



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 9, 2020 – 03:00 PM BST

PDB ID : 5KV8  
Title : Crystal structure of a hPIV haemagglutinin-neuraminidase-inhibitor complex  
Authors : Dirr, L.; El-Deeb, I.M.; Chavas, L.M.G.; Guillon, P.; von Itzstein, M.  
Deposited on : 2016-07-13  
Resolution : 1.95 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
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) : 2.13.1

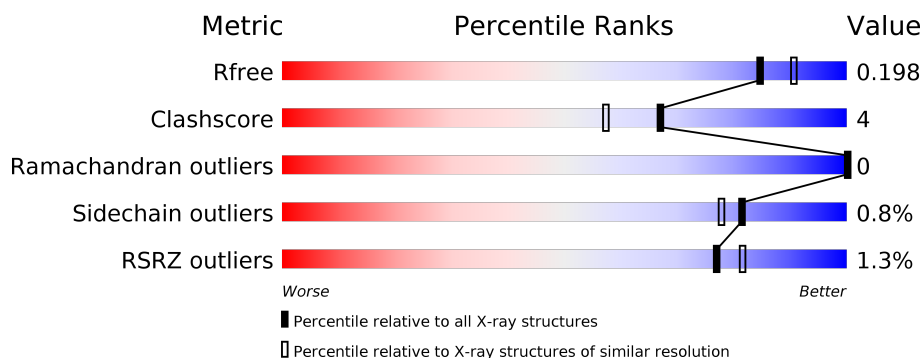
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.95 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	4310 (1.96-1.92)
Clashscore	141614	1023 (1.94-1.94)
Ramachandran outliers	138981	1007 (1.94-1.94)
Sidechain outliers	138945	1007 (1.94-1.94)
RSRZ outliers	127900	4250 (1.96-1.92)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	437	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span>%</span> <span>91%</span> <span>7% •</span> </div> </div>
1	B	437	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span>2%</span> <span>91%</span> <span>6% •</span> </div> </div>

## 2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 7608 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.

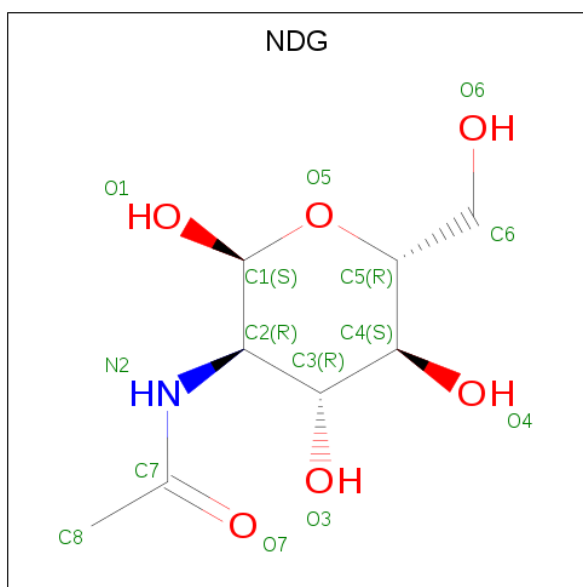
- Molecule 1 is a protein called Hemagglutinin-neuraminidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	430	Total	C	N	O	S	0	2	0
			3384	2139	582	643	20			
1	B	424	Total	C	N	O	S	0	0	0
			3326	2105	572	629	20			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	573	HIS	-	expression tag	UNP G8G134
A	574	HIS	-	expression tag	UNP G8G134
A	575	HIS	-	expression tag	UNP G8G134
A	576	HIS	-	expression tag	UNP G8G134
A	577	HIS	-	expression tag	UNP G8G134
A	578	HIS	-	expression tag	UNP G8G134
B	573	HIS	-	expression tag	UNP G8G134
B	574	HIS	-	expression tag	UNP G8G134
B	575	HIS	-	expression tag	UNP G8G134
B	576	HIS	-	expression tag	UNP G8G134
B	577	HIS	-	expression tag	UNP G8G134
B	578	HIS	-	expression tag	UNP G8G134

- Molecule 2 is 2-acetamido-2-deoxy-alpha-D-glucopyranose (three-letter code: NDG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 3 is beta-D-mannopyranose (three-letter code: BMA) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).

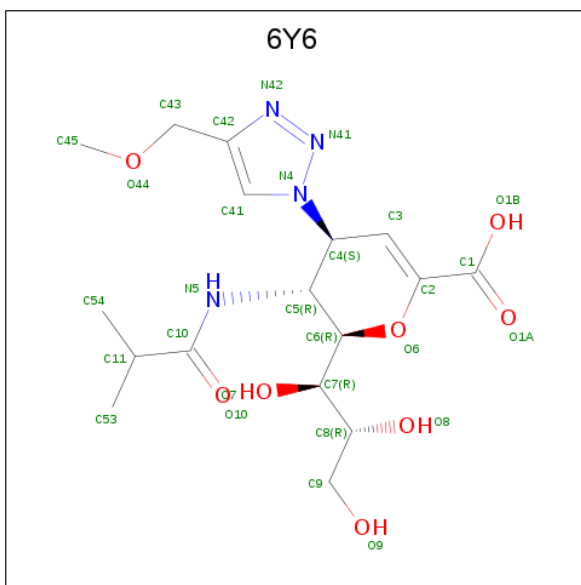


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			11	6	5		
3	B	1	Total	C	O	0	0
			11	6	5		

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Ca	0	0
			1	1		
4	A	1	Total	Ca	0	0
			1	1		

- Molecule 5 is 2,6-anhydro-3,4,5-trideoxy-4-[4-(methoxymethyl)-1H-1,2,3-triazol-1-yl]-5-[(2-methylpropanoyl)amino]-D-glycero-D-galacto -non-2-enonic acid (three-letter code: 6Y6) (formula: C<sub>17</sub>H<sub>26</sub>N<sub>4</sub>O<sub>8</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			29	17	4	8		
5	B	1	Total	C	N	O	0	0
			29	17	4	8		

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



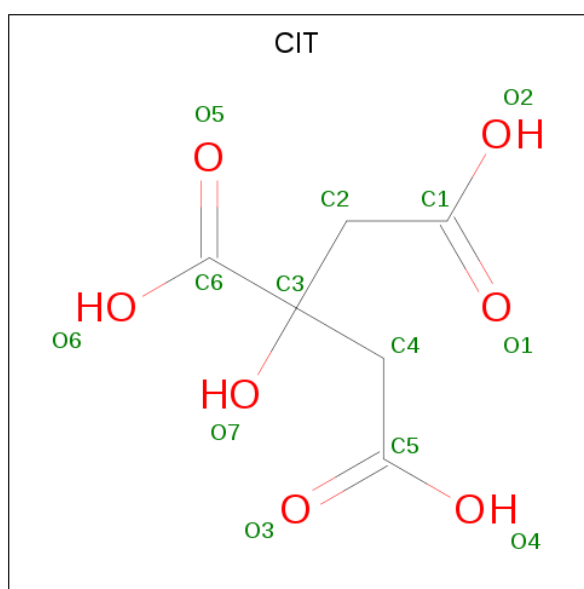
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			4	2	2		
6	A	1	Total	C	O	0	0
			4	2	2		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			4	2	2		
6	B	1	Total	C	O	0	0
			4	2	2		
6	B	1	Total	C	O	0	0
			4	2	2		
6	B	1	Total	C	O	0	0
			4	2	2		

- Molecule 7 is CITRIC ACID (three-letter code: CIT) (formula:  $C_6H_8O_7$ ).



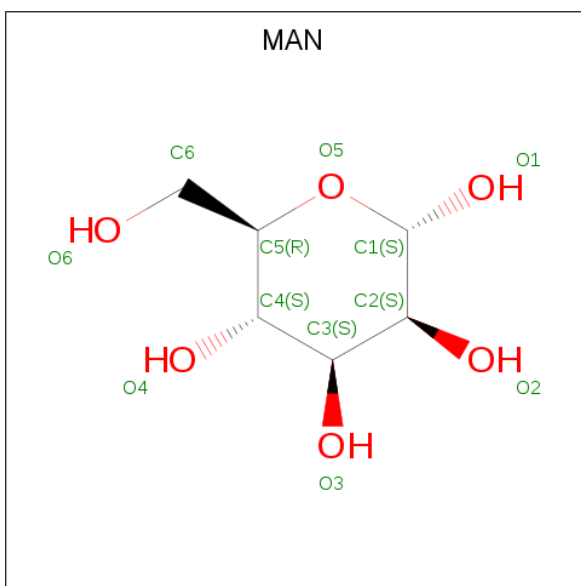
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			13	6	7		

- Molecule 8 is SULFATE ION (three-letter code: SO4) (formula:  $O_4S$ ).



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

- Molecule 9 is alpha-D-mannopyranose (three-letter code: MAN) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	B	1	Total	C	O	0	0
			11	6	5		

- Molecule 10 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
10	B	1	Total	C	O	0	0
			6	3	3		

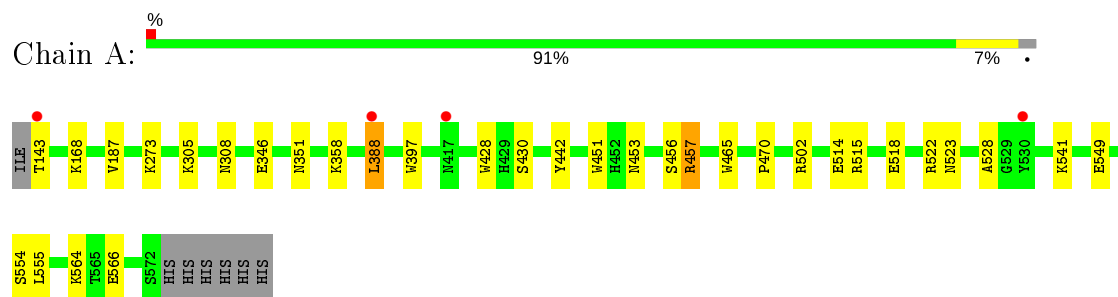
- Molecule 11 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	353	Total	O	0	0
			353	353		
11	B	292	Total	O	0	0
			292	292		

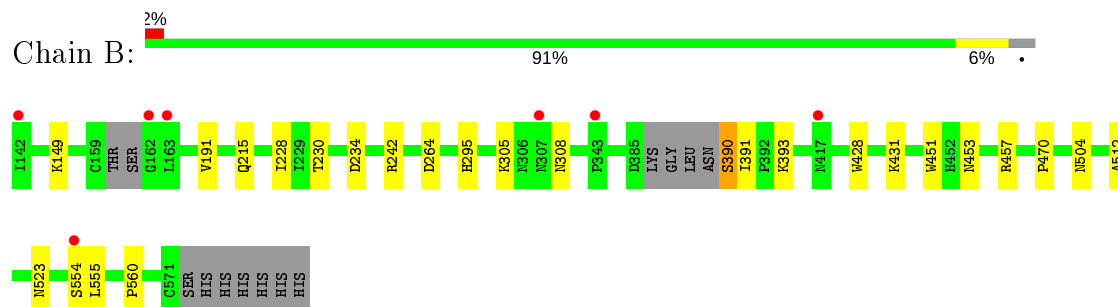
### 3 Residue-property plots [i](#)

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: Hemagglutinin-neuraminidase



#### • Molecule 1: Hemagglutinin-neuraminidase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	81.57Å 98.61Å 103.39Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.78 – 1.95 45.78 – 1.95	Depositor EDS
% Data completeness (in resolution range)	99.9 (45.78-1.95) 99.9 (45.78-1.95)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.97 (at 1.95Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.161 , 0.200 0.160 , 0.198	Depositor DCC
$R_{free}$ test set	3137 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.9	Xtriage
Anisotropy	0.212	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 53.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.013 for -h,l,k	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	7608	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.62% 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: GOL, BMA, 6Y6, CA, EDO, NDG, SO4, CIT, MAN

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	0.41	0/3466	0.58	1/4724 (0.0%)
1	B	0.40	0/3405	0.55	0/4639
All	All	0.41	0/6871	0.57	1/9363 (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	A	388	LEU	CA-CB-CG	5.08	126.99	115.30

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	A	3384	0	3337	26	1
1	B	3326	0	3285	18	1
2	A	56	0	48	8	0
2	B	56	0	48	7	1
3	A	11	0	10	1	0
3	B	11	0	10	1	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	29	0	0	0	0
5	B	29	0	0	0	0
6	A	12	0	18	0	0
6	B	12	0	18	0	0
7	A	13	0	5	1	0
8	B	5	0	0	1	0
9	B	11	0	10	2	0
10	B	6	0	8	0	0
11	A	353	0	0	7	1
11	B	292	0	0	5	0
All	All	7608	0	6797	53	3

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 4.

All (53) 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:A:555:LEU:O	11:A:701:HOH:O	1.92	0.87
1:A:143:THR:N	11:A:703:HOH:O	2.09	0.85
1:B:215:GLN:NE2	11:B:703:HOH:O	2.15	0.79
2:B:601:NDG:H4	2:B:602:NDG:C1	2.13	0.78
1:A:554:SER:OG	8:B:610:SO4:O4	1.99	0.77
1:A:351:ASN:HD21	2:A:601:NDG:C1	2.00	0.74
1:B:393:LYS:NZ	11:B:704:HOH:O	2.20	0.73
1:A:523:ASN:HD21	2:A:603:NDG:C1	2.02	0.72
1:B:308:ASN:HD21	2:B:611:NDG:C1	2.01	0.72
2:B:601:NDG:O3	11:B:701:HOH:O	2.10	0.70
1:A:308:ASN:HD21	2:A:610:NDG:C1	2.05	0.68
2:B:603:NDG:O6	11:B:702:HOH:O	2.13	0.65
2:B:601:NDG:H8C1	2:B:601:NDG:H3	1.78	0.65
2:A:601:NDG:H4	2:A:602:NDG:C1	2.31	0.60
1:A:305:LYS:O	1:A:308:ASN:HB2	2.01	0.59
1:B:554:SER:OG	1:B:555:LEU:N	2.36	0.57
1:B:305:LYS:O	1:B:308:ASN:HB2	2.06	0.55
1:B:191:VAL:HB	1:B:560:PRO:HD3	1.88	0.54
1:B:428:TRP:CD2	1:B:470:PRO:HA	2.42	0.54
1:B:228:ILE:HD11	1:B:242:ARG:NH2	2.22	0.54
2:A:602:NDG:H4	3:A:604:BMA:C1	2.37	0.54
3:B:604:BMA:O3	9:B:612:MAN:H3	2.09	0.53
1:A:168:LYS:NZ	1:A:514:GLU:OE2	2.43	0.52

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:549:GLU:OE2	11:A:705:HOH:O	2.18	0.51
1:A:351:ASN:ND2	2:A:601:NDG:C1	2.73	0.50
2:A:601:NDG:C4	2:A:602:NDG:C1	2.90	0.50
1:B:264:ASP:OD1	1:B:295:HIS:HD2	1.95	0.49
1:A:515:ARG:HD2	1:A:518:GLU:OE1	2.12	0.48
1:B:428:TRP:CG	1:B:470:PRO:HA	2.49	0.47
1:A:428:TRP:CG	1:A:470:PRO:HA	2.50	0.47
1:B:504:ASN:OD1	1:B:523:ASN:HA	2.15	0.46
1:A:541:LYS:NZ	1:A:566:GLU:OE1	2.39	0.46
1:A:451:TRP:CH2	1:A:453:ASN:HB2	2.51	0.46
1:A:428:TRP:CD2	1:A:470:PRO:HA	2.50	0.46
1:A:456:SER:HB3	1:A:465:TRP:CD2	2.50	0.46
1:B:512:ALA:HB3	9:B:612:MAN:H62	1.97	0.45
2:B:601:NDG:C8	2:B:601:NDG:H3	2.45	0.44
1:B:431:LYS:NZ	1:B:453:ASN:OD1	2.49	0.44
1:B:149:LYS:HE3	11:B:916:HOH:O	2.18	0.44
1:A:351:ASN:HD21	2:A:601:NDG:C2	2.30	0.44
1:A:397:TRP:CD2	1:A:442:TYR:HB3	2.53	0.44
1:A:187:VAL:HG22	11:A:887:HOH:O	2.17	0.43
1:B:390:SER:HB2	1:B:391:ILE:H	1.30	0.43
1:B:308:ASN:HD21	2:B:611:NDG:C2	2.30	0.43
1:A:522:ARG:HB3	1:A:523:ASN:H	1.65	0.42
1:A:273:LYS:HE3	11:A:937:HOH:O	2.19	0.42
1:A:346:GLU:HB3	11:A:774:HOH:O	2.20	0.41
1:A:502:ARG:HB3	1:A:528:ALA:O	2.21	0.41
1:A:430:SER:HB2	1:A:457:ARG:CZ	2.51	0.41
1:B:451:TRP:CH2	1:B:453:ASN:HB2	2.56	0.41
7:A:611:CIT:H42	1:B:230:THR:HB	2.03	0.41
1:A:358:LYS:HA	1:A:358:LYS:HD3	1.81	0.41
1:A:523:ASN:ND2	11:A:717:HOH:O	2.50	0.41

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:273:LYS:NZ	2:B:602:NDG:O6[2_544]	2.18	0.02
1:B:215:GLN:OE1	1:B:390:SER:OG[4_545]	2.18	0.02
11:A:972:HOH:O	11:A:994:HOH:O[2_554]	2.18	0.02

## 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	430/437 (98%)	412 (96%)	18 (4%)	0	100	100
1	B	418/437 (96%)	396 (95%)	22 (5%)	0	100	100
All	All	848/874 (97%)	808 (95%)	40 (5%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	390/395 (99%)	387 (99%)	3 (1%)	81	78
1	B	383/395 (97%)	380 (99%)	3 (1%)	81	78
All	All	773/790 (98%)	767 (99%)	6 (1%)	81	78

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	388	LEU
1	A	457	ARG
1	A	564	LYS
1	B	234	ASP
1	B	390	SER
1	B	457	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such

sidechains are listed below:

Mol	Chain	Res	Type
1	A	540	ASN
1	B	295	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

Of 24 ligands modelled in this entry, 2 are monoatomic - leaving 22 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
2	NDG	A	610	-	14,14,15	1.48	1 (7%)	17,19,21	0.80	1 (5%)
6	EDO	A	609	-	3,3,3	0.47	0	2,2,2	0.20	0
6	EDO	B	609	-	3,3,3	0.48	0	2,2,2	0.26	0
3	BMA	A	604	-	11,11,12	2.54	5 (45%)	15,15,17	3.14	4 (26%)
6	EDO	B	608	-	3,3,3	0.45	0	2,2,2	0.42	0
2	NDG	B	602	-	14,14,15	1.31	2 (14%)	17,19,21	1.63	3 (17%)
3	BMA	B	604	-	11,11,12	2.26	4 (36%)	15,15,17	2.02	3 (20%)
6	EDO	A	608	-	3,3,3	0.55	0	2,2,2	0.31	0
2	NDG	B	603	-	14,14,15	1.00	1 (7%)	17,19,21	0.76	0
2	NDG	B	611	-	14,14,15	1.46	1 (7%)	17,19,21	0.83	1 (5%)



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
9	MAN	B	612	-	11,11,12	1.89	3 (27%)	15,15,17	1.76	4 (26%)
2	NDG	A	601	-	14,14,15	2.43	2 (14%)	17,19,21	1.67	2 (11%)
2	NDG	B	601	-	14,14,15	1.59	1 (7%)	17,19,21	2.10	3 (17%)
5	6Y6	A	606	-	26,30,30	1.67	6 (23%)	24,42,42	1.66	5 (20%)
2	NDG	A	603	-	14,14,15	1.64	1 (7%)	17,19,21	1.15	2 (11%)
5	6Y6	B	606	-	26,30,30	1.67	5 (19%)	24,42,42	1.85	5 (20%)
7	CIT	A	611	-	3,12,12	1.02	0	3,17,17	1.38	0
8	SO4	B	610	-	4,4,4	0.15	0	6,6,6	0.09	0
6	EDO	A	607	-	3,3,3	0.51	0	2,2,2	0.42	0
10	GOL	B	613	-	5,5,5	0.29	0	5,5,5	0.56	0
2	NDG	A	602	-	14,14,15	1.94	2 (14%)	17,19,21	1.70	2 (11%)
6	EDO	B	607	-	3,3,3	0.43	0	2,2,2	0.45	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	EDO	A	608	-	-	1/1/1/1	-
2	NDG	A	610	-	-	1/6/23/26	0/1/1/1
6	EDO	A	609	-	-	1/1/1/1	-
2	NDG	A	603	-	-	2/6/23/26	0/1/1/1
6	EDO	B	609	-	-	1/1/1/1	-
2	NDG	B	603	-	-	0/6/23/26	0/1/1/1
2	NDG	B	611	-	-	0/6/23/26	0/1/1/1
10	GOL	B	613	-	-	0/4/4/4	-
9	MAN	B	612	-	-	0/2/19/22	0/1/1/1
3	BMA	A	604	-	-	1/2/19/22	0/1/1/1
6	EDO	B	608	-	-	0/1/1/1	-
2	NDG	A	602	-	-	2/6/23/26	0/1/1/1
6	EDO	B	607	-	-	1/1/1/1	-
5	6Y6	B	606	-	-	0/19/45/45	0/2/2/2
2	NDG	B	601	-	-	5/6/23/26	0/1/1/1
2	NDG	A	601	-	-	2/6/23/26	0/1/1/1
2	NDG	B	602	-	-	1/6/23/26	0/1/1/1
5	6Y6	A	606	-	-	0/19/45/45	0/2/2/2
7	CIT	A	611	-	-	0/6/16/16	-
3	BMA	B	604	-	-	1/2/19/22	0/1/1/1
6	EDO	A	607	-	-	1/1/1/1	-

All (34) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	601	NDG	O5-C1	-7.84	1.31	1.43
2	A	602	NDG	O5-C1	-5.84	1.34	1.43
2	A	603	NDG	O5-C1	-5.81	1.34	1.43
2	B	601	NDG	O5-C1	-5.44	1.35	1.43
2	B	611	NDG	O5-C1	-5.24	1.35	1.43
2	A	610	NDG	O5-C1	-5.16	1.35	1.43
3	A	604	BMA	C4-C5	5.14	1.63	1.53
5	A	606	6Y6	O6-C2	4.78	1.45	1.37
5	B	606	6Y6	O6-C2	4.78	1.45	1.37
3	B	604	BMA	C4-C5	4.65	1.62	1.53
2	A	601	NDG	C1-C2	-4.14	1.46	1.52
3	A	604	BMA	C4-C3	4.13	1.62	1.52
3	B	604	BMA	O5-C1	-4.00	1.37	1.43
2	A	602	NDG	C1-C2	-3.78	1.46	1.52
9	B	612	MAN	O5-C5	3.68	1.50	1.43
2	B	603	NDG	O5-C1	-3.59	1.38	1.43
2	B	602	NDG	O5-C1	-3.42	1.38	1.43
9	B	612	MAN	C1-C2	3.36	1.59	1.52
5	B	606	6Y6	N42-N41	3.35	1.40	1.34
3	A	604	BMA	C2-C3	-3.23	1.47	1.52
2	B	602	NDG	C1-C2	-3.22	1.47	1.52
5	A	606	6Y6	N42-N41	3.11	1.40	1.34
5	B	606	6Y6	C10-N5	-2.93	1.27	1.34
3	B	604	BMA	C2-C3	-2.92	1.48	1.52
5	A	606	6Y6	C10-N5	-2.87	1.27	1.34
5	B	606	6Y6	N41-N4	2.83	1.40	1.34
5	A	606	6Y6	N41-N4	2.67	1.39	1.34
5	A	606	6Y6	C3-C2	2.58	1.35	1.32
3	B	604	BMA	C4-C3	2.50	1.58	1.52
5	A	606	6Y6	C41-N4	-2.39	1.33	1.35
5	B	606	6Y6	C41-N4	-2.37	1.33	1.35
3	A	604	BMA	O5-C1	-2.37	1.39	1.43
3	A	604	BMA	C1-C2	-2.37	1.46	1.52
9	B	612	MAN	C2-C3	2.18	1.55	1.52

All (35) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	604	BMA	C1-C2-C3	-10.78	96.42	109.67
5	B	606	6Y6	C41-N4-C4	6.18	132.12	125.58
5	A	606	6Y6	C41-N4-C4	5.39	131.27	125.58
2	B	601	NDG	C2-N2-C7	5.35	130.53	122.90

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	602	NDG	C3-C4-C5	5.24	119.59	110.24
2	B	601	NDG	C3-C4-C5	5.02	119.19	110.24
2	A	601	NDG	C3-C4-C5	4.93	119.03	110.24
3	B	604	BMA	C1-O5-C5	-4.26	106.42	112.19
2	A	601	NDG	C1-O5-C5	-4.20	106.50	112.19
3	B	604	BMA	C3-C4-C5	3.88	117.15	110.24
2	B	602	NDG	C3-C4-C5	3.83	117.06	110.24
9	B	612	MAN	C1-O5-C5	3.77	117.30	112.19
2	A	602	NDG	C1-O5-C5	-3.76	107.10	112.19
3	A	604	BMA	C1-O5-C5	-3.59	107.33	112.19
2	B	602	NDG	C1-O5-C5	-3.55	107.39	112.19
9	B	612	MAN	O5-C1-C2	3.48	116.14	110.77
3	B	604	BMA	C1-C2-C3	-3.30	105.61	109.67
5	B	606	6Y6	O6-C2-C3	-3.29	119.88	124.33
2	B	602	NDG	O4-C4-C3	-3.00	103.41	110.35
2	A	603	NDG	O3-C3-C2	-2.98	103.30	109.47
5	A	606	6Y6	O6-C2-C3	-2.78	120.57	124.33
5	A	606	6Y6	C43-C42-C41	-2.66	123.42	128.45
5	B	606	6Y6	C8-C7-C6	-2.55	108.20	113.03
5	B	606	6Y6	C6-C5-N5	-2.51	106.75	110.91
3	A	604	BMA	O3-C3-C4	2.48	116.09	110.35
5	B	606	6Y6	C43-C42-C41	-2.47	123.79	128.45
9	B	612	MAN	O2-C2-C1	2.43	114.11	109.15
2	A	603	NDG	C3-C4-C5	2.42	114.56	110.24
9	B	612	MAN	O3-C3-C2	2.41	114.60	109.99
2	B	611	NDG	C3-C4-C5	2.18	114.12	110.24
2	B	601	NDG	C1-C2-N2	2.17	114.19	110.49
2	A	610	NDG	C3-C4-C5	2.16	114.10	110.24
5	A	606	6Y6	C6-C5-N5	-2.15	107.34	110.91
5	A	606	6Y6	C8-C7-C6	-2.02	109.19	113.03
3	A	604	BMA	C6-C5-C4	2.00	117.70	113.00

There are no chirality outliers.

All (20) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	603	NDG	O5-C5-C6-O6
2	A	601	NDG	C4-C5-C6-O6
2	A	603	NDG	C4-C5-C6-O6
2	A	601	NDG	O5-C5-C6-O6
2	B	601	NDG	C8-C7-N2-C2
2	B	601	NDG	O7-C7-N2-C2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
3	B	604	BMA	O5-C5-C6-O6
2	A	602	NDG	C4-C5-C6-O6
2	A	602	NDG	O5-C5-C6-O6
3	A	604	BMA	O5-C5-C6-O6
6	B	609	EDO	O1-C1-C2-O2
2	B	601	NDG	C4-C5-C6-O6
2	B	601	NDG	O5-C5-C6-O6
6	A	609	EDO	O1-C1-C2-O2
6	A	608	EDO	O1-C1-C2-O2
2	A	610	NDG	C4-C5-C6-O6
2	B	601	NDG	C3-C2-N2-C7
2	B	602	NDG	C4-C5-C6-O6
6	B	607	EDO	O1-C1-C2-O2
6	A	607	EDO	O1-C1-C2-O2

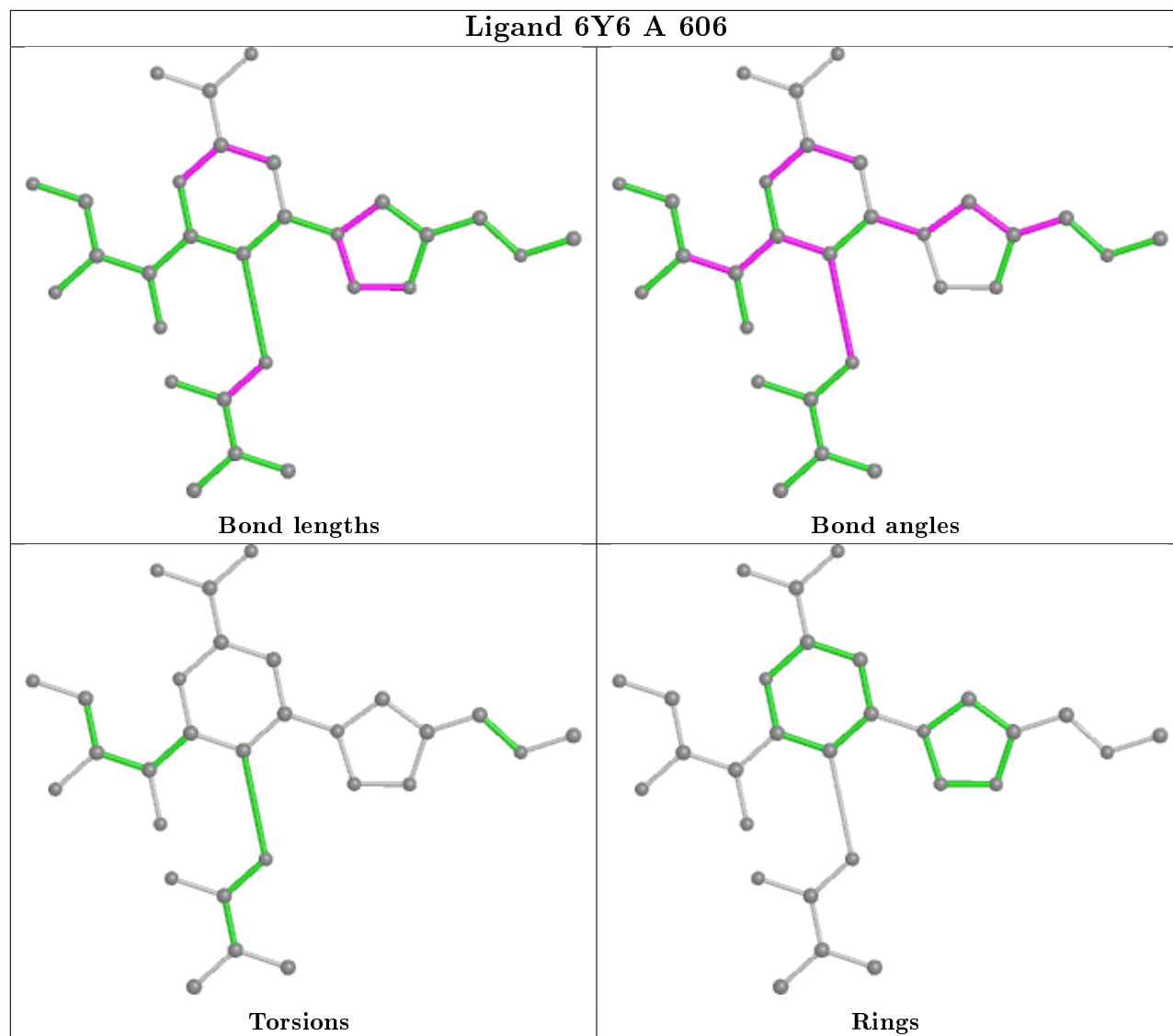
There are no ring outliers.

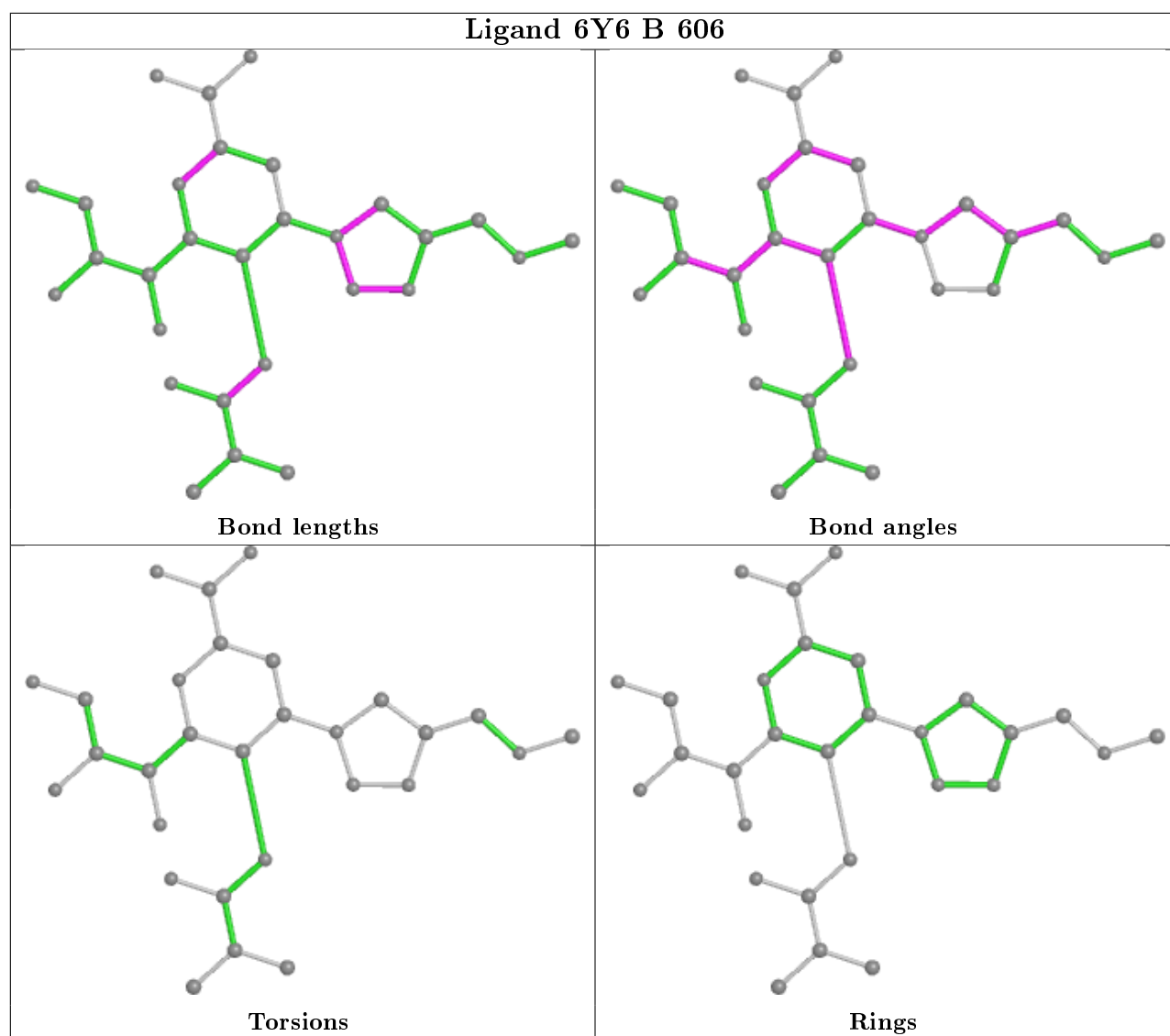
13 monomers are involved in 20 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	610	NDG	1	0
3	A	604	BMA	1	0
2	B	602	NDG	1	1
3	B	604	BMA	1	0
2	B	603	NDG	1	0
2	B	611	NDG	2	0
9	B	612	MAN	2	0
2	A	601	NDG	5	0
2	B	601	NDG	4	0
2	A	603	NDG	1	0
7	A	611	CIT	1	0
8	B	610	SO4	1	0
2	A	602	NDG	3	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	430/437 (98%)	-0.25	4 (0%) 84 87	15, 23, 39, 58	0
1	B	424/437 (97%)	-0.19	7 (1%) 70 75	16, 25, 44, 60	0
All	All	854/874 (97%)	-0.22	11 (1%) 77 81	15, 24, 42, 60	0

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	163	LEU	4.4
1	B	307	ASN	3.8
1	A	143	THR	3.7
1	A	388	LEU	3.4
1	A	417	ASN	3.2
1	B	554	SER	3.0
1	B	417	ASN	2.5
1	B	162	GLY	2.4
1	A	530[A]	TYR	2.4
1	B	343	PRO	2.3
1	B	142	ILE	2.1

### 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 ⓘ

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.

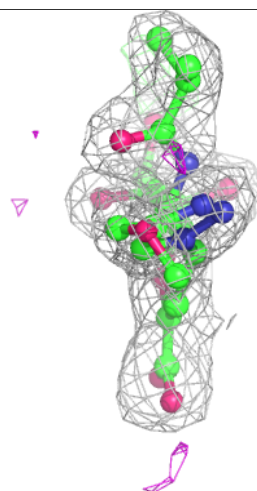
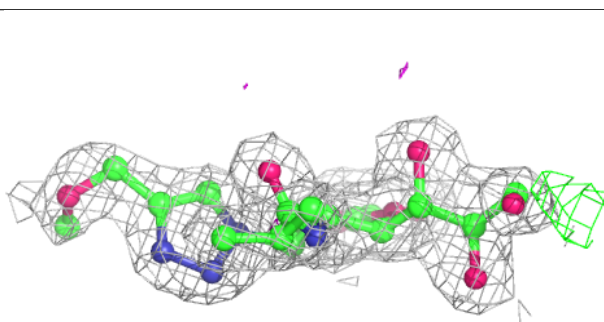
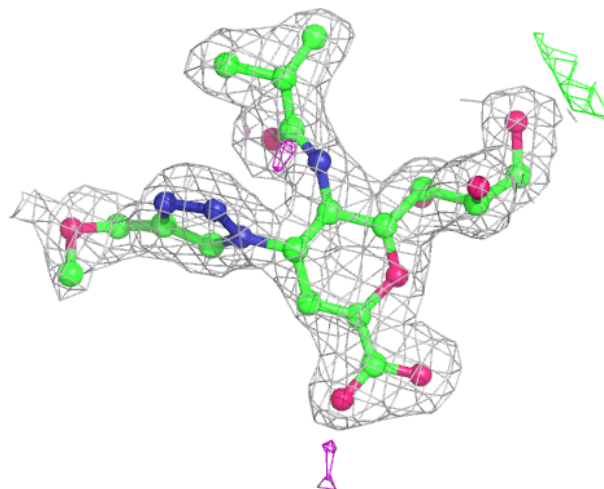
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	BMA	A	604	11/12	0.58	0.27	55,61,68,68	0
3	BMA	B	604	11/12	0.72	0.22	38,52,62,63	0
9	MAN	B	612	11/12	0.73	0.32	45,48,61,66	0
2	NDG	B	601	14/15	0.74	0.21	40,52,63,73	0
2	NDG	A	603	14/15	0.76	0.28	42,52,65,66	0
2	NDG	A	610	14/15	0.77	0.32	41,59,64,65	0
2	NDG	A	602	14/15	0.78	0.17	39,44,54,63	0
2	NDG	B	603	14/15	0.81	0.23	40,50,56,60	0
2	NDG	B	611	14/15	0.83	0.36	46,54,64,68	0
2	NDG	B	602	14/15	0.83	0.18	37,44,56,58	0
6	EDO	A	609	4/4	0.83	0.27	46,48,48,48	0
2	NDG	A	601	14/15	0.85	0.19	39,42,51,66	0
6	EDO	A	608	4/4	0.85	0.15	34,35,36,39	0
7	CIT	A	611	13/13	0.88	0.17	30,39,51,61	0
6	EDO	B	607	4/4	0.88	0.13	34,36,37,38	0
6	EDO	B	609	4/4	0.89	0.15	38,43,46,52	0
6	EDO	B	608	4/4	0.91	0.11	35,39,40,59	0
5	6Y6	B	606	29/29	0.91	0.14	24,31,43,48	0
5	6Y6	A	606	29/29	0.94	0.11	21,28,41,49	0
6	EDO	A	607	4/4	0.94	0.10	27,35,37,42	0
10	GOL	B	613	6/6	0.96	0.09	23,27,30,31	0
8	SO4	B	610	5/5	0.97	0.15	32,43,52,52	0
4	CA	A	605	1/1	0.99	0.04	23,23,23,23	0
4	CA	B	605	1/1	0.99	0.06	27,27,27,27	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.



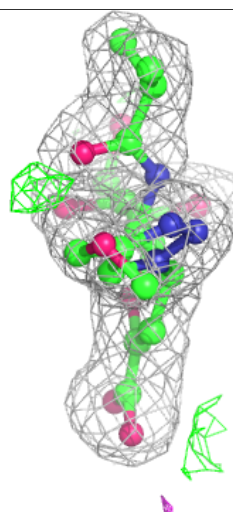
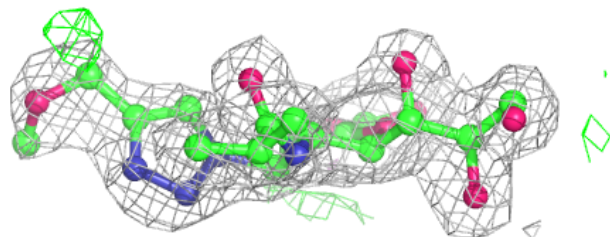
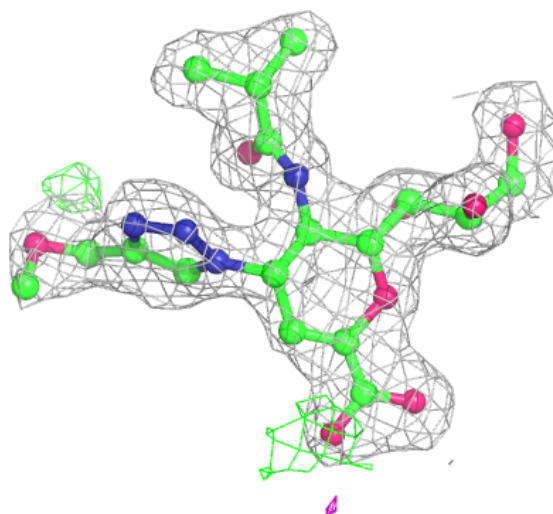
**Electron density around 6Y6 B 606:**

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



**Electron density around 6Y6 A 606:**

$2mF_o-DF_c$  (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.