



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 27, 2022 – 01:13 pm BST

PDB ID : 7AYY  
Title : Structure of the human 8-oxoguanine DNA Glycosylase hOGG1 in complex with activator TH10785  
Authors : Masuyer, G.; Davies, J.R.; Stenmark, P.  
Deposited on : 2020-11-13  
Resolution : 2.00 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 : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.29  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0267  
CCP4 : 7.1.010 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.29

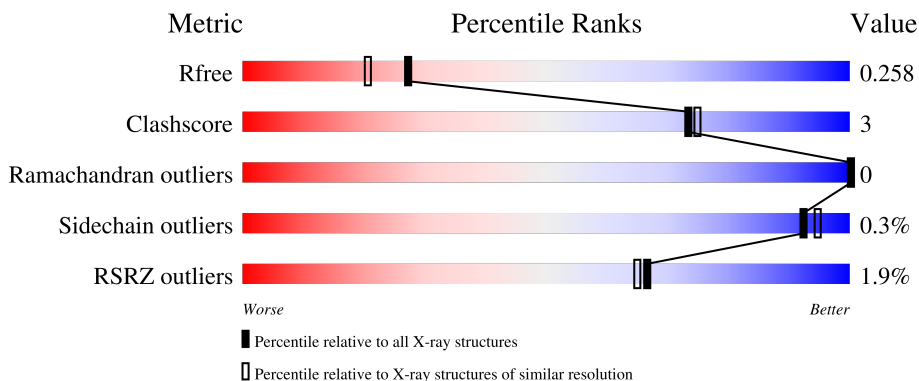
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	AAA	337	 3% 81% 10% 9%
1	BBB	337	 2% 85% 6% 9%
1	CCC	337	 2% 86% 6% 8%
1	DDD	337	 0% 85% 5% 10%
1	EEE	337	 0% 82% 8% 10%

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 13010 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 N-glycosylase/DNA lyase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	AAA	305	Total 2431	C 1546	N 436	O 437	S 12	0	0	0
1	BBB	307	Total 2439	C 1550	N 438	O 439	S 12	0	0	0
1	CCC	310	Total 2462	C 1565	N 443	O 442	S 12	0	0	0
1	DDD	304	Total 2415	C 1535	N 433	O 435	S 12	0	0	0
1	EEE	303	Total 2411	C 1533	N 432	O 434	S 12	0	0	0

There are 100 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	-9	MET	-	initiating methionine	UNP O15527
AAA	-8	GLY	-	expression tag	UNP O15527
AAA	-7	SER	-	expression tag	UNP O15527
AAA	-6	SER	-	expression tag	UNP O15527
AAA	-5	HIS	-	expression tag	UNP O15527
AAA	-4	HIS	-	expression tag	UNP O15527
AAA	-3	HIS	-	expression tag	UNP O15527
AAA	-2	HIS	-	expression tag	UNP O15527
AAA	-1	HIS	-	expression tag	UNP O15527
AAA	0	HIS	-	expression tag	UNP O15527
AAA	1	SER	-	expression tag	UNP O15527
AAA	2	SER	-	expression tag	UNP O15527
AAA	3	GLY	-	expression tag	UNP O15527
AAA	4	LEU	-	expression tag	UNP O15527
AAA	5	VAL	-	expression tag	UNP O15527
AAA	6	PRO	-	expression tag	UNP O15527
AAA	7	ARG	-	expression tag	UNP O15527
AAA	8	GLY	-	expression tag	UNP O15527
AAA	9	SER	-	expression tag	UNP O15527

*Continued on next page...*

*Continued from previous page...*

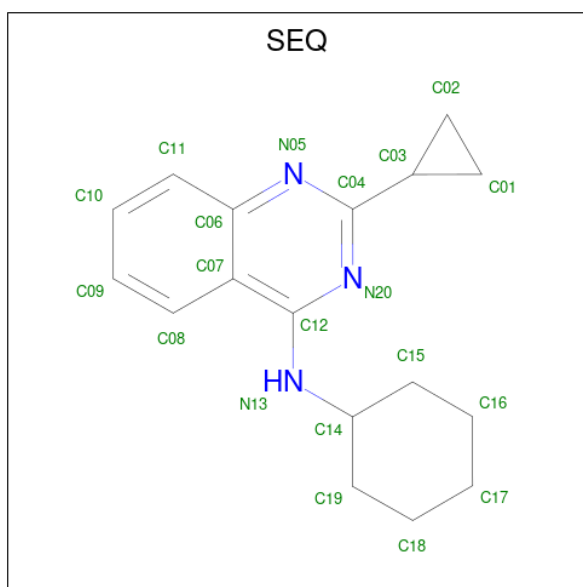
Chain	Residue	Modelled	Actual	Comment	Reference
AAA	10	HIS	-	expression tag	UNP O15527
BBB	-9	MET	-	initiating methionine	UNP O15527
BBB	-8	GLY	-	expression tag	UNP O15527
BBB	-7	SER	-	expression tag	UNP O15527
BBB	-6	SER	-	expression tag	UNP O15527
BBB	-5	HIS	-	expression tag	UNP O15527
BBB	-4	HIS	-	expression tag	UNP O15527
BBB	-3	HIS	-	expression tag	UNP O15527
BBB	-2	HIS	-	expression tag	UNP O15527
BBB	-1	HIS	-	expression tag	UNP O15527
BBB	0	HIS	-	expression tag	UNP O15527
BBB	1	SER	-	expression tag	UNP O15527
BBB	2	SER	-	expression tag	UNP O15527
BBB	3	GLY	-	expression tag	UNP O15527
BBB	4	LEU	-	expression tag	UNP O15527
BBB	5	VAL	-	expression tag	UNP O15527
BBB	6	PRO	-	expression tag	UNP O15527
BBB	7	ARG	-	expression tag	UNP O15527
BBB	8	GLY	-	expression tag	UNP O15527
BBB	9	SER	-	expression tag	UNP O15527
BBB	10	HIS	-	expression tag	UNP O15527
CCC	-9	MET	-	initiating methionine	UNP O15527
CCC	-8	GLY	-	expression tag	UNP O15527
CCC	-7	SER	-	expression tag	UNP O15527
CCC	-6	SER	-	expression tag	UNP O15527
CCC	-5	HIS	-	expression tag	UNP O15527
CCC	-4	HIS	-	expression tag	UNP O15527
CCC	-3	HIS	-	expression tag	UNP O15527
CCC	-2	HIS	-	expression tag	UNP O15527
CCC	-1	HIS	-	expression tag	UNP O15527
CCC	0	HIS	-	expression tag	UNP O15527
CCC	1	SER	-	expression tag	UNP O15527
CCC	2	SER	-	expression tag	UNP O15527
CCC	3	GLY	-	expression tag	UNP O15527
CCC	4	LEU	-	expression tag	UNP O15527
CCC	5	VAL	-	expression tag	UNP O15527
CCC	6	PRO	-	expression tag	UNP O15527
CCC	7	ARG	-	expression tag	UNP O15527
CCC	8	GLY	-	expression tag	UNP O15527
CCC	9	SER	-	expression tag	UNP O15527
CCC	10	HIS	-	expression tag	UNP O15527
DDD	-9	MET	-	initiating methionine	UNP O15527

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
DDD	-8	GLY	-	expression tag	UNP O15527
DDD	-7	SER	-	expression tag	UNP O15527
DDD	-6	SER	-	expression tag	UNP O15527
DDD	-5	HIS	-	expression tag	UNP O15527
DDD	-4	HIS	-	expression tag	UNP O15527
DDD	-3	HIS	-	expression tag	UNP O15527
DDD	-2	HIS	-	expression tag	UNP O15527
DDD	-1	HIS	-	expression tag	UNP O15527
DDD	0	HIS	-	expression tag	UNP O15527
DDD	1	SER	-	expression tag	UNP O15527
DDD	2	SER	-	expression tag	UNP O15527
DDD	3	GLY	-	expression tag	UNP O15527
DDD	4	LEU	-	expression tag	UNP O15527
DDD	5	VAL	-	expression tag	UNP O15527
DDD	6	PRO	-	expression tag	UNP O15527
DDD	7	ARG	-	expression tag	UNP O15527
DDD	8	GLY	-	expression tag	UNP O15527
DDD	9	SER	-	expression tag	UNP O15527
DDD	10	HIS	-	expression tag	UNP O15527
EEE	-9	MET	-	initiating methionine	UNP O15527
EEE	-8	GLY	-	expression tag	UNP O15527
EEE	-7	SER	-	expression tag	UNP O15527
EEE	-6	SER	-	expression tag	UNP O15527
EEE	-5	HIS	-	expression tag	UNP O15527
EEE	-4	HIS	-	expression tag	UNP O15527
EEE	-3	HIS	-	expression tag	UNP O15527
EEE	-2	HIS	-	expression tag	UNP O15527
EEE	-1	HIS	-	expression tag	UNP O15527
EEE	0	HIS	-	expression tag	UNP O15527
EEE	1	SER	-	expression tag	UNP O15527
EEE	2	SER	-	expression tag	UNP O15527
EEE	3	GLY	-	expression tag	UNP O15527
EEE	4	LEU	-	expression tag	UNP O15527
EEE	5	VAL	-	expression tag	UNP O15527
EEE	6	PRO	-	expression tag	UNP O15527
EEE	7	ARG	-	expression tag	UNP O15527
EEE	8	GLY	-	expression tag	UNP O15527
EEE	9	SER	-	expression tag	UNP O15527
EEE	10	HIS	-	expression tag	UNP O15527

- Molecule 2 is {N}-cyclohexyl-2-cyclopropyl-quinazolin-4-amine (three-letter code: SEQ) (formula: C<sub>17</sub>H<sub>21</sub>N<sub>3</sub>) (labeled as "Ligand of Interest" by depositor).



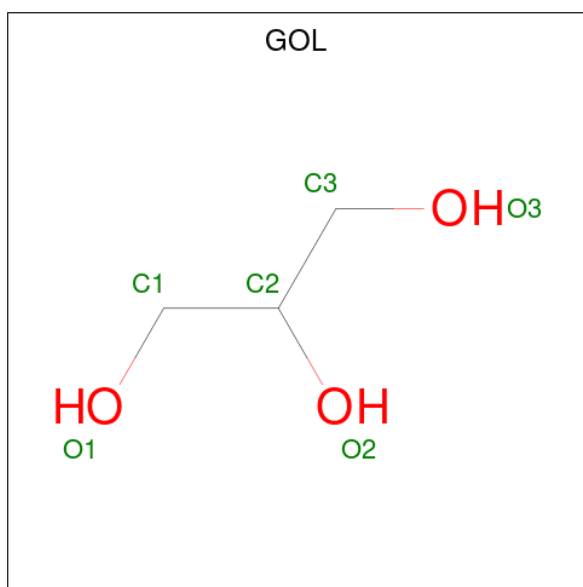
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	N		
2	AAA	1	20	17	3	0	0
2	BBB	1	20	17	3	0	0
2	CCC	1	20	17	3	0	0
2	DDD	1	20	17	3	0	0
2	EEE	1	20	17	3	0	0

- Molecule 3 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: C<sub>6</sub>H<sub>13</sub>NO<sub>4</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
3	AAA	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
3	BBB	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
3	EEE	1	Total	C	N	O	S	0	0
			12	6	1	4	1		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
4	DDD	1	Total	C	O	0	0
			6	3	3		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	EEE	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is water.

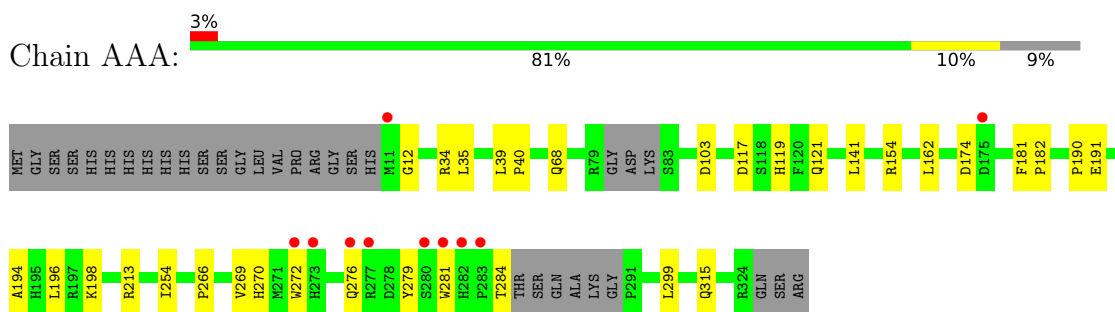
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	AAA	124	Total	O	0	0
			124	124		
5	BBB	125	Total	O	0	0
			125	125		
5	CCC	143	Total	O	0	0
			143	143		
5	DDD	168	Total	O	0	0
			168	168		
5	EEE	144	Total	O	0	0
			144	144		



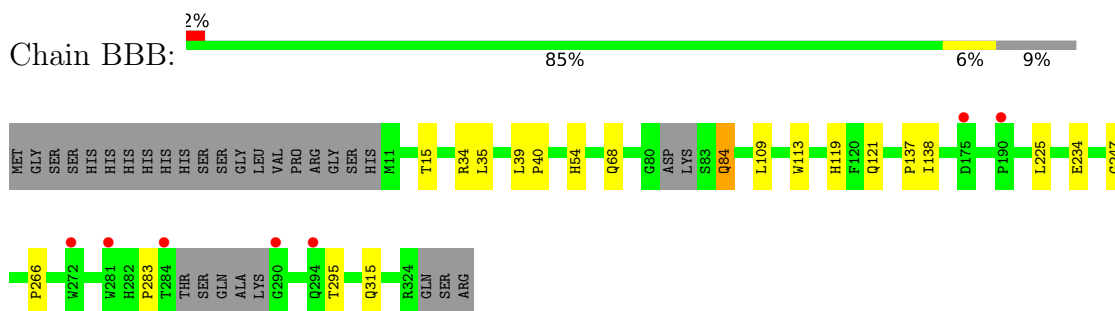
### 3 Residue-property plots [i](#)

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.

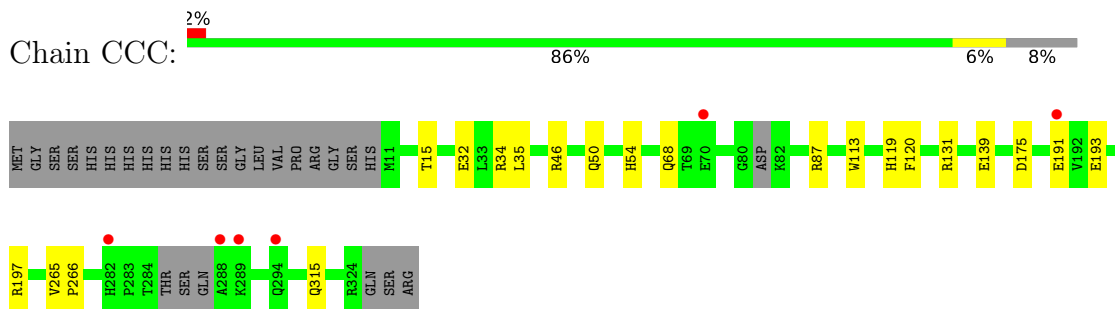
- Molecule 1: N-glycosylase/DNA lyase



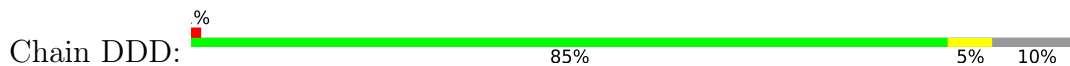
- Molecule 1: N-glycosylase/DNA lyase

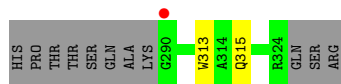
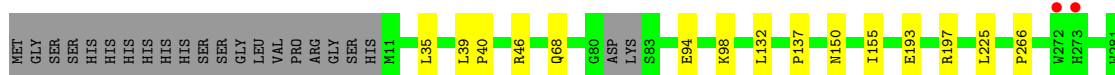


- Molecule 1: N-glycosylase/DNA lyase

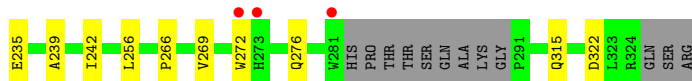
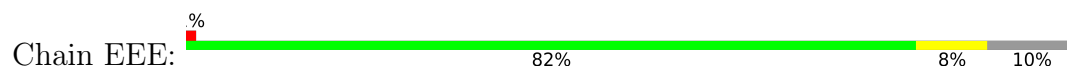


- Molecule 1: N-glycosylase/DNA lyase





● Molecule 1: N-glycosylase/DNA lyase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	86.03Å 86.03Å 427.60Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	84.48 – 2.00 84.34 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.8 (84.48-2.00) 99.8 (84.34-2.00)	Depositor EDS
$R_{merge}$	0.30	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.75 (at 2.00Å)	Xtrriage
Refinement program	REFMAC 5.8.0258	Depositor
R, $R_{free}$	0.207 , 0.250 0.216 , 0.258	Depositor DCC
$R_{free}$ test set	5445 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.7	Xtrriage
Anisotropy	0.155	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	13010	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 55.90 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.9594e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: GOL, SEQ, MES

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	AAA	0.66	0/2498	0.77	0/3398
1	BBB	0.66	0/2506	0.77	0/3409
1	CCC	0.65	0/2529	0.78	0/3438
1	DDD	0.66	0/2480	0.76	1/3372 (0.0%)
1	EEE	0.67	0/2476	0.76	0/3366
All	All	0.66	0/12489	0.77	1/16983 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	DDD	46	ARG	NE-CZ-NH2	-5.04	117.78	120.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	AAA	2431	0	2370	21	0
1	BBB	2439	0	2375	17	0
1	CCC	2462	0	2406	15	0
1	DDD	2415	0	2354	9	0
1	EEE	2411	0	2352	19	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	AAA	20	0	0	0	0
2	BBB	20	0	0	0	0
2	CCC	20	0	0	0	0
2	DDD	20	0	0	0	0
2	EEE	20	0	0	0	0
3	AAA	12	0	13	0	0
3	BBB	12	0	13	1	0
3	EEE	12	0	13	0	0
4	DDD	6	0	8	0	0
4	EEE	6	0	8	0	0
5	AAA	124	0	0	5	0
5	BBB	125	0	0	4	0
5	CCC	143	0	0	3	0
5	DDD	168	0	0	0	0
5	EEE	144	0	0	5	0
All	All	13010	0	11912	79	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 3.

All (79) 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:BBB:121:GLN:HG2	5:BBB:609:HOH:O	1.80	0.81
1:CCC:197:ARG:NH1	5:CCC:501:HOH:O	2.17	0.76
1:CCC:119:HIS:HE1	5:CCC:532:HOH:O	1.74	0.70
1:EEE:235:GLU:OE2	5:EEE:501:HOH:O	2.10	0.69
1:DDD:266:PRO:HD2	1:DDD:315:GLN:HE22	1.56	0.69
1:AAA:119:HIS:HE1	5:AAA:601:HOH:O	1.78	0.66
1:BBB:84:GLN:NE2	1:CCC:87:ARG:O	2.30	0.64
1:EEE:35:LEU:H	1:EEE:68:GLN:HE22	1.47	0.62
1:AAA:284:THR:HG21	1:EEE:27:PRO:HG2	1.81	0.62
1:BBB:84:GLN:HB2	5:BBB:613:HOH:O	2.01	0.60
1:AAA:34:ARG:HA	1:AAA:68:GLN:HE22	1.67	0.60
1:AAA:117:ASP:OD2	1:AAA:279:TYR:OH	2.17	0.60
1:DDD:35:LEU:H	1:DDD:68:GLN:HE22	1.50	0.59
1:EEE:35:LEU:H	1:EEE:68:GLN:NE2	2.02	0.58
1:CCC:32:GLU:O	1:CCC:131:ARG:HD2	2.05	0.57
1:BBB:119:HIS:HE1	5:BBB:519:HOH:O	1.88	0.56
1:EEE:128:GLN:NE2	5:EEE:506:HOH:O	2.39	0.56
1:BBB:266:PRO:HD2	1:BBB:315:GLN:HE22	1.72	0.55

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:AAA:103:ASP:HB2	5:AAA:595:HOH:O	2.07	0.54
1:CCC:175:ASP:OD1	1:CCC:175:ASP:N	2.40	0.54
1:EEE:239:ALA:O	1:EEE:242:ILE:HG12	2.07	0.54
1:EEE:266:PRO:HD2	1:EEE:315:GLN:HE22	1.72	0.53
1:DDD:35:LEU:H	1:DDD:68:GLN:NE2	2.07	0.53
1:BBB:137:PRO:HB2	1:BBB:225:LEU:HD13	1.91	0.53
1:BBB:35:LEU:H	1:BBB:68:GLN:HE22	1.57	0.53
1:CCC:266:PRO:HD2	1:CCC:315:GLN:HE22	1.75	0.52
1:BBB:35:LEU:H	1:BBB:68:GLN:NE2	2.07	0.52
1:EEE:59:LEU:HD11	1:EEE:99:TYR:CD1	2.45	0.52
1:BBB:234:GLU:OE2	5:BBB:501:HOH:O	2.19	0.52
1:DDD:132:LEU:HD22	1:DDD:313:TRP:CD1	2.45	0.52
1:CCC:35:LEU:H	1:CCC:68:GLN:NE2	2.09	0.51
1:CCC:34:ARG:HA	1:CCC:68:GLN:HE22	1.76	0.50
1:AAA:266:PRO:HD2	1:AAA:315:GLN:HE22	1.75	0.50
1:AAA:269:VAL:O	1:AAA:272:TRP:HB3	2.11	0.50
1:AAA:35:LEU:H	1:AAA:68:GLN:NE2	2.09	0.50
1:DDD:94:GLU:HG3	1:DDD:98:LYS:HE2	1.94	0.50
1:DDD:150:ASN:HB3	1:DDD:155:ILE:HG13	1.94	0.50
1:CCC:193:GLU:OE2	1:CCC:197:ARG:NH2	2.44	0.50
1:DDD:39:LEU:N	1:DDD:40:PRO:CD	2.76	0.48
1:EEE:132:LEU:HD21	1:EEE:256:LEU:HG	1.94	0.48
1:BBB:247:GLY:HA3	3:BBB:402:MES:HN4	1.78	0.47
1:CCC:35:LEU:H	1:CCC:68:GLN:HE22	1.62	0.47
1:CCC:265:VAL:HA	1:CCC:315:GLN:HE22	1.78	0.47
1:EEE:272:TRP:O	1:EEE:276:GLN:HG2	2.14	0.47
1:AAA:162:LEU:HD12	1:AAA:196:LEU:HD21	1.96	0.47
1:BBB:283:PRO:HB2	1:BBB:295:THR:HG22	1.97	0.46
1:BBB:39:LEU:N	1:BBB:40:PRO:CD	2.78	0.46
1:EEE:79:ARG:NH1	1:EEE:92:GLU:OE2	2.48	0.46
1:EEE:206:ARG:HD3	5:EEE:575:HOH:O	2.14	0.46
1:AAA:194:ALA:O	1:AAA:198:LYS:HG2	2.16	0.46
1:AAA:119:HIS:CE1	5:AAA:601:HOH:O	2.60	0.46
1:BBB:34:ARG:HA	1:BBB:68:GLN:HE22	1.81	0.46
1:AAA:181:PHE:CD1	1:AAA:182:PRO:HD2	2.51	0.45
1:BBB:15:THR:HA	1:BBB:54:HIS:CD2	2.51	0.45
1:DDD:193:GLU:OE2	1:DDD:197:ARG:NH2	2.44	0.45
1:EEE:39:LEU:N	1:EEE:40:PRO:CD	2.79	0.45
1:EEE:191:GLU:N	1:EEE:191:GLU:OE1	2.49	0.45
1:AAA:190:PRO:HA	1:AAA:213:ARG:NH2	2.32	0.45
1:BBB:138:ILE:HD13	1:BBB:225:LEU:HD11	1.99	0.44

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:AAA:154:ARG:HD3	5:AAA:557:HOH:O	2.18	0.44
1:AAA:281:TRP:CE2	1:AAA:299:LEU:HD13	2.53	0.44
1:EEE:34:ARG:HA	1:EEE:68:GLN:HE22	1.83	0.43
1:AAA:12:GLY:HA2	1:AAA:174:ASP:OD1	2.19	0.43
1:CCC:46:ARG:HG3	1:CCC:139:GLU:OE2	2.18	0.43
1:AAA:191:GLU:HB3	5:AAA:604:HOH:O	2.17	0.43
1:EEE:56:SER:HA	1:EEE:64:TRP:O	2.19	0.43
1:EEE:269:VAL:O	1:EEE:272:TRP:HB2	2.18	0.43
1:EEE:50:GLN:NE2	5:EEE:510:HOH:O	2.45	0.43
1:AAA:141:LEU:HD13	1:AAA:254:ILE:HD13	2.02	0.42
1:DDD:137:PRO:CB	1:DDD:225:LEU:HD23	2.50	0.42
1:AAA:39:LEU:N	1:AAA:40:PRO:CD	2.83	0.42
1:AAA:270:HIS:C	1:AAA:272:TRP:H	2.23	0.42
1:EEE:62:GLN:HB2	5:EEE:543:HOH:O	2.20	0.42
1:CCC:15:THR:HA	1:CCC:54:HIS:CD2	2.55	0.41
1:CCC:113:TRP:HE3	1:CCC:120:PHE:CZ	2.38	0.41
1:BBB:283:PRO:CB	1:BBB:295:THR:HG22	2.50	0.41
1:AAA:272:TRP:O	1:AAA:276:GLN:HG2	2.21	0.41
1:BBB:109:LEU:HB3	1:BBB:113:TRP:CH2	2.56	0.40
1:CCC:50:GLN:HB2	5:CCC:511:HOH:O	2.21	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

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

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	AAA	299/337 (89%)	288 (96%)	11 (4%)	0	100	100
1	BBB	301/337 (89%)	292 (97%)	9 (3%)	0	100	100
1	CCC	304/337 (90%)	291 (96%)	13 (4%)	0	100	100
1	DDD	298/337 (88%)	287 (96%)	11 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	EEE	297/337 (88%)	288 (97%)	9 (3%)	0	100	100
All	All	1499/1685 (89%)	1446 (96%)	53 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 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	AAA	258/284 (91%)	257 (100%)	1 (0%)	91	93
1	BBB	258/284 (91%)	257 (100%)	1 (0%)	91	93
1	CCC	260/284 (92%)	259 (100%)	1 (0%)	91	93
1	DDD	255/284 (90%)	255 (100%)	0	100	100
1	EEE	255/284 (90%)	254 (100%)	1 (0%)	91	93
All	All	1286/1420 (91%)	1282 (100%)	4 (0%)	92	95

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	121	GLN
1	BBB	84	GLN
1	CCC	191	GLU
1	EEE	322	ASP

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

### 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](#)

10 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	MES	BBB	402	-	12,12,12	0.83	0	14,16,16	0.66	0
2	SEQ	AAA	401	-	21,23,23	1.23	4 (19%)	30,32,32	2.56	10 (33%)
4	GOL	EEE	403	-	5,5,5	0.15	0	5,5,5	0.41	0
2	SEQ	CCC	401	-	21,23,23	0.82	0	30,32,32	2.77	6 (20%)
4	GOL	DDD	402	-	5,5,5	0.14	0	5,5,5	0.31	0
2	SEQ	DDD	401	-	21,23,23	0.95	0	30,32,32	2.31	7 (23%)
2	SEQ	EEE	401	-	21,23,23	1.12	1 (4%)	30,32,32	2.79	8 (26%)
3	MES	EEE	402	-	12,12,12	0.83	0	14,16,16	0.62	0
2	SEQ	BBB	401	-	21,23,23	1.05	1 (4%)	30,32,32	2.44	9 (30%)
3	MES	AAA	402	-	12,12,12	0.83	0	14,16,16	0.52	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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	MES	BBB	402	-	-	1/6/14/14	0/1/1/1
2	SEQ	AAA	401	-	-	0/8/18/18	0/4/4/4
4	GOL	EEE	403	-	-	2/4/4/4	-
2	SEQ	CCC	401	-	-	1/8/18/18	0/4/4/4

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	DDD	402	-	-	4/4/4/4	-
2	SEQ	DDD	401	-	-	0/8/18/18	0/4/4/4
2	SEQ	EEE	401	-	-	0/8/18/18	0/4/4/4
3	MES	EEE	402	-	-	2/6/14/14	0/1/1/1
2	SEQ	BBB	401	-	-	0/8/18/18	0/4/4/4
3	MES	AAA	402	-	-	0/6/14/14	0/1/1/1

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	AAA	401	SEQ	C12-N20	2.44	1.36	1.32
2	EEE	401	SEQ	C09-C08	2.42	1.42	1.36
2	AAA	401	SEQ	C12-C07	2.31	1.47	1.44
2	BBB	401	SEQ	C10-C11	2.20	1.41	1.36
2	AAA	401	SEQ	C09-C08	2.12	1.41	1.36
2	AAA	401	SEQ	C10-C11	2.01	1.41	1.36

All (40) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	CCC	401	SEQ	C02-C03-C04	11.65	132.72	119.43
2	AAA	401	SEQ	C02-C03-C04	8.41	129.02	119.43
2	BBB	401	SEQ	C02-C03-C04	8.25	128.84	119.43
2	DDD	401	SEQ	C02-C03-C04	7.69	128.20	119.43
2	EEE	401	SEQ	C01-C03-C04	7.29	127.74	119.43
2	EEE	401	SEQ	C04-N05-C06	6.92	121.23	116.54
2	EEE	401	SEQ	C02-C03-C04	6.57	126.92	119.43
2	AAA	401	SEQ	C04-N05-C06	5.57	120.32	116.54
2	BBB	401	SEQ	C01-C03-C04	5.06	125.20	119.43
2	BBB	401	SEQ	C04-N05-C06	4.90	119.86	116.54
2	DDD	401	SEQ	C04-N05-C06	4.42	119.54	116.54
2	EEE	401	SEQ	C12-N13-C14	-4.39	115.82	124.16
2	CCC	401	SEQ	C04-N05-C06	4.38	119.51	116.54
2	DDD	401	SEQ	C01-C03-C04	4.31	124.34	119.43
2	AAA	401	SEQ	C07-C06-N05	-4.03	118.54	122.81
2	EEE	401	SEQ	C07-C06-N05	-4.01	118.55	122.81
2	CCC	401	SEQ	C07-C06-N05	-4.00	118.57	122.81
2	CCC	401	SEQ	C12-N13-C14	-3.89	116.78	124.16
2	EEE	401	SEQ	C07-C12-N13	3.88	124.29	120.63
2	AAA	401	SEQ	C01-C03-C04	3.83	123.79	119.43
2	BBB	401	SEQ	C07-C06-N05	-3.72	118.87	122.81
2	CCC	401	SEQ	C12-C07-C06	3.67	118.19	115.88

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	BBB	401	SEQ	C12-N13-C14	-3.58	117.36	124.16
2	DDD	401	SEQ	C07-C06-N05	-3.35	119.25	122.81
2	AAA	401	SEQ	C12-C07-C06	3.30	117.95	115.88
2	AAA	401	SEQ	C12-N13-C14	-2.97	118.53	124.16
2	DDD	401	SEQ	C12-N13-C14	-2.93	118.60	124.16
2	AAA	401	SEQ	C18-C17-C16	2.90	120.11	111.18
2	CCC	401	SEQ	C07-C12-N13	2.78	123.25	120.63
2	EEE	401	SEQ	C04-N20-C12	2.78	122.00	117.20
2	DDD	401	SEQ	C12-C07-C06	2.77	117.62	115.88
2	AAA	401	SEQ	C07-C12-N13	2.75	123.22	120.63
2	EEE	401	SEQ	N05-C04-N20	-2.53	122.50	126.09
2	AAA	401	SEQ	C17-C18-C19	2.36	116.22	111.42
2	DDD	401	SEQ	C08-C07-C06	2.32	120.92	118.33
2	AAA	401	SEQ	C11-C06-N05	2.29	122.18	118.69
2	BBB	401	SEQ	C19-C14-N13	-2.19	107.19	110.60
2	BBB	401	SEQ	C08-C07-C06	2.16	120.74	118.33
2	BBB	401	SEQ	C07-C12-N13	2.03	122.54	120.63
2	BBB	401	SEQ	C11-C06-N05	2.01	121.77	118.69

There are no chirality outliers.

All (10) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	BBB	402	MES	C8-C7-N4-C5
4	DDD	402	GOL	O1-C1-C2-O2
4	DDD	402	GOL	O1-C1-C2-C3
4	DDD	402	GOL	C1-C2-C3-O3
4	EEE	403	GOL	C1-C2-C3-O3
2	CCC	401	SEQ	C01-C03-C04-N05
4	DDD	402	GOL	O2-C2-C3-O3
3	EEE	402	MES	C7-C8-S-O3S
4	EEE	403	GOL	O2-C2-C3-O3
3	EEE	402	MES	C7-C8-S-O1S

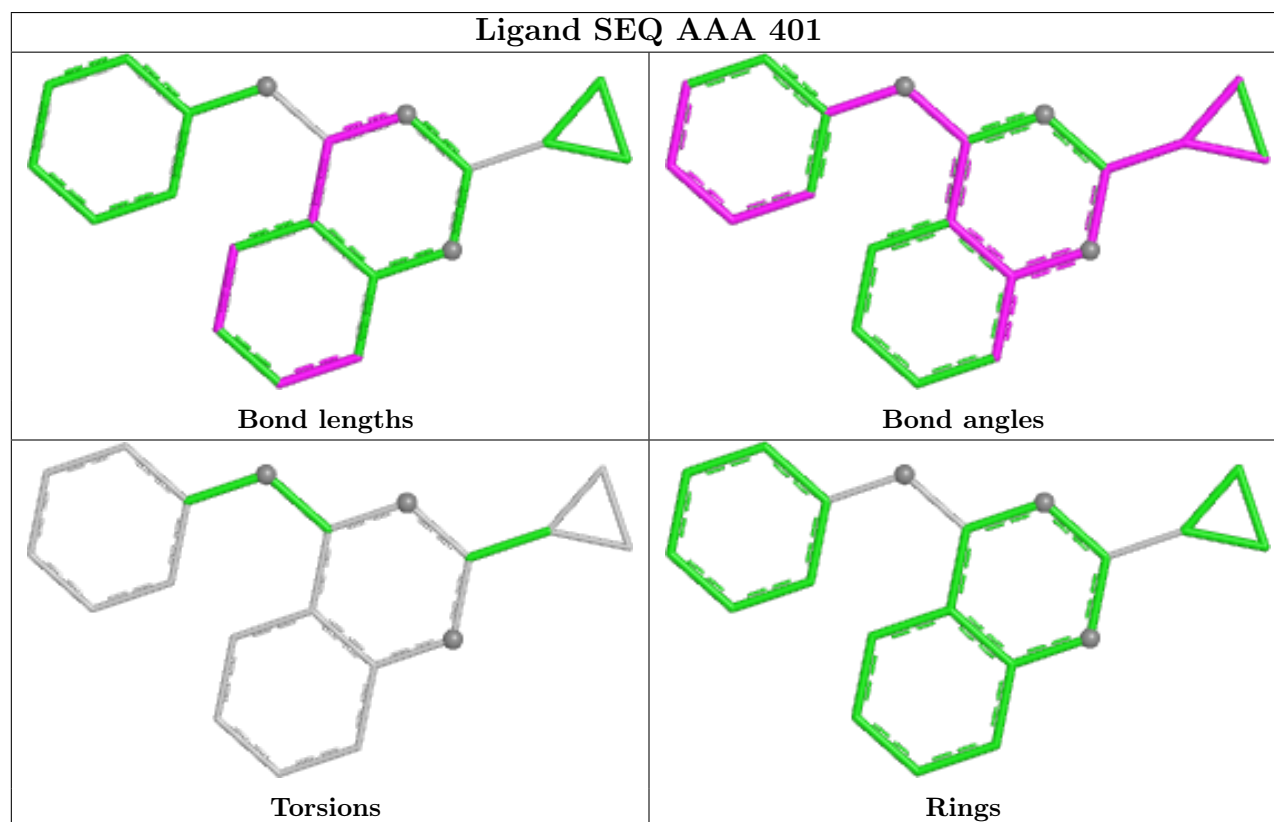
There are no ring outliers.

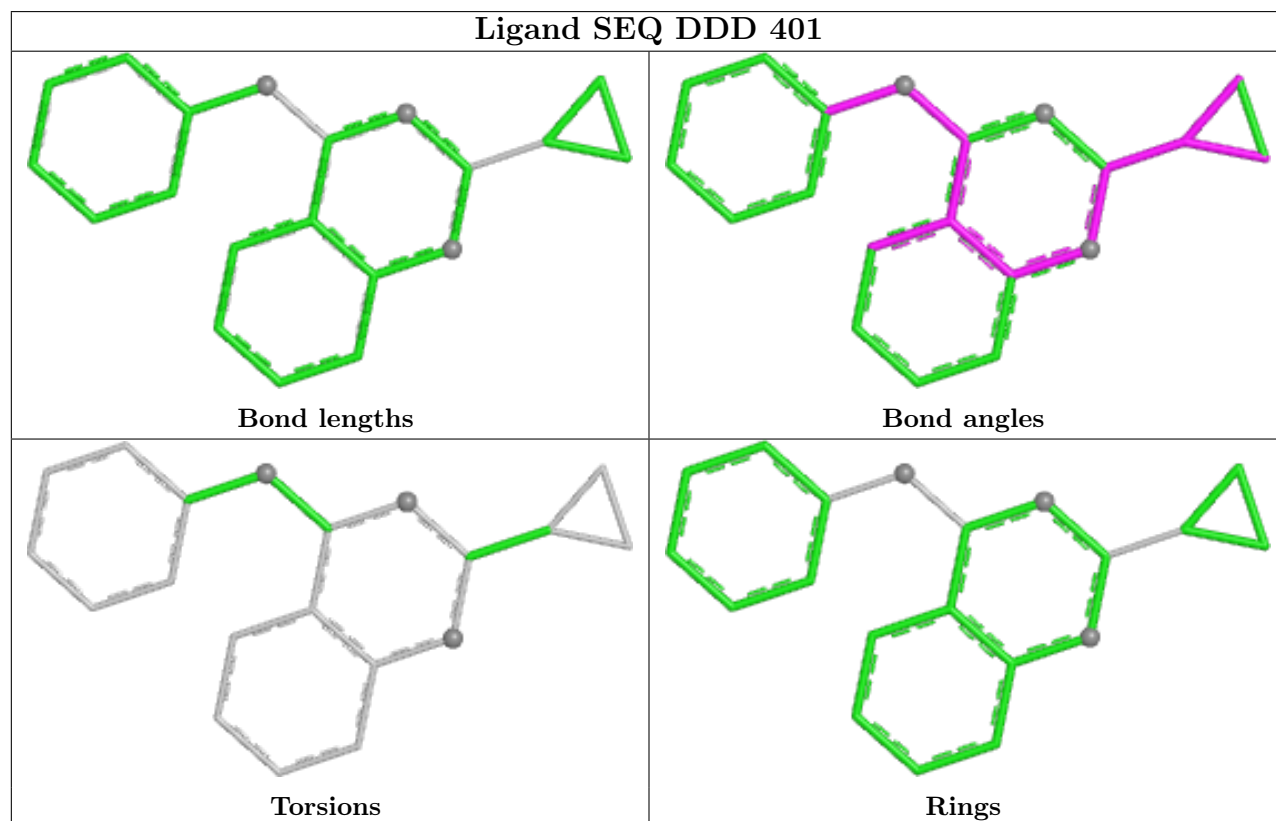
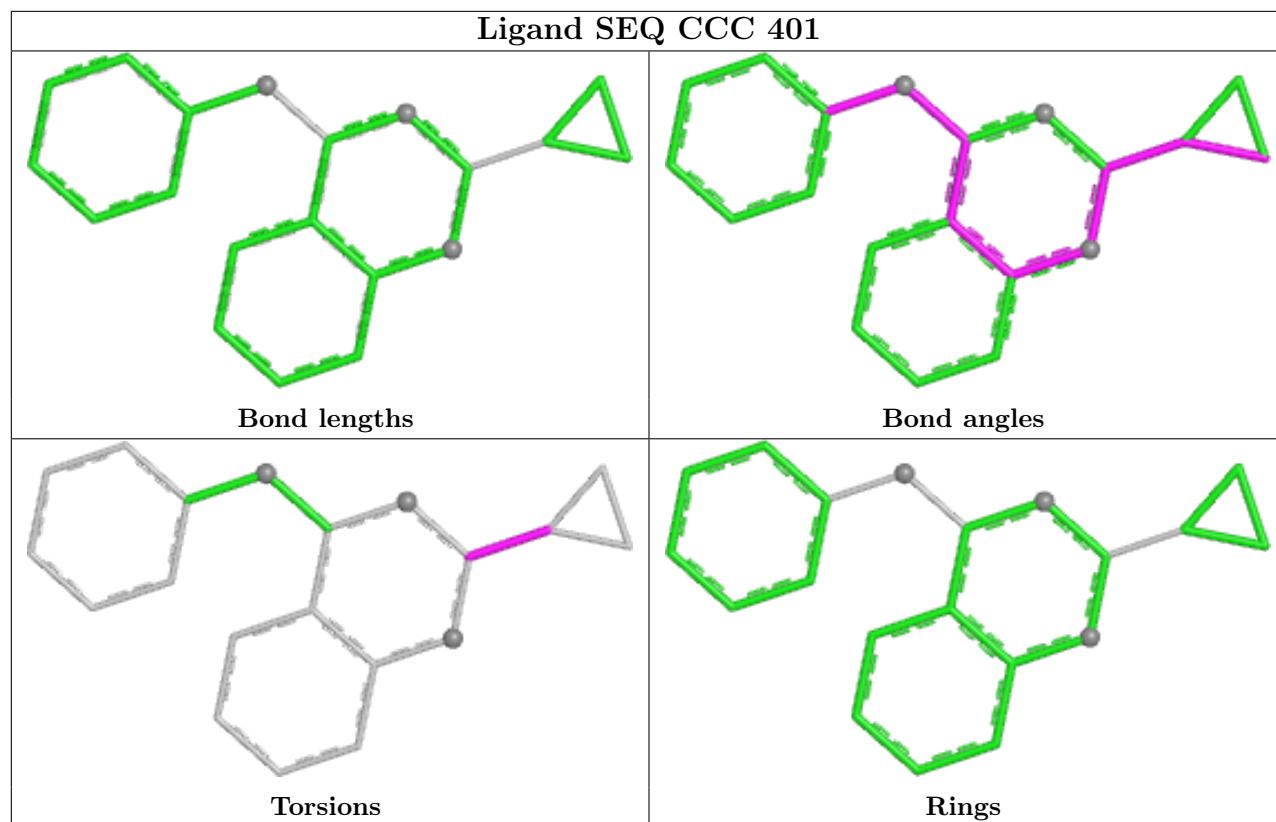
1 monomer is involved in 1 short contact:

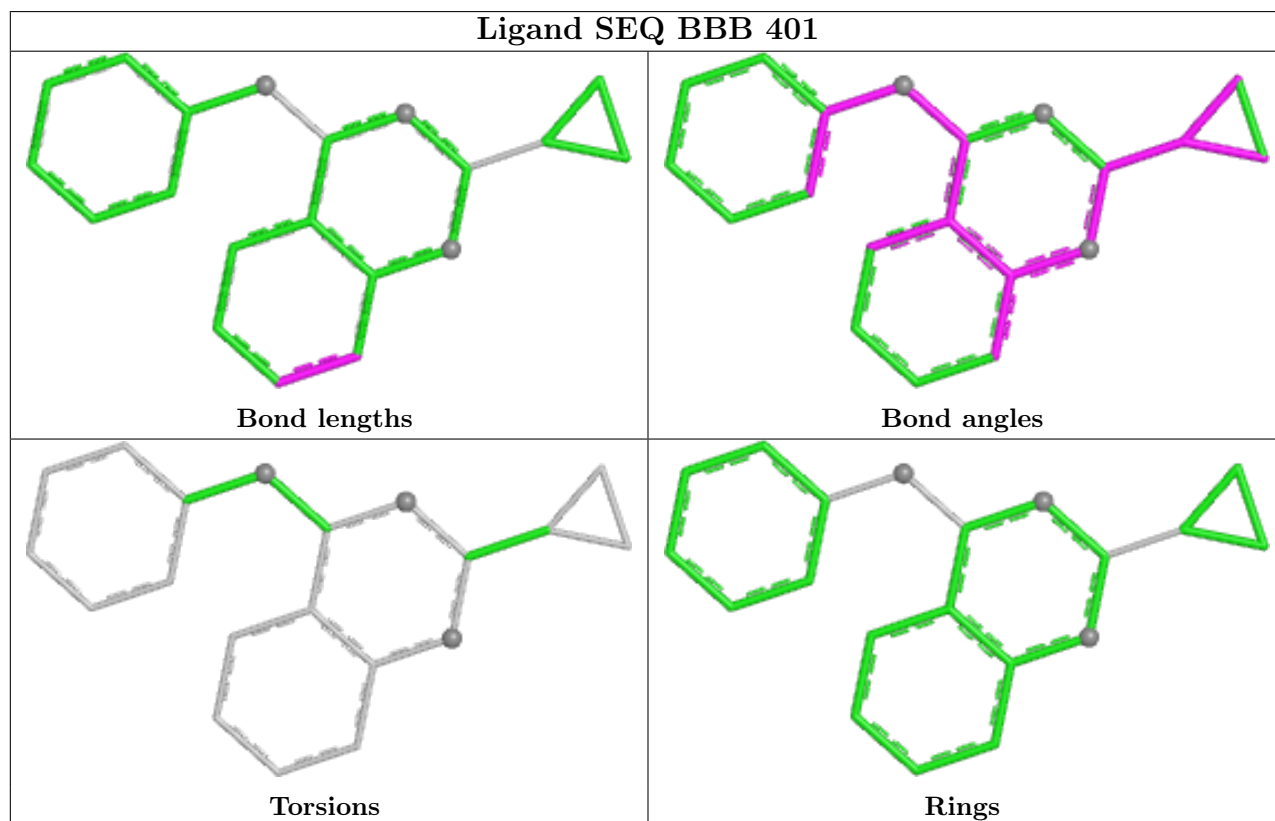
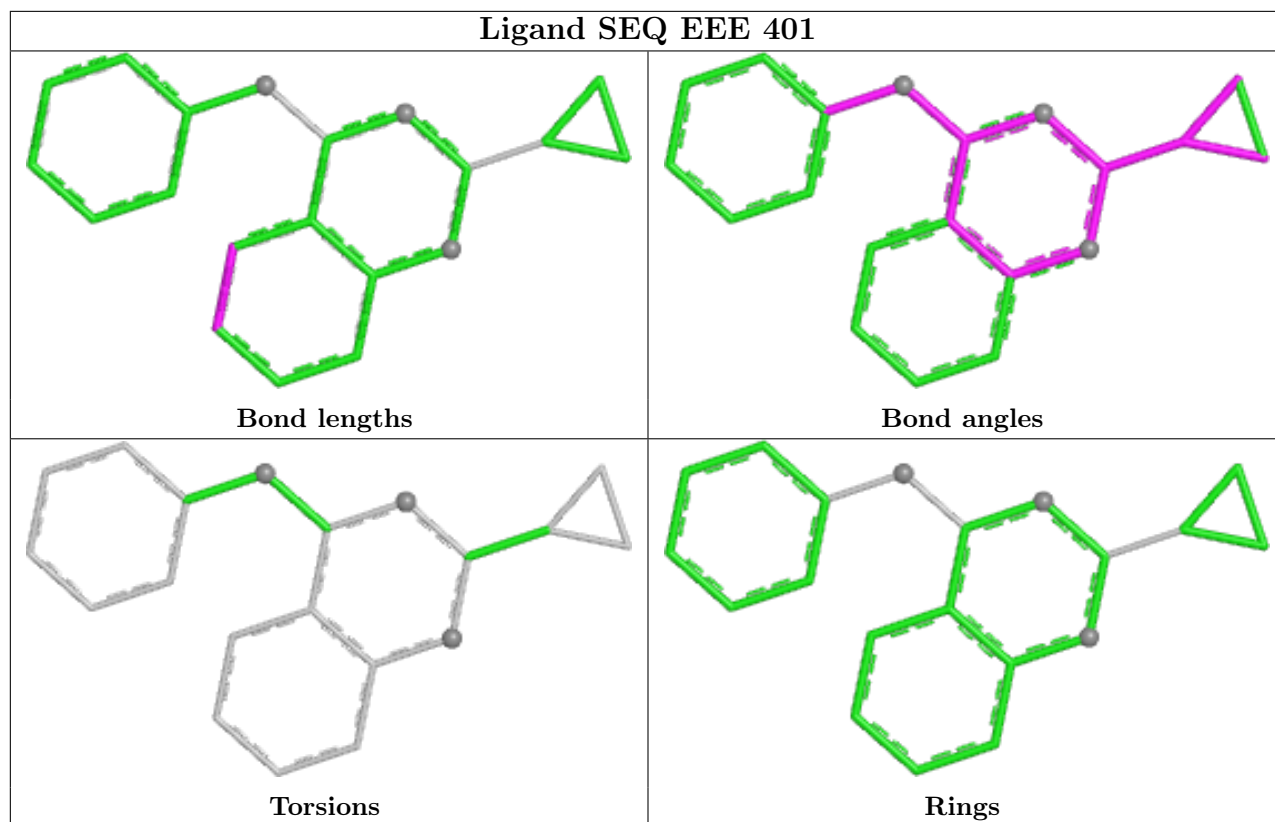
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	BBB	402	MES	1	0

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	AAA	305/337 (90%)	0.04	10 (3%) 46 45	17, 31, 60, 107	0
1	BBB	307/337 (91%)	-0.08	7 (2%) 60 59	16, 28, 50, 70	0
1	CCC	310/337 (91%)	-0.17	6 (1%) 66 65	14, 26, 48, 78	0
1	DDD	304/337 (90%)	-0.30	3 (0%) 82 81	14, 23, 43, 75	0
1	EEE	303/337 (89%)	-0.21	3 (0%) 82 81	15, 26, 51, 78	0
All	All	1529/1685 (90%)	-0.14	29 (1%) 66 65	14, 27, 51, 107	0

All (29) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	277	ARG	7.2
1	AAA	272	TRP	5.4
1	DDD	272	TRP	5.4
1	EEE	272	TRP	5.3
1	AAA	281	TRP	5.1
1	AAA	280	SER	5.0
1	BBB	290	GLY	4.5
1	CCC	288	ALA	4.2
1	AAA	175	ASP	4.0
1	AAA	11	MET	3.8
1	AAA	273	HIS	3.7
1	DDD	290	GLY	3.6
1	AAA	283	PRO	3.3
1	AAA	276	GLN	2.9
1	AAA	282	HIS	2.8
1	CCC	282	HIS	2.8
1	EEE	273	HIS	2.8
1	BBB	272	TRP	2.8
1	CCC	289	LYS	2.8
1	BBB	284	THR	2.6

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	CCC	70	GLU	2.6
1	CCC	294	GLN	2.5
1	DDD	273	HIS	2.4
1	EEE	281	TRP	2.4
1	BBB	281	TRP	2.3
1	BBB	175	ASP	2.2
1	BBB	294	GLN	2.1
1	CCC	191	GLU	2.1
1	BBB	190	PRO	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](#)

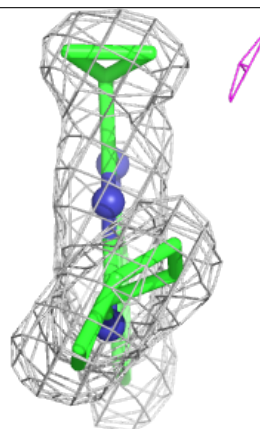
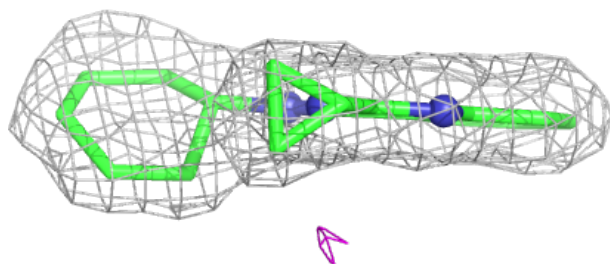
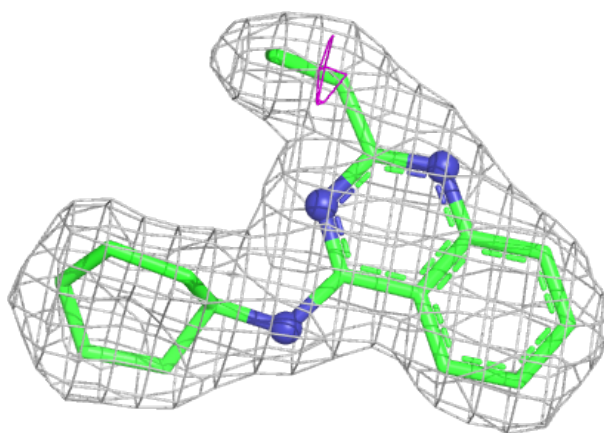
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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	MES	BBB	402	12/12	0.83	0.26	52,70,75,76	0
4	GOL	DDD	402	6/6	0.84	0.22	40,45,48,59	0
4	GOL	EEE	403	6/6	0.86	0.20	45,48,53,55	0
3	MES	EEE	402	12/12	0.89	0.18	40,55,58,59	0
3	MES	AAA	402	12/12	0.92	0.14	45,60,67,68	0
2	SEQ	AAA	401	20/20	0.94	0.13	22,29,32,33	0
2	SEQ	CCC	401	20/20	0.95	0.10	19,20,22,22	0
2	SEQ	BBB	401	20/20	0.95	0.10	20,21,25,25	0
2	SEQ	EEE	401	20/20	0.96	0.10	18,20,23,23	0
2	SEQ	DDD	401	20/20	0.96	0.09	20,21,25,25	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.

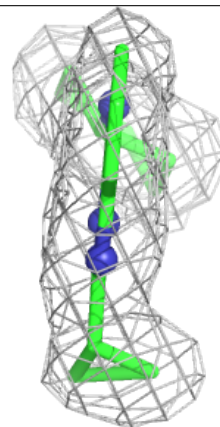
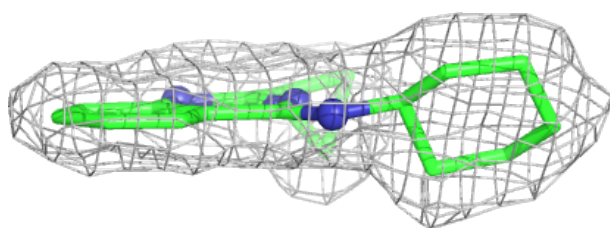
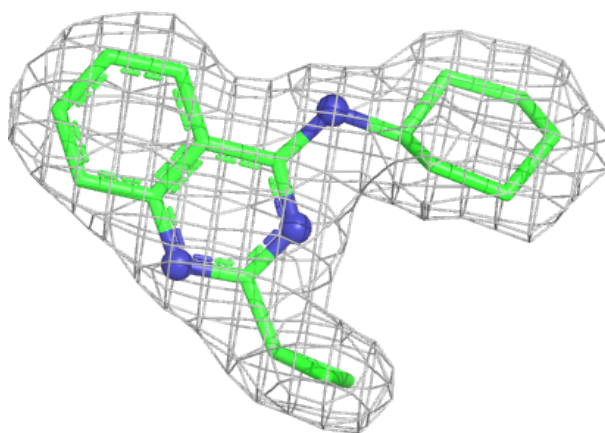
**Electron density around SEQ AAA 401:**

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



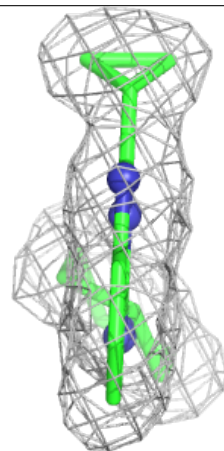
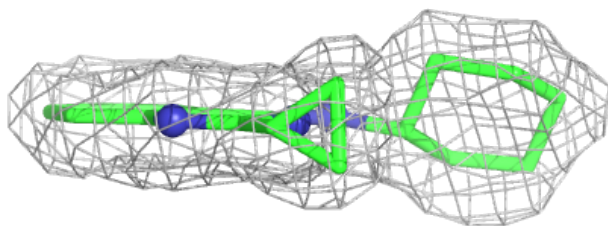
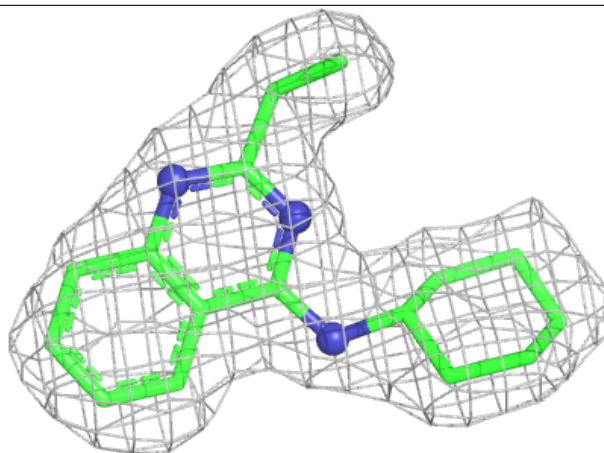
**Electron density around SEQ CCC 401:**

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



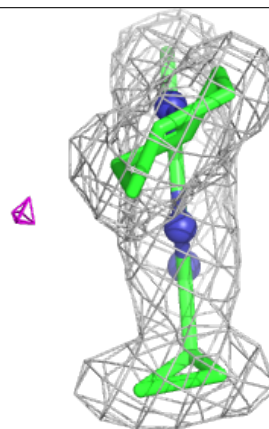
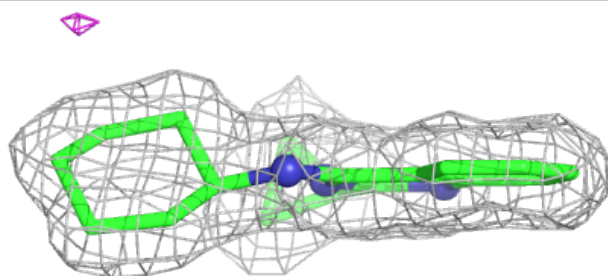
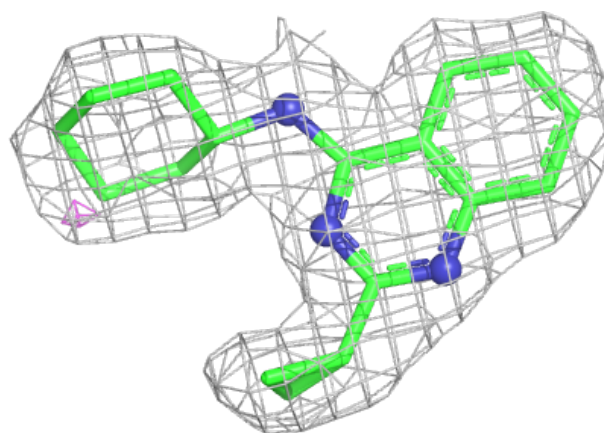
**Electron density around SEQ BBB 401:**

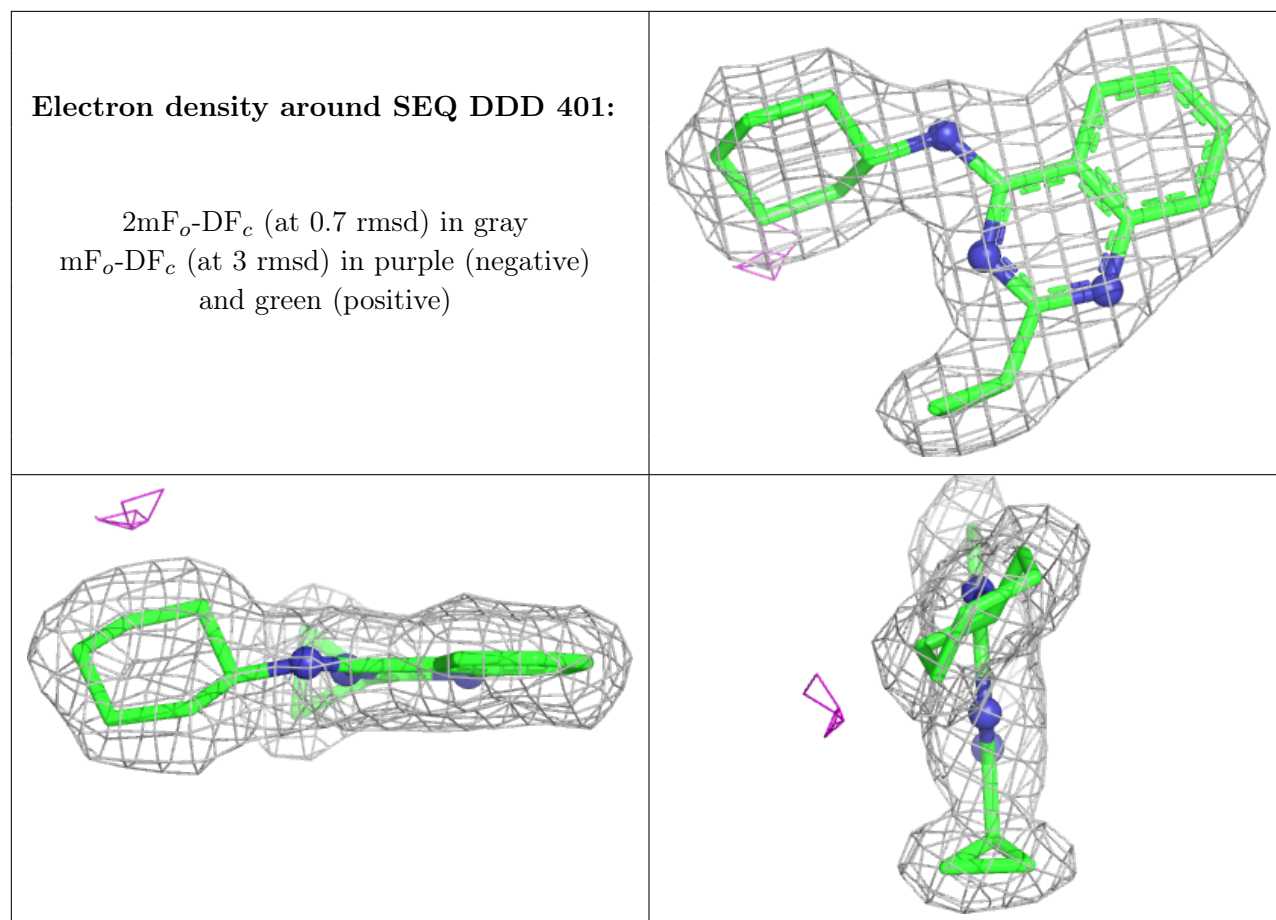
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around SEQ EEE 401:**

$2mF_o-DF_c$  (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.