



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 23, 2023 – 10:41 PM EDT

PDB ID : 3AEJ  
Title : Reaction intermediate structure of Entamoeba histolytica methionine gamma-lyase 1 tetramer containing Michaelis complex and methionine-pyridoxal-5'-phosphate  
Authors : Karaki, T.; Sato, D.; Shimizu, A.; Nozaki, T.; Harada, S.  
Deposited on : 2010-02-10  
Resolution : 2.59 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.36  
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) : 2.36

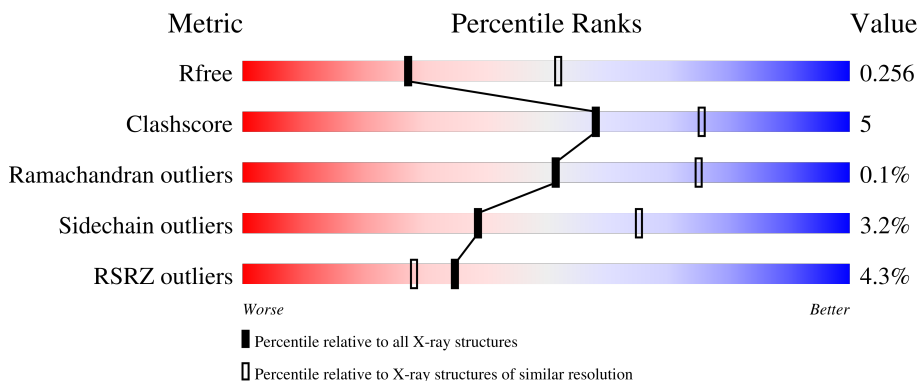
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*


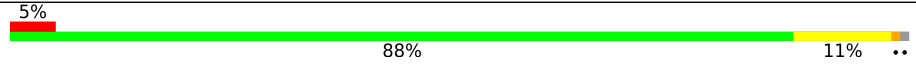
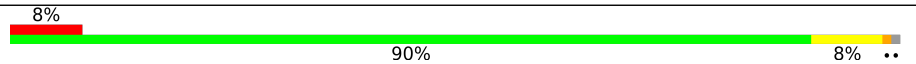

The reported resolution of this entry is 2.59 Å.

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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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  $\geq 3$ , 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	389	 2% 88% 11% ..
1	B	389	 5% 88% 11% ..
1	D	389	 8% 90% 8% ..
2	C	389	 3% 85% 13% ..

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 12308 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 Methionine gamma-lyase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	387	2952	1878	497	553	24	0	1	0
1	B	386	2937	1867	495	551	24	0	1	0
1	D	384	2920	1857	492	548	23	0	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	308	LEU	SER	SEE REMARK 999	UNP Q86D28
B	808	LEU	SER	SEE REMARK 999	UNP Q86D28
D	1808	LEU	SER	SEE REMARK 999	UNP Q86D28

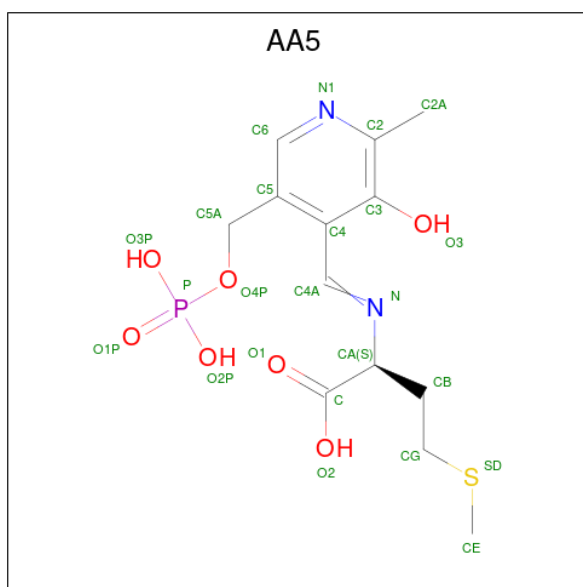
- Molecule 2 is a protein called Methionine gamma-lyase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	N	O	P				S
2	C	387	2967	1886	498	558	1	24	0	1	0

There is a discrepancy between the modelled and reference sequences:

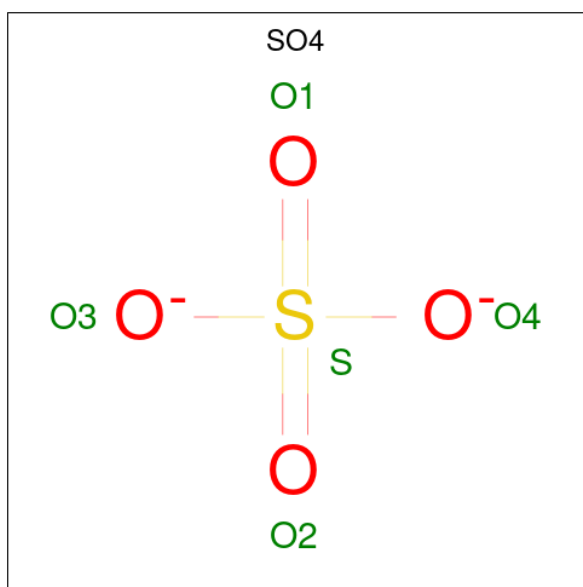
Chain	Residue	Modelled	Actual	Comment	Reference
C	1308	LEU	SER	SEE REMARK 999	UNP Q86D28

- Molecule 3 is N-[(3-HYDROXY-2-METHYL-5-[(TRIHYDROXYPHOSPHORANYL)OXY]METHYL}PYRIDIN-4-YL)METHYLENE]METHIONINE (three-letter code: AA5) (formula: C<sub>13</sub>H<sub>19</sub>N<sub>2</sub>O<sub>7</sub>PS).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	N	O	P			S
3	A	1	Total	C	N	O	P	S	0	0
			24	13	2	7	1	1		
3	B	1	Total	C	N	O	P	S	0	0
			24	13	2	7	1	1		
3	D	1	Total	C	N	O	P	S	0	0
			24	13	2	7	1	1		

- Molecule 4 is SULFATE ION (three-letter code: SO<sub>4</sub>) (formula: O<sub>4</sub>S).



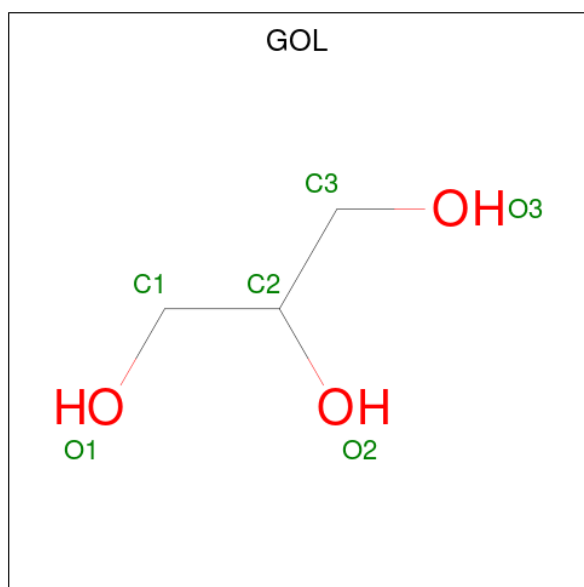
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
4	A	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

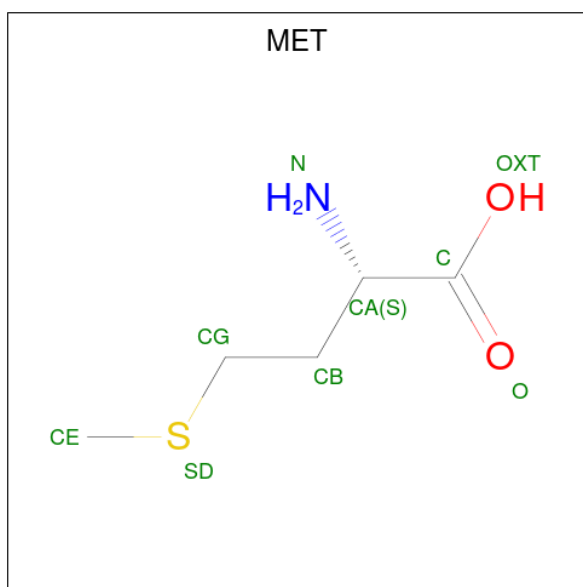
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	O	S	0	0
			5	4	1		
4	C	1	Total	O	S	0	0
			5	4	1		
4	D	1	Total	O	S	0	0
			5	4	1		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			6	3	3		
5	C	1	Total	C	O	0	0
			6	3	3		

- Molecule 6 is METHIONINE (three-letter code: MET) (formula:  $C_5H_{11}NO_2S$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
6	C	1	9	5	1	2	1	0	0


- Molecule 7 is water.

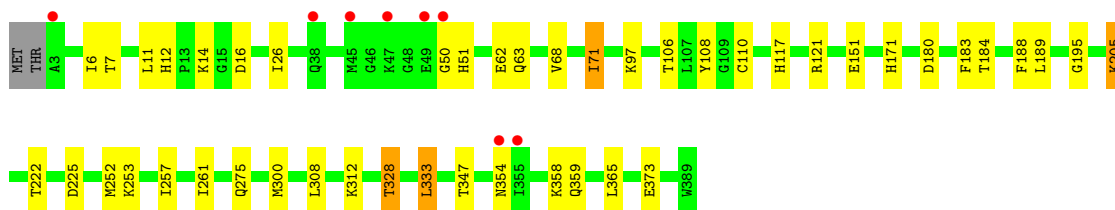
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	123	Total	O	0	0
			123	123		
7	B	78	Total	O	0	0
			78	78		
7	C	134	Total	O	0	0
			134	134		
7	D	84	Total	O	0	0
			84	84		

### 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.

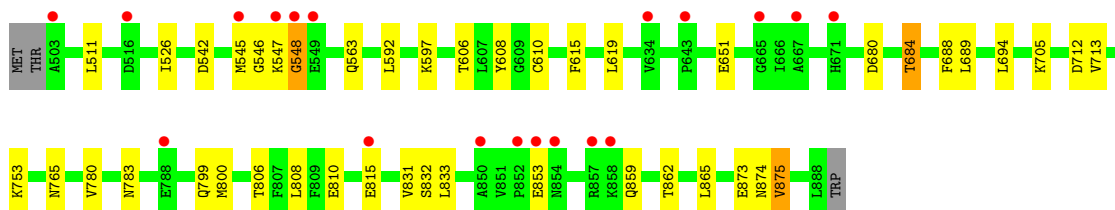
- Molecule 1: Methionine gamma-lyase

Chain A: 




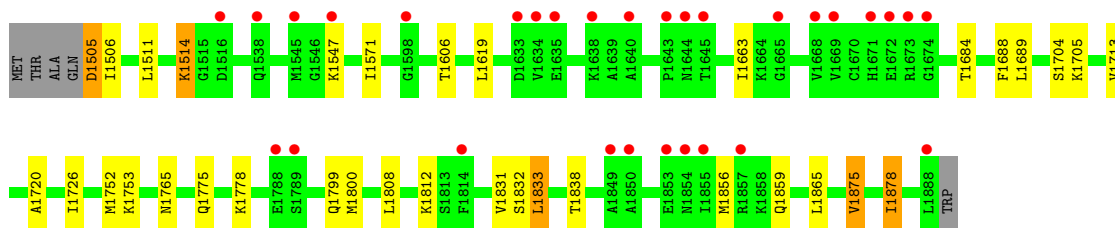
- Molecule 1: Methionine gamma-lyase

Chain B: 




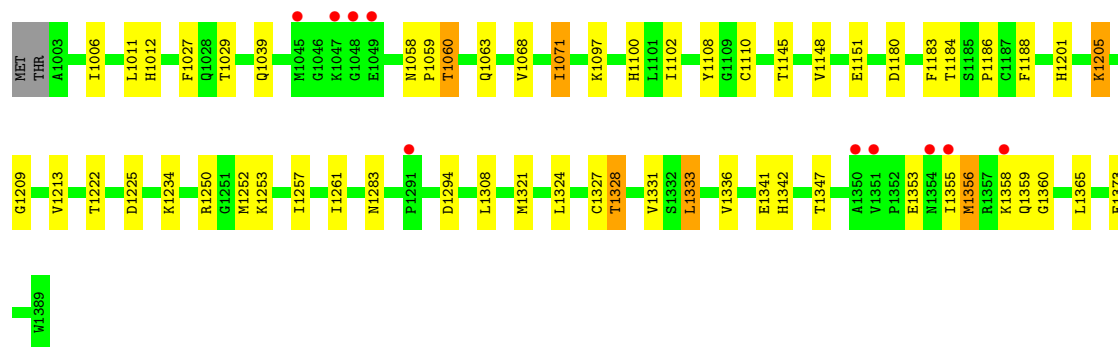
- Molecule 1: Methionine gamma-lyase

Chain D: 



- Molecule 2: Methionine gamma-lyase

Chain C: 





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	98.74Å 85.36Å 113.66Å 90.00° 101.92° 90.00°	Depositor
Resolution (Å)	39.04 – 2.59 39.04 – 2.59	Depositor EDS
% Data completeness (in resolution range)	99.4 (39.04-2.59) 99.4 (39.04-2.59)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.87 (at 2.58Å)	Xtrriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.209 , 0.256 0.208 , 0.256	Depositor DCC
$R_{free}$ test set	2924 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	26.7	Xtrriage
Anisotropy	0.355	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 29.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	12308	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 13.55% 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, AA5, LLP, GOL

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	0/3018	0.48	0/4080
1	B	0.32	0/3001	0.46	0/4057
1	D	0.32	0/2981	0.46	0/4030
2	C	0.32	0/3008	0.48	0/4066
All	All	0.32	0/12008	0.47	0/16233

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	2952	0	2957	32	0
1	B	2937	0	2947	30	0
1	D	2920	0	2929	22	0
2	C	2967	0	2960	42	0
3	A	24	0	16	4	0
3	B	24	0	15	2	0
3	D	24	0	15	2	0
4	A	5	0	0	0	0
4	B	5	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	C	5	0	0	0	0
4	D	5	0	0	0	0
5	A	6	0	8	1	0
5	C	6	0	8	0	0
6	C	9	0	8	4	0
7	A	123	0	0	3	0
7	B	78	0	0	1	0
7	C	134	0	0	3	0
7	D	84	0	0	0	0
All	All	12308	0	11863	115	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 (115) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:1353:GLU:HA	2:C:1356:MET:HB3	1.66	0.78
2:C:1071:ILE:HD11	2:C:1252:MET:HE1	1.69	0.74
1:D:1808:LEU:HD21	1:D:1865:LEU:HD13	1.69	0.73
2:C:1347:THR:HG22	6:C:2003:MET:HB3	1.70	0.73
1:A:261:ILE:HD11	2:C:1257:ILE:HG12	1.72	0.70
1:B:874:ASN:HB2	1:D:1505:ASP:HB2	1.72	0.70
1:A:71:ILE:HD11	1:A:252:MET:CE	2.22	0.69
1:B:651:GLU:HG2	1:B:680:ASP:HB3	1.72	0.69
2:C:1071:ILE:HD11	2:C:1252:MET:CE	2.24	0.68
2:C:1308:LEU:HD21	2:C:1365:LEU:HD13	1.74	0.68
1:D:1705:LYS:HZ2	3:D:2004:AA5:HA	1.58	0.68
1:A:205:LYS:HZ1	3:A:2001:AA5:HA	1.62	0.64
1:B:545:MET:H	1:B:546:GLY:HA2	1.62	0.64
1:D:1808:LEU:HD11	1:D:1865:LEU:HB3	1.80	0.63
1:A:11:LEU:HD12	2:C:1373:GLU:HG3	1.79	0.62
1:B:610:CYS:SG	2:C:1234:LYS:NZ	2.71	0.62
1:D:1511:LEU:HD23	1:D:1753:LYS:HD2	1.80	0.62
1:D:1606:THR:HA	1:D:1859:GLN:HE21	1.65	0.61
1:A:222:THR:HG22	1:A:225:ASP:H	1.66	0.60
1:D:1506:ILE:HG21	1:D:1571:ILE:HD13	1.82	0.60
1:A:71:ILE:HD11	1:A:252:MET:HE1	1.82	0.60
1:B:705:LYS:HZ1	3:B:2002:AA5:HA	1.66	0.60
1:D:1705:LYS:HE3	3:D:2004:AA5:H4A	1.83	0.59
1:A:205:LYS:HG2	1:A:333:LEU:HG	1.83	0.59

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:257:ILE:HG12	2:C:1261:ILE:HD11	1.85	0.59
1:B:873:GLU:HG3	1:D:1511:LEU:HD12	1.85	0.59
1:A:308:LEU:HD21	1:A:365:LEU:HD13	1.84	0.58
1:A:51:HIS:HE1	1:A:62:GLU:OE1	1.88	0.57
1:A:328:THR:HG22	7:A:3014:HOH:O	2.05	0.56
1:D:1684:THR:HG22	1:D:1688:PHE:HB2	1.86	0.56
1:A:373:GLU:HG3	2:C:1011:LEU:HD12	1.86	0.56
2:C:1068:VAL:HA	2:C:1071:ILE:HD11	1.86	0.56
2:C:1222:THR:HG22	2:C:1225:ASP:H	1.70	0.56
1:B:808:LEU:HD11	1:B:865:LEU:HB3	1.87	0.56
1:A:184:THR:HG22	1:A:188:PHE:HB2	1.87	0.55
2:C:1328:THR:HG22	7:C:3001:HOH:O	2.06	0.55
1:B:608:TYR:CE1	3:B:2002:AA5:HB1	2.41	0.55
2:C:1184:THR:HG22	2:C:1188:PHE:HB2	1.89	0.54
1:A:71:ILE:HD11	1:A:252:MET:HE2	1.90	0.54
1:B:545:MET:N	1:B:546:GLY:HA2	2.18	0.54
2:C:1308:LEU:HD11	2:C:1365:LEU:HB3	1.90	0.54
1:A:11:LEU:HD22	2:C:1336:VAL:HG12	1.90	0.53
1:B:808:LEU:HD21	1:B:865:LEU:HD13	1.91	0.53
1:D:1514:LYS:HD3	1:D:1514:LYS:H	1.75	0.52
1:A:151:GLU:HG2	1:A:180:ASP:HB3	1.92	0.52
1:D:1838:THR:HG21	1:D:1878:ILE:HD13	1.92	0.51
1:A:205:LYS:NZ	3:A:2001:AA5:HA	2.26	0.51
1:B:542:ASP:HB3	1:B:547:LYS:O	2.11	0.51
1:B:592:LEU:HD21	1:B:619:LEU:HD22	1.92	0.50
2:C:1027:PHE:H	2:C:1060:THR:HG21	1.76	0.50
1:B:511:LEU:HD23	1:B:753:LYS:HG3	1.93	0.50
1:B:705:LYS:HD2	1:B:833:LEU:HG	1.93	0.50
1:A:205:LYS:NZ	3:A:2001:AA5:H4A	2.27	0.50
1:B:597:LYS:HE3	2:C:1097:LYS:HE3	1.94	0.50
7:B:3010:HOH:O	2:C:1331:VAL:HG22	2.11	0.50
1:B:606:THR:HA	1:B:859:GLN:HE21	1.76	0.50
2:C:1068:VAL:HA	2:C:1071:ILE:CD1	2.42	0.49
2:C:1359:GLN:N	2:C:1360:GLY:HA2	2.26	0.49
1:D:1704:SER:HB2	1:D:1705:LYS:HE2	1.94	0.49
2:C:1027:PHE:H	2:C:1060:THR:CG2	2.25	0.49
2:C:1250:ARG:O	2:C:1253:LYS:HB2	2.13	0.49
1:B:862:THR:H	1:B:865:LEU:HD12	1.79	0.48
2:C:1108:TYR:HE1	6:C:2003:MET:HB2	1.78	0.48
1:A:6:ILE:HG12	1:A:71:ILE:HG23	1.95	0.48
1:B:542:ASP:O	1:B:546:GLY:HA2	2.14	0.48

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:68:VAL:HA	1:A:71:ILE:CD1	2.45	0.47
1:A:171:HIS:CE1	1:A:195:GLY:O	2.68	0.47
1:A:108:TYR:CE2	1:A:110:CYS:HB2	2.50	0.47
2:C:1205:LLP:H4'1	6:C:2003:MET:N	2.30	0.47
1:D:1663:ILE:H	1:D:1799:GLN:HE22	1.61	0.47
2:C:1102:ILE:O	2:C:1148:VAL:HA	2.14	0.46
2:C:1108:TYR:CE2	2:C:1110:CYS:HB2	2.50	0.46
1:D:1689:LEU:HD13	1:D:1800:MET:HG2	1.96	0.46
1:B:608:TYR:CE2	1:B:610:CYS:HB2	2.50	0.46
1:A:12:HIS:HE1	7:C:3001:HOH:O	1.99	0.46
1:A:7:THR:O	1:A:11:LEU:HG	2.16	0.45
1:D:1778:LYS:HA	1:D:1812:LYS:HE3	1.97	0.45
2:C:1006:ILE:HG12	2:C:1071:ILE:HG23	1.98	0.45
1:D:1720:ALA:HB3	1:D:1726:ILE:HB	1.98	0.45
1:D:1705:LYS:HD2	1:D:1833:LEU:HG	1.97	0.45
7:A:3014:HOH:O	2:C:1012:HIS:HE1	2.00	0.45
2:C:1321:MET:CE	2:C:1342:HIS:HB2	2.46	0.45
2:C:1100:HIS:HB3	2:C:1145:THR:HA	1.99	0.45
1:A:106:THR:HG22	1:A:359:GLN:HG2	1.99	0.45
2:C:1151:GLU:HG2	2:C:1180:ASP:HB3	1.99	0.44
1:B:689:LEU:HD13	1:B:800:MET:HG2	1.99	0.44
2:C:1186:PRO:HD3	2:C:1201:HIS:CE1	2.52	0.44
1:A:68:VAL:HA	1:A:71:ILE:HD11	2.00	0.44
2:C:1108:TYR:CE1	6:C:2003:MET:HB2	2.52	0.44
2:C:1205:LLP:HD3	2:C:1333:LEU:HG	1.98	0.44
1:A:347:THR:HG22	3:A:2001:AA5:HG2	2.00	0.43
1:B:684:THR:HG22	1:B:688:PHE:HB2	1.99	0.43
1:A:26:ILE:HB	1:B:526:ILE:HB	1.98	0.43
1:D:1514:LYS:H	1:D:1514:LYS:CD	2.31	0.43
2:C:1331:VAL:HG12	2:C:1341:GLU:HG3	2.00	0.43
1:B:712:ASP:HA	2:C:1029:THR:HA	2.00	0.43
2:C:1058:ASN:HA	2:C:1059:PRO:HD3	1.94	0.43
2:C:1222:THR:HG21	7:C:3403:HOH:O	2.18	0.43
1:A:253:LYS:HE3	7:A:3083:HOH:O	2.18	0.42
2:C:1209:GLY:O	2:C:1336:VAL:HG23	2.20	0.42
1:A:189:LEU:HD13	1:A:300:MET:HG2	2.01	0.42
1:B:694:LEU:HD13	1:B:799:GLN:HE21	1.84	0.42
1:D:1765:ASN:ND2	1:D:1875:VAL:HG22	2.35	0.42
1:B:765:ASN:ND2	1:B:875:VAL:HG22	2.35	0.41
1:A:312:LYS:HD3	5:A:2009:GOL:H31	2.02	0.41
1:B:806:THR:HG21	1:B:833:LEU:HD22	2.02	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:117:HIS:O	1:A:121:ARG:HG3	2.21	0.41
1:B:615:PHE:HA	1:B:619:LEU:HG	2.02	0.41
2:C:1324:LEU:HD13	2:C:1327:CYS:HB2	2.03	0.41
1:B:831:VAL:O	1:B:832:SER:HB3	2.19	0.41
1:B:547:LYS:HA	1:B:548:GLY:HA2	1.74	0.40
1:D:1831:VAL:O	1:D:1832:SER:HB3	2.22	0.40
1:B:780:VAL:HB	1:B:810:GLU:HB3	2.04	0.40
2:C:1355:ILE:HA	2:C:1358:LYS:HE2	2.03	0.40
1:D:1506:ILE:CG2	1:D:1571:ILE:HD13	2.51	0.40

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	386/389 (99%)	376 (97%)	9 (2%)	1 (0%)	41	64
1	B	385/389 (99%)	376 (98%)	8 (2%)	1 (0%)	41	64
1	D	382/389 (98%)	371 (97%)	11 (3%)	0	100	100
2	C	385/389 (99%)	377 (98%)	8 (2%)	0	100	100
All	All	1538/1556 (99%)	1500 (98%)	36 (2%)	2 (0%)	51	75

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	50	GLY
1	B	548	GLY

### 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	321/322 (100%)	309 (96%)	12 (4%)	34	60
1	B	320/322 (99%)	313 (98%)	7 (2%)	52	76
1	D	318/322 (99%)	307 (96%)	11 (4%)	36	62
2	C	320/321 (100%)	309 (97%)	11 (3%)	37	63
All	All	1279/1287 (99%)	1238 (97%)	41 (3%)	39	65

All (41) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	14	LYS
1	A	16	ASP
1	A	63	GLN
1	A	71	ILE
1	A	97	LYS
1	A	183	PHE
1	A	205	LYS
1	A	275	GLN
1	A	328	THR
1	A	333	LEU
1	A	354	ASN
1	A	358	LYS
1	B	563	GLN
1	B	684	THR
1	B	713	VAL
1	B	783	ASN
1	B	815	GLU
1	B	853	GLU
1	B	875	VAL
2	C	1039	GLN
2	C	1060	THR
2	C	1063	GLN
2	C	1071	ILE
2	C	1183	PHE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	C	1213	VAL
2	C	1283	ASN
2	C	1294	ASP
2	C	1328	THR
2	C	1333	LEU
2	C	1356	MET
1	D	1505	ASP
1	D	1514	LYS
1	D	1547	LYS
1	D	1619	LEU
1	D	1713	VAL
1	D	1752	MET
1	D	1775	GLN
1	D	1833	LEU
1	D	1856	MET
1	D	1875	VAL
1	D	1878	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (35) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	51	HIS
1	A	100	HIS
1	A	171	HIS
1	A	265	ASN
1	A	275	GLN
1	A	293	HIS
1	A	299	GLN
1	A	354	ASN
1	A	374	ASN
1	B	504	GLN
1	B	600	HIS
1	B	760	GLN
1	B	765	ASN
1	B	783	ASN
1	B	793	HIS
1	B	799	GLN
1	B	842	HIS
1	B	874	ASN
2	C	1039	GLN
2	C	1063	GLN
2	C	1100	HIS

*Continued on next page...*



Continued from previous page...

Mol	Chain	Res	Type
2	C	1265	ASN
2	C	1283	ASN
2	C	1293	HIS
2	C	1299	GLN
2	C	1323	HIS
2	C	1348	HIS
2	C	1359	GLN
2	C	1374	ASN
1	D	1600	HIS
1	D	1765	ASN
1	D	1793	HIS
1	D	1799	GLN
1	D	1854	ASN
1	D	1859	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	$\# Z  > 2$
2	LLP	C	1205	2	23,24,25	1.74	5 (21%)	25,32,34	1.74	4 (16%)

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
2	LLP	C	1205	2	-	7/16/17/19	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	1205	LLP	O3-C3	-5.74	1.23	1.37
2	C	1205	LLP	C2-N1	2.67	1.38	1.33
2	C	1205	LLP	C4-C4'	2.30	1.51	1.46
2	C	1205	LLP	C6-N1	2.19	1.39	1.34
2	C	1205	LLP	C4'-NZ	2.00	1.34	1.27

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	1205	LLP	OP4-C5'-C5	6.23	121.22	109.35
2	C	1205	LLP	C4-C4'-NZ	-3.18	109.72	124.31
2	C	1205	LLP	CE-NZ-C4'	-2.32	111.77	118.90
2	C	1205	LLP	C5-C6-N1	-2.21	120.14	123.82

There are no chirality outliers.

All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	1205	LLP	O-C-CA-CB
2	C	1205	LLP	C4-C4'-NZ-CE
2	C	1205	LLP	C3-C4-C4'-NZ
2	C	1205	LLP	CD-CE-NZ-C4'
2	C	1205	LLP	C5'-OP4-P-OP2
2	C	1205	LLP	C5-C4-C4'-NZ
2	C	1205	LLP	N-CA-CB-CG

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	1205	LLP	2	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

10 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	D	2008	-	4,4,4	0.13	0	6,6,6	0.06	0
6	MET	C	2003	-	7,8,8	0.92	1 (14%)	7,9,9	1.39	2 (28%)
4	SO4	C	2007	-	4,4,4	0.14	0	6,6,6	0.05	0
3	AA5	D	2004	-	24,24,24	2.16	2 (8%)	29,33,33	1.25	1 (3%)
3	AA5	B	2002	-	24,24,24	2.24	2 (8%)	29,33,33	1.57	2 (6%)
4	SO4	A	2005	-	4,4,4	0.14	0	6,6,6	0.05	0
3	AA5	A	2001	-	24,24,24	2.15	2 (8%)	29,33,33	1.34	2 (6%)
5	GOL	A	2009	-	5,5,5	0.36	0	5,5,5	0.40	0
4	SO4	B	2006	-	4,4,4	0.13	0	6,6,6	0.06	0
5	GOL	C	2010	-	5,5,5	0.36	0	5,5,5	0.26	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	MET	C	2003	-	-	2/8/8/8	-
3	AA5	D	2004	-	-	8/19/19/19	0/1/1/1
3	AA5	B	2002	-	-	6/19/19/19	0/1/1/1
5	GOL	A	2009	-	-	2/4/4/4	-
3	AA5	A	2001	-	-	9/19/19/19	0/1/1/1
5	GOL	C	2010	-	-	2/4/4/4	-

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	2002	AA5	C3-C2	8.40	1.49	1.40
3	A	2001	AA5	C3-C2	8.20	1.49	1.40
3	D	2004	AA5	C3-C2	8.13	1.49	1.40
3	B	2002	AA5	C4-C3	5.84	1.49	1.40
3	D	2004	AA5	C4-C3	5.38	1.48	1.40

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	2001	AA5	C4-C3	5.37	1.48	1.40
6	C	2003	MET	OXT-C	-2.38	1.22	1.30

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	2002	AA5	CA-N-C4A	5.46	125.19	117.31
3	B	2002	AA5	C4-C3-C2	-4.09	117.66	120.19
3	A	2001	AA5	C4-C3-C2	-3.88	117.79	120.19
3	D	2004	AA5	C4-C3-C2	-3.84	117.81	120.19
6	C	2003	MET	OXT-C-O	-3.01	117.25	124.09
3	A	2001	AA5	CA-N-C4A	2.19	120.47	117.31
6	C	2003	MET	OXT-C-CA	2.03	120.31	113.38

There are no chirality outliers.

All (29) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	2001	AA5	O1-C-CA-CB
3	A	2001	AA5	N-CA-CB-CG
3	A	2001	AA5	C-CA-CB-CG
3	A	2001	AA5	C3-C4-C4A-N
3	A	2001	AA5	C5A-O4P-P-O3P
3	B	2002	AA5	CB-CA-N-C4A
3	B	2002	AA5	N-CA-CB-CG
3	B	2002	AA5	C-CA-CB-CG
3	D	2004	AA5	N-CA-CB-CG
3	D	2004	AA5	C-CA-CB-CG
3	D	2004	AA5	C5A-O4P-P-O2P
3	D	2004	AA5	C5A-O4P-P-O3P
5	A	2009	GOL	O1-C1-C2-C3
3	D	2004	AA5	C3-C4-C4A-N
5	C	2010	GOL	O1-C1-C2-C3
3	A	2001	AA5	CB-CG-SD-CE
3	B	2002	AA5	CB-CG-SD-CE
5	A	2009	GOL	O1-C1-C2-O2
3	D	2004	AA5	C5A-O4P-P-O1P
3	D	2004	AA5	C5-C4-C4A-N
3	D	2004	AA5	CB-CG-SD-CE
3	A	2001	AA5	C5A-O4P-P-O2P
3	A	2001	AA5	O2-C-CA-CB
6	C	2003	MET	CA-CB-CG-SD

*Continued on next page...*

*Continued from previous page...*

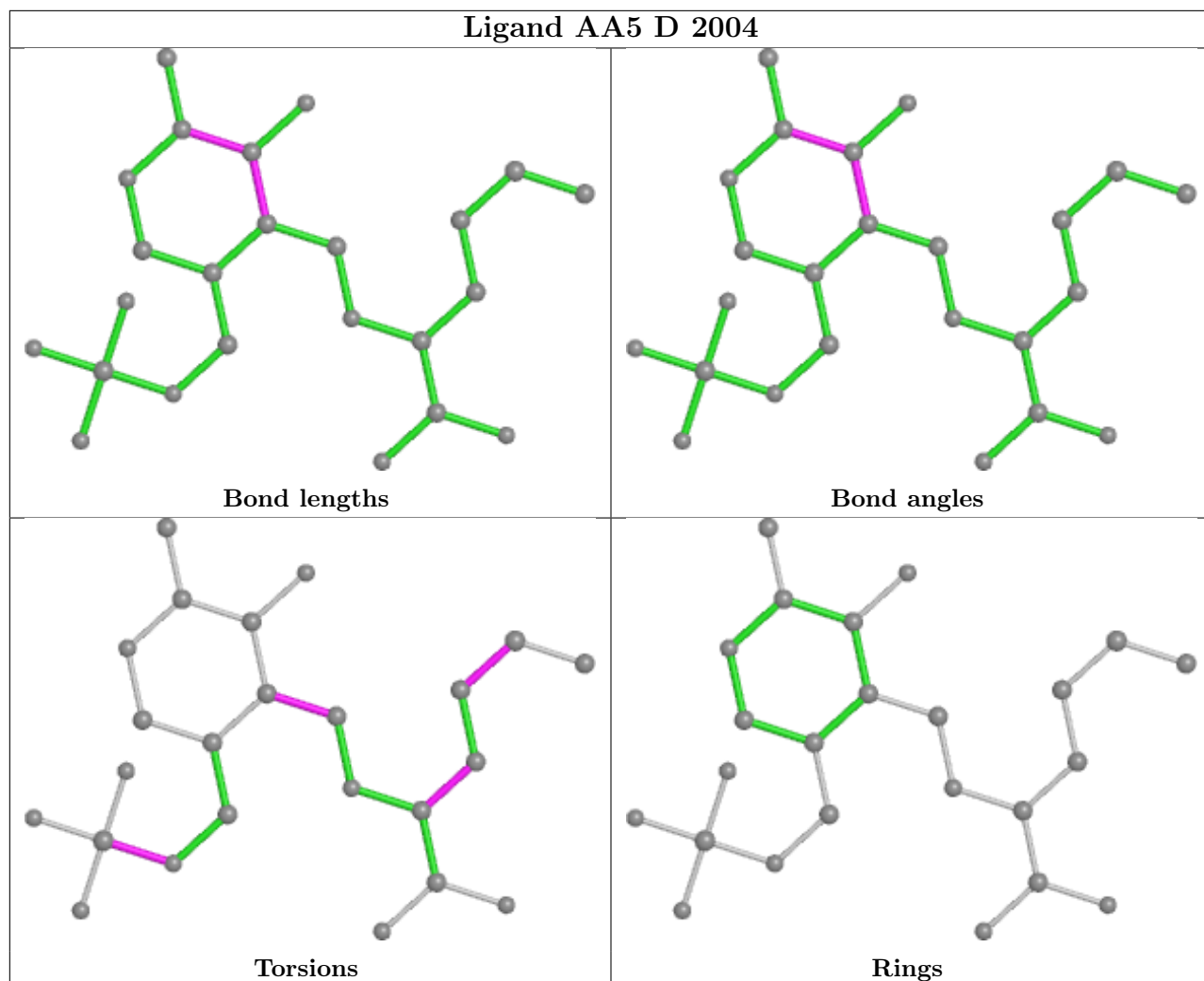
Mol	Chain	Res	Type	Atoms
3	B	2002	AA5	O1-C-CA-CB
3	B	2002	AA5	O2-C-CA-CB
6	C	2003	MET	OXT-C-CA-N
5	C	2010	GOL	O1-C1-C2-O2
3	A	2001	AA5	O1-C-CA-N

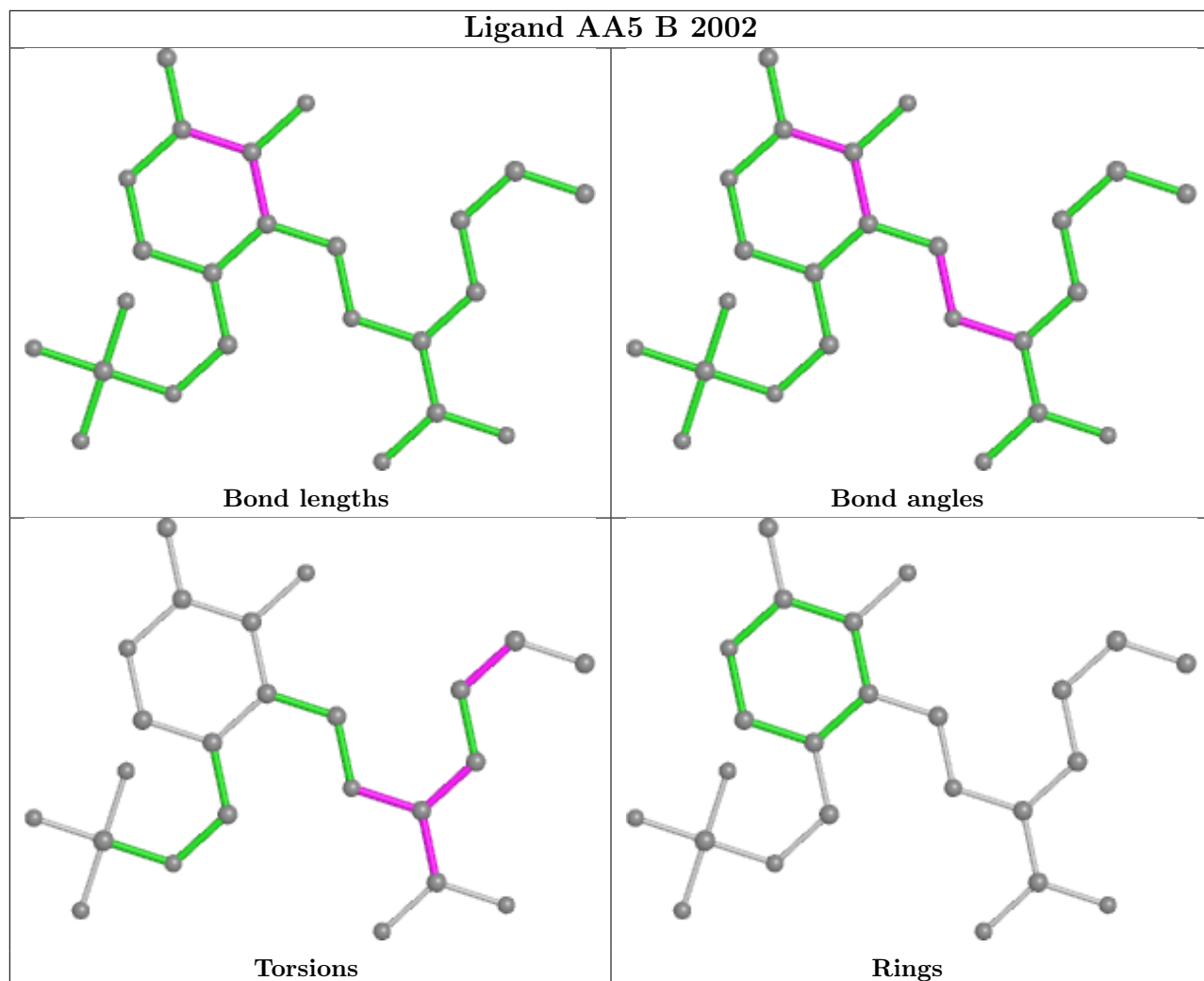
There are no ring outliers.

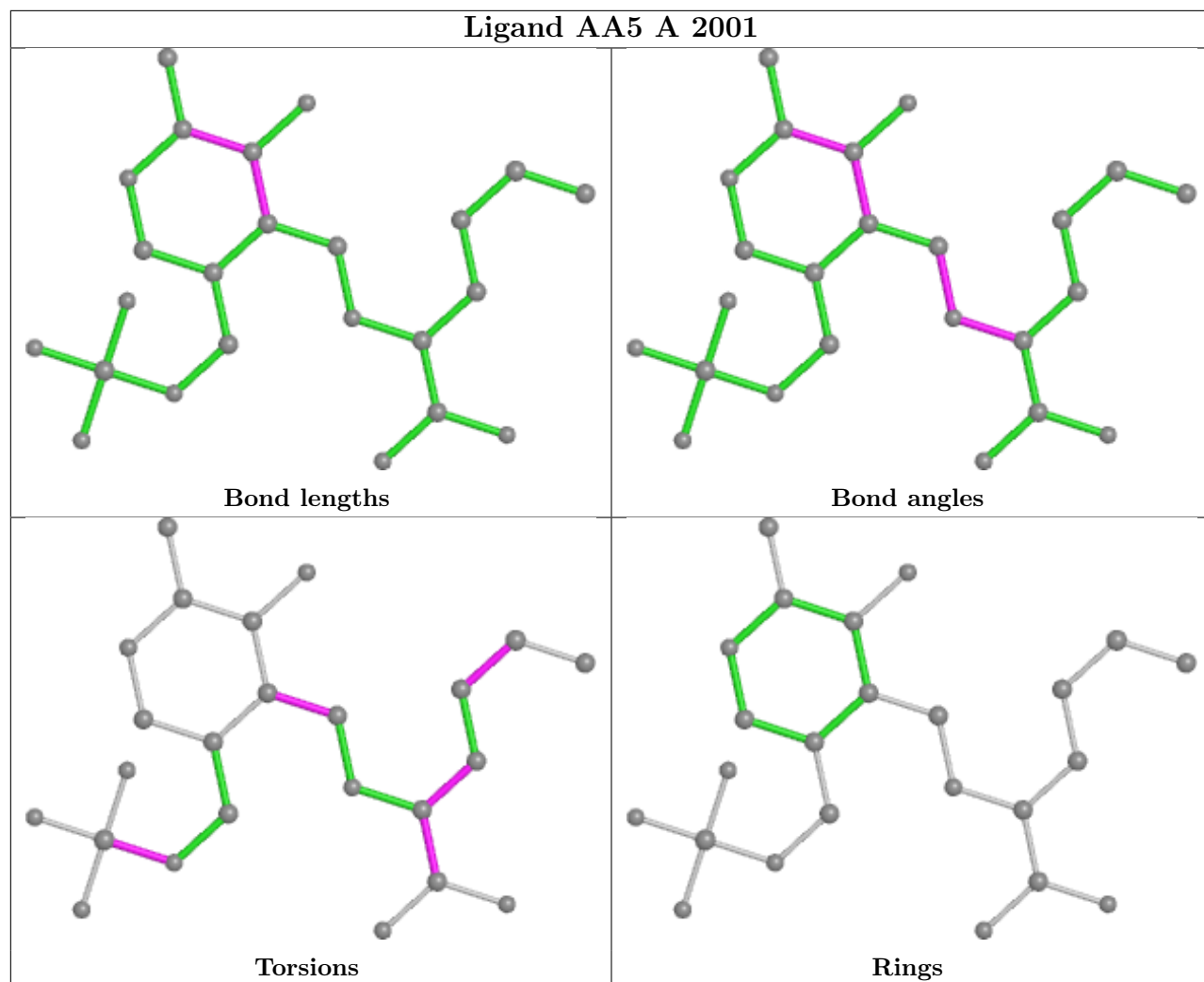
5 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	C	2003	MET	4	0
3	D	2004	AA5	2	0
3	B	2002	AA5	2	0
3	A	2001	AA5	4	0
5	A	2009	GOL	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	387/389 (99%)	0.01	8 (2%) 63 58	14, 20, 43, 62	1 (0%)
1	B	386/389 (99%)	0.25	19 (4%) 29 23	18, 31, 53, 66	1 (0%)
1	D	384/389 (98%)	0.31	30 (7%) 13 9	18, 32, 56, 66	0
2	C	386/389 (99%)	0.00	10 (2%) 56 50	14, 20, 41, 57	1 (0%)
All	All	1543/1556 (99%)	0.14	67 (4%) 35 28	14, 25, 51, 66	3 (0%)

All (67) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	1634	VAL	3.6
1	A	47	LYS	3.6
2	C	1047	LYS	3.4
1	B	548	GLY	3.4
1	D	1857	ARG	3.4
1	D	1671	HIS	3.3
1	B	853	GLU	3.3
1	B	516	ASP	3.2
1	B	854	ASN	3.2
1	B	545	MET	3.2
1	B	857	ARG	3.2
1	B	850	ALA	3.1
1	D	1854	ASN	3.1
2	C	1049	GLU	3.0
1	B	858	LYS	3.0
2	C	1045	MET	3.0
1	A	49	GLU	2.9
1	D	1643	PRO	2.9
2	C	1351	VAL	2.9
1	A	3	ALA	2.8
2	C	1354	ASN	2.8

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	B	788	GLU	2.7
1	A	354	ASN	2.7
1	B	634	VAL	2.6
1	D	1547	LYS	2.6
1	B	503	ALA	2.6
1	B	547	LYS	2.6
1	A	50	GLY	2.6
2	C	1048	GLY	2.5
1	D	1850	ALA	2.5
1	D	1888	LEU	2.5
1	D	1789	SER	2.5
1	B	852	PRO	2.4
1	D	1849	ALA	2.4
1	D	1516	ASP	2.4
1	D	1598	GLY	2.4
1	B	671	HIS	2.4
1	D	1635	GLU	2.4
1	D	1674	GLY	2.4
2	C	1350	ALA	2.3
1	B	665	GLY	2.3
1	A	38	GLN	2.3
2	C	1291	PRO	2.3
1	A	45	MET	2.3
2	C	1355	ILE	2.3
1	D	1665	GLY	2.3
1	D	1788	GLU	2.3
1	D	1644	ASN	2.3
1	D	1855	ILE	2.3
1	D	1814	PHE	2.3
1	D	1538	GLN	2.2
1	D	1673	ARG	2.2
1	D	1645	THR	2.2
1	D	1669	VAL	2.2
1	D	1672	GLU	2.2
1	D	1545	MET	2.2
1	D	1638	LYS	2.2
2	C	1358	LYS	2.1
1	D	1640	ALA	2.1
1	B	667	ALA	2.1
1	D	1668	VAL	2.1
1	D	1633	ASP	2.1
1	D	1853	GLU	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	815	GLU	2.1
1	A	355	ILE	2.1
1	B	643	PRO	2.0
1	B	549	GLU	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [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
2	LLP	C	1205	24/25	0.96	0.16	15,16,17,18	6

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

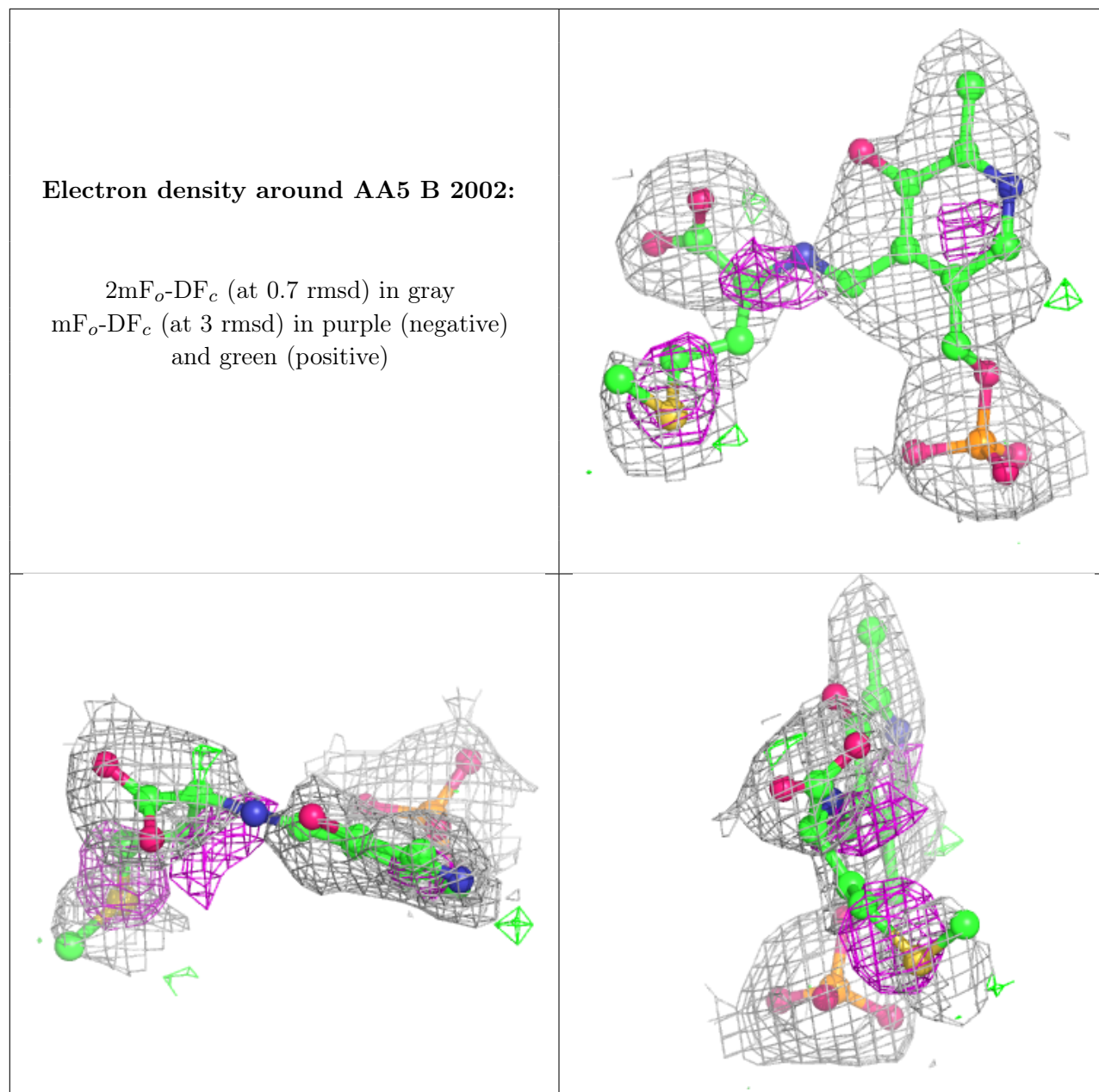
## 6.4 Ligands [i](#)

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(Å <sup>2</sup> )	Q<0.9
6	MET	C	2003	9/9	0.80	0.32	36,36,38,38	0
5	GOL	A	2009	6/6	0.86	0.17	36,36,37,37	0
3	AA5	B	2002	24/24	0.86	0.27	32,36,38,40	0
5	GOL	C	2010	6/6	0.87	0.16	46,46,46,46	0
3	AA5	D	2004	24/24	0.87	0.29	36,40,41,42	0
4	SO4	D	2008	5/5	0.89	0.33	65,65,65,65	0
3	AA5	A	2001	24/24	0.91	0.22	20,25,27,29	0
4	SO4	C	2007	5/5	0.91	0.28	59,59,60,60	0
4	SO4	B	2006	5/5	0.93	0.27	68,68,68,68	0
4	SO4	A	2005	5/5	0.95	0.21	63,63,64,64	0

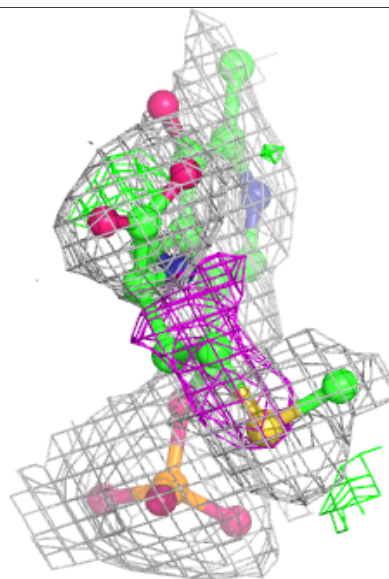
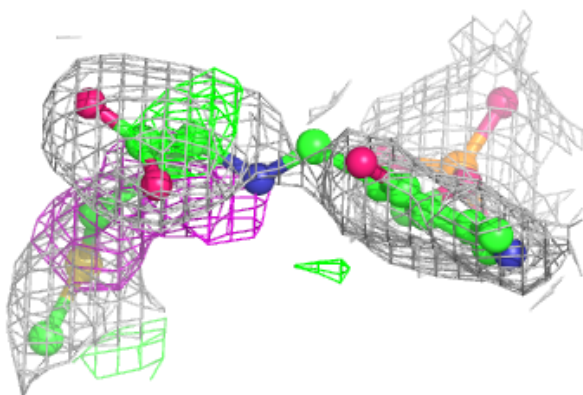
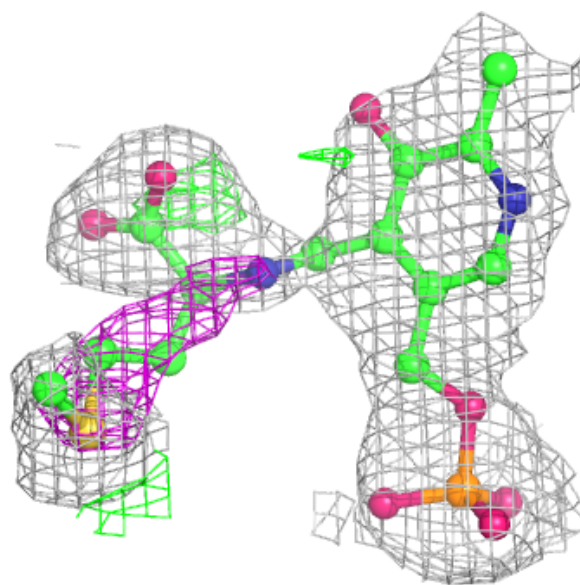
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different

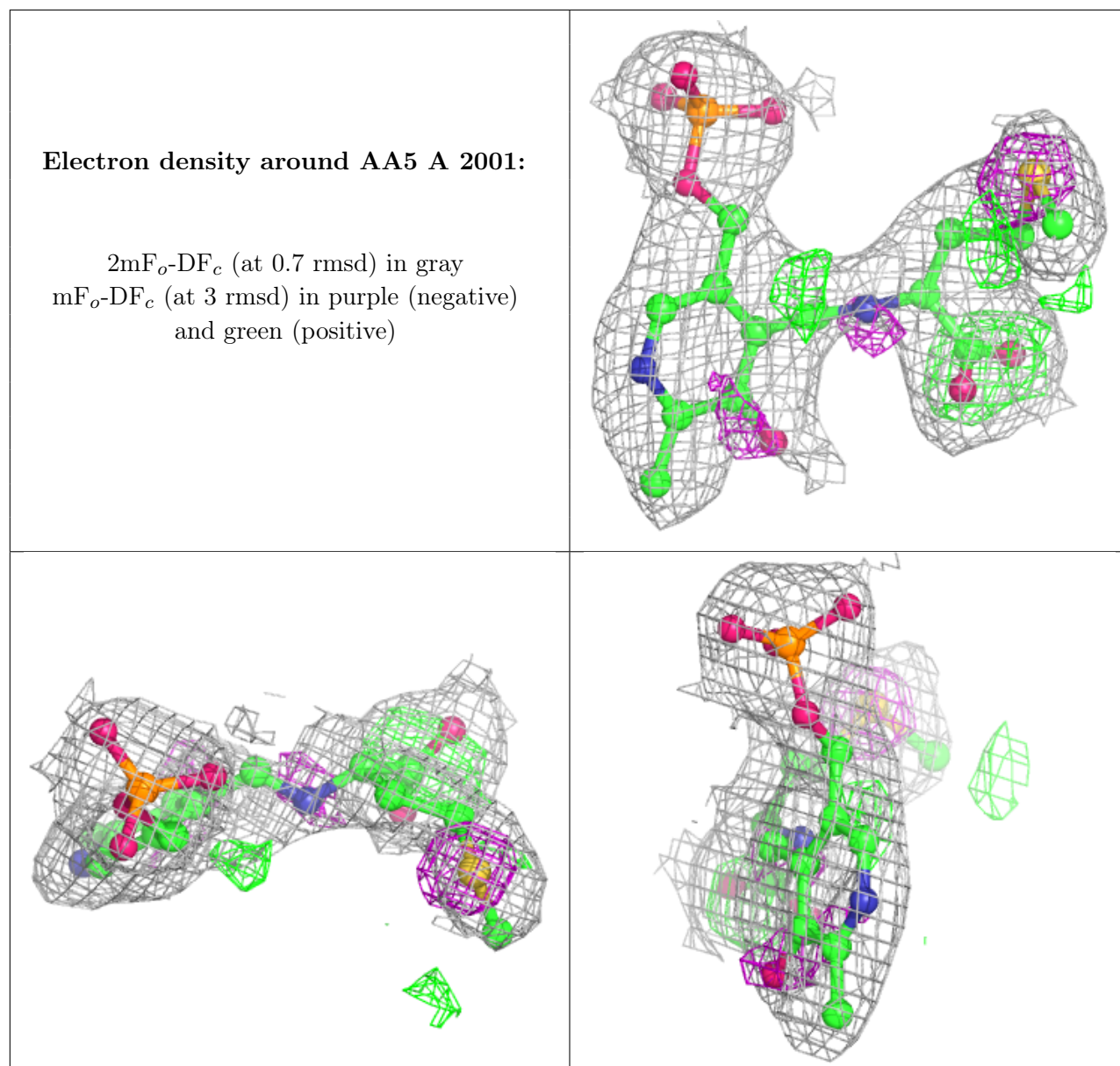
orientation to approximate a three-dimensional view.



**Electron density around AA5 D 2004:**

$2mF_o-DF_c$  (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.