:C:37:VAL:HA	1:C:51:LEU:HD22	2.00	0.42
1:C:73:THR:HG23	1:C:75:GLU:N	2.30	0.42
1:D:12:ASN:HD21	1:D:14:GLU:HG2	1.85	0.42
2:H:1[A]:DC:N4	3:I:9[A]:DG:O6	2.52	0.42
2:F:8[A]:DT:N3	3:G:2[A]:DA:N1	2.68	0.42
1:B:122:ALA:CB	1:B:164:LEU:HD22	2.50	0.41
1:B:13:LEU:N	6:B:312:HOH:O	2.52	0.41
2:H:2[B]:DT:N3	3:I:8[B]:DA:N1	2.69	0.41
1:A:67:LYS:O	1:A:68:SER:HB2	2.21	0.41
1:C:67:LYS:HD3	3:I:3[B]:DG:OP1	2.21	0.41
1:D:11:MET:HB2	1:D:11:MET:HE3	1.78	0.40
1:D:67:LYS:HB2	1:D:67:LYS:HE3	1.73	0.40

All (20) 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)
3:K:1[B]:DG:N2	3:K:9[B]:DG:N2[2_545]	0.31	1.89
3:K:2[B]:DA:N1	3:K:8[B]:DA:N1[2_545]	0.75	1.45
3:K:1[B]:DG:N1	3:K:9[B]:DG:N1[2_545]	0.79	1.41
3:K:1[B]:DG:N2	3:K:9[B]:DG:C2[2_545]	1.10	1.10
3:K:1[B]:DG:O6	3:K:9[B]:DG:O6[2_545]	1.30	0.90
3:K:2[B]:DA:N6	3:K:8[B]:DA:N6[2_545]	1.46	0.74
3:K:1[B]:DG:C2	3:K:9[B]:DG:N2[2_545]	1.51	0.69
3:K:1[B]:DG:C2	3:K:9[B]:DG:N1[2_545]	1.60	0.60
3:K:2[B]:DA:C2	3:K:8[B]:DA:N1[2_545]	1.67	0.53
3:K:1[B]:DG:C2	3:K:9[B]:DG:C2[2_545]	1.75	0.45
3:K:2[B]:DA:C2	3:K:8[B]:DA:C2[2_545]	1.76	0.44
3:K:1[B]:DG:N1	3:K:9[B]:DG:C6[2_545]	1.77	0.43
3:K:2[B]:DA:N1	3:K:8[B]:DA:C6[2_545]	1.83	0.37
3:K:2[B]:DA:N1	3:K:8[B]:DA:C2[2_545]	1.85	0.35
3:K:1[B]:DG:C6	3:K:9[B]:DG:O6[2_545]	1.97	0.23
3:K:1[B]:DG:N1	3:K:9[B]:DG:C2[2_545]	1.98	0.22
3:K:2[B]:DA:C6	3:K:8[B]:DA:N1[2_545]	1.99	0.21

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:K:1[B]:DG:N2	3:K:9[B]:DG:N1[2_545]	2.05	0.15
3:K:1[B]:DG:C6	3:K:9[B]:DG:N1[2_545]	2.10	0.10
3:K:1[B]:DG:N2	3:K:9[B]:DG:N3[2_545]	2.16	0.04

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	160/171 (94%)	159 (99%)	0	1 (1%)	21	11
1	B	155/171 (91%)	154 (99%)	1 (1%)	0	100	100
1	C	156/171 (91%)	153 (98%)	3 (2%)	0	100	100
1	D	167/171 (98%)	164 (98%)	2 (1%)	1 (1%)	21	11
1	E	156/171 (91%)	154 (99%)	2 (1%)	0	100	100
All	All	794/855 (93%)	784 (99%)	8 (1%)	2 (0%)	36	26

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	2	SER
1	A	68	SER

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	126/141 (89%)	124 (98%)	2 (2%)	55	44
1	B	126/141 (89%)	124 (98%)	2 (2%)	55	44
1	C	108/141 (77%)	100 (93%)	8 (7%)	13	1
1	D	126/141 (89%)	125 (99%)	1 (1%)	73	67
1	E	123/141 (87%)	119 (97%)	4 (3%)	33	15
All	All	609/705 (86%)	592 (97%)	17 (3%)	38	21

All (17) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	117	PHE
1	A	168	GLU
1	B	117	PHE
1	B	138	GLN
1	C	24	VAL
1	C	67	LYS
1	C	78	ARG
1	C	117	PHE
1	C	144	THR
1	C	152	LEU
1	C	167	VAL
1	C	168	GLU
1	D	67	LYS
1	E	16	LEU
1	E	61	LEU
1	E	117	PHE
1	E	169	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	54	GLN
1	E	96	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

20 non-standard protein/DNA/RNA residues 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	5CM	J	7[A]	2	18,21,22	0.25	0	24,30,33	0.50	0
2	5CM	F	4[B]	2	18,21,22	0.27	0	24,30,33	0.41	0
2	5CM	F	4[A]	2	18,21,22	0.22	0	24,30,33	0.38	0
3	5CM	I	4[B]	3	18,21,22	0.32	0	24,30,33	0.49	0
3	5CM	I	4[A]	3	18,21,22	0.28	0	24,30,33	0.40	0
2	5CM	F	7[B]	2	18,21,22	0.30	0	24,30,33	0.50	0
3	5CM	G	7[B]	3,2	18,21,22	0.29	0	24,30,33	0.57	0
3	5CM	G	7[A]	3,2	18,21,22	0.34	0	24,30,33	0.51	0
2	5CM	H	7[A]	2	18,21,22	0.24	0	24,30,33	0.51	0
2	5CM	H	7[B]	2	18,21,22	0.36	0	24,30,33	0.50	0
3	5CM	I	7[A]	3	18,21,22	0.33	0	24,30,33	0.50	0
3	5CM	I	7[B]	3	18,21,22	0.28	0	24,30,33	0.50	0
3	5CM	K	4[B]	3	18,21,22	0.24	0	24,30,33	0.44	0
3	5CM	G	4[A]	3	18,21,22	0.26	0	24,30,33	0.42	0
3	5CM	G	4[B]	3	18,21,22	0.28	0	24,30,33	0.42	0
3	5CM	K	7[B]	3	18,21,22	0.24	0	24,30,33	0.48	0
2	5CM	J	4[A]	2	18,21,22	0.23	0	24,30,33	0.44	0
2	5CM	H	4[B]	2	18,21,22	0.32	0	24,30,33	0.42	0
2	5CM	H	4[A]	2	18,21,22	0.36	0	24,30,33	0.47	0
2	5CM	F	7[A]	2	18,21,22	0.29	0	24,30,33	0.56	0

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	5CM	J	7[A]	2	-	0/7/21/22	0/2/2/2
2	5CM	F	4[B]	2	-	0/7/21/22	0/2/2/2
2	5CM	F	4[A]	2	-	0/7/21/22	0/2/2/2
3	5CM	I	4[B]	3	-	0/7/21/22	0/2/2/2
3	5CM	I	4[A]	3	-	0/7/21/22	0/2/2/2

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	5CM	F	7[B]	2	-	0/7/21/22	0/2/2/2
3	5CM	G	7[B]	3,2	-	0/7/21/22	0/2/2/2
3	5CM	G	7[A]	3,2	-	0/7/21/22	0/2/2/2
2	5CM	H	7[A]	2	-	0/7/21/22	0/2/2/2
2	5CM	H	7[B]	2	-	0/7/21/22	0/2/2/2
3	5CM	I	7[A]	3	-	0/7/21/22	0/2/2/2
3	5CM	I	7[B]	3	-	0/7/21/22	0/2/2/2
3	5CM	K	4[B]	3	-	0/7/21/22	0/2/2/2
3	5CM	G	4[A]	3	-	0/7/21/22	0/2/2/2
3	5CM	G	4[B]	3	-	0/7/21/22	0/2/2/2
3	5CM	K	7[B]	3	-	0/7/21/22	0/2/2/2
2	5CM	J	4[A]	2	-	0/7/21/22	0/2/2/2
2	5CM	H	4[B]	2	-	0/7/21/22	0/2/2/2
2	5CM	H	4[A]	2	-	0/7/21/22	0/2/2/2
2	5CM	F	7[A]	2	-	0/7/21/22	0/2/2/2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	I	4[B]	5CM	1	0
2	H	4[A]	5CM	1	0

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 13 ligands modelled in this entry, 5 are monoatomic - leaving 8 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	EDO	D	201	-	3,3,3	0.48	0	2,2,2	0.34	0
4	EDO	B	201	-	3,3,3	0.47	0	2,2,2	0.38	0
4	EDO	A	201	-	3,3,3	0.45	0	2,2,2	0.30	0
4	EDO	H	101	-	3,3,3	0.57	0	2,2,2	0.53	0
4	EDO	D	203	-	3,3,3	0.38	0	2,2,2	0.31	0
4	EDO	B	202	-	3,3,3	0.41	0	2,2,2	0.52	0
4	EDO	D	202	-	3,3,3	0.53	0	2,2,2	0.33	0
4	EDO	D	204	-	3,3,3	0.54	0	2,2,2	0.30	0

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	EDO	D	201	-	-	0/1/1/1	-
4	EDO	B	201	-	-	0/1/1/1	-
4	EDO	A	201	-	-	0/1/1/1	-
4	EDO	H	101	-	-	0/1/1/1	-
4	EDO	D	203	-	-	1/1/1/1	-
4	EDO	B	202	-	-	0/1/1/1	-
4	EDO	D	202	-	-	0/1/1/1	-
4	EDO	D	204	-	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	D	204	EDO	O1-C1-C2-O2
4	D	203	EDO	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	D	201	EDO	3	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

6.1 Protein, DNA and RNA chains

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	158/171 (92%)	0.46	6 (3%) 44 50	27, 48, 71, 101	4 (2%)
1	B	156/171 (91%)	0.20	2 (1%) 75 80	17, 40, 64, 89	1 (0%)
1	C	157/171 (91%)	2.10	80 (50%) 0 0	38, 81, 115, 137	1 (0%)
1	D	168/171 (98%)	0.60	13 (7%) 19 21	28, 47, 70, 104	1 (0%)
1	E	157/171 (91%)	0.58	9 (5%) 29 33	26, 49, 80, 104	1 (0%)
2	F	6/10 (60%)	-0.38	0 100 100	14, 18, 21, 22	6 (100%)
2	H	8/10 (80%)	1.39	3 (37%) 1 0	24, 28, 37, 37	8 (100%)
2	J	8/10 (80%)	0.72	0 100 100	38, 47, 60, 62	8 (100%)
3	G	7/9 (77%)	-0.30	0 100 100	16, 18, 22, 25	7 (100%)
3	I	7/9 (77%)	0.67	0 100 100	25, 28, 33, 35	7 (100%)
3	K	7/9 (77%)	0.75	0 100 100	23, 40, 53, 54	6 (85%)
All	All	839/912 (91%)	0.77	113 (13%) 7 7	14, 50, 98, 137	50 (5%)

All (113) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	79	LEU	6.3
1	C	69	TRP	5.6
1	C	104	ASP	5.3
1	C	71	ILE	4.6
1	D	135	TYR	4.4
1	C	136	TYR	4.3
1	C	56	ALA	4.3
1	C	129	ILE	4.2
1	C	107	ALA	4.2
1	C	122	ALA	4.2
1	C	103	TRP	4.1
1	C	160	TRP	4.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	78	ARG	3.9
1	C	70	ASP	3.9
1	C	123	ALA	3.9
1	C	128	ALA	3.9
1	C	64	ALA	3.8
1	A	12	ASN	3.8
1	C	143	TYR	3.7
1	C	109	VAL	3.7
1	C	66	GLN	3.6
1	C	144	THR	3.6
1	A	135[A]	TYR	3.6
1	C	156	GLU	3.6
1	C	53	VAL	3.6
1	D	2	SER	3.6
1	C	105	PHE	3.5
1	C	68	SER	3.5
1	A	64	ALA	3.5
1	C	141[A]	ARG	3.4
1	D	136	TYR	3.4
1	D	143	TYR	3.3
1	C	57	THR	3.3
1	C	106	ASP	3.3
1	C	65	SER	3.3
1	C	140	VAL	3.3
1	C	81	VAL	3.2
1	C	62	ALA	3.2
1	C	108	ALA	3.2
1	C	150	ALA	3.2
1	C	167	VAL	3.2
1	C	61	LEU	3.2
1	C	73	THR	3.1
1	C	151	LEU	3.0
1	C	102	SER	3.0
1	E	74	THR	3.0
1	C	52	LEU	3.0
1	C	135	TYR	3.0
2	H	2[A]	DT	3.0
1	C	137	SER	3.0
1	C	76	GLY	2.9
1	B	15	ASN	2.9
1	E	13	LEU	2.9
1	C	60	GLU	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	145	VAL	2.9
1	C	127	ALA	2.9
1	C	164	LEU	2.8
1	D	138	GLN	2.8
1	C	125	VAL	2.8
1	C	157	VAL	2.8
1	C	54	GLN	2.7
1	C	72	ARG	2.7
1	C	67	LYS	2.7
1	D	74	THR	2.7
1	C	100	ILE	2.7
1	C	166	SER	2.6
1	C	132	GLU	2.6
1	E	15	ASN	2.6
1	C	163	GLN	2.6
1	C	74	THR	2.6
1	C	155	SER	2.6
1	C	165	ARG	2.6
1	E	16	LEU	2.6
1	C	138	GLN	2.6
1	C	80	GLN	2.5
1	C	159	ASP	2.5
1	C	59	GLY	2.5
1	C	97	LEU	2.5
1	C	134	ALA	2.5
1	E	135	TYR	2.5
1	C	36	GLY	2.4
1	C	169	GLN	2.4
1	C	126	PRO	2.4
1	C	162	GLU	2.4
1	C	63	ASN	2.4
1	D	15	ASN	2.4
2	H	1[A]	DC	2.4
2	H	1[A]	DC	2.4
1	C	152	LEU	2.3
1	E	24	VAL	2.3
1	C	124	ARG	2.3
1	A	15	ASN	2.3
1	C	77	ASP	2.3
1	C	111	VAL	2.2
1	A	67	LYS	2.2
1	D	168	GLU	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	49	ALA	2.2
1	D	3	ARG	2.2
1	C	98	SER	2.2
1	C	142	GLY	2.2
1	D	142	GLY	2.2
1	E	17	THR	2.1
1	C	120	TRP	2.1
1	A	167	VAL	2.1
1	D	140	VAL	2.1
1	E	21	LEU	2.1
1	C	58	ASP	2.1
1	D	155	SER	2.1
1	D	76	GLY	2.1
1	C	110	ILE	2.0
1	C	99	THR	2.0
1	E	14	GLU	2.0
1	B	74	THR	2.0

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

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	5CM	H	4[A]	20/21	0.90	0.14	56,65,73,74	20
2	5CM	H	4[B]	20/21	0.90	0.14	39,44,48,51	20
3	5CM	K	7[B]	20/21	0.94	0.10	52,59,63,66	20
2	5CM	J	7[A]	20/21	0.95	0.10	69,70,72,76	20
3	5CM	I	4[A]	20/21	0.95	0.09	43,47,56,56	20
3	5CM	I	4[B]	20/21	0.95	0.09	55,60,70,77	20
3	5CM	K	4[B]	20/21	0.95	0.10	47,49,53,54	20
2	5CM	J	4[A]	20/21	0.95	0.10	59,60,62,65	20
2	5CM	H	7[B]	20/21	0.96	0.09	42,48,53,55	20
2	5CM	H	7[A]	20/21	0.96	0.09	39,44,52,52	20
3	5CM	I	7[B]	20/21	0.97	0.09	41,47,50,50	20
3	5CM	I	7[A]	20/21	0.97	0.09	46,51,54,55	20
2	5CM	F	7[B]	20/21	0.98	0.06	25,28,31,33	20
2	5CM	F	4[B]	20/21	0.98	0.06	27,31,33,35	20
2	5CM	F	4[A]	20/21	0.98	0.06	26,33,38,39	20
3	5CM	G	7[A]	20/21	0.98	0.05	30,31,33,34	20

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
3	5CM	G	7[B]	20/21	0.98	0.05	32,33,36,36	20
2	5CM	F	7[A]	20/21	0.98	0.06	27,29,32,32	20
3	5CM	G	4[A]	20/21	0.98	0.07	31,32,36,37	20
3	5CM	G	4[B]	20/21	0.98	0.07	31,33,41,41	20

6.3 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

6.4 Ligands [i](#)

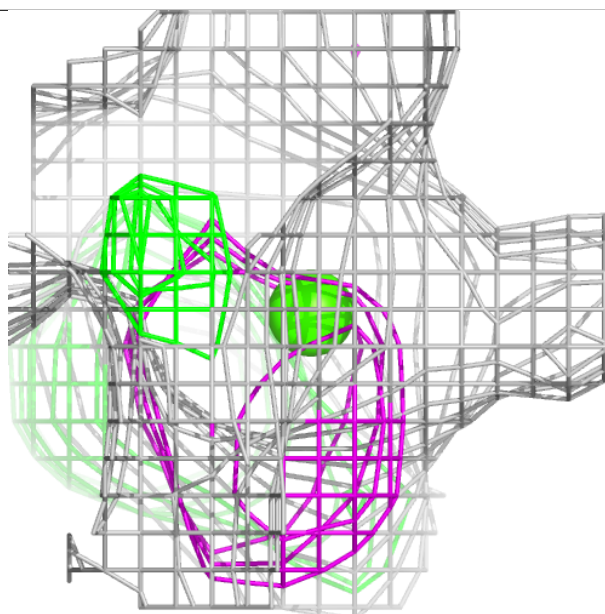
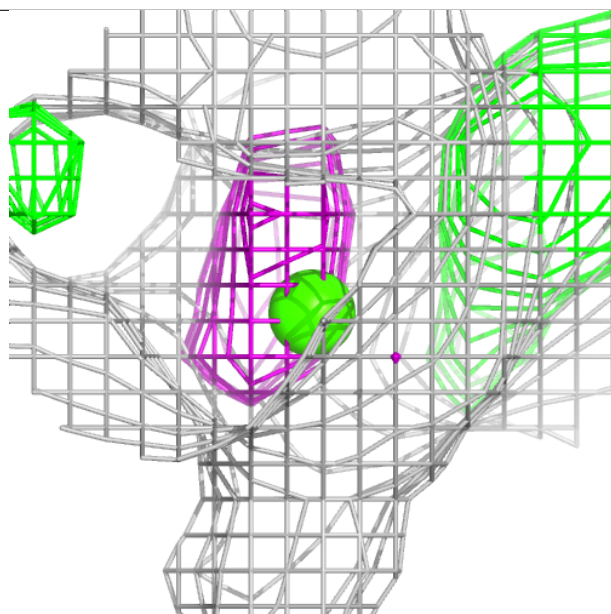
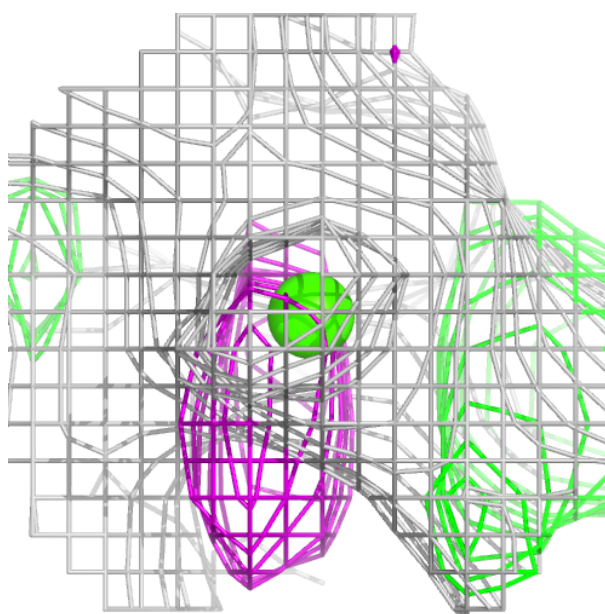
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists 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
4	EDO	A	201	4/4	0.73	0.18	58,67,72,77	0
4	EDO	H	101	4/4	0.75	0.15	41,42,51,51	0
4	EDO	B	202	4/4	0.80	0.24	67,69,71,82	0
4	EDO	D	202	4/4	0.82	0.19	55,57,60,65	0
4	EDO	D	201	4/4	0.82	0.15	51,59,61,66	0
4	EDO	D	204	4/4	0.85	0.16	49,49,63,63	0
4	EDO	B	201	4/4	0.88	0.15	63,69,70,71	0
4	EDO	D	203	4/4	0.89	0.12	50,51,55,62	0
5	CA	C	201	1/1	0.92	0.13	52,52,52,52	0
5	CA	D	205	1/1	0.98	0.04	34,34,34,34	0
5	CA	E	201	1/1	0.98	0.03	31,31,31,31	0
5	CA	B	203	1/1	0.99	0.02	27,27,27,27	0
5	CA	A	202	1/1	0.99	0.03	32,32,32,32	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.

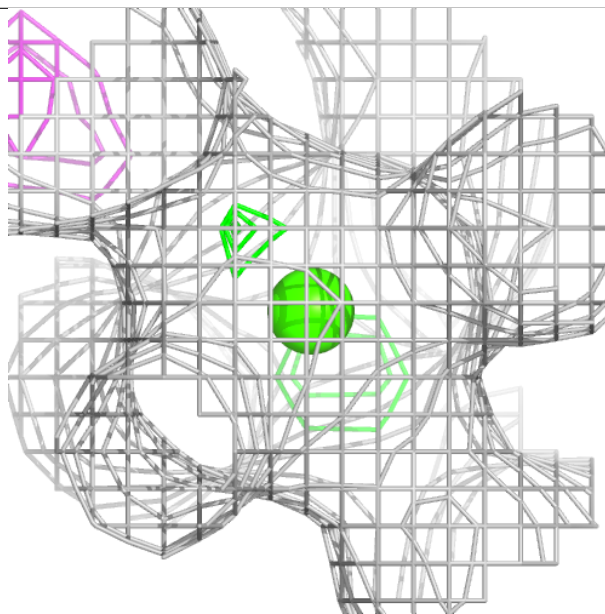
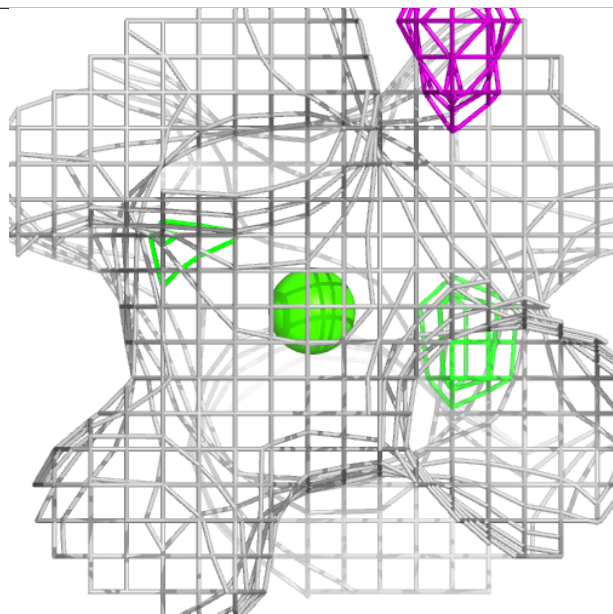
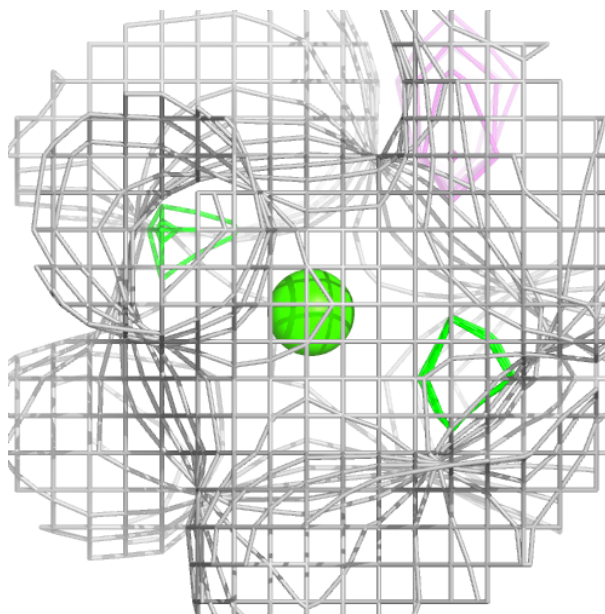
Electron density around CA C 201:

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



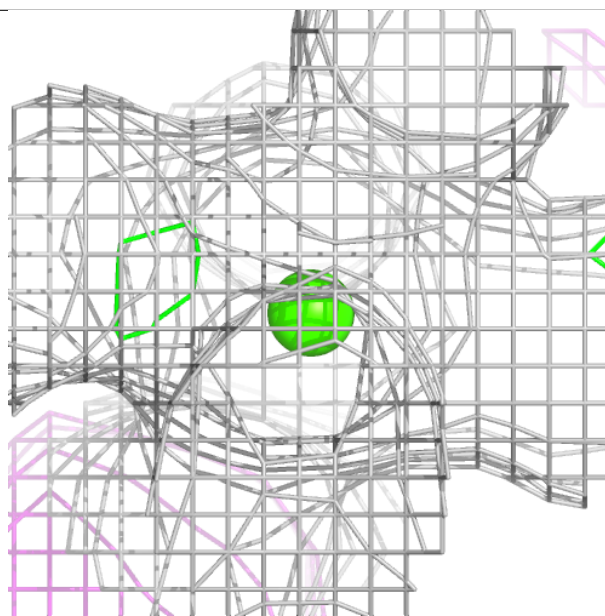
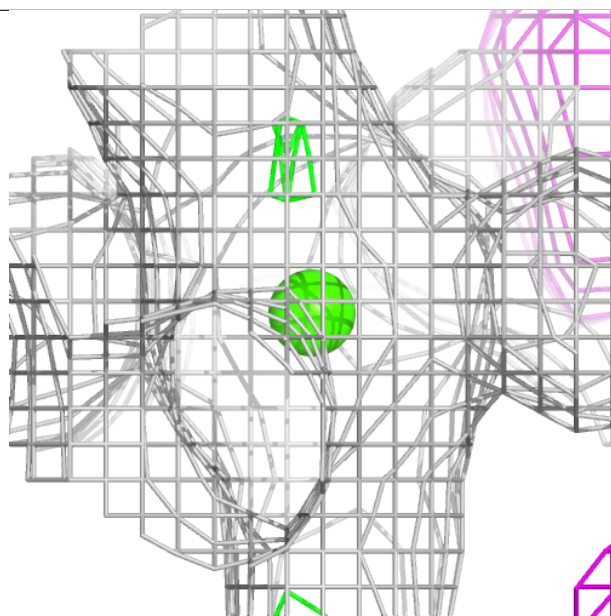
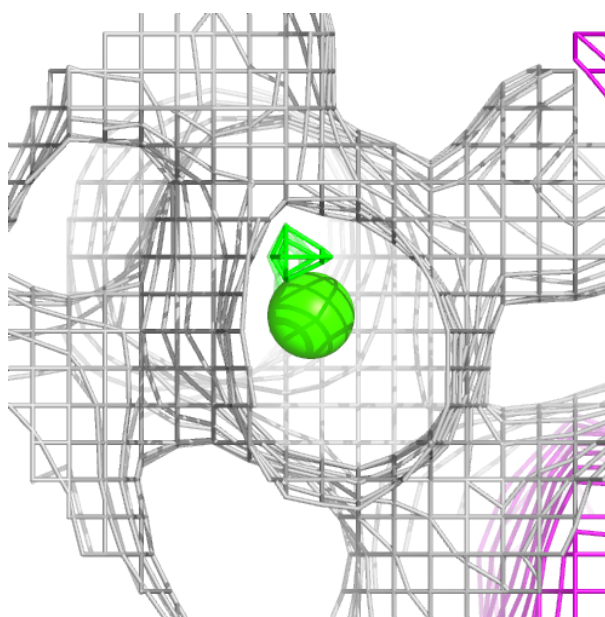
Electron density around CA D 205:

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



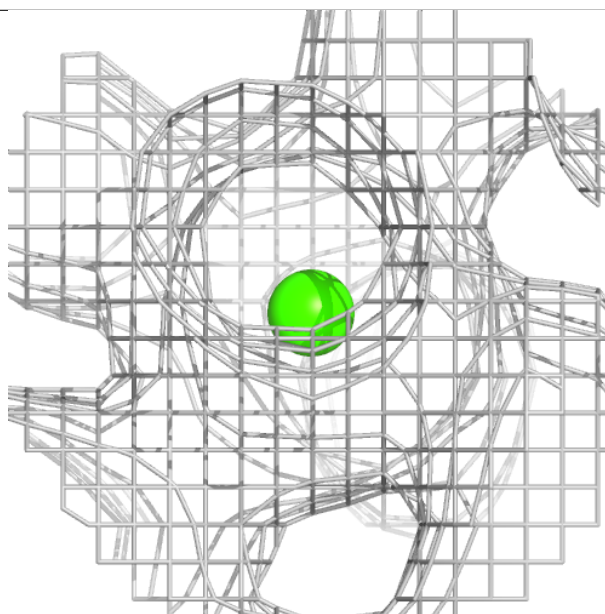
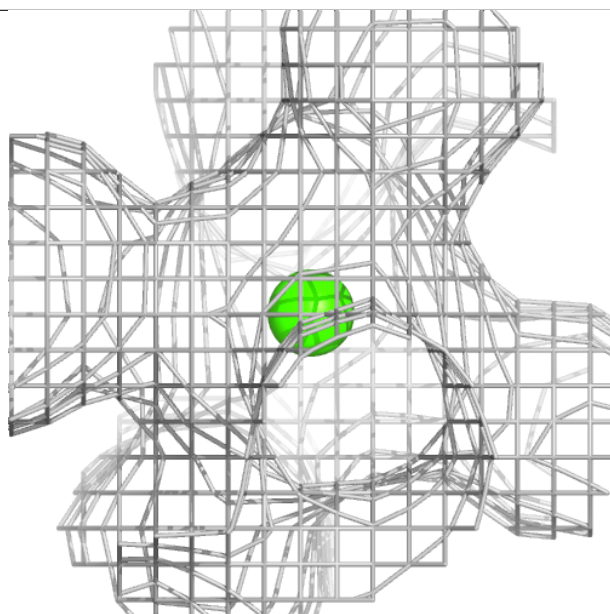
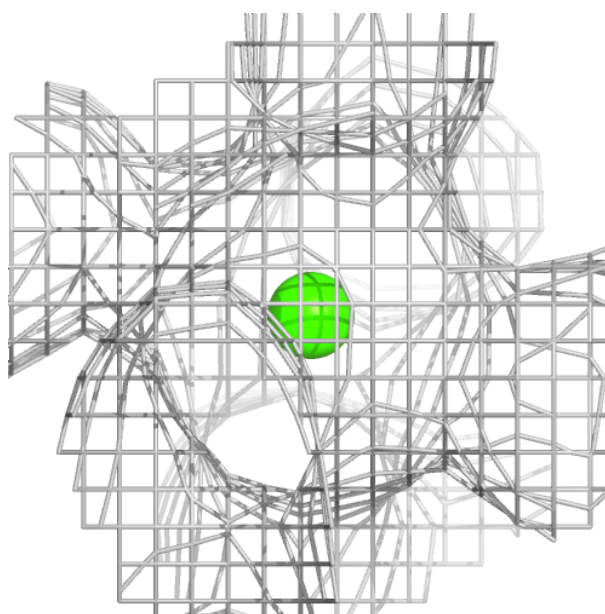
Electron density around CA E 201:

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



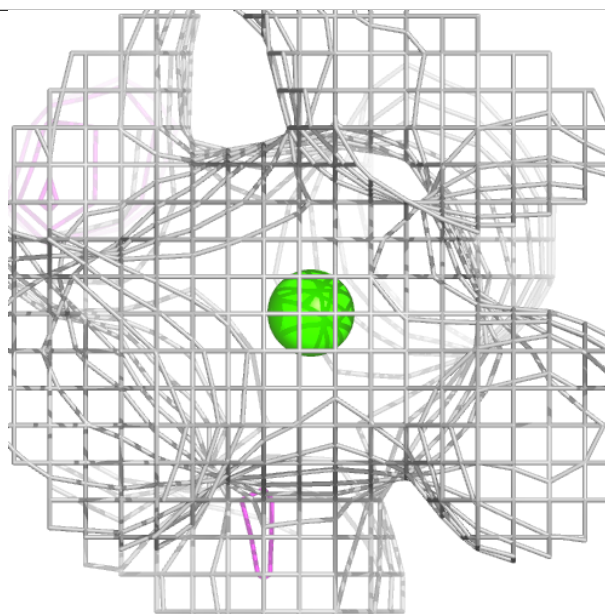
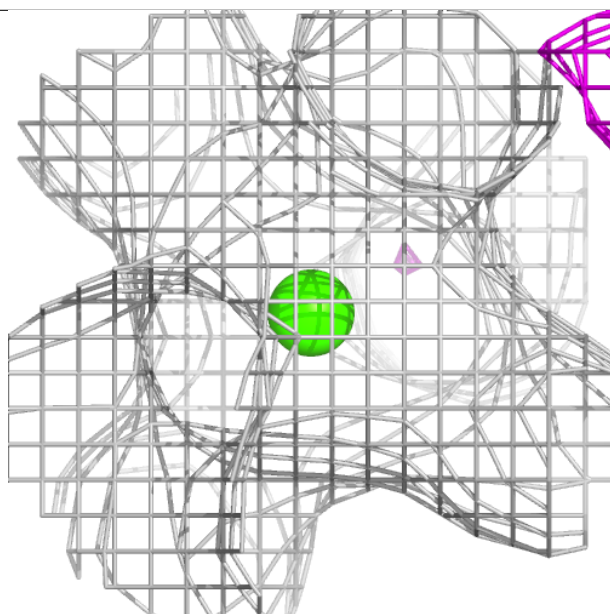
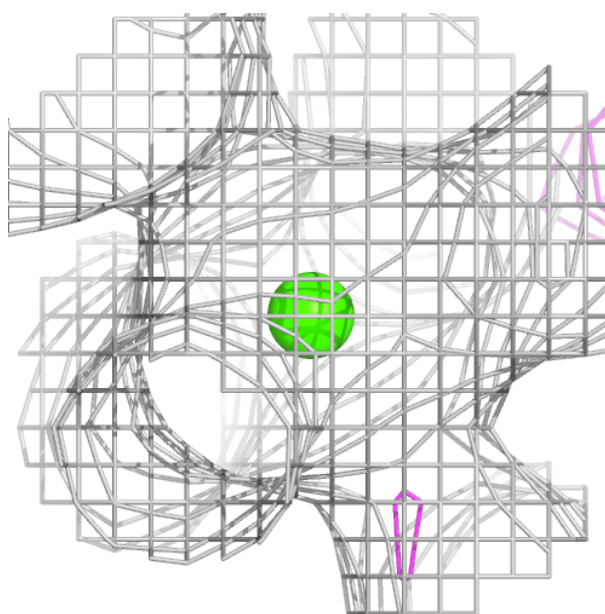
Electron density around CA B 203:

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



Electron density around CA A 202:

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



6.5 Other polymers ⓘ

There are no such residues in this entry.