WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Dec 18, 2024 – 06:07 PM EST

PDB ID	:	9ECN
Title	:	M. acetivorans MCR containing a 2-methylglutamine modification
Authors	:	Nair, S.K.; Borkar, J.
Deposited on	:	2024-11-14
Resolution	:	2.00 Å(reported)

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

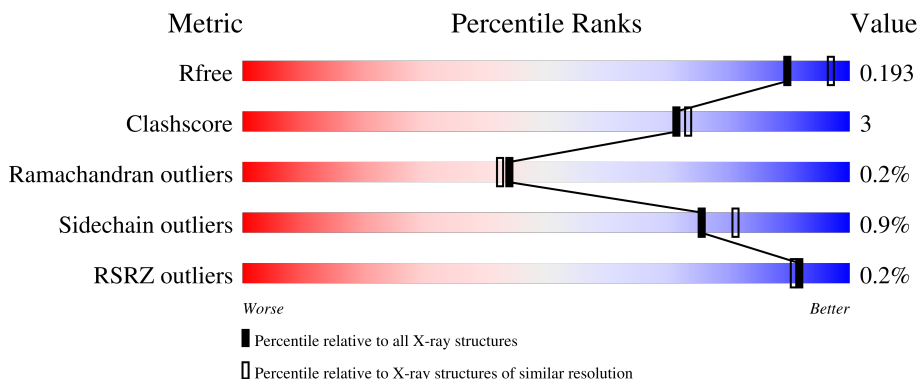|		
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.21
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

1   Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.

Metric	Percentile Ranks	Value
Rfree		0.193
Clashscore		3
Ramachandran outliers		0.2%
Sidechain outliers		0.9%
RSRZ outliers		0.2%

Worse　　　　　　　　　　　　　　　*Better*

▌Percentile relative to all X-ray structures

▯Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	9409 (2.00-2.00)
Clashscore	180529	10737 (2.00-2.00)
Ramachandran outliers	177936	10628 (2.00-2.00)
Sidechain outliers	177891	10627 (2.00-2.00)
RSRZ outliers	164620	9409 (2.00-2.00)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=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	570	92%　　　　　8%　・
1	B	570	92%　　　　　8%
2	C	434	91%　　　　　9%
2	D	434	92%　　　　　7%
3	E	321	71%　　　　6%　　23%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	321	73% · 23%

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
4	F43	A	1601	X	-	-	-
4	F43	A	1602	X	-	-	-
5	PEG	B	2001	-	-	X	-

2   Entry composition ⓘ

There are 8 unique types of molecules in this entry. The entry contains 20407 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 subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	568	Total 4347	C 2734	N 740	O 842	S 31	0	2	0
1	B	568	Total 4349	C 2735	N 740	O 843	S 31	0	3	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	432	Total 3148	C 1971	N 543	O 616	S 18	0	0	0
2	D	432	Total 3148	C 1971	N 543	O 616	S 18	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	247	Total 1926	C 1192	N 352	O 369	S 13	0	0	0
3	F	247	Total 1926	C 1192	N 352	O 369	S 13	0	0	0

There are 146 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	2928	MET	-	initiating methionine	UNP Q8THH0
E	2929	ASP	-	expression tag	UNP Q8THH0
E	2930	TYR	-	expression tag	UNP Q8THH0
E	2931	LYS	-	expression tag	UNP Q8THH0
E	2932	ASP	-	expression tag	UNP Q8THH0
E	2933	HIS	-	expression tag	UNP Q8THH0
E	2934	ASP	-	expression tag	UNP Q8THH0
E	2935	GLY	-	expression tag	UNP Q8THH0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	2936	ASP	-	expression tag	UNP Q8THH0
E	2937	TYR	-	expression tag	UNP Q8THH0
E	2938	LYS	-	expression tag	UNP Q8THH0
E	2939	ASP	-	expression tag	UNP Q8THH0
E	2940	HIS	-	expression tag	UNP Q8THH0
E	2941	ASP	-	expression tag	UNP Q8THH0
E	2942	ILE	-	expression tag	UNP Q8THH0
E	2943	ASP	-	expression tag	UNP Q8THH0
E	2944	TYR	-	expression tag	UNP Q8THH0
E	2945	LYS	-	expression tag	UNP Q8THH0
E	2946	ASP	-	expression tag	UNP Q8THH0
E	2947	ASP	-	expression tag	UNP Q8THH0
E	2948	ASP	-	expression tag	UNP Q8THH0
E	2949	ASP	-	expression tag	UNP Q8THH0
E	2950	LYS	-	expression tag	UNP Q8THH0
E	2951	GLY	-	expression tag	UNP Q8THH0
E	2952	SER	-	expression tag	UNP Q8THH0
E	2953	ALA	-	expression tag	UNP Q8THH0
E	2954	ALA	-	expression tag	UNP Q8THH0
E	2955	SER	-	expression tag	UNP Q8THH0
E	2956	TRP	-	expression tag	UNP Q8THH0
E	2957	SER	-	expression tag	UNP Q8THH0
E	2958	HIS	-	expression tag	UNP Q8THH0
E	2959	PRO	-	expression tag	UNP Q8THH0
E	2960	GLN	-	expression tag	UNP Q8THH0
E	2961	PHE	-	expression tag	UNP Q8THH0
E	2962	GLU	-	expression tag	UNP Q8THH0
E	2963	LYS	-	expression tag	UNP Q8THH0
E	2964	GLY	-	expression tag	UNP Q8THH0
E	2965	GLY	-	expression tag	UNP Q8THH0
E	2966	GLY	-	expression tag	UNP Q8THH0
E	2967	SER	-	expression tag	UNP Q8THH0
E	2968	GLY	-	expression tag	UNP Q8THH0
E	2969	GLY	-	expression tag	UNP Q8THH0
E	2970	GLY	-	expression tag	UNP Q8THH0
E	2971	SER	-	expression tag	UNP Q8THH0
E	2972	GLY	-	expression tag	UNP Q8THH0
E	2973	GLY	-	expression tag	UNP Q8THH0
E	2974	GLY	-	expression tag	UNP Q8THH0
E	2975	SER	-	expression tag	UNP Q8THH0
E	2976	TRP	-	expression tag	UNP Q8THH0
E	2977	SER	-	expression tag	UNP Q8THH0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	2978	HIS	-	expression tag	UNP Q8THH0
E	2979	PRO	-	expression tag	UNP Q8THH0
E	2980	GLN	-	expression tag	UNP Q8THH0
E	2981	PHE	-	expression tag	UNP Q8THH0
E	2982	GLU	-	expression tag	UNP Q8THH0
E	2983	LYS	-	expression tag	UNP Q8THH0
E	2984	SER	-	expression tag	UNP Q8THH0
E	2985	GLY	-	expression tag	UNP Q8THH0
E	2986	GLY	-	expression tag	UNP Q8THH0
E	2987	GLY	-	expression tag	UNP Q8THH0
E	2988	GLY	-	expression tag	UNP Q8THH0
E	2989	SER	-	expression tag	UNP Q8THH0
E	2990	GLY	-	expression tag	UNP Q8THH0
E	2991	GLY	-	expression tag	UNP Q8THH0
E	2992	GLY	-	expression tag	UNP Q8THH0
E	2993	SER	-	expression tag	UNP Q8THH0
E	2994	GLY	-	expression tag	UNP Q8THH0
E	2995	GLY	-	expression tag	UNP Q8THH0
E	2996	ASP	-	expression tag	UNP Q8THH0
E	2997	ASP	-	expression tag	UNP Q8THH0
E	2998	ASP	-	expression tag	UNP Q8THH0
E	2999	ASP	-	expression tag	UNP Q8THH0
E	3000	LYS	-	expression tag	UNP Q8THH0
F	2928	MET	-	initiating methionine	UNP Q8THH0
F	2929	ASP	-	expression tag	UNP Q8THH0
F	2930	TYR	-	expression tag	UNP Q8THH0
F	2931	LYS	-	expression tag	UNP Q8THH0
F	2932	ASP	-	expression tag	UNP Q8THH0
F	2933	HIS	-	expression tag	UNP Q8THH0
F	2934	ASP	-	expression tag	UNP Q8THH0
F	2935	GLY	-	expression tag	UNP Q8THH0
F	2936	ASP	-	expression tag	UNP Q8THH0
F	2937	TYR	-	expression tag	UNP Q8THH0
F	2938	LYS	-	expression tag	UNP Q8THH0
F	2939	ASP	-	expression tag	UNP Q8THH0
F	2940	HIS	-	expression tag	UNP Q8THH0
F	2941	ASP	-	expression tag	UNP Q8THH0
F	2942	ILE	-	expression tag	UNP Q8THH0
F	2943	ASP	-	expression tag	UNP Q8THH0
F	2944	TYR	-	expression tag	UNP Q8THH0
F	2945	LYS	-	expression tag	UNP Q8THH0
F	2946	ASP	-	expression tag	UNP Q8THH0

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	2947	ASP	-	expression tag	UNP Q8THH0
F	2948	ASP	-	expression tag	UNP Q8THH0
F	2949	ASP	-	expression tag	UNP Q8THH0
F	2950	LYS	-	expression tag	UNP Q8THH0
F	2951	GLY	-	expression tag	UNP Q8THH0
F	2952	SER	-	expression tag	UNP Q8THH0
F	2953	ALA	-	expression tag	UNP Q8THH0
F	2954	ALA	-	expression tag	UNP Q8THH0
F	2955	SER	-	expression tag	UNP Q8THH0
F	2956	TRP	-	expression tag	UNP Q8THH0
F	2957	SER	-	expression tag	UNP Q8THH0
F	2958	HIS	-	expression tag	UNP Q8THH0
F	2959	PRO	-	expression tag	UNP Q8THH0
F	2960	GLN	-	expression tag	UNP Q8THH0
F	2961	PHE	-	expression tag	UNP Q8THH0
F	2962	GLU	-	expression tag	UNP Q8THH0
F	2963	LYS	-	expression tag	UNP Q8THH0
F	2964	GLY	-	expression tag	UNP Q8THH0
F	2965	GLY	-	expression tag	UNP Q8THH0
F	2966	GLY	-	expression tag	UNP Q8THH0
F	2967	SER	-	expression tag	UNP Q8THH0
F	2968	GLY	-	expression tag	UNP Q8THH0
F	2969	GLY	-	expression tag	UNP Q8THH0
F	2970	GLY	-	expression tag	UNP Q8THH0
F	2971	SER	-	expression tag	UNP Q8THH0
F	2972	GLY	-	expression tag	UNP Q8THH0
F	2973	GLY	-	expression tag	UNP Q8THH0
F	2974	GLY	-	expression tag	UNP Q8THH0
F	2975	SER	-	expression tag	UNP Q8THH0
F	2976	TRP	-	expression tag	UNP Q8THH0
F	2977	SER	-	expression tag	UNP Q8THH0
F	2978	HIS	-	expression tag	UNP Q8THH0
F	2979	PRO	-	expression tag	UNP Q8THH0
F	2980	GLN	-	expression tag	UNP Q8THH0
F	2981	PHE	-	expression tag	UNP Q8THH0
F	2982	GLU	-	expression tag	UNP Q8THH0
F	2983	LYS	-	expression tag	UNP Q8THH0
F	2984	SER	-	expression tag	UNP Q8THH0
F	2985	GLY	-	expression tag	UNP Q8THH0
F	2986	GLY	-	expression tag	UNP Q8THH0
F	2987	GLY	-	expression tag	UNP Q8THH0
F	2988	GLY	-	expression tag	UNP Q8THH0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	2989	SER	-	expression tag	UNP Q8THH0
F	2990	GLY	-	expression tag	UNP Q8THH0
F	2991	GLY	-	expression tag	UNP Q8THH0
F	2992	GLY	-	expression tag	UNP Q8THH0
F	2993	SER	-	expression tag	UNP Q8THH0
F	2994	GLY	-	expression tag	UNP Q8THH0
F	2995	GLY	-	expression tag	UNP Q8THH0
F	2996	ASP	-	expression tag	UNP Q8THH0
F	2997	ASP	-	expression tag	UNP Q8THH0
F	2998	ASP	-	expression tag	UNP Q8THH0
F	2999	ASP	-	expression tag	UNP Q8THH0
F	3000	LYS	-	expression tag	UNP Q8THH0

- Molecule 4 is FACTOR 430 (three-letter code: F43) (formula: $C_{42}H_{51}N_6NiO_{13}$) (labeled as "Ligand of Interest" by depositor).
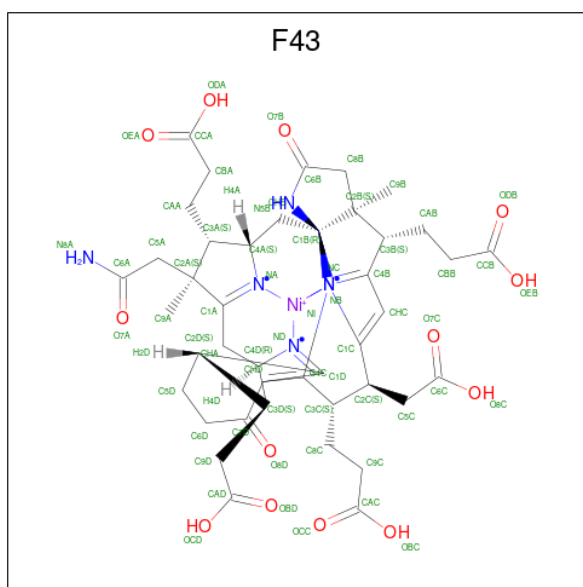
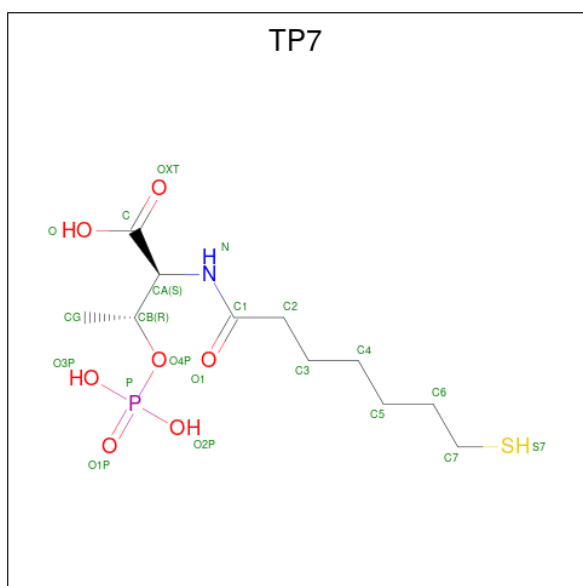| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	Ni	O	0	0
			62	42	6	1	13		
4	A	1	Total	C	N	Ni	O	0	0
			62	42	6	1	13		

- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).

PEG

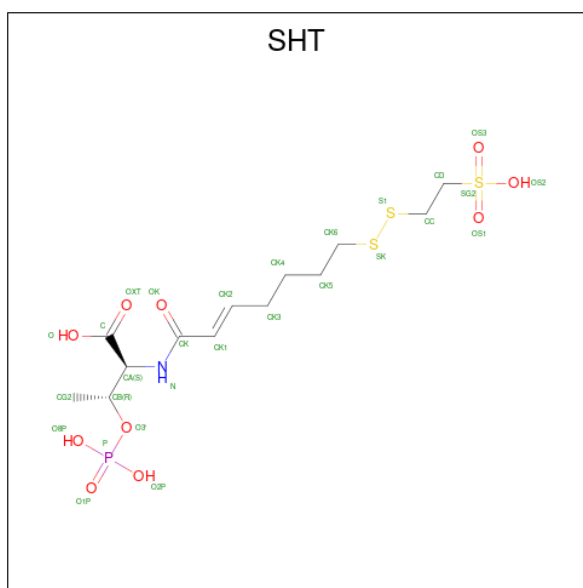| Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
5	B	1	7	4	3	0	0

- Molecule 6 is Coenzyme B (three-letter code: TP7) (formula: $C_{11}H_{22}NO_7PS$) (labeled as "Ligand of Interest" by depositor).

TP7

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

- Molecule 7 is O-PHOSPHONO-N-{(2E)-7-[(2-SULFOETHYL)DITHIO]HEPT-2-ENOYL}-L-THREONINE (three-letter code: SHT) (formula: $C_{13}H_{24}NO_{10}PS_3$) (labeled as "Ligand

of Interest" by depositor).

SHT

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
7	B	1	Total	C	N	O	P	S	0	0
			28	13	1	10	1	3		

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	305	Total 305	O 305	0	0
8	B	322	Total 322	O 322	0	0
8	C	242	Total 242	O 242	0	0
8	D	215	Total 215	O 215	0	0
8	E	143	Total 143	O 143	0	0
8	F	156	Total 156	O 156	0	0

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 subunit alpha

Chain A:

• Molecule 1: Methyl-coenzyme M reductase subunit alpha

Chain B:

• Molecule 2: Methyl-coenzyme M reductase subunit beta

Chain C:

• Molecule 2: Methyl-coenzyme M reductase subunit beta

Chain D:

● Molecule 3: Methyl-coenzyme M reductase subunit gamma

Chain E:

● Molecule 3: Methyl-coenzyme M reductase subunit gamma

Chain F:
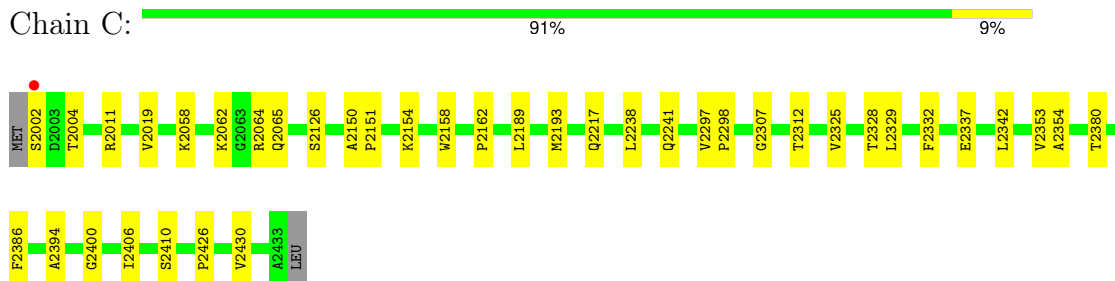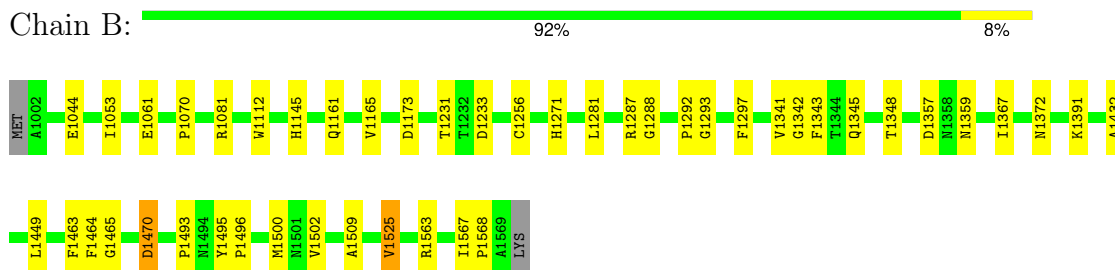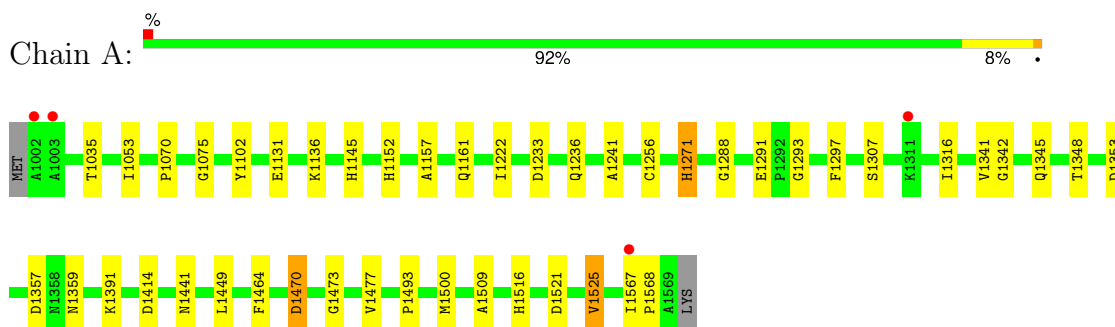
4   Data and refinement statistics   ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	109.47Å 82.94Å 122.47Å 90.00° 93.78° 90.00°	Depositor
Resolution (Å)	25.00 – 2.00 25.00 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.7 (25.00-2.00) 99.7 (25.00-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.50 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.8.0425	Depositor
R, R_{free}	0.152 , 0.193 0.152 , 0.193	Depositor DCC
R_{free} test set	7370 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	24.8	Xtriage
Anisotropy	0.421	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 37.4	EDS
L-test for twinning[2]	$< |L| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	20407	wwPDB-VP
Average B, all atoms (Å^2)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 29.71 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.4793e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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: SHT, TP7, PEG, DYA, AGM, MGN, F43, SMC, GL3, MHS

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	0.32	0/4393	0.70	0/5940
1	B	0.33	0/4398	0.69	3/5947 (0.1%)
2	C	0.33	0/3195	0.67	0/4324
2	D	0.33	0/3195	0.68	1/4324 (0.0%)
3	E	0.32	0/1965	0.69	1/2653 (0.0%)
3	F	0.33	0/1965	0.68	0/2653
All	All	0.33	0/19111	0.69	5/25841 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	1
All	All	0	2

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	2404	PHE	CB-CA-C	6.82	124.04	110.40
1	B	1081	ARG	CD-NE-CZ	5.54	131.35	123.60
1	B	1081	ARG	NE-CZ-NH2	5.51	123.06	120.30
1	B	1081	ARG	NE-CZ-NH1	-5.35	117.62	120.30
3	E	3101	PHE	CB-CA-C	5.20	120.79	110.40

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1470	DYA	Mainchain
1	B	1470	DYA	Mainchain

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	4347	0	4195	33	0
1	B	4349	0	4201	32	0
2	C	3148	0	3129	22	0
2	D	3148	0	3129	16	0
3	E	1926	0	1871	13	0
3	F	1926	0	1871	9	0
4	A	124	0	86	15	0
5	B	7	0	10	9	0
6	B	21	0	19	1	0
7	B	28	0	21	5	0
8	A	305	0	0	3	0
8	B	322	0	0	3	1
8	C	242	0	0	1	0
8	D	215	0	0	0	0
8	E	143	0	0	0	1
8	F	156	0	0	0	0
All	All	20407	0	18532	119	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 3.

The worst 5 of 119 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1288:GLY:H	5:B:2001:PEG:H32	1.42	0.85
1:B:1233:ASP:OD2	5:B:2001:PEG:H12	1.84	0.77
1:B:1287:ARG:HG3	5:B:2001:PEG:H31	1.75	0.68
1:B:1145:HIS:HE1	8:B:2347:HOH:O	1.79	0.66
3:F:3199:ASN:ND2	3:F:3201:GLY:H	1.94	0.66

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

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
8:B:2102:HOH:O	8:E:3425:HOH:O[2_556]	2.15	0.05

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	562/570 (99%)	542 (96%)	19 (3%)	1 (0%)	44	42
1	B	563/570 (99%)	547 (97%)	15 (3%)	1 (0%)	44	42
2	C	430/434 (99%)	425 (99%)	4 (1%)	1 (0%)	44	42
2	D	430/434 (99%)	424 (99%)	5 (1%)	1 (0%)	44	42
3	E	245/321 (76%)	239 (98%)	6 (2%)	0	100	100
3	F	245/321 (76%)	239 (98%)	6 (2%)	0	100	100
All	All	2475/2650 (93%)	2416 (98%)	55 (2%)	4 (0%)	44	42

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	2104	LYS
1	A	1525	VAL
1	B	1525	VAL
2	C	2400	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	440/440 (100%)	437 (99%)	3 (1%)	81	86
1	B	441/440 (100%)	434 (98%)	7 (2%)	58	64
2	C	325/327 (99%)	323 (99%)	2 (1%)	84	88
2	D	325/327 (99%)	324 (100%)	1 (0%)	91	94
3	E	199/251 (79%)	196 (98%)	3 (2%)	60	66
3	F	199/251 (79%)	196 (98%)	3 (2%)	60	66
All	All	1929/2036 (95%)	1910 (99%)	19 (1%)	75	78

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	E	3122	GLN
3	F	3176	ARG
3	F	3199	ASN
3	F	3122	GLN
1	B	1372	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
3	F	3199	ASN
3	F	3122	GLN
2	D	2018	ASN
3	E	3239	GLN
2	C	2402	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

12 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	DYA	B	1470	1	6,7,8	2.12	2 (33%)	6,8,10	1.69	1 (16%)
1	AGM	A	1285	1	10,11,12	0.43	0	7,13,15	0.74	0
1	SMC	A	1472	1	5,6,7	0.61	0	3,6,8	0.71	0
1	AGM	B	1285	1	10,11,12	0.41	0	7,13,15	0.81	0
1	GL3	A	1465	1	2,3,4	0.92	0	1,2,4	0.59	0
1	MGN	B	1420	1	6,9,10	0.50	0	7,12,14	0.62	0
1	GL3	B	1465	1	2,3,4	0.70	0	1,2,4	0.39	0
1	SMC	B	1472	1	5,6,7	0.58	0	3,6,8	0.88	0
1	MHS	B	1271	1	7,11,12	0.70	0	7,14,16	1.12	1 (14%)
1	MHS	A	1271	1	7,11,12	0.69	0	7,14,16	1.21	1 (14%)
1	MGN	A	1420	1	6,9,10	0.48	0	7,12,14	0.66	0
1	DYA	A	1470	1	6,7,8	2.11	2 (33%)	6,8,10	2.00	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	DYA	B	1470	1	-	2/4/6/8	-
1	AGM	A	1285	1	-	2/10/11/13	-
1	SMC	A	1472	1	-	1/3/5/7	-
1	AGM	B	1285	1	-	1/10/11/13	-
1	GL3	A	1465	1	-	0/1/1/2	-
1	MGN	B	1420	1	-	0/7/9/12	-
1	GL3	B	1465	1	-	1/1/1/2	-
1	SMC	B	1472	1	-	1/3/5/7	-
1	MHS	B	1271	1	-	0/5/6/8	0/1/1/1
1	MHS	A	1271	1	-	0/5/6/8	0/1/1/1
1	MGN	A	1420	1	-	0/7/9/12	-
1	DYA	A	1470	1	-	3/4/6/8	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1470	DYA	C-CA	-4.39	1.37	1.45
1	B	1470	DYA	C-CA	-4.28	1.37	1.45

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	1470	DYA	OD1-CG	-2.55	1.23	1.30
1	A	1470	DYA	OD1-CG	-2.44	1.24	1.30

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1470	DYA	CB-CA-N	-3.43	117.02	123.25
1	B	1470	DYA	CB-CA-N	-3.32	117.21	123.25
1	A	1470	DYA	OD2-CG-CB	-2.46	116.50	124.02
1	A	1271	MHS	CM-ND1-CG	2.23	127.41	124.44
1	B	1271	MHS	CB-CG-CD2	2.09	131.48	127.75

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	B	1465	GL3	S-C-CA-N
1	B	1472	SMC	CA-CB-SG-CS
1	A	1470	DYA	CA-CB-CG-OD1
1	A	1470	DYA	CA-CB-CG-OD2
1	B	1470	DYA	CA-CB-CG-OD2

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	B	1465	GL3	1	0
1	A	1271	MHS	1	0

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

5 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The

Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
7	SHT	B	2003	4	26,27,27	1.27	2 (7%)	32,36,36	0.72	0
5	PEG	B	2001	-	6,6,6	0.12	0	5,5,5	0.14	0
4	F43	A	1601	1,7	63,71,71	2.99	12 (19%)	73,118,118	2.39	21 (28%)
6	TP7	B	2002	-	19,20,20	0.72	0	24,26,26	0.57	0
4	F43	A	1602	1	63,71,71	3.03	12 (19%)	73,118,118	2.48	19 (26%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	SHT	B	2003	4	-	6/31/31/31	-
5	PEG	B	2001	-	-	1/4/4/4	-
4	F43	A	1601	1,7	4/4/30/32	8/28/185/185	-
6	TP7	B	2002	-	-	2/24/24/24	-
4	F43	A	1602	1	4/4/30/32	8/28/185/185	-

The worst 5 of 26 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1602	F43	C4D-ND	-11.94	1.31	1.49
4	A	1601	F43	C4D-ND	-11.84	1.31	1.49
4	A	1601	F43	NI-NB	10.44	2.14	1.89
4	A	1602	F43	NI-NB	9.84	2.13	1.89
4	A	1601	F43	NI-ND	8.46	2.09	1.89

The worst 5 of 40 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	1602	F43	C5D-C2D-C1D	12.74	127.06	110.43
4	A	1601	F43	C5D-C2D-C1D	11.82	125.86	110.43
4	A	1601	F43	C5D-C2D-C3D	6.41	129.96	118.26
4	A	1602	F43	C5D-C2D-C3D	5.94	129.08	118.26
4	A	1602	F43	C3D-C4D-ND	5.15	110.33	102.34

5 of 8 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	A	1601	F43	C4A
4	A	1601	F43	C2D
4	A	1601	F43	C3D
4	A	1601	F43	C4D
4	A	1602	F43	C4A

5 of 25 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1601	F43	C2D-C3D-C9D-CAD
4	A	1602	F43	C2D-C3D-C9D-CAD
7	B	2003	SHT	CC-CD-SG2-OS3
7	B	2003	SHT	CC-CD-SG2-OS1
7	B	2003	SHT	CC-CD-SG2-OS2

There are no ring outliers.

5 monomers are involved in 30 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	B	2003	SHT	5	0
5	B	2001	PEG	9	0
4	A	1601	F43	8	0
6	B	2002	TP7	1	0
4	A	1602	F43	7	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 SHT B 2003

Bond lengths

Bond angles

Torsions

Rings

Ligand F43 A 1601

Bond lengths

Bond angles

Torsions

Rings

Ligand TP7 B 2002

Bond lengths

Bond angles

Torsions

Rings

Ligand F43 A 1602

Bond lengths

Bond angles

Torsions

Rings
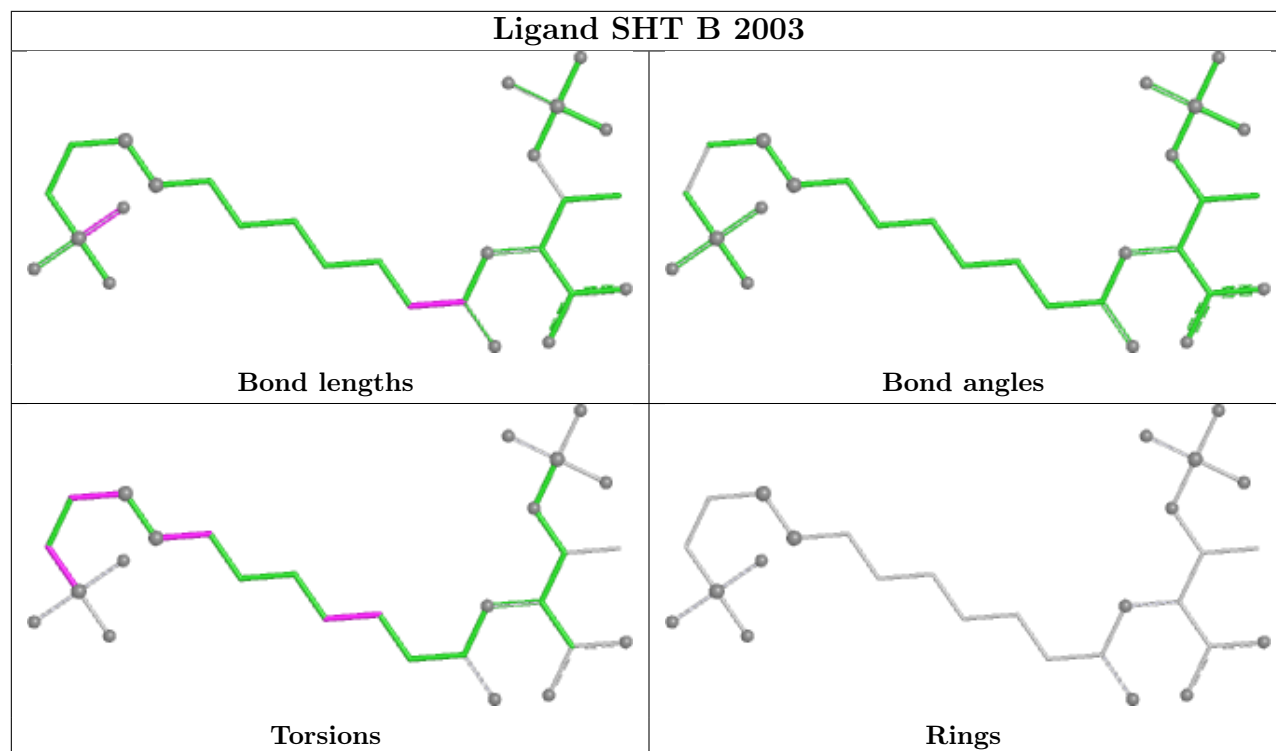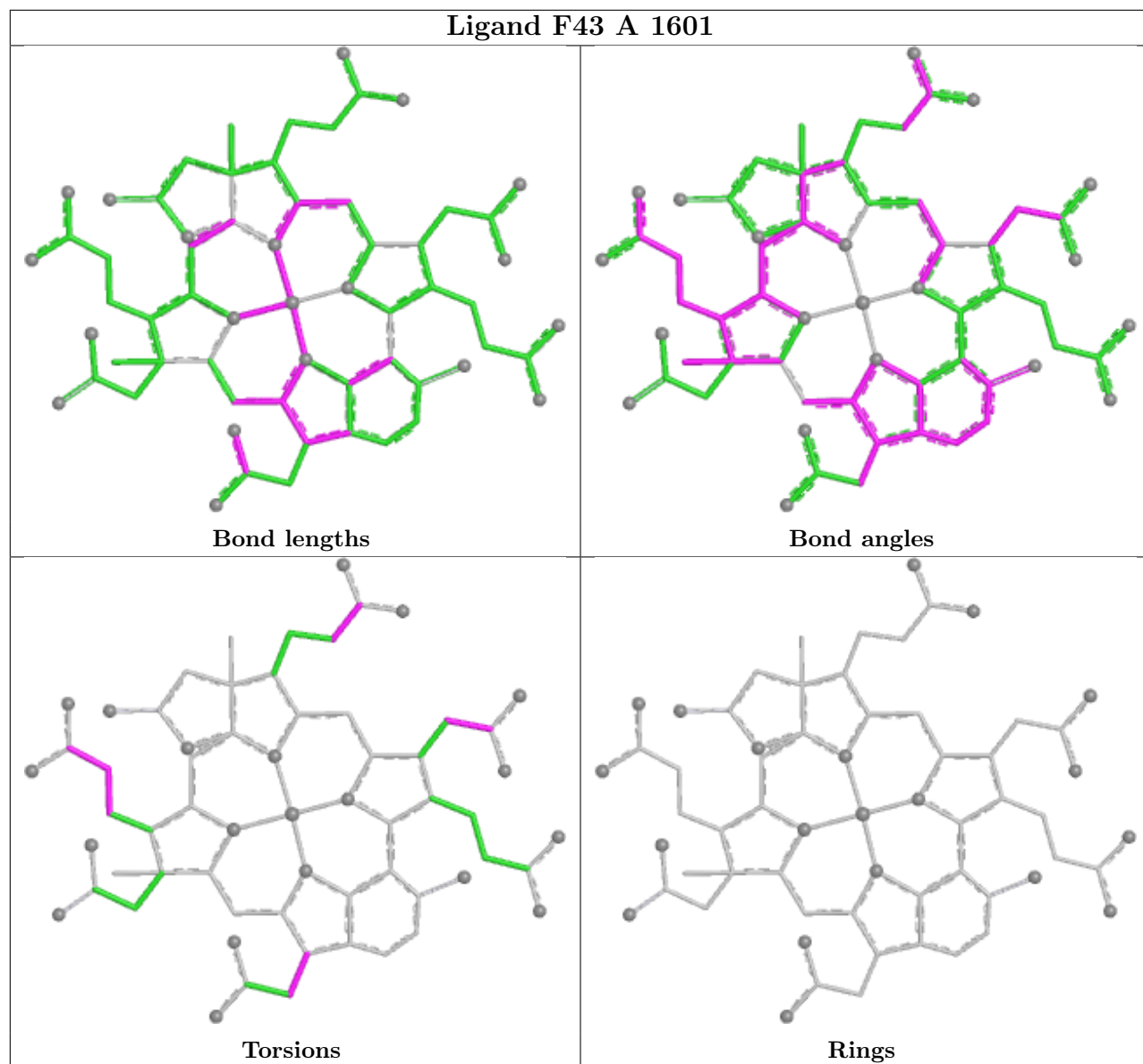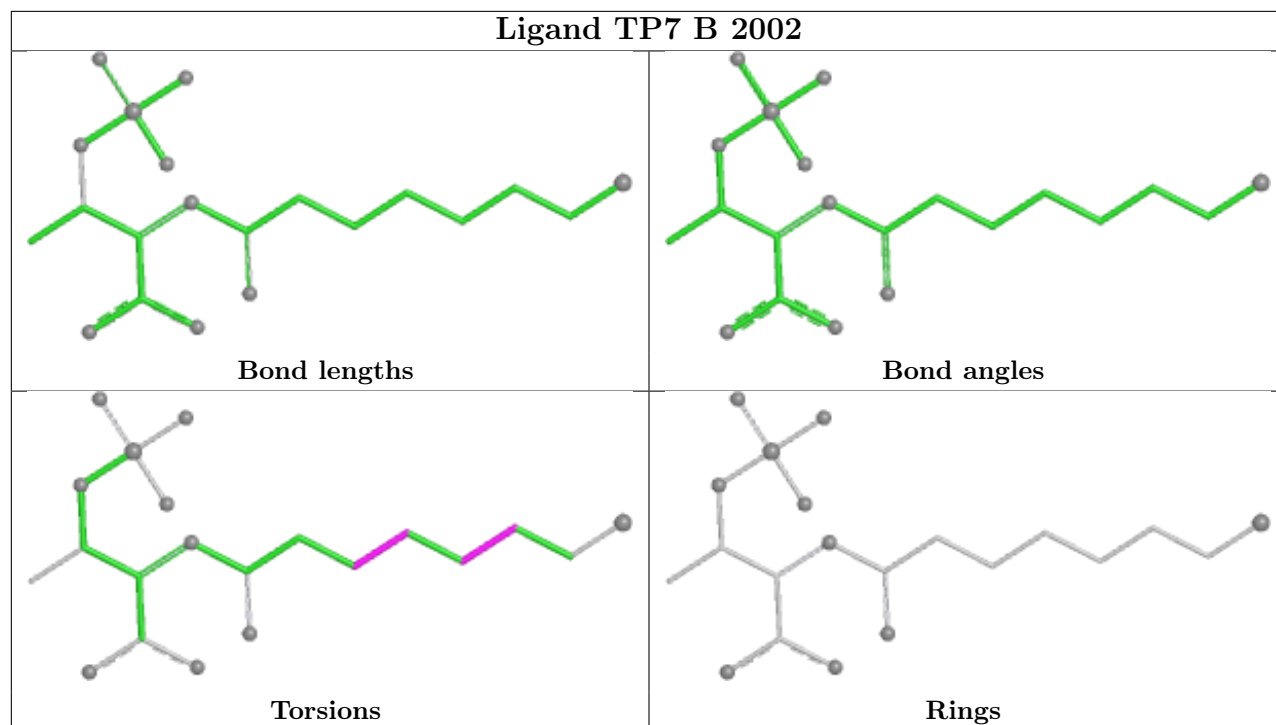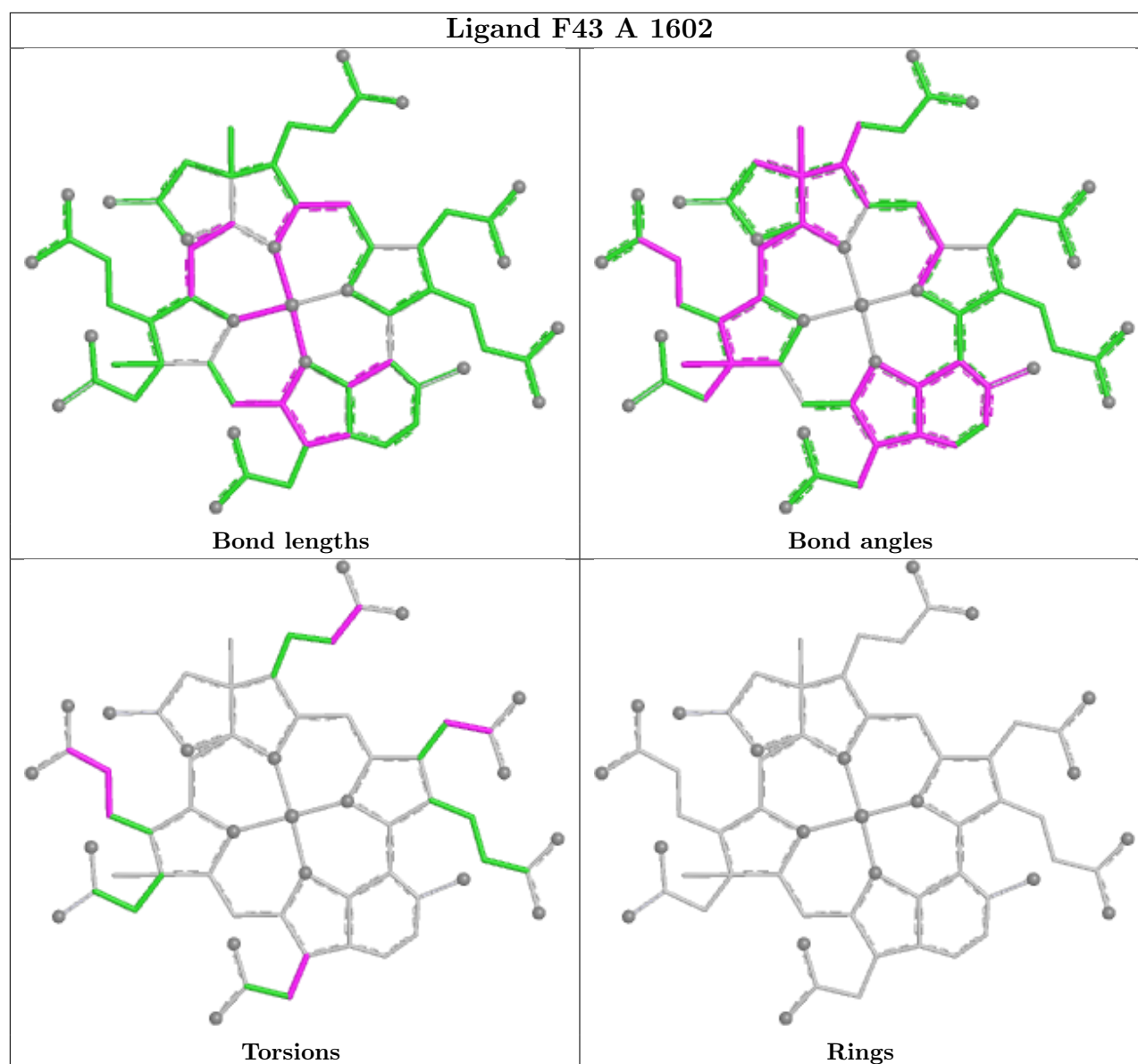
5.7 Other polymers ⓘ

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	562/570 (98%)	-0.71	4 (0%)	84	83	18, 25, 39, 55	2 (0%)
1	B	562/570 (98%)	-0.80	0	100	100	18, 23, 36, 53	3 (0%)
2	C	432/434 (99%)	-0.72	1 (0%)	92	91	17, 24, 36, 52	0
2	D	432/434 (99%)	-0.67	1 (0%)	92	91	19, 25, 40, 59	0
3	E	247/321 (76%)	-0.63	0	100	100	21, 27, 38, 56	0
3	F	247/321 (76%)	-0.71	0	100	100	19, 26, 38, 53	0
All	All	2482/2650 (93%)	-0.72	6 (0%)	92	91	17, 25, 38, 59	5 (0%)

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	2433	ALA	2.8
1	A	1567	ILE	2.7
1	A	1002	ALA	2.2
1	A	1003	ALA	2.2
2	C	2002	SER	2.1

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	1271	11/12	0.96	0.05	21,23,23,24	0
1	DYA	A	1470	8/9	0.96	0.05	21,23,24,24	0
1	SMC	A	1472	7/8	0.96	0.07	20,21,22,24	0
1	AGM	B	1285	12/13	0.96	0.05	19,20,21,22	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($Å^2$)	Q<0.9
1	DYA	B	1470	8/9	0.96	0.05	21,22,23,26	0
1	MHS	B	1271	11/12	0.97	0.05	20,21,23,23	0
1	AGM	A	1285	12/13	0.97	0.04	17,18,19,19	0
1	MGN	B	1420	10/11	0.97	0.05	20,21,21,22	0
1	MGN	A	1420	10/11	0.97	0.04	20,21,23,24	0
1	GL3	B	1465	4/5	0.98	0.05	19,20,21,22	0
1	SMC	B	1472	7/8	0.98	0.06	20,20,22,23	0
1	GL3	A	1465	4/5	0.99	0.04	20,20,21,21	0

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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
5	PEG	B	2001	7/7	0.86	0.13	49,52,55,58	0
4	F43	A	1602	62/62	0.96	0.07	21,24,26,29	0
4	F43	A	1601	62/62	0.96	0.06	18,23,26,31	0
6	TP7	B	2002	21/21	0.97	0.06	20,21,23,28	0
7	SHT	B	2003	28/28	0.97	0.06	20,23,38,41	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.
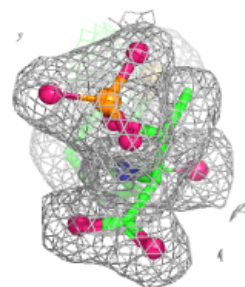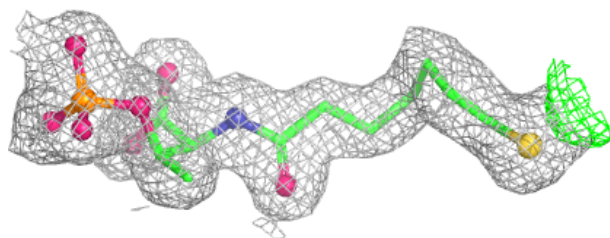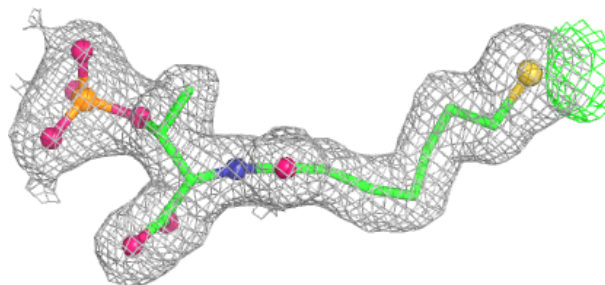
Electron density around F43 A 1602:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
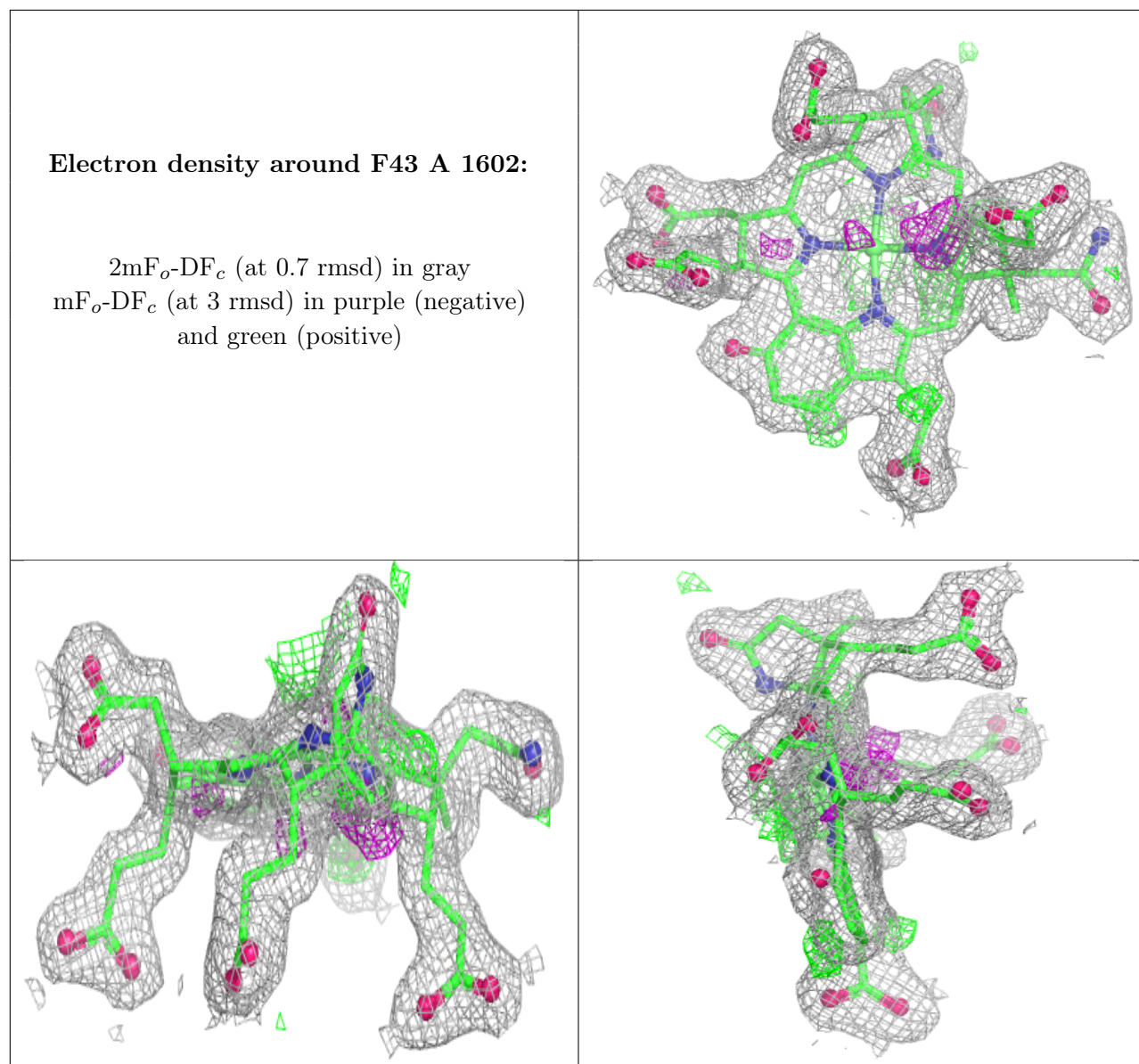
Electron density around F43 A 1601:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
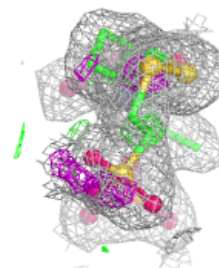mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
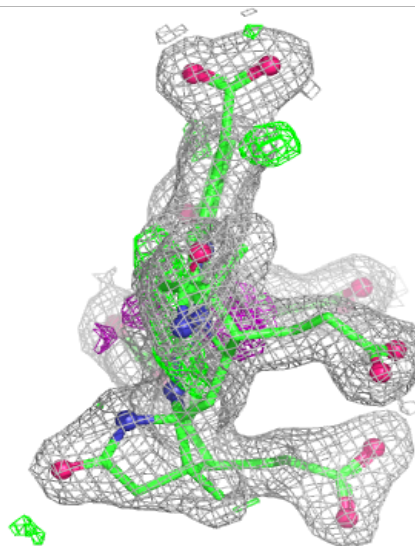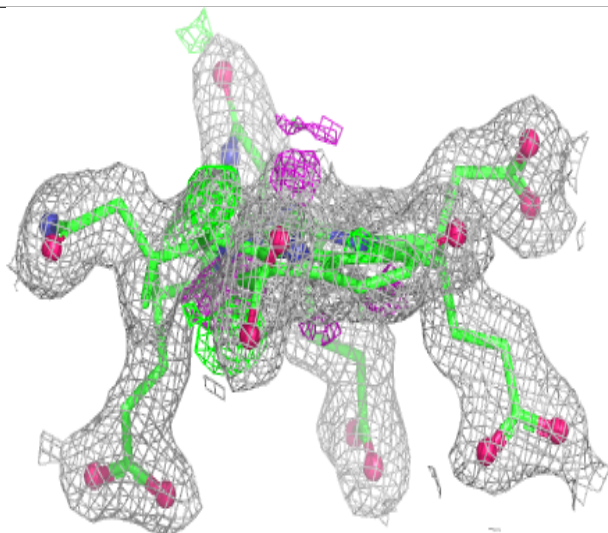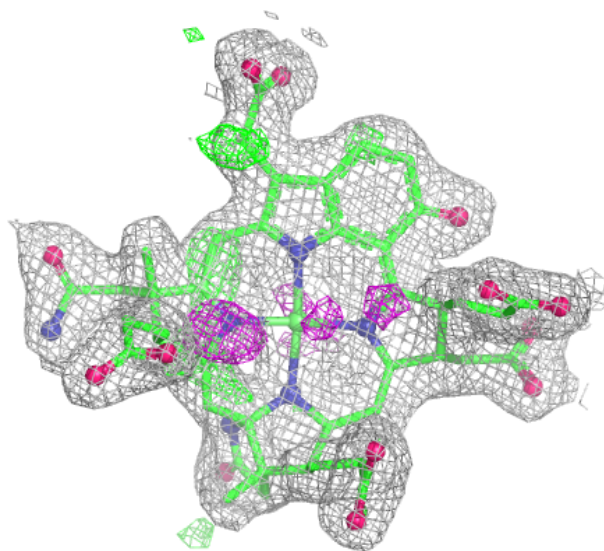
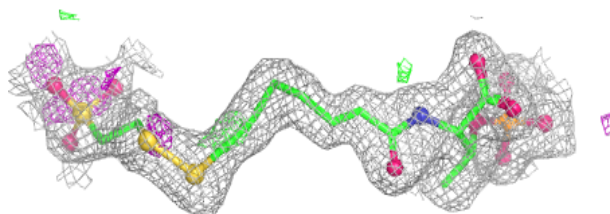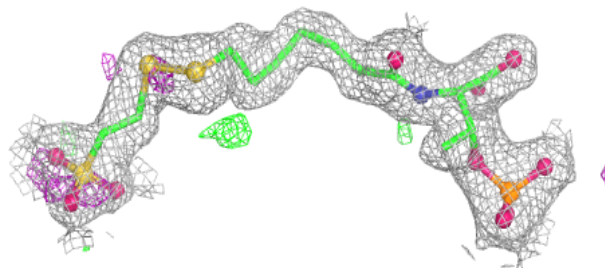**Electron density around TP7 B 2002:**

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

Electron density around SHT B 2003:

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

6.5 Other polymers (i)

There are no such residues in this entry.