



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 14, 2020 – 10:48 AM BST

PDB ID : 6UBV  
Title : The structure of the Streptococcus gordonii surface protein SspB in complex with TEV peptide provides clues to the adherence of oral streptococcal adherence to salivary agglutinin  
Authors : Schormann, N.; Deivanayagam, C.  
Deposited on : 2019-09-13  
Resolution : 2.70 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](mailto: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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.14.4.dev1  
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) : 2.14.4.dev1

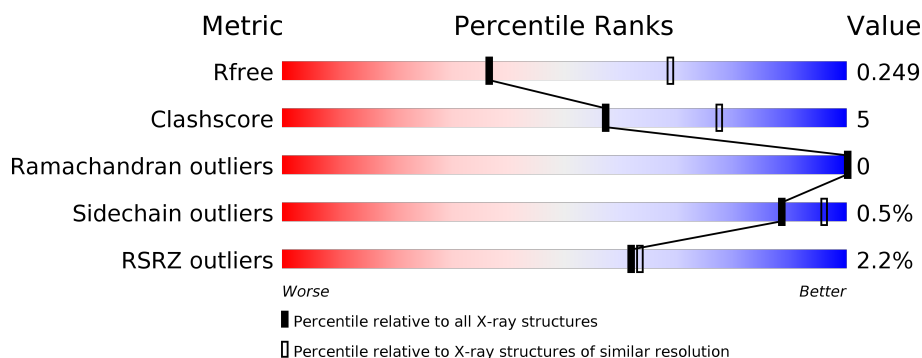
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	425	<div> <div>2%</div> <div> <div></div> <div>82%</div> <div>10%</div> <div>8%</div> </div> </div>
1	B	425	<div> <div>2%</div> <div> <div></div> <div>82%</div> <div>10%</div> <div>7%</div> </div> </div>
1	C	425	<div> <div>2%</div> <div> <div></div> <div>83%</div> <div>10%</div> <div>6%</div> </div> </div>
1	D	425	<div> <div>2%</div> <div> <div></div> <div>80%</div> <div>11%</div> <div>9%</div> </div> </div>

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	SNZ	A	902	-	-	-	X
3	SNZ	B	902	-	-	-	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 12619 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 Surface protein adhesin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	392	Total	C	N	O	S	0	0	0
			3011	1894	501	608	8			
1	B	394	Total	C	N	O	S	0	0	0
			3042	1916	504	614	8			
1	C	399	Total	C	N	O	S	0	0	0
			3079	1940	509	622	8			
1	D	388	Total	C	N	O	S	0	0	0
			2990	1881	496	605	8			

There are 88 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	433	MET	-	expression tag	UNP C9E3B4
A	434	ALA	-	expression tag	UNP C9E3B4
A	435	SER	-	expression tag	UNP C9E3B4
A	436	MET	-	expression tag	UNP C9E3B4
A	437	THR	-	expression tag	UNP C9E3B4
A	438	GLY	-	expression tag	UNP C9E3B4
A	439	GLY	-	expression tag	UNP C9E3B4
A	440	GLN	-	expression tag	UNP C9E3B4
A	441	GLN	-	expression tag	UNP C9E3B4
A	442	MET	-	expression tag	UNP C9E3B4
A	443	GLY	-	expression tag	UNP C9E3B4
A	444	ARG	-	expression tag	UNP C9E3B4
A	445	ILE	-	expression tag	UNP C9E3B4
A	446	GLN	-	expression tag	UNP C9E3B4
A	850	LEU	-	expression tag	UNP C9E3B4
A	851	GLU	-	expression tag	UNP C9E3B4
A	852	HIS	-	expression tag	UNP C9E3B4
A	853	HIS	-	expression tag	UNP C9E3B4
A	854	HIS	-	expression tag	UNP C9E3B4
A	855	HIS	-	expression tag	UNP C9E3B4
A	856	HIS	-	expression tag	UNP C9E3B4

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	857	HIS	-	expression tag	UNP C9E3B4
B	433	MET	-	expression tag	UNP C9E3B4
B	434	ALA	-	expression tag	UNP C9E3B4
B	435	SER	-	expression tag	UNP C9E3B4
B	436	MET	-	expression tag	UNP C9E3B4
B	437	THR	-	expression tag	UNP C9E3B4
B	438	GLY	-	expression tag	UNP C9E3B4
B	439	GLY	-	expression tag	UNP C9E3B4
B	440	GLN	-	expression tag	UNP C9E3B4
B	441	GLN	-	expression tag	UNP C9E3B4
B	442	MET	-	expression tag	UNP C9E3B4
B	443	GLY	-	expression tag	UNP C9E3B4
B	444	ARG	-	expression tag	UNP C9E3B4
B	445	ILE	-	expression tag	UNP C9E3B4
B	446	GLN	-	expression tag	UNP C9E3B4
B	850	LEU	-	expression tag	UNP C9E3B4
B	851	GLU	-	expression tag	UNP C9E3B4
B	852	HIS	-	expression tag	UNP C9E3B4
B	853	HIS	-	expression tag	UNP C9E3B4
B	854	HIS	-	expression tag	UNP C9E3B4
B	855	HIS	-	expression tag	UNP C9E3B4
B	856	HIS	-	expression tag	UNP C9E3B4
B	857	HIS	-	expression tag	UNP C9E3B4
C	433	MET	-	expression tag	UNP C9E3B4
C	434	ALA	-	expression tag	UNP C9E3B4
C	435	SER	-	expression tag	UNP C9E3B4
C	436	MET	-	expression tag	UNP C9E3B4
C	437	THR	-	expression tag	UNP C9E3B4
C	438	GLY	-	expression tag	UNP C9E3B4
C	439	GLY	-	expression tag	UNP C9E3B4
C	440	GLN	-	expression tag	UNP C9E3B4
C	441	GLN	-	expression tag	UNP C9E3B4
C	442	MET	-	expression tag	UNP C9E3B4
C	443	GLY	-	expression tag	UNP C9E3B4
C	444	ARG	-	expression tag	UNP C9E3B4
C	445	ILE	-	expression tag	UNP C9E3B4
C	446	GLN	-	expression tag	UNP C9E3B4
C	850	LEU	-	expression tag	UNP C9E3B4
C	851	GLU	-	expression tag	UNP C9E3B4
C	852	HIS	-	expression tag	UNP C9E3B4
C	853	HIS	-	expression tag	UNP C9E3B4
C	854	HIS	-	expression tag	UNP C9E3B4

*Continued on next page...*

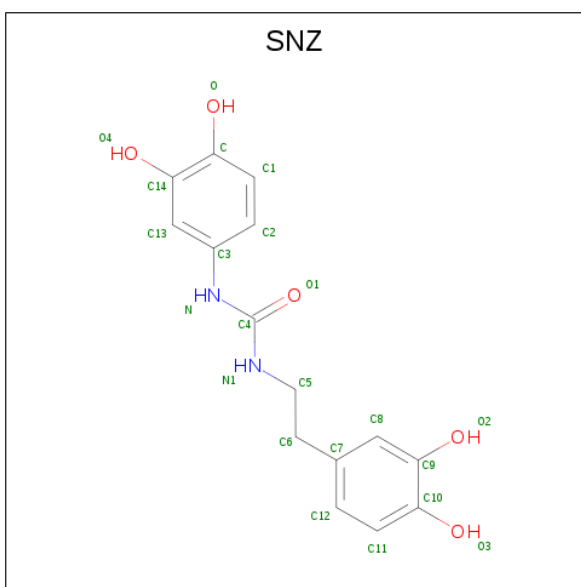
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	855	HIS	-	expression tag	UNP C9E3B4
C	856	HIS	-	expression tag	UNP C9E3B4
C	857	HIS	-	expression tag	UNP C9E3B4
D	433	MET	-	expression tag	UNP C9E3B4
D	434	ALA	-	expression tag	UNP C9E3B4
D	435	SER	-	expression tag	UNP C9E3B4
D	436	MET	-	expression tag	UNP C9E3B4
D	437	THR	-	expression tag	UNP C9E3B4
D	438	GLY	-	expression tag	UNP C9E3B4
D	439	GLY	-	expression tag	UNP C9E3B4
D	440	GLN	-	expression tag	UNP C9E3B4
D	441	GLN	-	expression tag	UNP C9E3B4
D	442	MET	-	expression tag	UNP C9E3B4
D	443	GLY	-	expression tag	UNP C9E3B4
D	444	ARG	-	expression tag	UNP C9E3B4
D	445	ILE	-	expression tag	UNP C9E3B4
D	446	GLN	-	expression tag	UNP C9E3B4
D	850	LEU	-	expression tag	UNP C9E3B4
D	851	GLU	-	expression tag	UNP C9E3B4
D	852	HIS	-	expression tag	UNP C9E3B4
D	853	HIS	-	expression tag	UNP C9E3B4
D	854	HIS	-	expression tag	UNP C9E3B4
D	855	HIS	-	expression tag	UNP C9E3B4
D	856	HIS	-	expression tag	UNP C9E3B4
D	857	HIS	-	expression tag	UNP C9E3B4

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca) (labeled as "Ligand of Interest" by author).

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

- Molecule 3 is N-(3,4-dihydroxyphenyl)-N'-[2-(3,4-dihydroxyphenyl)ethyl]urea (three-letter code: SNZ) (formula: C<sub>15</sub>H<sub>16</sub>N<sub>2</sub>O<sub>5</sub>) (labeled as "Ligand of Interest" by author).

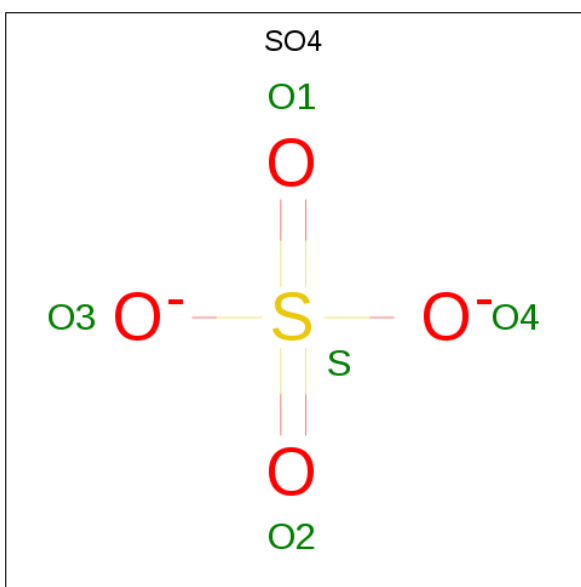


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total 22	C 15	N 2	O 5	0	0
3	B	1	Total 22	C 15	N 2	O 5	0	0
3	C	1	Total 22	C 15	N 2	O 5	0	0
3	D	1	Total 22	C 15	N 2	O 5	0	0

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na) (labeled as "Ligand of Interest" by author).

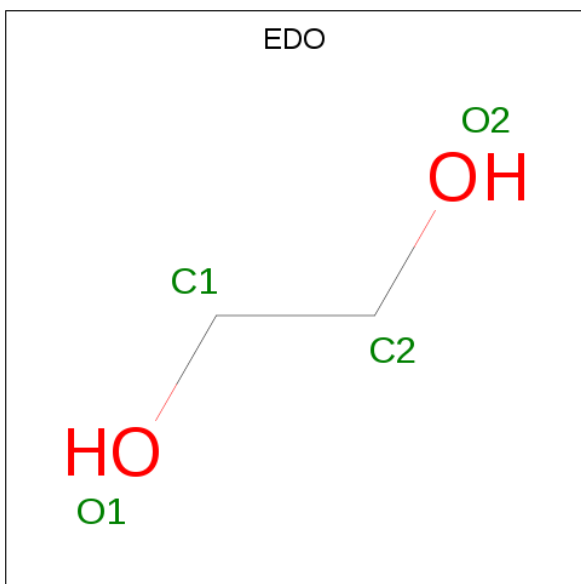
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Na 1 1	0	0
4	D	1	Total Na 1 1	0	0
4	C	1	Total Na 1 1	0	0

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S) (labeled as "Ligand of Interest" by author).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	O	S	0	0
			5	4	1		
5	D	1	Total	O	S	0	0
			5	4	1		

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ) (labeled as "Ligand of Interest" by author).



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

*Continued on next page...*

*Continued from previous page...*

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

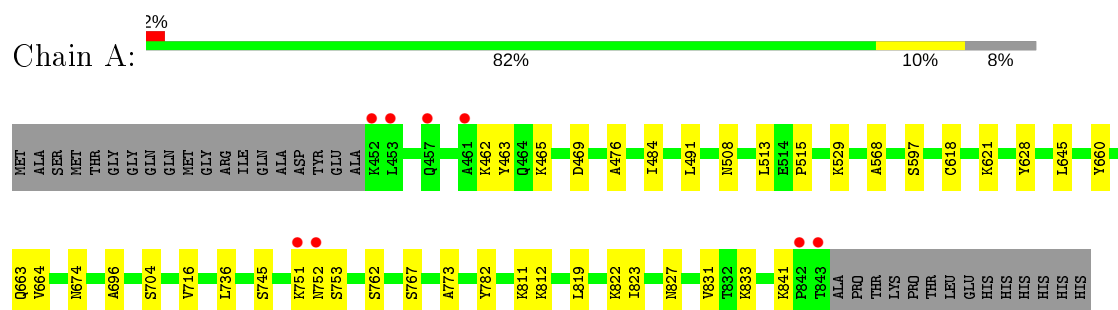
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	115	Total	O	0	0
			115	115		
7	B	69	Total	O	0	0
			69	69		
7	C	114	Total	O	0	0
			114	114		
7	D	82	Total	O	0	0
			82	82		

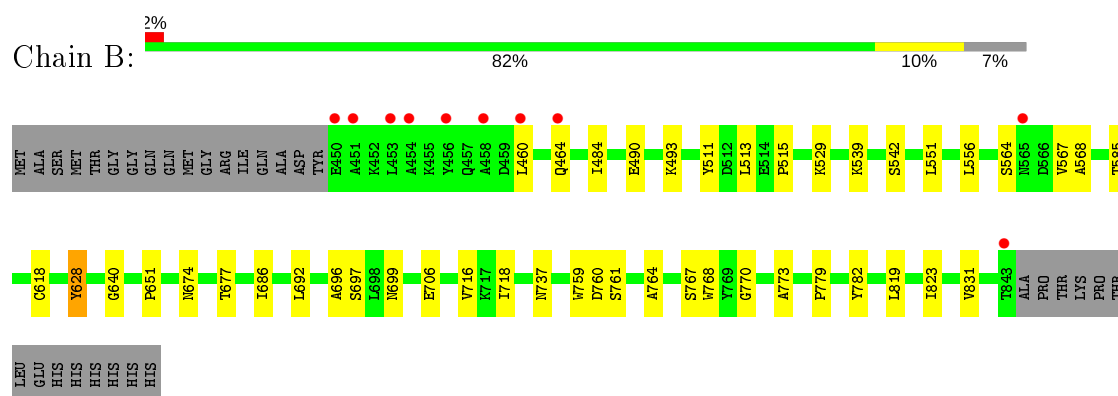
### 3 Residue-property plots [i](#)

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.

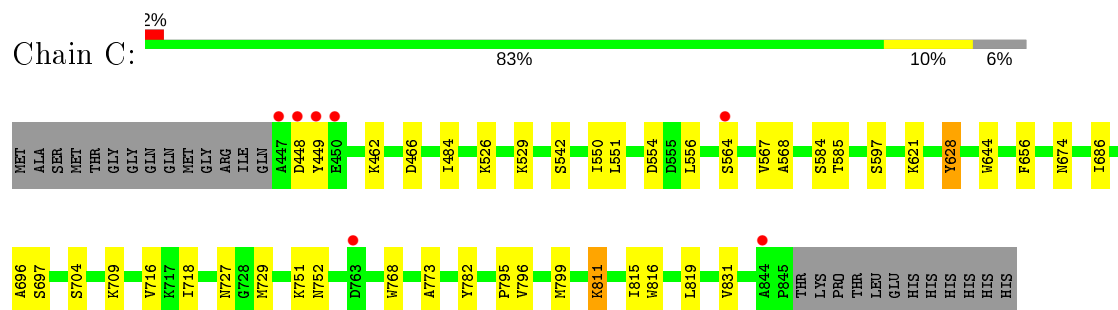
- Molecule 1: Surface protein adhesin



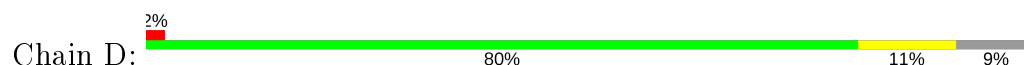
- Molecule 1: Surface protein adhesin

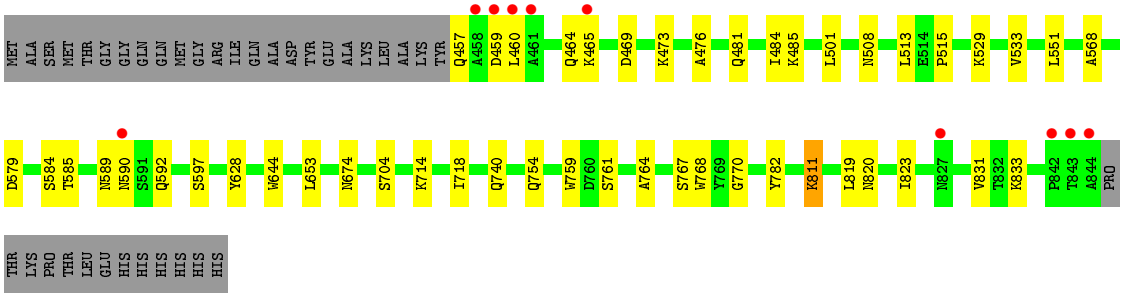


- Molecule 1: Surface protein adhesin



- Molecule 1: Surface protein adhesin





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	65.81Å 132.70Å 243.82Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.72 – 2.70 29.72 – 2.70	Depositor EDS
% Data completeness (in resolution range)	94.0 (29.72-2.70) 94.0 (29.72-2.70)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.71 (at 2.68Å)	Xtriage
Refinement program	PHENIX 1.16 _3549	Depositor
R, $R_{free}$	0.206 , 0.250 0.206 , 0.249	Depositor DCC
$R_{free}$ test set	2738 reflections (4.88%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.7	Xtriage
Anisotropy	0.138	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 26.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	12619	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.57% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: NA, CA, SNZ, EDO, SO4

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	0/3073	0.52	0/4162
1	B	0.33	0/3105	0.54	0/4205
1	C	0.32	0/3144	0.51	0/4260
1	D	0.32	0/3052	0.51	0/4135
All	All	0.32	0/12374	0.52	0/16762

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	3011	0	2936	25	0
1	B	3042	0	2978	30	0
1	C	3079	0	3009	36	0
1	D	2990	0	2921	34	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	22	0	0	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	22	0	0	6	0
3	C	22	0	0	6	0
3	D	22	0	0	2	0
4	A	1	0	0	0	0
4	C	1	0	0	0	0
4	D	1	0	0	0	0
5	B	5	0	0	0	0
5	D	5	0	0	1	0
6	B	4	0	6	0	0
6	C	4	0	6	2	0
6	D	4	0	6	0	0
7	A	115	0	0	1	0
7	B	69	0	0	1	0
7	C	114	0	0	0	0
7	D	82	0	0	0	0
All	All	12619	0	11862	118	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 5.

All (118) 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:B:460:LEU:O	1:B:464:GLN:HG2	1.55	1.06
1:C:696:ALA:HB1	3:C:902:SNZ:C1	2.07	0.85
1:B:542:SER:OG	1:B:556:LEU:HA	1.81	0.81
1:D:761:SER:HB2	1:D:764:ALA:HB2	1.66	0.77
1:C:529:LYS:HG2	1:C:568:ALA:HB2	1.65	0.77
1:B:761:SER:HB3	1:B:764:ALA:HB2	1.68	0.74
1:C:448:ASP:OD1	1:C:448:ASP:O	2.07	0.73
1:A:618:CYS:O	1:A:621:LYS:HB2	1.89	0.73
1:A:704:SER:OG	1:A:811:LYS:NZ	2.23	0.72
1:D:584:SER:HB2	1:D:589:ASN:C	2.12	0.70
1:B:760:ASP:HA	3:B:902:SNZ:C1	2.22	0.69
1:D:584:SER:HB2	1:D:589:ASN:O	1.94	0.68
1:C:564:SER:HB3	1:D:465:LYS:HE3	1.78	0.65
1:A:696:ALA:HB1	3:A:902:SNZ:C13	2.26	0.65
1:B:759:TRP:HE3	3:B:902:SNZ:O	1.80	0.64
1:A:463:TYR:CE2	1:A:841:LYS:HE2	2.32	0.64
1:A:773:ALA:HB2	1:A:823:ILE:HD11	1.80	0.63
1:D:513:LEU:HG	1:D:515:PRO:HD3	1.79	0.63

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:551:LEU:HB2	1:B:585:THR:HG22	1.82	0.62
1:C:727:ASN:HB3	1:D:590:ASN:HD22	1.65	0.62
1:C:564:SER:HB3	1:D:465:LYS:CE	2.29	0.62
1:A:664:VAL:HG12	1:A:812:LYS:HG3	1.81	0.61
1:C:815:ILE:HA	6:C:904:EDO:H12	1.84	0.60
1:C:716:VAL:HB	1:C:773:ALA:HB3	1.83	0.60
1:B:718:ILE:HG21	1:B:768:TRP:HB2	1.84	0.59
1:B:513:LEU:HG	1:B:515:PRO:HD3	1.84	0.59
1:D:704:SER:HB2	1:D:811:LYS:HD2	1.85	0.57
1:D:592:GLN:N	5:D:901:SO4:O4	2.29	0.57
1:A:767:SER:O	3:A:902:SNZ:O	2.24	0.56
1:B:759:TRP:O	3:B:902:SNZ:C1	2.54	0.56
1:C:697:SER:N	3:C:902:SNZ:C1	2.70	0.55
1:A:753:SER:HB2	7:A:1094:HOH:O	2.07	0.55
1:C:550:ILE:HG23	1:C:584:SER:HB3	1.89	0.55
1:D:551:LEU:HB2	1:D:585:THR:HG22	1.87	0.54
1:D:761:SER:O	1:D:767:SER:OG	2.26	0.54
1:B:716:VAL:HB	1:B:773:ALA:HB3	1.90	0.54
1:C:564:SER:CB	1:D:465:LYS:HE3	2.39	0.53
1:C:526:LYS:HD3	1:C:567:VAL:HG21	1.89	0.53
1:C:484:ILE:HD13	1:C:831:VAL:HG22	1.90	0.53
1:B:759:TRP:CE3	3:B:902:SNZ:O	2.60	0.53
1:C:718:ILE:HG21	1:C:768:TRP:HB2	1.90	0.53
1:C:697:SER:HB2	3:C:902:SNZ:C2	2.39	0.52
1:C:697:SER:H	3:C:902:SNZ:C	2.23	0.52
1:B:692:LEU:HD13	1:B:823:ILE:HD13	1.92	0.51
1:C:628:TYR:OH	1:C:819:LEU:HD12	2.10	0.51
1:C:697:SER:HB2	3:C:902:SNZ:C1	2.40	0.51
1:C:704:SER:CB	1:C:811:LYS:HD2	2.41	0.51
1:B:484:ILE:HD13	1:B:831:VAL:HG22	1.92	0.51
1:B:490:GLU:CD	1:B:493:LYS:HE3	2.32	0.51
1:C:816:TRP:H	6:C:904:EDO:H12	1.76	0.50
1:D:529:LYS:O	1:D:533:VAL:HG23	2.11	0.50
1:C:795:PRO:HB3	1:D:579:ASP:HA	1.92	0.50
1:C:727:ASN:HB3	1:D:590:ASN:ND2	2.26	0.50
1:B:767:SER:O	3:B:902:SNZ:O	2.30	0.49
1:D:674:ASN:O	1:D:782:TYR:HA	2.13	0.49
1:A:628:TYR:OH	1:A:819:LEU:HD12	2.12	0.49
1:B:628:TYR:OH	1:B:819:LEU:HD12	2.12	0.49
1:B:529:LYS:HG2	1:B:568:ALA:HB2	1.95	0.49
1:A:660:TYR:CE1	1:A:812:LYS:HG2	2.49	0.48

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:501:LEU:HD21	1:D:823:ILE:HD11	1.94	0.48
1:A:513:LEU:HG	1:A:515:PRO:HD3	1.96	0.48
1:C:462:LYS:NZ	1:C:466:ASP:OD2	2.44	0.48
1:C:704:SER:HB3	1:C:811:LYS:HD2	1.96	0.47
1:D:714:LYS:HE3	1:D:714:LYS:HB3	1.62	0.47
1:D:718:ILE:HG21	1:D:768:TRP:HB2	1.96	0.47
1:C:674:ASN:O	1:C:782:TYR:HA	2.16	0.46
1:C:709:LYS:HB3	1:C:729:MET:HG2	1.96	0.46
1:A:476:ALA:HB1	1:A:833:LYS:HD2	1.97	0.46
1:C:551:LEU:HB2	1:C:585:THR:HG22	1.98	0.46
1:A:491:LEU:HD11	1:A:827:ASN:O	2.15	0.46
1:C:697:SER:N	3:C:902:SNZ:C	2.79	0.46
1:D:469:ASP:OD1	1:D:473:LYS:HE3	2.16	0.46
1:D:460:LEU:O	1:D:464:GLN:HG2	2.16	0.46
1:B:564:SER:O	1:B:567:VAL:HG23	2.16	0.45
1:B:618:CYS:HB3	1:B:686:ILE:HG21	1.98	0.45
1:D:653:LEU:HD23	3:D:903:SNZ:C1	2.47	0.45
1:B:674:ASN:O	1:B:782:TYR:HA	2.17	0.45
1:B:696:ALA:HB1	3:B:902:SNZ:C13	2.47	0.45
1:D:740:GLN:OE1	1:D:754:GLN:NE2	2.50	0.45
1:A:465:LYS:NZ	1:A:469:ASP:OD2	2.51	0.44
1:D:597:SER:HB3	1:D:644:TRP:HE1	1.80	0.44
1:A:597:SER:HA	1:A:645:LEU:O	2.17	0.44
1:A:674:ASN:O	1:A:782:TYR:HA	2.17	0.44
1:A:628:TYR:CE1	1:A:674:ASN:HB3	2.53	0.44
1:B:640:GLY:HA3	7:B:1034:HOH:O	2.17	0.44
1:C:621:LYS:HD2	1:C:686:ILE:HG12	2.00	0.44
1:C:751:LYS:HD3	1:C:752:ASN:N	2.33	0.44
1:B:511:TYR:OH	1:B:651:PRO:HB2	2.18	0.43
1:D:820:ASN:ND2	3:D:903:SNZ:O4	2.47	0.43
1:A:529:LYS:HG2	1:A:568:ALA:HB2	2.00	0.43
1:A:716:VAL:HB	1:A:773:ALA:HB3	1.99	0.43
1:C:796:VAL:HG12	1:C:799:MET:HE3	2.01	0.43
1:A:508:ASN:HA	1:A:831:VAL:HB	2.01	0.43
1:B:539:LYS:O	1:B:542:SER:HB2	2.19	0.43
1:D:759:TRP:CZ3	1:D:770:GLY:HA3	2.54	0.42
1:B:697:SER:OG	1:B:697:SER:O	2.36	0.42
1:B:759:TRP:CZ3	1:B:770:GLY:HA3	2.55	0.42
1:B:677:THR:HG23	1:B:779:PRO:O	2.20	0.42
1:D:508:ASN:HA	1:D:831:VAL:HB	2.01	0.42
1:D:628:TYR:OH	1:D:819:LEU:HD12	2.19	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:476:ALA:HB1	1:D:833:LYS:HD2	2.01	0.42
1:A:484:ILE:HD13	1:A:831:VAL:HG22	2.00	0.42
1:C:628:TYR:CE1	1:C:674:ASN:HB3	2.54	0.42
1:B:699:ASN:O	1:B:706:GLU:HB2	2.20	0.42
1:C:564:SER:HB3	1:D:465:LYS:HE2	2.01	0.41
1:D:484:ILE:HD13	1:D:831:VAL:HG22	2.02	0.41
1:B:490:GLU:OE2	1:B:493:LYS:HE3	2.20	0.41
1:D:457:GLN:HB2	1:D:459:ASP:OD2	2.20	0.41
1:C:597:SER:HB3	1:C:644:TRP:HE1	1.86	0.41
1:C:554:ASP:HA	1:C:656:PHE:HZ	1.86	0.41
1:A:751:LYS:HD2	1:A:752:ASN:H	1.85	0.41
1:B:737:ASN:O	1:B:759:TRP:HZ2	2.04	0.41
1:D:481:GLN:O	1:D:485:LYS:HG3	2.21	0.41
1:A:762:SER:HB2	1:A:822:LYS:NZ	2.37	0.40
1:D:529:LYS:HG2	1:D:568:ALA:HB2	2.03	0.40
1:A:462:LYS:HE2	1:A:462:LYS:HB3	2.00	0.40
1:C:542:SER:HB2	1:C:556:LEU:HA	2.04	0.40
1:A:736:LEU:O	1:A:745:SER:HB2	2.22	0.40

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	390/425 (92%)	378 (97%)	12 (3%)	0	100	100
1	B	392/425 (92%)	381 (97%)	11 (3%)	0	100	100
1	C	397/425 (93%)	387 (98%)	10 (2%)	0	100	100
1	D	386/425 (91%)	377 (98%)	9 (2%)	0	100	100
All	All	1565/1700 (92%)	1523 (97%)	42 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	328/358 (92%)	327 (100%)	1 (0%)	92	98
1	B	333/358 (93%)	332 (100%)	1 (0%)	92	98
1	C	336/358 (94%)	333 (99%)	3 (1%)	78	92
1	D	328/358 (92%)	327 (100%)	1 (0%)	92	98
All	All	1325/1432 (92%)	1319 (100%)	6 (0%)	88	96

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	663	GLN
1	B	628	TYR
1	C	449	TYR
1	C	628	TYR
1	C	811	LYS
1	D	811	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry

Of 16 ligands modelled in this entry, 7 are monoatomic - leaving 9 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$
5	SO4	D	901	-	4,4,4	0.87	0	6,6,6	0.17	0
6	EDO	C	904	2	3,3,3	0.06	0	2,2,2	0.19	0
6	EDO	B	904	2	3,3,3	0.06	0	2,2,2	0.18	0
6	EDO	D	905	2	3,3,3	0.05	0	2,2,2	0.19	0
3	SNZ	C	902	-	23,23,23	0.44	0	31,31,31	0.75	0
3	SNZ	B	902	-	23,23,23	0.26	0	31,31,31	0.47	0
3	SNZ	A	902	-	23,23,23	0.18	0	31,31,31	0.46	0
3	SNZ	D	903	-	23,23,23	0.21	0	31,31,31	0.98	4 (12%)
5	SO4	B	903	-	4,4,4	0.14	0	6,6,6	0.07	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
6	EDO	C	904	2	-	1/1/1/1	-
6	EDO	B	904	2	-	1/1/1/1	-
6	EDO	D	905	2	-	0/1/1/1	-
3	SNZ	C	902	-	-	5/10/10/10	0/2/2/2
3	SNZ	B	902	-	-	5/10/10/10	0/2/2/2
3	SNZ	A	902	-	-	6/10/10/10	0/2/2/2
3	SNZ	D	903	-	-	5/10/10/10	0/2/2/2

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	903	SNZ	C2-C3-N	-2.35	112.51	120.40
3	D	903	SNZ	C3-N-C4	2.29	131.30	126.61

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	903	SNZ	C13-C3-N	2.28	127.62	120.18
3	D	903	SNZ	C5-C6-C7	2.19	117.94	112.87

There are no chirality outliers.

All (23) torsion outliers are listed below:

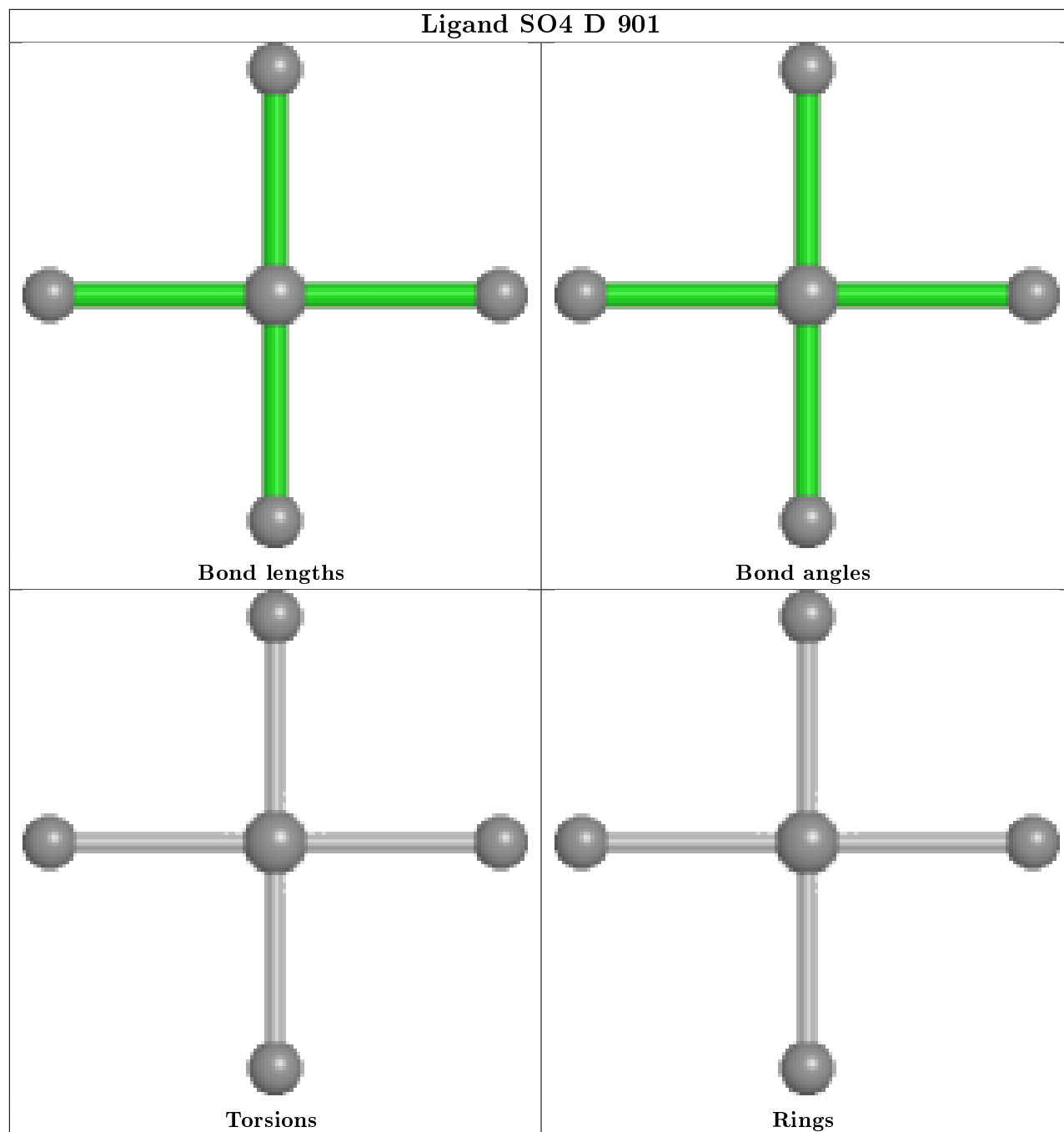
Mol	Chain	Res	Type	Atoms
3	C	902	SNZ	N1-C4-N-C3
3	C	902	SNZ	O1-C4-N-C3
3	C	902	SNZ	N-C4-N1-C5
3	C	902	SNZ	O1-C4-N1-C5
3	B	902	SNZ	N1-C5-C6-C7
3	A	902	SNZ	N1-C5-C6-C7
3	D	903	SNZ	N-C4-N1-C5
3	D	903	SNZ	N1-C5-C6-C7
3	A	902	SNZ	N1-C4-N-C3
3	B	902	SNZ	N1-C4-N-C3
3	A	902	SNZ	O1-C4-N-C3
3	D	903	SNZ	O1-C4-N1-C5
3	B	902	SNZ	O1-C4-N-C3
3	A	902	SNZ	N-C4-N1-C5
3	B	902	SNZ	C6-C5-N1-C4
3	A	902	SNZ	C6-C5-N1-C4
6	C	904	EDO	O1-C1-C2-O2
3	C	902	SNZ	N1-C5-C6-C7
6	B	904	EDO	O1-C1-C2-O2
3	D	903	SNZ	C5-C6-C7-C8
3	D	903	SNZ	C5-C6-C7-C12
3	A	902	SNZ	O1-C4-N1-C5
3	B	902	SNZ	N-C4-N1-C5

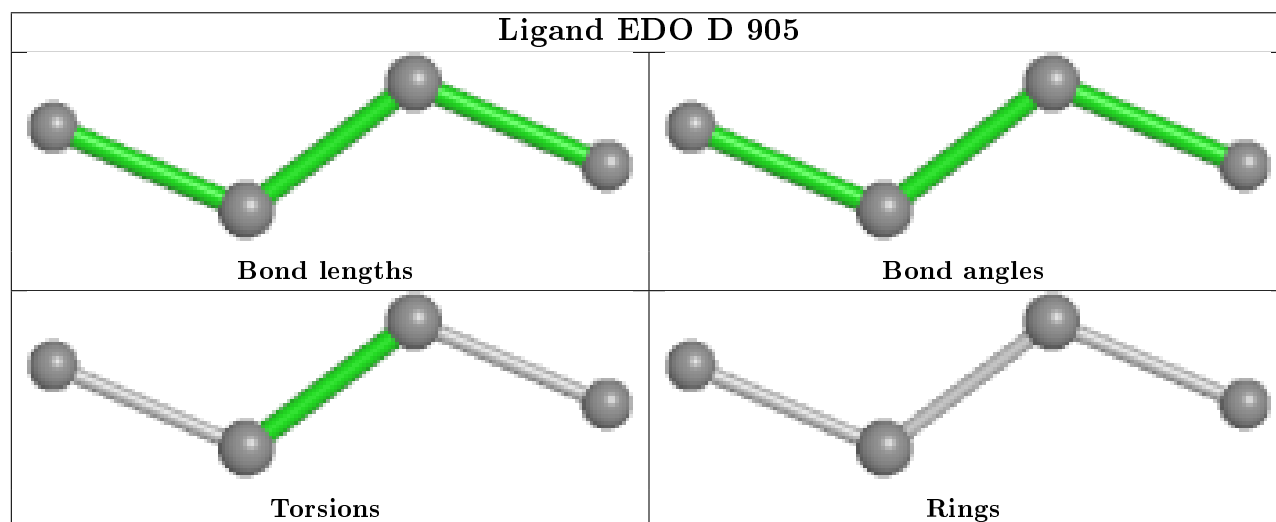
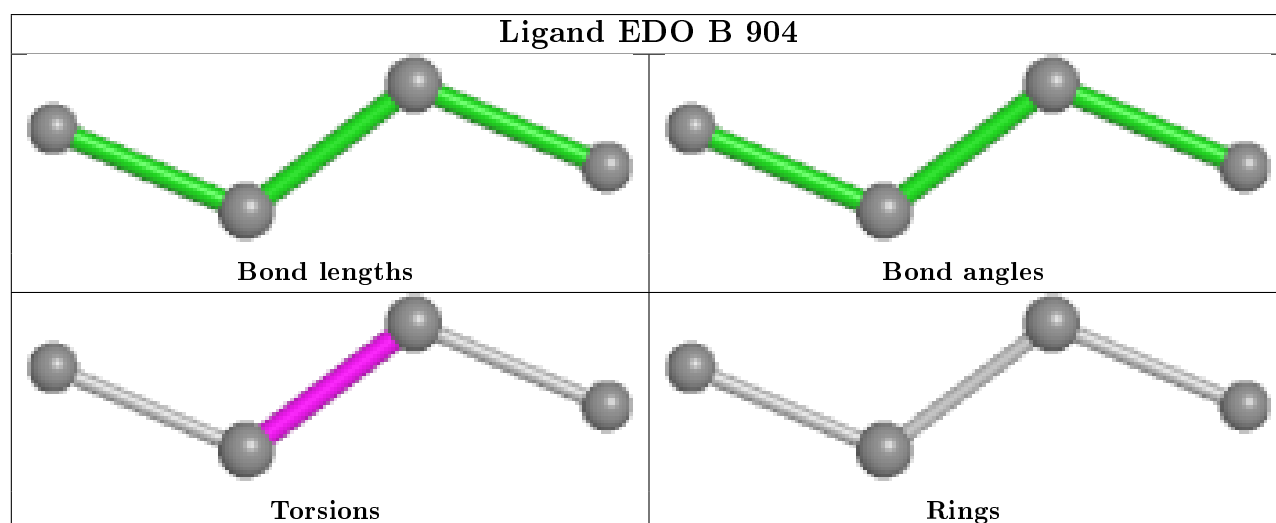
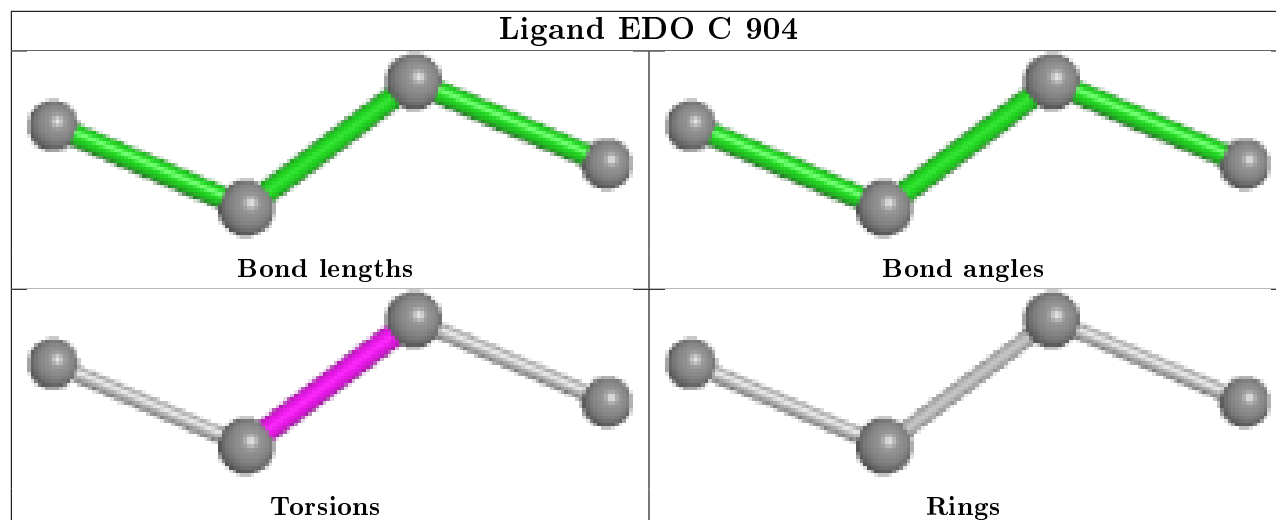
There are no ring outliers.

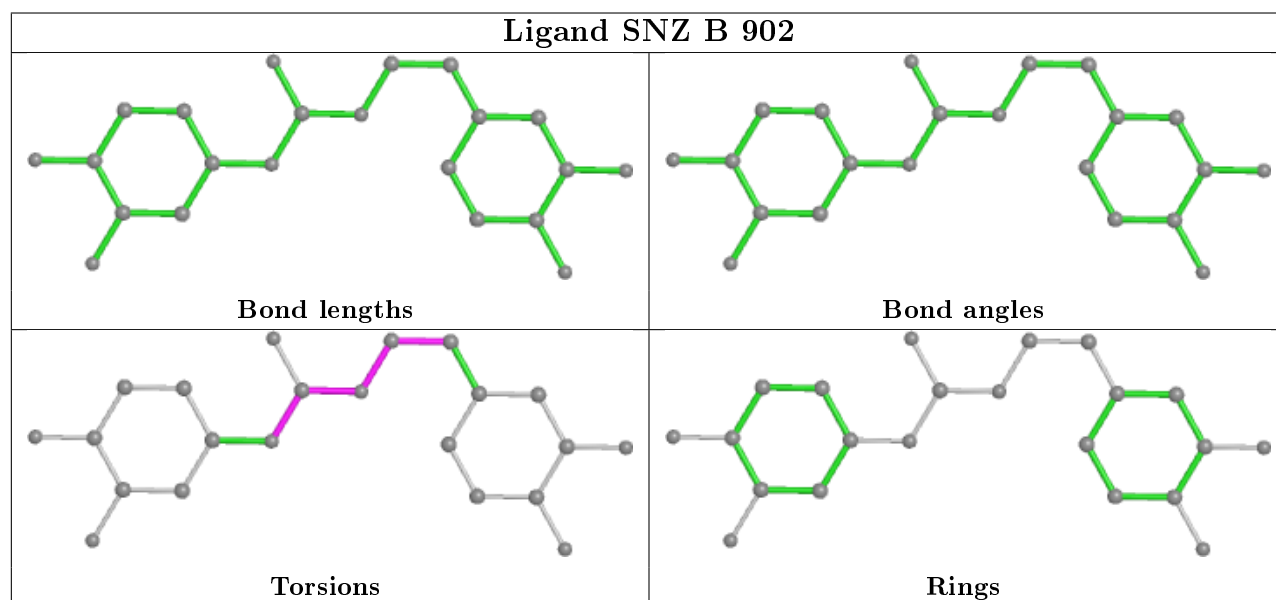
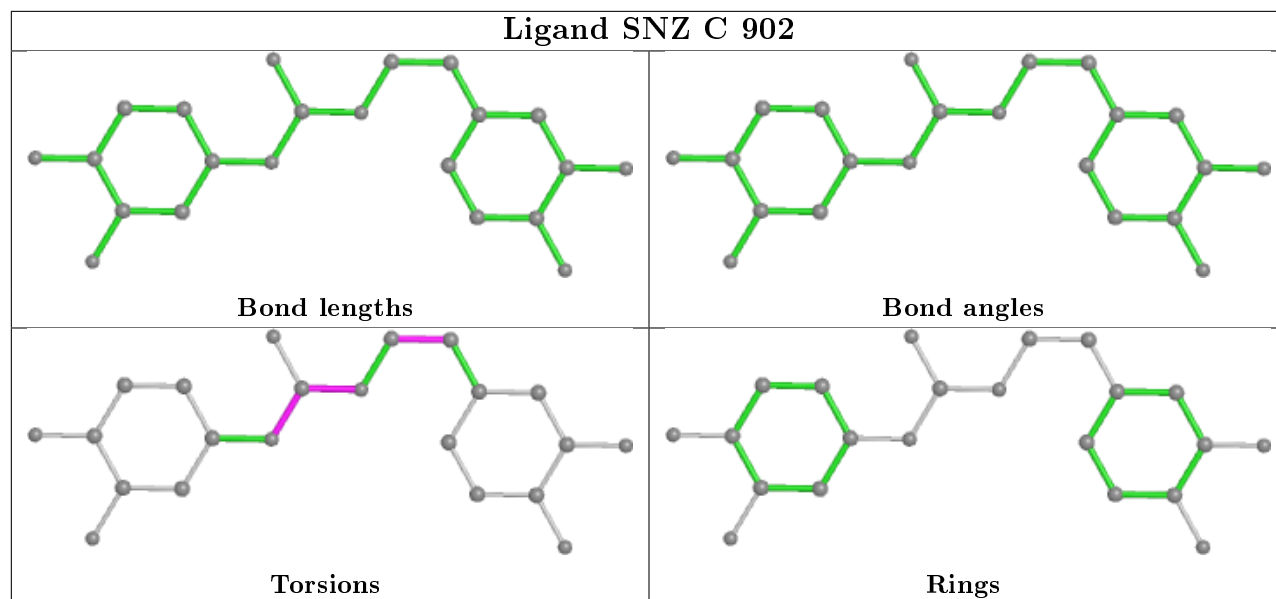
6 monomers are involved in 19 short contacts:

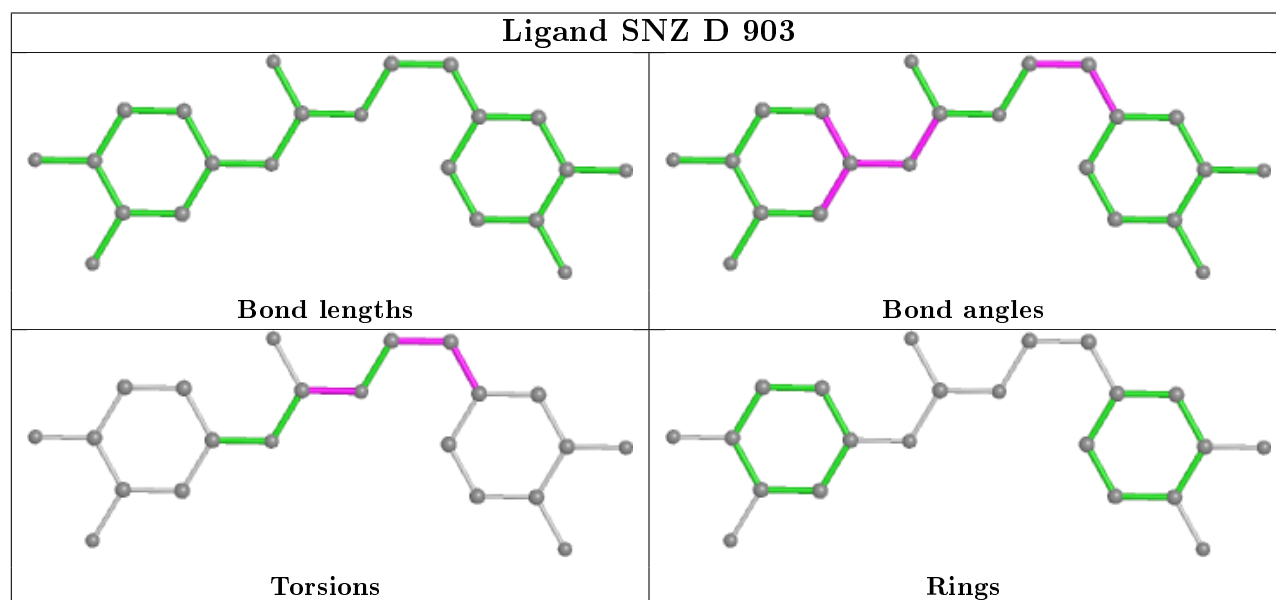
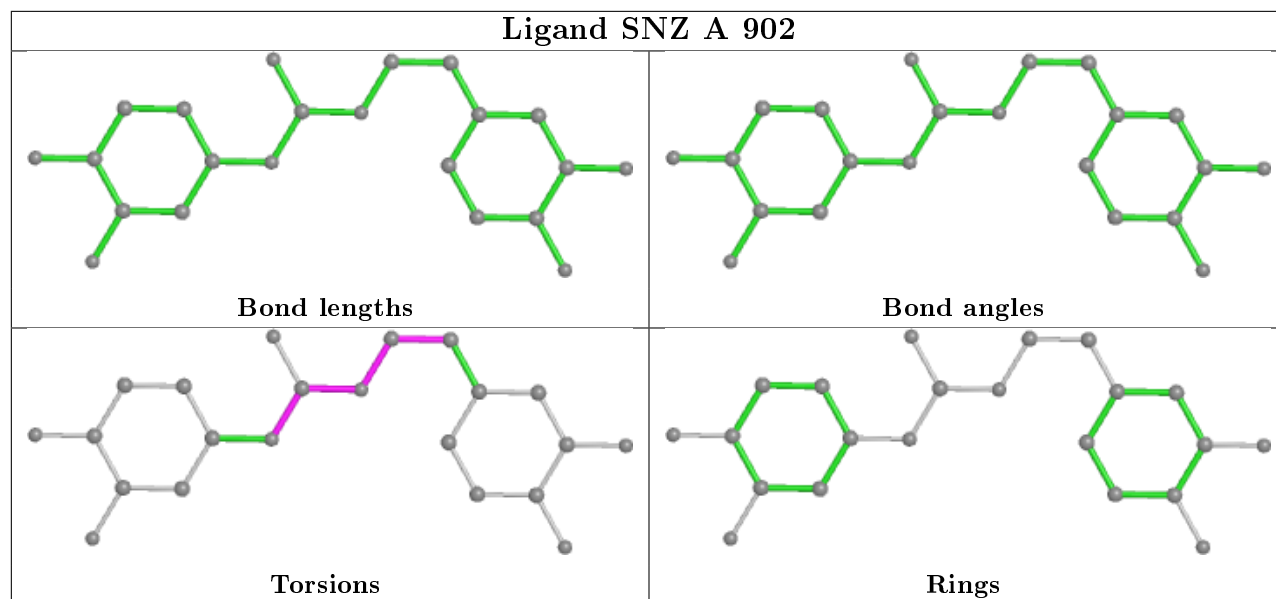
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	D	901	SO4	1	0
6	C	904	EDO	2	0
3	C	902	SNZ	6	0
3	B	902	SNZ	6	0
3	A	902	SNZ	2	0
3	D	903	SNZ	2	0

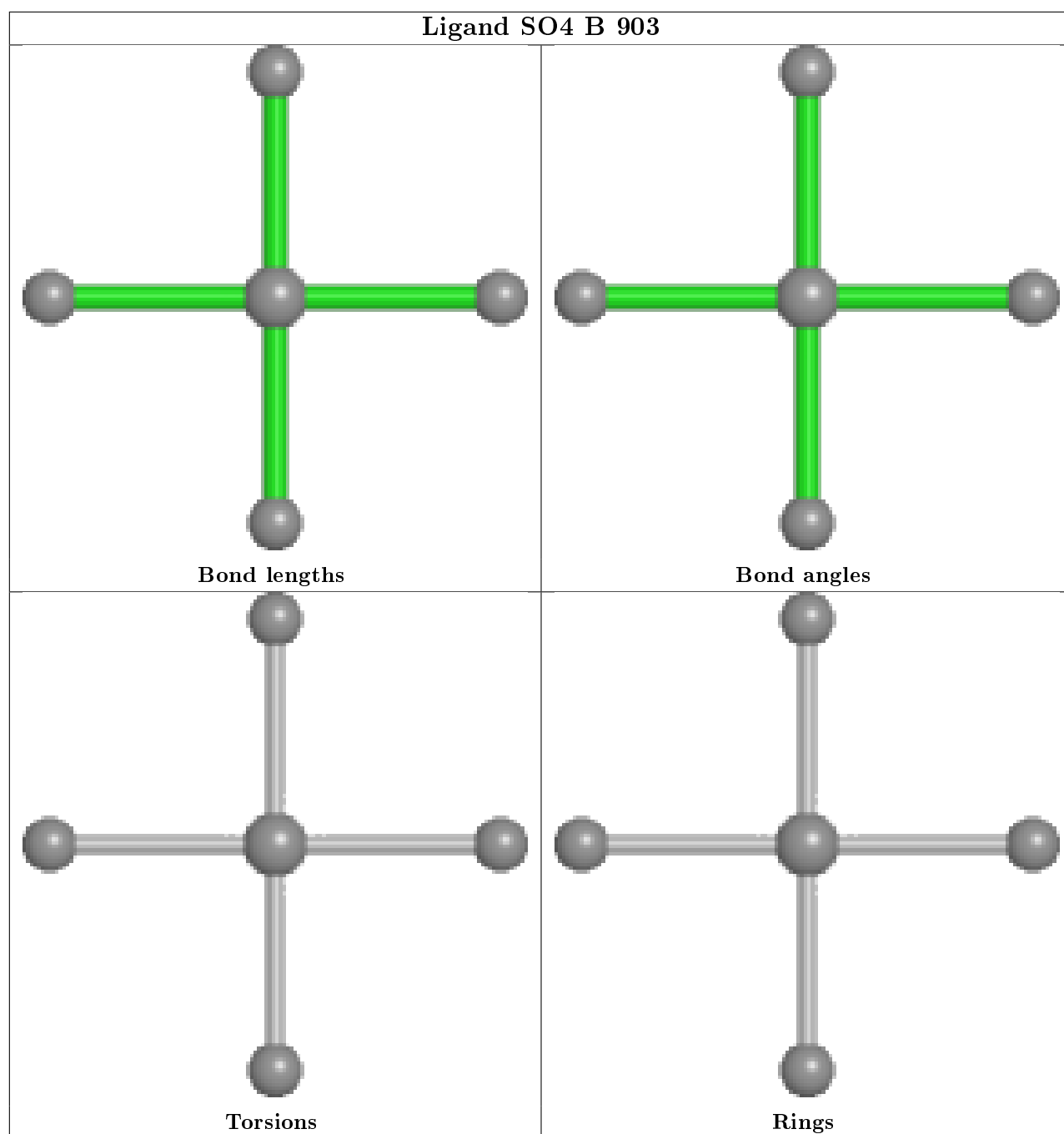
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.











## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data

### 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	392/425 (92%)	-0.24	8 (2%) 65 67	11, 21, 52, 90	0
1	B	394/425 (92%)	-0.11	10 (2%) 57 59	12, 27, 51, 88	0
1	C	399/425 (93%)	-0.21	7 (1%) 68 70	9, 21, 52, 79	0
1	D	388/425 (91%)	-0.16	10 (2%) 56 57	11, 22, 51, 91	0
All	All	1573/1700 (92%)	-0.18	35 (2%) 62 63	9, 23, 52, 91	0

All (35) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	451	ALA	6.8
1	D	460	LEU	5.7
1	D	844	ALA	5.1
1	D	842	PRO	4.7
1	C	564	SER	4.1
1	B	450	GLU	4.0
1	D	461	ALA	4.0
1	B	458	ALA	3.9
1	D	459	ASP	3.9
1	D	843	THR	3.7
1	D	590	ASN	3.2
1	C	450	GLU	3.1
1	B	565	ASN	3.1
1	D	458	ALA	3.1
1	C	448	ASP	3.1
1	B	454	ALA	3.0
1	C	449	TYR	3.0
1	A	843	THR	2.9
1	C	844	ALA	2.8
1	B	456	TYR	2.7
1	A	453	LEU	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	457	GLN	2.7
1	A	751	LYS	2.6
1	B	453	LEU	2.6
1	D	465	LYS	2.6
1	D	827	ASN	2.5
1	A	842	PRO	2.5
1	A	461	ALA	2.4
1	B	464	GLN	2.3
1	A	752	ASN	2.2
1	C	763	ASP	2.2
1	B	843	THR	2.1
1	C	447	ALA	2.1
1	A	452	LYS	2.0
1	B	460	LEU	2.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
3	SNZ	B	902	22/22	0.75	0.42	40,54,68,74	9
3	SNZ	C	902	22/22	0.80	0.38	35,43,73,84	9
3	SNZ	A	902	22/22	0.80	0.47	33,51,58,60	0
3	SNZ	D	903	22/22	0.80	0.38	28,53,67,72	9
2	CA	A	901	1/1	0.84	0.10	58,58,58,58	0
6	EDO	C	904	4/4	0.88	0.23	27,28,33,35	0
2	CA	D	902	1/1	0.89	0.12	41,41,41,41	0
2	CA	C	901	1/1	0.89	0.12	46,46,46,46	0
4	NA	D	904	1/1	0.90	0.24	33,33,33,33	0

*Continued on next page...*

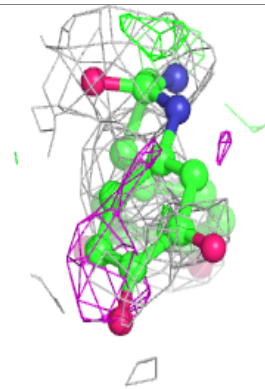
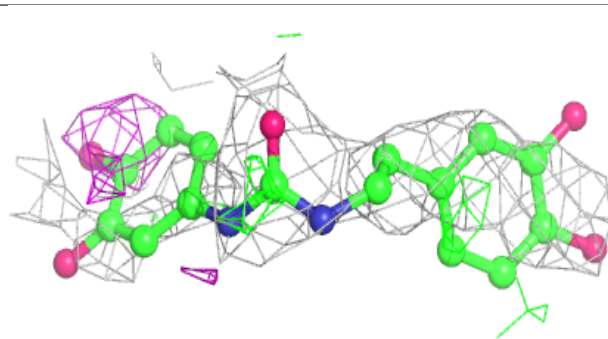
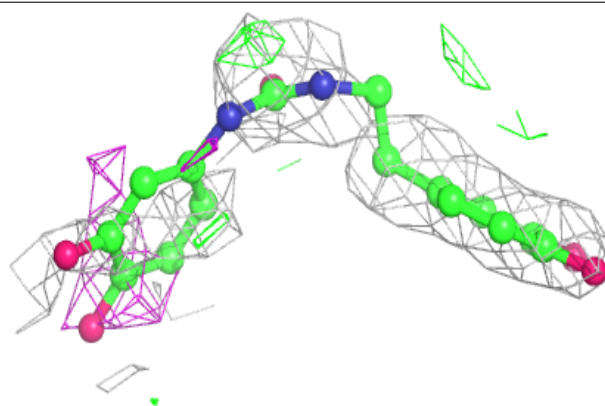
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	CA	B	901	1/1	0.91	0.10	52,52,52,52	0
6	EDO	B	904	4/4	0.92	0.18	23,32,46,48	0
4	NA	C	903	1/1	0.92	0.22	34,34,34,34	0
5	SO4	B	903	5/5	0.92	0.26	46,52,63,81	0
5	SO4	D	901	5/5	0.92	0.17	31,44,51,70	0
4	NA	A	903	1/1	0.95	0.19	27,27,27,27	0
6	EDO	D	905	4/4	0.96	0.17	24,31,31,33	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.

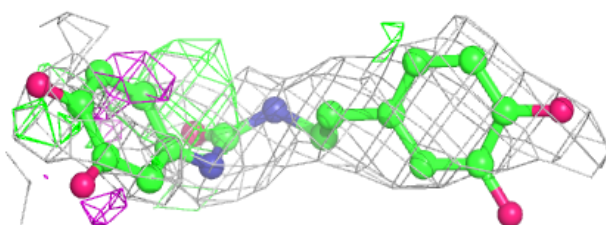
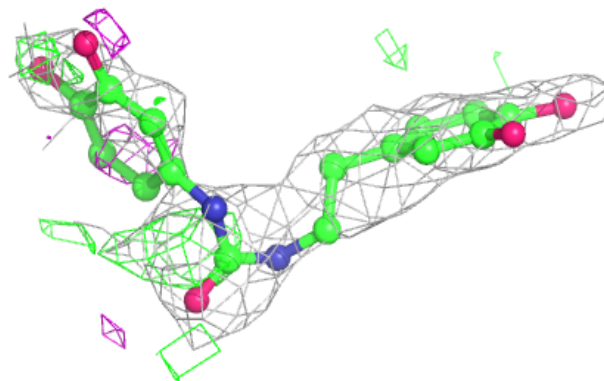
**Electron density around SNZ B 902:**

2mF<sub>o</sub>-DF<sub>c</sub> (at 0.7 rmsd) in gray  
mF<sub>o</sub>-DF<sub>c</sub> (at 3 rmsd) in purple (negative)  
and green (positive)

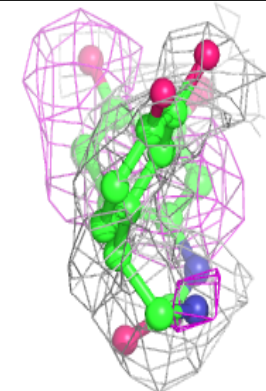
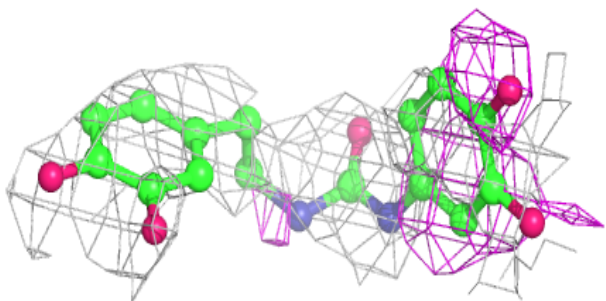
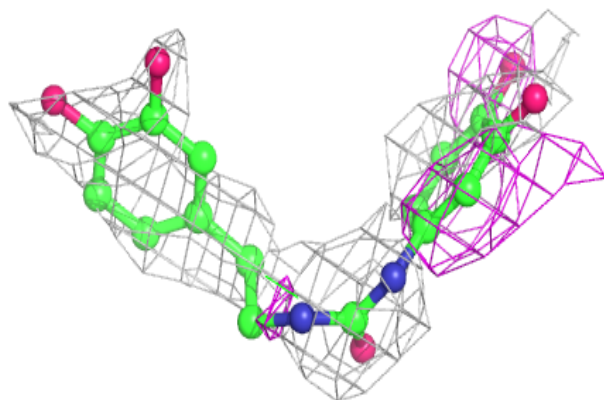


**Electron density around SNZ C 902:**

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

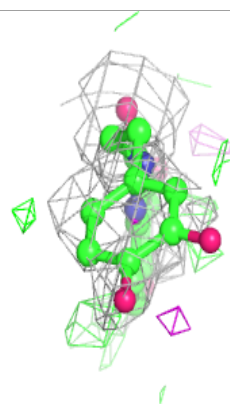
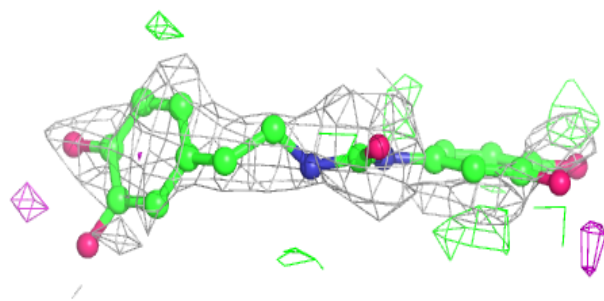
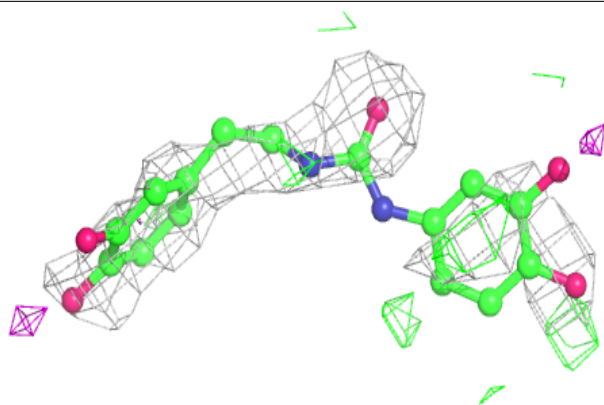
**Electron density around SNZ A 902:**

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



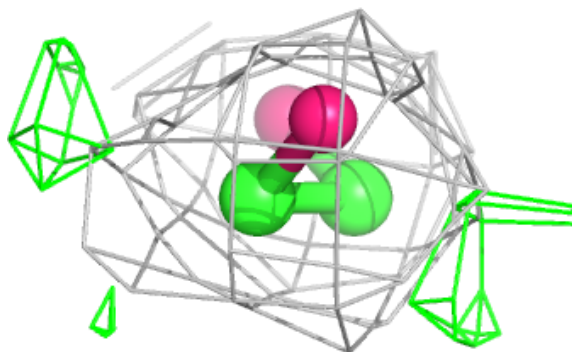
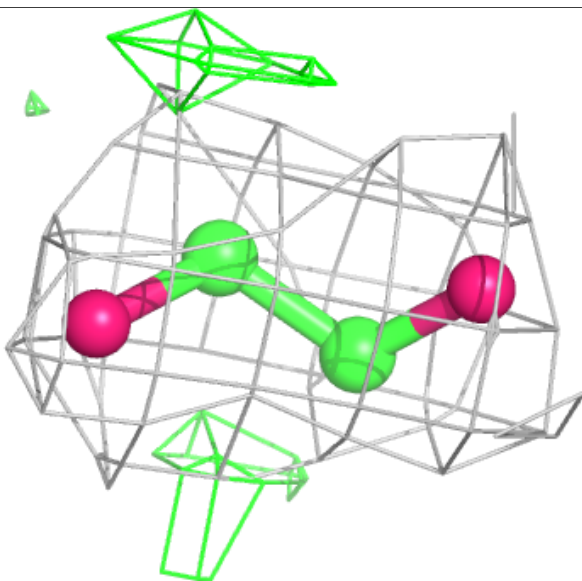
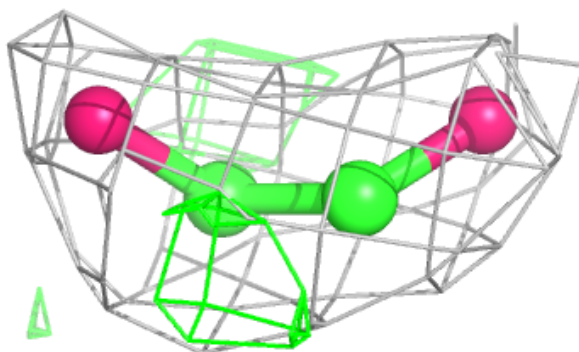
**Electron density around SNZ D 903:**

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



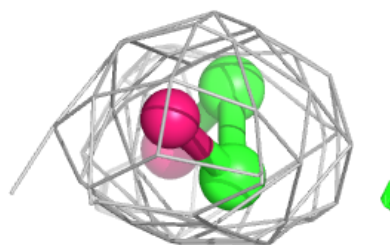
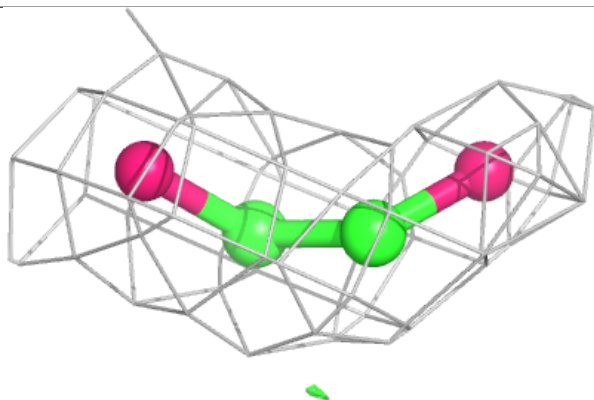
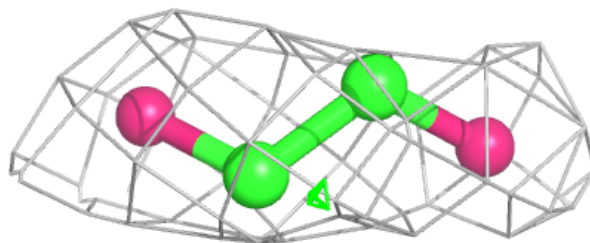
**Electron density around EDO C 904:**

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



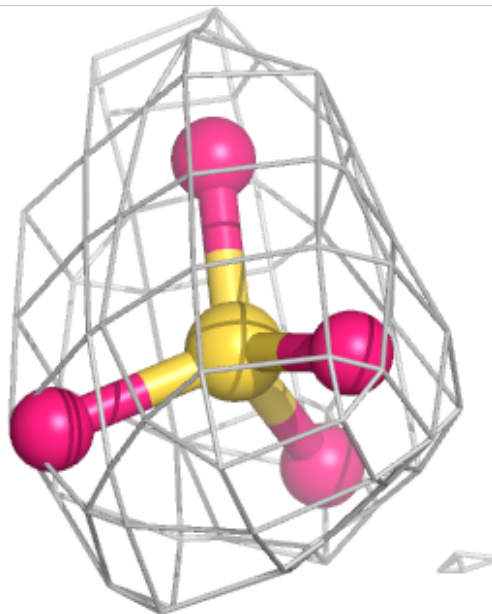
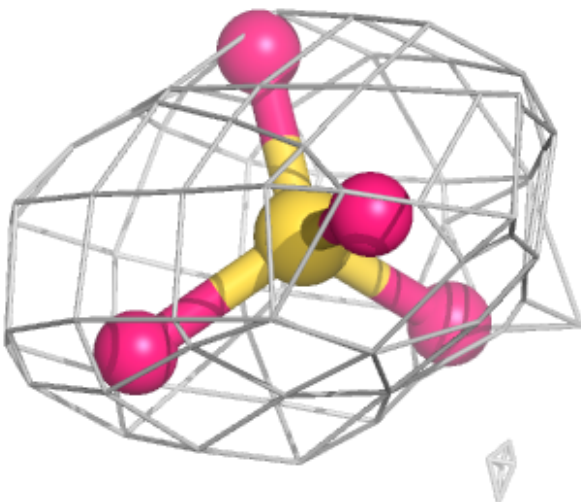
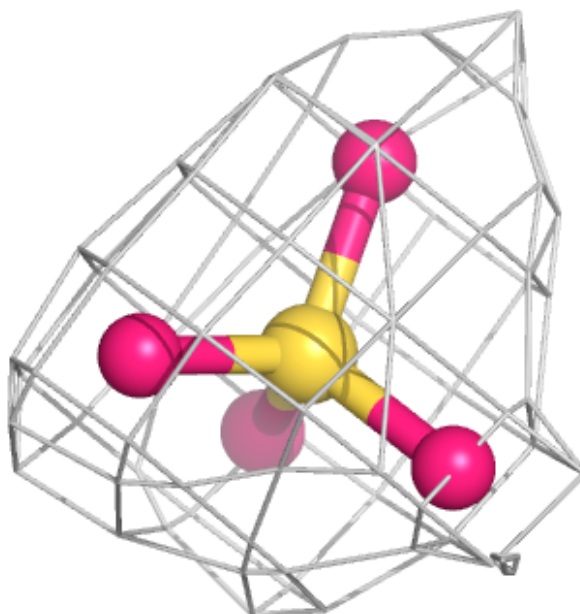
**Electron density around EDO B 904:**

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



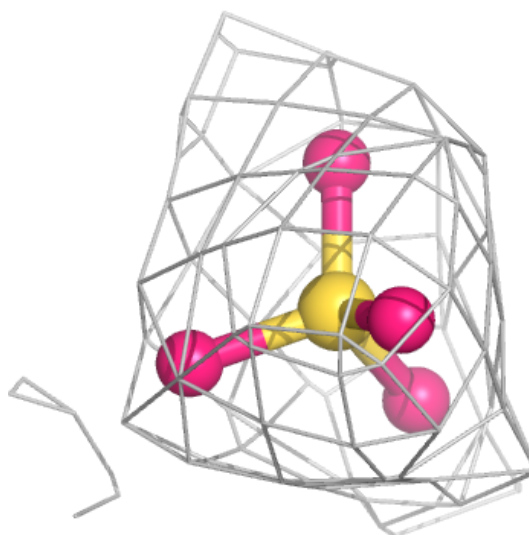
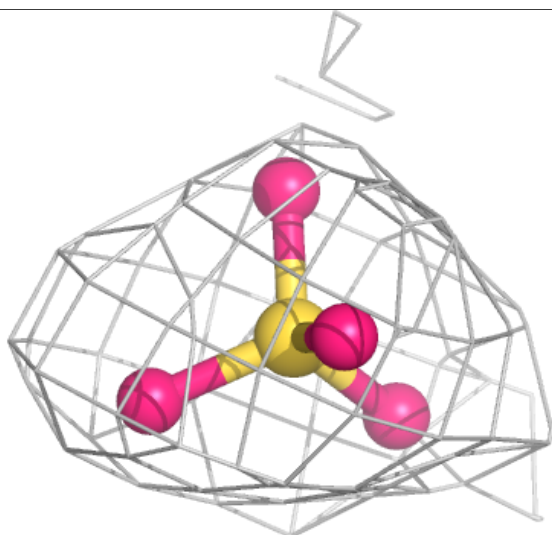
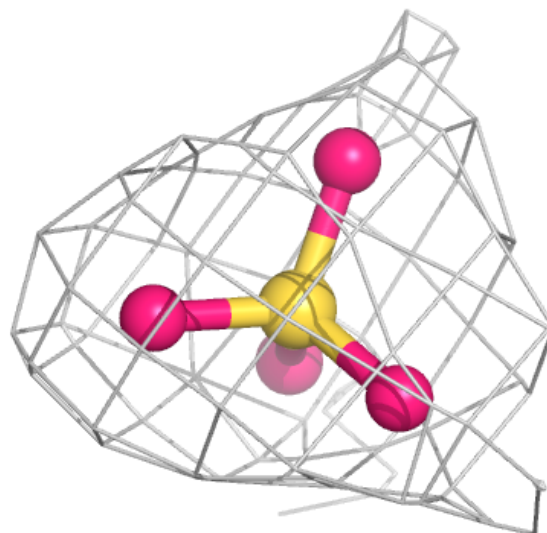
**Electron density around SO4 B 903:**

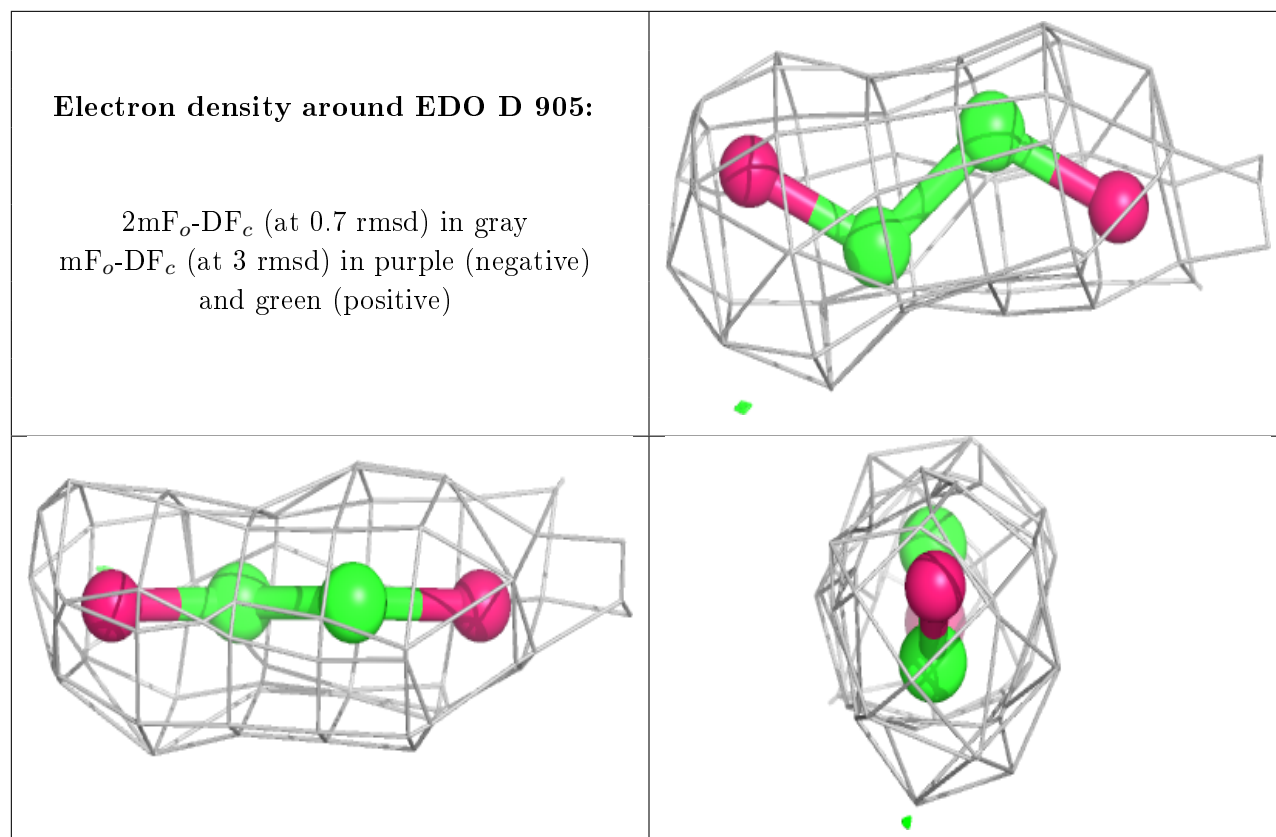
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around SO4 D 901:**

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





## 6.5 Other polymers [i](#)

There are no such residues in this entry.