



# Full wwPDB X-ray Structure Validation Report ⓘ

May 22, 2020 – 12:49 pm BST

PDB ID : 6SAN  
Title : SALSA / DMBT1 / GP340 SRCR domain 8 soaked in calcium and magnesium  
Authors : Reichhardt, M.P.; Johnson, S.; Loimaranta, V.; Lea, S.M.  
Deposited on : 2019-07-17  
Resolution : 1.36 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 : 2.11  
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) : 2.11

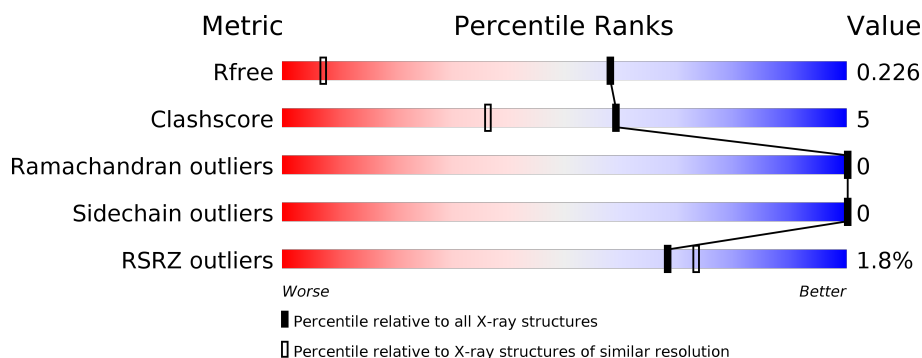
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.36 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	1509 (1.38-1.34)
Clashscore	141614	1551 (1.38-1.34)
Ramachandran outliers	138981	1530 (1.38-1.34)
Sidechain outliers	138945	1530 (1.38-1.34)
RSRZ outliers	127900	1487 (1.38-1.34)

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	137	<div> <div>%</div> <div> </div> <div>73% 9% 18%</div> </div>
1	B	137	<div> <div>%</div> <div> </div> <div>80% 18%</div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 3584 atoms, of which 1531 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 Deleted in malignant brain tumors 1 protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	112	Total	C	H	N	O	S	0	1	0
			1587	505	754	154	165	9			
1	B	112	Total	C	H	N	O	S	0	3	0
			1605	510	763	156	167	9			

There are 56 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	968	MET	-	initiating methionine	UNP Q9UGM3
A	969	LYS	-	expression tag	UNP Q9UGM3
A	970	LEU	-	expression tag	UNP Q9UGM3
A	971	CYS	-	expression tag	UNP Q9UGM3
A	972	ILE	-	expression tag	UNP Q9UGM3
A	973	LEU	-	expression tag	UNP Q9UGM3
A	974	LEU	-	expression tag	UNP Q9UGM3
A	975	ALA	-	expression tag	UNP Q9UGM3
A	976	VAL	-	expression tag	UNP Q9UGM3
A	977	VAL	-	expression tag	UNP Q9UGM3
A	978	ALA	-	expression tag	UNP Q9UGM3
A	979	PHE	-	expression tag	UNP Q9UGM3
A	980	VAL	-	expression tag	UNP Q9UGM3
A	981	GLY	-	expression tag	UNP Q9UGM3
A	982	LEU	-	expression tag	UNP Q9UGM3
A	983	SER	-	expression tag	UNP Q9UGM3
A	984	LEU	-	expression tag	UNP Q9UGM3
A	985	GLY	-	expression tag	UNP Q9UGM3
A	1095	VAL	-	expression tag	UNP Q9UGM3
A	1096	ASN	-	expression tag	UNP Q9UGM3
A	1097	ILE	-	expression tag	UNP Q9UGM3
A	1098	ASP	-	expression tag	UNP Q9UGM3
A	1099	HIS	-	expression tag	UNP Q9UGM3
A	1100	HIS	-	expression tag	UNP Q9UGM3
A	1101	HIS	-	expression tag	UNP Q9UGM3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	1102	HIS	-	expression tag	UNP Q9UGM3
A	1103	HIS	-	expression tag	UNP Q9UGM3
A	1104	HIS	-	expression tag	UNP Q9UGM3
B	968	MET	-	initiating methionine	UNP Q9UGM3
B	969	LYS	-	expression tag	UNP Q9UGM3
B	970	LEU	-	expression tag	UNP Q9UGM3
B	971	CYS	-	expression tag	UNP Q9UGM3
B	972	ILE	-	expression tag	UNP Q9UGM3
B	973	LEU	-	expression tag	UNP Q9UGM3
B	974	LEU	-	expression tag	UNP Q9UGM3
B	975	ALA	-	expression tag	UNP Q9UGM3
B	976	VAL	-	expression tag	UNP Q9UGM3
B	977	VAL	-	expression tag	UNP Q9UGM3
B	978	ALA	-	expression tag	UNP Q9UGM3
B	979	PHE	-	expression tag	UNP Q9UGM3
B	980	VAL	-	expression tag	UNP Q9UGM3
B	981	GLY	-	expression tag	UNP Q9UGM3
B	982	LEU	-	expression tag	UNP Q9UGM3
B	983	SER	-	expression tag	UNP Q9UGM3
B	984	LEU	-	expression tag	UNP Q9UGM3
B	985	GLY	-	expression tag	UNP Q9UGM3
B	1095	VAL	-	expression tag	UNP Q9UGM3
B	1096	ASN	-	expression tag	UNP Q9UGM3
B	1097	ILE	-	expression tag	UNP Q9UGM3
B	1098	ASP	-	expression tag	UNP Q9UGM3
B	1099	HIS	-	expression tag	UNP Q9UGM3
B	1100	HIS	-	expression tag	UNP Q9UGM3
B	1101	HIS	-	expression tag	UNP Q9UGM3
B	1102	HIS	-	expression tag	UNP Q9UGM3
B	1103	HIS	-	expression tag	UNP Q9UGM3
B	1104	HIS	-	expression tag	UNP Q9UGM3

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by author).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	3	Total Mg 3 3	0	0
2	A	3	Total Mg 3 3	0	0

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Cl 1 1	0	0
3	A	1	Total Cl 1 1	0	0

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C H O 13 3 7 3	0	0
4	B	1	Total C H O 13 3 7 3	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	194	Total O 194 194	0	0
5	B	164	Total O 164 164	0	0

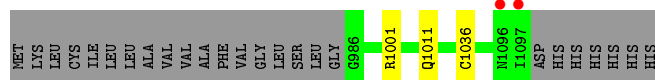
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: Deleted in malignant brain tumors 1 protein



- Molecule 1: Deleted in malignant brain tumors 1 protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	27.23Å 46.64Å 93.63Å 90.00° 97.37° 90.00°	Depositor
Resolution (Å)	30.95 – 1.36 30.95 – 1.36	Depositor EDS
% Data completeness (in resolution range)	98.3 (30.95-1.36) 93.6 (30.95-1.36)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.20 (at 1.36Å)	Xtriage
Refinement program	PHENIX dev_3126	Depositor
R, $R_{free}$	0.189 , 0.226 0.189 , 0.226	Depositor DCC
$R_{free}$ test set	2366 reflections (4.81%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	14.1	Xtriage
Anisotropy	0.281	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 42.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.41$ , $\langle L^2 \rangle = 0.23$	Xtriage
Estimated twinning fraction	0.199 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	3584	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.20% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, MG, CL

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	0.39	0/855	0.63	0/1163
1	B	0.40	0/868	0.60	0/1180
All	All	0.39	0/1723	0.62	0/2343

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

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	833	754	754	11	0
1	B	842	763	759	4	0
2	A	3	0	0	0	0
2	B	3	0	0	0	0
3	A	1	0	0	1	0
3	B	1	0	0	0	0
4	A	6	7	8	1	0
4	B	6	7	8	0	0
5	A	194	0	0	8	7
5	B	164	0	0	4	6
All	All	2053	1531	1529	15	13



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 (15) 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:1036:CYS:SG	5:B:2188:HOH:O	2.36	0.83
1:A:1095:VAL:O	5:A:2101:HOH:O	2.06	0.73
1:A:1000:ASP:OD1	5:A:2102:HOH:O	2.08	0.70
1:B:1001:ARG:NH2	5:B:2106:HOH:O	2.30	0.64
1:A:1003:GLN:O	5:A:2103:HOH:O	2.16	0.60
1:B:1011[B]:GLN:OE1	5:B:2101:HOH:O	2.17	0.57
1:A:1028:ASN:ND2	5:A:2108:HOH:O	2.34	0.57
1:A:1047:ARG:NH1	5:A:2110:HOH:O	2.41	0.54
1:A:1006:VAL:O	3:A:2004:CL:CL	2.66	0.51
1:A:1055:ILE:HD12	4:A:2005:GOL:H2	1.92	0.50
1:A:995:LEU:O	5:A:2104:HOH:O	2.20	0.49
1:A:1004:GLY:HA3	5:A:2103:HOH:O	2.14	0.47
1:A:1096:ASN:OD1	1:A:1097:ILE:N	2.48	0.47
1:B:1011[B]:GLN:NE2	5:B:2110:HOH:O	2.50	0.45
1:A:986:GLY:HA3	5:A:2135:HOH:O	2.22	0.40

All (13) 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)
5:A:2240:HOH:O	5:A:2268:HOH:O[1_455]	1.94	0.26
5:B:2191:HOH:O	5:B:2218:HOH:O[2_656]	1.95	0.25
5:A:2122:HOH:O	5:A:2225:HOH:O[1_655]	1.96	0.24
5:A:2255:HOH:O	5:A:2281:HOH:O[1_455]	2.00	0.20
5:B:2216:HOH:O	5:B:2232:HOH:O[2_556]	2.02	0.18
5:B:2103:HOH:O	5:B:2221:HOH:O[1_455]	2.02	0.18
5:B:2201:HOH:O	5:B:2215:HOH:O[1_455]	2.06	0.14
5:A:2122:HOH:O	5:A:2234:HOH:O[1_655]	2.07	0.13
5:A:2101:HOH:O	5:A:2216:HOH:O[1_455]	2.09	0.11
5:A:2245:HOH:O	5:A:2283:HOH:O[2_645]	2.11	0.09
5:B:2186:HOH:O	5:B:2224:HOH:O[1_455]	2.16	0.04
5:A:2259:HOH:O	5:A:2267:HOH:O[1_655]	2.17	0.03
5:B:2138:HOH:O	5:B:2207:HOH:O[1_655]	2.19	0.01

## 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	111/137 (81%)	109 (98%)	2 (2%)	0	100	100
1	B	113/137 (82%)	110 (97%)	3 (3%)	0	100	100
All	All	224/274 (82%)	219 (98%)	5 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	90/110 (82%)	90 (100%)	0	100	100
1	B	91/110 (83%)	91 (100%)	0	100	100
All	All	181/220 (82%)	181 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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

### 5.3.3 RNA

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 10 ligands modelled in this entry, 8 are monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
4	GOL	A	2005	-	5,5,5	0.80	0	5,5,5	1.02	0
4	GOL	B	2005	-	5,5,5	0.87	0	5,5,5	0.86	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	A	2005	-	-	0/4/4/4	-
4	GOL	B	2005	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	2005	GOL	C1-C2-C3-O3
4	B	2005	GOL	O2-C2-C3-O3

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	2005	GOL	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	112/137 (81%)	0.08	2 (1%) 68 73	9, 15, 23, 53	0
1	B	112/137 (81%)	0.20	2 (1%) 68 73	12, 17, 27, 47	0
All	All	224/274 (81%)	0.14	4 (1%) 68 73	9, 16, 27, 53	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1097	ILE	10.4
1	B	1097	ILE	6.2
1	A	1096	ASN	3.4
1	B	1096	ASN	2.1

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

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

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. 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.

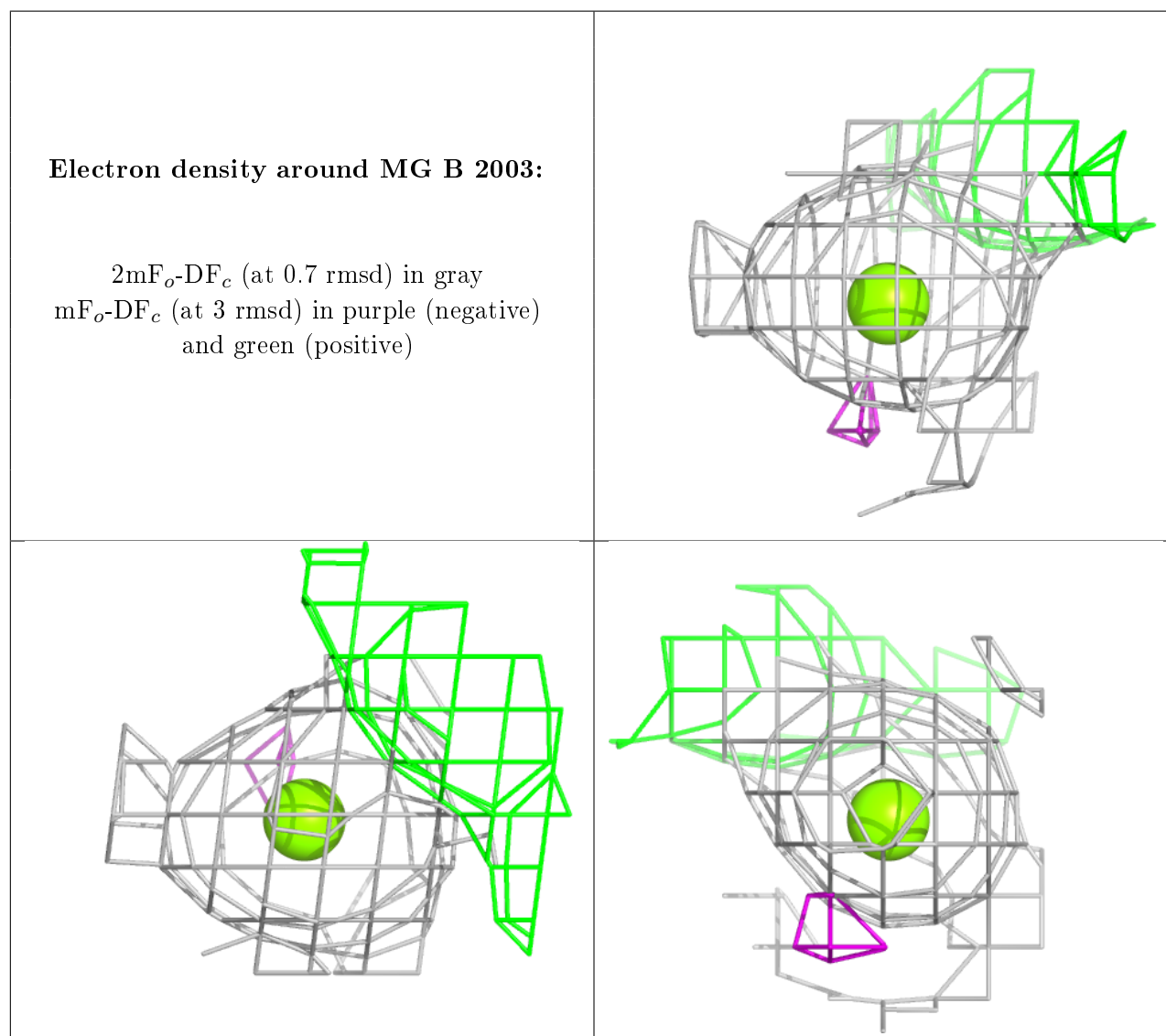
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	GOL	A	2005	6/6	0.84	0.17	31,35,43,43	0

*Continued on next page...*

*Continued from previous page...*

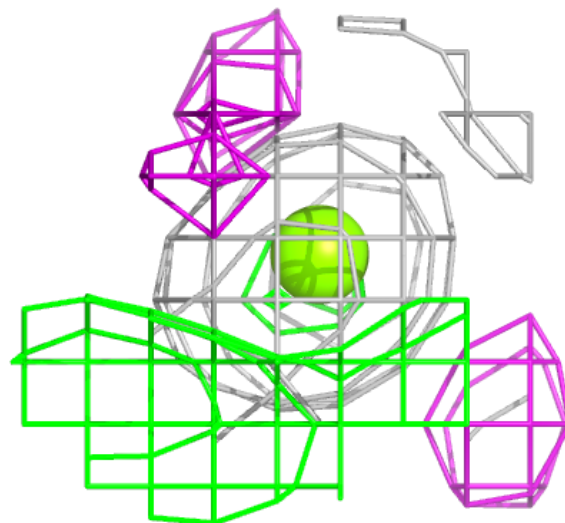
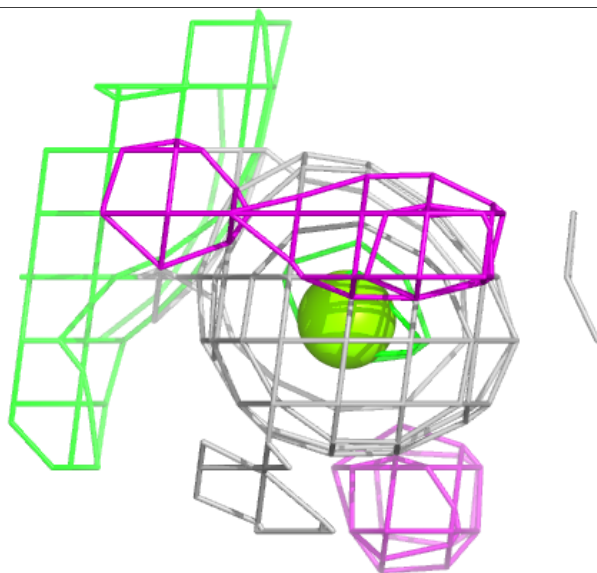
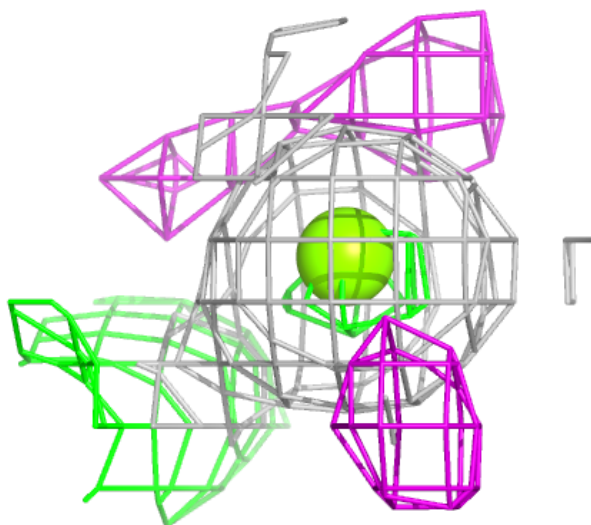
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	GOL	B	2005	6/6	0.89	0.13	21,29,34,36	0
2	MG	B	2003	1/1	0.94	0.11	24,24,24,24	0
2	MG	A	2003	1/1	0.97	0.14	24,24,24,24	0
2	MG	B	2001	1/1	0.99	0.11	15,15,15,15	0
3	CL	B	2004	1/1	0.99	0.03	26,26,26,26	0
2	MG	B	2002	1/1	0.99	0.05	16,16,16,16	0
2	MG	A	2002	1/1	0.99	0.07	14,14,14,14	0
3	CL	A	2004	1/1	0.99	0.07	20,20,20,20	0
2	MG	A	2001	1/1	1.00	0.04	16,16,16,16	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.



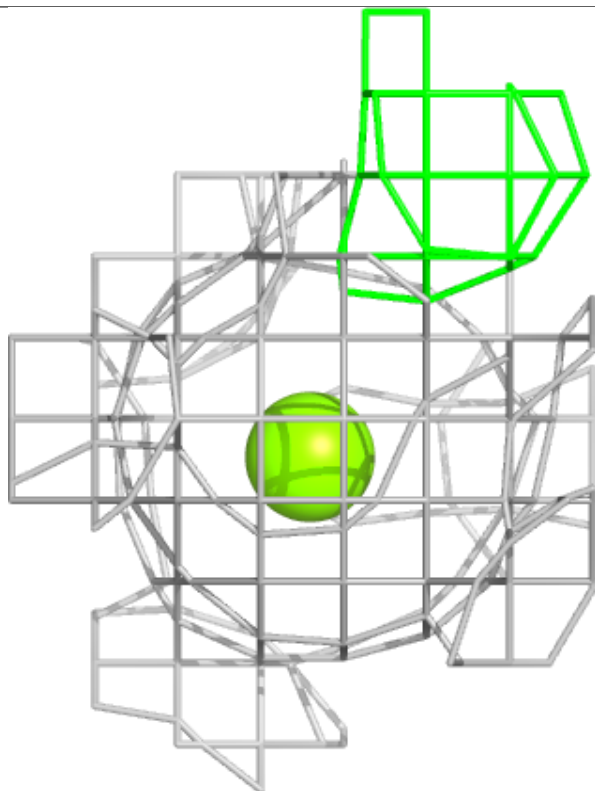
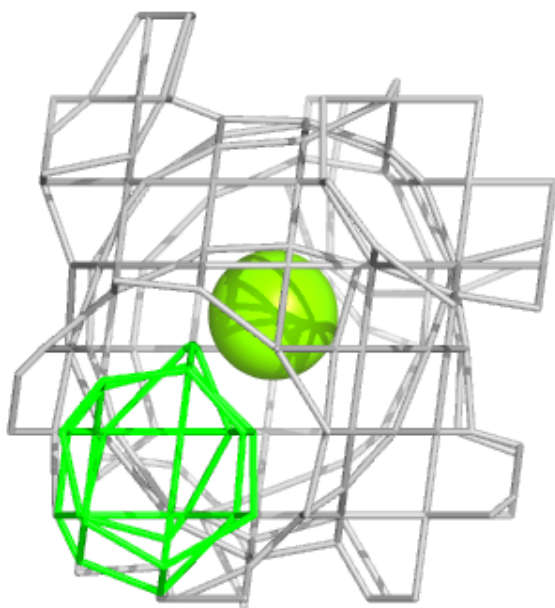
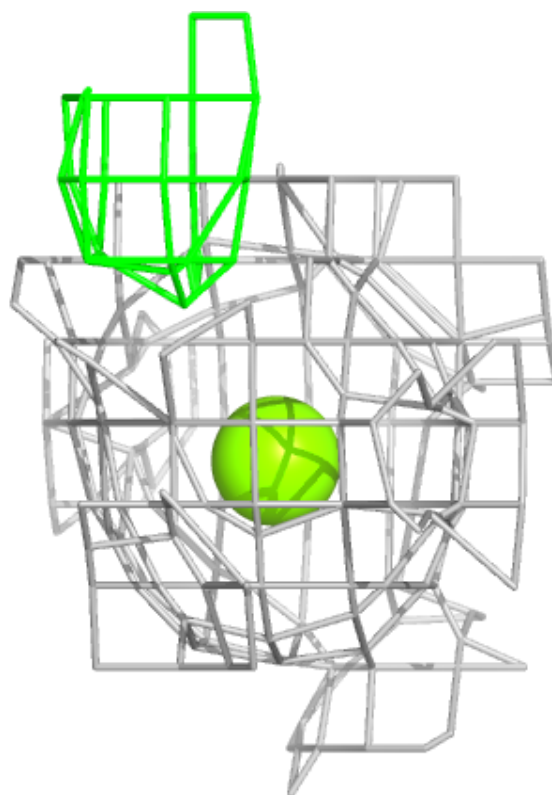
**Electron density around MG A 2003:**

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



**Electron density around MG B 2001:**

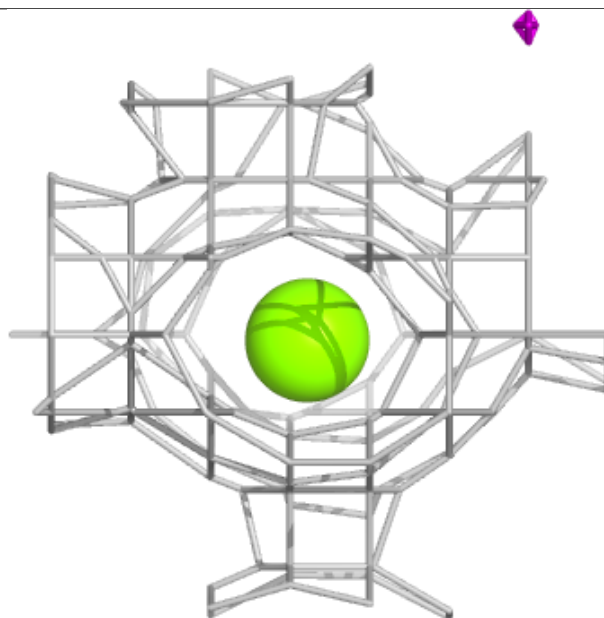
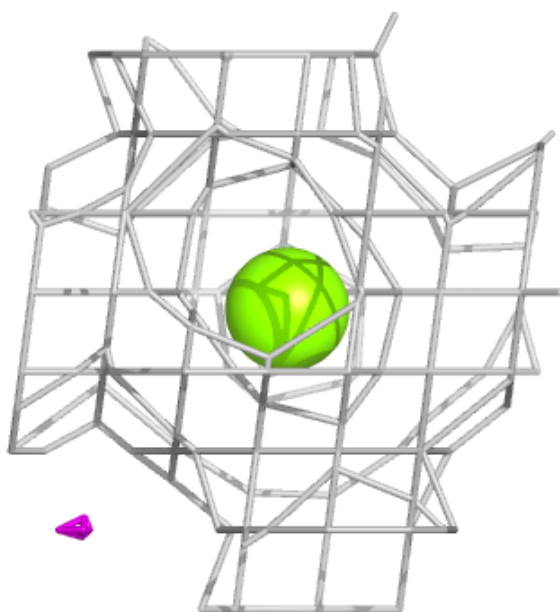
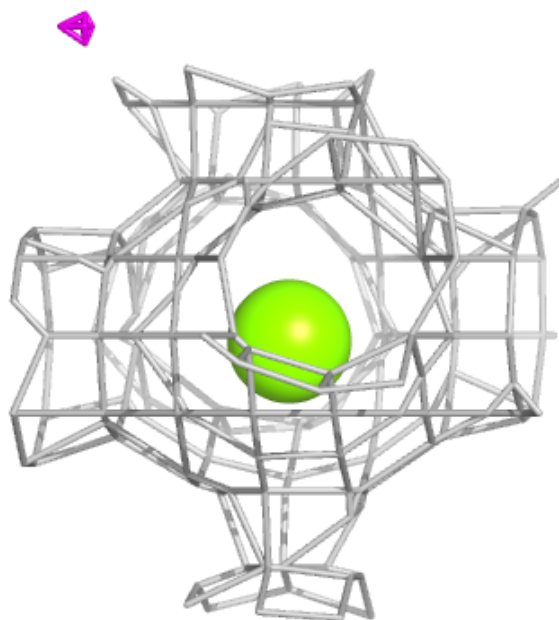
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





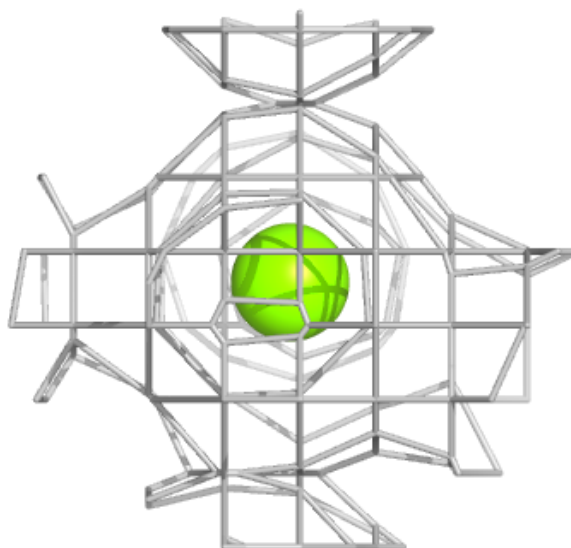
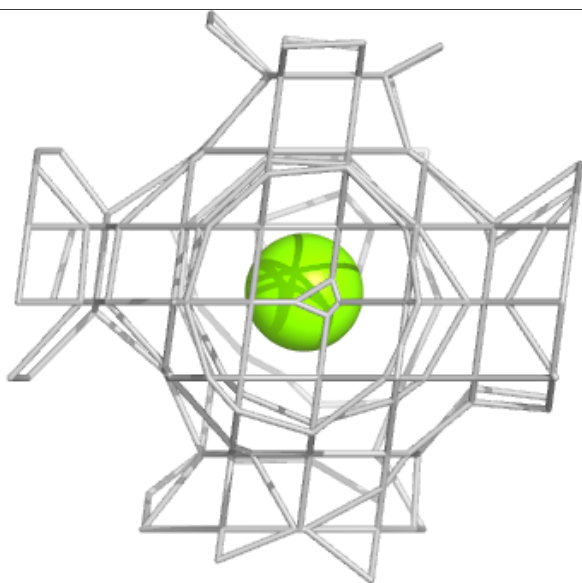
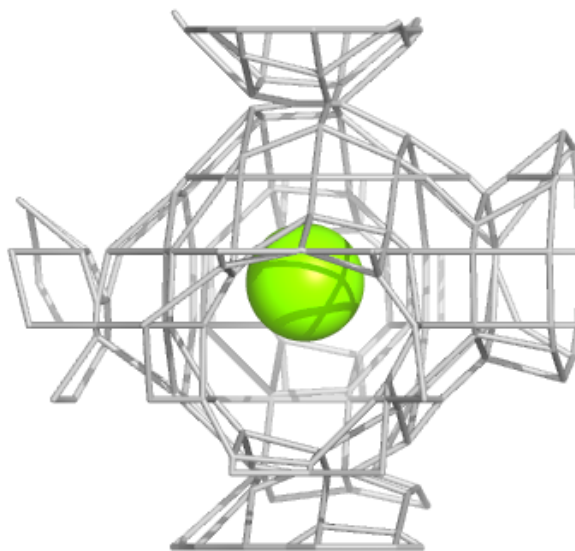
**Electron density around MG B 2002:**

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



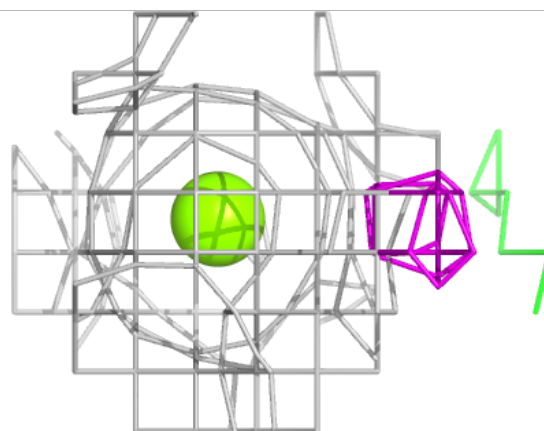
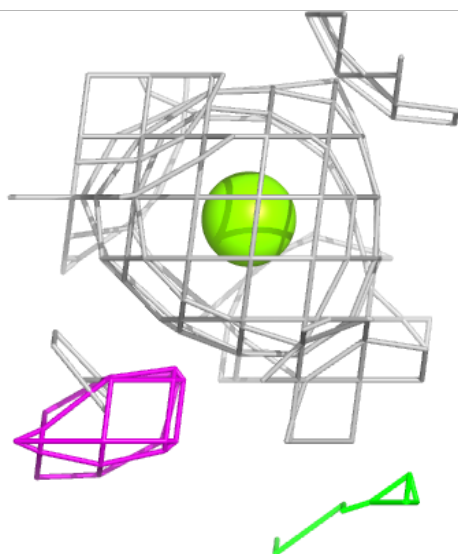
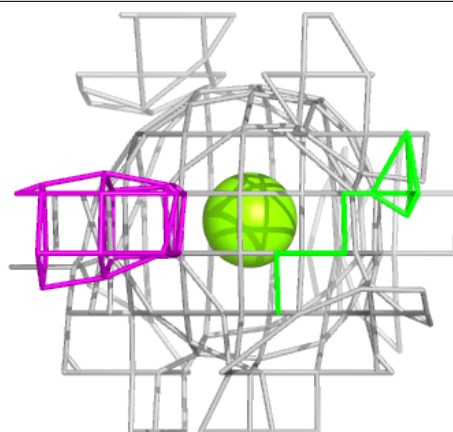
**Electron density around MG A 2002:**

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



**Electron density around MG A 2001:**

$2mF_o-DF_c$  (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.