



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 9, 2020 – 12:08 PM JST

PDB ID : 6LB3  
Title : Crystal structure of PA4674 in complex with its operator DNA (18bp) from *Pseudomonas aeruginosa*  
Authors : Liu, Y.; Zhang, H.; Gao, Z.; Dong, Y.  
Deposited on : 2019-11-13  
Resolution : 2.50 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 : 2.14.6  
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) : 2.14.6

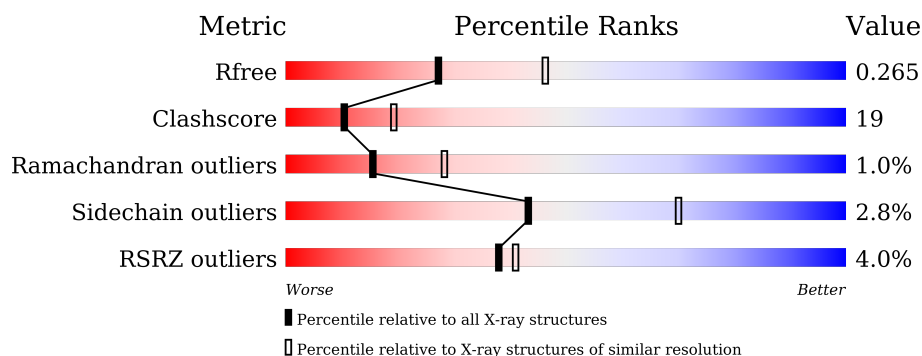
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	109	<div> <div></div> <div>64% 17% 17%</div> </div>
1	B	109	<div> <div>2%</div> <div>65% 19% 15%</div> </div>
1	C	109	<div> <div></div> <div>63% 22% 15%</div> </div>
1	D	109	<div> <div>3%</div> <div>61% 24% 16%</div> </div>
1	E	109	<div> <div>5%</div> <div>57% 20% 6% 17%</div> </div>
1	F	109	<div> <div>3%</div> <div>63% 22% 15%</div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	G	109	
1	H	109	
2	I	18	
2	K	18	
2	M	18	
3	J	18	
3	L	18	
3	N	18	

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 7646 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 HTH cro/C1-type domain-containing protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	91	Total	C	N	O	S	0	0	0
			722	457	126	136	3			
1	B	93	Total	C	N	O	S	0	0	0
			734	464	128	138	4			
1	C	93	Total	C	N	O	S	0	0	0
			734	464	128	138	4			
1	D	92	Total	C	N	O	S	0	0	0
			730	462	127	137	4			
1	E	90	Total	C	N	O	S	0	0	0
			715	451	125	135	4			
1	F	93	Total	C	N	O	S	0	0	0
			738	468	128	138	4			
1	G	88	Total	C	N	O	S	0	0	0
			692	438	120	131	3			
1	H	95	Total	C	N	O	S	0	0	0
			747	473	130	140	4			

There are 64 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	102	GLY	-	expression tag	UNP Q9HVC1
A	103	SER	-	expression tag	UNP Q9HVC1
A	104	HIS	-	expression tag	UNP Q9HVC1
A	105	HIS	-	expression tag	UNP Q9HVC1
A	106	HIS	-	expression tag	UNP Q9HVC1
A	107	HIS	-	expression tag	UNP Q9HVC1
A	108	HIS	-	expression tag	UNP Q9HVC1
A	109	HIS	-	expression tag	UNP Q9HVC1
B	102	GLY	-	expression tag	UNP Q9HVC1
B	103	SER	-	expression tag	UNP Q9HVC1
B	104	HIS	-	expression tag	UNP Q9HVC1
B	105	HIS	-	expression tag	UNP Q9HVC1
B	106	HIS	-	expression tag	UNP Q9HVC1

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	107	HIS	-	expression tag	UNP Q9HVC1
B	108	HIS	-	expression tag	UNP Q9HVC1
B	109	HIS	-	expression tag	UNP Q9HVC1
C	102	GLY	-	expression tag	UNP Q9HVC1
C	103	SER	-	expression tag	UNP Q9HVC1
C	104	HIS	-	expression tag	UNP Q9HVC1
C	105	HIS	-	expression tag	UNP Q9HVC1
C	106	HIS	-	expression tag	UNP Q9HVC1
C	107	HIS	-	expression tag	UNP Q9HVC1
C	108	HIS	-	expression tag	UNP Q9HVC1
C	109	HIS	-	expression tag	UNP Q9HVC1
D	102	GLY	-	expression tag	UNP Q9HVC1
D	103	SER	-	expression tag	UNP Q9HVC1
D	104	HIS	-	expression tag	UNP Q9HVC1
D	105	HIS	-	expression tag	UNP Q9HVC1
D	106	HIS	-	expression tag	UNP Q9HVC1
D	107	HIS	-	expression tag	UNP Q9HVC1
D	108	HIS	-	expression tag	UNP Q9HVC1
D	109	HIS	-	expression tag	UNP Q9HVC1
E	102	GLY	-	expression tag	UNP Q9HVC1
E	103	SER	-	expression tag	UNP Q9HVC1
E	104	HIS	-	expression tag	UNP Q9HVC1
E	105	HIS	-	expression tag	UNP Q9HVC1
E	106	HIS	-	expression tag	UNP Q9HVC1
E	107	HIS	-	expression tag	UNP Q9HVC1
E	108	HIS	-	expression tag	UNP Q9HVC1
E	109	HIS	-	expression tag	UNP Q9HVC1
F	102	GLY	-	expression tag	UNP Q9HVC1
F	103	SER	-	expression tag	UNP Q9HVC1
F	104	HIS	-	expression tag	UNP Q9HVC1
F	105	HIS	-	expression tag	UNP Q9HVC1
F	106	HIS	-	expression tag	UNP Q9HVC1
F	107	HIS	-	expression tag	UNP Q9HVC1
F	108	HIS	-	expression tag	UNP Q9HVC1
F	109	HIS	-	expression tag	UNP Q9HVC1
G	102	GLY	-	expression tag	UNP Q9HVC1
G	103	SER	-	expression tag	UNP Q9HVC1
G	104	HIS	-	expression tag	UNP Q9HVC1
G	105	HIS	-	expression tag	UNP Q9HVC1
G	106	HIS	-	expression tag	UNP Q9HVC1
G	107	HIS	-	expression tag	UNP Q9HVC1
G	108	HIS	-	expression tag	UNP Q9HVC1

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
G	109	HIS	-	expression tag	UNP Q9HVC1
H	102	GLY	-	expression tag	UNP Q9HVC1
H	103	SER	-	expression tag	UNP Q9HVC1
H	104	HIS	-	expression tag	UNP Q9HVC1
H	105	HIS	-	expression tag	UNP Q9HVC1
H	106	HIS	-	expression tag	UNP Q9HVC1
H	107	HIS	-	expression tag	UNP Q9HVC1
H	108	HIS	-	expression tag	UNP Q9HVC1
H	109	HIS	-	expression tag	UNP Q9HVC1

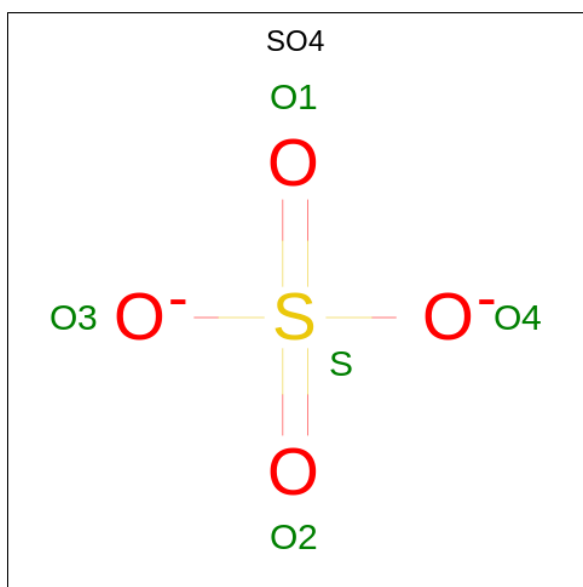
- Molecule 2 is a DNA chain called DNA (5'-D(P\*AP\*CP\*GP\*CP\*TP\*TP\*AP\*AP\*CP\*GP\*TP\*TP\*AP\*AP\*GP\*GP\*GP\*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	I	18	Total	C	N	O	P	0	0	0
			372	177	69	108	18			
2	K	18	Total	C	N	O	P	0	0	0
			372	177	69	108	18			
2	M	7	Total	C	N	O	P	0	0	0
			142	68	25	42	7			

- Molecule 3 is a DNA chain called DNA (5'-D(P\*AP\*CP\*CP\*CP\*TP\*TP\*AP\*AP\*CP\*GP\*TP\*TP\*AP\*AP\*GP\*CP\*GP\*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	J	18	Total	C	N	O	P	0	0	0
			366	175	65	108	18			
3	L	18	Total	C	N	O	P	0	0	0
			366	175	65	108	18			
3	N	7	Total	C	N	O	P	0	0	0
			145	69	27	42	7			

- Molecule 4 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
4	A	1	Total	O	S	0	0
			5	4	1		
4	G	1	Total	O	S	0	0
			5	4	1		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	4	Total	O	0	0
			4	4		
5	B	4	Total	O	0	0
			4	4		
5	C	5	Total	O	0	0
			5	5		
5	D	11	Total	O	0	0
			11	11		
5	E	11	Total	O	0	0
			11	11		
5	F	4	Total	O	0	0
			4	4		
5	G	5	Total	O	0	0
			5	5		
5	H	1	Total	O	0	0
			1	1		
5	I	2	Total	O	0	0
			2	2		
5	J	2	Total	O	0	0
			2	2		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	K	6	Total	O	0	0
			6	6		
5	L	6	Total	O	0	0
			6	6		



### 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.

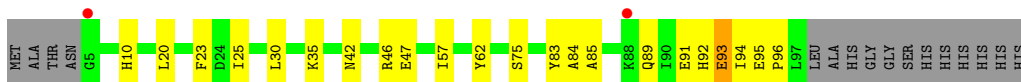
- Molecule 1: HTH cro/C1-type domain-containing protein

Chain A: 



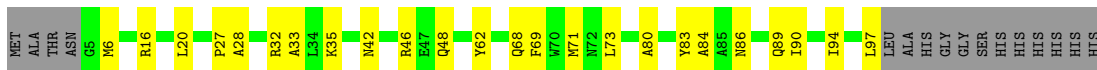
- Molecule 1: HTH cro/C1-type domain-containing protein

Chain B: 



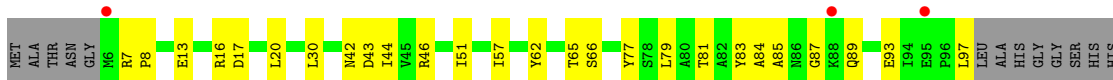
- Molecule 1: HTH cro/C1-type domain-containing protein

Chain C: 



- Molecule 1: HTH cro/C1-type domain-containing protein

Chain D: 



HIS  
HIS  
HIS  
HIS

- Molecule 1: HTH cro/C1-type domain-containing protein

Chain E: 



HIS  
HIS  
HIS

- Molecule 1: HTH cro/C1-type domain-containing protein

Chain F: 

MET ALA THR ASN GLY M6 E13 R16 F19 L20 D24 L30 A31 R32 A33 L34 S37 A38 P39 V45 D54 R58 R61 Y62 F63 L79 Y83 A84 A85 K88 Q89 I90 E93 I94 L98 ALA HIS GLY SER HIS HIS HIS HIS

HIS

- Molecule 1: HTH cro/C1-type domain-containing protein

Chain G: 

MET ALA THR ASN MET ARG P8 I9 H10 L20 M21 R32 E47 I51 R58 Q74 L79 Y83 A84 N86 K87 K88 Q89 I90 E91 H92 E93 I94 E95 PRO LEU ALA HIS GLY GLY HIS HIS HIS HIS

- Molecule 1: HTH cro/C1-type domain-containing protein

Chain H: 

MET ALA THR ASN M6 E13 R16 D17 E18 F19 L20 M21 F23 A28 R32 A33 L34 K35 Y36 S37 A38 P39 T40 Y45 R46 E47 R58 R61 Y62 F63 D64 Y83 A84 A85 N86 I90 E93 I94 L98 A99 HIS GLY SER HIS HIS HIS HIS

HIS  
HIS  
HIS

- Molecule 2: DNA (5'-D(P\*AP\*CP\*GP\*CP\*TP\*TP\*AP\*AP\*CP\*GP\*TP\*TP\*AP\*AP\*GP\*GP\*GP\*T)-3')

Chain I: 

A1 G15 G16 G17 T18

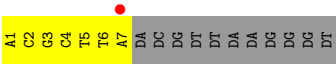
- Molecule 2: DNA (5'-D(P\*AP\*CP\*GP\*CP\*TP\*TP\*AP\*AP\*CP\*GP\*TP\*TP\*AP\*AP\*GP\*GP\*GP\*T)-3')

Chain K: 

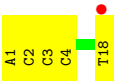
A1 A7 G15 G16 G17 T18

- Molecule 2: DNA (5'-D(P\*AP\*CP\*GP\*CP\*TP\*TP\*AP\*AP\*CP\*GP\*TP\*TP\*AP\*AP\*GP\*GP\*GP\*T)-3')

Chain M: 



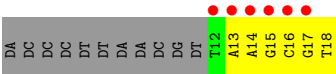
● Molecule 3: DNA (5'-D(P\*AP\*CP\*CP\*CP\*TP\*TP\*AP\*AP\*CP\*GP\*TP\*TP\*AP\*AP\*GP\*C  
P\*GP\*T)-3')



● Molecule 3: DNA (5'-D(P\*AP\*CP\*CP\*CP\*TP\*TP\*AP\*AP\*CP\*GP\*TP\*TP\*AP\*AP\*GP\*C  
P\*GP\*T)-3')



● Molecule 3: DNA (5'-D(P\*AP\*CP\*CP\*CP\*TP\*TP\*AP\*AP\*CP\*GP\*TP\*TP\*AP\*AP\*GP\*C  
P\*GP\*T)-3')



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	57.28Å 95.57Å 128.86Å 90.00° 96.29° 90.00°	Depositor
Resolution (Å)	47.20 – 2.50 47.20 – 2.50	Depositor EDS
% Data completeness (in resolution range)	98.5 (47.20-2.50) 98.5 (47.20-2.50)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.13 (at 2.51Å)	Xtriage
Refinement program	PHENIX 1.10.1_2155	Depositor
R, $R_{free}$	0.218 , 0.262 0.222 , 0.265	Depositor DCC
$R_{free}$ test set	1995 reflections (4.22%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	52.8	Xtriage
Anisotropy	0.291	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 45.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	7646	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.37% 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: 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	0.50	0/737	0.66	0/998
1	B	0.48	0/749	0.66	0/1013
1	C	0.40	0/749	0.59	0/1013
1	D	0.45	0/745	0.66	0/1008
1	E	0.51	0/729	0.74	0/985
1	F	0.48	0/753	0.61	0/1019
1	G	0.51	0/706	0.63	0/954
1	H	0.52	1/762 (0.1%)	0.70	0/1031
2	I	0.95	0/417	1.06	0/642
2	K	1.07	0/417	1.08	0/642
2	M	0.85	0/158	1.17	0/241
3	J	0.96	0/409	1.01	0/628
3	L	0.97	0/409	1.06	0/628
3	N	0.76	0/162	0.98	0/248
All	All	0.64	1/7902 (0.0%)	0.79	0/11050

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	H	19	PHE	C-N	5.61	1.47	1.34

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	722	0	709	27	0
1	B	734	0	721	26	0
1	C	734	0	721	21	0
1	D	730	0	718	24	0
1	E	715	0	700	34	0
1	F	738	0	729	20	0
1	G	692	0	676	14	0
1	H	747	0	737	21	0
2	I	372	0	204	9	0
2	K	372	0	204	10	0
2	M	142	0	80	41	0
3	J	366	0	204	5	0
3	L	366	0	204	19	0
3	N	145	0	80	46	0
4	A	5	0	0	1	0
4	G	5	0	0	0	0
5	A	4	0	0	0	0
5	B	4	0	0	0	0
5	C	5	0	0	0	0
5	D	11	0	0	0	0
5	E	11	0	0	0	0
5	F	4	0	0	0	0
5	G	5	0	0	0	0
5	H	1	0	0	0	0
5	I	2	0	0	0	0
5	J	2	0	0	0	0
5	K	6	0	0	0	0
5	L	6	0	0	1	0
All	All	7646	0	6687	259	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 19.

The worst 5 of 259 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)
3:N:16:DC:H2''	3:N:17:DG:C8	1.22	1.62
3:N:16:DC:C2'	3:N:17:DG:C8	2.02	1.40
1:H:37:SER:OG	1:H:39:PRO:HD2	1.34	1.27

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:35:LYS:NZ	1:B:93:GLU:HB3	1.50	1.25
2:M:1:DA:C2'	2:M:2:DC:H5''	1.75	1.16

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	89/109 (82%)	80 (90%)	8 (9%)	1 (1%)	14	26
1	B	91/109 (84%)	86 (94%)	5 (6%)	0	100	100
1	C	91/109 (84%)	87 (96%)	3 (3%)	1 (1%)	14	26
1	D	90/109 (83%)	83 (92%)	6 (7%)	1 (1%)	14	26
1	E	88/109 (81%)	80 (91%)	7 (8%)	1 (1%)	14	26
1	F	91/109 (84%)	86 (94%)	5 (6%)	0	100	100
1	G	86/109 (79%)	77 (90%)	8 (9%)	1 (1%)	13	24
1	H	93/109 (85%)	82 (88%)	9 (10%)	2 (2%)	6	10
All	All	719/872 (82%)	661 (92%)	51 (7%)	7 (1%)	15	28

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	H	18	GLU
1	G	87	GLY
1	H	6	MET
1	C	86	ASN
1	A	86	ASN

### 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	75/88 (85%)	73 (97%)	2 (3%)	44	71
1	B	76/88 (86%)	74 (97%)	2 (3%)	46	72
1	C	76/88 (86%)	75 (99%)	1 (1%)	69	87
1	D	76/88 (86%)	76 (100%)	0	100	100
1	E	74/88 (84%)	68 (92%)	6 (8%)	11	23
1	F	77/88 (88%)	75 (97%)	2 (3%)	46	72
1	G	71/88 (81%)	70 (99%)	1 (1%)	67	86
1	H	77/88 (88%)	74 (96%)	3 (4%)	32	57
All	All	602/704 (86%)	585 (97%)	17 (3%)	43	70

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	61	ARG
1	E	71	MET
1	G	89	GLN
1	E	20	LEU
1	H	17	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	86	ASN
1	D	89	GLN
1	G	42	ASN
1	C	86	ASN
1	E	89	GLN

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

2 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	$\# Z  > 2$
4	SO4	A	201	-	4,4,4	0.23	0	6,6,6	0.56	0
4	SO4	G	201	-	4,4,4	0.14	0	6,6,6	0.21	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

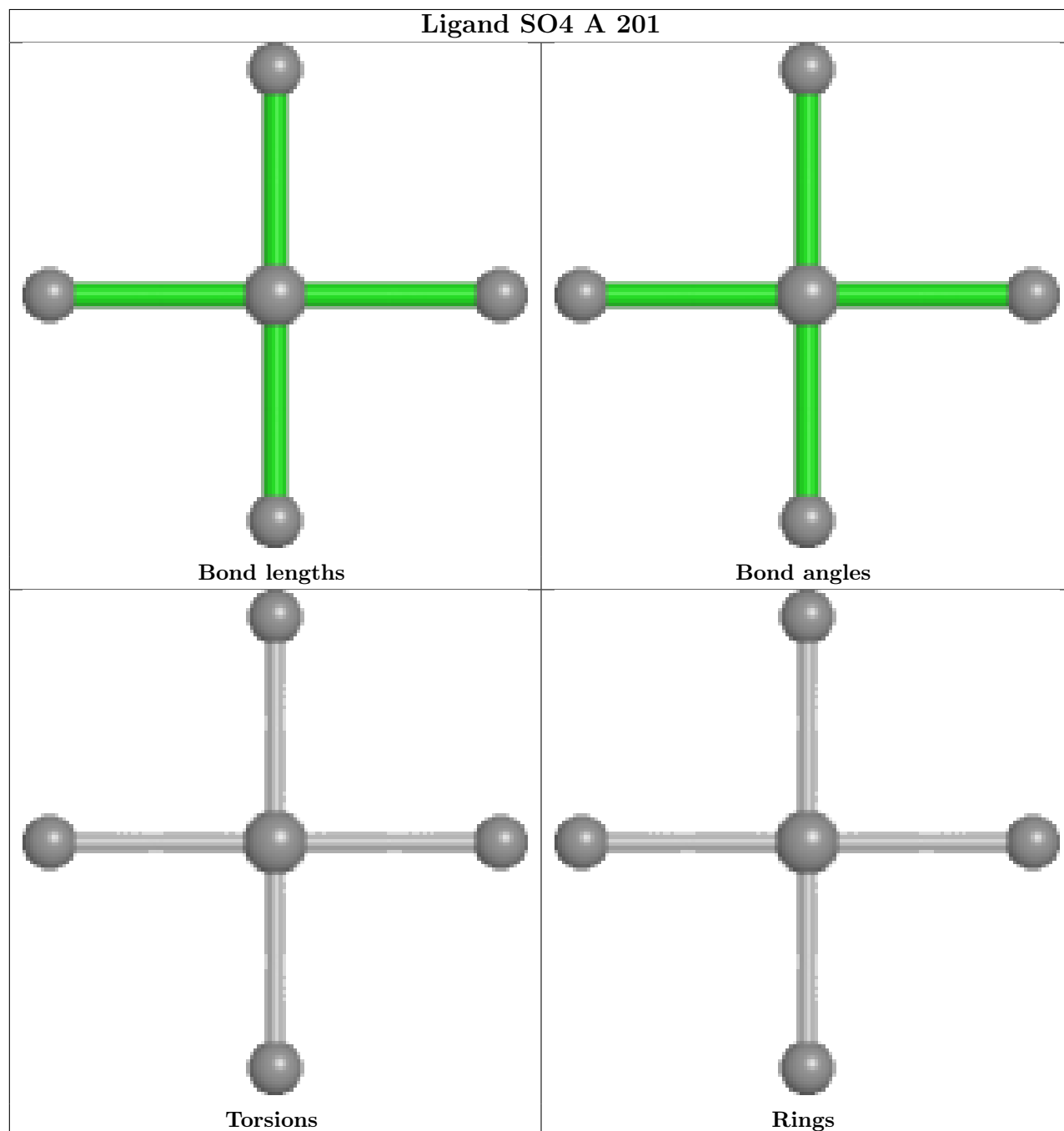
There are no ring outliers.

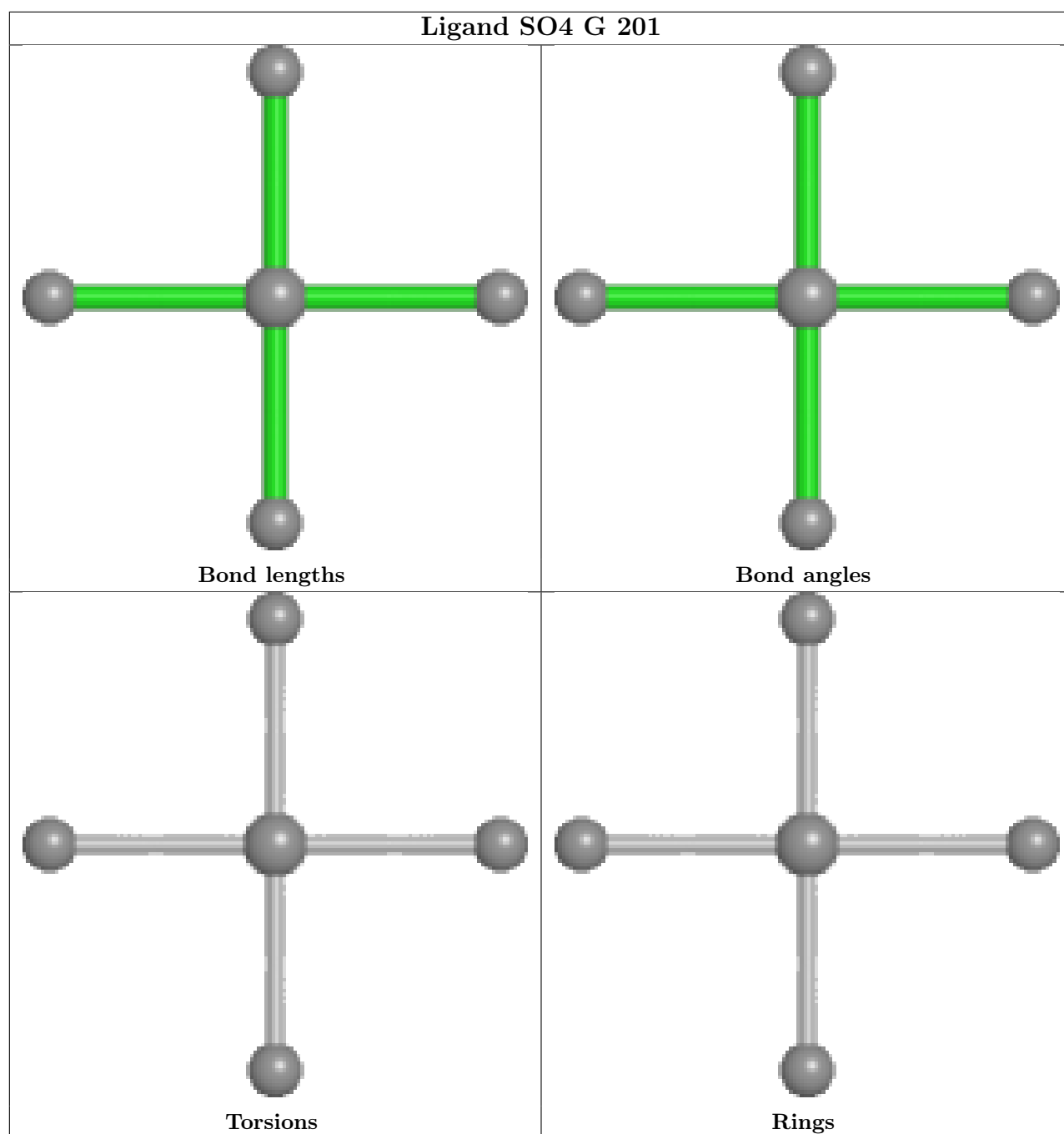
1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	201	SO4	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight  $> 250$  and outliers as shown on the validation Tables will also be included. For torsion angles, if less 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	91/109 (83%)	0.25	0 100 100	47, 60, 84, 87	0
1	B	93/109 (85%)	0.44	2 (2%) 62 65	46, 59, 84, 89	0
1	C	93/109 (85%)	0.21	0 100 100	45, 60, 73, 80	0
1	D	92/109 (84%)	0.24	3 (3%) 46 50	42, 54, 85, 104	0
1	E	90/109 (82%)	0.28	5 (5%) 24 25	44, 52, 89, 95	0
1	F	93/109 (85%)	0.29	3 (3%) 47 51	52, 66, 85, 103	0
1	G	88/109 (80%)	0.45	6 (6%) 17 17	47, 60, 93, 100	0
1	H	95/109 (87%)	0.55	6 (6%) 20 21	57, 72, 87, 94	0
2	I	18/18 (100%)	-0.07	0 100 100	51, 58, 98, 113	0
2	K	18/18 (100%)	-0.18	0 100 100	47, 54, 82, 91	0
2	M	7/18 (38%)	0.92	1 (14%) 2 2	98, 105, 114, 117	0
3	J	18/18 (100%)	-0.13	1 (5%) 24 25	52, 62, 91, 96	0
3	L	18/18 (100%)	-0.01	0 100 100	47, 53, 87, 104	0
3	N	7/18 (38%)	3.41	6 (85%) 0 0	103, 127, 138, 139	0
All	All	821/980 (83%)	0.33	33 (4%) 38 41	42, 61, 90, 139	0

The worst 5 of 33 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	92	HIS	6.2
3	N	15	DG	5.7
3	N	17	DG	4.7
1	H	98	LEU	4.5
1	E	6	MET	4.4

## 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 monosaccharides in this entry.

## 6.4 Ligands [i](#)

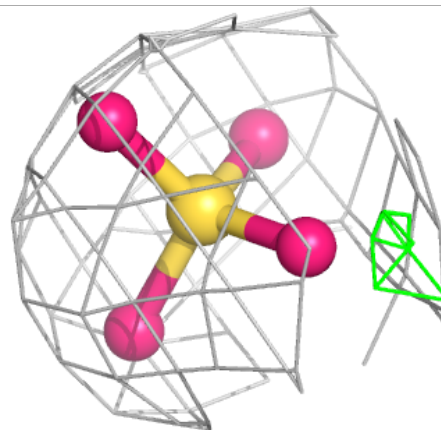
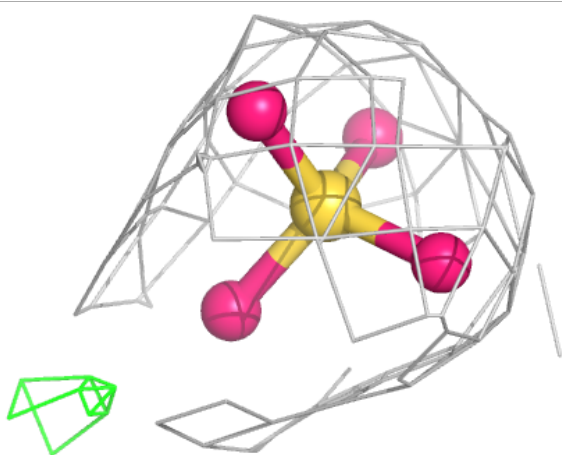
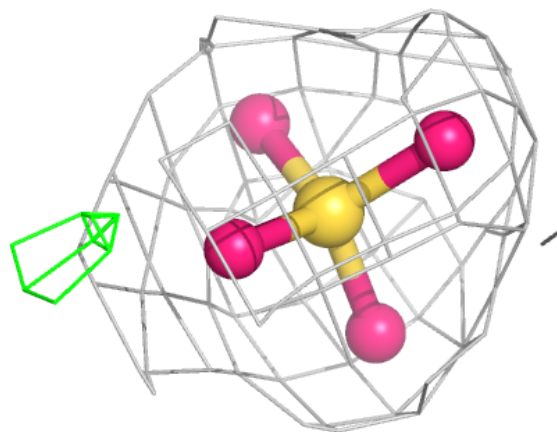
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	SO4	G	201	5/5	0.95	0.15	58,70,88,88	0
4	SO4	A	201	5/5	0.98	0.12	49,59,61,62	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.

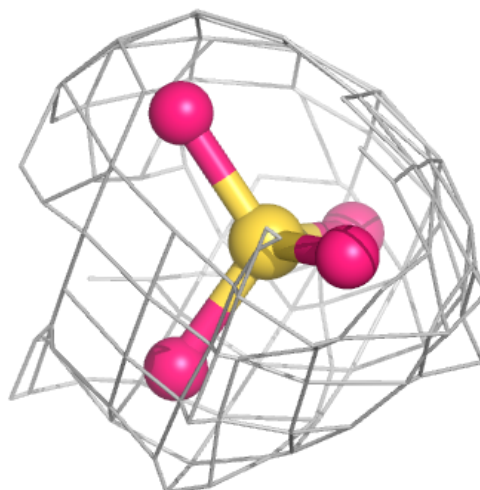
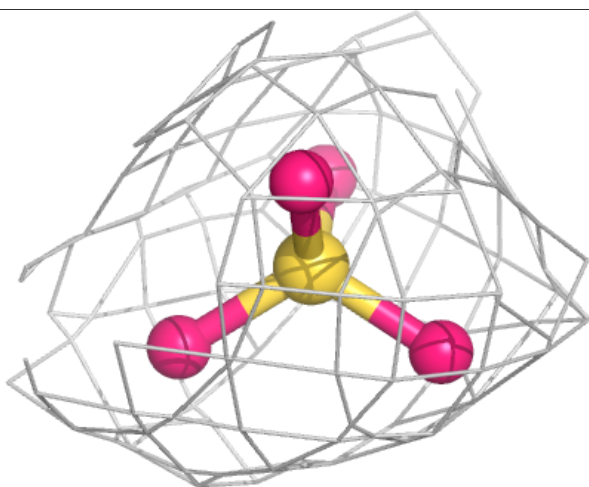
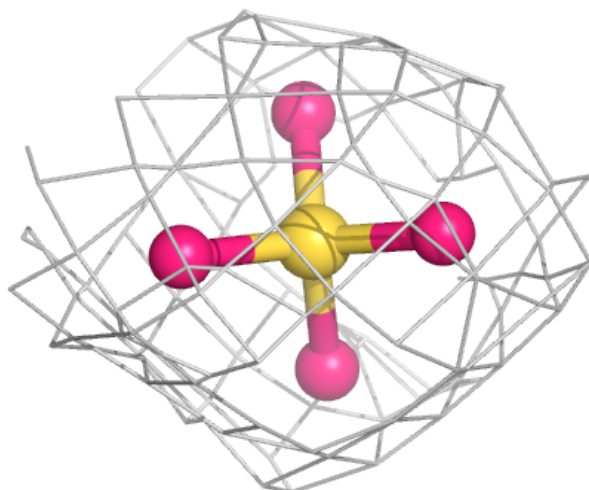
**Electron density around SO4 G 201:**

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



**Electron density around SO4 A 201:**

$2mF_o-DF_c$  (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.