



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 19, 2024 – 11:18 AM EDT

PDB ID : 4XPH  
Title : X-ray structure of Drosophila dopamine transporter with subsiteB mutations (D121G/S426M) bound to 3,4dichlorophenethylamine  
Authors : Penmatsa, A.; Wang, K.; Gouaux, E.  
Deposited on : 2015-01-17  
Resolution : 2.90 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 : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 1.20.1  
EDS : 3.0  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.003 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.39

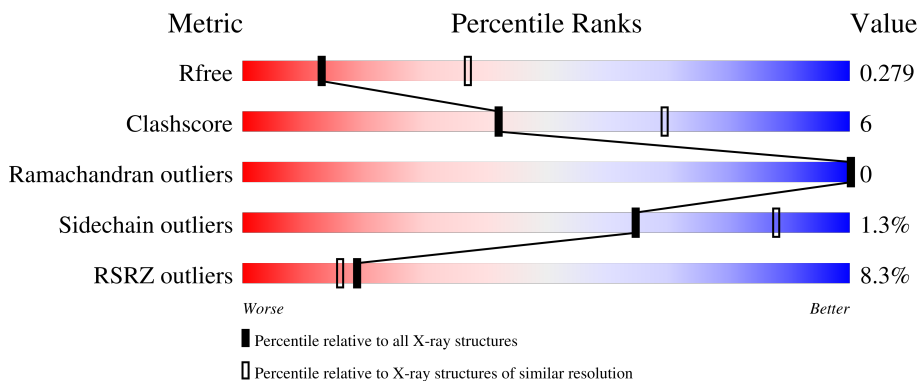
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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	2335 (2.90-2.90)
Clashscore	180529	2564 (2.90-2.90)
Ramachandran outliers	177936	2514 (2.90-2.90)
Sidechain outliers	177891	2516 (2.90-2.90)
RSRZ outliers	164620	2337 (2.90-2.90)

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	534	 10% 83% 16%
2	L	214	 6% 88% 11%
3	H	219	 8% 86% 14%
4	B	2	 50% 50%

## 2 Entry composition [i](#)

There are 10 unique types of molecules in this entry. The entry contains 7605 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 Transporter.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	534	4229	2836	657	716	20	0	1	0

There are 45 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	74	ALA	VAL	engineered mutation	UNP Q9NB97
A	121	GLY	ASP	engineered mutation	UNP Q9NB97
A	?	-	SER	deletion	UNP Q9NB97
A	?	-	GLN	deletion	UNP Q9NB97
A	?	-	ASN	deletion	UNP Q9NB97
A	?	-	ALA	deletion	UNP Q9NB97
A	?	-	SER	deletion	UNP Q9NB97
A	?	-	ARG	deletion	UNP Q9NB97
A	?	-	VAL	deletion	UNP Q9NB97
A	?	-	PRO	deletion	UNP Q9NB97
A	?	-	VAL	deletion	UNP Q9NB97
A	?	-	ILE	deletion	UNP Q9NB97
A	?	-	GLY	deletion	UNP Q9NB97
A	?	-	ASN	deletion	UNP Q9NB97
A	?	-	TYR	deletion	UNP Q9NB97
A	?	-	SER	deletion	UNP Q9NB97
A	?	-	ASP	deletion	UNP Q9NB97
A	?	-	LEU	deletion	UNP Q9NB97
A	?	-	TYR	deletion	UNP Q9NB97
A	?	-	ALA	deletion	UNP Q9NB97
A	?	-	MET	deletion	UNP Q9NB97
A	?	-	GLY	deletion	UNP Q9NB97
A	?	-	ASN	deletion	UNP Q9NB97
A	?	-	GLN	deletion	UNP Q9NB97
A	?	-	SER	deletion	UNP Q9NB97
A	?	-	LEU	deletion	UNP Q9NB97
A	?	-	LEU	deletion	UNP Q9NB97

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	TYR	deletion	UNP Q9NB97
A	?	-	ASN	deletion	UNP Q9NB97
A	?	-	GLU	deletion	UNP Q9NB97
A	?	-	THR	deletion	UNP Q9NB97
A	?	-	TYR	deletion	UNP Q9NB97
A	?	-	MET	deletion	UNP Q9NB97
A	?	-	ASN	deletion	UNP Q9NB97
A	?	-	GLY	deletion	UNP Q9NB97
A	?	-	SER	deletion	UNP Q9NB97
A	?	-	SER	deletion	UNP Q9NB97
A	?	-	LEU	deletion	UNP Q9NB97
A	?	-	ASP	deletion	UNP Q9NB97
A	?	-	THR	deletion	UNP Q9NB97
A	?	-	SER	deletion	UNP Q9NB97
A	?	-	ALA	deletion	UNP Q9NB97
A	?	-	VAL	deletion	UNP Q9NB97
A	415	ALA	LEU	engineered mutation	UNP Q9NB97
A	426	MET	SER	engineered mutation	UNP Q9NB97

- Molecule 2 is a protein called Antibody fragment light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	L	214	1630	1013	271	338	8	0	0	0

- Molecule 3 is a protein called Antibody fragment heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	H	219	1626	1023	275	320	8	0	0	0

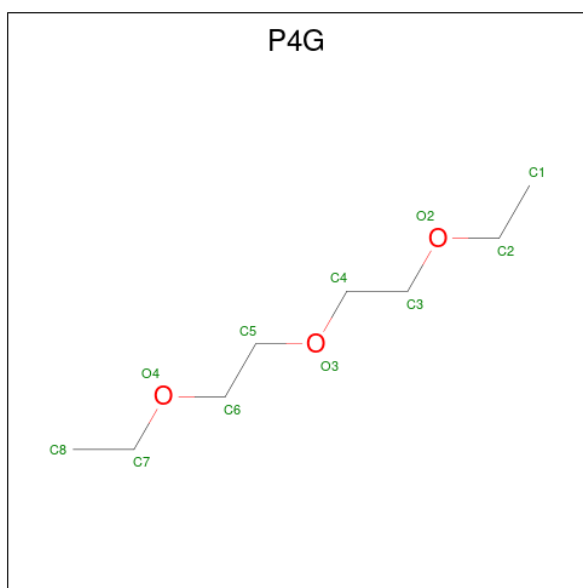
- Molecule 4 is an oligosaccharide called alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
			Total	C	O			
4	B	2	23	12	11	0	0	0

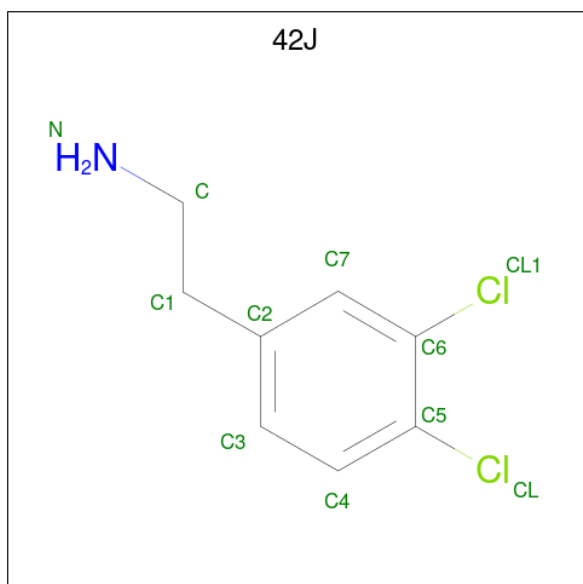
- Molecule 5 is 1-ETHOXY-2-(2-ETHOXYETHOXY)ETHANE (three-letter code: P4G)

(formula:  $C_8H_{18}O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			11	8	3		

- Molecule 6 is 2-(3,4-dichlorophenyl)ethanamine (three-letter code: 42J) (formula:  $C_8H_9Cl_2N$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	Cl	N	0	0
			11	8	2	1		

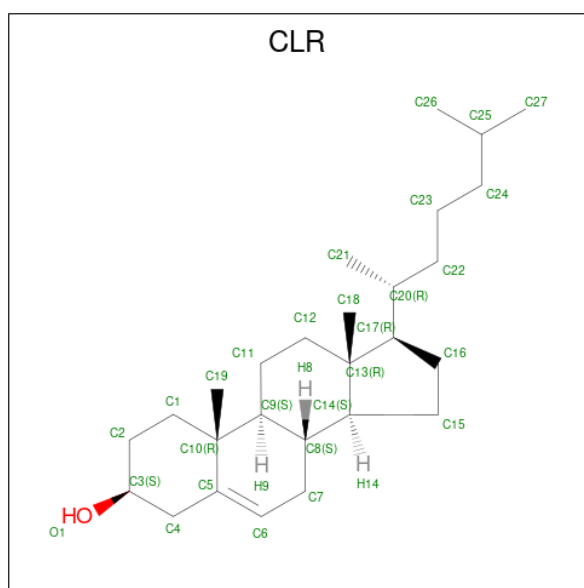
- Molecule 7 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	2	Total Na 2 2	0	0
7	L	1	Total Na 1 1	0	0

- Molecule 8 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	1	Total Cl 1 1	0	0

- Molecule 9 is CHOLESTEROL (three-letter code: CLR) (formula: C<sub>27</sub>H<sub>46</sub>O).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	1	Total C O 28 27 1	0	0
9	A	1	Total C O 28 27 1	0	0

- Molecule 10 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	A	9	Total O 9 9	0	0
10	L	4	Total O 4 4	0	0

*Continued on next page...*

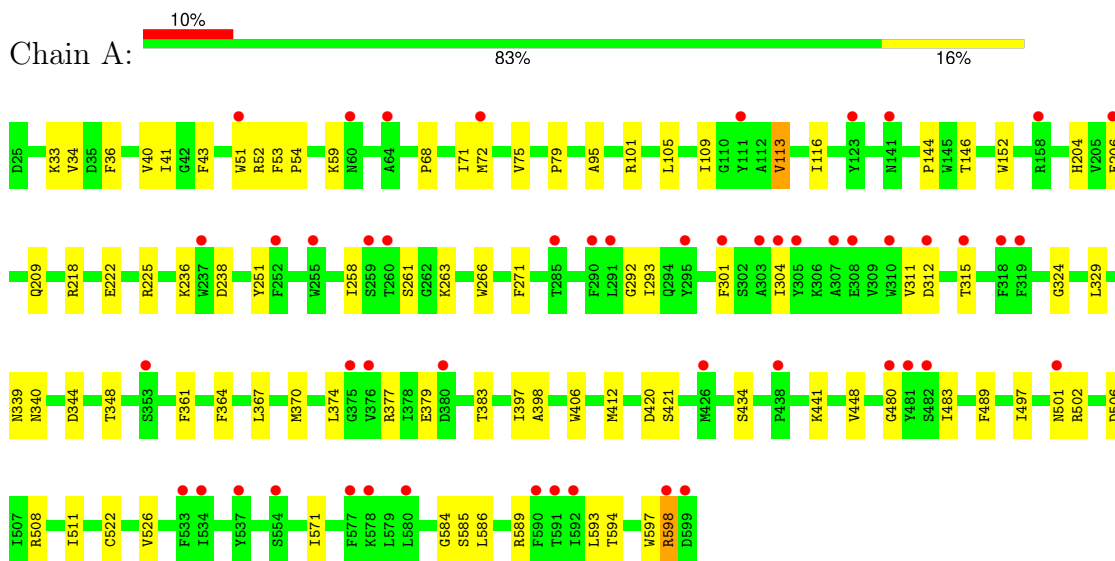
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
10	H	2	Total	O	0	0
			2	2		

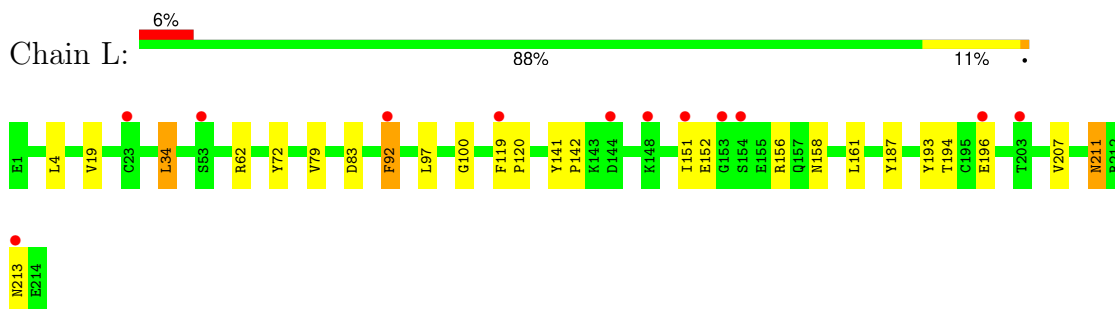
### 3 Residue-property plots [i](#)

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.

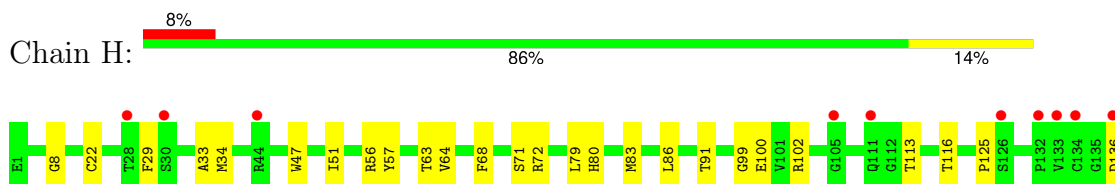
- Molecule 1: Transporter



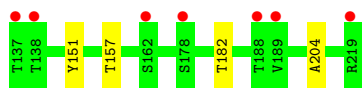
- Molecule 2: Antibody fragment light chain



- Molecule 3: Antibody fragment heavy chain







- Molecule 4: alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose

Chain B: 50% 50%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	97.97Å 144.66Å 165.67Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.11 – 2.90 48.11 – 2.90	Depositor EDS
% Data completeness (in resolution range)	95.2 (48.11-2.90) 96.8 (48.11-2.90)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.25 (at 2.86Å)	Xtrriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.240 , 0.277 0.247 , 0.279	Depositor DCC
$R_{free}$ test set	2605 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	92.3	Xtrriage
Anisotropy	0.202	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 44.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.53$ , $\langle L^2 \rangle = 0.37$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	7605	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	92.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.34% 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: 42J, P4G, CL, GLC, NA, BGC, CLR

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	0.25	0/4374	0.44	0/5973
2	L	0.27	0/1668	0.45	0/2267
3	H	0.24	0/1665	0.44	0/2272
All	All	0.25	0/7707	0.44	0/10512

There are no bond length outliers.

There are no bond angle outliers.

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	4229	0	4160	60	0
2	L	1630	0	1539	18	0
3	H	1626	0	1559	19	0
4	B	23	0	21	1	0
5	A	11	0	18	0	0
6	A	11	0	9	1	0
7	A	2	0	0	0	0
7	L	1	0	0	0	0
8	A	1	0	0	0	0
9	A	56	0	92	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
10	A	9	0	0	0	0
10	H	2	0	0	0	0
10	L	4	0	0	1	0
All	All	7605	0	7398	93	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 6.

All (93) 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:95:ALA:HA	1:A:329:LEU:HD23	1.71	0.72
1:A:364:PHE:HA	1:A:367:LEU:HB2	1.73	0.69
1:A:497:ILE:HG23	1:A:594:THR:HG23	1.76	0.68
1:A:146:THR:HG22	1:A:398:ALA:HB1	1.75	0.67
1:A:222:GLU:HG3	1:A:236:LYS:HZ1	1.60	0.66
1:A:584:GLY:C	1:A:589:ARG:HH22	1.99	0.65
1:A:258:ILE:HA	1:A:261:SER:HB3	1.80	0.64
1:A:236:LYS:HG2	1:A:238:ASP:OD1	1.98	0.64
1:A:585:SER:OG	1:A:586:LEU:N	2.32	0.62
1:A:33:LYS:NZ	1:A:339:ASN:OD1	2.31	0.60
2:L:62:ARG:NH1	2:L:83:ASP:OD1	2.36	0.59
1:A:33:LYS:NZ	1:A:344:ASP:OD2	2.36	0.59
3:H:22:CYS:HB3	3:H:79:LEU:HB3	1.85	0.58
2:L:196:GLU:HG2	2:L:207:VAL:HG13	1.86	0.58
1:A:41:ILE:HD13	1:A:348:THR:HG23	1.85	0.58
1:A:508:ARG:NE	3:H:100:GLU:O	2.36	0.57
1:A:40:VAL:HG12	1:A:348:THR:HG21	1.86	0.57
1:A:75:VAL:HB	1:A:526:VAL:HG11	1.87	0.57
1:A:480:GLY:O	1:A:483:ILE:N	2.33	0.56
2:L:194:THR:HG22	2:L:196:GLU:HG3	1.87	0.56
3:H:29:PHE:O	3:H:72:ARG:NH2	2.38	0.56
1:A:506:ASP:OD1	3:H:56:ARG:NH2	2.40	0.55
2:L:156:ARG:NH1	2:L:158:ASN:O	2.40	0.54
1:A:293:ILE:HD12	1:A:361:PHE:HD1	1.73	0.54
1:A:370:MET:HG2	1:A:374:LEU:HD12	1.89	0.53
1:A:222:GLU:HB3	1:A:225:ARG:HD2	1.92	0.52
1:A:263:LYS:HA	1:A:266:TRP:CD1	2.45	0.52
2:L:34:LEU:HD22	2:L:72:TYR:CG	2.45	0.52
2:L:187:TYR:O	2:L:193:TYR:OH	2.25	0.52
1:A:43:PHE:HA	1:A:421:SER:HA	1.92	0.51

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:H:33:ALA:O	3:H:99:GLY:HA2	2.11	0.51
3:H:8:GLY:O	3:H:113:THR:OG1	2.27	0.51
1:A:206:GLU:O	1:A:209:GLN:NE2	2.44	0.51
3:H:34:MET:HG2	3:H:72:ARG:HH12	1.76	0.50
1:A:101:ARG:HB3	1:A:597:TRP:HB2	1.94	0.50
1:A:109:ILE:HG12	1:A:571:ILE:HD12	1.94	0.50
2:L:97:LEU:HD12	3:H:47:TRP:CE2	2.47	0.50
1:A:397:ILE:HG23	1:A:406:TRP:HB2	1.94	0.49
1:A:152:TRP:O	1:A:218:ARG:HD3	2.13	0.49
1:A:251:TYR:CE1	1:A:448:VAL:HG23	2.48	0.49
1:A:420:ASP:N	1:A:420:ASP:OD1	2.44	0.48
1:A:116:ILE:HG21	1:A:483:ILE:HD13	1.95	0.48
1:A:271:PHE:CE2	1:A:412:MET:HG3	2.48	0.48
1:A:367:LEU:HA	1:A:370:MET:HE2	1.94	0.48
2:L:211:ASN:HB2	2:L:213:ASN:OD1	2.13	0.48
1:A:113:VAL:HG22	1:A:324:GLY:O	2.13	0.47
1:A:75:VAL:O	1:A:79:PRO:HG2	2.14	0.47
1:A:340:ASN:HA	1:A:511:ILE:HG22	1.96	0.47
1:A:598:ARG:HD2	1:A:598:ARG:HA	1.74	0.47
3:H:83:MET:HB3	3:H:86:LEU:HD21	1.96	0.47
1:A:144:PRO:HB3	4:B:2:GLC:O6	2.15	0.46
1:A:222:GLU:CG	1:A:236:LYS:HZ1	2.28	0.46
1:A:292:GLY:HA3	1:A:364:PHE:O	2.15	0.46
3:H:125:PRO:HB3	3:H:151:TYR:HB3	1.97	0.46
1:A:584:GLY:CA	1:A:589:ARG:HH22	2.28	0.45
3:H:51:ILE:HD13	3:H:72:ARG:HB2	1.98	0.45
1:A:204:HIS:NE2	1:A:377:ARG:HD2	2.31	0.45
1:A:584:GLY:O	1:A:589:ARG:NH1	2.49	0.45
1:A:489:PHE:CD2	1:A:571:ILE:HG21	2.52	0.45
1:A:585:SER:O	1:A:589:ARG:HG2	2.17	0.44
2:L:92:PHE:CE1	2:L:97:LEU:HD13	2.53	0.44
1:A:311:VAL:O	1:A:315:THR:HG22	2.18	0.43
2:L:152:GLU:O	10:L:401:HOH:O	2.21	0.43
3:H:157:THR:HB	3:H:204:ALA:HB3	2.00	0.43
1:A:71:ILE:O	1:A:75:VAL:HG22	2.18	0.43
1:A:105:LEU:HB2	1:A:593:LEU:HB3	2.00	0.43
1:A:506:ASP:OD1	3:H:57:TYR:OH	2.26	0.43
2:L:4:LEU:HB2	2:L:100:GLY:HA2	1.98	0.43
1:A:301:PHE:O	1:A:304:ILE:HG22	2.17	0.43
1:A:370:MET:O	1:A:374:LEU:HB2	2.19	0.43
1:A:434:SER:HB3	1:A:441:LYS:HG2	2.01	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:501:ASN:OD1	1:A:502:ARG:N	2.52	0.43
9:A:606:CLR:H211	9:A:606:CLR:H232	1.77	0.43
3:H:71:SER:HB3	3:H:80:HIS:HB2	2.01	0.43
2:L:92:PHE:CE2	3:H:102:ARG:HG2	2.55	0.42
1:A:53:PHE:HB3	1:A:54:PRO:HD3	2.00	0.42
2:L:161:LEU:HD13	3:H:182:THR:HB	2.01	0.42
9:A:607:CLR:H211	9:A:607:CLR:H232	1.67	0.42
1:A:34:VAL:HG13	9:A:607:CLR:H42	2.02	0.42
2:L:141:TYR:CG	2:L:142:PRO:HA	2.55	0.42
1:A:51:TRP:CD1	1:A:52:ARG:N	2.88	0.41
3:H:91:THR:HG23	3:H:116:THR:HA	2.02	0.41
6:A:602:42J:H2	6:A:602:42J:H7	1.62	0.41
3:H:64:VAL:HB	3:H:68:PHE:CG	2.55	0.41
1:A:36:PHE:O	1:A:40:VAL:HG23	2.20	0.41
2:L:92:PHE:HE2	3:H:102:ARG:HG2	1.85	0.41
2:L:19:VAL:HG13	2:L:79:VAL:HG21	2.03	0.41
1:A:222:GLU:HG3	1:A:236:LYS:NZ	2.34	0.40
2:L:119:PHE:HA	2:L:120:PRO:HD2	1.93	0.40
2:L:151:ILE:HG13	2:L:152:GLU:H	1.87	0.40
1:A:522:CYS:HA	1:A:526:VAL:HB	2.04	0.40
1:A:59:LYS:NZ	1:A:312:ASP:OD2	2.52	0.40
1:A:68:PRO:HD3	1:A:304:ILE:HD13	2.04	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	533/534 (100%)	514 (96%)	19 (4%)	0	100	100
2	L	212/214 (99%)	204 (96%)	8 (4%)	0	100	100
3	H	217/219 (99%)	212 (98%)	5 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	962/967 (100%)	930 (97%)	32 (3%)	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	A	434/438 (99%)	428 (99%)	6 (1%)	62	86
2	L	184/187 (98%)	181 (98%)	3 (2%)	58	84
3	H	177/187 (95%)	175 (99%)	2 (1%)	70	90
All	All	795/812 (98%)	784 (99%)	11 (1%)	65	86

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	72[A]	MET
1	A	72[B]	MET
1	A	113	VAL
1	A	379	GLU
1	A	383	THR
1	A	598	ARG
2	L	34	LEU
2	L	92	PHE
2	L	211	ASN
3	H	63	THR
3	H	136	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](#)

2 monosaccharides 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
4	BGC	B	1	4	12,12,12	1.49	2 (16%)	17,17,17	1.09	2 (11%)
4	GLC	B	2	4	11,11,12	2.08	2 (18%)	15,15,17	0.83	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
4	BGC	B	1	4	-	0/2/22/22	0/1/1/1
4	GLC	B	2	4	-	0/2/19/22	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	2	GLC	O2-C2	5.85	1.55	1.43
4	B	1	BGC	C4-C3	3.09	1.60	1.52
4	B	2	GLC	O5-C1	2.90	1.48	1.43
4	B	1	BGC	O1-C1	2.02	1.45	1.39

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	1	BGC	O5-C1-C2	-2.42	106.05	110.30
4	B	1	BGC	C1-C2-C3	-2.01	106.25	110.36



There are no chirality outliers.

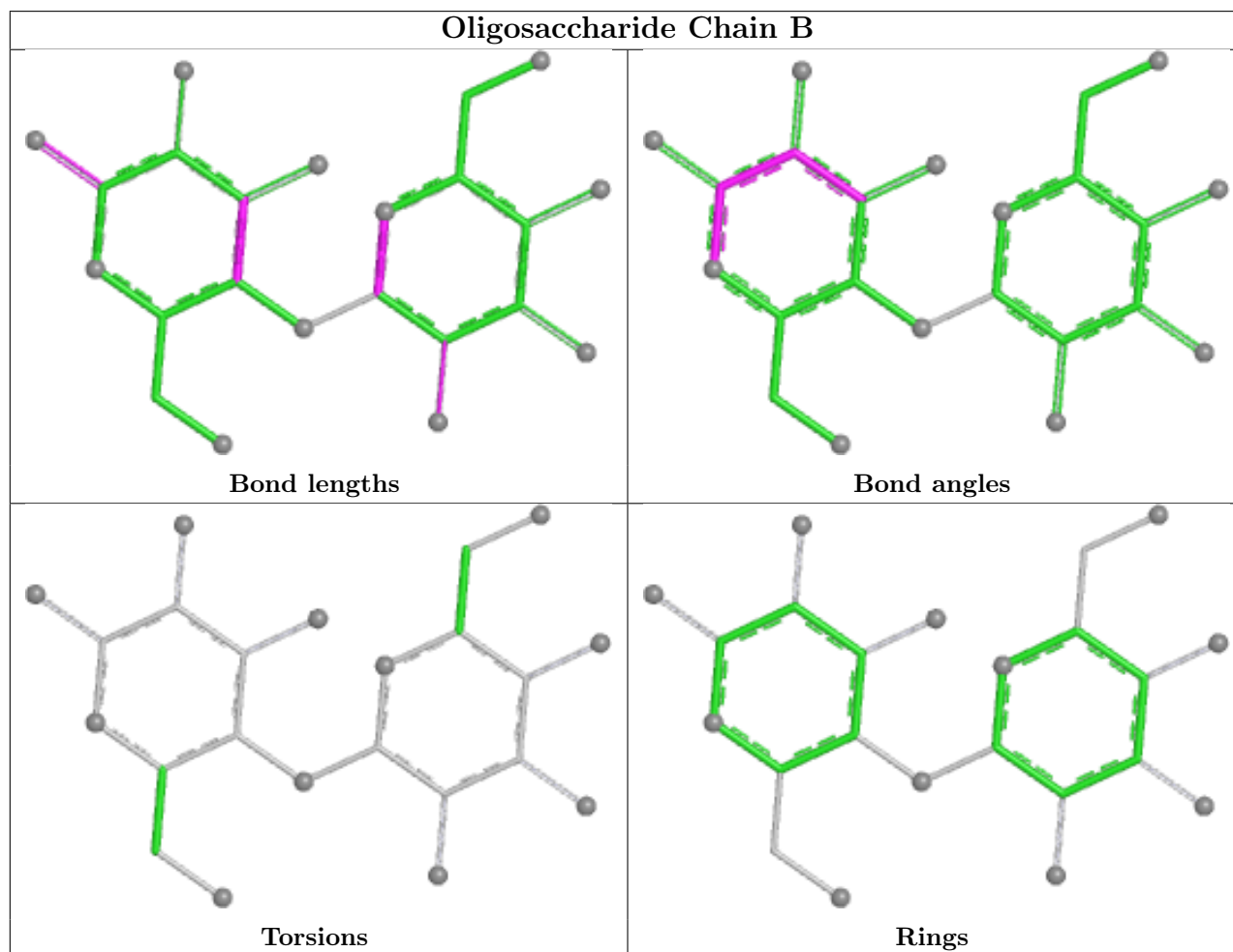
There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	2	GLC	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 oligosaccharide.



## 5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 4 are monoatomic - leaving 4 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
5	P4G	A	601	-	10,10,10	0.40	0	9,9,9	0.33	0
9	CLR	A	607	-	31,31,31	0.67	0	48,48,48	1.22	4 (8%)
6	42J	A	602	-	11,11,11	1.16	2 (18%)	14,14,14	0.82	0
9	CLR	A	606	-	31,31,31	0.65	0	48,48,48	1.10	4 (8%)

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
5	P4G	A	601	-	-	4/8/8/8	-
9	CLR	A	607	-	-	2/10/68/68	0/4/4/4
6	42J	A	602	-	-	3/3/3/3	0/1/1/1
9	CLR	A	606	-	-	3/10/68/68	0/4/4/4

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	602	42J	C5-CL	2.73	1.80	1.73
6	A	602	42J	C6-CL1	2.58	1.79	1.73

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	A	606	CLR	C4-C5-C10	3.47	120.87	116.42
9	A	607	CLR	C8-C7-C6	-3.21	108.32	112.76
9	A	607	CLR	C4-C5-C10	2.94	120.19	116.42
9	A	606	CLR	C8-C7-C6	-2.67	109.06	112.76
9	A	607	CLR	C7-C6-C5	-2.07	121.52	125.02
9	A	606	CLR	C19-C10-C9	-2.07	109.34	111.66
9	A	606	CLR	C4-C5-C6	-2.05	117.79	120.57
9	A	607	CLR	C13-C17-C20	-2.03	116.36	119.50

There are no chirality outliers.

All (12) torsion outliers are listed below:

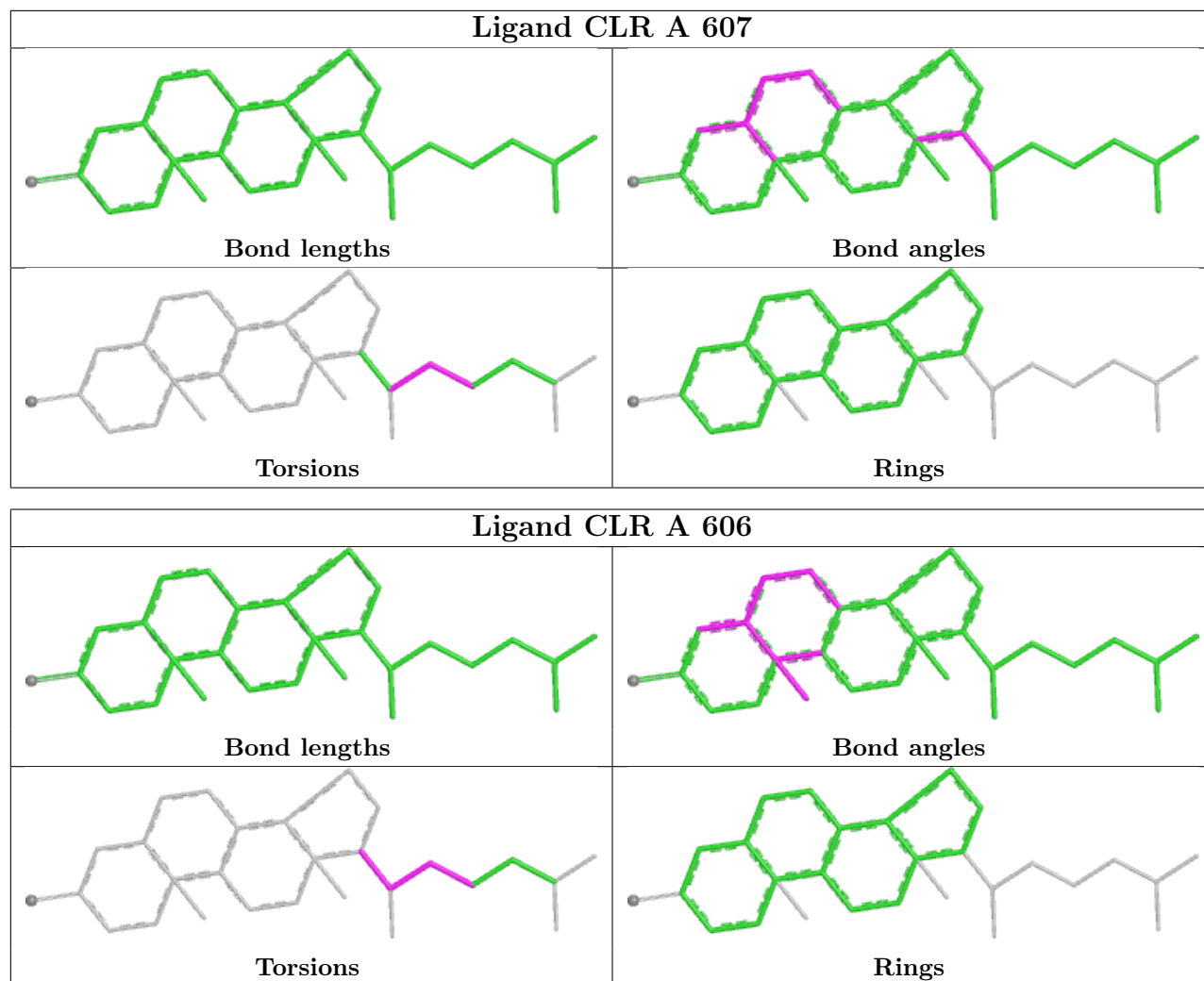
Mol	Chain	Res	Type	Atoms
6	A	602	42J	N-C-C1-C2
9	A	606	CLR	C20-C22-C23-C24
5	A	601	P4G	O2-C3-C4-O3
5	A	601	P4G	C1-C2-O2-C3
5	A	601	P4G	C8-C7-O4-C6
6	A	602	42J	C-C1-C2-C3
6	A	602	42J	C-C1-C2-C7
9	A	606	CLR	C21-C20-C22-C23
9	A	607	CLR	C21-C20-C22-C23
9	A	607	CLR	C20-C22-C23-C24
9	A	606	CLR	C13-C17-C20-C21
5	A	601	P4G	C5-C6-O4-C7

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	A	607	CLR	2	0
6	A	602	42J	1	0
9	A	606	CLR	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	534/534 (100%)	0.60	51 (9%) 15 13	63, 90, 124, 147	1 (0%)
2	L	214/214 (100%)	0.47	12 (5%) 31 27	68, 84, 112, 136	0
3	H	219/219 (100%)	0.54	17 (7%) 20 18	69, 87, 126, 164	0
All	All	967/967 (100%)	0.56	80 (8%) 19 16	63, 89, 120, 164	1 (0%)

All (80) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	481	TYR	4.9
1	A	290	PHE	4.7
1	A	537	TYR	4.3
3	H	219	ARG	3.7
1	A	534	ILE	3.7
1	A	307	ALA	3.7
3	H	162	SER	3.6
2	L	53	SER	3.6
2	L	153	GLY	3.5
2	L	154	SER	3.5
2	L	148	LYS	3.5
1	A	255	TRP	3.5
2	L	196	GLU	3.5
3	H	28	THR	3.4
1	A	206	GLU	3.4
1	A	315	THR	3.3
3	H	105	GLY	3.2
1	A	111	TYR	3.2
3	H	134	CYS	3.1
1	A	303	ALA	3.1
1	A	304	ILE	3.1
3	H	133	VAL	3.1
3	H	189	VAL	3.0

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
2	L	92	PHE	3.0
1	A	319	PHE	2.9
1	A	141	ASN	2.9
2	L	151	ILE	2.8
1	A	305	TYR	2.8
1	A	51	TRP	2.7
1	A	482	SER	2.7
1	A	291	LEU	2.7
1	A	285	THR	2.6
1	A	123	TYR	2.6
3	H	30	SER	2.6
1	A	312	ASP	2.6
1	A	590	PHE	2.5
1	A	158	ARG	2.5
1	A	375	GLY	2.5
3	H	137	THR	2.5
3	H	132	PRO	2.5
1	A	308	GLU	2.5
2	L	213	ASN	2.5
3	H	111	GLN	2.5
1	A	501	ASN	2.4
3	H	126	SER	2.4
1	A	376	VAL	2.4
1	A	438	PRO	2.4
1	A	480	GLY	2.4
1	A	353	SER	2.3
1	A	260	THR	2.3
1	A	301	PHE	2.3
1	A	318	PHE	2.3
1	A	64	ALA	2.3
3	H	138	THR	2.3
1	A	252	PHE	2.3
3	H	178	SER	2.3
1	A	237	TRP	2.2
1	A	592	ILE	2.2
1	A	533	PHE	2.2
1	A	426	MET	2.2
1	A	259	SER	2.2
1	A	578	LYS	2.1
1	A	599	ASP	2.1
3	H	136	ASP	2.1
1	A	591	THR	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	L	203	THR	2.1
1	A	295	TYR	2.1
1	A	72[A]	MET	2.1
1	A	60	ASN	2.1
2	L	23	CYS	2.1
3	H	44	ARG	2.1
1	A	380	ASP	2.0
3	H	188	THR	2.0
1	A	554	SER	2.0
1	A	598	ARG	2.0
2	L	119	PHE	2.0
2	L	144	ASP	2.0
1	A	310	TRP	2.0
1	A	577	PHE	2.0
1	A	580	LEU	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](#)

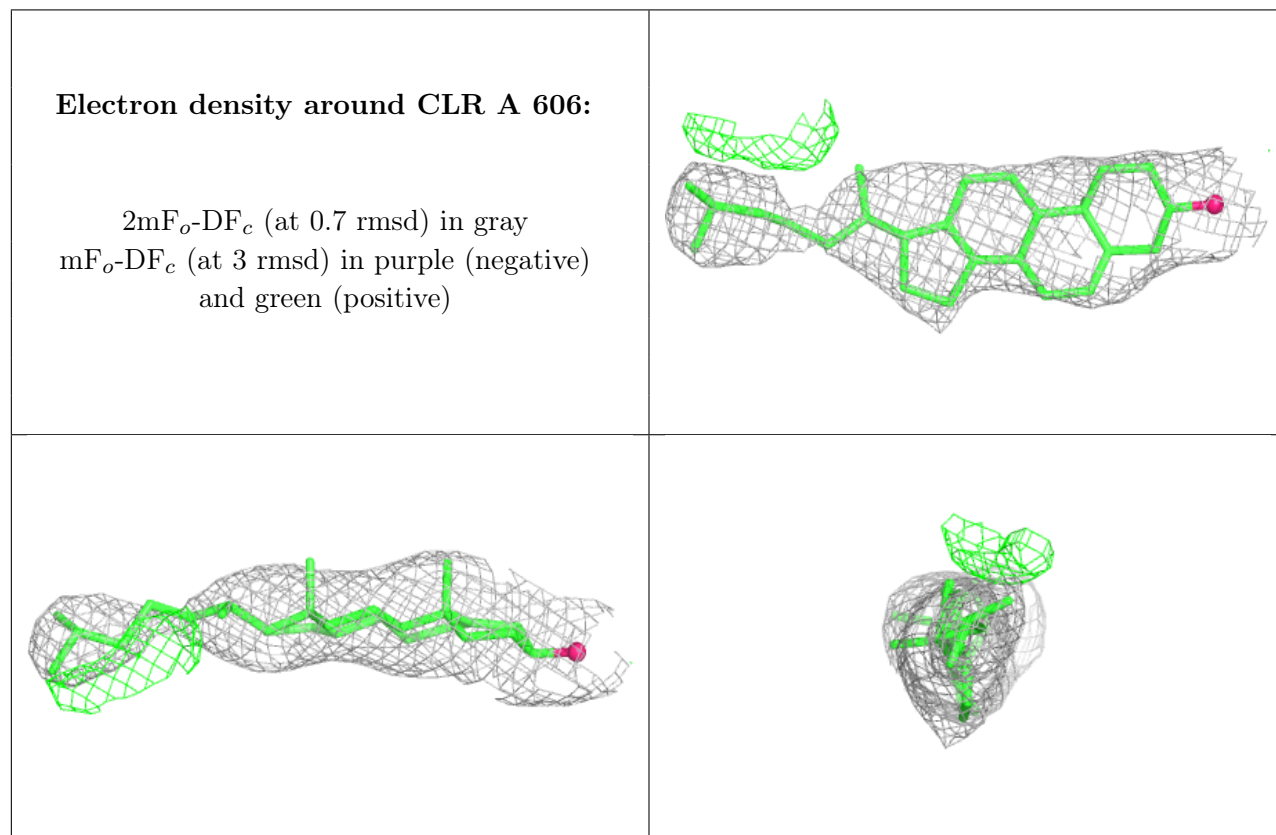
SUGAR-RSR INFOmissingINFO

## 6.4 Ligands [i](#)

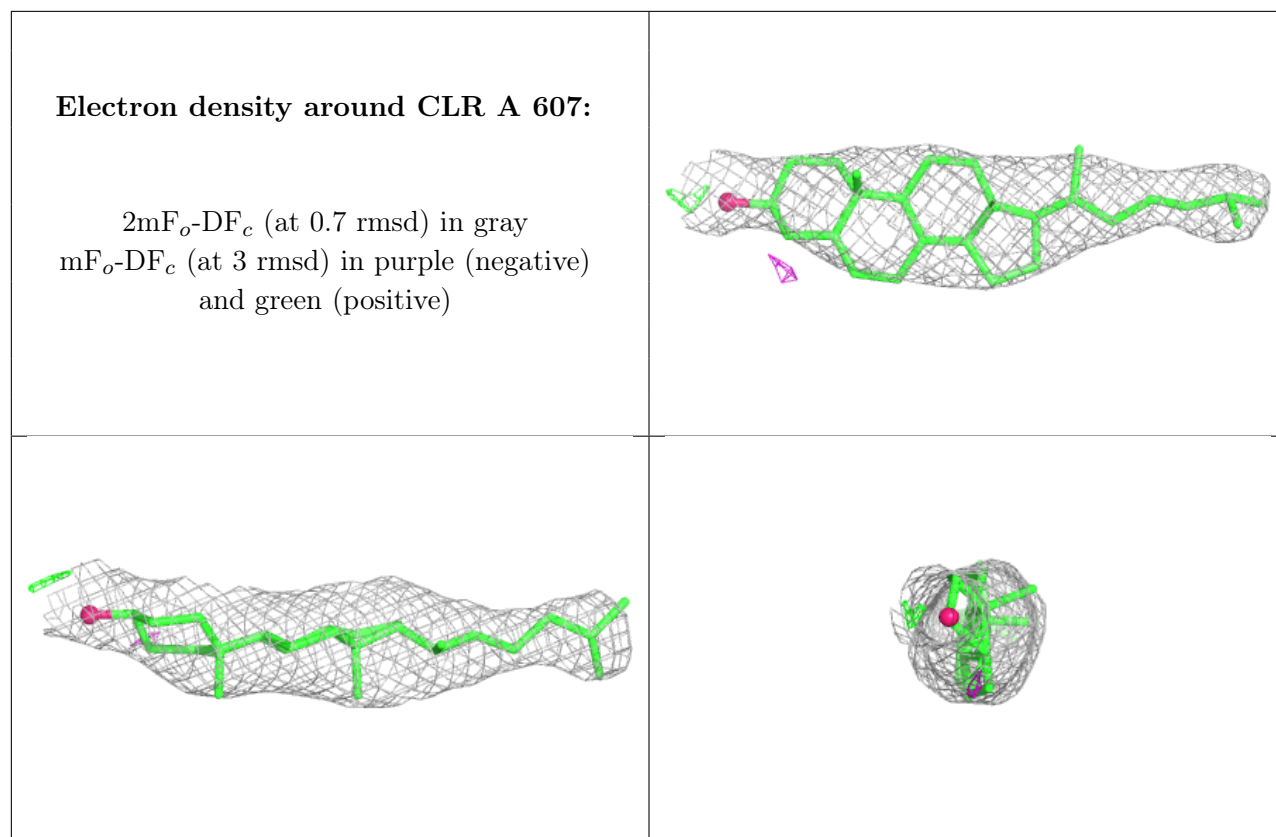
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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
5	P4G	A	601	11/11	0.81	0.37	124,137,142,144	0
9	CLR	A	606	28/28	0.90	0.20	88,111,115,120	0
7	NA	L	301	1/1	0.91	0.23	81,81,81,81	0
6	42J	A	602	11/11	0.91	0.17	83,87,102,110	0
9	CLR	A	607	28/28	0.92	0.17	76,102,107,108	0
7	NA	A	604	1/1	0.94	0.10	78,78,78,78	0
7	NA	A	603	1/1	0.94	0.07	85,85,85,85	0
8	CL	A	605	1/1	0.97	0.12	82,82,82,82	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.







## 6.5 Other polymers [i](#)

There are no such residues in this entry.