



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 16, 2017 – 05:11 PM EDT

PDB ID : 3CU7  
Title : Human Complement Component 5  
Authors : Fredslund, F.; Andersen, G.R.  
Deposited on : unknown  
Resolution : 3.10 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20030345
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	rb-20030345

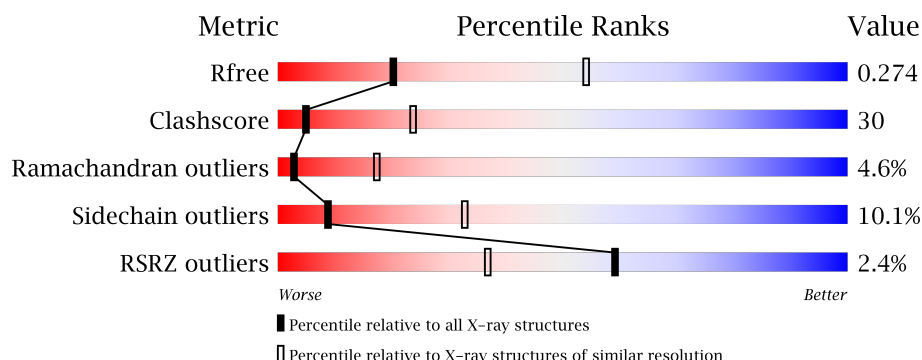
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.10 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))
$R_{free}$	100719	1001 (3.12-3.08)
Clashscore	112137	1099 (3.12-3.08)
Ramachandran outliers	110173	1057 (3.12-3.08)
Sidechain outliers	110143	1057 (3.12-3.08)
RSRZ outliers	101464	1006 (3.12-3.08)

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	1676	<div> <div>3%</div> <div> <div></div> <div>45%</div> <div>44%</div> <div>8%</div> <div>••</div> </div> </div>
1	B	1676	<div> <div>%</div> <div> <div></div> <div>42%</div> <div>39%</div> <div>6%</div> <div>12%</div> </div> </div>

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 24655 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.

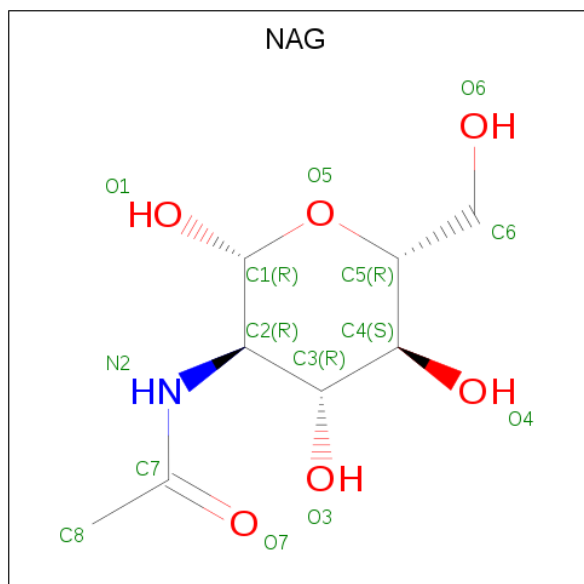
- Molecule 1 is a protein called Complement C5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1625	Total	C	N	O	S	0	0	0
			12861	8239	2111	2458	53			
1	B	1481	Total	C	N	O	S	0	0	0
			11701	7493	1930	2232	46			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	802	ILE	VAL	SEE REMARK 999	UNP P01031
B	802	ILE	VAL	SEE REMARK 999	UNP P01031

- Molecule 2 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	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
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		

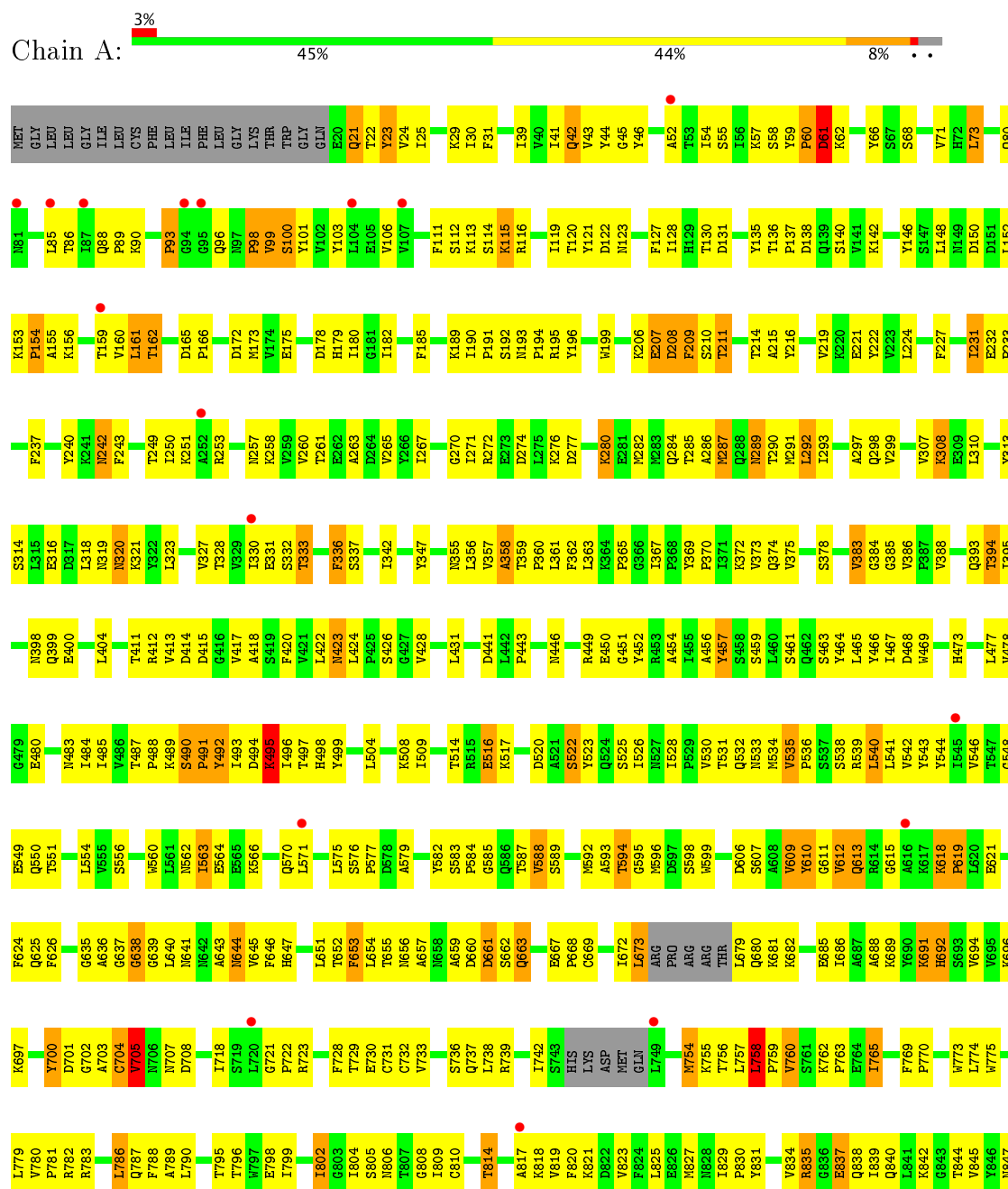
- Molecule 3 is CADMIUM ION (three-letter code: CD) (formula: Cd).

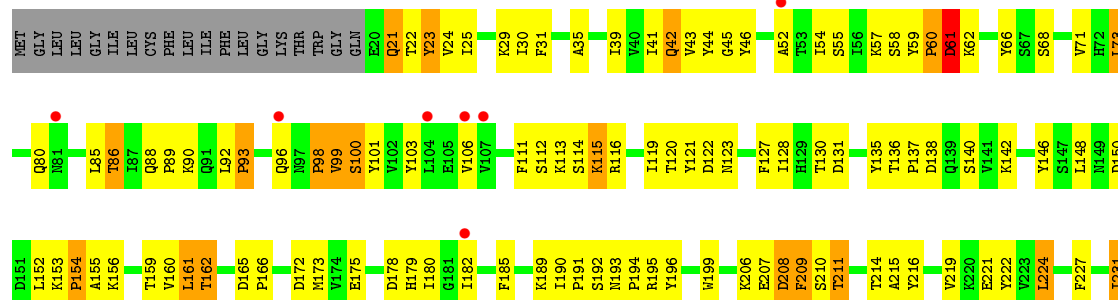
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	4	Total	Cd	0	0
			4	4		
3	A	5	Total	Cd	0	0
			5	5		

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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: Complement C5





A1822	A1252	D1165	K1091	H1002	K925	Y848	L779	K697	Q625	E549	E480	Q393	Y313	E232
L1323	Y1253	T1166	Y1092	L1003	K926	R849	V780	Y700	F626	Q550	H481	T394	S314	P233
H1324	A1254	K1170	N1096	A1008	L927	S850	P781	D701	G635	T551	L482	I395	L315	
Y1326	L1256				R928	G852	R783	G702	A636	A552	I483	N398	D317	F237
H1332	T1257	N1173	S1099	L1012	V929	M853	L786	A703	G638	S556	I484	Q399	L318	Y240
F1333	L1175	F1174	I1100	V1015	R930	Q854	Q787	C704	G639		N241	E400	N319	N241
L1334	L1176	L1176	L1104	V1016	P931	F855	F788	Y705	L640	W560	T487		N320	N242
G1335	E1177	E1177	L1105	P1017	E932	R856	A789	N707	N641	L561	N488	L404	K321	F243
	L1261		W1106	V1018	G933	V857	L790	D708	N642	N562	K489		Y322	
			L1107	V1019	V934	S860	L795		N643	E564	S490	T411	L323	T249
			L1108	F1020	K935	R861	T796	L718	A643	R491	N492	R412	V327	K250
			E1109	E937	R936	V862	T796	S719	N644	Y492	I493	D414	T328	A252
			N1110	S938	E937	W862		L720	V645	I493	D494	D415	Y329	R253
			Y1111	S939	S938	C866	E798	G721	F646	K495	I496	G416	I330	
			Q1112	Y939	E798	T867	L799	P722	H647	N497	T497	V417	E331	K258
			L1113	S940	E799	S868	N806	C731	N855	A418	S332	A418	S332	K258
			N1114	G941	L799	E869	N807	G732	N856	S419	V259	F420	T333	V259
			N1115	V942	L799	S870	G808	V733	A657	F421	H498	F421	F336	V260
				T943	L944	P871	L809		N858	Y582	I509	S426	Y347	V265
				L944		VAL	C810	S736	A859	S583	T514	G427	S350	Y266
				W1031	I949	ILE	N806	Q737	D660	G585	E516	V428	P351	I267
				L1025		ASP	N807	L738	N856	Q586	E516	L431	N355	G270
				G1028		GLN	L809	R739	Q663	W588	K517	K436	L356	I271
				N1029		THR	C810		Q663	S589	D520	P443	V357	R272
				W1031		THR				M592	A521	M446	A358	D274
				F1034		ASP	L817	I742	E667	M592	I521		T359	D277
				L1039		GLY	K318	S743	P668	A593	T594		P360	
				L1040		THR	R819	HIS	C669	A593	Y523		L361	
				E1041		THR	R820	LVS	I672	Q524	Q524		F362	K280
				L1049		THR	R821	ASP	L673	Q525	S525		L363	E281
				L1053		THR	R822	GLN	ARG	G595	S525		G451	E282
				I1056		THR	R823	GLN	ARG	M596	S526		P365	M283
				M1057		THR	R824	L749	PRG	D597	I527		R364	M283
				M1058		THR	L825	G750	ARG	S598	I528		Q366	Q284
				Y1059		THR	E826	R751	THR	W599	P529		I367	T285
				D1063		THR	R827	L752		D606	V530		P368	T285
				Y1064		THR	N828	R752	L679	S607	T531		Y369	A286
				S1065		THR	R829	H753	Q680	Q681	Q532		P370	N287
				Y1066		THR	P830	H754	K681	A456	Q532		I371	Q288
				S1067		THR	R831	K755	K682	Y457	N633		G372	N289
				V1068		THR	E834	L756	I683	S458	S459		K374	M291
				W1069		THR	R835	L757	E884	G611	G611		Q374	M291
				K1070		THR	R836	P759	E885	Y610	Y635		Y375	L292
				S1075		THR	E837	V760	I686	Y612	S537		Y464	L293
				T1076		THR	R838	W760	A687	Q613	S538		L465	I293
				W1077		THR	Q838	I765		R614	R539		L466	
				L1078		THR	R839	I765	A688	G615	L540		I467	A297
				R1084		THR	Q840	F769	K689	A616	L541		D468	Q298
				Y951		THR	L841	P770	Y690	K617	Y542		W469	V299
				V951		THR	K842	K691	Y690	K617	Y542		G384	
				G956		THR	E843	W773	E885	P619	Y544		G385	V307
				L1000		THR	T844	W773	S883	L620	I545		V386	K308
				T1001		THR	E845	W775	V694	E621	V546		P387	E309
						THR	W846	W775	G696		V546		V388	L310
						THR	N847		K696	F624	G548			

[illegible]



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	144.26 Å   144.26 Å   241.00 Å 90.00°   90.00°   120.00°	Depositor
Resolution (Å)	29.29 – 3.10 29.29 – 3.11	Depositor EDS
% Data completeness (in resolution range)	97.5 (29.29-3.10) 97.4 (29.29-3.11)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	0.04	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.55 (at 3.11 Å)	Xtriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, $R_{free}$	0.236 , 0.281 0.229 , 0.274	Depositor DCC
$R_{free}$ test set	4916 reflections (4.99%)	DCC
Wilson B-factor (Å <sup>2</sup> )	94.3	Xtriage
Anisotropy	0.350	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 91.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.008 for -h,-k,l 0.468 for h,-h-k,-l 0.009 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	24655	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	140.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.47% 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: NAG, CD

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	0.26	0/13137	0.47	1/17820 (0.0%)
1	B	0.26	0/11954	0.47	1/16219 (0.0%)
All	All	0.26	0/25091	0.47	2/34039 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	1488	LEU	CA-CB-CG	5.05	126.92	115.30
1	A	1488	LEU	CA-CB-CG	5.05	126.91	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	12861	0	12817	794	0
1	B	11701	0	11669	700	0
2	A	42	0	38	2	0
2	B	42	0	38	2	0
3	A	5	0	0	0	0
3	B	4	0	0	0	0
All	All	24655	0	24562	1484	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 30.

The worst 5 of 1484 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:1538:GLU:HG2	1:A:1539:LEU:HG	1.31	1.12
1:A:253:ARG:HH22	1:A:257:ASN:HA	1.13	1.11
1:B:253:ARG:HH22	1:B:257:ASN:HA	1.12	1.10
1:A:1279:ARG:HG3	1:A:1284:PHE:HB2	1.30	1.08
1:B:1279:ARG:HG3	1:B:1284:PHE:HB2	1.31	1.06

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	1615/1676 (96%)	1282 (79%)	257 (16%)	76 (5%)	3	17
1	B	1471/1676 (88%)	1191 (81%)	214 (14%)	66 (4%)	3	17
All	All	3086/3352 (92%)	2473 (80%)	471 (15%)	142 (5%)	3	17

5 of 142 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	60	PRO
1	A	261	THR
1	A	308	LYS
1	A	336	PHE
1	A	490	SER

### 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	1441/1484 (97%)	1294 (90%)	147 (10%)	8	32
1	B	1314/1484 (88%)	1183 (90%)	131 (10%)	9	33
All	All	2755/2968 (93%)	2477 (90%)	278 (10%)	9	33

5 of 278 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1500	ARG
1	B	116	ARG
1	B	1338	VAL
1	A	1548	ARG
1	A	1626	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 82 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1459	HIS
1	B	80	GLN
1	B	1234	HIS
1	A	1463	GLN
1	A	1626	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 [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 15 ligands modelled in this entry, 9 are monoatomic - leaving 6 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	NAG	A	2001	1,2	14,14,15	0.50	0	15,19,21	0.73	0
2	NAG	A	2002	2	14,14,15	0.47	0	15,19,21	0.79	1 (6%)
2	NAG	A	2003	1	14,14,15	0.57	0	15,19,21	0.91	1 (6%)
2	NAG	B	2001	1,2	14,14,15	0.50	0	15,19,21	0.72	0
2	NAG	B	2002	2	14,14,15	0.47	0	15,19,21	0.81	1 (6%)
2	NAG	B	2003	1	14,14,15	0.60	0	15,19,21	1.09	1 (6%)

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
2	NAG	A	2001	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	2002	2	-	0/6/23/26	0/1/1/1
2	NAG	A	2003	1	-	0/6/23/26	0/1/1/1
2	NAG	B	2001	1,2	-	0/6/23/26	0/1/1/1
2	NAG	B	2002	2	-	2/6/23/26	0/1/1/1
2	NAG	B	2003	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	2002	NAG	C1-O5-C5	2.14	115.12	112.17

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	2003	NAG	O5-C1-C2	2.20	114.54	111.47
2	A	2002	NAG	C1-O5-C5	2.21	115.22	112.17
2	B	2003	NAG	O5-C1-C2	2.91	115.52	111.47

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	2002	NAG	O7-C7-N2-C2
2	B	2002	NAG	C8-C7-N2-C2

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	2001	NAG	2	0
2	B	2001	NAG	2	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 [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	1625/1676 (96%)	0.17	49 (3%)	51 27	73, 136, 224, 290	0
1	B	1481/1676 (88%)	0.10	25 (1%)	70 49	75, 131, 205, 267	0
All	All	3106/3352 (92%)	0.14	74 (2%)	59 37	73, 133, 215, 290	0

The worst 5 of 74 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	871	PRO	10.0
1	B	871	PRO	6.1
1	A	1534	GLN	4.9
1	A	1635	TYR	4.1
1	B	1518	LYS	4.1

### 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 carbohydrates 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. 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( $\text{\AA}^2$ )	Q<0.9
3	CD	B	2006	1/1	0.76	0.19	-	183,183,183,183	0
2	NAG	A	2001	14/15	0.81	0.33	-	257,259,270,271	0
2	NAG	B	2003	14/15	0.81	0.57	-	246,252,259,262	0
3	CD	A	2007	1/1	0.94	0.21	-	184,184,184,184	0
3	CD	B	2007	1/1	0.78	0.13	-	264,264,264,264	0
3	CD	A	2006	1/1	0.68	0.18	-	189,189,189,189	0
3	CD	B	2004	1/1	0.99	0.18	-	239,239,239,239	0
2	NAG	B	2001	14/15	0.89	0.38	-	263,269,274,275	0
2	NAG	A	2003	14/15	0.85	0.25	-	215,223,233,237	0
3	CD	B	2005	1/1	0.75	0.20	-	181,181,181,181	0
2	NAG	A	2002	14/15	0.79	0.24	-	249,258,261,262	0
3	CD	A	2005	1/1	0.97	0.16	-	243,243,243,243	0
2	NAG	B	2002	14/15	0.58	0.51	-	251,259,262,262	0
3	CD	A	2008	1/1	0.73	0.11	-	282,282,282,282	0
3	CD	A	2004	1/1	0.95	0.35	-	123,123,123,123	1

## 6.5 Other polymers [i](#)

There are no such residues in this entry.