



# Full wwPDB X-ray Structure Validation Report ⓘ

May 15, 2020 – 10:55 pm BST

PDB ID : 4AAY  
Title : Crystal Structure of the arsenite oxidase protein complex from Rhizobium species strain NT-26  
Authors : Oke, M.; Santini, J.M.; Naismith, J.H.  
Deposited on : 2011-12-05  
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 : 2.11  
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

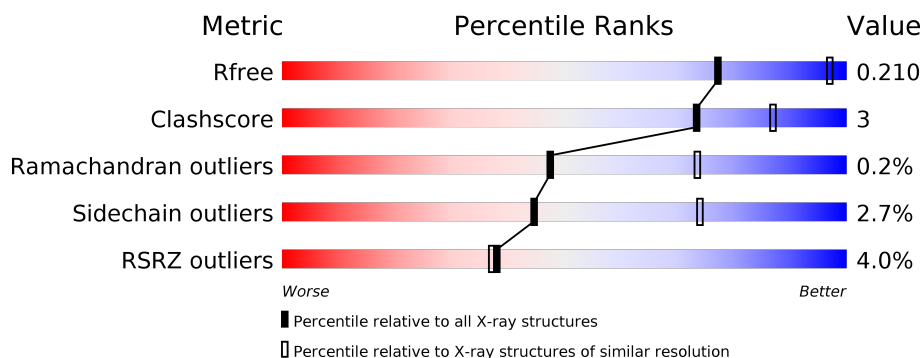
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 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	845	<div> <div>2%</div> <div> <div></div> <div>92%</div> <div>8%</div> </div> </div>
1	C	845	<div> <div>3%</div> <div> <div></div> <div>90%</div> <div>9%</div> </div> </div>
1	E	845	<div> <div>3%</div> <div> <div></div> <div>92%</div> <div>8%</div> </div> </div>
1	G	845	<div> <div>3%</div> <div> <div></div> <div>92%</div> <div>7%</div> </div> </div>
2	B	175	<div> <div>6%</div> <div> <div></div> <div>70%</div> <div>5%</div> <div>25%</div> </div> </div>
2	D	175	<div> <div>13%</div> <div> <div></div> <div>67%</div> <div>6%</div> <div>25%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	175	<div> <div>5%</div> <div>69%</div> <div>5%</div> <div>25%</div> </div>
2	H	175	<div> <div>14%</div> <div>69%</div> <div>5%</div> <div>25%</div> </div>

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 30609 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.

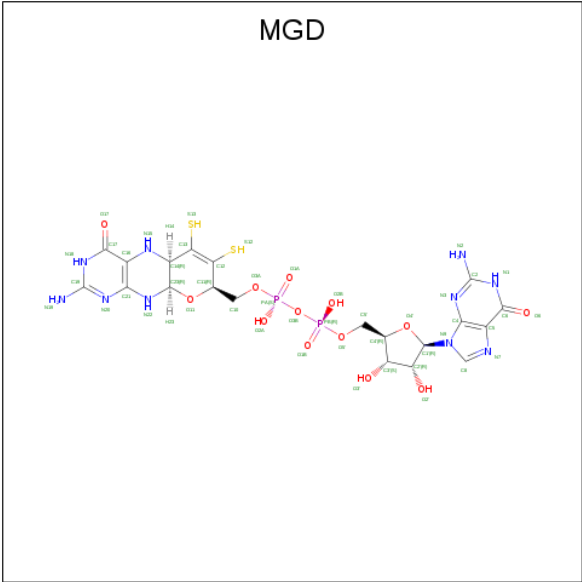
- Molecule 1 is a protein called AROA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	843	Total	C	N	O	S	0	0	0
			6545	4088	1167	1253	37			
1	C	843	Total	C	N	O	S	0	1	0
			6551	4092	1168	1254	37			
1	E	843	Total	C	N	O	S	0	0	0
			6545	4088	1167	1253	37			
1	G	843	Total	C	N	O	S	0	0	0
			6545	4088	1167	1253	37			

- Molecule 2 is a protein called AROB.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	132	Total	C	N	O	S	0	0	0
			995	629	166	198	2			
2	D	132	Total	C	N	O	S	0	0	0
			995	629	166	198	2			
2	F	132	Total	C	N	O	S	0	0	0
			995	629	166	198	2			
2	H	132	Total	C	N	O	S	0	0	0
			995	629	166	198	2			

- Molecule 3 is 2-AMINO-5,6-DIMERCAPTO-7-METHYL-3,7,8A,9-TETRAHYDRO-8-OXA-1,3,9,10-TETRAAZA-ANTHRACEN-4-ONE GUANOSINE DINUCLEOTIDE (three-letter code: MGD) (formula: C<sub>20</sub>H<sub>26</sub>N<sub>10</sub>O<sub>13</sub>P<sub>2</sub>S<sub>2</sub>).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
3	A	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
3	C	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
3	C	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
3	E	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
3	E	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
3	G	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
3	G	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		

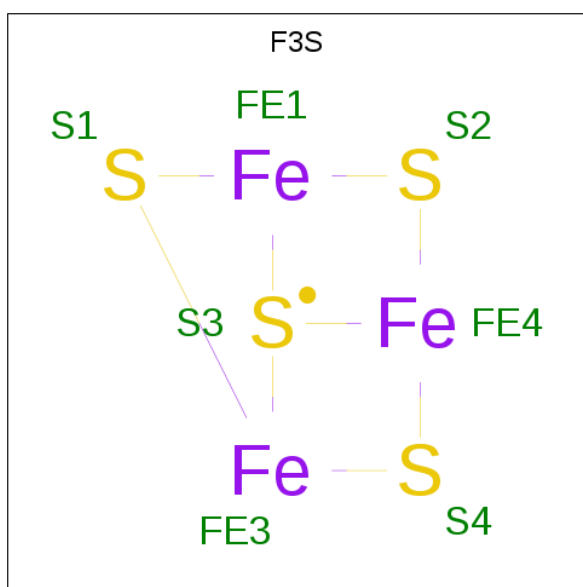
- Molecule 4 is OXYGEN ATOM (three-letter code: O) (formula: O).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	G	1	Total	O	0	0
			1	1		
4	A	1	Total	O	0	0
			1	1		
4	C	1	Total	O	0	0
			1	1		
4	E	1	Total	O	0	0
			1	1		

- Molecule 5 is MOLYBDENUM(IV) ION (three-letter code: 4MO) (formula: Mo).

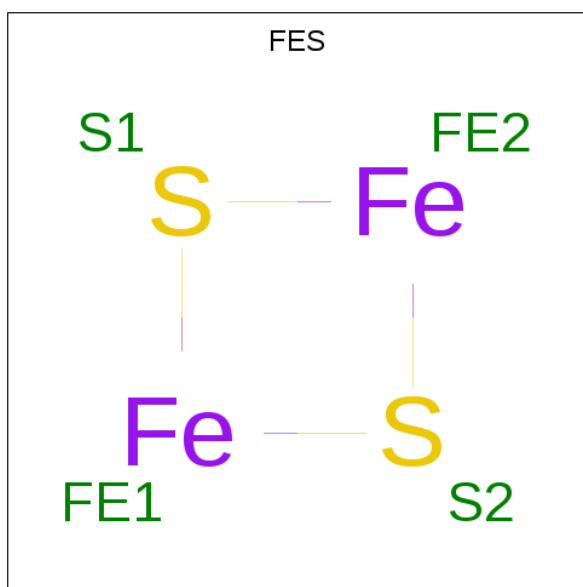
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	G	1	Total	Mo	0	0
			1	1		
5	A	1	Total	Mo	0	0
			1	1		
5	C	1	Total	Mo	0	0
			1	1		
5	E	1	Total	Mo	0	0
			1	1		

- Molecule 6 is FE3-S4 CLUSTER (three-letter code: F3S) (formula: Fe<sub>3</sub>S<sub>4</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	Fe	S	0	0
			7	3	4		
6	C	1	Total	Fe	S	0	0
			7	3	4		
6	E	1	Total	Fe	S	0	0
			7	3	4		
6	G	1	Total	Fe	S	0	0
			7	3	4		

- Molecule 7 is FE2/S2 (INORGANIC) CLUSTER (three-letter code: FES) (formula: Fe<sub>2</sub>S<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	B	1	Total	Fe	S	0	0
			4	2	2		
7	D	1	Total	Fe	S	0	0
			4	2	2		
7	F	1	Total	Fe	S	0	0
			4	2	2		
7	H	1	Total	Fe	S	0	0
			4	2	2		

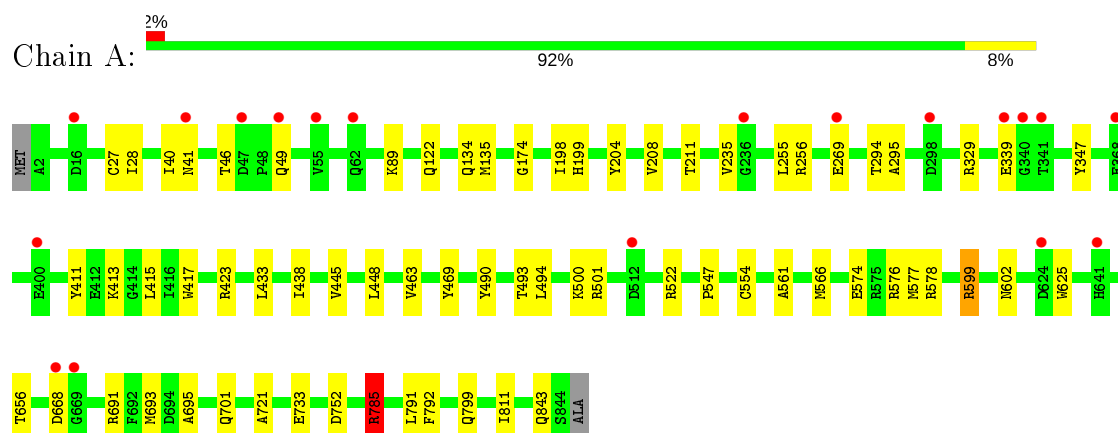
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	6	Total	O	0	0
			6	6		
8	C	2	Total	O	0	0
			2	2		
8	E	5	Total	O	0	0
			5	5		
8	G	2	Total	O	0	0
			2	2		

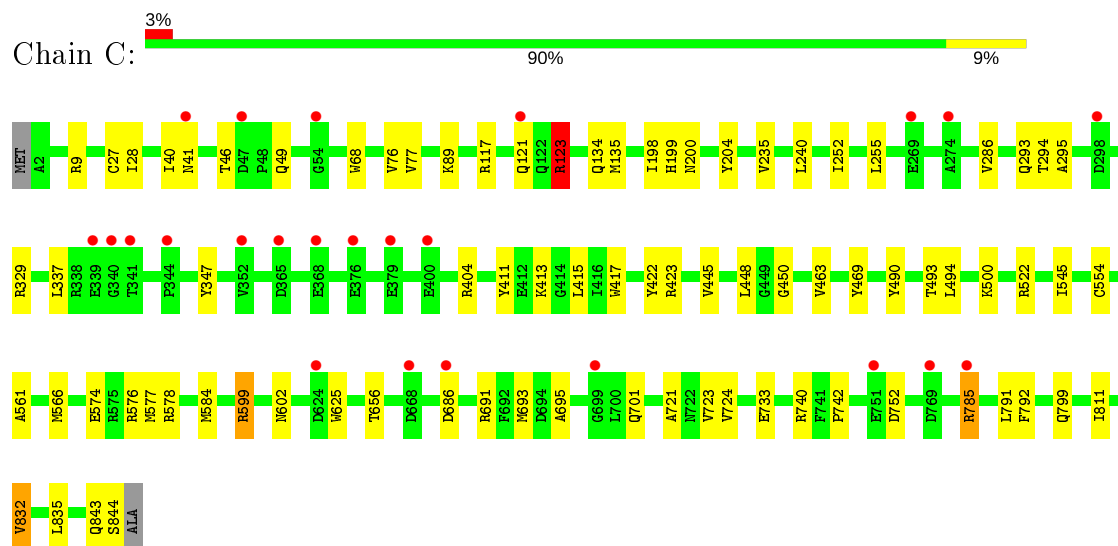
### 3 Residue-property plots [i](#)

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 ( $\text{RSRZ} > 2$ ). 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.

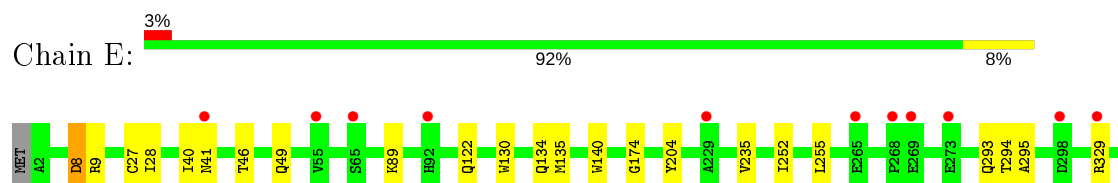
#### • Molecule 1: AROA



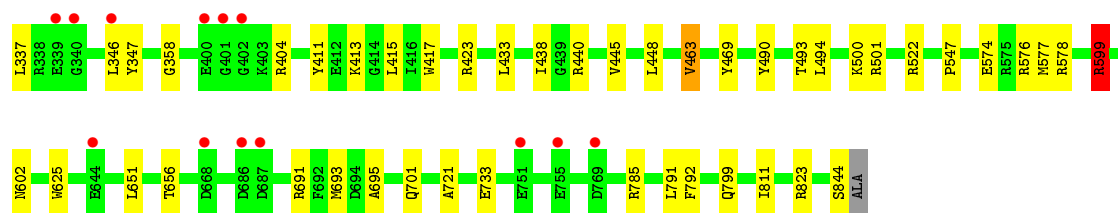
#### • Molecule 1: AROA



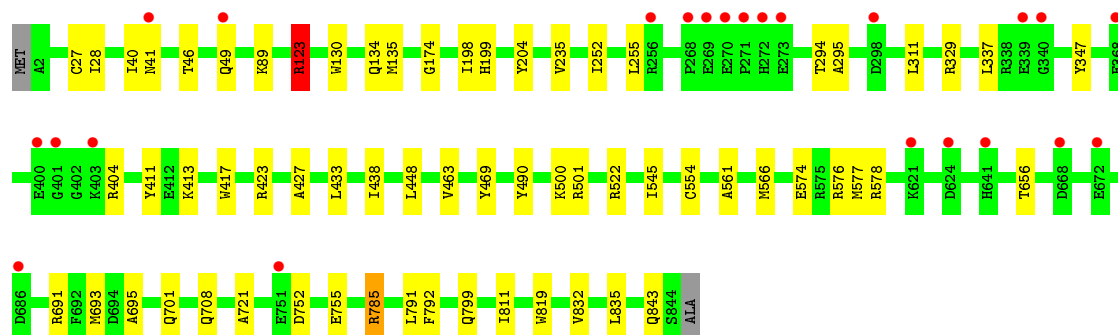
#### • Molecule 1: AROA



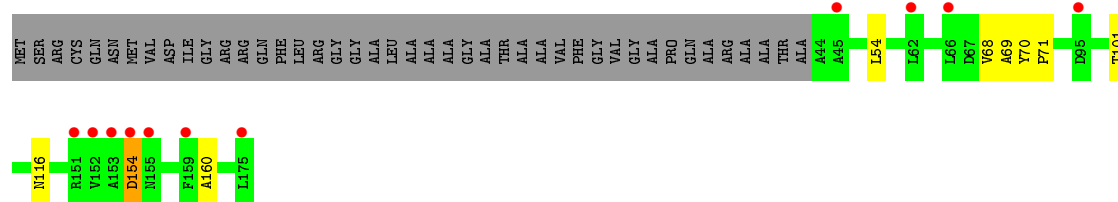




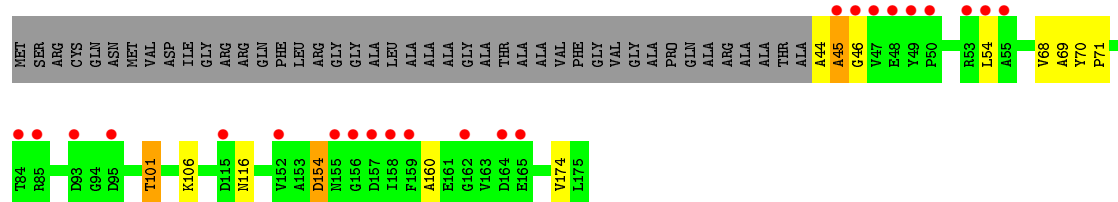
• Molecule 1: AROA



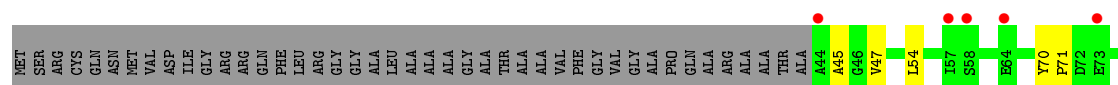
• Molecule 2: AROB

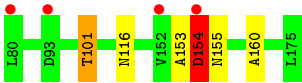


• Molecule 2: AROB

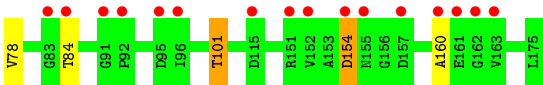
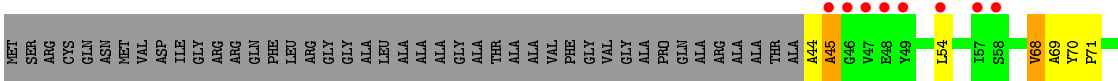


• Molecule 2: AROB





● Molecule 2: AROB



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	149.13Å 232.96Å 141.87Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	141.87 – 2.70 43.52 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.9 (141.87-2.70) 99.9 (43.52-2.70)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.56 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.6.0119	Depositor
R, $R_{free}$	0.193 , 0.211 0.193 , 0.210	Depositor DCC
$R_{free}$ test set	6818 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	31.4	Xtriage
Anisotropy	0.097	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 28.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.024 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	30609	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.14% 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: 4MO, F3S, FES, O, MGD

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	0.60	0/6693	0.67	4/9068 (0.0%)
1	C	0.59	1/6702 (0.0%)	0.68	4/9080 (0.0%)
1	E	0.58	1/6693 (0.0%)	0.69	6/9068 (0.1%)
1	G	0.58	2/6693 (0.0%)	0.68	4/9068 (0.0%)
2	B	0.47	0/1018	0.61	0/1387
2	D	0.47	0/1018	0.67	1/1387 (0.1%)
2	F	0.50	0/1018	0.69	1/1387 (0.1%)
2	H	0.49	0/1018	0.63	0/1387
All	All	0.58	4/30853 (0.0%)	0.67	20/41832 (0.0%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
2	D	0	2
2	F	0	1
All	All	0	3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	140	TRP	CD2-CE2	5.51	1.48	1.41
1	C	68	TRP	CD2-CE2	5.21	1.47	1.41
1	G	819	TRP	CD2-CE2	5.09	1.47	1.41
1	G	130	TRP	CD2-CE2	5.02	1.47	1.41

All (20) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	599	ARG	NE-CZ-NH1	11.08	125.84	120.30
1	E	599	ARG	NE-CZ-NH2	-9.94	115.33	120.30
1	C	123	ARG	NE-CZ-NH1	9.19	124.89	120.30
1	G	123	ARG	NE-CZ-NH2	-9.17	115.71	120.30
1	G	123	ARG	NE-CZ-NH1	9.07	124.83	120.30
1	C	123	ARG	NE-CZ-NH2	-8.42	116.09	120.30
2	F	154	ASP	CB-CG-OD2	7.32	124.88	118.30
1	E	8	ASP	CB-CG-OD2	-6.31	112.62	118.30
1	A	785	ARG	NE-CZ-NH1	6.16	123.38	120.30
1	E	599	ARG	CD-NE-CZ	6.08	132.11	123.60
2	D	46	GLY	N-CA-C	-5.86	98.45	113.10
1	A	785	ARG	NE-CZ-NH2	-5.77	117.41	120.30
1	E	785	ARG	NE-CZ-NH2	-5.65	117.47	120.30
1	C	134	GLN	CB-CA-C	-5.45	99.49	110.40
1	A	668	ASP	CB-CG-OD1	5.40	123.16	118.30
1	A	501	ARG	NE-CZ-NH2	-5.36	117.62	120.30
1	E	501	ARG	NE-CZ-NH2	-5.20	117.70	120.30
1	C	123	ARG	CB-CG-CD	5.18	125.06	111.60
1	G	123	ARG	CB-CG-CD	5.05	124.73	111.60
1	G	501	ARG	NE-CZ-NH2	-5.03	117.78	120.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	44	ALA	Peptide
2	D	45	ALA	Peptide
2	F	153	ALA	Peptide

## 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	6545	0	6296	43	0
1	C	6551	0	6304	46	1
1	E	6545	0	6296	51	0
1	G	6545	0	6296	36	0
2	B	995	0	947	5	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	995	0	947	15	0
2	F	995	0	947	5	0
2	H	995	0	947	8	0
3	A	94	0	44	1	0
3	C	94	0	44	1	0
3	E	94	0	44	1	0
3	G	94	0	44	1	0
4	A	1	0	0	0	0
4	C	1	0	0	1	0
4	E	1	0	0	0	0
4	G	1	0	0	0	0
5	A	1	0	0	0	0
5	C	1	0	0	0	0
5	E	1	0	0	0	0
5	G	1	0	0	0	0
6	A	7	0	0	0	0
6	C	7	0	0	0	0
6	E	7	0	0	0	0
6	G	7	0	0	0	0
7	B	4	0	0	0	0
7	D	4	0	0	1	0
7	F	4	0	0	0	0
7	H	4	0	0	0	0
8	A	6	0	0	0	0
8	C	2	0	0	0	0
8	E	5	0	0	0	0
8	G	2	0	0	0	0
All	All	30609	0	29156	195	1

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

All (195) 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:E:130:TRP:HB2	1:E:135:MET:CE	1.85	1.04
1:E:130:TRP:HB2	1:E:135:MET:HE1	1.02	1.00
1:E:130:TRP:CB	1:E:135:MET:HE1	1.93	0.99
1:A:423:ARG:HA	1:A:693:MET:HE2	1.48	0.95
1:A:423:ARG:HA	1:A:693:MET:CE	1.97	0.93
1:G:423:ARG:HA	1:G:693:MET:CE	2.00	0.91

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:G:752:ASP:OD1	1:G:785:ARG:NH1	2.09	0.85
1:C:752:ASP:OD1	1:C:785:ARG:NH1	2.10	0.84
1:G:423:ARG:HA	1:G:693:MET:HE3	1.60	0.83
1:C:9:ARG:NH2	2:D:45:ALA:HB1	1.94	0.82
1:C:423:ARG:HA	1:C:693:MET:HE2	1.61	0.82
1:A:256:ARG:HG3	1:A:294:THR:HG23	1.62	0.81
2:D:45:ALA:HB3	1:E:9:ARG:CZ	2.11	0.80
1:C:599:ARG:NH1	1:C:602:ASN:OD1	2.16	0.79
1:A:339:GLU:HG2	1:G:755:GLU:HA	1.65	0.79
1:A:599:ARG:NH1	1:A:602:ASN:OD1	2.16	0.78
1:E:130:TRP:HE3	1:E:135:MET:CE	1.98	0.77
1:A:843:GLN:OE1	2:B:116:ASN:ND2	2.16	0.76
1:C:423:ARG:HA	1:C:693:MET:CE	2.17	0.75
1:E:423:ARG:HA	1:E:693:MET:HE2	1.70	0.74
1:G:423:ARG:HA	1:G:693:MET:HE1	1.71	0.73
1:E:599:ARG:NH1	1:E:602:ASN:OD1	2.22	0.73
1:E:130:TRP:HE3	1:E:135:MET:HE2	1.55	0.72
2:D:45:ALA:HB3	1:E:9:ARG:NE	2.05	0.71
1:E:423:ARG:HA	1:E:693:MET:CE	2.21	0.70
1:E:415:LEU:C	1:E:415:LEU:HD23	2.13	0.69
1:C:204:TYR:HB3	1:C:463:VAL:HG11	1.76	0.68
2:D:45:ALA:CB	1:E:9:ARG:CZ	2.72	0.68
1:A:135:MET:CE	1:A:554:CYS:O	2.42	0.67
1:C:252:ILE:HD12	1:C:294:THR:HG21	1.78	0.66
1:E:130:TRP:CE3	1:E:135:MET:CE	2.79	0.66
1:E:252:ILE:HD12	1:E:294:THR:HG21	1.78	0.66
1:E:599:ARG:HH11	1:E:599:ARG:HG3	1.61	0.65
1:E:130:TRP:CE3	1:E:135:MET:HE2	2.32	0.64
1:C:9:ARG:HH21	2:D:45:ALA:HB1	1.61	0.64
1:A:423:ARG:HA	1:A:693:MET:HE1	1.81	0.62
1:A:256:ARG:HG3	1:A:294:THR:CG2	2.29	0.62
2:H:44:ALA:O	2:H:45:ALA:HB3	2.00	0.62
1:C:561:ALA:O	1:C:566:MET:CE	2.48	0.61
1:G:40:ILE:O	1:G:41:ASN:OD1	2.19	0.60
1:E:130:TRP:CB	1:E:135:MET:CE	2.63	0.60
1:A:204:TYR:HB3	1:A:463:VAL:HG11	1.84	0.59
1:A:40:ILE:O	1:A:41:ASN:OD1	2.20	0.59
1:A:135:MET:HE3	1:A:554:CYS:O	2.03	0.58
1:G:252:ILE:HD12	1:G:294:THR:HG21	1.85	0.58
1:C:77:VAL:HB	1:C:584:MET:HE3	1.84	0.58
1:C:832:VAL:HG22	1:C:835:LEU:HD12	1.86	0.57

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:433:LEU:HD23	1:E:438:ILE:HD12	1.87	0.56
2:B:54:LEU:HD11	2:B:160:ALA:HB2	1.87	0.56
1:A:561:ALA:O	1:A:566:MET:CE	2.53	0.56
1:A:198:ILE:HG22	1:A:199:HIS:N	2.21	0.56
2:D:68:VAL:HG13	2:D:69:ALA:N	2.21	0.55
1:A:413:LYS:HD2	1:A:417:TRP:NE1	2.22	0.54
1:C:413:LYS:HD2	1:C:417:TRP:NE1	2.23	0.54
1:E:130:TRP:HE3	1:E:135:MET:HE3	1.72	0.54
1:E:413:LYS:HD2	1:E:417:TRP:NE1	2.23	0.54
1:A:135:MET:HE1	1:A:554:CYS:O	2.07	0.54
2:D:174:VAL:O	2:D:174:VAL:HG13	2.07	0.54
1:E:599:ARG:NH2	1:E:625:TRP:O	2.40	0.54
1:A:599:ARG:NH2	1:A:625:TRP:O	2.41	0.54
1:A:433:LEU:HD23	1:A:438:ILE:HD12	1.90	0.54
1:G:433:LEU:HD23	1:G:438:ILE:HD12	1.89	0.53
1:G:413:LYS:HD2	1:G:417:TRP:NE1	2.23	0.52
1:A:339:GLU:HG2	1:G:755:GLU:CA	2.37	0.52
1:A:561:ALA:O	1:A:566:MET:HE3	2.10	0.52
1:C:599:ARG:NH2	1:C:625:TRP:O	2.42	0.52
1:A:198:ILE:HG22	1:A:199:HIS:H	1.75	0.52
1:G:561:ALA:O	1:G:566:MET:CE	2.56	0.52
1:E:40:ILE:O	1:E:41:ASN:HB2	2.10	0.52
1:G:123:ARG:HG2	1:G:545:ILE:O	2.10	0.51
1:C:28:ILE:HG21	1:C:448:LEU:O	2.11	0.51
1:A:255:LEU:HB3	1:A:295:ALA:HB2	1.91	0.51
1:G:123:ARG:HH11	1:G:123:ARG:HG2	1.76	0.51
1:C:123:ARG:HG2	1:C:545:ILE:O	2.10	0.51
1:A:28:ILE:HG21	1:A:448:LEU:O	2.10	0.50
1:C:561:ALA:O	1:C:566:MET:HE3	2.10	0.50
2:D:68:VAL:CG1	2:D:69:ALA:N	2.74	0.50
1:C:445:VAL:HG13	1:C:445:VAL:O	2.12	0.50
1:G:578:ARG:HD2	1:G:656:THR:O	2.12	0.50
1:C:415:LEU:HD23	1:C:415:LEU:C	2.32	0.50
1:G:721:ALA:HB2	1:G:791:LEU:HD13	1.93	0.50
2:H:44:ALA:O	2:H:45:ALA:CB	2.59	0.50
1:C:337:LEU:HD13	1:C:691:ARG:HH21	1.75	0.50
2:D:54:LEU:HD11	2:D:160:ALA:HB2	1.93	0.50
1:G:574:GLU:OE1	1:G:576:ARG:NH1	2.45	0.50
1:C:574:GLU:OE1	1:C:576:ARG:NH1	2.45	0.50
1:A:574:GLU:OE1	1:A:576:ARG:NH1	2.45	0.50
1:E:578:ARG:HD2	1:E:656:THR:O	2.12	0.49

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:70:TYR:CD1	2:B:71:PRO:HA	2.48	0.49
1:E:28:ILE:HG21	1:E:448:LEU:O	2.11	0.49
1:A:339:GLU:HG2	1:G:755:GLU:O	2.12	0.49
2:B:68:VAL:CG1	2:B:69:ALA:N	2.76	0.49
1:E:574:GLU:OE1	1:E:576:ARG:NH1	2.45	0.49
1:C:578:ARG:HD2	1:C:656:THR:O	2.13	0.49
1:G:578:ARG:CD	1:G:656:THR:O	2.61	0.49
2:H:70:TYR:CD1	2:H:71:PRO:HA	2.48	0.49
1:E:578:ARG:CD	1:E:656:THR:O	2.60	0.49
1:E:255:LEU:HB3	1:E:295:ALA:HB2	1.95	0.49
1:C:135:MET:CE	1:C:554:CYS:O	2.60	0.48
1:A:578:ARG:HD2	1:A:656:THR:O	2.13	0.48
1:E:337:LEU:HD13	1:E:691:ARG:HH21	1.77	0.48
2:H:54:LEU:HD11	2:H:160:ALA:HB2	1.94	0.48
1:G:832:VAL:HG22	1:G:835:LEU:HD22	1.96	0.48
1:C:255:LEU:HB3	1:C:295:ALA:HB2	1.96	0.48
1:E:721:ALA:HB2	1:E:791:LEU:HD13	1.94	0.48
1:C:578:ARG:CD	1:C:656:THR:O	2.62	0.47
1:C:742:PRO:HD2	1:C:835:LEU:HD11	1.96	0.47
1:E:293:GLN:NE2	1:E:844:SER:HB3	2.29	0.47
1:A:578:ARG:CD	1:A:656:THR:O	2.62	0.47
1:G:255:LEU:HB3	1:G:295:ALA:HB2	1.95	0.47
2:D:70:TYR:CD1	2:D:71:PRO:HA	2.48	0.47
1:C:123:ARG:HG2	1:C:123:ARG:HH11	1.77	0.47
2:D:101:THR:CG2	2:D:101:THR:O	2.63	0.47
1:E:433:LEU:CD2	1:E:438:ILE:HD12	2.44	0.47
1:A:493:THR:HG23	1:A:494:LEU:O	2.14	0.47
1:G:198:ILE:HG22	1:G:199:HIS:N	2.30	0.47
2:H:68:VAL:CG1	2:H:69:ALA:N	2.77	0.47
2:F:70:TYR:CD1	2:F:71:PRO:HA	2.49	0.47
1:C:721:ALA:HB2	1:C:791:LEU:HD13	1.95	0.47
1:C:40:ILE:O	1:C:41:ASN:HB2	2.14	0.46
1:C:693:MET:HG2	1:C:695:ALA:HB2	1.97	0.46
1:C:135:MET:HE2	1:C:554:CYS:O	2.16	0.46
1:G:337:LEU:HD13	1:G:691:ARG:HH21	1.81	0.46
1:E:423:ARG:HA	1:E:693:MET:HE1	1.97	0.46
1:A:577:MET:O	1:A:578:ARG:HG3	2.16	0.46
1:C:493:THR:HG23	1:C:494:LEU:O	2.16	0.46
2:D:154:ASP:N	2:D:154:ASP:OD1	2.49	0.46
1:E:204:TYR:HB3	1:E:463:VAL:HG11	1.97	0.46
2:F:54:LEU:HD11	2:F:160:ALA:HB2	1.97	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:G:135:MET:CE	1:G:554:CYS:O	2.64	0.45
2:H:154:ASP:N	2:H:154:ASP:OD1	2.49	0.45
1:G:28:ILE:HG21	1:G:448:LEU:O	2.16	0.45
1:E:490:TYR:O	1:E:500:LYS:HE2	2.16	0.45
1:C:740:ARG:O	1:C:835:LEU:HD22	2.16	0.45
1:C:490:TYR:O	1:C:500:LYS:HE2	2.17	0.45
1:C:76:VAL:HG23	1:C:117:ARG:HD2	1.99	0.45
1:G:198:ILE:HG22	1:G:199:HIS:H	1.82	0.45
1:C:293:GLN:NE2	1:C:844:SER:HB2	2.32	0.44
2:D:101:THR:HG23	2:D:101:THR:O	2.18	0.44
2:H:101:THR:O	2:H:101:THR:CG2	2.66	0.44
1:C:240:LEU:HD23	1:C:286:VAL:HG23	1.99	0.44
1:G:490:TYR:O	1:G:500:LYS:HE2	2.17	0.44
1:A:490:TYR:O	1:A:500:LYS:HE2	2.18	0.44
1:G:135:MET:HE2	1:G:554:CYS:O	2.17	0.44
1:A:445:VAL:O	1:A:445:VAL:HG13	2.18	0.43
1:E:577:MET:O	1:E:578:ARG:HG3	2.18	0.43
1:G:693:MET:HG2	1:G:695:ALA:HB2	2.01	0.43
1:C:577:MET:O	1:C:578:ARG:HG3	2.18	0.43
1:A:693:MET:HG2	1:A:695:ALA:HB2	2.00	0.43
1:C:235:VAL:O	1:C:411:TYR:HA	2.19	0.43
2:D:45:ALA:CB	1:E:9:ARG:NH2	2.81	0.43
1:A:413:LYS:HD2	1:A:417:TRP:CE2	2.54	0.43
2:H:54:LEU:HD13	2:H:68:VAL:HG11	2.00	0.43
1:E:358:GLY:HA3	1:E:693:MET:HE3	2.01	0.43
1:E:493:THR:HG23	1:E:494:LEU:O	2.19	0.43
1:C:740:ARG:O	1:C:835:LEU:CD2	2.66	0.43
1:C:9:ARG:NE	2:F:45:ALA:HA	2.34	0.43
1:C:200:ASN:ND2	4:C:2003:O:O	2.50	0.42
1:G:577:MET:O	1:G:578:ARG:HG3	2.18	0.42
1:G:204:TYR:HB3	1:G:463:VAL:HG11	2.00	0.42
1:A:235:VAL:O	1:A:411:TYR:HA	2.20	0.42
2:B:154:ASP:OD1	2:B:154:ASP:N	2.52	0.42
1:E:130:TRP:CA	1:E:135:MET:CE	2.98	0.42
1:E:235:VAL:O	1:E:411:TYR:HA	2.19	0.42
1:C:450:GLY:HA2	3:C:2002:MGD:H11	2.01	0.42
1:E:130:TRP:CE3	1:E:135:MET:HE3	2.50	0.42
1:G:433:LEU:CD2	1:G:438:ILE:HD12	2.49	0.42
1:A:174:GLY:HA2	3:A:2001:MGD:C6	2.49	0.42
1:C:422:TYR:CD2	1:C:693:MET:HE1	2.55	0.42
1:E:693:MET:HG2	1:E:695:ALA:HB2	2.01	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:445:VAL:HG13	1:E:445:VAL:O	2.20	0.42
1:E:8:ASP:HB3	2:F:47:VAL:HG21	2.01	0.42
1:G:311:LEU:CD1	1:G:427:ALA:HB2	2.49	0.42
1:A:122:GLN:O	1:A:547:PRO:HD3	2.20	0.42
1:C:413:LYS:HD2	1:C:417:TRP:CE2	2.55	0.42
1:G:174:GLY:HA2	3:G:2001:MGD:C6	2.50	0.42
1:E:174:GLY:HA2	3:E:2001:MGD:C6	2.50	0.41
1:A:208:VAL:HB	1:A:211:THR:HG22	2.02	0.41
1:A:339:GLU:CG	1:G:755:GLU:O	2.69	0.41
1:A:415:LEU:C	1:A:415:LEU:HD23	2.40	0.41
1:E:130:TRP:HA	1:E:135:MET:HE2	2.02	0.41
1:G:235:VAL:O	1:G:411:TYR:HA	2.20	0.41
1:C:198:ILE:HG22	1:C:199:HIS:N	2.36	0.41
1:G:413:LYS:HD2	1:G:417:TRP:CE2	2.56	0.41
1:A:752:ASP:OD1	1:A:785:ARG:HG3	2.20	0.41
2:F:101:THR:O	2:F:101:THR:CG2	2.69	0.41
1:A:198:ILE:CG2	1:A:199:HIS:N	2.84	0.41
1:A:269:GLU:OE1	1:E:823:ARG:NH1	2.54	0.41
1:A:721:ALA:HB2	1:A:791:LEU:HD13	2.03	0.41
2:D:106:LYS:HB2	7:D:2006:FES:S2	2.61	0.41
1:E:122:GLN:O	1:E:547:PRO:HD3	2.21	0.40
1:C:723:VAL:HG23	1:C:724:VAL:HG23	2.03	0.40
1:E:423:ARG:HG2	1:E:693:MET:CE	2.51	0.40
1:E:651:LEU:HD23	1:E:651:LEU:HA	1.94	0.40
1:E:490:TYR:O	1:E:500:LYS:CE	2.70	0.40

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

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:686:ASP:O	1:C:686:ASP:O[2_545]	1.73	0.47

## 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	841/845 (100%)	810 (96%)	30 (4%)	1 (0%)	51	78
1	C	842/845 (100%)	812 (96%)	29 (3%)	1 (0%)	51	78
1	E	841/845 (100%)	815 (97%)	25 (3%)	1 (0%)	51	78
1	G	841/845 (100%)	809 (96%)	31 (4%)	1 (0%)	51	78
2	B	130/175 (74%)	125 (96%)	5 (4%)	0	100	100
2	D	130/175 (74%)	126 (97%)	4 (3%)	0	100	100
2	F	130/175 (74%)	121 (93%)	7 (5%)	2 (2%)	10	26
2	H	130/175 (74%)	125 (96%)	4 (3%)	1 (1%)	19	43
All	All	3885/4080 (95%)	3743 (96%)	135 (4%)	7 (0%)	47	73

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	F	154	ASP
2	H	45	ALA
2	F	155	ASN
1	E	811	ILE
1	A	811	ILE
1	G	811	ILE
1	C	811	ILE

### 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	677/678 (100%)	661 (98%)	16 (2%)	49	77
1	C	678/678 (100%)	658 (97%)	20 (3%)	42	71
1	E	677/678 (100%)	659 (97%)	18 (3%)	44	74
1	G	677/678 (100%)	660 (98%)	17 (2%)	47	76
2	B	105/130 (81%)	103 (98%)	2 (2%)	57	82

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	D	105/130 (81%)	102 (97%)	3 (3%)	42	71
2	F	105/130 (81%)	102 (97%)	3 (3%)	42	71
2	H	105/130 (81%)	100 (95%)	5 (5%)	25	53
All	All	3129/3232 (97%)	3045 (97%)	84 (3%)	44	74

All (84) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	27	CYS
1	A	46	THR
1	A	49	GLN
1	A	89	LYS
1	A	134	GLN
1	A	329	ARG
1	A	347	TYR
1	A	469	TYR
1	A	522	ARG
1	A	599	ARG
1	A	691	ARG
1	A	701	GLN
1	A	733	GLU
1	A	785	ARG
1	A	792	PHE
1	A	799	GLN
2	B	101	THR
2	B	154	ASP
1	C	27	CYS
1	C	46	THR
1	C	49	GLN
1	C	89	LYS
1	C	121[A]	GLN
1	C	121[B]	GLN
1	C	123	ARG
1	C	329	ARG
1	C	347	TYR
1	C	404	ARG
1	C	469	TYR
1	C	522	ARG
1	C	599	ARG
1	C	701	GLN
1	C	733	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	785	ARG
1	C	792	PHE
1	C	799	GLN
1	C	832	VAL
1	C	843	GLN
2	D	101	THR
2	D	116	ASN
2	D	154	ASP
1	E	27	CYS
1	E	46	THR
1	E	49	GLN
1	E	89	LYS
1	E	134	GLN
1	E	329	ARG
1	E	346	LEU
1	E	347	TYR
1	E	404	ARG
1	E	440	ARG
1	E	463	VAL
1	E	469	TYR
1	E	522	ARG
1	E	599	ARG
1	E	701	GLN
1	E	733	GLU
1	E	792	PHE
1	E	799	GLN
2	F	101	THR
2	F	116	ASN
2	F	154	ASP
1	G	27	CYS
1	G	46	THR
1	G	49	GLN
1	G	89	LYS
1	G	123	ARG
1	G	134	GLN
1	G	329	ARG
1	G	347	TYR
1	G	404	ARG
1	G	469	TYR
1	G	522	ARG
1	G	701	GLN
1	G	708	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	G	785	ARG
1	G	792	PHE
1	G	799	GLN
1	G	843	GLN
2	H	68	VAL
2	H	78	VAL
2	H	84	THR
2	H	101	THR
2	H	154	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	293	GLN
1	C	542	ASN
1	E	293	GLN
1	G	425	ASN

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 24 ligands modelled in this entry, 8 are monoatomic - leaving 16 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$
3	MGD	G	2001	5	41,52,52	1.31	4 (9%)	43,81,81	2.11	16 (37%)
6	F3S	A	2005	1	0,9,9	0.00	-	-		
6	F3S	G	2005	1	0,9,9	0.00	-	-		
6	F3S	C	2005	1	0,9,9	0.00	-	-		
7	FES	H	2006	2	0,4,4	0.00	-	-		
3	MGD	E	2001	5	41,52,52	1.25	4 (9%)	43,81,81	2.02	16 (37%)
3	MGD	C	2001	5	41,52,52	1.27	4 (9%)	43,81,81	2.28	19 (44%)
3	MGD	A	2001	5	41,52,52	1.21	4 (9%)	43,81,81	2.14	13 (30%)
3	MGD	E	2002	5	41,52,52	1.28	7 (17%)	43,81,81	2.37	15 (34%)
7	FES	D	2006	2	0,4,4	0.00	-	-		
3	MGD	G	2002	5	41,52,52	1.32	5 (12%)	43,81,81	2.26	12 (27%)
7	FES	B	2006	2	0,4,4	0.00	-	-		
3	MGD	A	2002	5	41,52,52	1.27	5 (12%)	43,81,81	2.21	15 (34%)
3	MGD	C	2002	5	41,52,52	1.36	5 (12%)	43,81,81	2.12	13 (30%)
7	FES	F	2006	2	0,4,4	0.00	-	-		
6	F3S	E	2005	1	0,9,9	0.00	-	-		

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	F3S	A	2005	1	-	-	0/3/3/3
6	F3S	G	2005	1	-	-	0/3/3/3
6	F3S	C	2005	1	-	-	0/3/3/3
7	FES	F	2006	2	-	-	0/1/1/1
7	FES	H	2006	2	-	-	0/1/1/1
3	MGD	E	2001	5	-	5/18/66/66	0/6/6/6
3	MGD	C	2001	5	-	4/18/66/66	0/6/6/6
3	MGD	A	2001	5	-	4/18/66/66	0/6/6/6
3	MGD	E	2002	5	-	2/18/66/66	0/6/6/6
7	FES	D	2006	2	-	-	0/1/1/1
3	MGD	G	2002	5	-	3/18/66/66	0/6/6/6
7	FES	B	2006	2	-	-	0/1/1/1
3	MGD	A	2002	5	-	3/18/66/66	0/6/6/6

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	MGD	C	2002	5	-	3/18/66/66	0/6/6/6
3	MGD	G	2001	5	-	5/18/66/66	0/6/6/6
6	F3S	E	2005	1	-	-	0/3/3/3

All (38) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	2001	MGD	C17-C16	5.02	1.48	1.41
3	C	2002	MGD	C17-C16	4.88	1.48	1.41
3	G	2002	MGD	C17-C16	4.83	1.48	1.41
3	E	2002	MGD	C6-C5	4.37	1.48	1.41
3	G	2001	MGD	C17-C16	4.13	1.47	1.41
3	C	2001	MGD	C17-C16	4.12	1.47	1.41
3	A	2002	MGD	C17-C16	4.10	1.47	1.41
3	G	2002	MGD	C16-C21	3.83	1.48	1.41
3	A	2001	MGD	C17-C16	3.68	1.46	1.41
3	A	2002	MGD	C6-C5	3.67	1.47	1.41
3	C	2002	MGD	C6-C5	3.53	1.47	1.41
3	A	2001	MGD	C6-C5	3.40	1.47	1.41
3	G	2001	MGD	C6-C5	3.35	1.47	1.41
3	C	2002	MGD	C16-C21	3.20	1.47	1.41
3	C	2001	MGD	C6-C5	3.13	1.46	1.41
3	G	2002	MGD	C6-C5	3.12	1.46	1.41
3	E	2001	MGD	C16-C21	3.11	1.47	1.41
3	G	2001	MGD	C16-C21	3.07	1.47	1.41
3	C	2001	MGD	C16-C21	2.96	1.47	1.41
3	E	2002	MGD	C17-C16	2.93	1.45	1.41
3	A	2002	MGD	C16-C21	2.90	1.46	1.41
3	C	2001	MGD	C16-N15	-2.87	1.32	1.38
3	G	2001	MGD	C16-N15	-2.71	1.32	1.38
3	C	2002	MGD	C5-C4	2.69	1.48	1.40
3	E	2002	MGD	O11-C11	-2.41	1.40	1.43
3	E	2001	MGD	C6-C5	2.34	1.45	1.41
3	E	2001	MGD	C16-N15	-2.31	1.33	1.38
3	A	2001	MGD	C16-N15	-2.26	1.33	1.38
3	A	2002	MGD	O11-C11	-2.25	1.40	1.43
3	E	2002	MGD	C5-C4	2.23	1.46	1.40
3	G	2002	MGD	C5-C4	2.23	1.46	1.40
3	A	2002	MGD	C16-N15	-2.17	1.33	1.38
3	A	2001	MGD	C16-C21	2.08	1.45	1.41
3	E	2002	MGD	O4'-C1'	2.06	1.44	1.41
3	C	2002	MGD	O11-C11	-2.05	1.41	1.43

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	G	2002	MGD	C2'-C1'	-2.05	1.50	1.53
3	E	2002	MGD	C16-C21	2.05	1.45	1.41
3	E	2002	MGD	C16-N15	-2.05	1.33	1.38

All (119) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	2002	MGD	O11-C23-N22	-7.45	100.91	108.57
3	G	2002	MGD	O11-C23-N22	-6.08	102.32	108.57
3	A	2002	MGD	O11-C23-C14	5.86	112.87	108.96
3	C	2002	MGD	O11-C23-N22	-5.45	102.97	108.57
3	C	2002	MGD	C5-C6-N1	-5.30	116.18	123.43
3	E	2002	MGD	C5-C6-N1	-4.96	116.65	123.43
3	A	2002	MGD	C5-C6-N1	-4.95	116.66	123.43
3	C	2002	MGD	C6-N1-C2	4.93	123.76	115.93
3	A	2001	MGD	C6-C5-C4	-4.85	116.17	120.80
3	G	2001	MGD	C2-N3-C4	4.66	120.68	115.36
3	C	2001	MGD	C5-C6-N1	-4.62	117.12	123.43
3	C	2002	MGD	C17-N18-C19	4.55	123.16	115.93
3	E	2002	MGD	O11-C23-C14	4.52	111.98	108.96
3	C	2001	MGD	C17-N18-C19	4.51	123.09	115.93
3	G	2001	MGD	C6-C5-C4	-4.44	116.56	120.80
3	G	2002	MGD	C2-N3-C4	4.42	120.40	115.36
3	A	2002	MGD	C17-N18-C19	4.42	122.95	115.93
3	C	2001	MGD	C6-N1-C2	4.41	122.94	115.93
3	A	2001	MGD	C5-C6-N1	-4.33	117.51	123.43
3	G	2002	MGD	C17-C16-N15	4.28	122.71	119.12
3	A	2002	MGD	C2-N3-C4	4.19	120.15	115.36
3	A	2001	MGD	C17-C16-N15	4.18	122.63	119.12
3	C	2001	MGD	C17-C16-N15	4.13	122.58	119.12
3	A	2001	MGD	C1'-N9-C4	-4.08	119.48	126.64
3	G	2002	MGD	C6-C5-C4	-4.08	116.91	120.80
3	G	2002	MGD	C6-N1-C2	4.08	122.40	115.93
3	E	2002	MGD	C2-N3-C4	4.04	119.98	115.36
3	A	2001	MGD	C6-N1-C2	4.02	122.32	115.93
3	C	2001	MGD	C2-N3-C4	3.92	119.83	115.36
3	E	2002	MGD	C17-N18-C19	3.91	122.14	115.93
3	E	2002	MGD	C6-N1-C2	3.91	122.14	115.93
3	G	2001	MGD	C5-C6-N1	-3.89	118.11	123.43
3	E	2001	MGD	C6-C5-C4	-3.85	117.12	120.80
3	G	2002	MGD	C17-N18-C19	3.84	122.04	115.93
3	C	2002	MGD	C17-C16-N15	3.80	122.31	119.12

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	2001	MGD	C6-C5-C4	-3.80	117.17	120.80
3	C	2001	MGD	N3-C2-N1	-3.78	122.19	127.22
3	C	2001	MGD	C1'-N9-C4	-3.76	120.03	126.64
3	A	2001	MGD	C2-N3-C4	3.75	119.64	115.36
3	G	2002	MGD	N3-C2-N1	-3.72	122.26	127.22
3	E	2001	MGD	C17-N18-C19	3.69	121.80	115.93
3	A	2002	MGD	C4-C5-N7	-3.69	105.56	109.40
3	E	2001	MGD	C2-N3-C4	3.67	119.55	115.36
3	A	2001	MGD	C17-C16-C21	3.64	117.80	114.57
3	G	2001	MGD	C17-N18-C19	3.63	121.70	115.93
3	G	2001	MGD	PA-O3B-PB	-3.60	120.47	132.83
3	E	2001	MGD	C5-C6-N1	-3.56	118.56	123.43
3	G	2002	MGD	C5-C6-N1	-3.55	118.57	123.43
3	E	2001	MGD	C17-C16-C21	3.54	117.71	114.57
3	E	2001	MGD	C6-N1-C2	3.53	121.55	115.93
3	C	2001	MGD	C19-N20-C21	3.53	122.45	114.54
3	A	2002	MGD	C17-C16-N15	3.52	122.07	119.12
3	E	2001	MGD	C17-C16-N15	3.50	122.06	119.12
3	G	2001	MGD	N2-C2-N1	3.49	122.68	117.25
3	A	2002	MGD	C6-N1-C2	3.42	121.36	115.93
3	G	2001	MGD	C17-C16-C21	3.37	117.56	114.57
3	G	2001	MGD	C6-N1-C2	3.34	121.24	115.93
3	A	2002	MGD	O11-C23-N22	-3.24	105.24	108.57
3	C	2002	MGD	C6-C5-C4	-3.23	117.71	120.80
3	E	2002	MGD	C19-N20-C21	3.22	121.76	114.54
3	A	2001	MGD	C19-N20-C21	3.20	121.70	114.54
3	E	2001	MGD	C1'-N9-C4	-3.19	121.04	126.64
3	G	2001	MGD	N3-C2-N1	-3.16	123.01	127.22
3	E	2001	MGD	N3-C2-N1	-3.12	123.06	127.22
3	G	2002	MGD	O11-C23-C14	3.12	111.04	108.96
3	A	2001	MGD	PA-O3B-PB	-3.12	122.13	132.83
3	E	2002	MGD	C4-C5-N7	-3.11	106.15	109.40
3	G	2001	MGD	C1'-N9-C4	-3.07	121.24	126.64
3	C	2001	MGD	C17-C16-C21	3.06	117.28	114.57
3	A	2001	MGD	N3-C2-N1	-3.05	123.15	127.22
3	A	2001	MGD	C17-N18-C19	2.99	120.68	115.93
3	G	2002	MGD	C19-N20-C21	2.99	121.24	114.54
3	C	2002	MGD	C19-N20-C21	2.97	121.19	114.54
3	E	2002	MGD	C6-C5-C4	-2.91	118.02	120.80
3	C	2001	MGD	C16-C21-N22	2.84	120.73	118.13
3	A	2002	MGD	C19-N20-C21	2.83	120.89	114.54
3	E	2002	MGD	N18-C19-N20	-2.82	121.00	125.42

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	G	2001	MGD	C19-N20-C21	2.81	120.84	114.54
3	C	2001	MGD	PA-O3B-PB	-2.78	123.27	132.83
3	E	2001	MGD	C19-N20-C21	2.77	120.74	114.54
3	E	2002	MGD	C17-C16-N15	2.73	121.42	119.12
3	C	2001	MGD	C4-C5-N7	-2.69	106.59	109.40
3	A	2001	MGD	C16-C21-N22	2.69	120.59	118.13
3	C	2002	MGD	N3-C2-N1	-2.68	123.65	127.22
3	G	2002	MGD	O4'-C1'-C2'	-2.68	103.01	106.93
3	E	2001	MGD	O11-C23-N22	2.64	111.28	108.57
3	E	2001	MGD	PA-O3B-PB	-2.58	123.99	132.83
3	C	2001	MGD	N18-C19-N20	-2.56	121.40	125.42
3	A	2002	MGD	C17-C16-C21	2.54	116.83	114.57
3	E	2001	MGD	O4'-C4'-C3'	2.44	109.94	105.11
3	E	2002	MGD	O4'-C1'-C2'	-2.44	103.36	106.93
3	C	2002	MGD	O4'-C1'-C2'	-2.43	103.38	106.93
3	C	2001	MGD	N19-C19-N18	2.40	120.99	117.25
3	C	2002	MGD	C2-N3-C4	2.38	118.07	115.36
3	A	2002	MGD	C6-C5-C4	-2.36	118.54	120.80
3	E	2001	MGD	C16-C17-N18	-2.34	117.35	124.01
3	E	2002	MGD	N3-C2-N1	-2.31	124.14	127.22
3	G	2001	MGD	C16-C17-N18	-2.28	117.52	124.01
3	G	2001	MGD	C17-C16-N15	2.23	120.99	119.12
3	C	2002	MGD	N18-C19-N20	-2.21	121.95	125.42
3	G	2001	MGD	C4-C5-N7	-2.21	107.09	109.40
3	C	2001	MGD	N2-C2-N1	2.21	120.68	117.25
3	A	2002	MGD	N3-C2-N1	-2.19	124.31	127.22
3	C	2001	MGD	C16-C17-N18	-2.18	117.82	124.01
3	E	2001	MGD	C16-C21-N22	2.16	120.11	118.13
3	E	2002	MGD	C17-C16-C21	2.15	116.48	114.57
3	A	2002	MGD	O4'-C1'-C2'	-2.14	103.80	106.93
3	G	2002	MGD	C16-C21-N22	2.14	120.09	118.13
3	G	2001	MGD	N2-C2-N3	-2.13	114.31	117.79
3	A	2002	MGD	N18-C19-N20	-2.13	122.08	125.42
3	A	2002	MGD	C16-C17-N18	-2.11	118.01	124.01
3	E	2002	MGD	O2A-PA-O1A	2.10	122.64	112.24
3	C	2001	MGD	O2B-PB-O1B	2.09	122.56	112.24
3	C	2002	MGD	C16-C17-N18	-2.09	118.08	124.01
3	G	2001	MGD	O2A-PA-O1A	2.05	122.38	112.24
3	E	2001	MGD	N2-C2-N1	2.02	120.40	117.25
3	A	2001	MGD	O2A-PA-O1A	2.02	122.21	112.24
3	C	2002	MGD	C1'-N9-C4	-2.02	123.10	126.64
3	C	2001	MGD	O4'-C4'-C3'	2.00	109.08	105.11

There are no chirality outliers.

All (29) torsion outliers are listed below:

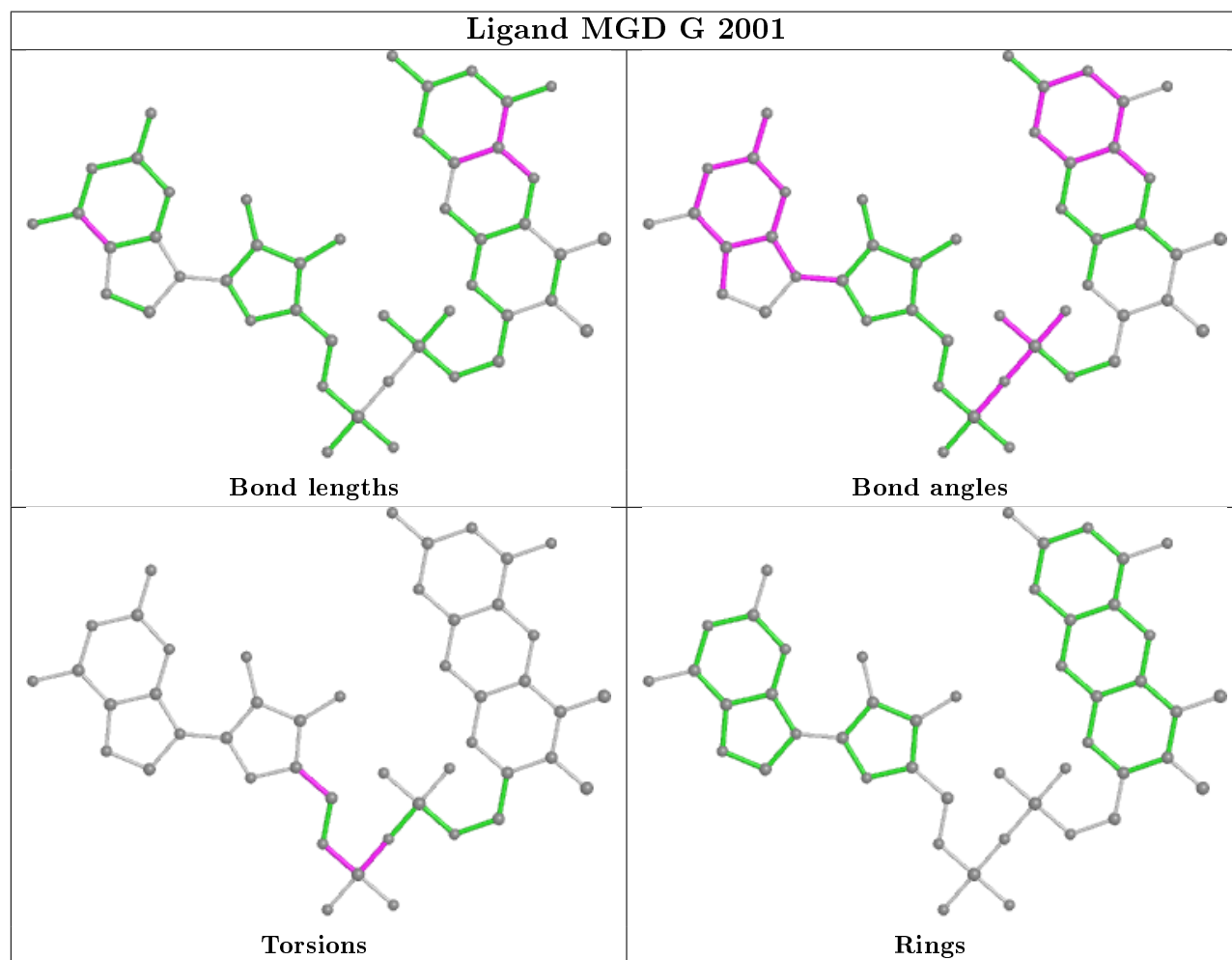
Mol	Chain	Res	Type	Atoms
3	A	2001	MGD	C5'-O5'-PB-O3B
3	C	2001	MGD	PA-O3B-PB-O5'
3	C	2001	MGD	C5'-O5'-PB-O3B
3	A	2002	MGD	PA-O3B-PB-O5'
3	A	2002	MGD	C5'-O5'-PB-O1B
3	A	2002	MGD	C5'-O5'-PB-O3B
3	G	2001	MGD	PA-O3B-PB-O5'
3	G	2001	MGD	C5'-O5'-PB-O1B
3	G	2001	MGD	C5'-O5'-PB-O3B
3	G	2002	MGD	PA-O3B-PB-O5'
3	G	2002	MGD	C5'-O5'-PB-O1B
3	G	2002	MGD	C5'-O5'-PB-O3B
3	E	2001	MGD	C5'-O5'-PB-O3B
3	G	2001	MGD	O4'-C4'-C5'-O5'
3	A	2001	MGD	PA-O3B-PB-O5'
3	E	2002	MGD	PA-O3B-PB-O5'
3	E	2001	MGD	PA-O3B-PB-O5'
3	C	2002	MGD	PA-O3B-PB-O5'
3	E	2001	MGD	O4'-C4'-C5'-O5'
3	C	2002	MGD	C5'-O5'-PB-O3B
3	A	2001	MGD	C5'-O5'-PB-O1B
3	C	2001	MGD	C5'-O5'-PB-O1B
3	E	2001	MGD	C5'-O5'-PB-O1B
3	A	2001	MGD	O4'-C4'-C5'-O5'
3	G	2001	MGD	C3'-C4'-C5'-O5'
3	C	2001	MGD	O4'-C4'-C5'-O5'
3	E	2002	MGD	C5'-O5'-PB-O1B
3	C	2002	MGD	C5'-O5'-PB-O1B
3	E	2001	MGD	C3'-C4'-C5'-O5'

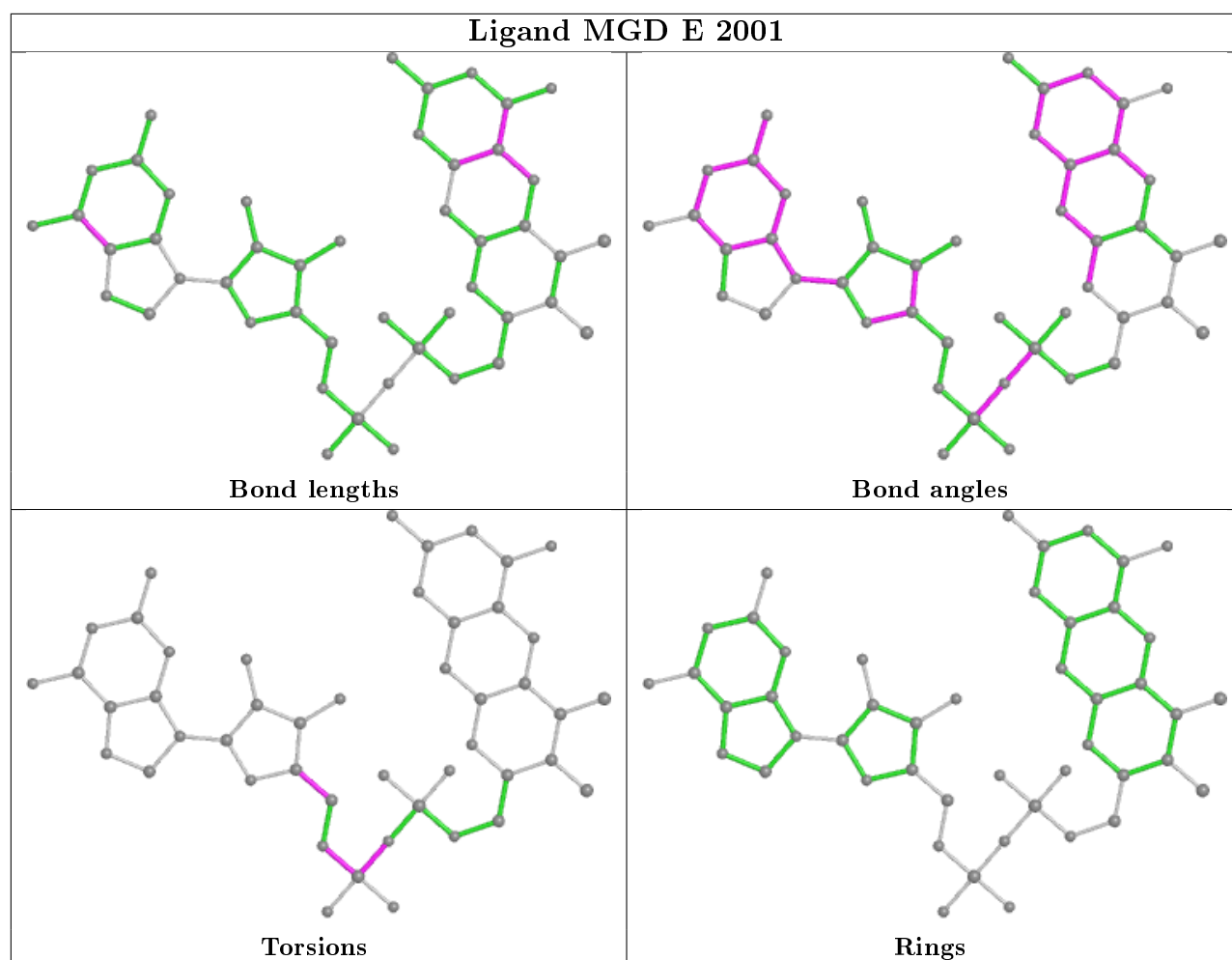
There are no ring outliers.

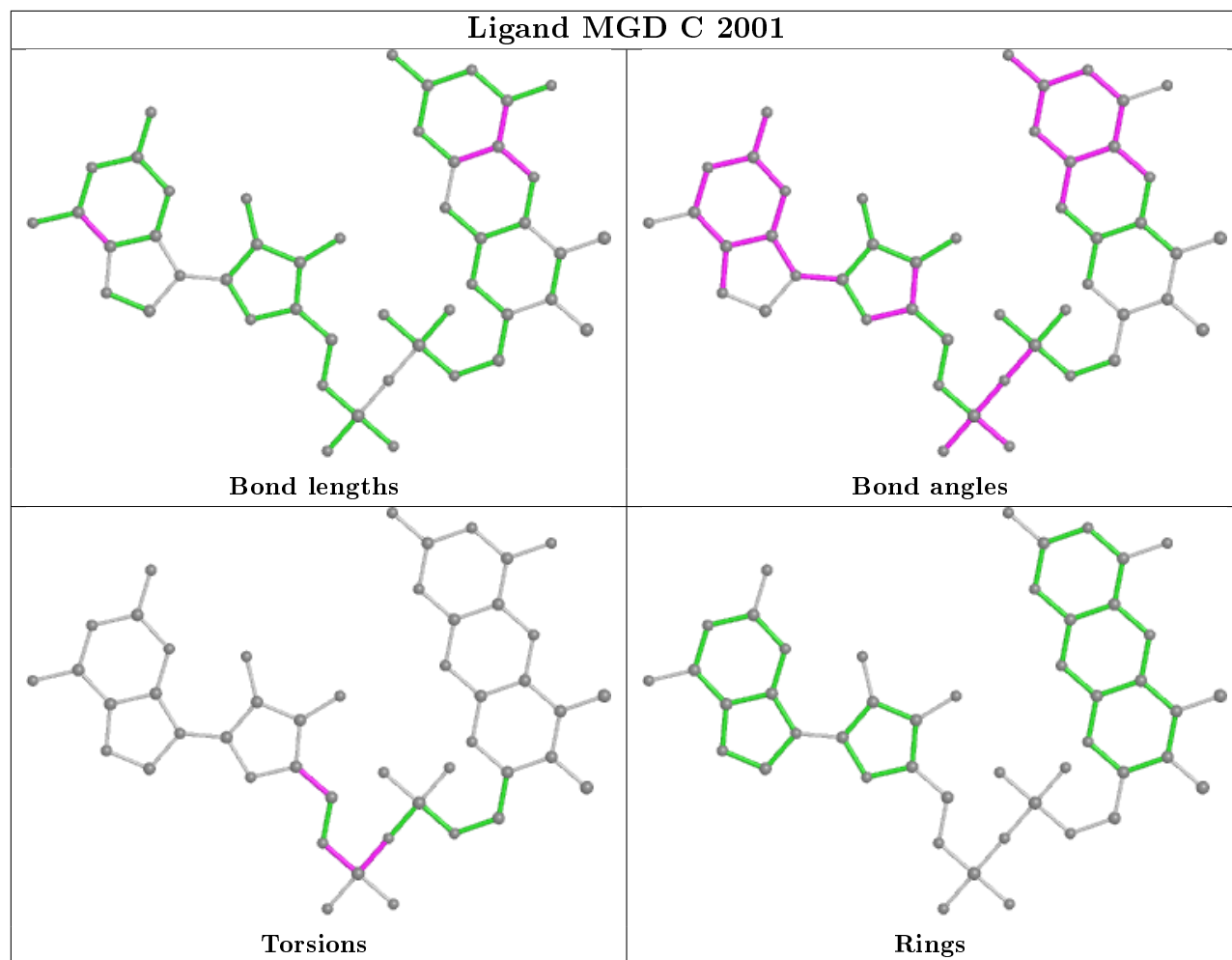
5 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	G	2001	MGD	1	0
3	E	2001	MGD	1	0
3	A	2001	MGD	1	0
7	D	2006	FES	1	0
3	C	2002	MGD	1	0

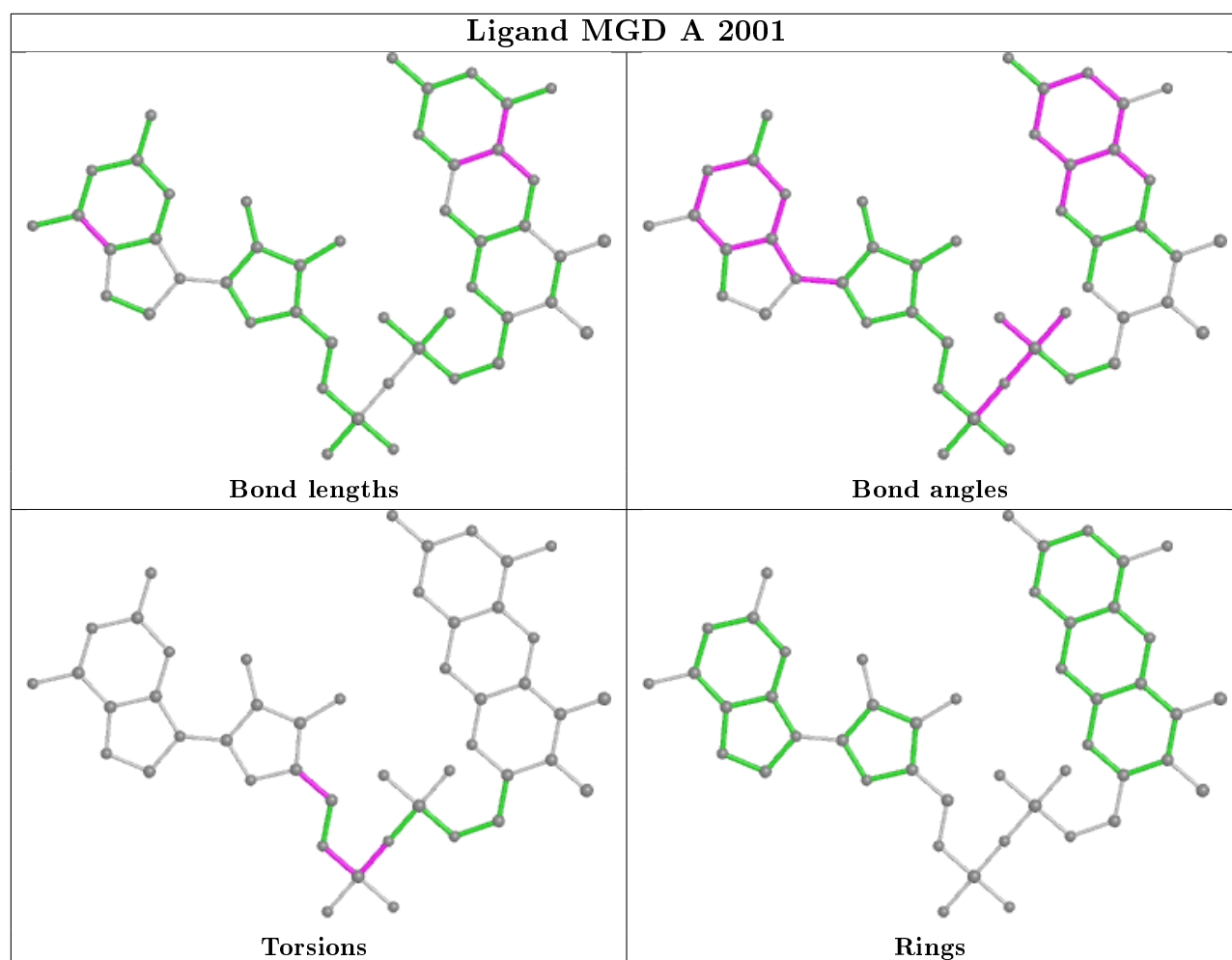
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.

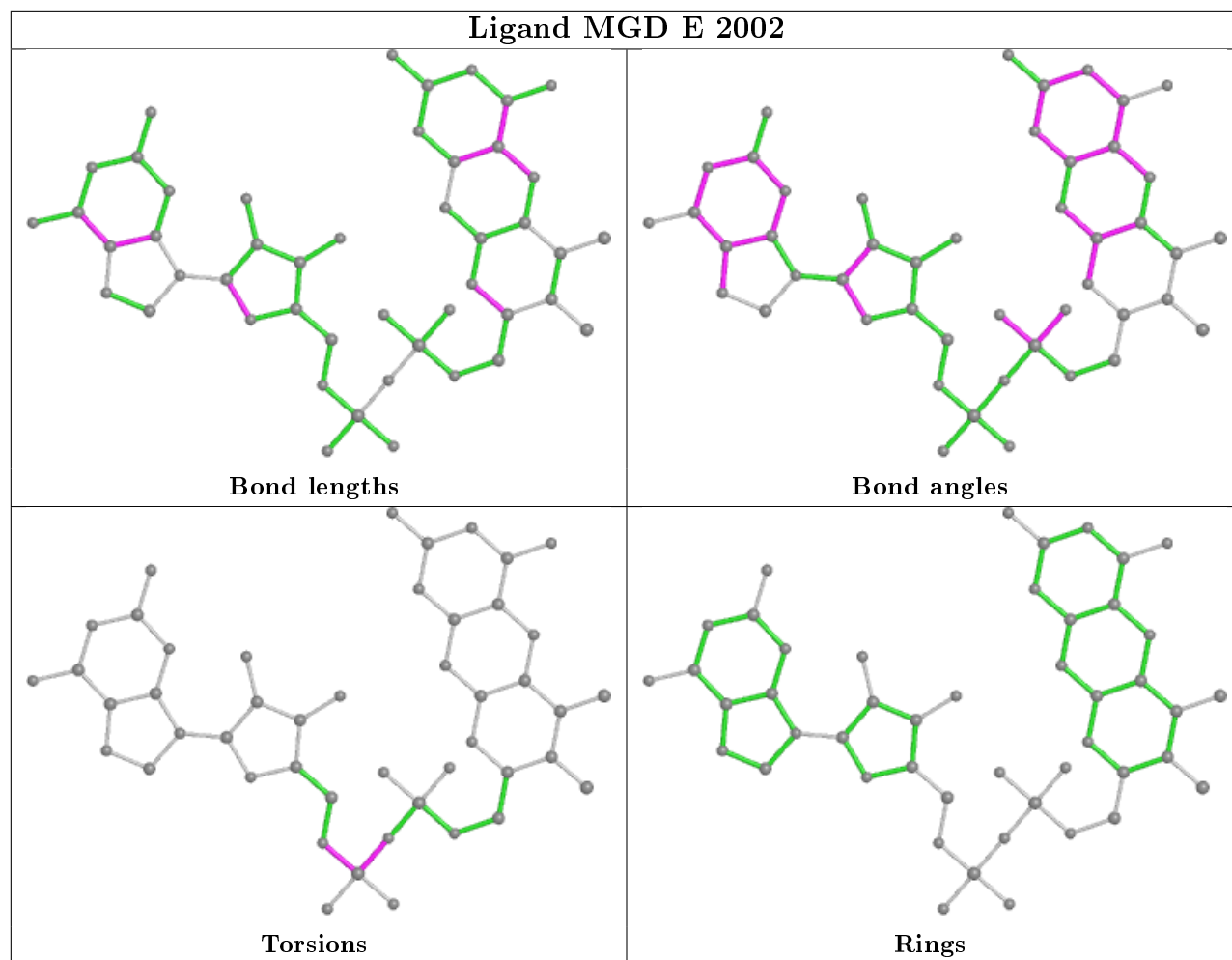


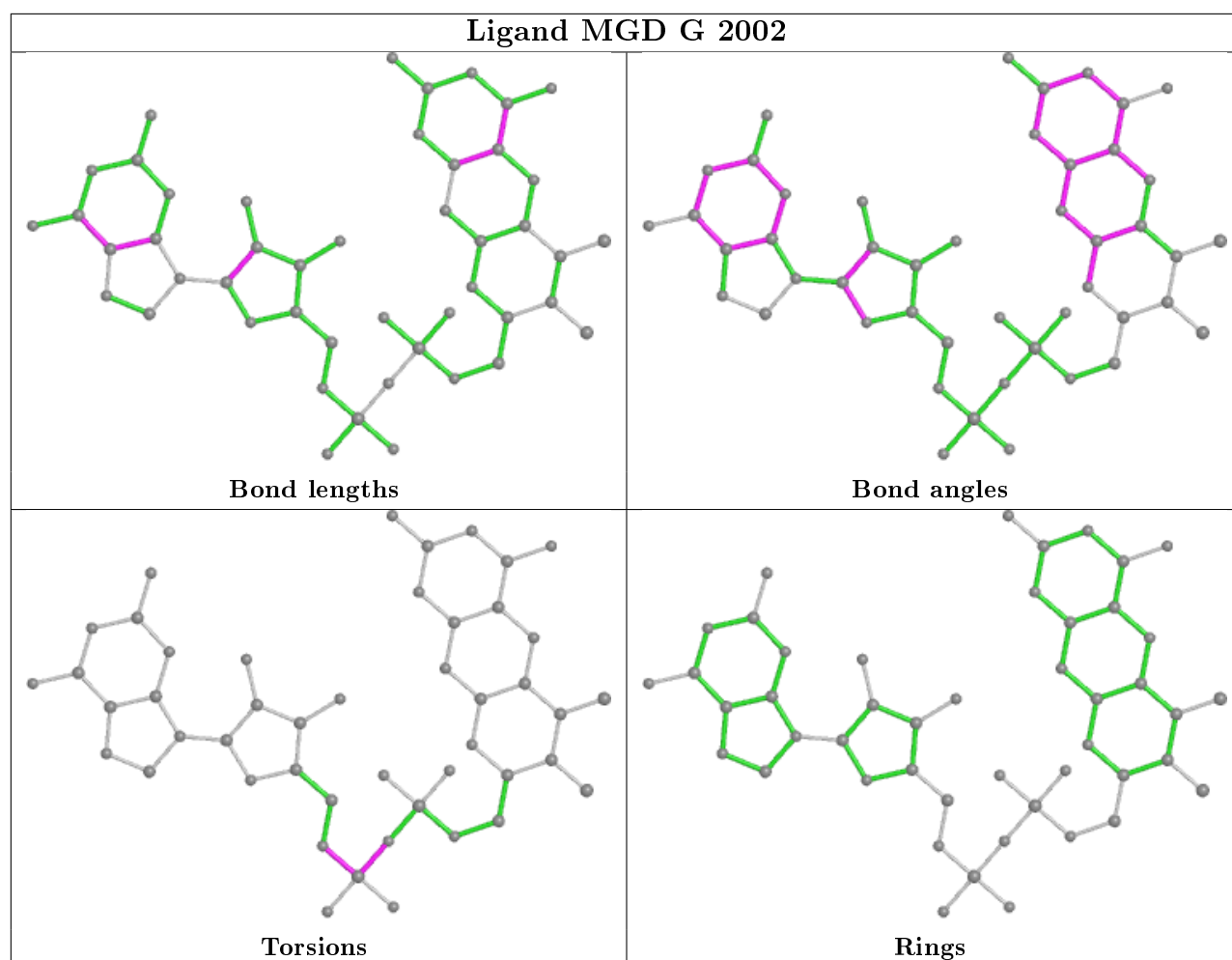


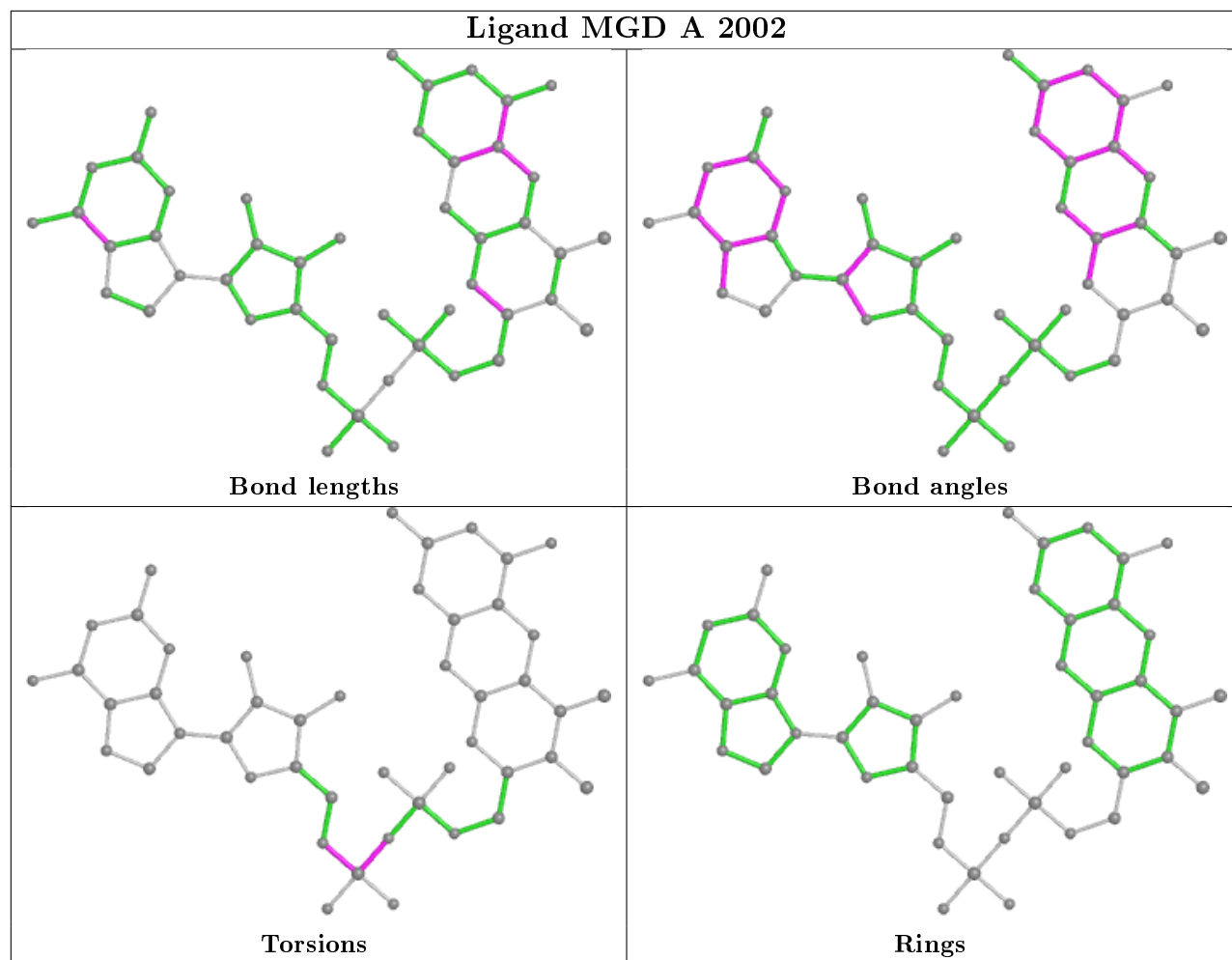


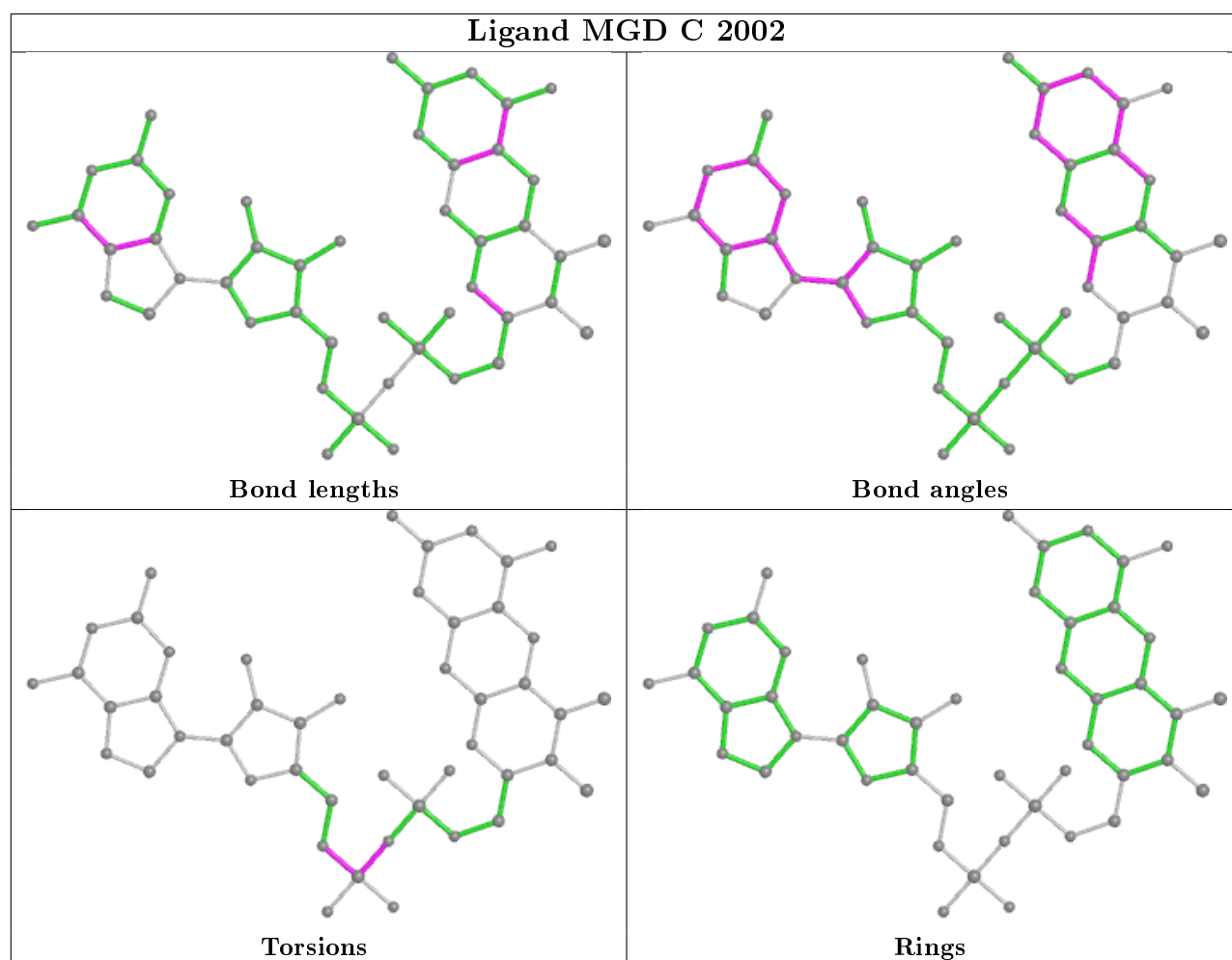












## 5.7 Other polymers [i](#)

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	843/845 (99%)	0.06	19 (2%) 60 62	44, 49, 70, 93	0
1	C	843/845 (99%)	0.10	24 (2%) 53 54	45, 49, 72, 104	0
1	E	843/845 (99%)	0.09	24 (2%) 53 54	45, 51, 77, 105	0
1	G	843/845 (99%)	0.11	23 (2%) 54 55	44, 53, 83, 112	0
2	B	132/175 (75%)	0.53	11 (8%) 11 9	49, 69, 105, 133	0
2	D	132/175 (75%)	0.94	23 (17%) 1 1	47, 68, 240, 343	0
2	F	132/175 (75%)	0.56	9 (6%) 17 15	49, 69, 96, 105	0
2	H	132/175 (75%)	0.83	24 (18%) 1 1	49, 78, 168, 322	0
All	All	3900/4080 (95%)	0.17	157 (4%) 38 37	44, 52, 85, 343	0

All (157) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	46	GLY	8.7
2	D	45	ALA	8.6
2	H	46	GLY	7.9
2	H	45	ALA	7.4
2	D	47	VAL	6.6
2	H	47	VAL	6.2
2	D	164	ASP	5.9
1	G	269	GLU	5.9
1	E	269	GLU	5.8
1	G	400	GLU	5.6
1	E	400	GLU	5.5
2	D	165	GLU	5.2
1	C	400	GLU	5.1
2	H	152	VAL	5.0
1	A	269	GLU	4.8
1	E	340	GLY	4.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	340	GLY	4.3
2	D	53	ARG	4.2
2	H	57	ILE	4.2
2	F	152	VAL	4.1
2	H	162	GLY	4.1
1	G	268	PRO	4.1
1	E	268	PRO	4.0
2	D	49	TYR	4.0
2	F	154	ASP	4.0
1	C	341	THR	4.0
1	C	686	ASP	3.9
1	A	400	GLU	3.7
1	G	340	GLY	3.7
2	B	155	ASN	3.7
1	A	340	GLY	3.5
2	H	95	ASP	3.5
1	G	641	HIS	3.4
1	E	273	GLU	3.4
1	A	668	ASP	3.4
2	B	175	LEU	3.4
2	D	54	LEU	3.4
1	C	269	GLU	3.3
1	C	339	GLU	3.2
1	G	686	ASP	3.2
1	E	751	GLU	3.1
1	C	352	VAL	3.1
1	G	272	HIS	3.1
1	C	785	ARG	3.1
1	E	41	ASN	3.1
1	E	686	ASP	3.1
1	E	339	GLU	3.1
1	E	401	GLY	3.1
1	A	339	GLU	3.1
1	G	271	PRO	3.1
1	C	121[A]	GLN	3.0
1	G	273	GLU	3.0
1	G	668	ASP	3.0
1	E	346	LEU	3.0
2	D	152	VAL	3.0
1	A	641	HIS	3.0
2	D	48	GLU	3.0
2	H	151	ARG	3.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	G	256	ARG	2.9
1	E	265	GLU	2.9
2	B	153	ALA	2.9
2	F	57	ILE	2.9
2	D	55	ALA	2.9
1	A	47	ASP	2.9
1	G	401	GLY	2.8
2	H	91	GLY	2.8
1	A	62	GLN	2.8
2	H	160	ALA	2.7
2	D	95	ASP	2.7
1	C	47	ASP	2.7
2	B	154	ASP	2.7
2	F	93	ASP	2.7
1	G	368	GLU	2.7
1	G	49	GLN	2.7
1	C	365	ASP	2.7
1	C	624	ASP	2.7
2	H	49	TYR	2.6
2	D	50	PRO	2.6
1	A	49	GLN	2.6
1	A	41	ASN	2.6
1	C	41	ASN	2.6
2	H	48	GLU	2.6
2	D	93	ASP	2.6
2	D	84	THR	2.6
2	D	158	ILE	2.6
1	G	672	GLU	2.6
2	F	58	SER	2.6
1	G	339	GLU	2.5
1	E	329	ARG	2.5
2	D	157	ASP	2.5
2	F	64	GLU	2.5
1	G	270	GLU	2.5
1	C	368	GLU	2.5
1	E	644	GLU	2.5
2	B	66	LEU	2.4
1	C	668	ASP	2.4
2	H	115	ASP	2.4
1	C	54	GLY	2.4
1	E	92	HIS	2.4
1	G	751	GLU	2.4

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	H	161	GLU	2.4
1	A	512	ASP	2.4
1	E	687	ASP	2.4
1	C	376	GLU	2.4
2	F	80	LEU	2.4
1	G	403	LYS	2.4
1	A	55	VAL	2.4
2	B	62	LEU	2.3
2	H	96	ILE	2.3
2	D	159	PHE	2.3
1	C	699	GLY	2.3
1	A	624	ASP	2.3
1	G	624	ASP	2.3
1	C	274	ALA	2.3
1	C	769	ASP	2.3
2	B	151	ARG	2.3
1	A	341	THR	2.3
1	A	16	ASP	2.3
1	E	668	ASP	2.3
1	E	298	ASP	2.3
1	C	344	PRO	2.3
1	G	298	ASP	2.2
1	A	368	GLU	2.2
1	C	298	ASP	2.2
1	E	755	GLU	2.2
2	D	115	ASP	2.2
2	F	73	GLU	2.2
1	G	621	LYS	2.2
2	H	92	PRO	2.2
2	B	159	PHE	2.2
1	G	41	ASN	2.2
2	H	155	ASN	2.2
2	H	54	LEU	2.2
1	E	769	ASP	2.2
2	F	44	ALA	2.1
1	A	236	GLY	2.1
2	H	58	SER	2.1
1	A	298	ASP	2.1
2	D	156	GLY	2.1
2	H	83	GLY	2.1
2	B	152	VAL	2.1
2	H	163	VAL	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	H	157	ASP	2.1
1	E	229	ALA	2.1
2	D	155	ASN	2.1
1	C	379	GLU	2.1
1	A	669	GLY	2.1
2	B	45	ALA	2.0
2	D	85	ARG	2.0
1	E	402	GLY	2.0
2	D	162	GLY	2.0
1	C	751	GLU	2.0
1	E	65	SER	2.0
2	B	95	ASP	2.0
2	H	154	ASP	2.0
2	H	84	THR	2.0
1	E	55	VAL	2.0

## 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( $\text{\AA}^2$ )	Q<0.9
7	FES	F	2006	4/4	0.86	0.12	74,76,77,84	0
7	FES	H	2006	4/4	0.89	0.15	81,81,82,88	0
7	FES	B	2006	4/4	0.94	0.15	74,79,82,85	0
4	O	G	2003	1/1	0.95	0.23	56,56,56,56	0
3	MGD	C	2002	47/47	0.95	0.20	50,52,55,55	0
6	F3S	E	2005	7/7	0.96	0.09	56,56,57,57	0
3	MGD	A	2001	47/47	0.96	0.18	51,53,54,55	0
3	MGD	E	2002	47/47	0.96	0.20	52,53,56,57	0

*Continued on next page...*

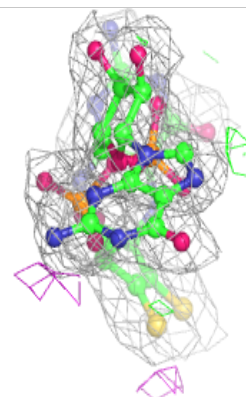
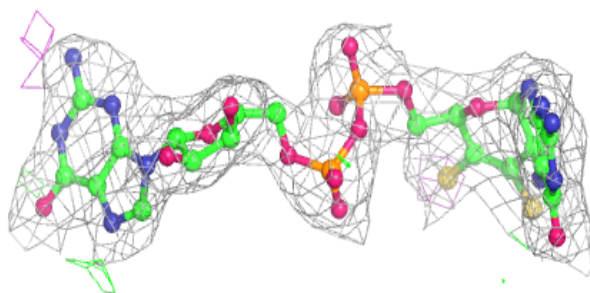
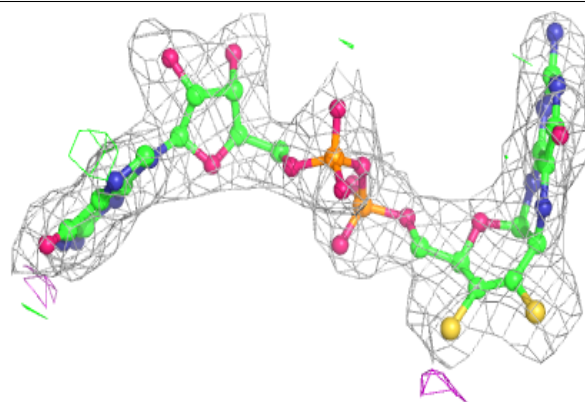
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	MGD	G	2002	47/47	0.96	0.18	53,55,58,59	0
6	F3S	A	2005	7/7	0.96	0.09	53,54,55,55	0
6	F3S	G	2005	7/7	0.96	0.09	56,57,57,57	0
4	O	C	2003	1/1	0.96	0.20	56,56,56,56	0
3	MGD	E	2001	47/47	0.96	0.17	53,54,57,57	0
3	MGD	A	2002	47/47	0.96	0.20	48,51,51,52	0
3	MGD	G	2001	47/47	0.96	0.17	52,55,57,58	0
4	O	A	2003	1/1	0.97	0.17	55,55,55,55	0
7	FES	D	2006	4/4	0.97	0.14	79,80,83,86	0
3	MGD	C	2001	47/47	0.97	0.18	52,53,54,55	0
6	F3S	C	2005	7/7	0.97	0.08	54,55,56,57	0
4	O	E	2003	1/1	0.98	0.14	57,57,57,57	0
5	4MO	E	2004	1/1	0.99	0.09	57,57,57,57	0
5	4MO	A	2004	1/1	1.00	0.09	53,53,53,53	0
5	4MO	C	2004	1/1	1.00	0.11	54,54,54,54	0
5	4MO	G	2004	1/1	1.00	0.09	56,56,56,56	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.

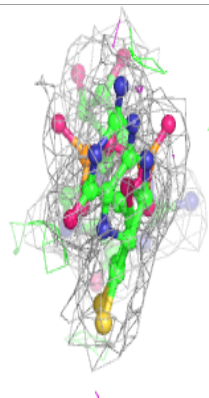
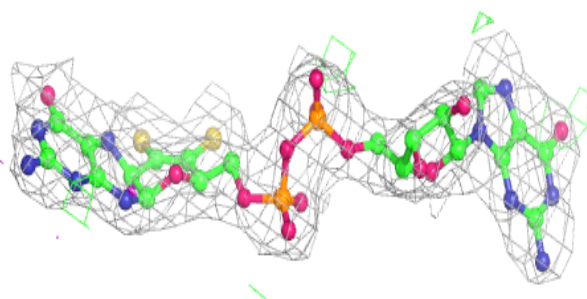
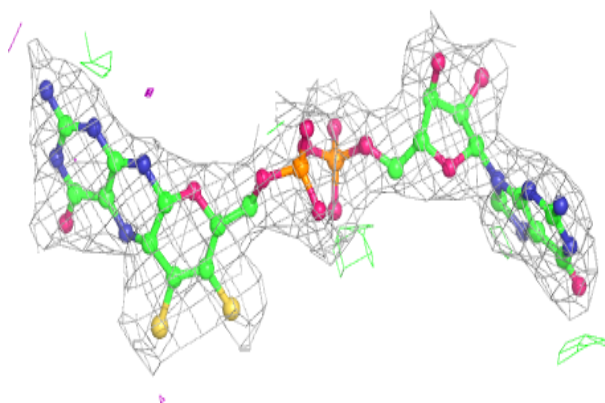
**Electron density around MGD C 2002:**

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)

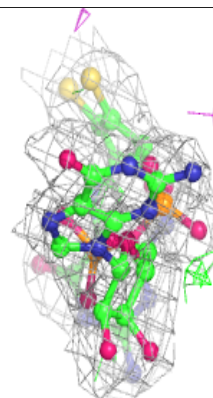
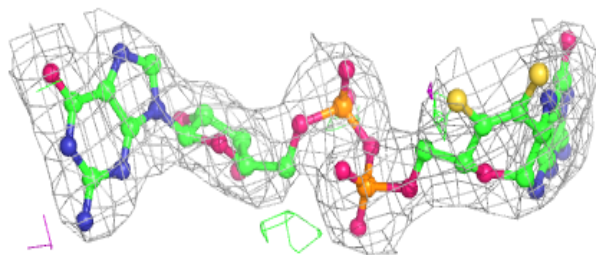
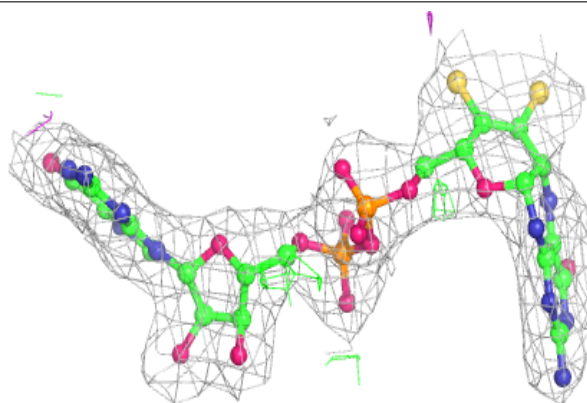


**Electron density around MGD A 2001:**

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

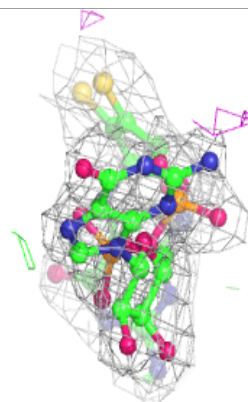
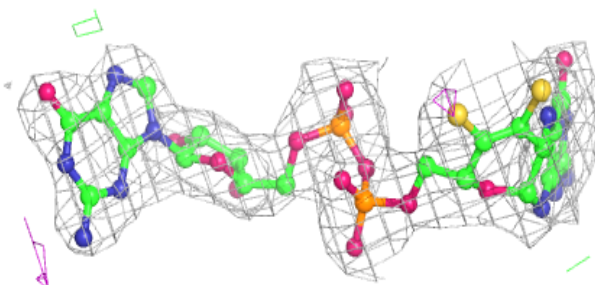
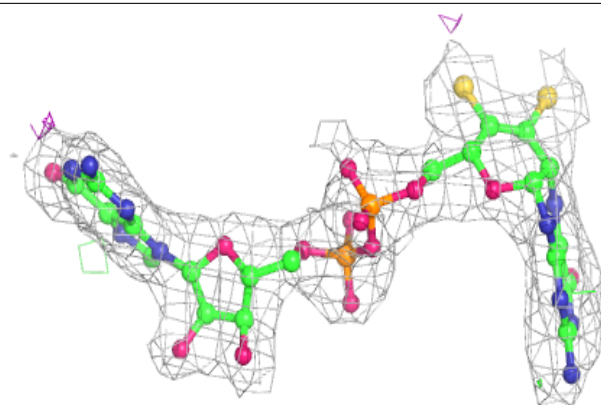
**Electron density around MGD E 2002:**

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

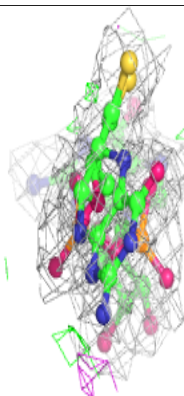
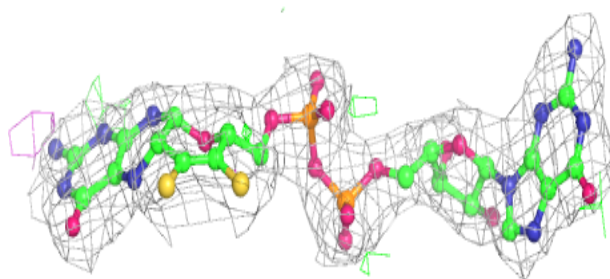
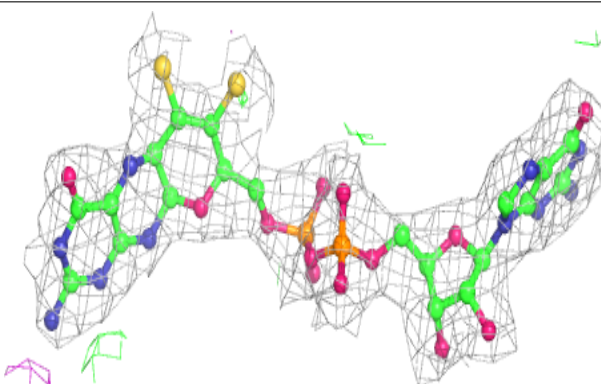


**Electron density around MGD G 2002:**

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

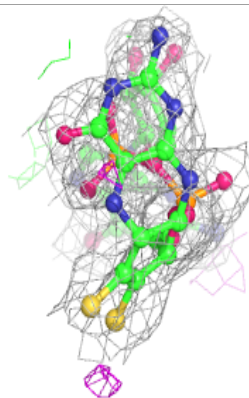
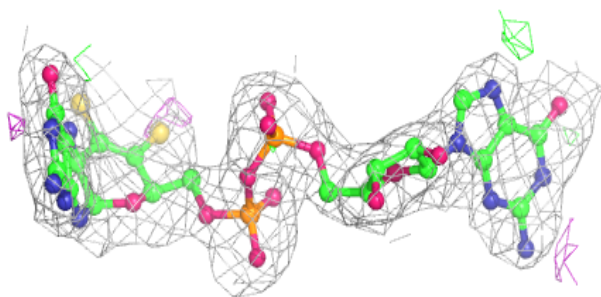
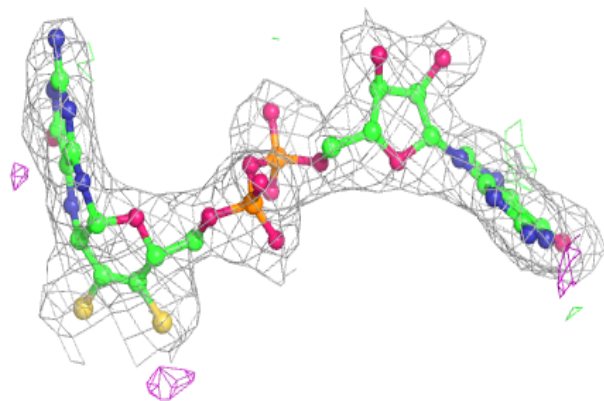
**Electron density around MGD E 2001:**

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

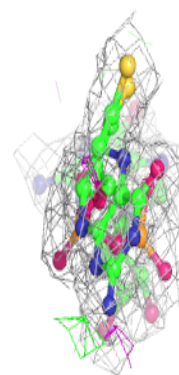
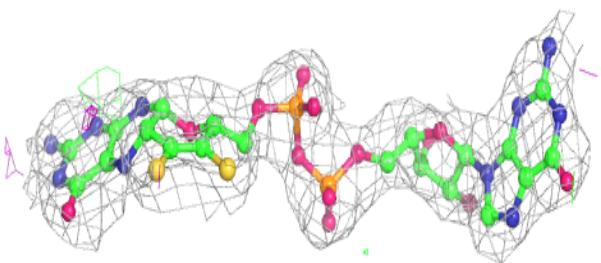
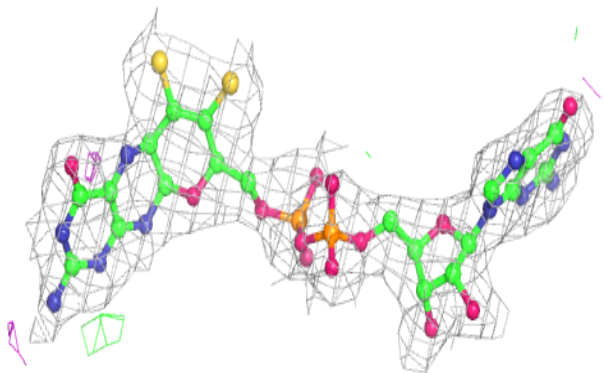


**Electron density around MGD A 2002:**

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

**Electron density around MGD G 2001:**

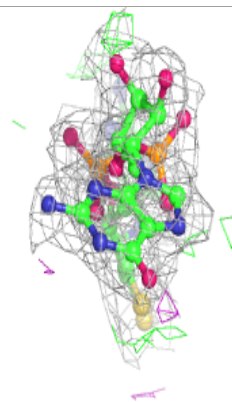
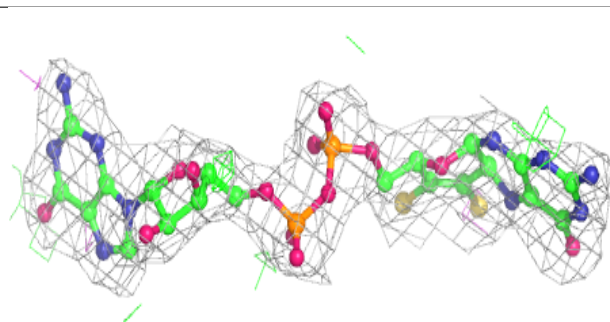
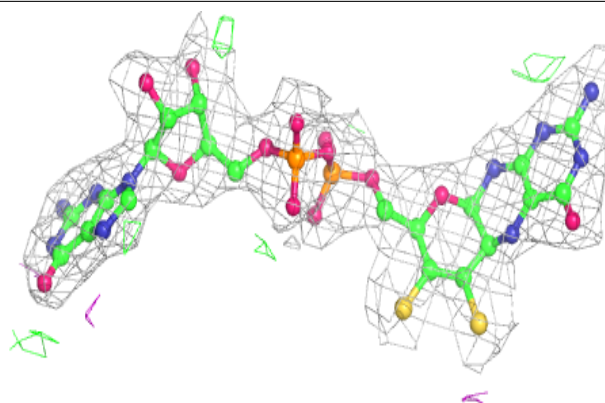
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





**Electron density around MGD C 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.