



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 03:37 PM GMT

PDB ID : 4CQX  
Title : H5 (tyTy) Del133/Ile155Thr Mutant Haemagglutinin in Complex with Human Receptor Analogue 6'SLN  
Authors : Xiong, X.; Xiao, H.; Martin, S.R.; Coombs, P.J.; Liu, J.; Collins, P.J.; Vachieri, S.G.; Walker, P.A.; Lin, Y.P.; McCauley, J.W.; Gamblin, S.J.; Skehel, J.J.  
Deposited on : 2014-02-21  
Resolution : 2.30 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

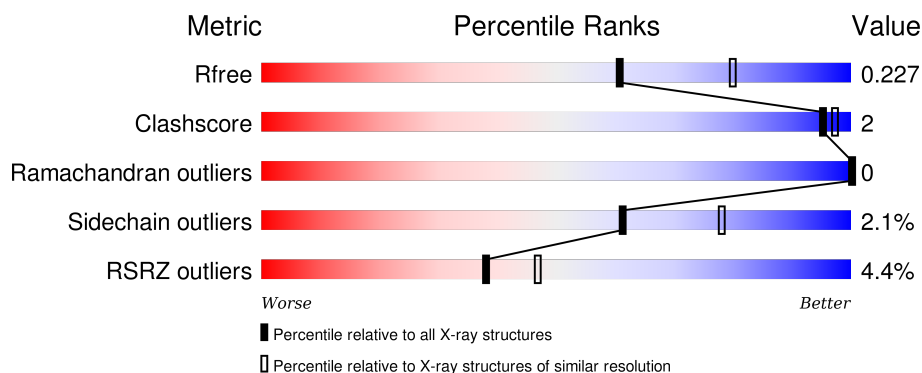
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



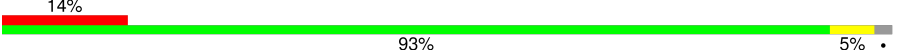
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (2.30-2.30)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	327	<div> <div>94%</div> <div> <div></div> <div></div> <div></div> <div></div> </div> </div>
1	C	327	<div> <div>%</div> <div>94%</div> <div> <div></div> <div></div> <div></div> <div></div> </div> </div>
1	E	327	<div> <div>%</div> <div>94%</div> <div> <div></div> <div></div> <div></div> <div></div> </div> </div>
2	B	166	<div> <div></div> <div>90%</div> <div>8%</div> <div> <div></div> <div></div> <div></div> <div></div> </div> </div>
2	D	166	<div> <div>20%</div> <div>94%</div> <div> <div></div> <div></div> <div></div> <div></div> </div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	166	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	NAG	A	1322	X	-	-	-
3	NAG	E	1322	X	-	-	-
5	PO4	A	1327	-	-	-	X
5	PO4	A	1328	-	-	-	X
5	PO4	C	1326	-	-	-	X
5	PO4	C	1327	-	-	-	X
5	PO4	E	1328	-	-	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 12610 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 HAEMAGGLUTININ HA1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	322	Total	C	N	O	S	0	0	0
			2552	1607	443	488	14			
1	C	320	Total	C	N	O	S	0	0	0
			2537	1598	441	484	14			
1	E	322	Total	C	N	O	S	0	0	0
			2549	1606	443	486	14			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	ASP	-	EXPRESSION TAG	UNP Q207Z6
A	0	PRO	-	EXPRESSION TAG	UNP Q207Z6
A	.	-	ALA	DELETION	UNP Q207Z6
A	150	THR	ILE	ENGINEERED MUTATION	UNP Q207Z6
A	322	ARG	GLY	CONFLICT	UNP Q207Z6
A	324	THR	ARG	CONFLICT	UNP Q207Z6
C	-1	ASP	-	EXPRESSION TAG	UNP Q207Z6
C	0	PRO	-	EXPRESSION TAG	UNP Q207Z6
C	.	-	ALA	DELETION	UNP Q207Z6
C	150	THR	ILE	ENGINEERED MUTATION	UNP Q207Z6
C	322	ARG	GLY	CONFLICT	UNP Q207Z6
C	324	THR	ARG	CONFLICT	UNP Q207Z6
E	-1	ASP	-	EXPRESSION TAG	UNP Q207Z6
E	0	PRO	-	EXPRESSION TAG	UNP Q207Z6
E	.	-	ALA	DELETION	UNP Q207Z6
E	150	THR	ILE	ENGINEERED MUTATION	UNP Q207Z6
E	322	ARG	GLY	CONFLICT	UNP Q207Z6
E	324	THR	ARG	CONFLICT	UNP Q207Z6

- Molecule 2 is a protein called HAEMAGGLUTININ HA2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	163	Total	C	N	O	S	0	0	0
			1320	822	229	261	8			
2	D	163	Total	C	N	O	S	0	0	0
			1298	808	228	254	8			
2	F	163	Total	C	N	O	S	0	0	0
			1306	815	226	257	8			

- Molecule 3 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

*Continued on next page...*

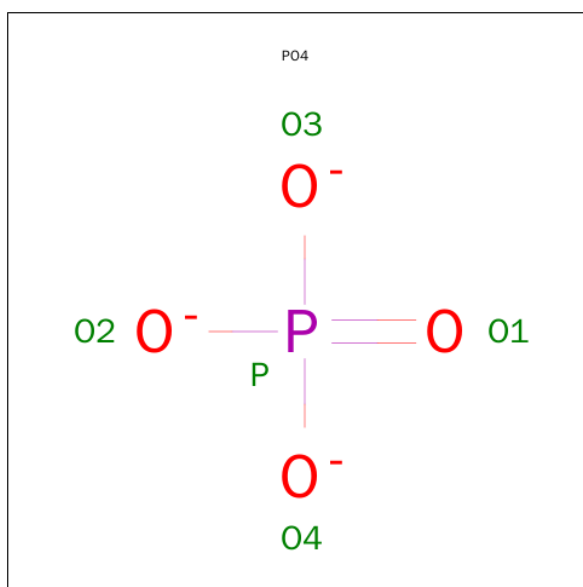
Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	E	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 4 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	3	Total	C	N	O	0	0
			46	25	2	19		
4	C	3	Total	C	N	O	0	0
			46	25	2	19		
4	E	3	Total	C	N	O	0	0
			46	25	2	19		

- Molecule 5 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	O	P	0	0
			5	4	1		
5	A	1	Total	O	P	0	0
			5	4	1		
5	C	1	Total	O	P	0	0
			5	4	1		
5	C	1	Total	O	P	0	0
			5	4	1		
5	E	1	Total	O	P	0	0
			5	4	1		

- Molecule 6 is water.

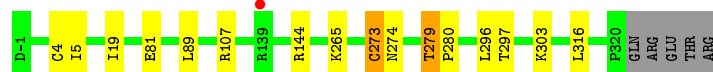
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	182	Total 182	O 182	0	0
6	B	84	Total 84	O 84	0	0
6	C	195	Total 195	O 195	0	0
6	D	45	Total 45	O 45	0	0
6	E	217	Total 217	O 217	0	0
6	F	36	Total 36	O 36	0	0

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

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: HAEMAGGLUTININ HA1

Chain A: 

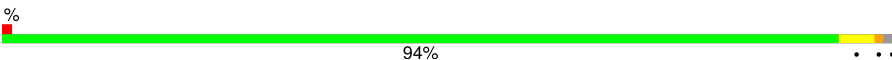


- Molecule 1: HAEMAGGLUTININ HA1

Chain C: 




- Molecule 1: HAEMAGGLUTININ HA1

Chain E: 



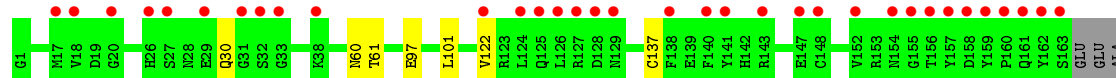
- Molecule 2: HAEMAGGLUTININ HA2

Chain B: 



- Molecule 2: HAEMAGGLUTININ HA2

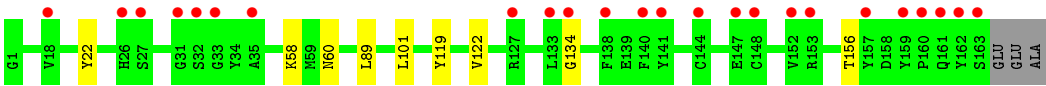
Chain D: 



- Molecule 2: HAEMAGGLUTININ HA2

Chain F: 





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	87.82Å 117.10Å 101.28Å 90.00° 92.58° 90.00°	Depositor
Resolution (Å)	101.17 – 2.30 70.21 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.5 (101.17-2.30) 99.5 (70.21-2.30)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.05 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.8.0046	Depositor
R, $R_{free}$	0.186 , 0.226 0.189 , 0.227	Depositor DCC
$R_{free}$ test set	4539 reflections (5.28%)	DCC
Wilson B-factor (Å <sup>2</sup> )	35.7	Xtriage
Anisotropy	0.658	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 42.5	EDS
Estimated twinning fraction	0.030 for h,-k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	1 of 90456 reflections (0.001%)	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	12610	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.70% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: PO4, GAL, NAG, SIA

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	0.35	0/2614	0.58	0/3554
1	C	0.37	0/2598	0.58	0/3531
1	E	0.39	0/2611	0.61	0/3550
2	B	0.39	0/1347	0.54	0/1813
2	D	0.36	0/1324	0.51	0/1784
2	F	0.36	0/1333	0.54	0/1796
All	All	0.37	0/11827	0.57	0/16028

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	2552	0	2483	8	0
1	C	2537	0	2473	7	0
1	E	2549	0	2480	9	0
2	B	1320	0	1224	10	0
2	D	1298	0	1196	4	0
2	F	1306	0	1205	5	0
3	A	42	0	39	4	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	28	0	26	0	0
3	E	56	0	52	0	0
4	A	46	0	40	0	0
4	C	46	0	40	0	0
4	E	46	0	40	1	0
5	A	10	0	0	0	0
5	C	10	0	0	0	0
5	E	5	0	0	1	0
6	A	182	0	0	0	0
6	B	84	0	0	0	0
6	C	195	0	0	1	0
6	D	45	0	0	1	0
6	E	217	0	0	0	0
6	F	36	0	0	0	0
All	All	12610	0	11298	36	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 2.

All (36) 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:5:ILE:HD11	2:B:122:VAL:HG21	1.63	0.78
1:E:279:THR:HG22	1:E:281:ILE:H	1.48	0.78
1:E:279:THR:HB	1:E:282:GLY:O	1.87	0.73
2:B:156:THR:O	2:B:156:THR:HG22	1.90	0.72
1:A:303:LYS:HE2	2:B:60:ASN:HA	1.85	0.58
2:B:51:LYS:HE3	1:E:20:MET:HE2	1.85	0.58
3:A:1322:NAG:H3	3:A:1322:NAG:C8	2.34	0.58
2:B:124:LEU:HD22	2:F:134:GLY:HA2	1.87	0.57
1:C:19:ILE:HD12	2:D:101:LEU:HB3	1.87	0.56
2:D:60:ASN:HB3	6:D:2013:HOH:O	2.06	0.55
3:A:1322:NAG:H3	3:A:1322:NAG:H83	1.91	0.53
2:B:156:THR:CG2	2:B:156:THR:O	2.60	0.49
1:C:166:THR:HG22	6:C:2125:HOH:O	2.14	0.47
1:A:273:CYS:SG	1:A:274:ASN:N	2.87	0.47
4:E:1327:NAG:H1	4:E:1327:NAG:H82	1.96	0.47
1:E:5:ILE:HG13	2:F:119:TYR:HA	1.98	0.46
1:E:5:ILE:HD11	2:F:122:VAL:HG21	1.97	0.45
1:A:81:GLU:O	1:A:265:LYS:HA	2.16	0.45
1:C:121:SER:CB	1:C:161:ARG:HH22	2.29	0.45

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:19:ILE:HD12	2:F:101:LEU:HB3	1.99	0.45
1:C:234:PRO:O	1:C:235:ASN:HB2	2.16	0.45
1:C:279:THR:HG22	1:C:297:THR:HG22	1.99	0.45
2:B:72:ASN:OD1	2:B:75:ARG:NH2	2.50	0.45
2:D:97:GLU:HB3	2:F:58:LYS:HE3	2.00	0.43
3:A:1322:NAG:C8	3:A:1322:NAG:C3	2.96	0.43
1:C:121:SER:OG	1:C:161:ARG:NH2	2.41	0.43
1:E:35:LYS:N	5:E:1328:PO4:O2	2.41	0.43
1:A:280:PRO:HD3	1:A:296:LEU:O	2.19	0.42
1:C:5:ILE:HD11	2:D:122:VAL:HG21	2.01	0.42
1:E:90:CYS:HB2	1:E:133:ALA:O	2.19	0.42
2:B:119:TYR:OH	2:B:132:GLU:HG3	2.19	0.42
1:A:279:THR:HG22	1:A:297:THR:HG22	2.02	0.41
3:A:1322:NAG:H83	3:A:1322:NAG:C3	2.49	0.41
1:A:19:ILE:HD12	2:B:101:LEU:HB3	2.01	0.41
1:E:83:ILE:HD12	1:E:267:GLU:OE2	2.20	0.41
1:A:4:CYS:O	2:B:24:TYR:HA	2.21	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	320/327 (98%)	310 (97%)	10 (3%)	0	100	100
1	C	318/327 (97%)	307 (96%)	11 (4%)	0	100	100
1	E	320/327 (98%)	311 (97%)	9 (3%)	0	100	100
2	B	161/166 (97%)	159 (99%)	2 (1%)	0	100	100
2	D	161/166 (97%)	157 (98%)	4 (2%)	0	100	100
2	F	161/166 (97%)	157 (98%)	4 (2%)	0	100	100
All	All	1441/1479 (97%)	1401 (97%)	40 (3%)	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	289/294 (98%)	283 (98%)	6 (2%)	61	78
1	C	287/294 (98%)	283 (99%)	4 (1%)	74	86
1	E	288/294 (98%)	282 (98%)	6 (2%)	61	78
2	B	139/141 (99%)	135 (97%)	4 (3%)	50	66
2	D	134/141 (95%)	131 (98%)	3 (2%)	60	77
2	F	136/141 (96%)	132 (97%)	4 (3%)	50	66
All	All	1273/1305 (98%)	1246 (98%)	27 (2%)	61	78

All (27) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	89	LEU
1	A	107	ARG
1	A	144	ARG
1	A	273	CYS
1	A	279	THR
1	A	316	LEU
2	B	22	TYR
2	B	29	GLU
2	B	30	GLN
2	B	60	ASN
1	C	4	CYS
1	C	95	PHE
1	C	107	ARG
1	C	316	LEU
2	D	30	GLN
2	D	61	THR
2	D	137	CYS
1	E	174	LEU
1	E	278	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	279	THR
1	E	281	ILE
1	E	296	LEU
1	E	316	LEU
2	F	22	TYR
2	F	60	ASN
2	F	89	LEU
2	F	156	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	87	ASN
2	B	154	ASN
1	C	2	GLN
1	C	15	GLN
2	F	161	GLN

### 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 ⓘ

9 carbohydrates 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	SIA	A	1324	4	16,20,21	0.34	0	18,28,31	1.07	2 (11%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	GAL	A	1325	4	11,11,12	0.42	0	14,15,17	0.97	1 (7%)
4	NAG	A	1326	4	15,15,15	0.38	0	17,21,21	0.71	1 (5%)
4	SIA	C	1323	4	16,20,21	0.38	0	18,28,31	1.19	2 (11%)
4	GAL	C	1324	4	11,11,12	0.41	0	14,15,17	0.99	1 (7%)
4	NAG	C	1325	4	15,15,15	0.48	0	17,21,21	1.04	3 (17%)
4	SIA	E	1325	4	16,20,21	0.25	0	18,28,31	0.98	2 (11%)
4	GAL	E	1326	4	11,11,12	0.46	0	14,15,17	0.75	0
4	NAG	E	1327	4	15,15,15	0.63	0	17,21,21	1.94	5 (29%)

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	SIA	A	1324	4	-	0/14/34/38	0/1/1/1
4	GAL	A	1325	4	-	0/2/19/22	0/1/1/1
4	NAG	A	1326	4	-	0/6/26/26	0/1/1/1
4	SIA	C	1323	4	-	0/14/34/38	0/1/1/1
4	GAL	C	1324	4	-	0/2/19/22	0/1/1/1
4	NAG	C	1325	4	-	0/6/26/26	0/1/1/1
4	SIA	E	1325	4	-	0/14/34/38	0/1/1/1
4	GAL	E	1326	4	-	0/2/19/22	0/1/1/1
4	NAG	E	1327	4	-	0/6/26/26	0/1/1/1

There are no bond length outliers.

All (17) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	1323	SIA	C4-C5-N5	-3.15	103.56	110.41
4	A	1324	SIA	C3-C4-C5	-2.82	108.34	111.47
4	A	1324	SIA	C4-C5-N5	-2.36	105.29	110.41
4	E	1325	SIA	C3-C4-C5	-2.10	109.14	111.47
4	A	1326	NAG	C3-C2-N2	-2.07	106.37	110.66
4	E	1325	SIA	C4-C5-N5	-2.07	105.90	110.41
4	C	1325	NAG	C3-C2-N2	-2.03	106.46	110.66
4	C	1325	NAG	O5-C5-C6	2.03	111.48	106.36
4	C	1325	NAG	C4-C3-C2	2.23	113.52	110.43
4	C	1323	SIA	C7-C6-C5	2.29	117.80	114.32
4	E	1327	NAG	O5-C5-C4	2.45	114.29	109.68

*Continued on next page...*



Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	1324	GAL	C1-C2-C3	2.87	112.93	109.54
4	A	1325	GAL	C1-C2-C3	2.89	112.96	109.54
4	E	1327	NAG	C8-C7-N2	3.06	121.96	116.11
4	E	1327	NAG	C3-C4-C5	3.42	116.15	110.20
4	E	1327	NAG	C2-N2-C7	3.64	132.44	123.10
4	E	1327	NAG	C4-C3-C2	3.92	115.87	110.43

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	E	1327	NAG	1	0

## 5.6 Ligand geometry

14 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
3	NAG	A	1321	1	14,14,15	0.45	0	15,19,21	1.13	1 (6%)
3	NAG	A	1322	1	14,14,15	0.87	1 (7%)	15,19,21	2.32	4 (26%)
3	NAG	A	1323	1	14,14,15	0.32	0	15,19,21	2.31	3 (20%)
5	PO4	A	1327	-	4,4,4	0.41	0	6,6,6	0.28	0
5	PO4	A	1328	-	4,4,4	0.45	0	6,6,6	0.28	0
3	NAG	C	1321	1	14,14,15	0.75	0	15,19,21	1.46	3 (20%)
3	NAG	C	1322	1	14,14,15	0.48	0	15,19,21	1.32	2 (13%)
5	PO4	C	1326	-	4,4,4	0.31	0	6,6,6	0.27	0
5	PO4	C	1327	-	4,4,4	0.52	0	6,6,6	0.28	0
3	NAG	E	1321	1	14,14,15	0.55	0	15,19,21	1.27	2 (13%)
3	NAG	E	1322	1	14,14,15	0.54	0	15,19,21	1.46	3 (20%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	NAG	E	1323	1	14,14,15	0.54	0	15,19,21	1.14	2 (13%)
3	NAG	E	1324	1	14,14,15	0.44	0	15,19,21	1.42	2 (13%)
5	PO4	E	1328	-	4,4,4	0.48	0	6,6,6	0.27	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
3	NAG	A	1321	1	-	0/6/23/26	0/1/1/1
3	NAG	A	1322	1	1/1/5/7	0/6/23/26	0/1/1/1
3	NAG	A	1323	1	-	0/6/23/26	0/1/1/1
5	PO4	A	1327	-	-	0/0/0/0	0/0/0/0
5	PO4	A	1328	-	-	0/0/0/0	0/0/0/0
3	NAG	C	1321	1	-	0/6/23/26	0/1/1/1
3	NAG	C	1322	1	-	0/6/23/26	0/1/1/1
5	PO4	C	1326	-	-	0/0/0/0	0/0/0/0
5	PO4	C	1327	-	-	0/0/0/0	0/0/0/0
3	NAG	E	1321	1	-	0/6/23/26	0/1/1/1
3	NAG	E	1322	1	1/1/5/7	0/6/23/26	0/1/1/1
3	NAG	E	1323	1	-	0/6/23/26	0/1/1/1
3	NAG	E	1324	1	-	0/6/23/26	0/1/1/1
5	PO4	E	1328	-	-	0/0/0/0	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1322	NAG	C1-C2	2.52	1.55	1.52

All (22) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1322	NAG	O7-C7-C8	-3.25	116.09	122.06
3	E	1322	NAG	C4-C3-C2	-2.68	107.06	111.23
3	C	1321	NAG	O7-C7-C8	-2.56	117.36	122.06
3	C	1322	NAG	C3-C4-C5	-2.42	105.97	110.20
3	A	1321	NAG	O7-C7-C8	-2.21	118.02	122.06
3	E	1324	NAG	C3-C4-C5	-2.17	106.41	110.20
3	A	1323	NAG	C6-C5-C4	-2.17	107.65	113.02
3	A	1323	NAG	C3-C2-N2	-2.15	105.41	110.56
3	E	1323	NAG	C3-C4-C5	-2.01	106.69	110.20

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	1321	NAG	O5-C5-C6	2.06	111.80	107.35
3	E	1322	NAG	C1-O5-C5	2.16	114.98	112.25
3	E	1322	NAG	C8-C7-N2	2.21	120.33	116.11
3	C	1321	NAG	C4-C3-C2	2.24	114.71	111.23
3	E	1323	NAG	C1-O5-C5	2.52	115.45	112.25
3	A	1322	NAG	C1-O5-C5	3.09	116.17	112.25
3	C	1321	NAG	C2-N2-C7	3.39	127.39	123.04
3	E	1321	NAG	C1-O5-C5	3.41	116.58	112.25
3	C	1322	NAG	C1-O5-C5	3.58	116.79	112.25
3	A	1322	NAG	C8-C7-N2	4.12	124.00	116.11
3	E	1324	NAG	C1-O5-C5	4.27	117.67	112.25
3	A	1322	NAG	C2-N2-C7	5.93	130.65	123.04
3	A	1323	NAG	C1-O5-C5	7.91	122.28	112.25

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	E	1322	NAG	C1
3	A	1322	NAG	C1

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1322	NAG	4	0
5	E	1328	PO4	1	0

## 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	322/327 (98%)	-0.17	1 (0%) 94 96	28, 47, 65, 82	0
1	C	320/327 (97%)	-0.04	3 (0%) 85 89	29, 43, 63, 112	0
1	E	322/327 (98%)	-0.08	2 (0%) 90 93	27, 39, 55, 102	0
2	B	163/166 (98%)	-0.02	0 100 100	25, 44, 71, 83	0
2	D	163/166 (98%)	1.04	34 (20%) 1 2	28, 64, 149, 167	0
2	F	163/166 (98%)	0.65	24 (14%) 3 5	25, 57, 156, 176	0
All	All	1453/1479 (98%)	0.12	64 (4%) 38 47	25, 44, 109, 176	0

All (64) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	33	GLY	6.9
2	D	157	TYR	6.1
2	D	141	TYR	5.5
2	F	160	PRO	5.3
2	D	159	TYR	5.0
2	F	27	SER	4.8
2	D	160	PRO	4.4
2	D	27	SER	4.2
2	D	143	ARG	4.2
2	D	147	GLU	4.1
1	C	3	ILE	4.1
2	D	161	GLN	4.0
2	D	17	MET	4.0
2	F	159	TYR	3.9
2	D	152	VAL	3.9
2	D	124	LEU	3.7
2	D	140	PHE	3.7
2	D	18	VAL	3.6
2	D	125	GLN	3.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	F	144	CYS	3.3
2	F	33	GLY	3.3
2	D	154	ASN	3.3
2	D	158	ASP	3.2
1	C	320	PRO	3.2
2	F	147	GLU	3.2
2	F	26	HIS	3.2
2	D	129	ASN	3.2
2	D	26	HIS	3.2
2	D	32	SER	3.2
2	D	156	THR	3.2
2	F	141	TYR	3.1
2	D	163	SER	3.0
2	F	153	ARG	2.9
2	F	157	TYR	2.9
2	D	138	PHE	2.9
1	A	139	ARG	2.9
2	F	134	GLY	2.9
2	D	155	GLY	2.8
2	F	152	VAL	2.8
2	D	127	ARG	2.7
1	C	5	ILE	2.7
2	D	148	CYS	2.7
2	D	29	GLU	2.7
2	F	133	LEU	2.6
2	F	163	SER	2.6
2	D	20	GLY	2.6
2	D	162	TYR	2.5
2	F	31	GLY	2.5
2	F	127	ARG	2.4
2	D	122	VAL	2.3
1	E	1	ASP	2.3
2	D	31	GLY	2.3
2	F	140	PHE	2.2
2	F	18	VAL	2.2
2	F	32	SER	2.2
2	F	161	GLN	2.1
2	D	126	LEU	2.1
2	F	35	ALA	2.1
1	E	273	CYS	2.1
2	D	38	LYS	2.1
2	F	148	CYS	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	F	138	PHE	2.0
2	D	128	ASP	2.0
2	F	162	TYR	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](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
4	SIA	E	1325	20/21	0.96	0.12	-0.30	35,39,48,50	0
4	SIA	C	1323	20/21	0.98	0.12	-0.31	34,37,47,48	0
4	SIA	A	1324	20/21	0.96	0.10	-0.41	44,48,53,53	0
4	GAL	E	1326	11/12	0.89	0.16	-	52,76,84,88	0
4	NAG	C	1325	15/15	0.83	0.24	-	88,98,108,108	0
4	GAL	A	1325	11/12	0.85	0.17	-	58,81,91,92	0
4	NAG	E	1327	15/15	0.71	0.32	-	99,107,111,111	0
4	NAG	A	1326	15/15	0.83	0.36	-	100,109,117,118	0
4	GAL	C	1324	11/12	0.88	0.12	-	52,75,77,79	0

## 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
5	PO4	E	1328	5/5	0.86	0.34	16.76	97,100,104,112	0
5	PO4	A	1327	5/5	0.79	0.42	12.09	108,112,116,122	0
5	PO4	A	1328	5/5	0.93	0.33	10.50	101,102,104,107	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
5	PO4	C	1327	5/5	0.92	0.28	8.31	80,81,85,93	0
5	PO4	C	1326	5/5	0.86	0.29	5.56	85,90,91,92	0
3	NAG	E	1324	14/15	0.78	0.17	1.04	77,91,94,97	0
3	NAG	A	1321	14/15	0.96	0.12	-0.40	39,48,54,61	0
3	NAG	A	1323	14/15	0.90	0.11	-0.43	46,51,56,57	0
3	NAG	C	1322	14/15	0.94	0.09	-3.04	44,51,54,54	0
3	NAG	C	1321	14/15	0.90	0.19	-	65,75,80,81	0
3	NAG	E	1322	14/15	0.88	0.23	-	64,77,87,89	0
3	NAG	A	1322	14/15	0.86	0.14	-	63,76,85,88	0
3	NAG	E	1321	14/15	0.93	0.13	-	50,56,58,60	0
3	NAG	E	1323	14/15	0.95	0.11	-	41,47,52,61	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.