



# Full wwPDB X-ray Structure Validation Report ⓘ

Jul 21, 2025 – 11:07 am BST

PDB ID : 9G53 / pdb\_00009g53  
Title : The structure of *Aspergillus fumigatus* UDP-GlcNAc pyrophosphorylase in complex with a fragment  
Authors : Yan, K.  
Deposited on : 2024-07-16  
Resolution : 1.96 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	:	4-5-2 with Phenix2.0rc1
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.44

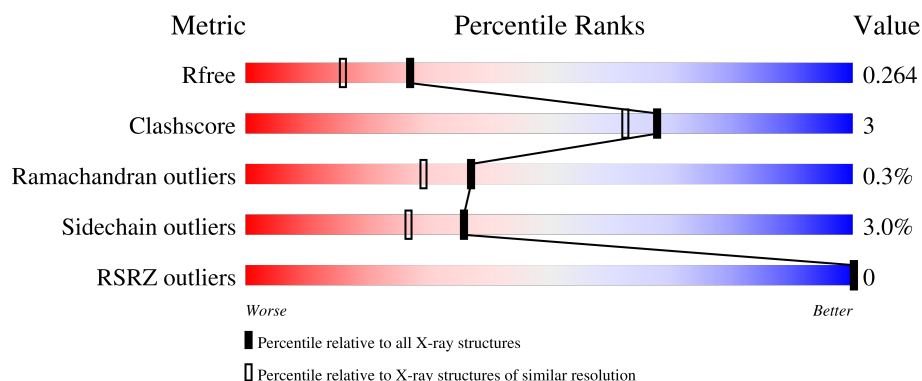
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.96 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))
$R_{free}$	164625	3187 (1.96-1.96)
Clashscore	180529	3412 (1.96-1.96)
Ramachandran outliers	177936	3390 (1.96-1.96)
Sidechain outliers	177891	3390 (1.96-1.96)
RSRZ outliers	164620	3186 (1.96-1.96)

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	484	
1	B	484	
2	C	484	
2	D	484	

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 15150 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 UDP-N-acetylglucosamine diphosphorylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	472	Total	C	N	O	S	0	1	0
			3669	2346	624	687	12			
1	A	473	Total	C	N	O	S	0	4	0
			3684	2355	623	693	13			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	26	GLY	-	expression tag	UNP A0A229XUD0
A	26	GLY	-	expression tag	UNP A0A229XUD0

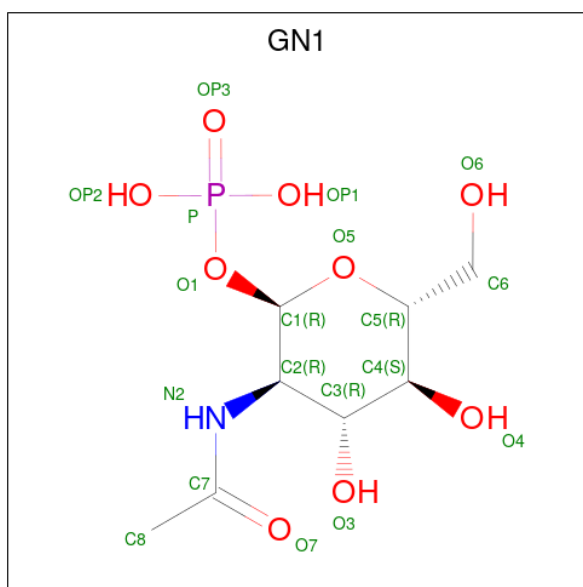
- Molecule 2 is a protein called UDP-N-acetylglucosamine diphosphorylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	467	Total	C	N	O	S	0	1	0
			3609	2304	614	678	13			
2	D	462	Total	C	N	O	S	0	2	0
			3585	2291	608	672	14			

There are 2 discrepancies between the modelled and reference sequences:

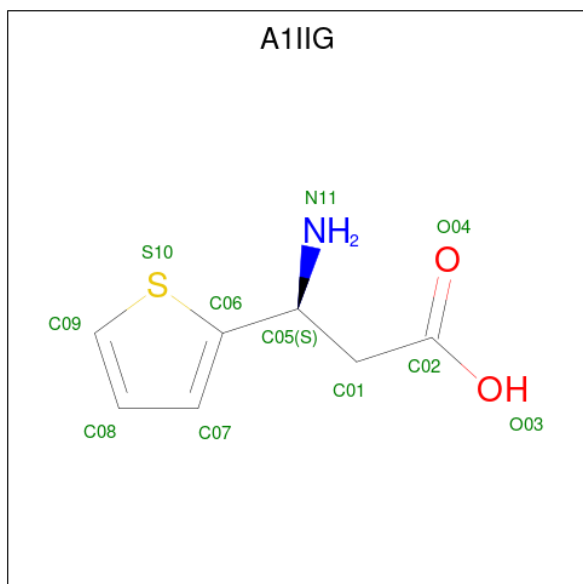
Chain	Residue	Modelled	Actual	Comment	Reference
C	26	GLY	-	expression tag	UNP A0A229XUD0
D	26	GLY	-	expression tag	UNP A0A229XUD0

- Molecule 3 is 2-acetamido-2-deoxy-1-O-phosphono-alpha-D-glucopyranose (CCD ID: GN1) (formula: C<sub>8</sub>H<sub>16</sub>NO<sub>9</sub>P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	B	1	Total	C	N	O	P	0	0
			19	8	1	9	1		
3	A	1	Total	C	N	O	P	0	0
			19	8	1	9	1		

- Molecule 4 is (3 {S})-3-azanyl-3-thiophen-2-yl-propanoic acid (CCD ID: A1IIG) (formula:  $C_7H_9NO_2S$ ) (labeled as "Ligand of Interest" by depositor).



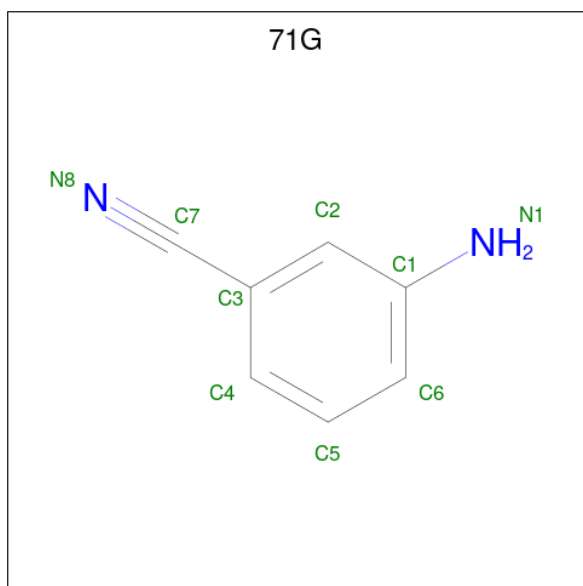
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	B	1	Total	C	N	O	S	0	0
			11	7	1	2	1		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	S	0	0
			11	7	1	2	1		

- Molecule 5 is 3-aminobenzonitrile (CCD ID: 71G) (formula:  $C_7H_6N_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	C	N	0	0
			9	7	2		
5	D	1	Total	C	N	0	0
			9	7	2		
5	D	1	Total	C	N	0	0
			9	7	2		

- Molecule 6 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	1	Total	Mg	0	0
			1	1		

- Molecule 7 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	C	1	Total	Cl	0	0
			1	1		

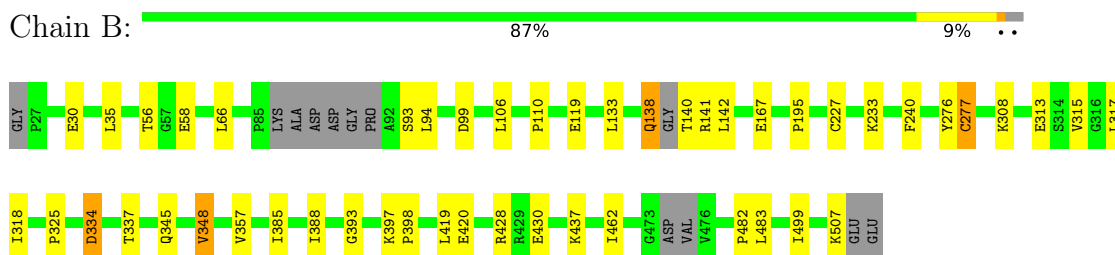
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	B	151	Total 151	O 151	0	0
8	A	145	Total 145	O 145	0	0
8	C	99	Total 99	O 99	0	0
8	D	119	Total 119	O 119	0	0

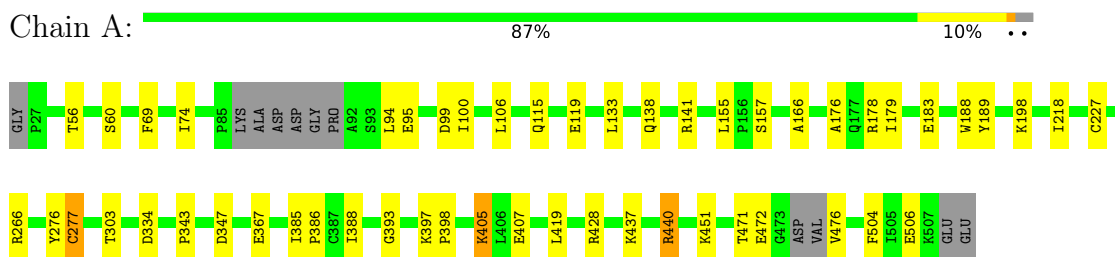
### 3 Residue-property plots

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.

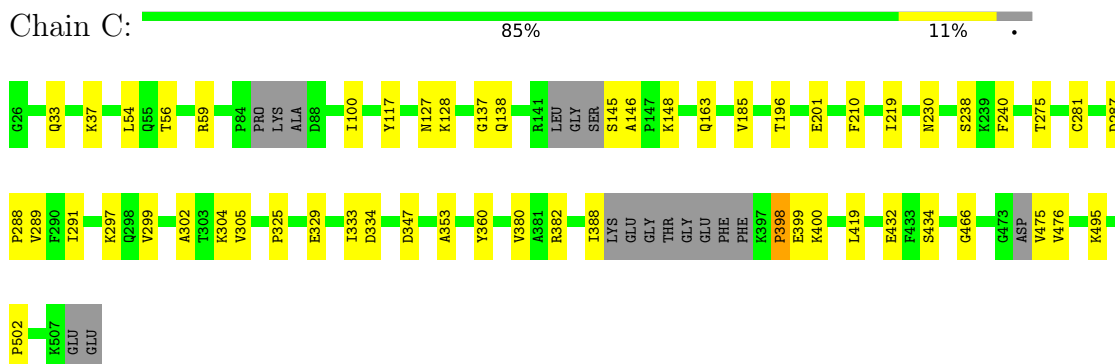
- Molecule 1: UDP-N-acetylglucosamine diphosphorylase



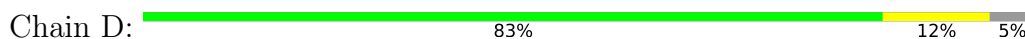
- Molecule 1: UDP-N-acetylglucosamine diphosphorylase



- Molecule 2: UDP-N-acetylglucosamine diphosphorylase



- Molecule 2: UDP-N-acetylglucosamine diphosphorylase







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	55.67Å 136.77Å 142.55Å 90.00° 90.03° 90.00°	Depositor
Resolution (Å)	48.50 – 1.96 48.50 – 1.96	Depositor EDS
% Data completeness (in resolution range)	99.3 (48.50-1.96) 99.2 (48.50-1.96)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.15 (at 1.97Å)	Xtriage
Refinement program	PHENIX (1.19.2_4158: ???)	Depositor
R, $R_{free}$	0.216 , 0.263 0.217 , 0.264	Depositor DCC
$R_{free}$ test set	7362 reflections (4.86%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	36.8	Xtriage
Anisotropy	0.613	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 50.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.000 for -h,l,k 0.000 for -h,-l,-k 0.445 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	15150	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.15% 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: 71G, GN1, A1IIG, MG, CL, YCM

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z  > 5$	RMSZ	$\# Z  > 5$
1	A	0.32	0/3764	0.48	0/5085
1	B	0.32	0/3739	0.48	0/5049
2	C	0.30	0/3688	0.46	0/4987
2	D	0.30	0/3666	0.49	0/4953
All	All	0.31	0/14857	0.47	0/20074

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
1	A	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	277	YCM	Mainchain

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

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3684	0	3635	27	0
1	B	3669	0	3624	23	0
2	C	3609	0	3547	23	0
2	D	3585	0	3542	28	0
3	A	19	0	0	0	0
3	B	19	0	0	0	0
4	A	11	0	0	0	0
4	B	11	0	0	0	0
5	B	9	0	0	0	0
5	D	18	0	0	0	0
6	B	1	0	0	0	0
7	C	1	0	0	0	0
8	A	145	0	0	2	0
8	B	151	0	0	0	0
8	C	99	0	0	1	0
8	D	119	0	0	1	0
All	All	15150	0	14348	100	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 (100) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:277[A]:CYS:SG	8:D:720:HOH:O	2.39	0.78
2:D:230:ASN:OD1	2:D:382:ARG:NH1	2.23	0.69
1:A:100:ILE:H	1:A:100:ILE:HD12	1.60	0.67
2:D:148:LYS:HD3	2:D:279:ASP:HB3	1.76	0.66
2:D:275:THR:HG22	2:D:359:HIS:HB2	1.79	0.65
2:D:197:ARG:HD2	2:D:221:GLU:HG3	1.80	0.64
1:B:110:PRO:HB3	1:A:183:GLU:HB2	1.77	0.64
1:A:440:ARG:NH1	8:A:703:HOH:O	2.30	0.63
1:B:388:ILE:HG23	1:B:393:GLY:HA2	1.81	0.63
2:D:222:GLN:NE2	2:D:247:ASP:OD1	2.34	0.61
2:C:238:SER:HA	2:C:388:ILE:HD11	1.84	0.59
1:A:138:GLN:OE1	1:A:138:GLN:N	2.37	0.58
1:B:227:CYS:HB3	1:B:385:ILE:HD13	1.85	0.57
2:C:137:GLY:HA3	2:C:196:THR:HG21	1.86	0.57
2:D:54:LEU:HD13	2:D:240:PHE:HB3	1.85	0.57
2:C:475:VAL:HB	2:C:502:PRO:HG3	1.86	0.57
1:B:345:GLN:HB3	1:B:348:VAL:HG13	1.87	0.55
2:D:305:VAL:HG11	2:D:353:ALA:HB1	1.89	0.55

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:141:ARG:HB2	1:B:437:LYS:O	2.06	0.55
1:B:133:LEU:HD13	1:B:276:TYR:CZ	2.42	0.54
1:A:388:ILE:HG23	1:A:393:GLY:HA2	1.89	0.54
2:C:137:GLY:HA2	2:C:148:LYS:H	1.72	0.54
2:C:304:LYS:NZ	2:C:432:GLU:OE1	2.40	0.54
1:A:227:CYS:HB3	1:A:385:ILE:HD13	1.90	0.52
1:A:471:THR:HG21	1:A:476:VAL:HG23	1.91	0.52
1:A:106:LEU:HD21	1:A:428:ARG:HG2	1.90	0.52
1:B:308:LYS:HD2	1:B:317:LEU:HD21	1.92	0.52
1:A:141:ARG:HB2	1:A:437:LYS:O	2.09	0.51
2:C:163:GLN:HB2	2:C:210:PHE:CE1	2.45	0.51
2:C:297:LYS:HB2	2:C:299:VAL:HG13	1.94	0.50
2:C:201:GLU:HG3	2:C:219:ILE:HD12	1.93	0.49
2:D:114:GLU:O	2:D:118:GLU:HG3	2.13	0.49
2:C:54:LEU:HD13	2:C:240:PHE:HB3	1.94	0.49
2:D:141:ARG:HD3	2:D:437:LYS:O	2.13	0.48
1:B:167:GLU:HG2	1:B:483:LEU:HD21	1.95	0.48
2:C:230:ASN:OD1	2:C:382:ARG:NH1	2.46	0.48
2:C:302:ALA:HB3	2:C:360:TYR:HB3	1.94	0.48
2:C:466:GLY:HA3	2:C:495:LYS:HG3	1.96	0.48
1:A:176:ALA:O	1:A:179:ILE:HG12	2.13	0.48
2:D:148:LYS:NZ	2:D:436:LEU:O	2.45	0.48
2:D:200:THR:HG22	2:D:219:ILE:HD13	1.94	0.48
2:C:305:VAL:HG11	2:C:353:ALA:HB1	1.95	0.48
2:D:385:ILE:HD11	2:D:403:GLY:HA3	1.96	0.48
2:D:128:LYS:HD2	2:D:271:GLU:HG3	1.96	0.48
1:A:157:SER:HB3	1:A:504:PHE:CZ	2.49	0.47
2:D:435:PRO:HB2	2:D:437:LYS:HE2	1.96	0.47
2:D:201:GLU:HG2	2:D:219:ILE:HD12	1.97	0.47
1:B:133:LEU:HD13	1:B:276:TYR:CE2	2.50	0.47
1:A:95:GLU:HB3	8:A:820:HOH:O	2.14	0.47
1:A:386:PRO:HA	1:A:398:PRO:HD3	1.98	0.46
2:D:365:PHE:HZ	2:D:413:VAL:HG21	1.81	0.46
1:A:115:GLN:O	1:A:119:GLU:HG2	2.16	0.46
2:D:347:ASP:OD1	2:D:347:ASP:N	2.45	0.46
1:B:430:GLU:H	1:B:430:GLU:CD	2.24	0.45
2:C:33:GLN:NE2	8:C:707:HOH:O	2.36	0.45
1:A:405:LYS:HD2	1:A:407:GLU:HG3	1.98	0.45
1:A:133:LEU:HD13	1:A:276:TYR:CZ	2.51	0.45
1:B:462:ILE:HG21	1:B:499:ILE:HD12	1.99	0.45
1:B:507:LYS:HE2	1:B:507:LYS:HB3	1.70	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:100:ILE:O	2:C:297:LYS:NZ	2.48	0.44
1:A:178:ARG:HE	1:A:178:ARG:HB3	1.57	0.44
2:D:281[A]:CYS:SG	2:D:434:SER:HB3	2.57	0.44
2:C:127:ASN:OD1	2:C:185:VAL:N	2.46	0.44
2:D:127:ASN:OD1	2:D:185:VAL:N	2.47	0.44
1:A:189:TYR:CE1	1:A:218:ILE:HD13	2.52	0.44
2:C:398:PRO:O	2:C:400:LYS:N	2.49	0.43
1:B:106:LEU:HD21	1:B:428:ARG:HG2	2.01	0.43
1:A:155:LEU:HD23	1:A:451:LYS:HG3	2.01	0.43
2:D:249:ASN:HB3	2:D:410:VAL:HB	2.00	0.43
1:B:140:THR:C	1:B:142:LEU:H	2.27	0.43
1:A:133:LEU:HD13	1:A:276:TYR:CE2	2.53	0.43
1:A:347:ASP:OD1	1:A:347:ASP:N	2.41	0.43
1:B:313:GLU:HG2	1:B:315:VAL:HG13	2.01	0.43
2:C:325:PRO:HD3	2:C:419:LEU:HD23	2.00	0.43
1:B:325:PRO:HG3	1:B:419:LEU:HD23	2.02	0.42
1:A:397:LYS:HA	1:A:398:PRO:HD3	1.91	0.42
2:D:236:MET:O	2:D:387:CYS:HB2	2.20	0.42
2:C:117:TYR:CE1	2:C:288:PRO:HG2	2.54	0.42
2:D:198:LYS:HE3	2:D:198:LYS:HB3	1.88	0.42
1:A:198:LYS:HB2	1:A:198:LYS:HE2	1.92	0.42
1:B:334:ASP:OD1	1:B:337:THR:OG1	2.35	0.41
2:C:281[A]:CYS:SG	2:C:434:SER:HB3	2.60	0.41
2:D:109:ASP:OD1	2:D:111:LYS:HG2	2.20	0.41
2:D:191:MET:HE1	2:D:252:ILE:HB	2.01	0.41
1:B:58:GLU:HB3	1:B:240:PHE:CG	2.55	0.41
1:B:277:YCM:HB3	1:B:357:VAL:HG11	2.02	0.41
1:A:166:ALA:HB2	1:A:188:TRP:CZ3	2.56	0.41
2:C:145:SER:HB2	2:C:146:ALA:H	1.69	0.41
2:C:287:ASP:O	2:C:291:ILE:HG12	2.21	0.41
1:A:266:ARG:HD3	1:A:367:GLU:OE2	2.21	0.41
2:C:117:TYR:HB2	2:C:289:VAL:HG22	2.02	0.41
2:D:135:ALA:O	2:D:137:GLY:N	2.54	0.41
2:D:325:PRO:HD3	2:D:419:LEU:HD13	2.03	0.41
1:B:35:LEU:HD21	1:B:66:LEU:HB2	2.03	0.41
1:A:69:PHE:HB3	1:A:74:ILE:HD11	2.03	0.40
1:B:138:GLN:HE22	1:B:195:PRO:HG2	1.85	0.40
1:B:397:LYS:HA	1:B:398:PRO:HD3	1.95	0.40
1:A:100:ILE:H	1:A:100:ILE:CD1	2.31	0.40
2:D:48:PHE:HD1	2:D:51:VAL:HG21	1.87	0.40
1:B:167:GLU:HB3	1:B:482:PRO:HB2	2.03	0.40

There are no symmetry-related clashes.

### 5.3 Torsion angles [i](#)

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

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	470/484 (97%)	456 (97%)	13 (3%)	1 (0%)	44	37
1	B	464/484 (96%)	453 (98%)	11 (2%)	0	100	100
2	C	458/484 (95%)	442 (96%)	13 (3%)	3 (1%)	19	10
2	D	454/484 (94%)	443 (98%)	10 (2%)	1 (0%)	44	37
All	All	1846/1936 (95%)	1794 (97%)	47 (2%)	5 (0%)	37	29

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	136	GLY
2	C	138	GLN
1	A	343	PRO
2	C	399	GLU
2	C	398	PRO

#### 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	387/407 (95%)	376 (97%)	11 (3%)	38	29
1	B	384/407 (94%)	372 (97%)	12 (3%)	35	25
2	C	380/408 (93%)	369 (97%)	11 (3%)	37	28

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	D	380/408 (93%)	368 (97%)	12 (3%)	34	24
All	All	1531/1630 (94%)	1485 (97%)	46 (3%)	36	27

All (46) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	30	GLU
1	B	56	THR
1	B	93	SER
1	B	94	LEU
1	B	99	ASP
1	B	119	GLU
1	B	138	GLN
1	B	233	LYS
1	B	318	ILE
1	B	334	ASP
1	B	348	VAL
1	B	420	GLU
1	A	56	THR
1	A	60	SER
1	A	94	LEU
1	A	99	ASP
1	A	303	THR
1	A	334	ASP
1	A	405	LYS
1	A	419	LEU
1	A	440	ARG
1	A	472	GLU
1	A	506	GLU
2	C	37	LYS
2	C	56	THR
2	C	59	ARG
2	C	128	LYS
2	C	275	THR
2	C	329	GLU
2	C	333	ILE
2	C	334	ASP
2	C	347	ASP
2	C	380	VAL
2	C	476	VAL
2	D	52	ASP
2	D	56	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	D	198	LYS
2	D	206	GLN
2	D	333	ILE
2	D	341	LYS
2	D	371	LEU
2	D	374	HIS
2	D	380	VAL
2	D	385	ILE
2	D	406	LEU
2	D	497	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	115	GLN
1	B	222	GLN
1	B	374	HIS
1	B	378	HIS
1	A	222	GLN
1	A	230	ASN
2	C	64	HIS
2	C	165	GLN
2	C	322	ASN
2	C	408	GLN
2	D	177	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

2 non-standard protein/DNA/RNA residues 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
1	YCM	A	277	1	7,9,10	1.84	1 (14%)	4,10,12	0.89	0
1	YCM	B	277	1	7,9,10	1.80	1 (14%)	4,10,12	1.33	1 (25%)

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
1	YCM	A	277	1	-	2/6/8/10	-
1	YCM	B	277	1	-	2/6/8/10	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	277	YCM	CE-NZ2	4.26	1.46	1.32
1	B	277	YCM	CE-NZ2	4.13	1.46	1.32

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	277	YCM	CE-CD-SG	-2.13	107.33	113.59

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	B	277	YCM	SG-CD-CE-OZ1
1	B	277	YCM	SG-CD-CE-NZ2
1	A	277	YCM	SG-CD-CE-NZ2
1	A	277	YCM	SG-CD-CE-OZ1

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	B	277	YCM	1	0

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 9 ligands modelled in this entry, 2 are monoatomic - leaving 7 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	GN1	B	601	-	18,19,19	1.49	3 (16%)	27,28,28	1.53	5 (18%)
5	71G	D	601	-	9,9,9	1.44	2 (22%)	11,11,11	0.83	0
5	71G	D	602	-	9,9,9	1.49	2 (22%)	11,11,11	0.58	0
4	A1IIG	B	602	-	10,11,11	0.81	0	10,14,14	2.85	2 (20%)
3	GN1	A	601	-	18,19,19	1.51	4 (22%)	27,28,28	1.40	3 (11%)
5	71G	B	603	-	9,9,9	1.34	2 (22%)	11,11,11	0.74	0
4	A1IIG	A	602	-	10,11,11	0.84	0	10,14,14	2.93	3 (30%)

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
3	GN1	B	601	-	-	1/10/31/31	0/1/1/1
5	71G	D	601	-	-	0/2/2/2	0/1/1/1
5	71G	D	602	-	-	0/2/2/2	0/1/1/1
4	A1IIG	B	602	-	-	0/4/8/8	0/1/1/1
3	GN1	A	601	-	-	3/10/31/31	0/1/1/1
5	71G	B	603	-	-	0/2/2/2	0/1/1/1
4	A1IIG	A	602	-	-	0/4/8/8	0/1/1/1

All (13) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	D	602	71G	C3-C7	3.83	1.53	1.44
3	B	601	GN1	C7-N2	3.30	1.45	1.34
5	B	603	71G	C3-C7	3.27	1.51	1.44
3	A	601	GN1	P-O1	3.24	1.65	1.59
5	D	601	71G	C3-C7	3.21	1.51	1.44
3	A	601	GN1	C7-N2	3.05	1.44	1.34
3	B	601	GN1	P-O1	2.91	1.64	1.59
5	D	601	71G	C1-N1	2.69	1.47	1.38
3	A	601	GN1	P-OP1	-2.37	1.45	1.54
3	B	601	GN1	P-OP2	-2.34	1.45	1.54
5	D	602	71G	C1-N1	2.25	1.46	1.38
3	A	601	GN1	P-OP2	-2.19	1.46	1.54
5	B	603	71G	C1-N1	2.15	1.45	1.38

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	602	A1IIG	C08-C09-S10	-8.09	106.42	112.98
4	B	602	A1IIG	C08-C09-S10	-8.04	106.45	112.98
3	A	601	GN1	C4-C3-C2	-4.43	103.86	110.34
3	B	601	GN1	C4-C3-C2	-4.29	104.06	110.34
3	B	601	GN1	O5-C1-O1	-2.97	107.48	111.36
3	A	601	GN1	C1-O5-C5	2.76	119.11	113.69
4	B	602	A1IIG	C07-C06-S10	-2.51	105.50	111.36
4	A	602	A1IIG	C07-C06-S10	-2.39	105.77	111.36
3	B	601	GN1	C1-O5-C5	2.35	118.31	113.69
4	A	602	A1IIG	C07-C06-C05	-2.29	126.30	128.98
3	B	601	GN1	C2-N2-C7	-2.19	117.86	123.18
3	A	601	GN1	O5-C1-O1	-2.17	108.53	111.36
3	B	601	GN1	C8-C7-N2	2.00	119.49	116.10

There are no chirality outliers.

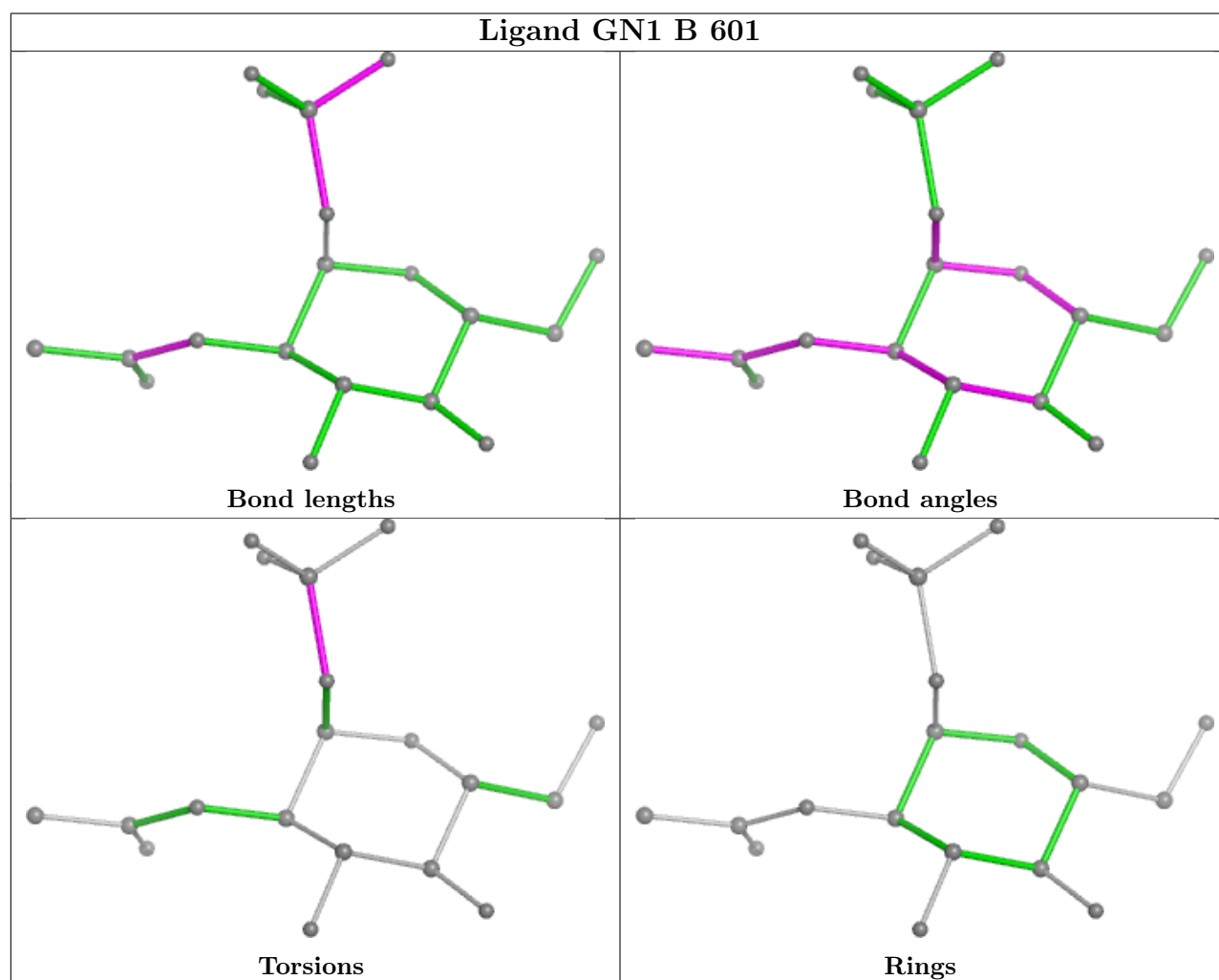
All (4) torsion outliers are listed below:

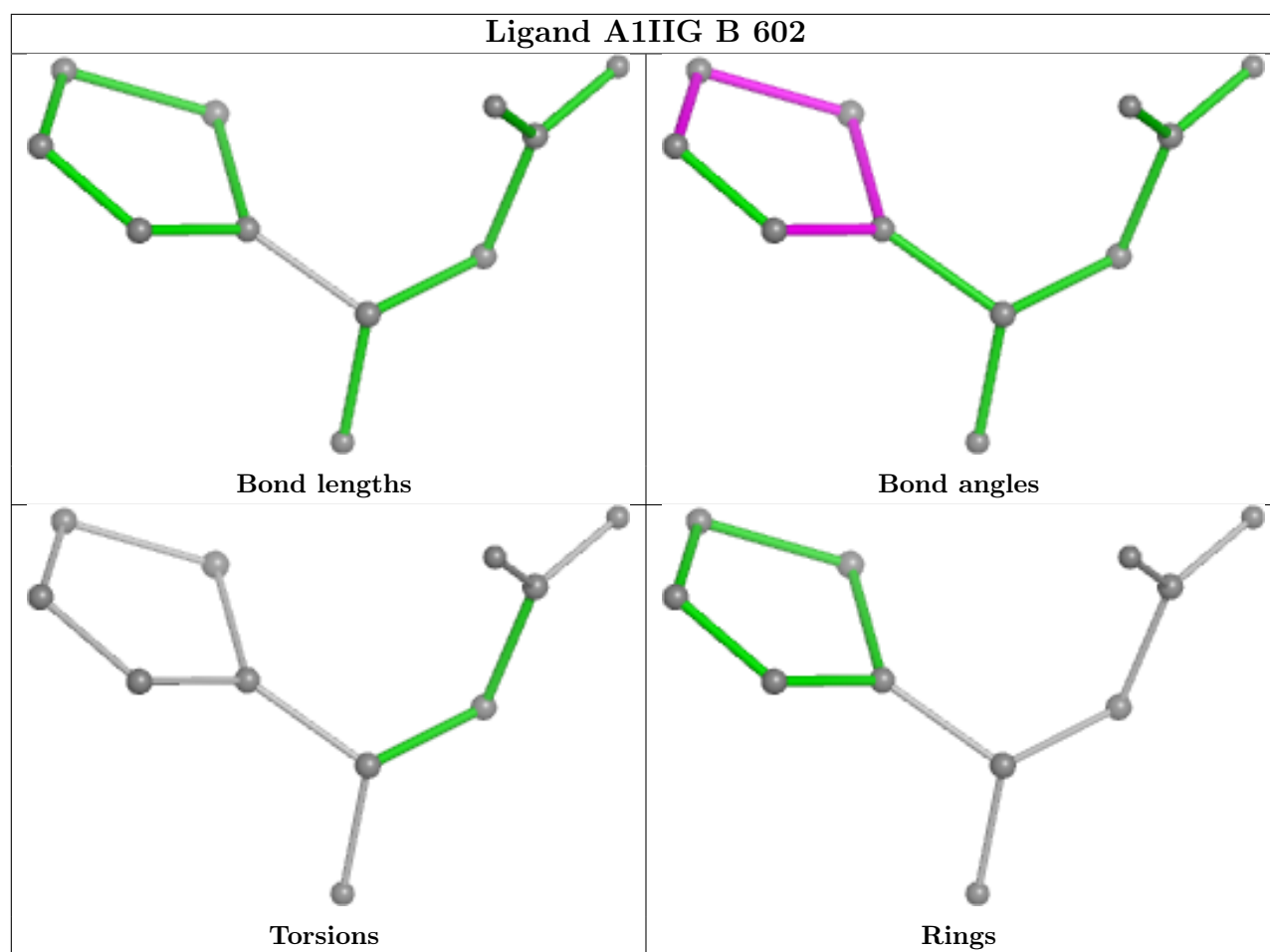
Mol	Chain	Res	Type	Atoms
3	B	601	GN1	C1-O1-P-OP3
3	A	601	GN1	C1-O1-P-OP1
3	A	601	GN1	C1-O1-P-OP2
3	A	601	GN1	C1-O1-P-OP3

There are no ring outliers.

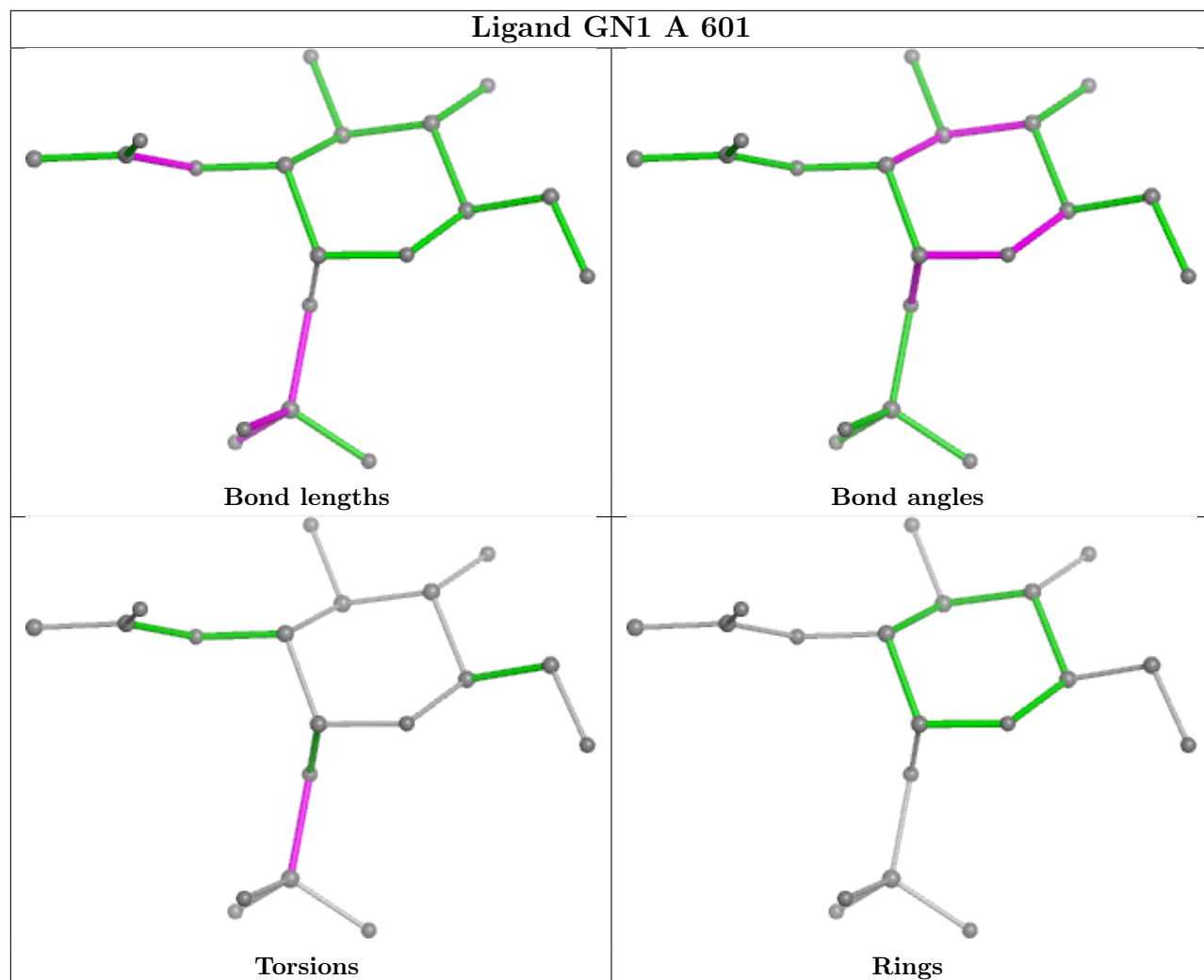
No monomer is involved in short contacts.

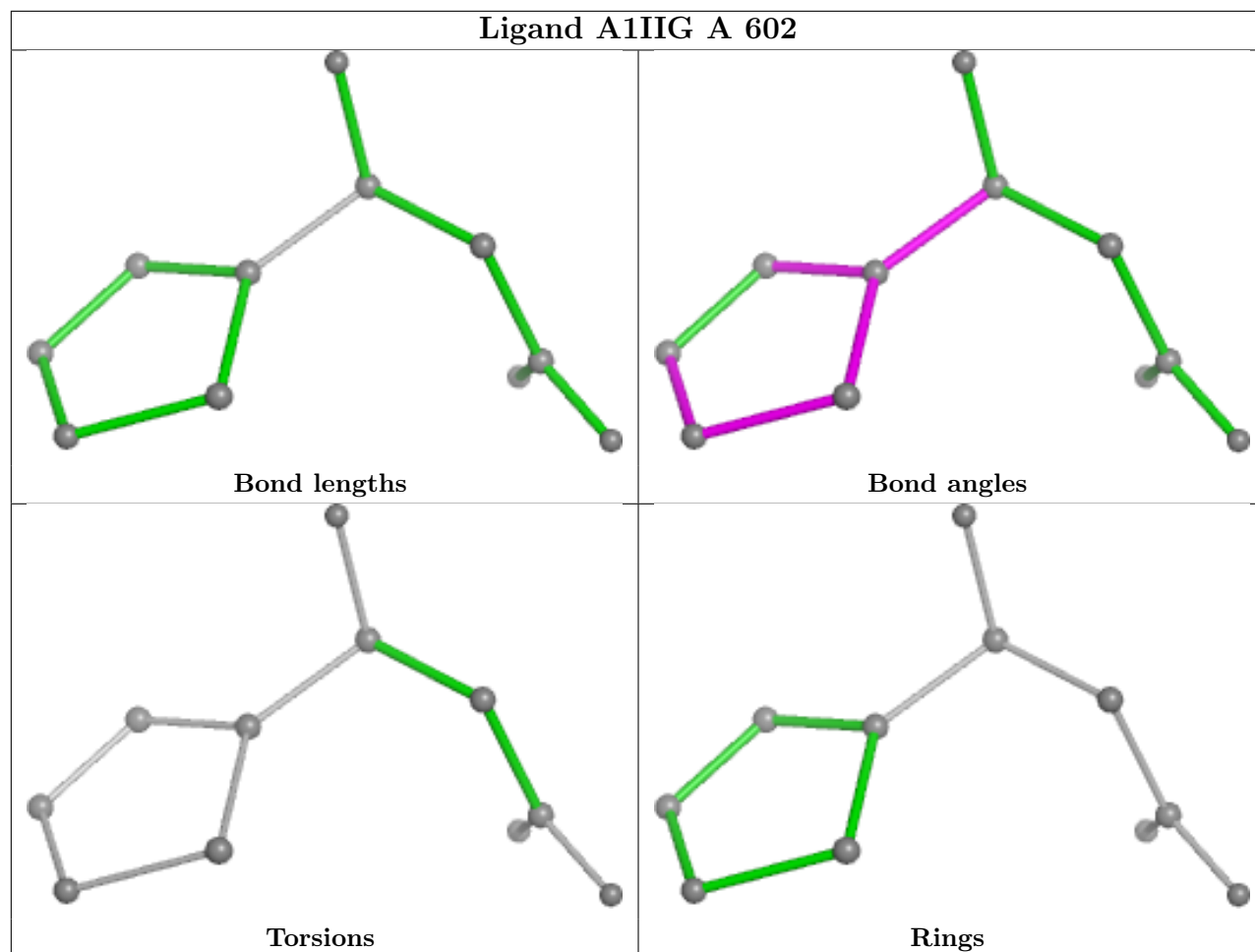
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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





## Ligand GN1 A 601





## 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	472/484 (97%)	-0.95	0 100 100	21, 40, 66, 80	4 (0%)
1	B	471/484 (97%)	-0.94	0 100 100	22, 40, 66, 76	1 (0%)
2	C	467/484 (96%)	-0.86	0 100 100	25, 44, 70, 95	1 (0%)
2	D	462/484 (95%)	-0.86	0 100 100	24, 44, 67, 99	2 (0%)
All	All	1872/1936 (96%)	-0.90	0 100 100	21, 42, 67, 99	8 (0%)

There are no RSRZ outliers to report.

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

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
1	YCM	B	277	10/11	0.99	0.05	28,29,34,37	0
1	YCM	A	277	10/11	0.99	0.06	30,32,37,38	0

### 6.3 Carbohydrates [i](#)

There are no oligosaccharides 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



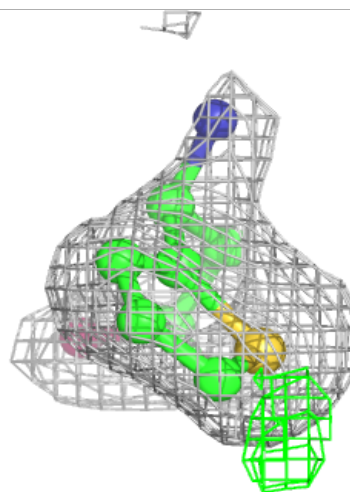
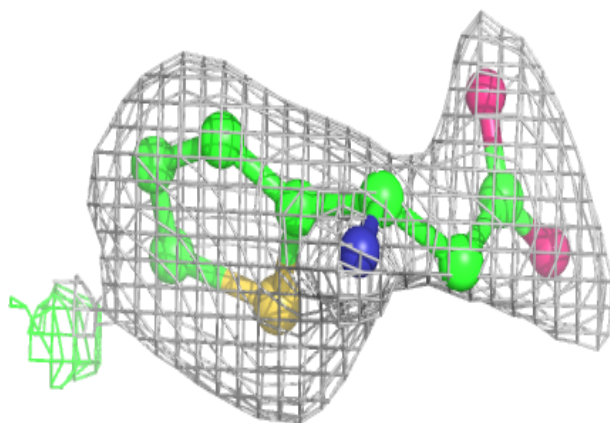
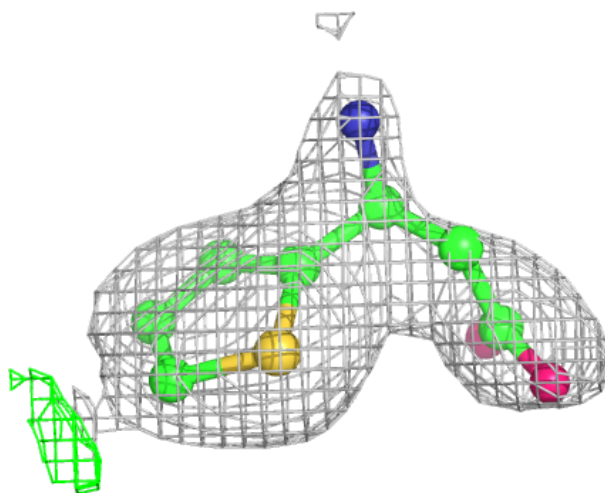
labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	71G	B	603	9/9	0.98	0.05	35,37,41,41	0
5	71G	D	602	9/9	0.98	0.05	33,40,47,52	0
6	MG	B	604	1/1	0.98	0.04	57,57,57,57	0
4	A1IIG	A	602	11/11	0.99	0.04	37,44,54,56	0
3	GN1	B	601	19/19	0.99	0.04	28,31,39,40	0
5	71G	D	601	9/9	0.99	0.04	34,39,43,44	0
3	GN1	A	601	19/19	0.99	0.04	27,31,41,42	0
4	A1IIG	B	602	11/11	0.99	0.04	41,43,55,56	0
7	CL	C	601	1/1	0.99	0.04	60,60,60,60	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.

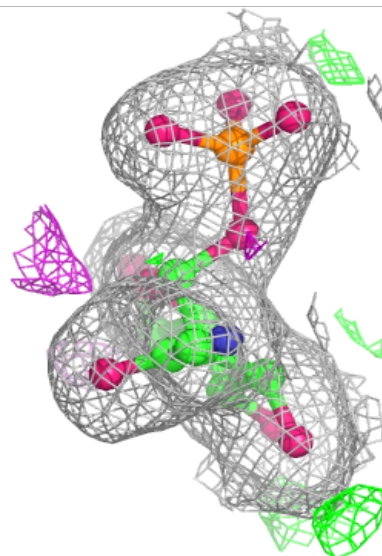
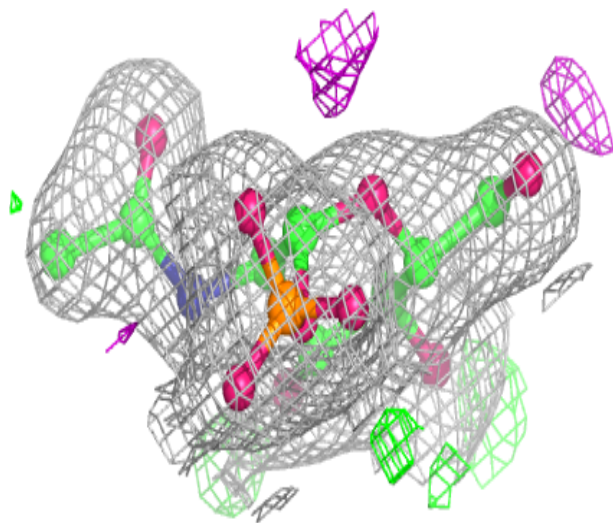
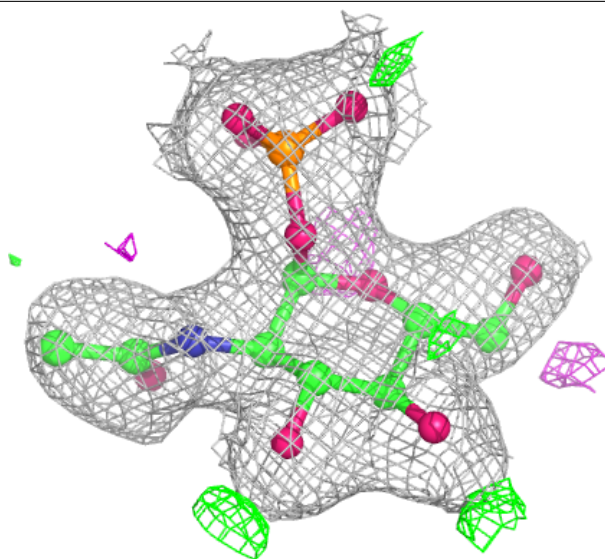
#### Electron density around A1IIG A 602:

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)



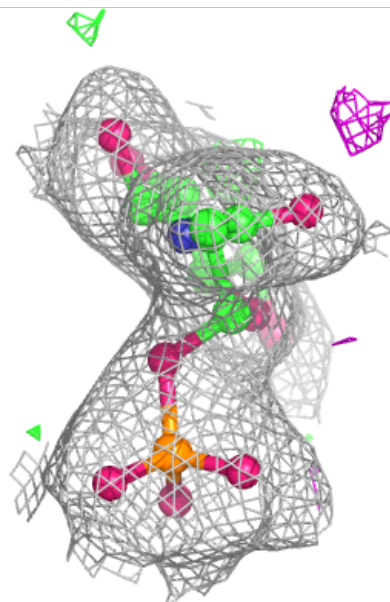
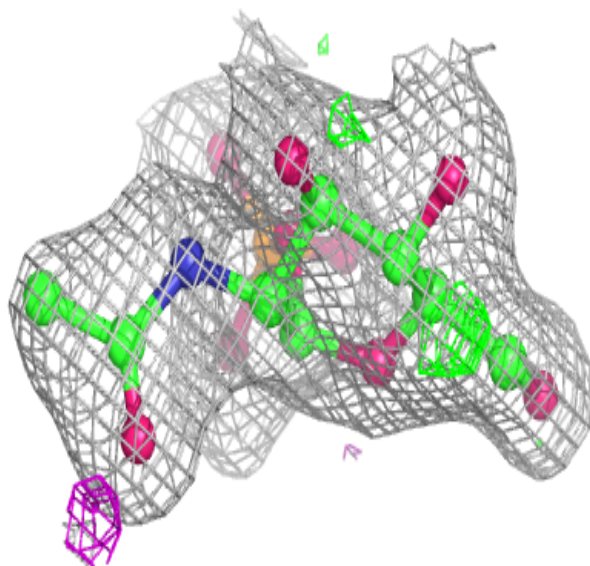
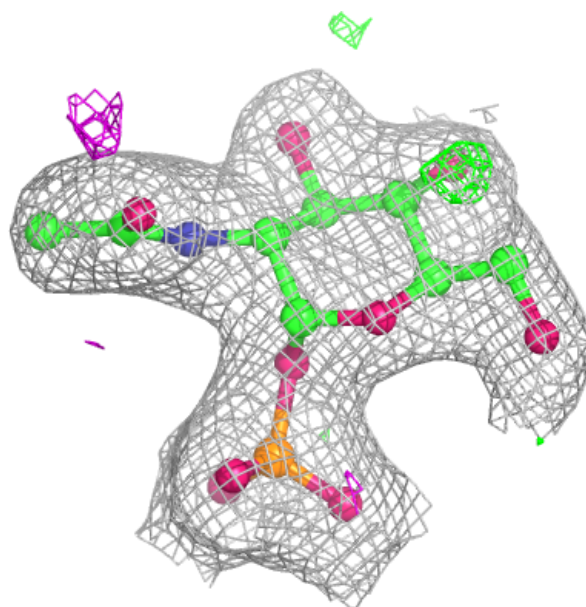
**Electron density around GN1 B 601:**

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



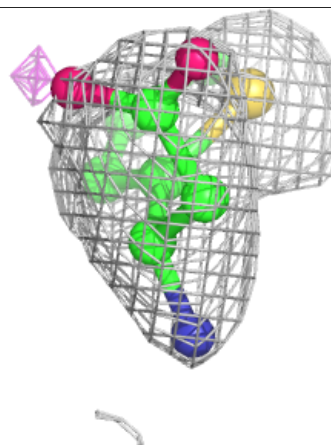
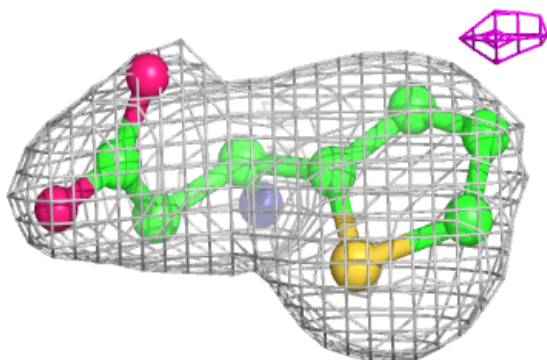
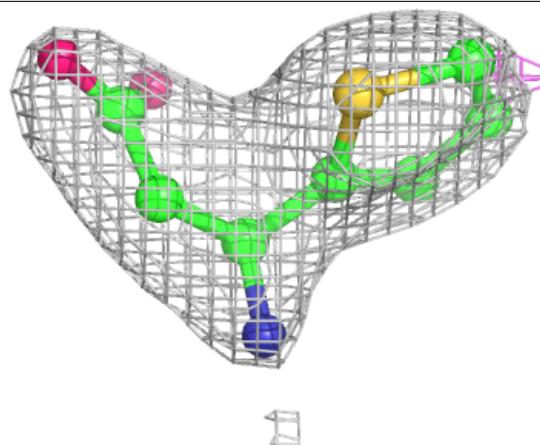
**Electron density around GN1 A 601:**

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



**Electron density around A1IIG B 602:**

$2mF_o - DF_c$  (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.