



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 5, 2017 – 09:28 AM EST

PDB ID : 4E0S
Title : Crystal Structure of C5b-6
Authors : Aleshin, A.E.; Stec, B.; DiScipio, R.; Liddington, R.C.
Deposited on : unknown
Resolution : 4.21 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

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

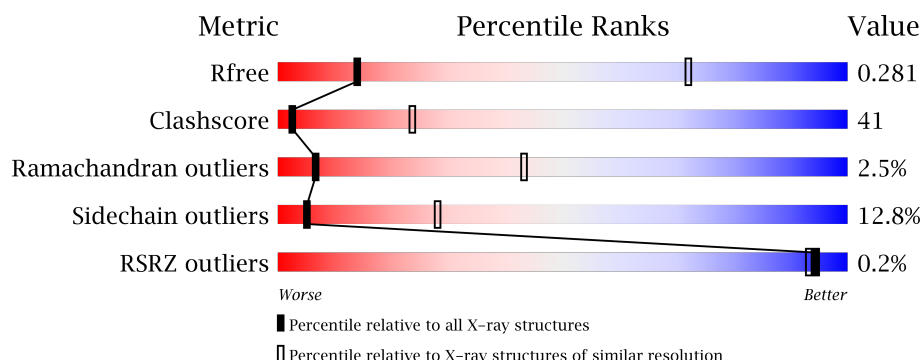
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 4.21 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	1180 (4.84-3.60)
Clashscore	112137	1027 (4.76-3.66)
Ramachandran outliers	110173	1006 (4.80-3.64)
Sidechain outliers	110143	1015 (4.80-3.62)
RSRZ outliers	101464	1003 (4.80-3.62)

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 ≥ 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	
2	B	913	

2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 19475 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
			Total	C	N	O	S			
1	A	1552	12306	7891	2011	2359	45	0	0	0

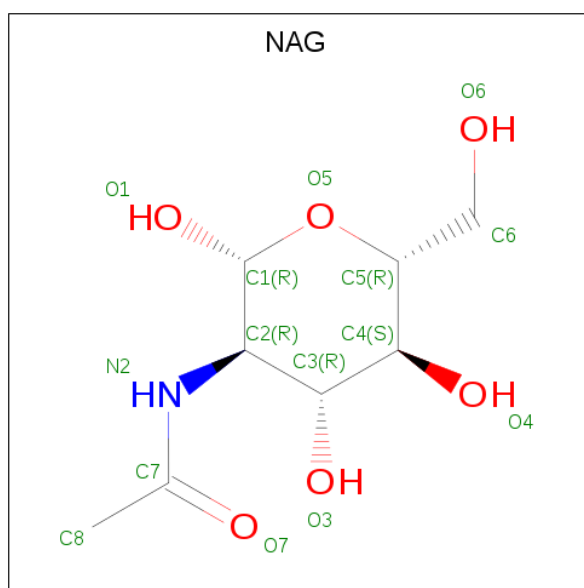
There is a discrepancy between the modelled and reference sequences:

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

- Molecule 2 is a protein called Complement component C6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	898	7046	4353	1239	1383	71	0	0	0

- Molecule 3 is 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	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		

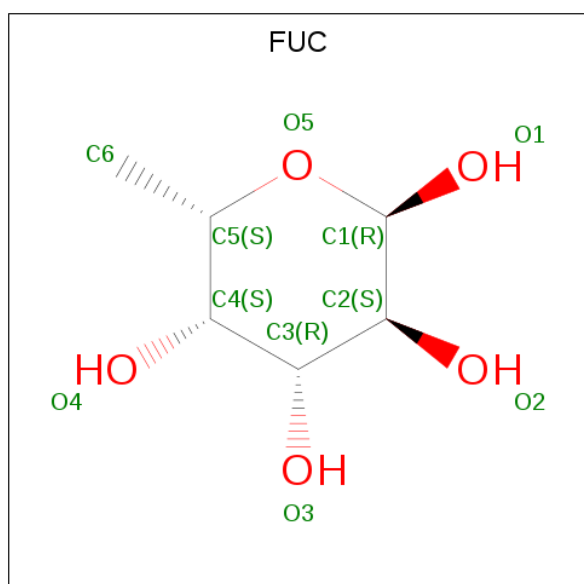
- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Na	0	0
			1	1		

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

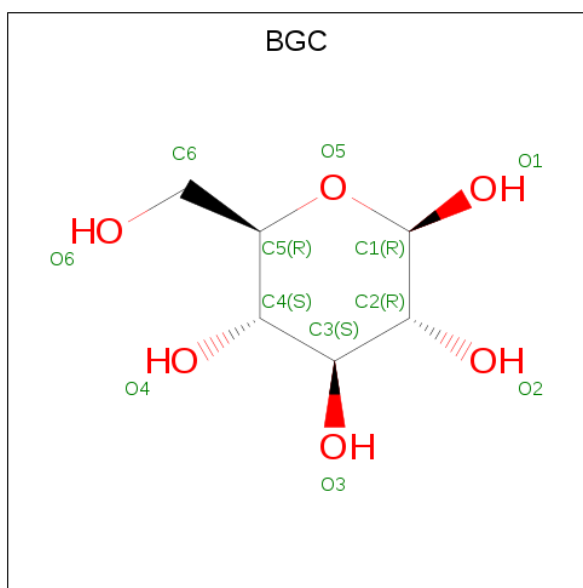
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	Ca	0	0
			1	1		

- Molecule 6 is ALPHA-L-FUCOSE (three-letter code: FUC) (formula: C₆H₁₂O₅).



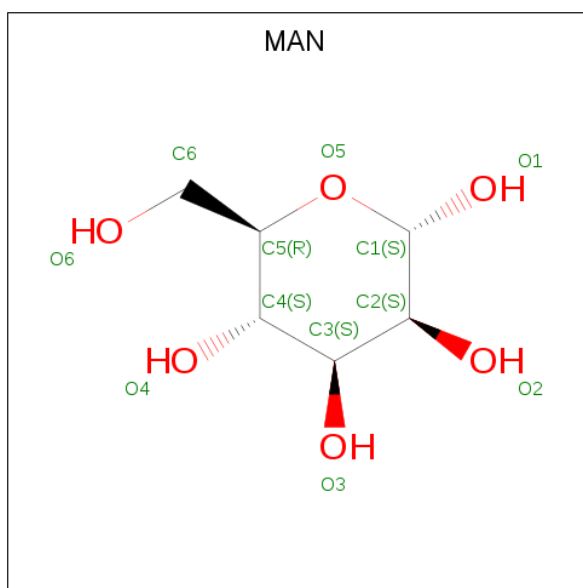
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	B	1	Total	C	O	0	0
			10	6	4		

- Molecule 7 is BETA-D-GLUCOSE (three-letter code: BGC) (formula: $C_6H_{12}O_6$).



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

- Molecule 8 is ALPHA-D-MANNOSE (three-letter code: MAN) (formula: $C_6H_{12}O_6$).

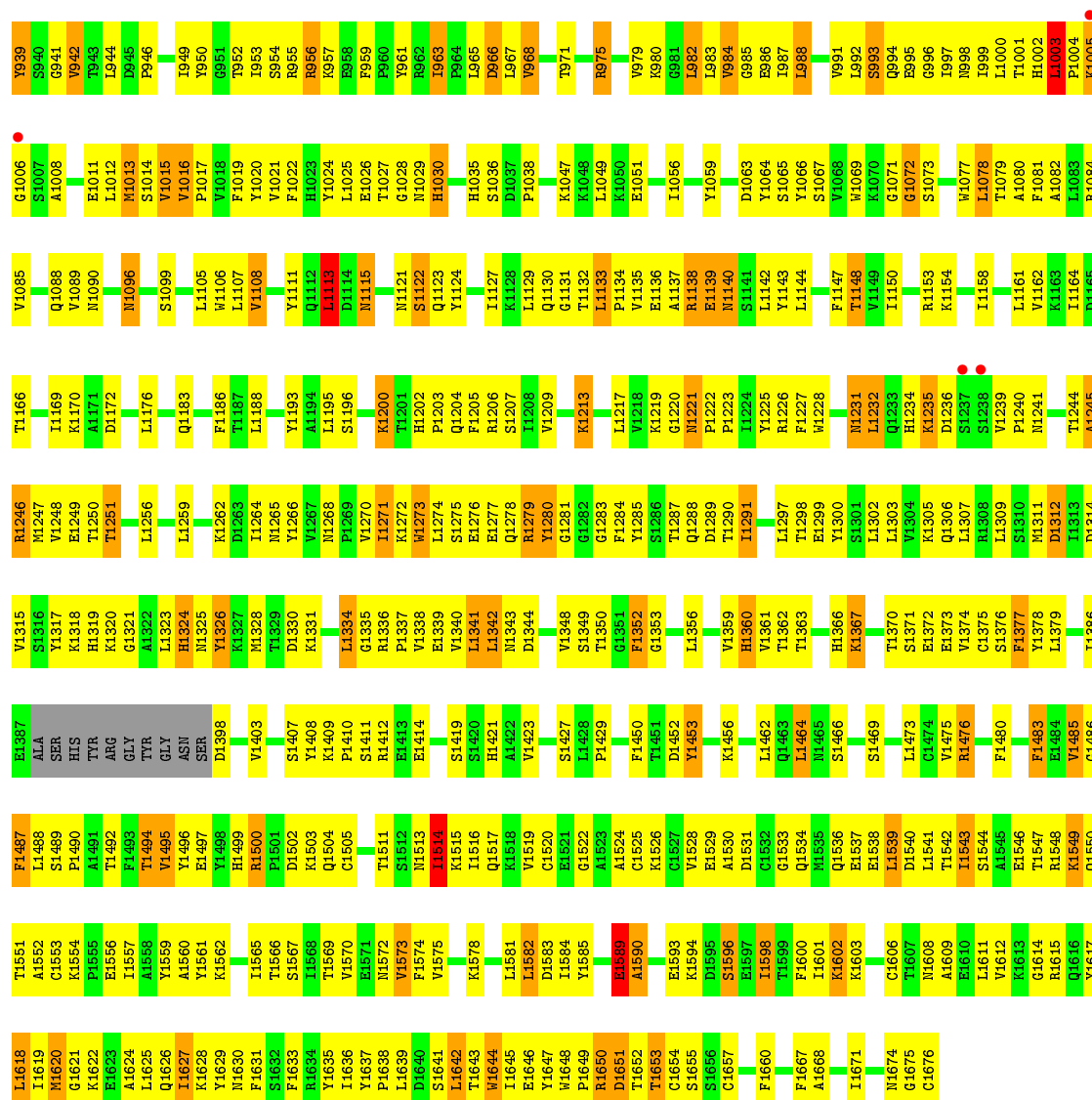


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	B	1	Total	C	O	0	0
			11	6	5		
8	B	1	Total	C	O	0	0
			11	6	5		

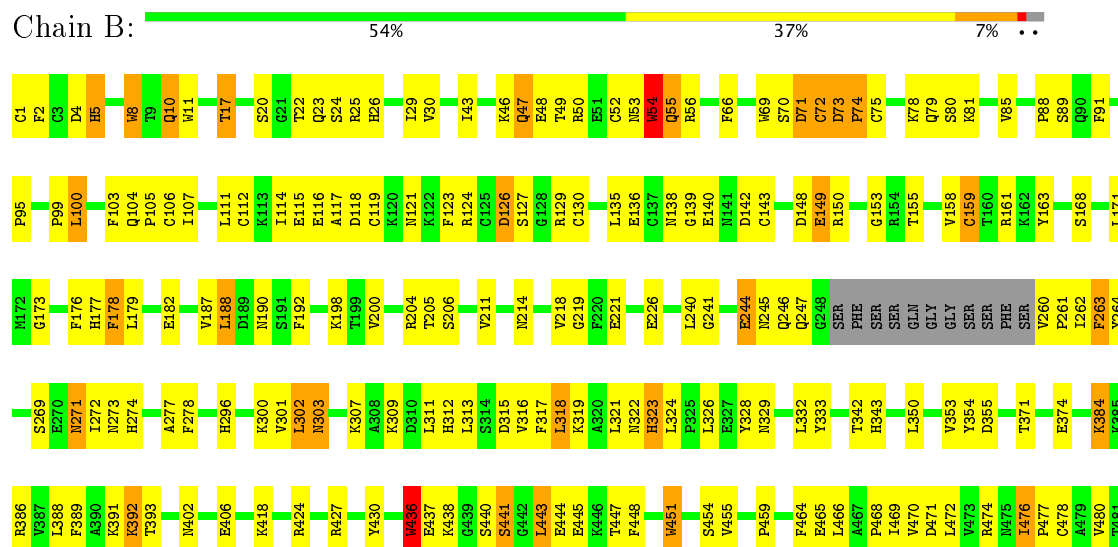
Continued on next page...

Continued from previous page...

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



• Molecule 2: Complement component C6



L906	R482	R566	R637	Q727	R823
H907	R483	E567	R638	N728	R824
	N485	C568	Q639	S729	G825
		N569	L640	W730	L826
		N570	Y641	T731	E827
		P571		P732	R828
	A489	A572	E645	P733	T829
		P573		I734	R830
	Y493	Q574	L652	S735	L831
	A494		T653	N736	
	A495	G577	G654	S737	R838
	F497		F655	L738	E839
	D498	E581	E656	T739	S840
	P499	G582		C740	
	Q501	E583	Y660	E741	Y843
		R585	Q661	K742	D844
		Q586	Y662	D743	T845
	P506			T744	C946
	N507		D668	L745	Y847
	R508	D589	G669	THR	D848
	R510	C590	T670	LYS	R849
		T591	W671	LEU	E850
		F592		LYS	
	T516	S593	V676	G750	A854
		I594		H751	S855
	Y521	N595	Q679		
		E596	R680	L754	C859
	Q528	N597	T681	G755	Y860
	E529	N598	E682		C861
	N530	G599	Q683	S773	L862
		Q600	T684		
	R533	P601	R685	S776	C867
		C602	P686		P868
	D537	I603	V687	V781	Y875
	Y538	N604	Y688	F782	C876
	K539	D605	Q689		Y877
	S540	D606	E690	D785	R878
		E607	V691		R879
		E508	L692	Y789	C880
	D544	M609	T693	F790	S881
	G545	K610			S882
	Q546	E611	L694	A794	
	N547	V612	F697		
		D613	G698	F797	E885
	C549	L614	R699	L798	R886
	N550	P615	L700		T887
	S551		Y701	K801	L888
	S552	E618	R702	C802	R889
	N553	A619	I703	L803	I890
	S554	D620		N804	C891
		S621	L709	N805	E892
	D557	G622		Q806	Y893
	A558		K713	Q807	G894
	T559	V627	G714	L808	T895
	Y561		F715	R809	
	R562	E630	Y716	F810	C898
	S563	N631	V717	L811	
	R564		A718	H812	N903
	T565	R635	G719	L822	E904
		N636			I905

4 Data and refinement statistics

Property	Value	Source
Space group	I 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	158.95Å 227.53Å 278.16Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.93 – 4.21 29.93 – 4.21	Depositor EDS
% Data completeness (in resolution range)	81.8 (29.93-4.21) 82.1 (29.93-4.21)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.98 (at 4.26Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.218 , 0.278 0.219 , 0.281	Depositor DCC
R_{free} test set	1529 reflections (5.35%)	DCC
Wilson B-factor (Å ²)	113.5	Xtriage
Anisotropy	0.396	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.20 , 124.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	19475	wwPDB-VP
Average B, all atoms (Å ²)	192.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.57% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: BGC, NAG, NA, CA, FUC, 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	0.58	6/12576 (0.0%)	0.80	4/17068 (0.0%)
2	B	0.58	5/7193 (0.1%)	0.78	5/9708 (0.1%)
All	All	0.58	11/19769 (0.1%)	0.79	9/26776 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	2
2	B	0	3
All	All	0	5

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	917	TRP	CD2-CE2	5.98	1.48	1.41
1	A	1273	TRP	CD2-CE2	5.97	1.48	1.41
1	A	797	TRP	CD2-CE2	5.75	1.48	1.41
1	A	1077	TRP	CD2-CE2	5.30	1.47	1.41
2	B	436	TRP	CD2-CE2	5.27	1.47	1.41

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	570	ASN	C-N-CD	-8.57	101.73	120.60
2	B	731	THR	C-N-CD	-7.08	105.03	120.60
1	A	794	LEU	CB-CG-CD2	-5.78	101.18	111.00
2	B	627	VAL	CB-CA-C	-5.50	100.96	111.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	560	TYR	CA-CB-CG	5.41	123.69	113.40

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	90	LYS	Peptide
1	A	985	GLY	Peptide
2	B	391	LYS	Peptide
2	B	599	GLY	Peptide
2	B	731	THR	Peptide

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	12306	0	12238	1116	0
2	B	7046	0	6708	482	0
3	A	28	0	25	1	0
3	B	28	0	25	3	0
4	A	1	0	0	0	0
5	B	1	0	0	0	0
6	B	10	0	9	0	0
7	B	11	0	10	0	0
8	B	44	0	40	7	0
All	All	19475	0	19055	1562	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 41.

The worst 5 of 1562 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:1598:ILE:HG21	1:A:1637:TYR:CE2	1.45	1.52
1:A:21:GLN:NE2	1:A:45:GLY:HA2	1.32	1.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1013:MET:SD	1:A:1129:LEU:HG	1.65	1.35
1:A:1127:ILE:HD12	1:A:1130:GLN:NE2	1.36	1.35
1:A:1539:LEU:HD22	1:A:1540:ASP:N	1.48	1.27

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	1544/1676 (92%)	1339 (87%)	170 (11%)	35 (2%)	7	45
2	B	892/913 (98%)	770 (86%)	96 (11%)	26 (3%)	5	40
All	All	2436/2589 (94%)	2109 (87%)	266 (11%)	61 (2%)	6	43

5 of 61 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	60	PRO
1	A	89	PRO
1	A	191	PRO
1	A	335	GLY
1	A	490	SER

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	1379/1484 (93%)	1183 (86%)	196 (14%)	4	25
2	B	797/810 (98%)	715 (90%)	82 (10%)	8	36
All	All	2176/2294 (95%)	1898 (87%)	278 (13%)	5	28

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

Mol	Chain	Res	Type
1	A	1206	ARG
1	A	1453	TYR
2	B	713	LYS
1	A	1232	LEU
1	A	1331	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 71 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1231	ASN
1	A	1499	HIS
2	B	636	ASN
1	A	1265	ASN
1	A	1332	ASN

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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 12 ligands modelled in this entry, 2 are monoatomic - leaving 10 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
3	NAG	A	2001	1,3	14,14,15	0.57	0	15,19,21	1.64	4 (26%)
3	NAG	A	2002	3	14,14,15	0.53	0	15,19,21	1.12	1 (6%)
3	NAG	B	1002	3,2	14,14,15	0.42	0	15,19,21	2.02	2 (13%)
3	NAG	B	1003	3	14,14,15	0.57	0	15,19,21	1.09	2 (13%)
6	FUC	B	1004	2,7	9,10,11	0.69	0	13,14,16	1.60	3 (23%)
7	BGC	B	1005	6	11,11,12	0.73	0	13,15,17	0.95	0
8	MAN	B	1006	2	11,11,12	0.56	0	13,15,17	1.00	1 (7%)
8	MAN	B	1007	2	11,11,12	0.69	0	13,15,17	0.94	0
8	MAN	B	1008	2	11,11,12	0.67	0	13,15,17	1.62	2 (15%)
8	MAN	B	1009	2	11,11,12	0.61	0	13,15,17	1.32	3 (23%)

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	2001	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	2002	3	-	0/6/23/26	0/1/1/1
3	NAG	B	1002	3,2	-	0/6/23/26	0/1/1/1
3	NAG	B	1003	3	-	0/6/23/26	0/1/1/1
6	FUC	B	1004	2,7	-	0/0/17/20	0/1/1/1
7	BGC	B	1005	6	-	0/2/19/22	0/1/1/1
8	MAN	B	1006	2	-	0/2/19/22	0/1/1/1
8	MAN	B	1007	2	-	0/2/19/22	0/1/1/1
8	MAN	B	1008	2	-	0/2/19/22	0/1/1/1
8	MAN	B	1009	2	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	2001	NAG	O5-C1-C2	-3.88	106.07	111.47
8	B	1009	MAN	O5-C1-C2	-2.80	106.41	110.79
3	A	2001	NAG	O4-C4-C3	-2.53	104.86	110.36
8	B	1008	MAN	O5-C1-C2	-2.42	107.00	110.79
3	B	1003	NAG	O4-C4-C3	-2.18	105.62	110.36

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	2002	NAG	1	0
3	B	1002	NAG	3	0
8	B	1007	MAN	2	0
8	B	1008	MAN	4	0
8	B	1009	MAN	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 [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, 95th 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 ²)	Q<0.9
1	A	1552/1676 (92%)	-0.64	4 (0%) 93 91	110, 177, 258, 386	0
2	B	898/913 (98%)	-0.55	0 100 100	121, 201, 282, 389	0
All	All	2450/2589 (94%)	-0.61	4 (0%) 94 93	110, 187, 269, 389	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1238	SER	5.0
1	A	1006	GLY	2.7
1	A	1237	SER	2.3
1	A	1005	LYS	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 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, 95th 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(\AA^2)	Q<0.9
8	MAN	B	1008	11/12	0.95	0.23	0.54	147,167,176,188	0
8	MAN	B	1009	11/12	0.94	0.25	0.30	178,190,201,215	0
8	MAN	B	1006	11/12	0.88	0.25	0.12	245,260,277,281	0
3	NAG	A	2001	14/15	0.93	0.21	0.06	152,160,185,186	0
8	MAN	B	1007	11/12	0.89	0.22	-0.36	205,212,234,240	0
5	CA	B	1001	1/1	0.91	0.06	-2.09	167,167,167,167	1
3	NAG	B	1003	14/15	0.80	0.53	-	269,286,304,315	0
4	NA	A	2003	1/1	0.81	0.33	-	111,111,111,111	0
3	NAG	A	2002	14/15	0.83	0.27	-	207,226,257,262	0
3	NAG	B	1002	14/15	0.81	0.39	-	205,240,279,295	0
7	BGC	B	1005	11/12	0.90	0.29	-	187,198,214,224	0
6	FUC	B	1004	10/11	0.97	0.21	-	174,188,192,196	0

6.5 Other polymers [i](#)

There are no such residues in this entry.