



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 25, 2024 – 04:49 AM EST

PDB ID : 7KHV  
Title : CpOGA IN COMPLEX WITH LIGAND 54  
Authors : Shaffer, P.L.  
Deposited on : 2020-10-22  
Resolution : 2.30 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

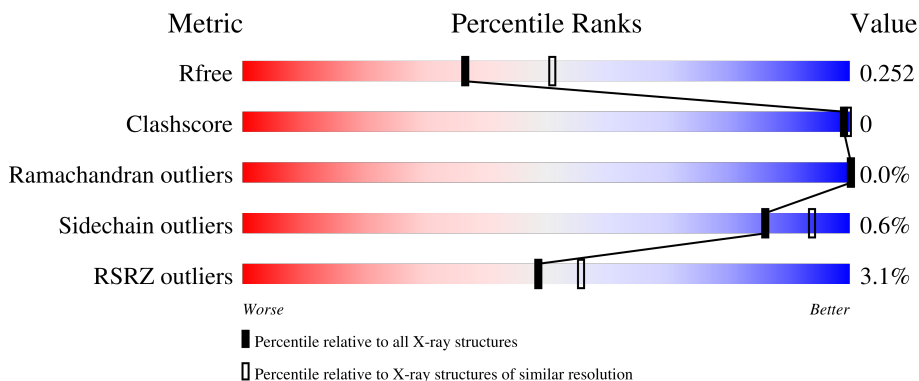
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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.



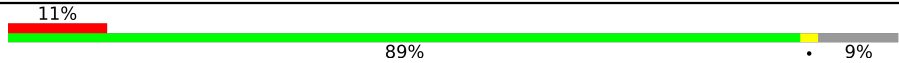
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	594	97%
1	B	594	97%
1	C	594	97%
1	D	594	97%
1	E	594	98%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	594	 <p>A horizontal bar chart representing the quality of chain. The bar is divided into three segments: a red segment on the left labeled '11%', a large green segment in the middle labeled '89%', and a small grey segment on the right labeled '9%'. A small black dot is visible at the end of the grey segment.</p>

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 55791 atoms, of which 26384 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 O-GlcNAcase NagJ.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	586	9067	2915	4431	760	943	18	124	0	0
1	B	585	9110	2928	4449	765	950	18	116	4	0
1	C	585	9079	2919	4436	760	946	18	137	2	0
1	D	585	9053	2911	4425	758	941	18	122	0	0
1	E	586	9067	2915	4431	760	943	18	140	0	0
1	F	539	8367	2698	4080	704	867	18	209	0	0

There are 18 discrepancies between the modelled and reference sequences:

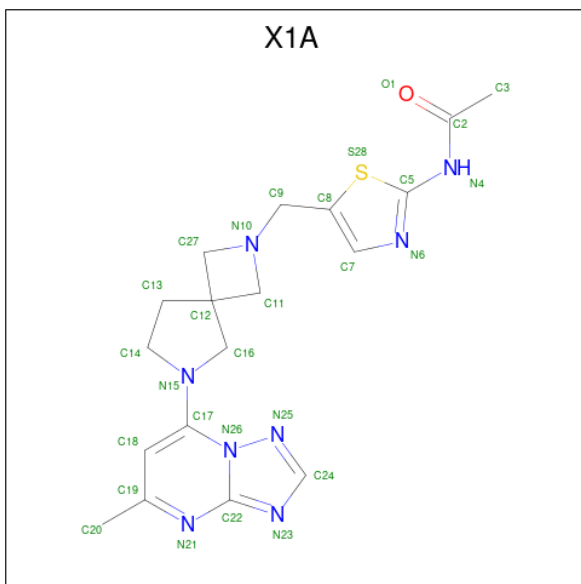
Chain	Residue	Modelled	Actual	Comment	Reference
A	298	ASN	ASP	engineered mutation	UNP Q0TR53
A	331	CYS	VAL	engineered mutation	UNP Q0TR53
A	388	ASP	ASN	engineered mutation	UNP Q0TR53
B	298	ASN	ASP	engineered mutation	UNP Q0TR53
B	331	CYS	VAL	engineered mutation	UNP Q0TR53
B	388	ASP	ASN	engineered mutation	UNP Q0TR53
C	298	ASN	ASP	engineered mutation	UNP Q0TR53
C	331	CYS	VAL	engineered mutation	UNP Q0TR53
C	388	ASP	ASN	engineered mutation	UNP Q0TR53
D	298	ASN	ASP	engineered mutation	UNP Q0TR53
D	331	CYS	VAL	engineered mutation	UNP Q0TR53
D	388	ASP	ASN	engineered mutation	UNP Q0TR53
E	298	ASN	ASP	engineered mutation	UNP Q0TR53
E	331	CYS	VAL	engineered mutation	UNP Q0TR53
E	388	ASP	ASN	engineered mutation	UNP Q0TR53
F	298	ASN	ASP	engineered mutation	UNP Q0TR53
F	331	CYS	VAL	engineered mutation	UNP Q0TR53

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	388	ASP	ASN	engineered mutation	UNP Q0TR53

- Molecule 2 is N-(5-{[6-(5-methyl[1,2,4]triazolo[1,5-a]pyrimidin-7-yl)-2,6-diazaspiro[3.4]octan-2-yl]methyl}-1,3-thiazol-2-yl)acetamide (three-letter code: X1A) (formula: C<sub>18</sub>H<sub>22</sub>N<sub>8</sub>OS).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	H	N	O			S
2	A	1	Total	C	H	N	O	S	0	0
			50	18	22	8	1	1		
2	B	1	Total	C	H	N	O	S	0	0
			50	18	22	8	1	1		
2	C	1	Total	C	H	N	O	S	0	0
			50	18	22	8	1	1		
2	D	1	Total	C	H	N	O	S	0	0
			50	18	22	8	1	1		
2	E	1	Total	C	H	N	O	S	0	0
			50	18	22	8	1	1		
2	F	1	Total	C	H	N	O	S	0	0
			50	18	22	8	1	1		

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Ca	0	0
			1	1		
3	D	1	Total	Ca	0	0
			1	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
3	E	1	1	1	0	0

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

Continued on next page...

*Continued from previous page...*

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

*Continued on next page...*

*Continued from previous page...*

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

- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	2	Total	Cl	0	0
			2	2		
5	C	1	Total	Cl	0	0
			1	1		
5	D	1	Total	Cl	0	0
			1	1		
5	F	1	Total	Cl	0	0
			1	1		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	324	Total	O	0	0
			324	324		
6	B	367	Total	O	0	0
			367	367		
6	C	332	Total	O	0	0
			332	332		
6	D	243	Total	O	0	0
			243	243		
6	E	159	Total	O	0	0
			159	159		
6	F	140	Total	O	0	0
			140	140		



### 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: O-GlcNAcase NagJ

Chain A:  97%



- Molecule 1: O-GlcNAcase NagJ

Chain B:  97%



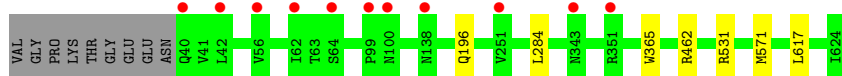
- Molecule 1: O-GlcNAcase NagJ

Chain C:  97%



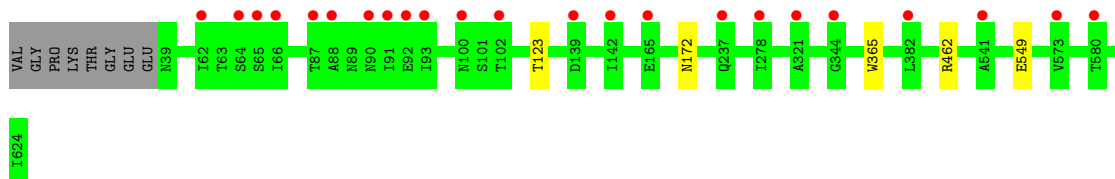
- Molecule 1: O-GlcNAcase NagJ

Chain D:  97%

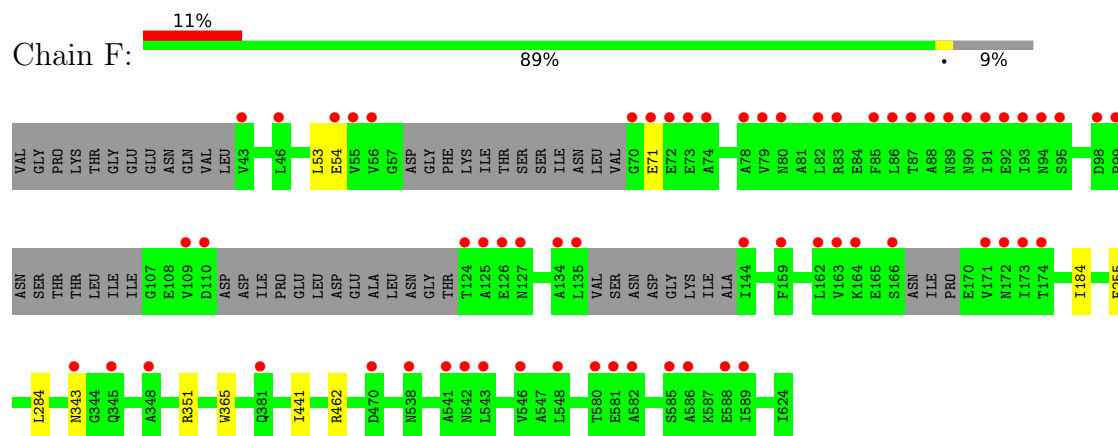


- Molecule 1: O-GlcNAcase NagJ

Chain E:  98%



- Molecule 1: O-GlcNAcase NagJ



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	149.49Å 178.10Å 407.05Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	203.53 – 2.30 48.51 – 2.30	Depositor EDS
% Data completeness (in resolution range)	94.2 (203.53-2.30) 94.3 (48.51-2.30)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.84 (at 2.29Å)	Xtrriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.206 , 0.247 0.211 , 0.252	Depositor DCC
$R_{free}$ test set	1081 reflections (0.48%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	26.8	Xtrriage
Anisotropy	0.774	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 42.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	55791	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.06% 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: X1A, SO4, CA, CL

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.51	0/4732	0.71	3/6424 (0.0%)
1	B	0.50	0/4757	0.72	4/6458 (0.1%)
1	C	0.50	0/4739	0.72	2/6433 (0.0%)
1	D	0.52	0/4724	0.70	2/6413 (0.0%)
1	E	0.54	1/4732 (0.0%)	0.73	1/6424 (0.0%)
1	F	0.56	1/4375 (0.0%)	0.72	3/5927 (0.1%)
All	All	0.52	2/28059 (0.0%)	0.72	15/38079 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	549	GLU	CG-CD	5.72	1.60	1.51
1	F	54	GLU	CG-CD	-5.14	1.44	1.51

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	53	LEU	CB-CG-CD2	6.39	121.87	111.00
1	B	83	ARG	NE-CZ-NH1	5.89	123.24	120.30
1	F	462	ARG	NE-CZ-NH1	5.75	123.18	120.30
1	A	502	ARG	NE-CZ-NH1	5.74	123.17	120.30
1	A	238	ARG	NE-CZ-NH1	5.58	123.09	120.30
1	E	462	ARG	NE-CZ-NH1	5.53	123.07	120.30
1	B	502	ARG	NE-CZ-NH1	5.48	123.04	120.30
1	D	462	ARG	NE-CZ-NH1	5.47	123.03	120.30
1	C	462	ARG	NE-CZ-NH1	5.38	122.99	120.30
1	A	462	ARG	NE-CZ-NH1	5.34	122.97	120.30
1	F	351	ARG	NE-CZ-NH1	5.19	122.89	120.30
1	B	531	ARG	NE-CZ-NH1	5.16	122.88	120.30
1	C	502	ARG	NE-CZ-NH1	5.11	122.85	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	531	ARG	NE-CZ-NH1	5.07	122.83	120.30
1	B	462	ARG	NE-CZ-NH1	5.06	122.83	120.30

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4636	4431	4422	2	0
1	B	4661	4449	4436	2	0
1	C	4643	4436	4425	3	0
1	D	4628	4425	4416	1	0
1	E	4636	4431	4422	1	0
1	F	4287	4080	4066	2	0
2	A	28	22	0	1	0
2	B	28	22	0	0	0
2	C	28	22	0	1	0
2	D	28	22	0	0	0
2	E	28	22	0	1	0
2	F	28	22	0	1	0
3	A	1	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0
4	A	30	0	0	0	0
4	B	25	0	0	0	0
4	C	35	0	0	0	0
4	D	40	0	0	0	0
4	E	20	0	0	0	0
4	F	25	0	0	0	0
5	B	2	0	0	0	0
5	C	1	0	0	0	0
5	D	1	0	0	0	0
5	F	1	0	0	0	0
6	A	324	0	0	0	0
6	B	367	0	0	1	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	C	332	0	0	0	0
6	D	243	0	0	0	0
6	E	159	0	0	0	0
6	F	140	0	0	0	0
All	All	29407	26384	26187	15	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 0.

All (15) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:453[B]:ASN:ND2	6:B:801:HOH:O	1.88	1.04
2:A:701:X1A:S28	2:A:701:X1A:O1	2.66	0.54
2:C:701:X1A:O1	2:C:701:X1A:S28	2.67	0.52
1:C:436:LEU:HD21	1:C:551:CYS:HB3	1.93	0.49
2:F:701:X1A:O1	2:F:701:X1A:S28	2.74	0.45
1:C:68:LEU:HG	1:C:71:GLU:HG2	2.00	0.44
1:B:502:ARG:HD3	1:B:502:ARG:O	2.18	0.44
1:F:255:PHE:CD1	1:F:284:LEU:HD22	2.53	0.43
2:E:701:X1A:O1	2:E:701:X1A:S28	2.76	0.43
1:A:502:ARG:HD3	1:A:502:ARG:O	2.18	0.43
1:D:571:MET:HE1	1:D:617:LEU:HA	2.01	0.42
1:A:255:PHE:CD1	1:A:284:LEU:HD22	2.56	0.41
1:E:123:THR:HG21	1:E:172:ASN:OD1	2.20	0.41
1:C:502:ARG:O	1:C:502:ARG:HD3	2.21	0.41
1:F:184:ILE:HD11	1:F:441:ILE:HG23	2.03	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	584/594 (98%)	567 (97%)	17 (3%)	0	100	100
1	B	587/594 (99%)	573 (98%)	14 (2%)	0	100	100
1	C	585/594 (98%)	567 (97%)	17 (3%)	1 (0%)	47	58
1	D	583/594 (98%)	567 (97%)	16 (3%)	0	100	100
1	E	584/594 (98%)	565 (97%)	19 (3%)	0	100	100
1	F	527/594 (89%)	511 (97%)	16 (3%)	0	100	100
All	All	3450/3564 (97%)	3350 (97%)	99 (3%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	41	VAL

### 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	501/507 (99%)	495 (99%)	6 (1%)	71	84
1	B	504/507 (99%)	503 (100%)	1 (0%)	93	97
1	C	502/507 (99%)	498 (99%)	4 (1%)	81	91
1	D	500/507 (99%)	497 (99%)	3 (1%)	86	94
1	E	501/507 (99%)	500 (100%)	1 (0%)	93	97
1	F	459/507 (90%)	456 (99%)	3 (1%)	84	92
All	All	2967/3042 (98%)	2949 (99%)	18 (1%)	86	94

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	196	GLN
1	A	257	ILE
1	A	284	LEU
1	A	355	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	365	TRP
1	A	420	ASN
1	B	365	TRP
1	C	257	ILE
1	C	284	LEU
1	C	365	TRP
1	C	556	ASP
1	D	196	GLN
1	D	284	LEU
1	D	365	TRP
1	E	365	TRP
1	F	71	GLU
1	F	343	ASN
1	F	365	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	298	ASN
1	A	453	ASN
1	B	100	ASN
1	B	196	GLN
1	B	298	ASN
1	B	510	ASN
1	C	196	GLN
1	C	298	ASN
1	C	390	ASN
1	E	298	ASN
1	F	52	ASN
1	F	298	ASN
1	F	442	HIS
1	F	610	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](#)

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



## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 49 ligands modelled in this entry, 8 are monoatomic - leaving 41 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	SO4	E	704	-	4,4,4	0.32	0	6,6,6	0.20	0
4	SO4	B	707	-	4,4,4	0.30	0	6,6,6	0.20	0
2	X1A	A	701	-	23,32,32	1.52	3 (13%)	18,48,48	2.71	10 (55%)
4	SO4	B	705	-	4,4,4	0.34	0	6,6,6	0.25	0
4	SO4	D	704	-	4,4,4	0.36	0	6,6,6	0.12	0
4	SO4	F	704	-	4,4,4	0.32	0	6,6,6	0.11	0
4	SO4	A	706	-	4,4,4	0.37	0	6,6,6	0.16	0
4	SO4	F	707	-	4,4,4	0.37	0	6,6,6	0.09	0
2	X1A	C	701	-	23,32,32	1.45	3 (13%)	18,48,48	3.03	11 (61%)
4	SO4	A	708	-	4,4,4	0.36	0	6,6,6	0.19	0
4	SO4	C	709	-	4,4,4	0.33	0	6,6,6	0.20	0
4	SO4	C	703	-	4,4,4	0.35	0	6,6,6	0.20	0
4	SO4	A	704	-	4,4,4	0.40	0	6,6,6	0.24	0
4	SO4	D	706	-	4,4,4	0.33	0	6,6,6	0.21	0
2	X1A	F	701	-	23,32,32	1.53	2 (8%)	18,48,48	2.93	10 (55%)
4	SO4	C	707	-	4,4,4	0.31	0	6,6,6	0.10	0
2	X1A	B	701	-	23,32,32	1.43	2 (8%)	18,48,48	2.68	10 (55%)
4	SO4	A	703	-	4,4,4	0.42	0	6,6,6	0.13	0
4	SO4	A	705	-	4,4,4	0.32	0	6,6,6	0.25	0
4	SO4	B	706	-	4,4,4	0.39	0	6,6,6	0.27	0
2	X1A	D	701	-	23,32,32	1.52	2 (8%)	18,48,48	2.74	10 (55%)
4	SO4	B	704	-	4,4,4	0.40	0	6,6,6	0.10	0
4	SO4	D	709	-	4,4,4	0.35	0	6,6,6	0.16	0
4	SO4	B	708	-	4,4,4	0.34	0	6,6,6	0.29	0
4	SO4	C	706	-	4,4,4	0.32	0	6,6,6	0.24	0
4	SO4	D	708	-	4,4,4	0.30	0	6,6,6	0.22	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	SO4	C	708	-	4,4,4	0.39	0	6,6,6	0.16	0
4	SO4	D	705	-	4,4,4	0.32	0	6,6,6	0.14	0
2	X1A	E	701	-	23,32,32	1.39	1 (4%)	18,48,48	2.99	9 (50%)
4	SO4	D	711	-	4,4,4	0.37	0	6,6,6	0.10	0
4	SO4	A	707	-	4,4,4	0.40	0	6,6,6	0.08	0
4	SO4	C	704	-	4,4,4	0.32	0	6,6,6	0.28	0
4	SO4	D	707	-	4,4,4	0.31	0	6,6,6	0.26	0
4	SO4	E	706	-	4,4,4	0.36	0	6,6,6	0.14	0
4	SO4	F	705	-	4,4,4	0.32	0	6,6,6	0.18	0
4	SO4	E	705	-	4,4,4	0.34	0	6,6,6	0.18	0
4	SO4	F	703	-	4,4,4	0.37	0	6,6,6	0.09	0
4	SO4	E	703	-	4,4,4	0.37	0	6,6,6	0.14	0
4	SO4	C	705	-	4,4,4	0.32	0	6,6,6	0.21	0
4	SO4	D	710	-	4,4,4	0.38	0	6,6,6	0.19	0
4	SO4	F	706	-	4,4,4	0.30	0	6,6,6	0.18	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
2	X1A	C	701	-	-	1/9/33/33	0/5/5/5
2	X1A	E	701	-	-	1/9/33/33	0/5/5/5
2	X1A	A	701	-	-	1/9/33/33	0/5/5/5
2	X1A	F	701	-	-	2/9/33/33	0/5/5/5
2	X1A	B	701	-	-	1/9/33/33	0/5/5/5
2	X1A	D	701	-	-	1/9/33/33	0/5/5/5

All (13) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	701	X1A	C27-C12	-5.29	1.49	1.54
2	F	701	X1A	C27-C12	-5.14	1.49	1.54
2	D	701	X1A	C27-C12	-5.06	1.49	1.54
2	B	701	X1A	C27-C12	-4.98	1.49	1.54
2	E	701	X1A	C27-C12	-4.86	1.50	1.54
2	C	701	X1A	C27-C12	-4.55	1.50	1.54
2	B	701	X1A	C7-N6	-2.75	1.32	1.36
2	C	701	X1A	C7-N6	-2.44	1.32	1.36
2	D	701	X1A	C7-N6	-2.38	1.32	1.36
2	A	701	X1A	C7-N6	-2.35	1.32	1.36

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	701	X1A	C13-C14	2.14	1.56	1.52
2	C	701	X1A	C13-C14	2.08	1.56	1.52
2	A	701	X1A	C13-C14	2.06	1.56	1.52

All (60) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	701	X1A	C14-N15-C16	-5.83	104.01	112.36
2	C	701	X1A	C14-N15-C16	-5.73	104.16	112.36
2	E	701	X1A	C14-N15-C16	-5.45	104.56	112.36
2	D	701	X1A	C14-N15-C16	-5.37	104.67	112.36
2	A	701	X1A	C14-N15-C16	-5.29	104.78	112.36
2	A	701	X1A	C16-N15-C17	5.05	129.78	123.60
2	E	701	X1A	C16-N15-C17	4.98	129.70	123.60
2	B	701	X1A	C14-N15-C16	-4.84	105.42	112.36
2	B	701	X1A	C16-N15-C17	4.80	129.47	123.60
2	C	701	X1A	C14-C13-C12	-4.71	96.99	104.25
2	F	701	X1A	C7-C8-S28	-4.66	107.37	112.00
2	A	701	X1A	C14-C13-C12	-4.45	97.39	104.25
2	D	701	X1A	C16-N15-C17	4.43	129.02	123.60
2	C	701	X1A	C7-C8-S28	-4.42	107.61	112.00
2	C	701	X1A	C17-C18-C19	4.35	119.54	116.91
2	E	701	X1A	C7-C8-S28	-4.34	107.68	112.00
2	C	701	X1A	C16-N15-C17	4.25	128.80	123.60
2	A	701	X1A	C7-C8-S28	-4.16	107.86	112.00
2	E	701	X1A	C14-C13-C12	-4.09	97.95	104.25
2	D	701	X1A	C7-C8-S28	-4.06	107.97	112.00
2	D	701	X1A	C14-C13-C12	-4.01	98.07	104.25
2	F	701	X1A	C14-C13-C12	-3.97	98.13	104.25
2	C	701	X1A	C13-C14-N15	3.94	106.99	103.61
2	E	701	X1A	C17-C18-C19	3.91	119.27	116.91
2	F	701	X1A	C16-N15-C17	3.86	128.32	123.60
2	E	701	X1A	C13-C14-N15	3.77	106.85	103.61
2	B	701	X1A	C14-C13-C12	-3.72	98.51	104.25
2	F	701	X1A	C13-C14-N15	3.71	106.79	103.61
2	E	701	X1A	C3-C2-N4	3.53	120.16	114.98
2	B	701	X1A	C13-C14-N15	3.35	106.49	103.61
2	F	701	X1A	C3-C2-N4	3.35	119.90	114.98
2	F	701	X1A	C14-N15-C17	3.33	127.67	123.60
2	E	701	X1A	O1-C2-C3	-3.31	115.92	122.06
2	A	701	X1A	C13-C14-N15	3.30	106.44	103.61
2	F	701	X1A	C17-C18-C19	3.28	118.89	116.91

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	701	X1A	C13-C14-N15	3.25	106.40	103.61
2	B	701	X1A	C7-C8-S28	-3.15	108.86	112.00
2	C	701	X1A	C3-C2-N4	3.15	119.61	114.98
2	B	701	X1A	C20-C19-C18	-3.14	117.02	121.81
2	D	701	X1A	C17-C18-C19	3.11	118.79	116.91
2	F	701	X1A	O1-C2-C3	-3.01	116.46	122.06
2	B	701	X1A	O1-C2-C3	-3.00	116.48	122.06
2	B	701	X1A	C3-C2-N4	2.93	119.28	114.98
2	E	701	X1A	C19-N21-C22	2.92	120.73	117.65
2	D	701	X1A	O1-C2-C3	-2.89	116.69	122.06
2	C	701	X1A	C14-N15-C17	2.73	126.94	123.60
2	A	701	X1A	C3-C2-N4	2.68	118.92	114.98
2	B	701	X1A	C17-C18-C19	2.67	118.53	116.91
2	D	701	X1A	C3-C2-N4	2.66	118.88	114.98
2	C	701	X1A	O1-C2-C3	-2.58	117.27	122.06
2	F	701	X1A	C19-N21-C22	2.52	120.31	117.65
2	D	701	X1A	C19-N21-C22	2.30	120.08	117.65
2	A	701	X1A	C17-C18-C19	2.24	118.26	116.91
2	C	701	X1A	C20-C19-C18	-2.23	118.40	121.81
2	A	701	X1A	C19-N21-C22	2.18	119.95	117.65
2	D	701	X1A	C14-N15-C17	2.18	126.27	123.60
2	B	701	X1A	C19-N21-C22	2.15	119.92	117.65
2	A	701	X1A	O1-C2-C3	-2.11	118.13	122.06
2	A	701	X1A	C20-C19-C18	-2.06	118.67	121.81
2	C	701	X1A	C19-N21-C22	2.01	119.77	117.65

There are no chirality outliers.

All (7) torsion outliers are listed below:

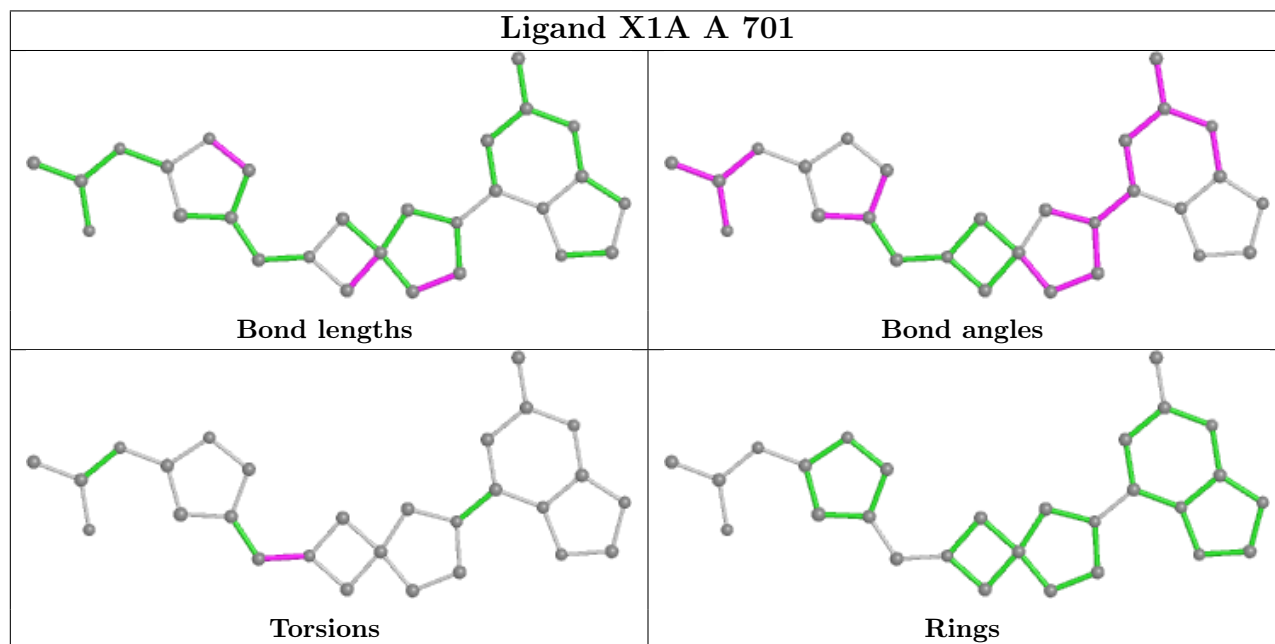
Mol	Chain	Res	Type	Atoms
2	A	701	X1A	C8-C9-N10-C27
2	B	701	X1A	C8-C9-N10-C27
2	C	701	X1A	C8-C9-N10-C27
2	D	701	X1A	C8-C9-N10-C27
2	E	701	X1A	C8-C9-N10-C27
2	F	701	X1A	C8-C9-N10-C11
2	F	701	X1A	C8-C9-N10-C27

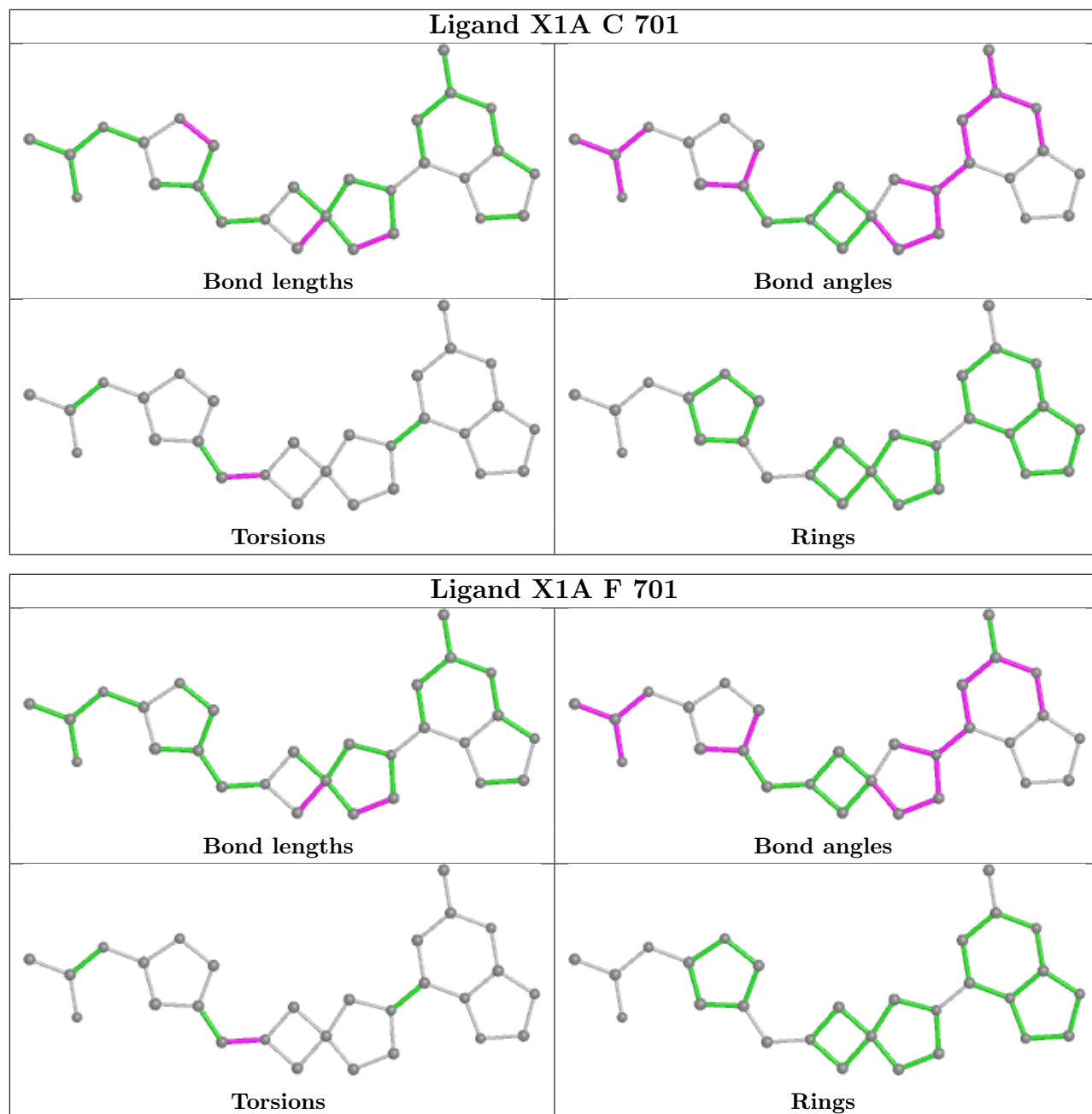
There are no ring outliers.

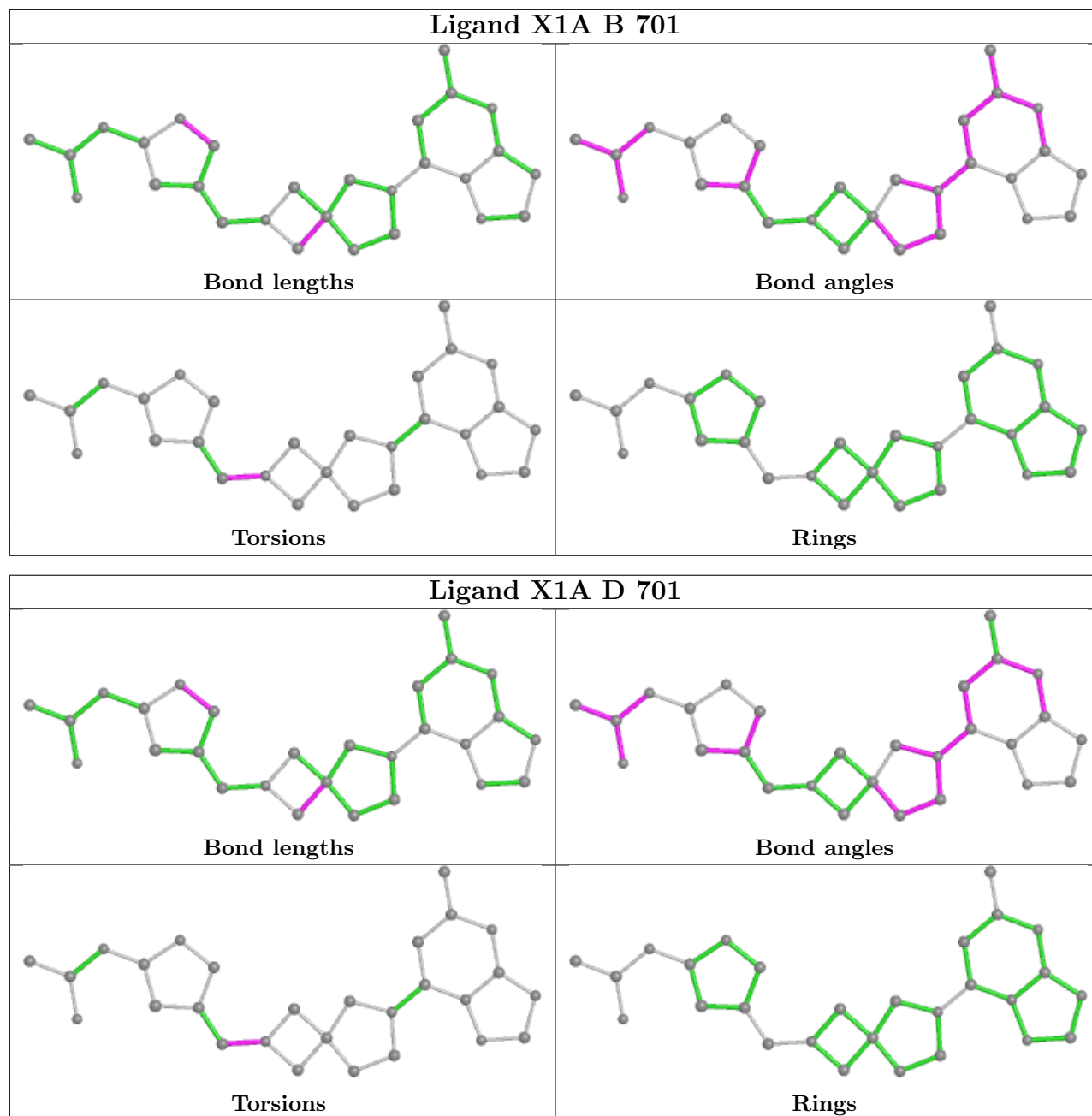
4 monomers are involved in 4 short contacts:

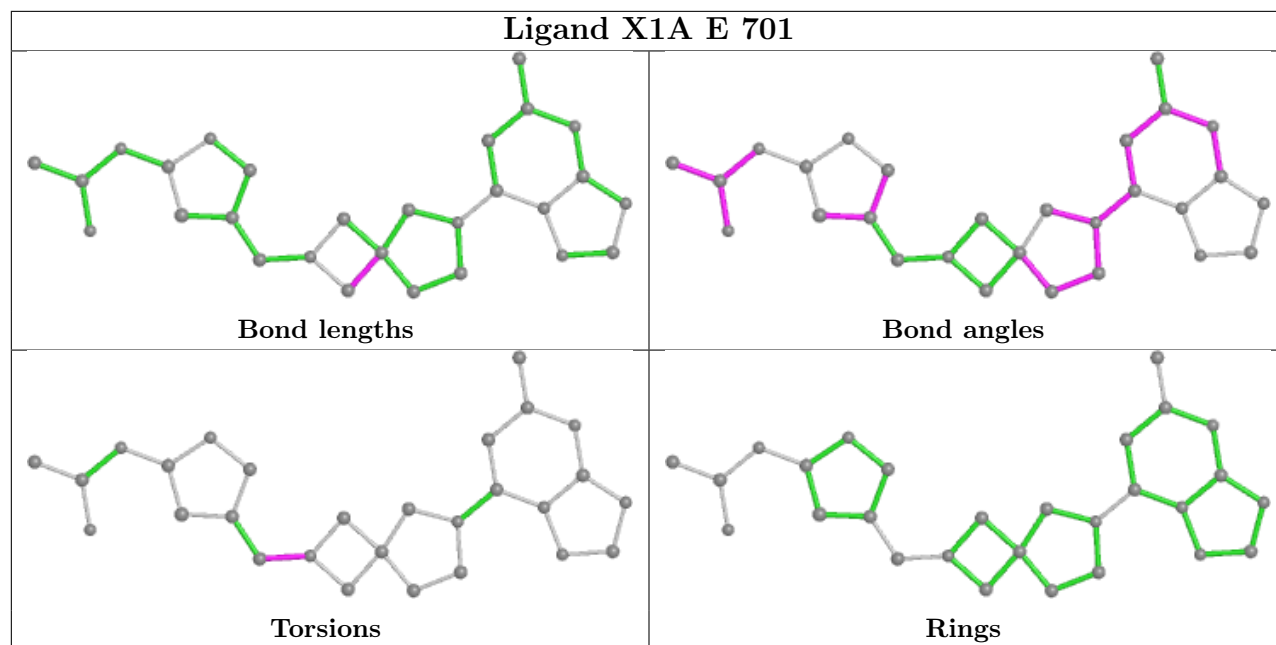
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	701	X1A	1	0
2	C	701	X1A	1	0
2	F	701	X1A	1	0
2	E	701	X1A	1	0

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 i

### 6.1 Protein, DNA and RNA chains i

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	586/594 (98%)	-0.21	2 (0%) 94   96	17, 27, 51, 81	6 (1%)
1	B	585/594 (98%)	-0.14	0 100   100	14, 24, 46, 73	4 (0%)
1	C	585/594 (98%)	0.04	8 (1%) 75   80	15, 28, 50, 77	9 (1%)
1	D	585/594 (98%)	0.12	11 (1%) 66   73	22, 34, 57, 88	7 (1%)
1	E	586/594 (98%)	0.32	23 (3%) 39   46	26, 42, 64, 83	9 (1%)
1	F	539/594 (90%)	0.53	64 (11%) 4   6	21, 44, 87, 115	21 (3%)
All	All	3466/3564 (97%)	0.10	108 (3%) 49   56	14, 33, 63, 115	56 (1%)

All (108) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	55	VAL	6.4
1	F	95	SER	5.5
1	F	135	LEU	4.8
1	F	93	ILE	4.8
1	F	56	VAL	4.8
1	F	173	ILE	4.7
1	F	134	ALA	4.7
1	F	94	ASN	4.5
1	F	91	ILE	4.5
1	F	163	VAL	4.5
1	F	98	ASP	4.5
1	F	87	THR	4.4
1	F	162	LEU	4.3
1	F	589	ILE	4.3
1	F	43	VAL	4.2
1	F	86	LEU	4.1
1	F	174	THR	4.1
1	D	56	VAL	4.1
1	F	88	ALA	4.0

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	F	82	LEU	4.0
1	F	343	ASN	3.9
1	F	79	VAL	3.8
1	E	93	ILE	3.8
1	E	90	ASN	3.8
1	F	85	PHE	3.8
1	F	71	GLU	3.7
1	F	92	GLU	3.7
1	E	88	ALA	3.7
1	D	40	GLN	3.6
1	F	70	GLY	3.6
1	F	159	PHE	3.4
1	F	580	THR	3.4
1	F	171	VAL	3.4
1	F	99	PRO	3.3
1	E	66	ILE	3.2
1	F	90	ASN	3.1
1	E	91	ILE	3.1
1	F	89	ASN	3.1
1	C	112	ASP	3.1
1	F	172	ASN	3.1
1	F	585	SER	3.1
1	F	348	ALA	3.0
1	D	64	SER	3.0
1	F	72	GLU	3.0
1	F	127	ASN	2.9
1	F	54	GLU	2.9
1	D	343	ASN	2.9
1	E	165	GLU	2.9
1	E	64	SER	2.9
1	F	542	ASN	2.8
1	F	83	ARG	2.8
1	F	538	ASN	2.7
1	E	62	ILE	2.7
1	E	580	THR	2.7
1	C	40	GLN	2.7
1	F	345	GLN	2.6
1	E	321	ALA	2.6
1	C	110	ASP	2.6
1	F	109	VAL	2.6
1	F	73	GLU	2.6
1	F	166	SER	2.6

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	F	548	LEU	2.6
1	F	144	ILE	2.6
1	E	87	THR	2.5
1	C	111	ASP	2.5
1	A	100	ASN	2.5
1	F	78	ALA	2.5
1	E	102	THR	2.5
1	E	100	ASN	2.5
1	F	581	GLU	2.5
1	E	237	GLN	2.5
1	E	541	ALA	2.5
1	F	80	ASN	2.5
1	F	541	ALA	2.4
1	D	100	ASN	2.4
1	E	142	ILE	2.4
1	F	46	LEU	2.4
1	F	582	ALA	2.4
1	F	110	ASP	2.4
1	F	164	LYS	2.3
1	F	124	THR	2.3
1	A	90	ASN	2.3
1	C	397	TYR	2.3
1	F	126	GLU	2.3
1	F	588	GLU	2.3
1	C	268	ASP	2.3
1	D	99	PRO	2.2
1	D	251	VAL	2.2
1	E	278	ILE	2.2
1	C	232	PRO	2.2
1	E	382	LEU	2.1
1	D	351	ARG	2.1
1	F	125	ALA	2.1
1	E	139	ASP	2.1
1	F	470	ASP	2.1
1	F	74	ALA	2.1
1	F	546	VAL	2.1
1	D	62	ILE	2.1
1	E	65	SER	2.1
1	D	42	LEU	2.0
1	C	485	MET	2.0
1	E	92	GLU	2.0
1	E	573	VAL	2.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	F	381	GLN	2.0
1	F	543	LEU	2.0
1	D	138	ASN	2.0
1	E	344	GLY	2.0
1	F	586	ALA	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 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
3	CA	D	702	1/1	0.83	0.11	65,65,65,65	0
4	SO4	E	703	5/5	0.83	0.27	88,91,105,105	0
4	SO4	C	704	5/5	0.86	0.19	70,74,83,84	0
4	SO4	A	703	5/5	0.87	0.21	64,68,105,109	0
4	SO4	E	704	5/5	0.87	0.21	71,89,114,115	0
4	SO4	C	703	5/5	0.88	0.22	65,83,103,108	0
3	CA	A	702	1/1	0.88	0.08	57,57,57,57	0
3	CA	E	702	1/1	0.90	0.14	70,70,70,70	0
4	SO4	F	703	5/5	0.91	0.27	72,75,88,92	0
4	SO4	A	707	5/5	0.93	0.18	47,68,77,77	0
4	SO4	B	704	5/5	0.93	0.20	59,65,90,91	0
4	SO4	D	706	5/5	0.93	0.19	75,94,114,115	0
4	SO4	D	704	5/5	0.94	0.24	75,87,91,94	0
4	SO4	A	704	5/5	0.94	0.14	48,62,64,72	0
4	SO4	B	707	5/5	0.94	0.20	45,63,65,73	0
4	SO4	C	708	5/5	0.94	0.17	52,63,69,94	0
4	SO4	E	705	5/5	0.94	0.14	64,72,87,89	0
4	SO4	C	709	5/5	0.94	0.19	61,73,82,84	0

*Continued on next page...*

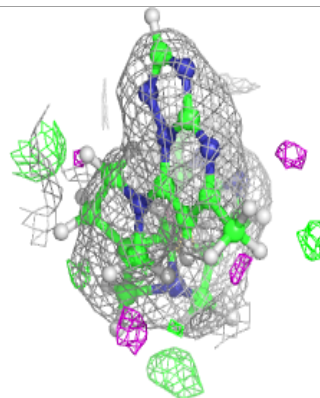
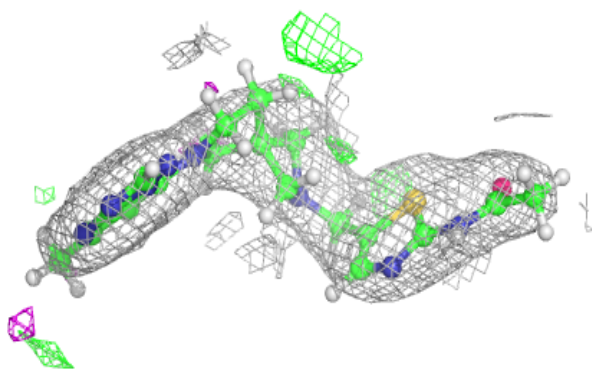
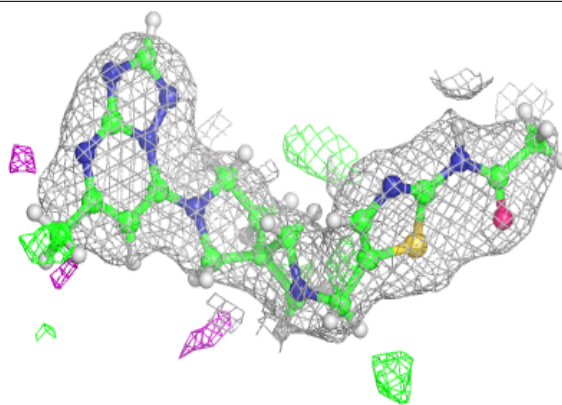
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	SO4	F	707	5/5	0.94	0.24	69,73,81,82	0
4	SO4	D	710	5/5	0.95	0.22	56,62,74,78	0
4	SO4	C	705	5/5	0.95	0.17	58,63,75,78	0
4	SO4	B	706	5/5	0.95	0.13	44,47,67,71	0
2	X1A	F	701	28/28	0.95	0.16	29,38,46,48	0
4	SO4	E	706	5/5	0.95	0.12	56,63,67,80	0
2	X1A	E	701	28/28	0.95	0.16	24,34,41,42	0
4	SO4	B	705	5/5	0.95	0.16	62,64,65,74	0
2	X1A	B	701	28/28	0.96	0.16	17,24,30,32	0
4	SO4	C	706	5/5	0.96	0.20	41,45,45,52	5
2	X1A	C	701	28/28	0.96	0.19	21,30,39,42	0
4	SO4	A	705	5/5	0.96	0.14	65,66,74,76	0
2	X1A	D	701	28/28	0.96	0.13	23,29,34,37	0
2	X1A	A	701	28/28	0.96	0.14	15,25,33,35	0
4	SO4	F	704	5/5	0.96	0.21	65,65,78,88	0
4	SO4	F	705	5/5	0.96	0.12	52,55,69,71	0
4	SO4	F	706	5/5	0.96	0.16	62,65,67,84	0
4	SO4	D	708	5/5	0.96	0.10	40,48,59,60	0
4	SO4	D	709	5/5	0.97	0.10	52,59,63,64	0
4	SO4	B	708	5/5	0.97	0.10	36,43,46,47	0
4	SO4	D	711	5/5	0.97	0.13	52,60,76,80	0
5	CL	F	702	1/1	0.97	0.11	46,46,46,46	0
4	SO4	D	707	5/5	0.98	0.14	34,41,43,61	0
4	SO4	A	706	5/5	0.98	0.12	41,44,46,57	0
4	SO4	C	707	5/5	0.98	0.13	45,52,62,65	0
4	SO4	D	705	5/5	0.98	0.19	68,73,83,83	0
5	CL	C	702	1/1	0.98	0.19	39,39,39,39	0
5	CL	D	703	1/1	0.98	0.07	59,59,59,59	0
4	SO4	A	708	5/5	0.98	0.10	42,46,50,54	0
5	CL	B	703	1/1	0.99	0.08	46,46,46,46	0
5	CL	B	702	1/1	1.00	0.13	34,34,34,34	0

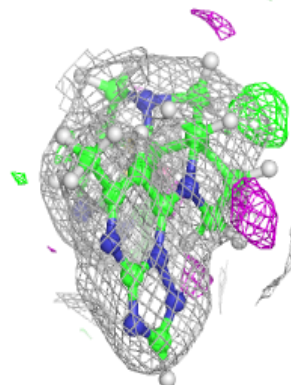
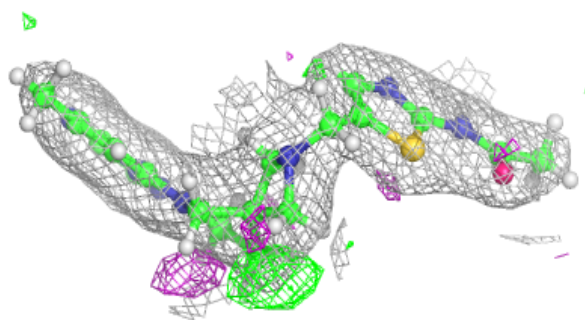
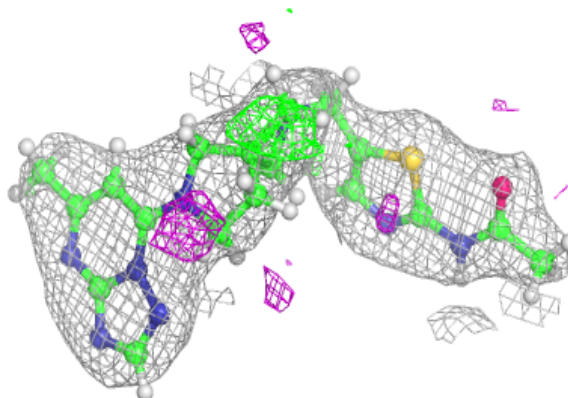
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around X1A F 701:**

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

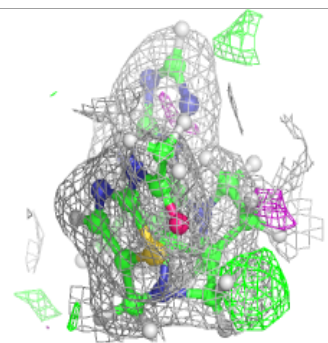
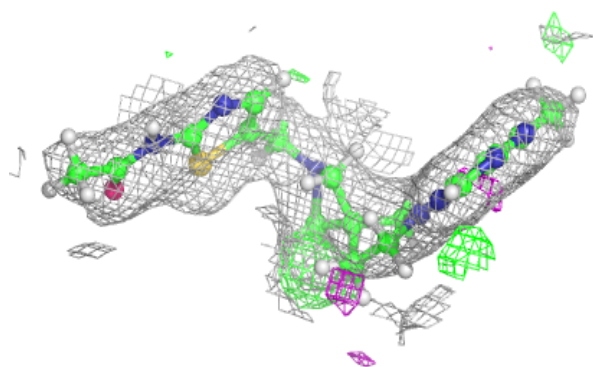
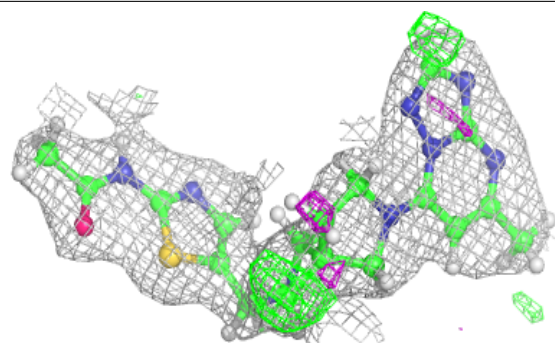
**Electron density around X1A E 701:**

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

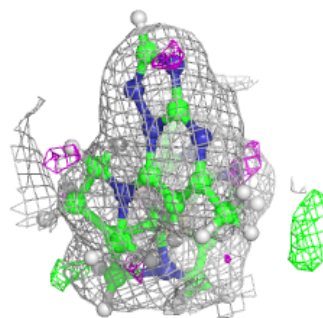
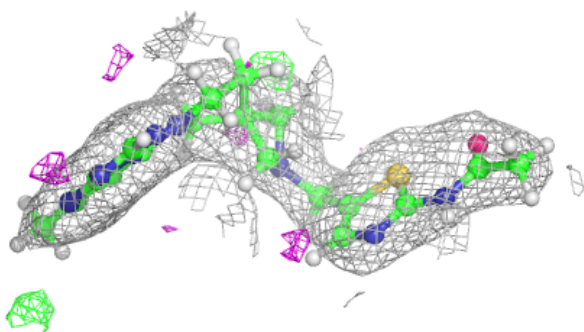
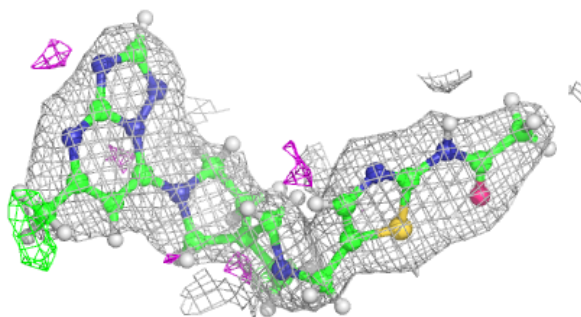


**Electron density around X1A B 701:**

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

**Electron density around X1A C 701:**

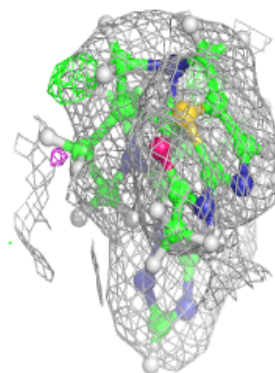
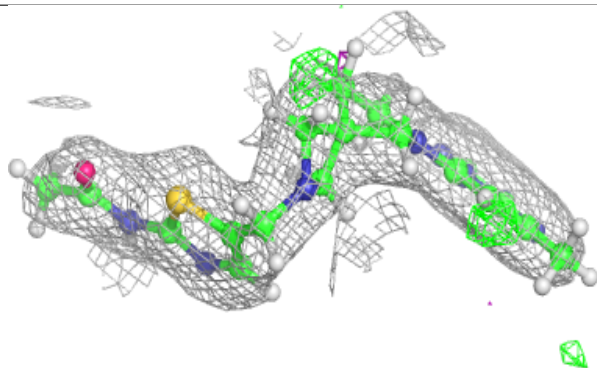
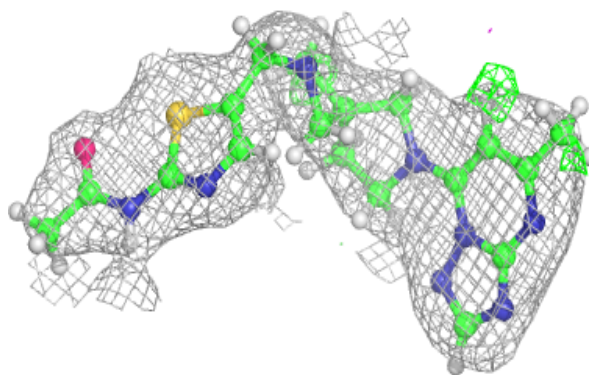
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



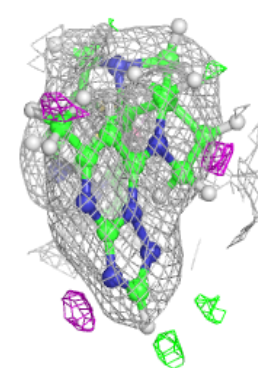
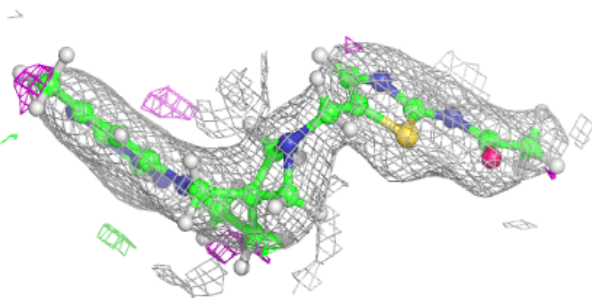
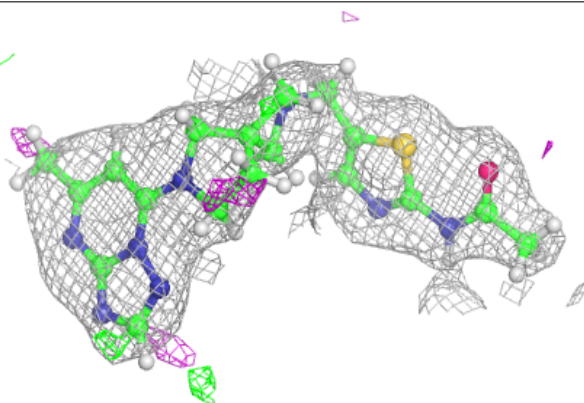


**Electron density around X1A D 701:**

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

**Electron density around X1A A 701:**

$2mF_o-DF_c$  (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.