WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jun 17, 2024 – 07:47 PM EDT

PDB ID	:	3M2V
Title	:	Structural Insight into Methyl-Coenzyme M Reductase Chemistry using Coenzyme B Analogues
Authors	:	Cedervall, P.E.; Dey, M.; Ragsdale, S.W.; Wilmot, C.M.
Deposited on	:	2010-03-08
Resolution	:	1.80 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

|		
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
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)	:	2.37.1

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 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.
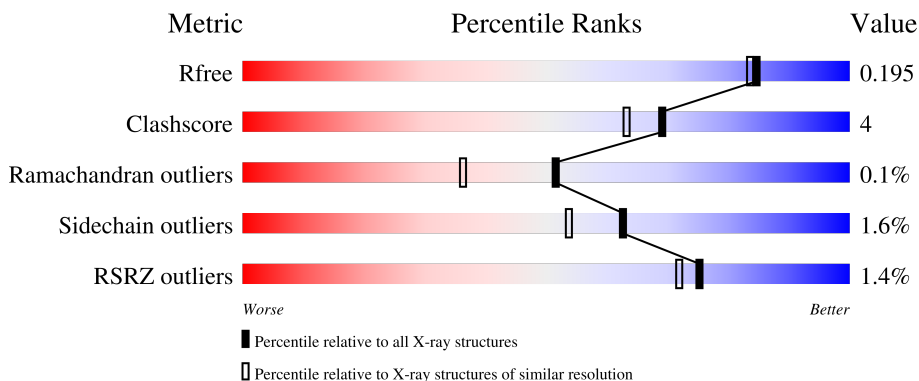
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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 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 <=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	549	91% 8% •
1	D	549	% 89% 11% •
2	B	442	89% 9% •
2	E	442	2% 91% 9%
3	C	248	3% 87% 12% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	248	6% ▬ 87% 11% ·

2 Entry composition ⓘ

There are 12 unique types of molecules in this entry. The entry contains 21893 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 Methyl-coenzyme M reductase I subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	548	Total	C	N	O	S	0	23	0
			4408	2785	731	871	21			
1	D	548	Total	C	N	O	S	0	19	0
			4359	2767	721	851	20			

- Molecule 2 is a protein called Methyl-coenzyme M reductase I subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	442	Total	C	N	O	S	0	18	0
			3422	2177	559	664	22			
2	E	442	Total	C	N	O	S	0	19	0
			3431	2182	563	665	21			

- Molecule 3 is a protein called Methyl-coenzyme M reductase I subunit gamma.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	248	Total	C	N	O	S	0	8	0
			2059	1273	362	412	12			
3	F	247	Total	C	N	O	S	0	11	0
			2071	1280	367	412	12			

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Mg	0	0
			1	1		
4	B	2	Total	Mg	0	0
			2	2		
4	C	1	Total	Mg	0	0
			1	1		
4	F	1	Total	Mg	0	0
			1	1		

- Molecule 5 is FACTOR 430 (three-letter code: F43) (formula: $C_{42}H_{51}N_6NiO_{13}$).
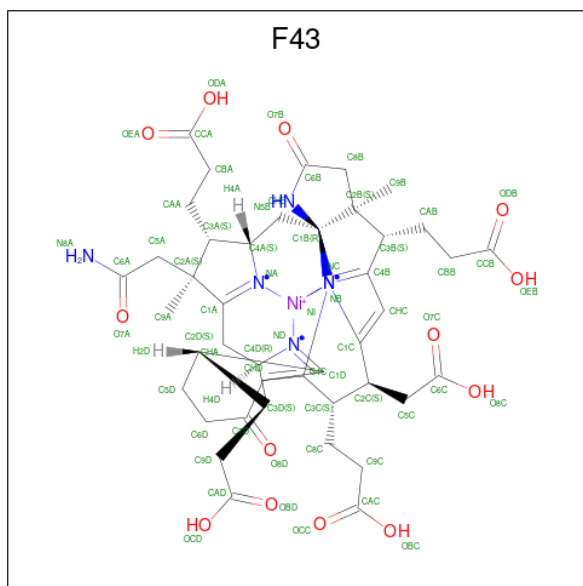
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	Ni	O	0	0
			62	42	6	1	13		
5	D	1	Total	C	N	Ni	O	0	0
			62	42	6	1	13		

- Molecule 6 is Coenzyme B (three-letter code: TP7) (formula: $C_{11}H_{22}NO_7PS$).
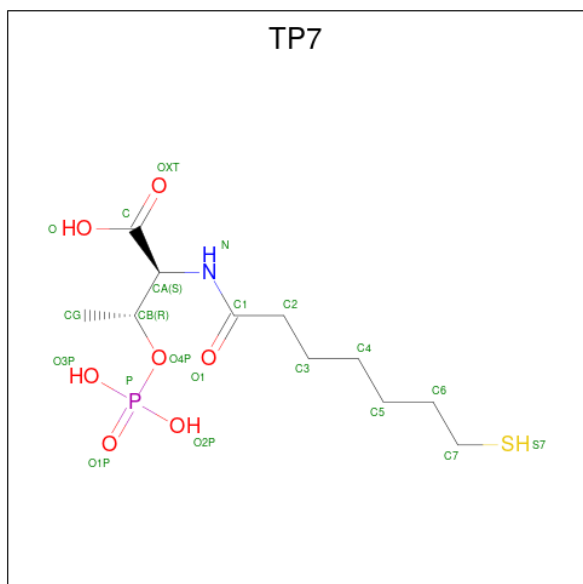
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
6	A	1	Total	C	N	O	P	S	0	1
			21	11	1	7	1	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
6	D	1	Total 21	C 11	N 1	O 7	P 1	S 1	0	1

- Molecule 7 is 1-THIOETHANESULFONIC ACID (three-letter code: COM) (formula: $C_2H_6O_3S_2$).
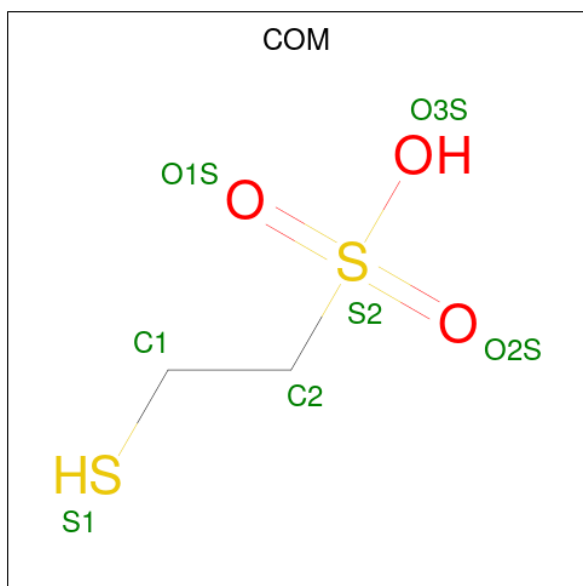
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	A	1	Total 7	C 2	O 3	S 2	0	0
7	D	1	Total 7	C 2	O 3	S 2	0	0

- Molecule 8 is O-phosphono-N-(8-sulfanyloctanoyl)-L-threonine (three-letter code: XP8) (formula: $C_{12}H_{24}NO_7PS$).
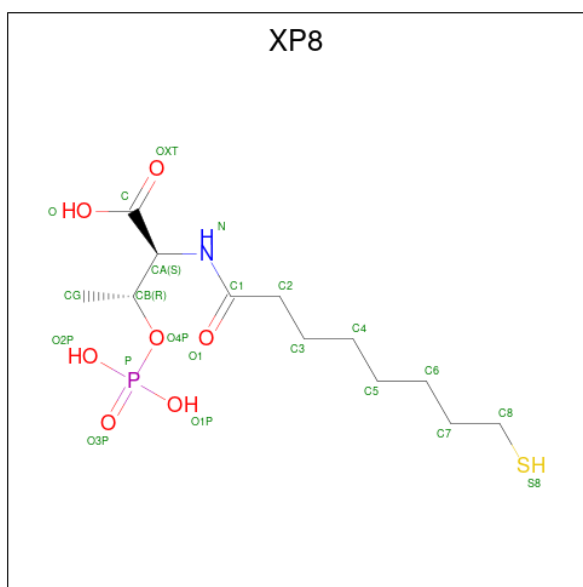
Mol	Chain	Residues	Atoms							ZeroOcc	AltConf
8	A	1	Total	C	N	O	P	S		0	1
			22	12	1	7	1	1			
8	D	1	Total	C	N	O	P	S		0	1
			22	12	1	7	1	1			

- Molecule 9 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).
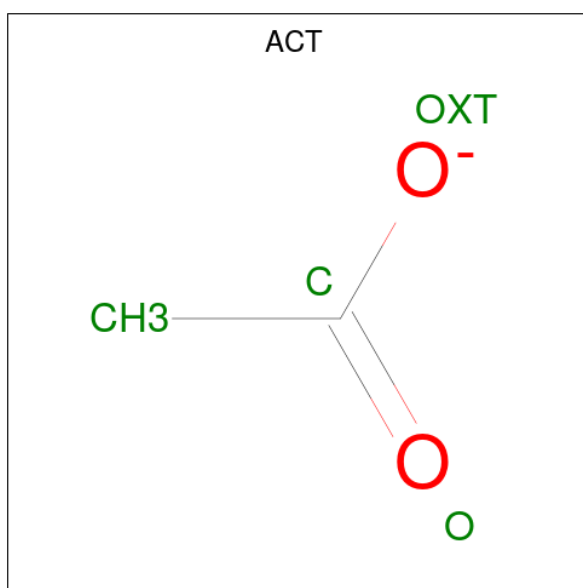
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	A	1	Total	C	O	0	1
			4	2	2		

- Molecule 10 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).

EDO

O2

OH

C1

C2

HO

O1

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

- Molecule 11 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	D	1	Total Zn 1 1		0	0

- Molecule 12 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	A	427	Total O 435 435		0	15
12	B	345	Total O 349 349		0	5
12	C	200	Total O 201 201		0	8
12	D	415	Total O 421 421		0	9

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	E	312	Total 314	O 314	0	6
12	F	172	Total 173	O 173	0	6

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.

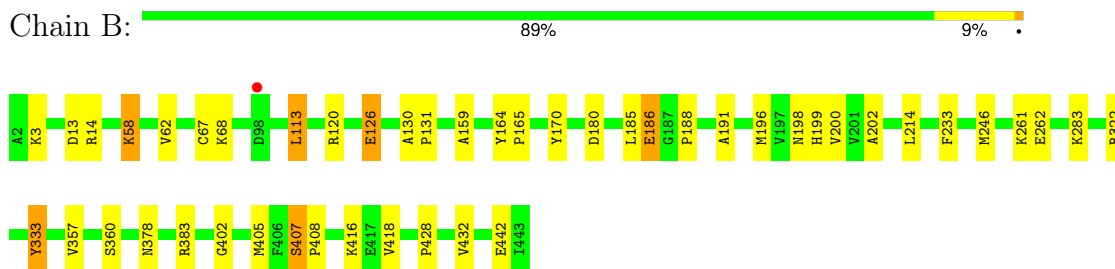• Molecule 1: Methyl-coenzyme M reductase I subunit alpha
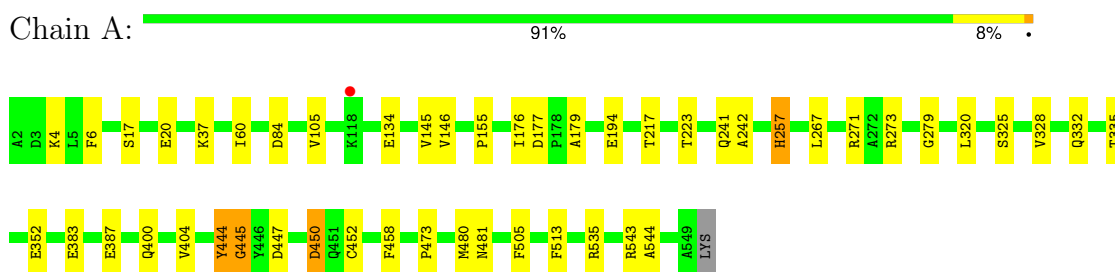
Chain A:

• Molecule 1: Methyl-coenzyme M reductase I subunit alpha
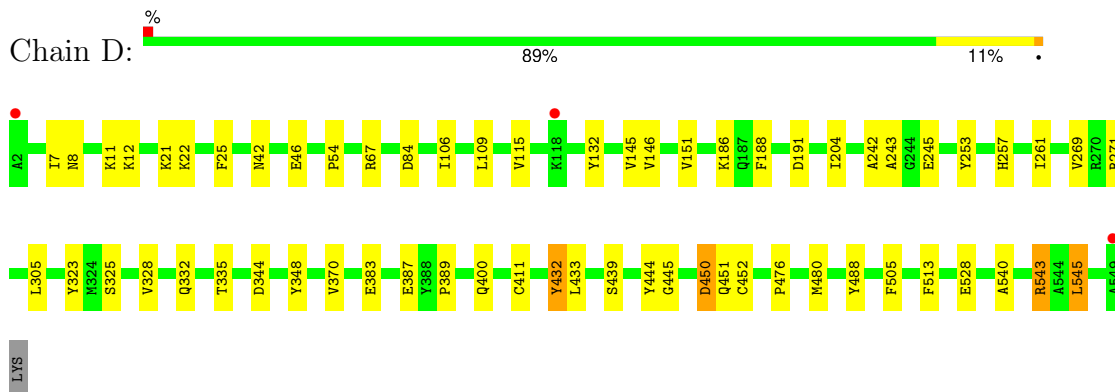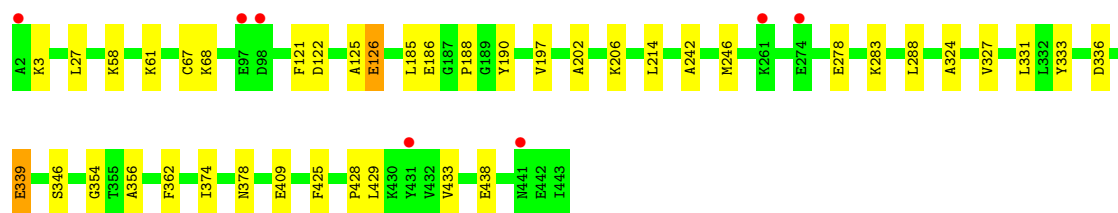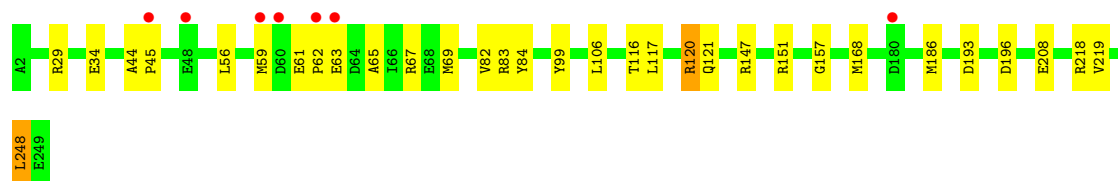
Chain D:

• Molecule 2: Methyl-coenzyme M reductase I subunit beta
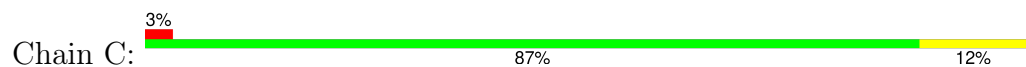
Chain B:

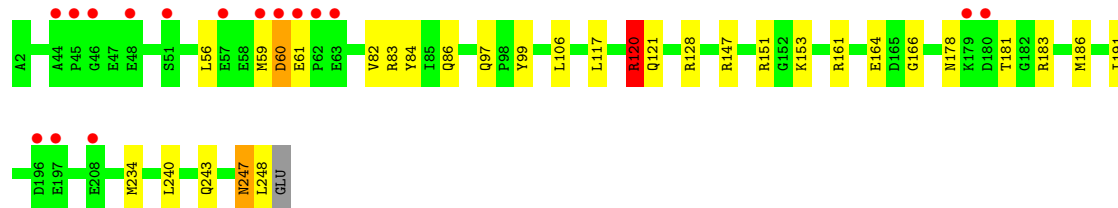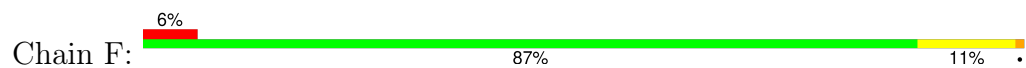• Molecule 2: Methyl-coenzyme M reductase I subunit beta

Chain E:

• Molecule 3: Methyl-coenzyme M reductase I subunit gamma

Chain C:

• Molecule 3: Methyl-coenzyme M reductase I subunit gamma

Chain F:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	82.02Å 118.26Å 122.39Å 90.00° 91.84° 90.00°	Depositor
Resolution (Å)	19.93 – 1.80 19.93 – 1.80	Depositor EDS
% Data completeness (in resolution range)	98.0 (19.93-1.80) 98.1 (19.93-1.80)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.32 (at 1.80Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.150 , 0.195 0.150 , 0.195	Depositor DCC
R_{free} test set	10625 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	15.2	Xtriage
Anisotropy	0.039	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 54.0	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.000 for -h,l,k 0.010 for -h,-l,-k 0.017 for h,-k,-l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	21893	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 3.35% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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: MHS, MGN, ZN, GL3, AGM, ACT, F43, EDO, TP7, SMC, XP8, MG, COM

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	1.39	7/4492 (0.2%)	1.02	3/6099 (0.0%)
1	D	1.44	14/4459 (0.3%)	1.06	5/6052 (0.1%)
2	B	1.34	7/3527 (0.2%)	1.01	4/4770 (0.1%)
2	E	1.28	7/3524 (0.2%)	0.99	2/4766 (0.0%)
3	C	1.23	3/2111 (0.1%)	1.05	7/2841 (0.2%)
3	F	1.19	3/2120 (0.1%)	1.03	3/2854 (0.1%)
All	All	1.34	41/20233 (0.2%)	1.03	24/27382 (0.1%)

The worst 5 of 41 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	62	VAL	CB-CG1	6.38	1.66	1.52
1	D	370	VAL	CB-CG1	6.15	1.65	1.52
1	D	25	PHE	CD1-CE1	5.96	1.51	1.39
1	D	411	CYS	CB-SG	5.92	1.92	1.82
1	A	458	PHE	CE1-CZ	5.85	1.48	1.37

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	122	ASP	CB-CG-OD2	-6.07	112.83	118.30
1	D	488	TYR	CB-CG-CD2	-5.99	117.41	121.00
3	C	147	ARG	NE-CZ-NH1	5.80	123.20	120.30
1	A	450	ASP	N-CA-CB	5.65	120.78	110.60
3	F	161	ARG	NE-CZ-NH1	5.56	123.08	120.30

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	4408	0	4185	29	0
1	D	4359	0	4196	37	0
2	B	3422	0	3450	36	0
2	E	3431	0	3455	21	0
3	C	2059	0	1985	21	0
3	F	2071	0	2003	27	0
4	A	1	0	0	0	0
4	B	2	0	0	0	0
4	C	1	0	0	0	0
4	F	1	0	0	0	0
5	A	62	0	43	2	0
5	D	62	0	43	3	0
6	A	21	0	19	3	0
6	D	21	0	19	1	0
7	A	7	0	5	1	0
7	D	7	0	5	0	0
8	A	22	0	21	1	0
8	D	22	0	21	1	0
9	A	4	0	3	0	0
10	A	8	0	12	0	0
10	D	8	0	12	3	0
11	D	1	0	0	0	0
12	A	435	0	0	3	0
12	B	349	0	0	3	0
12	C	201	0	0	3	0
12	D	421	0	0	8	0
12	E	314	0	0	3	0
12	F	173	0	0	4	0
All	All	21893	0	19477	146	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 4.

The worst 5 of 146 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:196[A]:MET:CE	2:B:198:ASN:HB2	1.71	1.18
1:D:432:TYR:HB2	3:F:234[B]:MET:HE3	1.18	1.16
2:B:196[A]:MET:HE1	2:B:198:ASN:HB2	1.38	1.02
1:D:432:TYR:CB	3:F:234[B]:MET:HE3	1.90	1.02
2:B:196[A]:MET:HE3	2:B:198:ASN:HB2	1.45	0.94

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	564/549 (103%)	543 (96%)	20 (4%)	1 (0%)	47	33
1	D	560/549 (102%)	543 (97%)	16 (3%)	1 (0%)	47	33
2	B	459/442 (104%)	448 (98%)	10 (2%)	1 (0%)	47	33
2	E	459/442 (104%)	449 (98%)	10 (2%)	0	100	100
3	C	254/248 (102%)	243 (96%)	11 (4%)	0	100	100
3	F	256/248 (103%)	246 (96%)	10 (4%)	0	100	100
All	All	2552/2478 (103%)	2472 (97%)	77 (3%)	3 (0%)	51	36

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	325	SER
1	A	325	SER
2	B	402	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	456/434 (105%)	452 (99%)	4 (1%)	78	75
1	D	452/434 (104%)	444 (98%)	8 (2%)	59	48
2	B	360/341 (106%)	355 (99%)	5 (1%)	67	59
2	E	360/341 (106%)	354 (98%)	6 (2%)	60	51
3	C	224/216 (104%)	220 (98%)	4 (2%)	59	48
3	F	225/216 (104%)	219 (97%)	6 (3%)	44	31
All	All	2077/1982 (105%)	2044 (98%)	33 (2%)	62	54

5 of 33 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	F	61	GLU
3	F	120	ARG
3	F	247	ASN
3	C	248	LEU
3	C	186	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	111	HIS
2	E	318	GLN
3	F	247	ASN
3	F	121	GLN
3	C	121	GLN

5.3.3   RNA ⓘ

There are no RNA molecules in this entry.

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

10 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	$\#|Z| > 2$
1	AGM	D	271	1	10,11,12	1.62	3 (30%)	7,13,15	2.15	4 (57%)
1	SMC	D	452	1	5,6,7	0.77	0	3,6,8	2.18	1 (33%)
1	MGN	A	400	1	6,9,10	2.12	1 (16%)	7,12,14	1.74	1 (14%)
1	MHS	D	257	1	7,11,12	2.33	1 (14%)	7,14,16	3.64	4 (57%)
1	SMC	A	452	1	5,6,7	0.95	0	3,6,8	2.20	1 (33%)
1	AGM	A	271	1	10,11,12	1.38	2 (20%)	7,13,15	2.63	4 (57%)
1	MHS	A	257	1	7,11,12	2.25	2 (28%)	7,14,16	3.89	4 (57%)
1	MGN	D	400	1	6,9,10	1.67	1 (16%)	7,12,14	1.79	1 (14%)
1	GL3	D	445	1	2,3,4	3.48	1 (50%)	1,2,4	1.07	0
1	GL3	A	445	1	2,3,4	3.17	1 (50%)	1,2,4	0.12	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
1	AGM	D	271	1	-	1/10/11/13	-
1	SMC	D	452	1	-	1/3/5/7	-
1	MGN	A	400	1	-	0/7/9/12	-
1	MHS	D	257	1	-	0/5/6/8	0/1/1/1
1	SMC	A	452	1	-	1/3/5/7	-
1	AGM	A	271	1	-	3/10/11/13	-
1	MHS	A	257	1	-	0/5/6/8	0/1/1/1
1	MGN	D	400	1	-	0/7/9/12	-
1	GL3	D	445	1	-	1/1/1/2	-
1	GL3	A	445	1	-	1/1/1/2	-

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	257	MHS	CE1-NE2	5.50	1.47	1.34
1	A	257	MHS	CE1-NE2	5.07	1.46	1.34

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	445	GL3	C-S	-4.90	1.60	1.80
1	A	400	MGN	CB1-CA	4.61	1.60	1.55
1	A	445	GL3	C-S	-4.44	1.62	1.80

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	257	MHS	NE2-CE1-ND1	-8.69	99.36	112.26
1	D	257	MHS	NE2-CE1-ND1	-8.19	100.11	112.26
1	A	400	MGN	CB1-CG-CD	-4.26	103.62	112.13
1	A	271	AGM	NE1-CZ-NH2	-4.13	113.50	120.58
1	D	400	MGN	CB1-CG-CD	-4.00	104.15	112.13

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	452	SMC	CA-CB-SG-CS
1	D	452	SMC	CA-CB-SG-CS
1	A	445	GL3	S-C-CA-N
1	D	445	GL3	S-C-CA-N
1	A	271	AGM	CE2-CD-NE1-CZ

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	257	MHS	1	0
1	A	445	GL3	1	0

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 19 ligands modelled in this entry, 6 are monoatomic - leaving 13 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	$\#|Z| > 2$
9	ACT	A	555[A]	4	3,3,3	0.85	0	3,3,3	1.48	0
10	EDO	D	555	-	3,3,3	0.49	0	2,2,2	0.51	0
7	COM	A	553	5	6,6,6	2.82	2 (33%)	8,8,8	2.32	2 (25%)
6	TP7	D	1[A]	-	19,20,20	0.91	1 (5%)	24,26,26	0.98	0
5	F43	D	551	7,1	63,71,71	2.43	15 (23%)	73,118,118	1.92	22 (30%)
7	COM	D	552	5	6,6,6	2.68	2 (33%)	8,8,8	2.22	3 (37%)
5	F43	A	1	7,1	63,71,71	2.47	16 (25%)	73,118,118	1.62	13 (17%)
8	XP8	D	553[B]	-	20,21,21	0.79	0	25,27,27	0.85	1 (4%)
10	EDO	A	556	-	3,3,3	0.88	0	2,2,2	0.46	0
10	EDO	D	554	-	3,3,3	0.33	0	2,2,2	0.29	0
10	EDO	A	557	-	3,3,3	0.45	0	2,2,2	0.40	0
6	TP7	A	552[A]	-	19,20,20	1.05	1 (5%)	24,26,26	1.18	1 (4%)
8	XP8	A	554[B]	-	20,21,21	0.98	1 (5%)	25,27,27	1.03	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
10	EDO	D	555	-	-	1/1/1/1	-
7	COM	A	553	5	-	0/4/4/4	-
6	TP7	D	1[A]	-	-	2/24/24/24	-
5	F43	D	551	7,1	-	10/28/185/185	-
7	COM	D	552	5	-	1/4/4/4	-
5	F43	A	1	7,1	-	9/28/185/185	-
8	XP8	D	553[B]	-	-	2/25/25/25	-
10	EDO	A	556	-	-	1/1/1/1	-
10	EDO	D	554	-	-	1/1/1/1	-
10	EDO	A	557	-	-	1/1/1/1	-
6	TP7	A	552[A]	-	-	3/24/24/24	-
8	XP8	A	554[B]	-	-	2/25/25/25	-

The worst 5 of 38 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	D	551	F43	NI-NA	9.67	2.12	1.89
5	A	1	F43	NI-NB	9.39	2.12	1.89
5	A	1	F43	NI-NA	9.32	2.11	1.89
5	D	551	F43	NI-NB	8.88	2.10	1.89
5	A	1	F43	NI-ND	7.23	2.06	1.89

The worst 5 of 42 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	D	551	F43	C5D-C2D-C1D	5.92	118.15	110.43
7	A	553	COM	O1S-S2-C2	5.71	115.35	106.73
5	D	551	F43	O8D-C7D-C6D	-4.16	114.05	120.87
5	D	551	F43	C2C-C5C-C6C	-4.05	106.38	113.95
5	D	551	F43	C6D-C7D-CHD	3.91	124.16	116.94

There are no chirality outliers.

5 of 33 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	D	552	COM	C1-C2-S2-O2S
8	D	553[B]	XP8	C6-C7-C8-S8
5	D	551	F43	C3A-CAA-CBA-CCA
5	A	1	F43	C3A-CAA-CBA-CCA
10	D	554	EDO	O1-C1-C2-O2

There are no ring outliers.
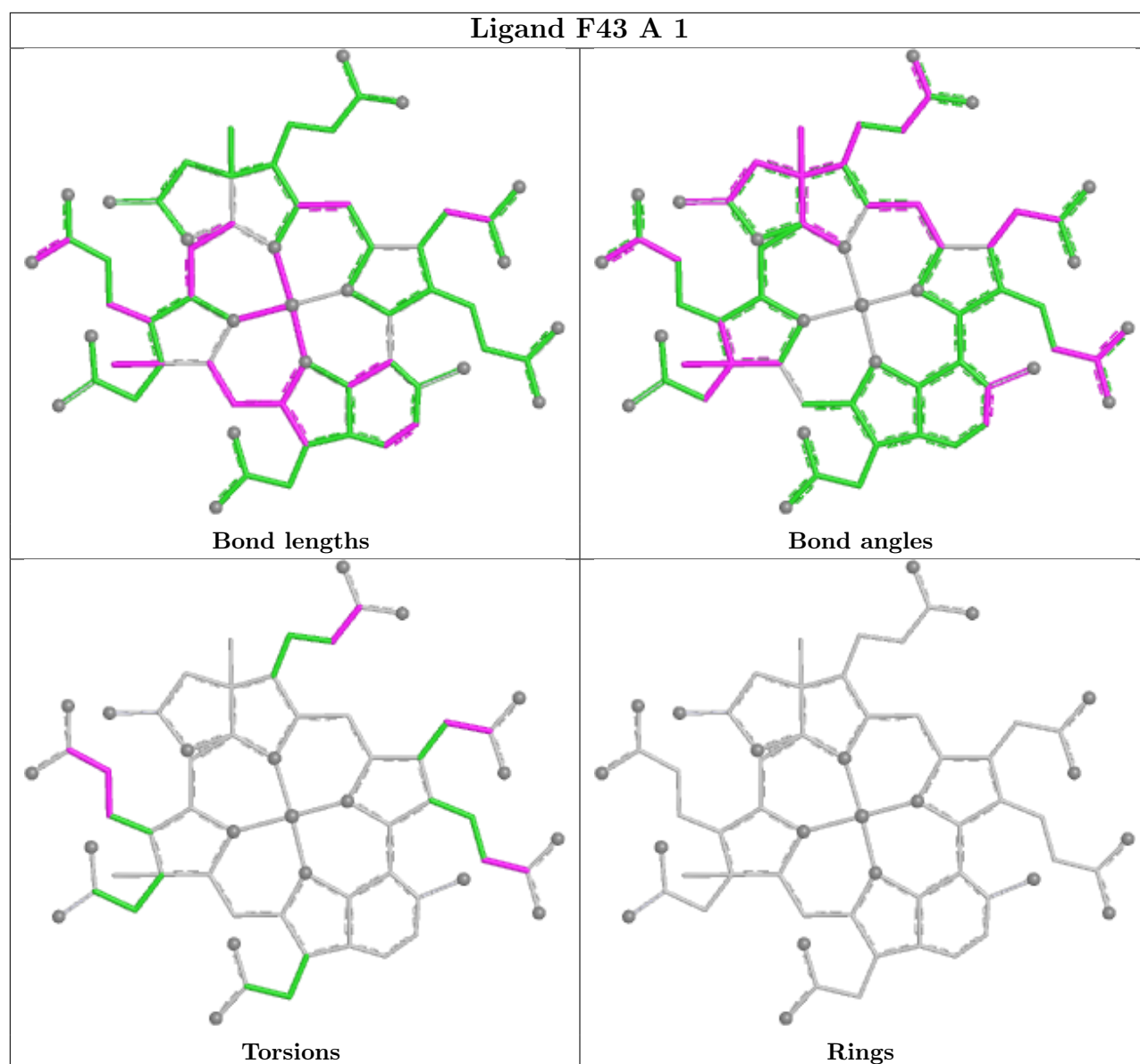
9 monomers are involved in 14 short contacts:

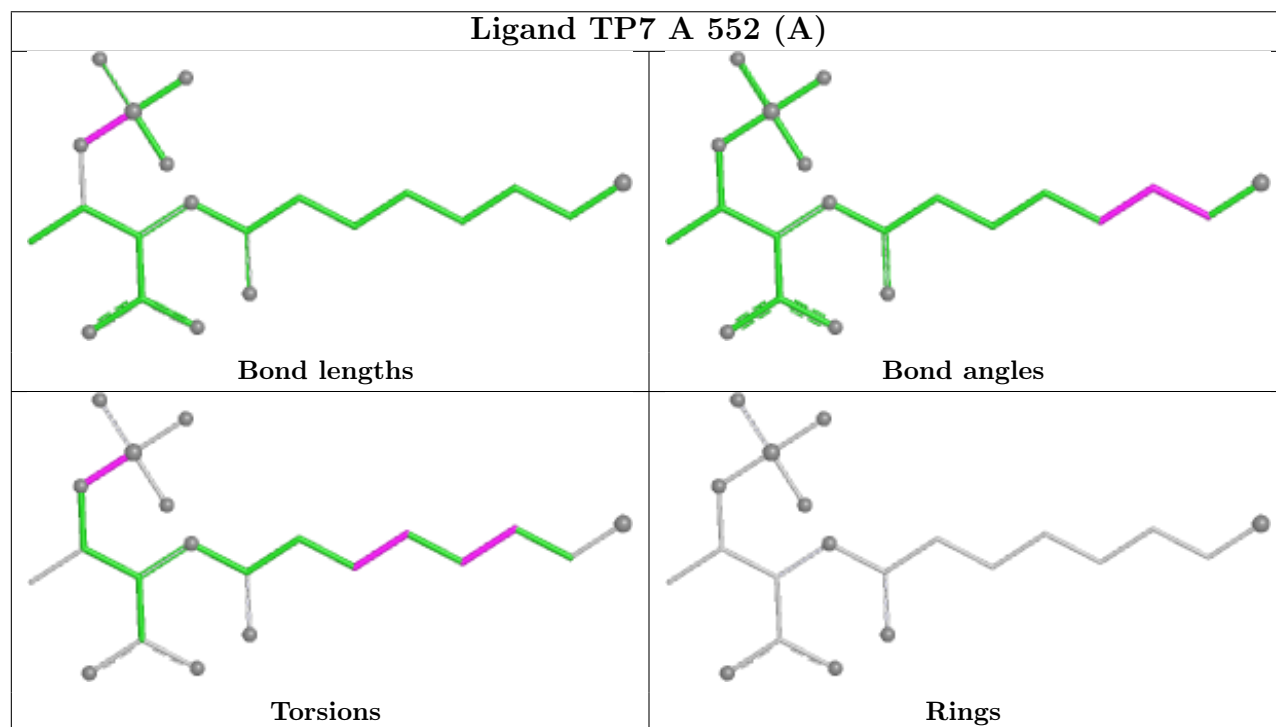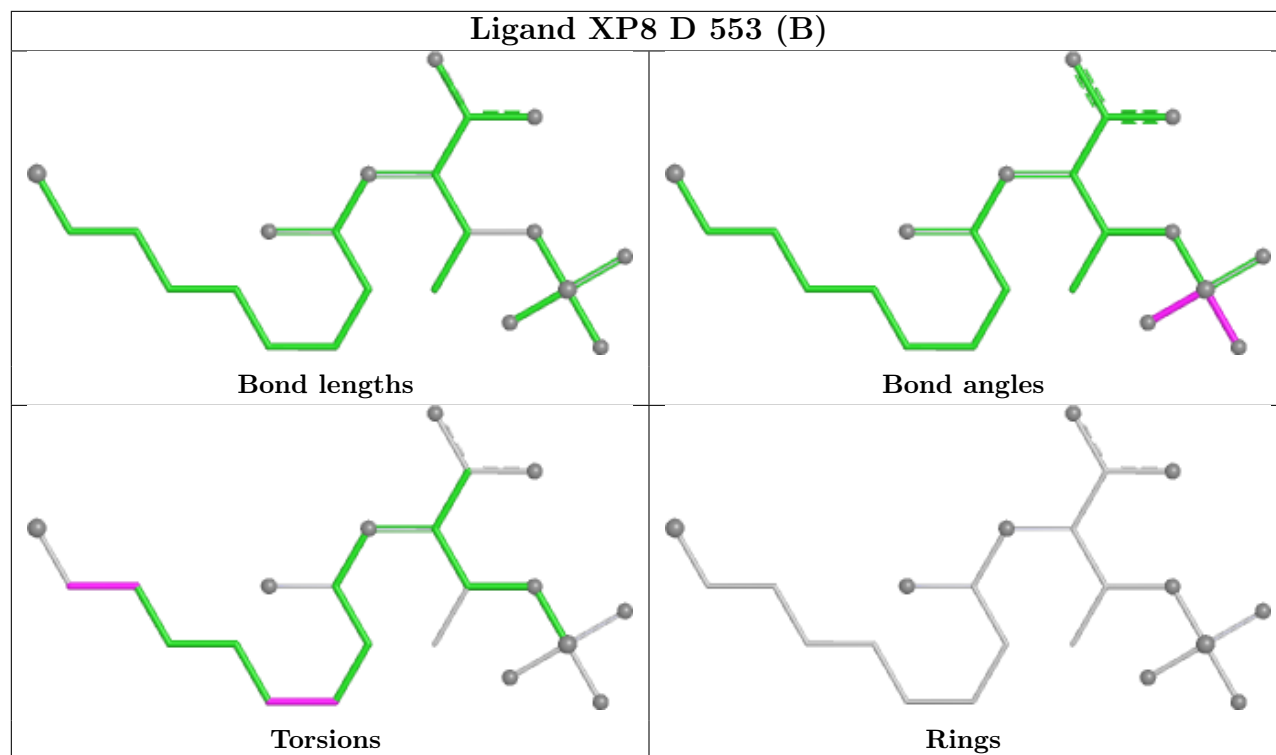| Mol	Chain	Res	Type	Clashes	Symm-Clashes
10	D	555	EDO	1	0
7	A	553	COM	1	0
6	D	1[A]	TP7	1	0
5	D	551	F43	3	0
5	A	1	F43	2	0
8	D	553[B]	XP8	1	0
10	D	554	EDO	2	0
6	A	552[A]	TP7	3	0
8	A	554[B]	XP8	1	0

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
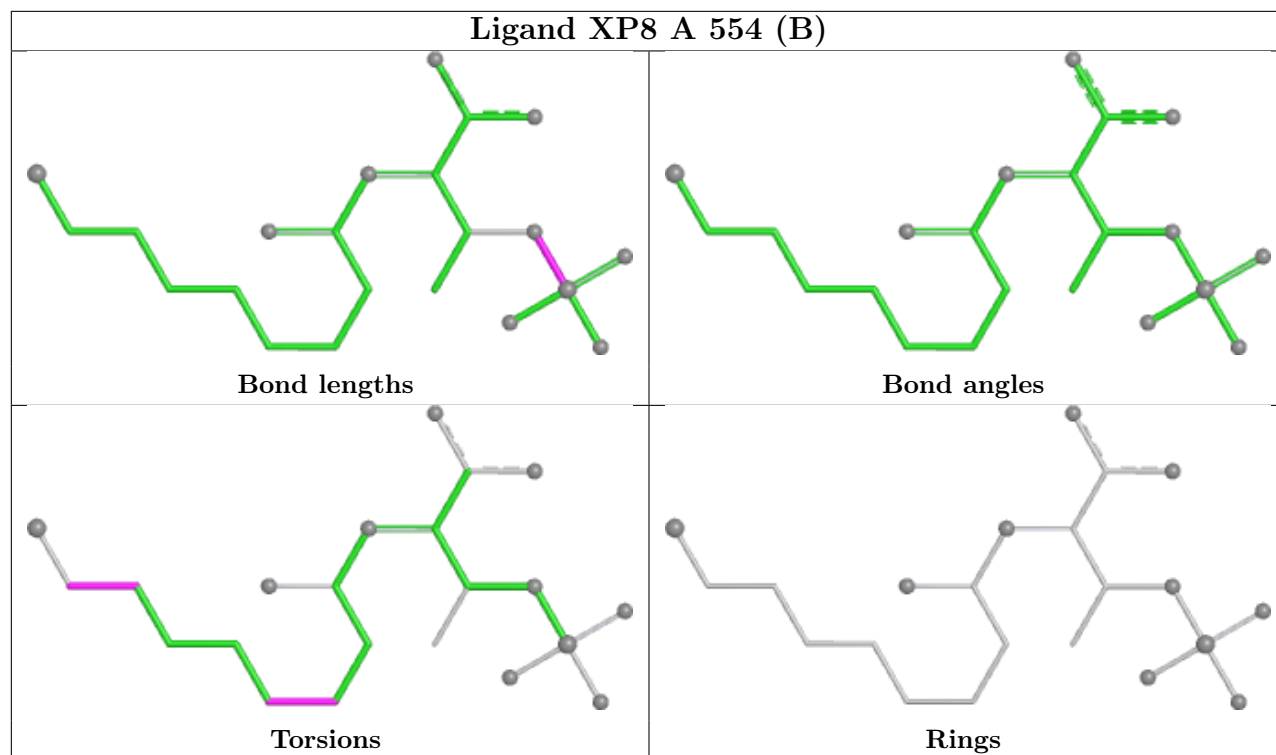
also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted 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.

Ligand TP7 D 1 (A)

Bond lengths

Bond angles

Torsions

Rings

Ligand F43 D 551

Bond lengths

Bond angles

Torsions

Rings

Ligand F43 A 1

Bond lengths

Bond angles

Torsions

Rings

Ligand XP8 D 553 (B)

Bond lengths

Bond angles

Torsions

Rings

Ligand TP7 A 552 (A)

Bond lengths

Bond angles

Torsions

Rings

Ligand XP8 A 554 (B)

Bond lengths

Bond angles

Torsions

Rings
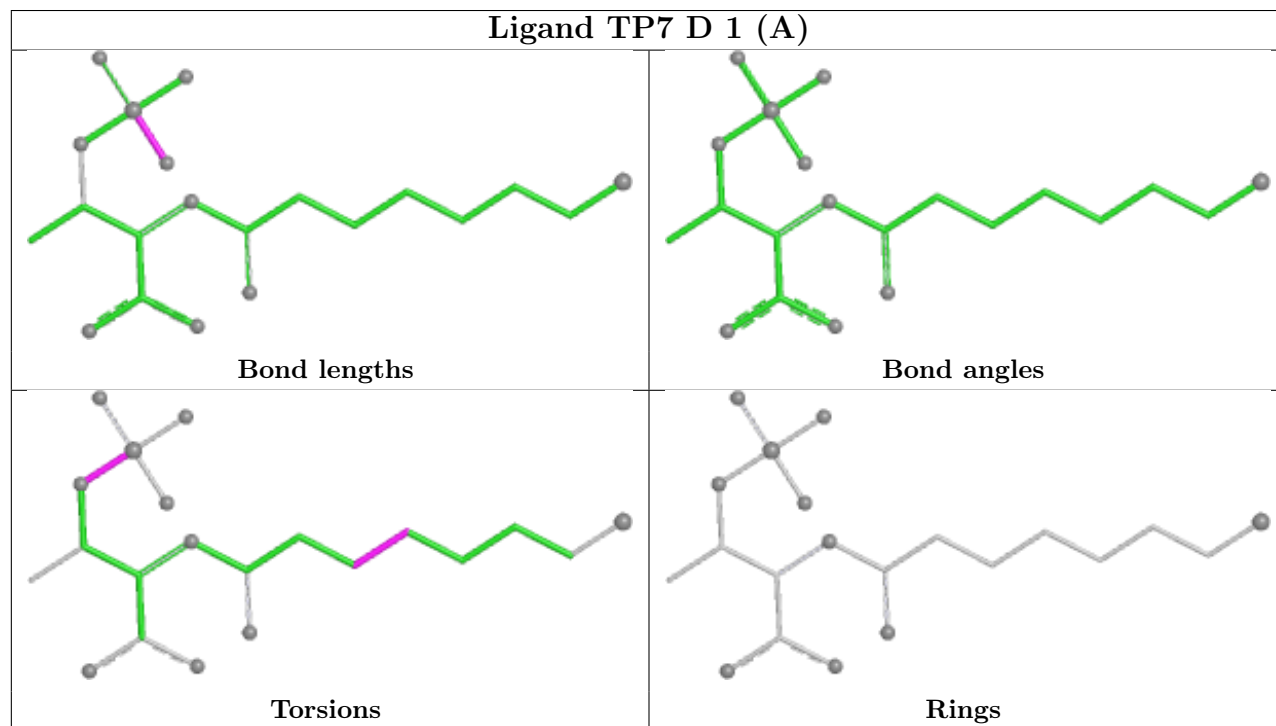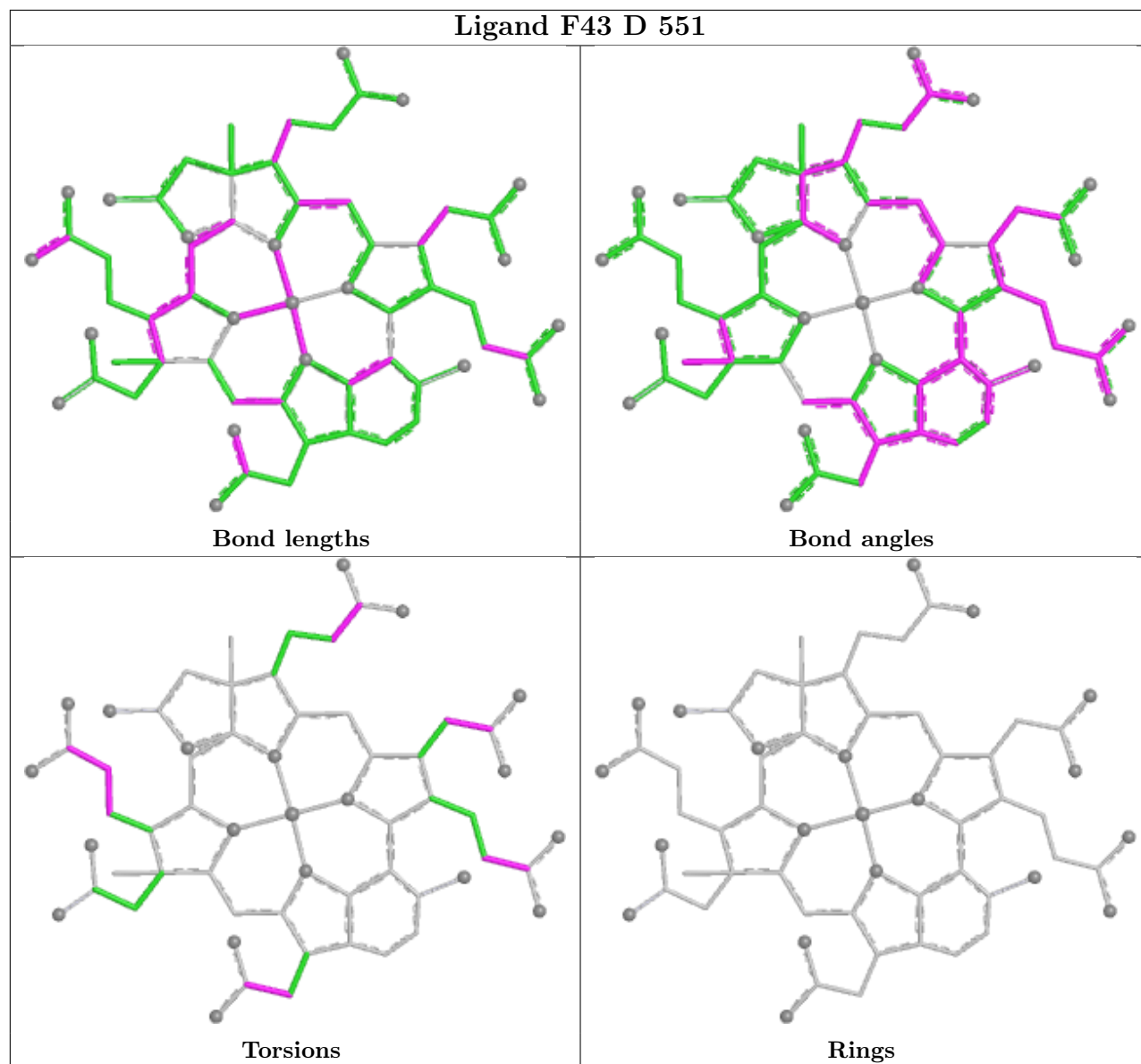
5.7 Other polymers ⓘ

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 ⓘ

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	543/549 (98%)	-0.78	1 (0%)	95	93	6, 11, 24, 39	0
1	D	543/549 (98%)	-0.76	3 (0%)	89	87	6, 11, 24, 40	0
2	B	442/442 (100%)	-0.56	1 (0%)	95	93	8, 16, 27, 43	0
2	E	442/442 (100%)	-0.35	7 (1%)	72	68	8, 19, 37, 53	0
3	C	248/248 (100%)	-0.41	7 (2%)	53	47	9, 18, 37, 58	0
3	F	247/248 (99%)	-0.13	16 (6%)	18	15	11, 22, 41, 63	0
All	All	2465/2478 (99%)	-0.56	35 (1%)	75	72	6, 15, 31, 63	0

The worst 5 of 35 RSRZ outliers are listed below:

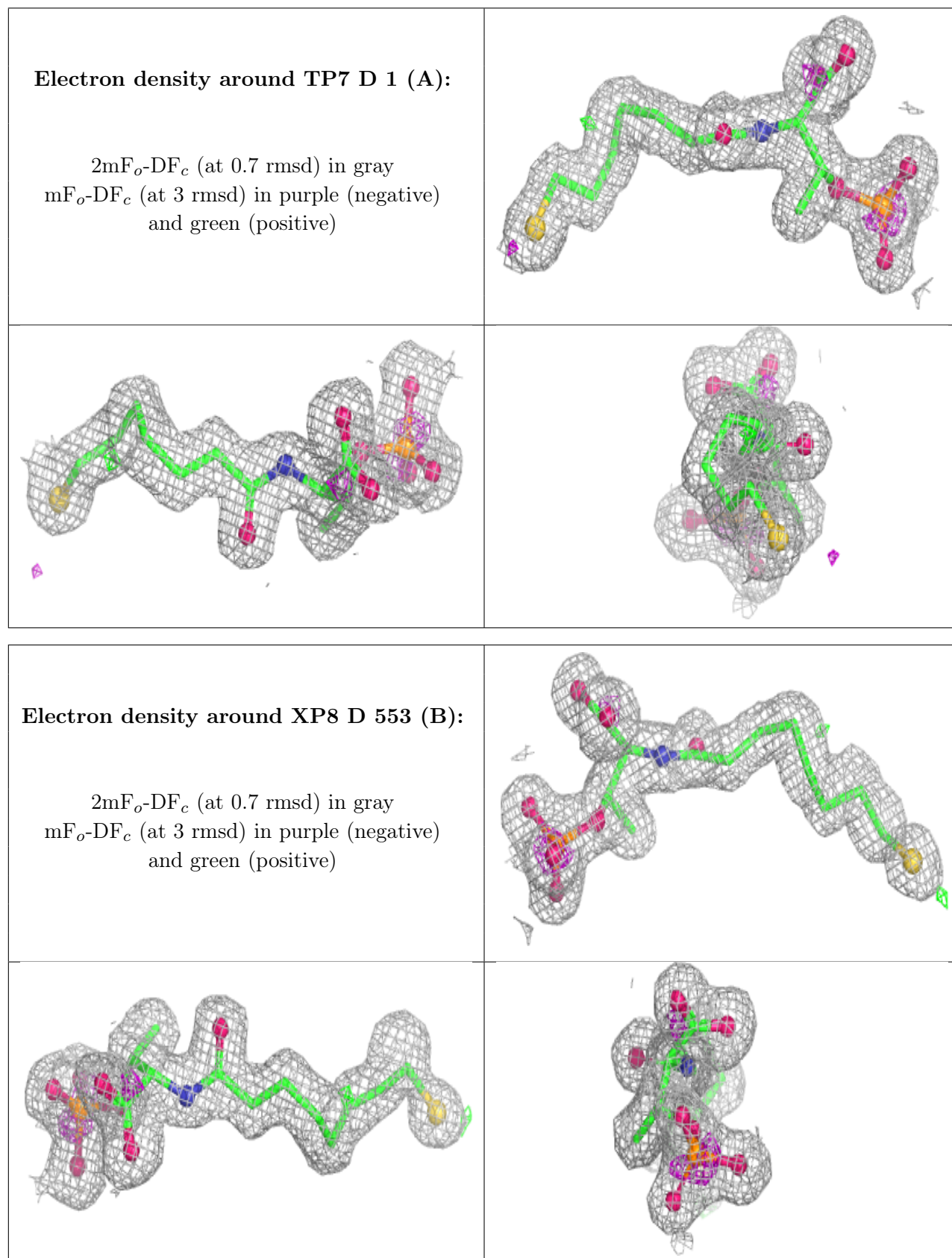| Mol	Chain	Res	Type	RSRZ
3	F	62	PRO	5.3
3	C	60	ASP	4.8
3	F	60	ASP	4.5
3	C	45	PRO	4.2
1	D	549	ALA	4.0

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
1	MHS	A	257	11/12	0.96	0.06	15,18,21,22	0
1	MHS	D	257	11/12	0.96	0.08	15,18,20,21	0
1	MGN	A	400	10/11	0.98	0.05	6,6,11,11	0
1	SMC	A	452	7/8	0.98	0.06	7,8,11,12	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
1	AGM	A	271	12/13	0.98	0.06	7,9,10,11	0
1	AGM	D	271	12/13	0.98	0.07	7,9,11,13	0
1	SMC	D	452	7/8	0.98	0.06	9,10,13,16	0
1	MGN	D	400	10/11	0.99	0.04	7,9,10,12	0
1	GL3	A	445	4/5	0.99	0.06	6,6,7,10	0
1	GL3	D	445	4/5	1.00	0.04	7,8,9,10	0

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
10	EDO	A	557	4/4	0.81	0.13	40,42,43,44	0
10	EDO	D	555	4/4	0.81	0.14	39,43,43,45	0
10	EDO	A	556	4/4	0.86	0.17	31,31,31,34	0
10	EDO	D	554	4/4	0.89	0.10	33,38,38,38	0
4	MG	B	1	1/1	0.89	0.12	41,41,41,41	0
4	MG	B	444	1/1	0.91	0.10	30,30,30,30	1
4	MG	A	551	1/1	0.94	0.18	22,22,22,22	1
4	MG	F	250	1/1	0.96	0.12	28,28,28,28	0
6	TP7	D	1[A]	21/21	0.96	0.09	12,14,17,19	21
8	XP8	D	553[B]	22/22	0.96	0.09	12,14,17,19	22
9	ACT	A	555[A]	4/4	0.96	0.10	21,21,21,22	4
6	TP7	A	552[A]	21/21	0.97	0.08	10,14,16,17	21
4	MG	C	250	1/1	0.97	0.12	32,32,32,32	0
8	XP8	A	554[B]	22/22	0.97	0.09	11,14,18,18	22
7	COM	D	552	7/7	0.99	0.07	14,16,18,22	0
5	F43	A	1	62/62	0.99	0.05	5,9,14,19	0
5	F43	D	551	62/62	0.99	0.05	3,8,12,16	0
7	COM	A	553	7/7	0.99	0.06	12,15,17,20	0
11	ZN	D	556	1/1	1.00	0.03	14,14,14,14	1

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.
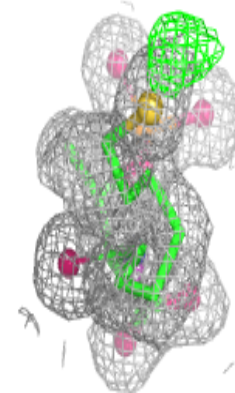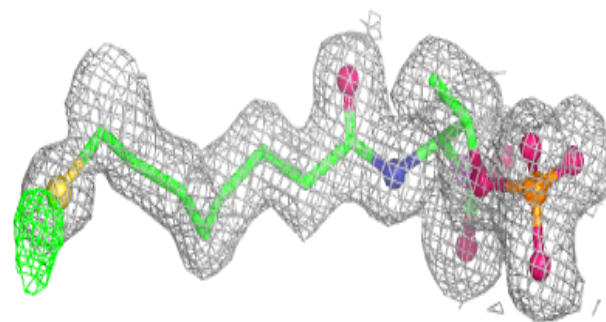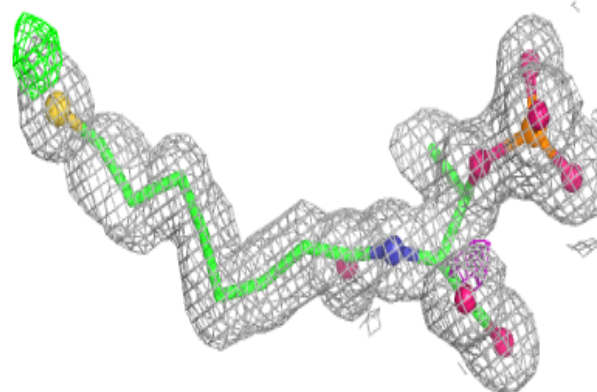
Electron density around TP7 D 1 (A):

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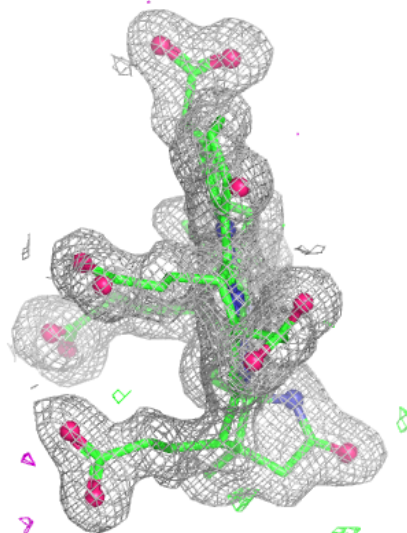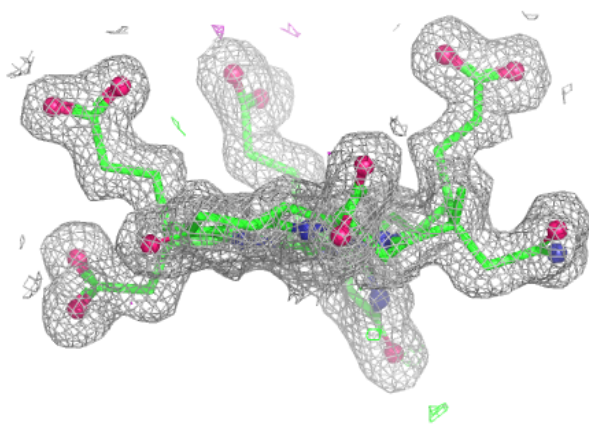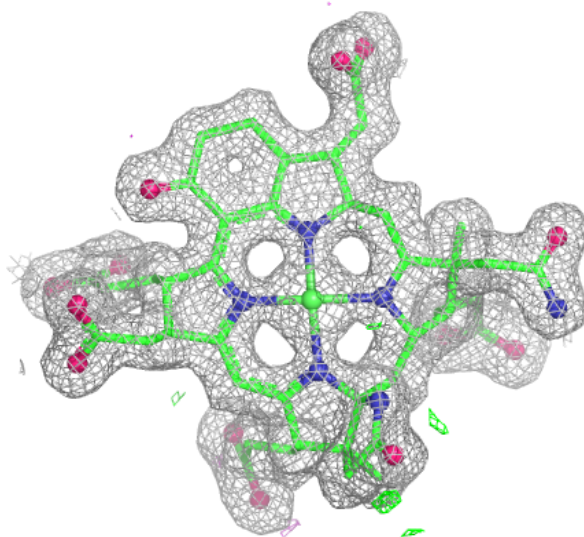mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

Electron density around XP8 D 553 (B):

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

Electron density around TP7 A 552 (A):

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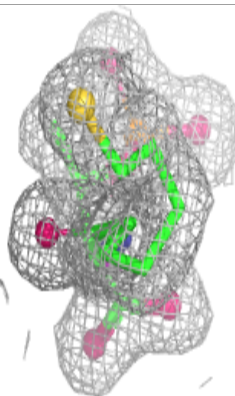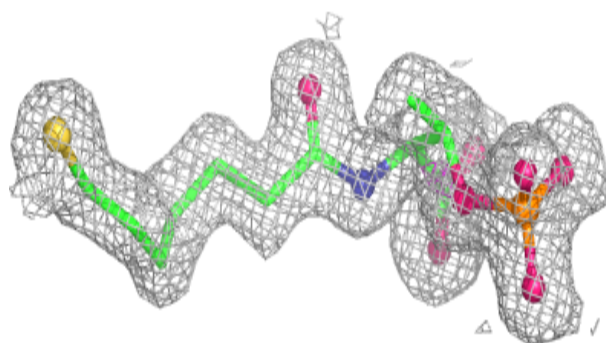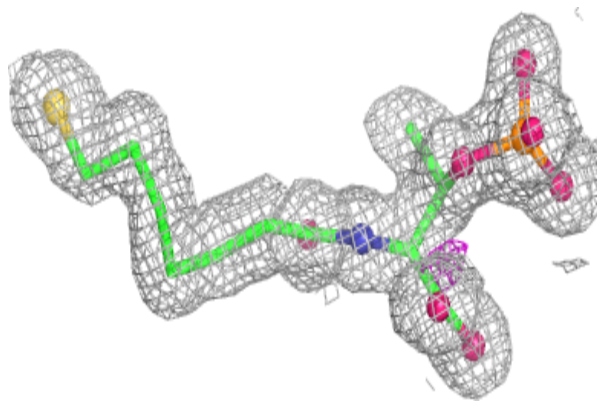
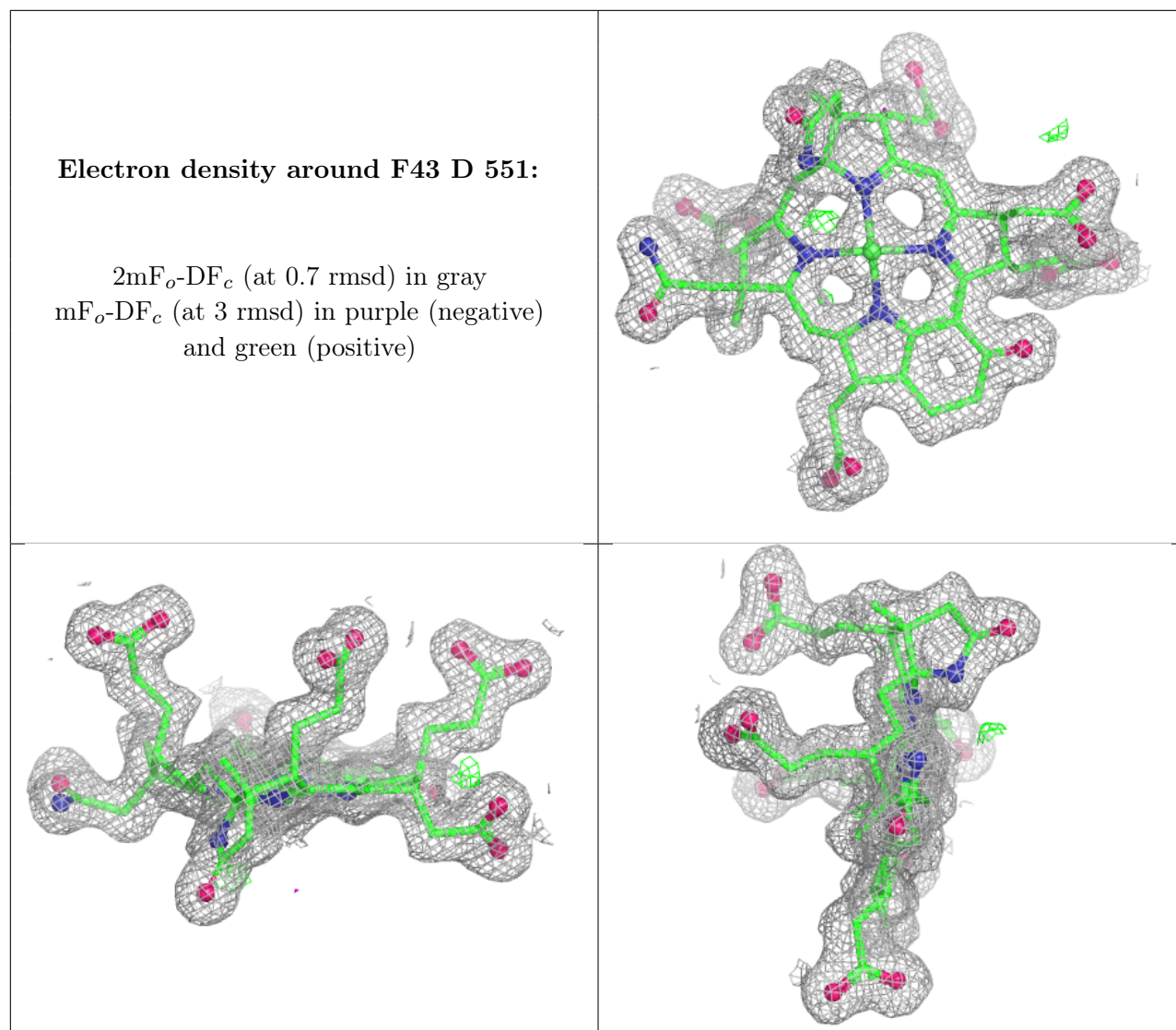

Electron density around XP8 A 554 (B):

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

Electron density around F43 A 1:

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

Electron density around F43 D 551:

$2mF_o$-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.