



# wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 5, 2023 – 12:55 PM JST

PDB ID : 8HQ1  
Title : Crystal Structure Of Human Lgi1-Adam22 Complex In Space Group C2  
Authors : Liu, H.L.; Lin, Z.H.; Xu, F.  
Deposited on : 2022-12-13  
Resolution : 4.17 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

<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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

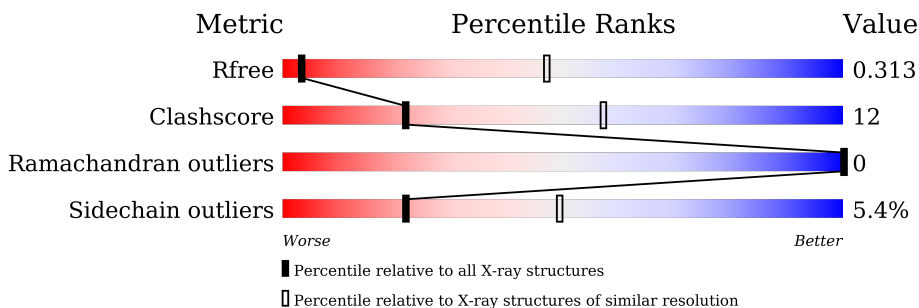
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 4.17 Å.

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



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1034 (4.60-3.76)
Clashscore	141614	1030 (4.54-3.80)
Ramachandran outliers	138981	1006 (4.58-3.78)
Sidechain outliers	138945	1037 (4.60-3.76)

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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	486	77% 21% .
1	B	486	75% 22% .
1	C	486	75% 23% .
2	D	526	70% 27% ..
2	E	526	72% 24% ..
2	F	526	70% 26% ..

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 23997 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 Disintegrin and metalloproteinase domain-containing protein 22.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	486	3728	2303	643	731	51	0	0	0
1	B	486	3728	2303	643	731	51	0	0	0
1	C	486	3728	2303	643	731	51	0	0	0

- Molecule 2 is a protein called Leucine-rich glioma-inactivated protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	D	515	4183	2694	696	778	15	0	0	0
2	E	515	4183	2694	696	778	15	0	0	0
2	F	515	4183	2694	696	778	15	0	0	0

There are 21 discrepancies between the modelled and reference sequences:

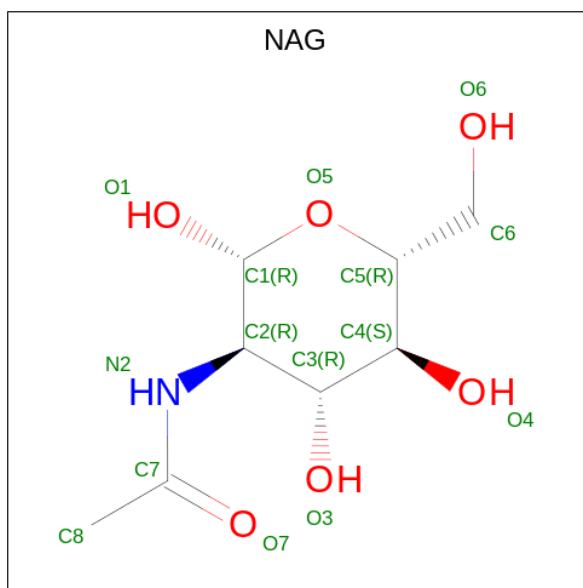
Chain	Residue	Modelled	Actual	Comment	Reference
D	32	HIS	-	expression tag	UNP O95970
D	33	HIS	-	expression tag	UNP O95970
D	34	HIS	-	expression tag	UNP O95970
D	35	HIS	-	expression tag	UNP O95970
D	36	HIS	-	expression tag	UNP O95970
D	37	HIS	-	expression tag	UNP O95970
D	38	HIS	-	expression tag	UNP O95970
E	32	HIS	-	expression tag	UNP O95970
E	33	HIS	-	expression tag	UNP O95970
E	34	HIS	-	expression tag	UNP O95970
E	35	HIS	-	expression tag	UNP O95970
E	36	HIS	-	expression tag	UNP O95970

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	37	HIS	-	expression tag	UNP O95970
E	38	HIS	-	expression tag	UNP O95970
F	32	HIS	-	expression tag	UNP O95970
F	33	HIS	-	expression tag	UNP O95970
F	34	HIS	-	expression tag	UNP O95970
F	35	HIS	-	expression tag	UNP O95970
F	36	HIS	-	expression tag	UNP O95970
F	37	HIS	-	expression tag	UNP O95970
F	38	HIS	-	expression tag	UNP O95970

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (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	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		
3	B	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	C	1	Total C N O 14 8 1 5	0	0
3	C	1	Total C N O 14 8 1 5	0	0
3	C	1	Total C N O 14 8 1 5	0	0
3	D	1	Total C N O 14 8 1 5	0	0
3	D	1	Total C N O 14 8 1 5	0	0
3	D	1	Total C N O 14 8 1 5	0	0
3	E	1	Total C N O 14 8 1 5	0	0
3	E	1	Total C N O 14 8 1 5	0	0
3	E	1	Total C N O 14 8 1 5	0	0
3	F	1	Total C N O 14 8 1 5	0	0
3	F	1	Total C N O 14 8 1 5	0	0
3	F	1	Total C N O 14 8 1 5	0	0

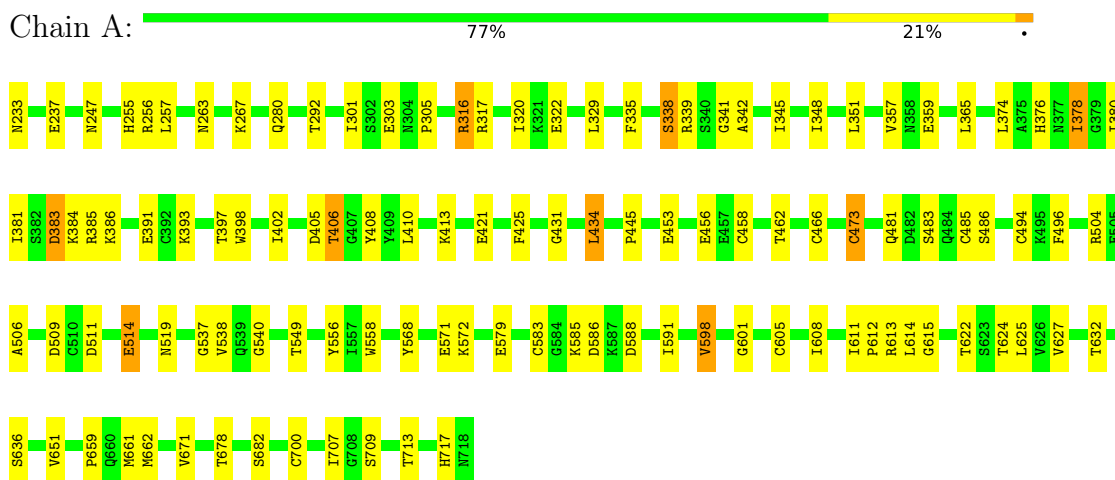
- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	3	Total Ca 3 3	0	0
4	B	3	Total Ca 3 3	0	0
4	C	3	Total Ca 3 3	0	0
4	D	1	Total Ca 1 1	0	0
4	E	1	Total Ca 1 1	0	0
4	F	1	Total Ca 1 1	0	0

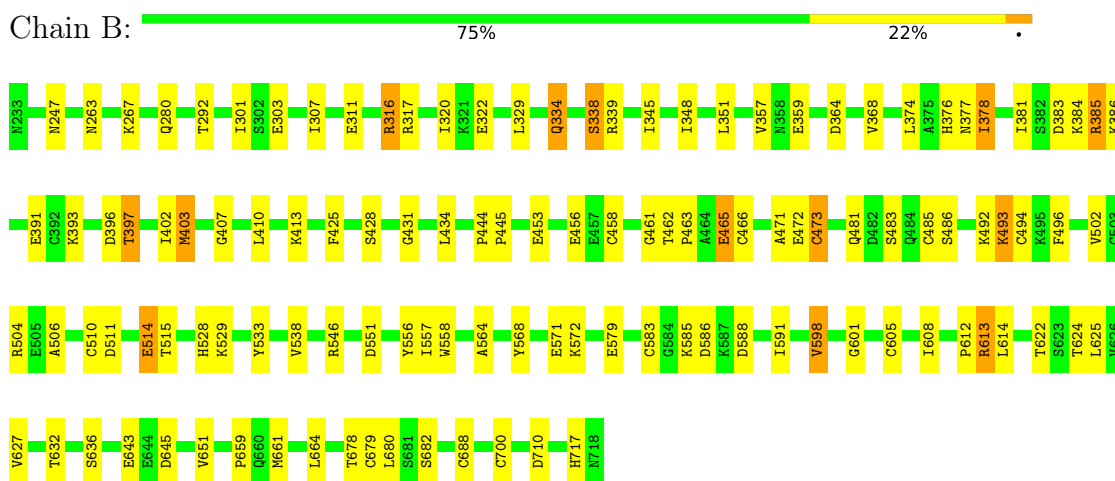
### 3 Residue-property plots [i](#)

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. 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. 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.

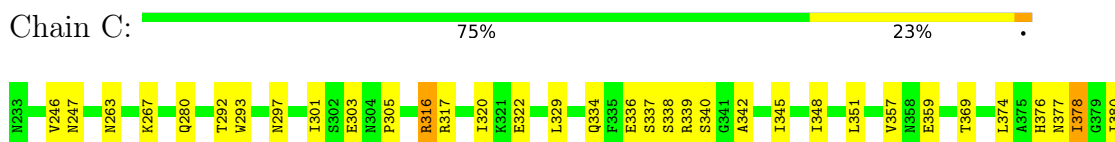
- Molecule 1: Disintegrin and metalloproteinase domain-containing protein 22



- Molecule 1: Disintegrin and metalloproteinase domain-containing protein 22

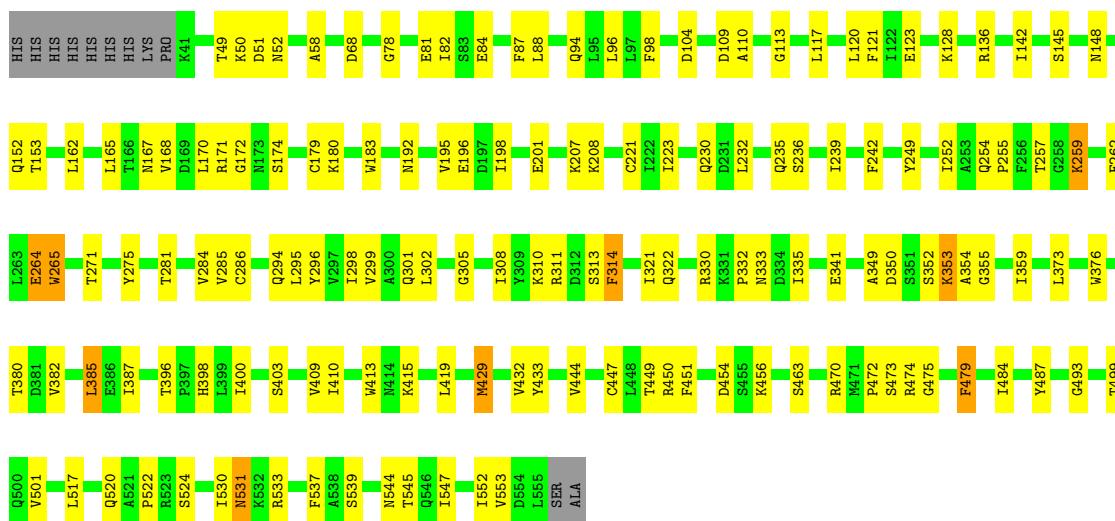


- Molecule 1: Disintegrin and metalloproteinase domain-containing protein 22

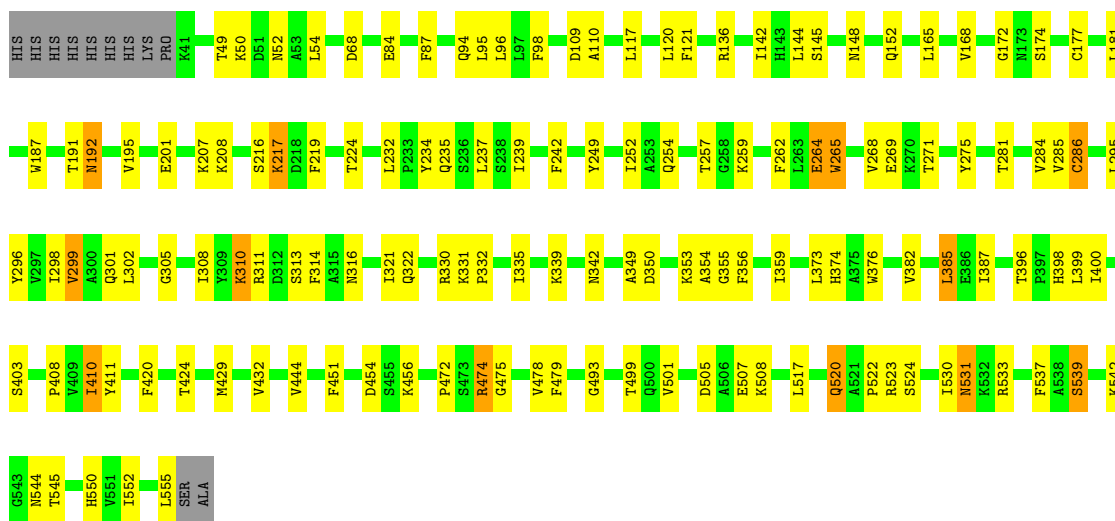




• Molecule 2: Leucine-rich glioma-inactivated protein 1

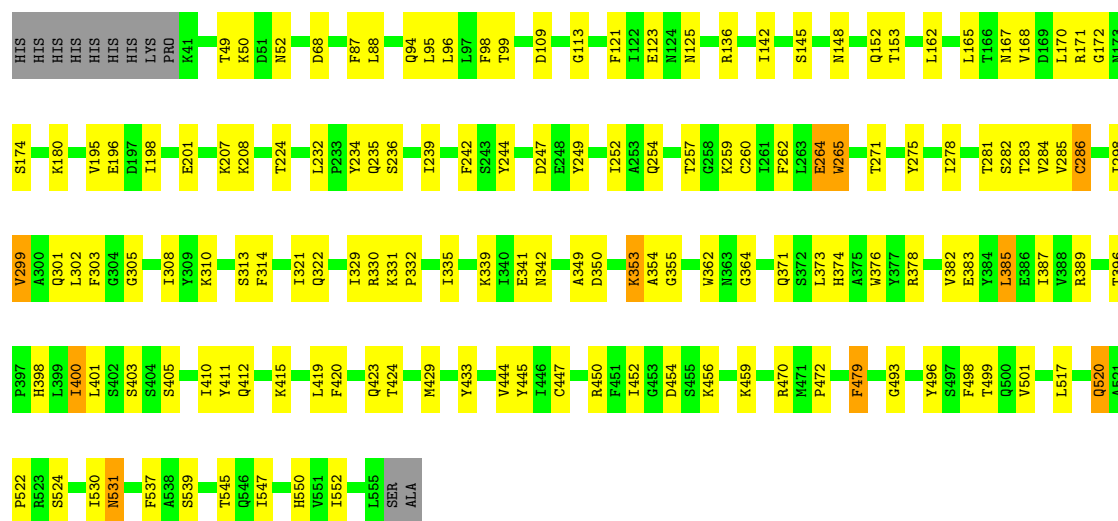


• Molecule 2: Leucine-rich glioma-inactivated protein 1



• Molecule 2: Leucine-rich glioma-inactivated protein 1

Chain F:  70% 26% ..





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	373.50Å 232.37Å 80.03Å 90.00° 94.97° 90.00°	Depositor
Resolution (Å)	37.92 – 4.17 40.11 – 4.17	Depositor EDS
% Data completeness (in resolution range)	99.0 (37.92-4.17) 99.0 (40.11-4.17)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	0.04	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.05 (at 4.13Å)	Xtrriage
Refinement program	PHENIX (1.10.1_2155)	Depositor
R, $R_{free}$	0.266 , 0.314 0.266 , 0.313	Depositor DCC
$R_{free}$ test set	2558 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	172.5	Xtrriage
Anisotropy	0.811	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 203.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	23997	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	264.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.62% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, CA

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	0.30	0/3792	0.52	0/5105
1	B	0.31	0/3792	0.54	0/5105
1	C	0.31	0/3792	0.55	0/5105
2	D	0.28	0/4290	0.50	0/5821
2	E	0.27	0/4290	0.47	0/5821
2	F	0.27	0/4290	0.49	0/5821
All	All	0.29	0/24246	0.51	0/32778

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	3728	0	3580	85	0
1	B	3728	0	3579	92	0
1	C	3728	0	3579	104	0
2	D	4183	0	4101	99	0
2	E	4183	0	4101	96	0
2	F	4183	0	4101	109	0
3	A	42	0	39	2	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	42	0	39	1	0
3	C	42	0	39	0	0
3	D	42	0	39	0	0
3	E	42	0	39	0	0
3	F	42	0	39	1	0
4	A	3	0	0	0	0
4	B	3	0	0	0	0
4	C	3	0	0	0	0
4	D	1	0	0	0	0
4	E	1	0	0	0	0
4	F	1	0	0	0	0
All	All	23997	0	23275	556	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 12.

The worst 5 of 556 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:B:345:ILE:HD12	1:B:381:ILE:CG2	1.66	1.25
1:C:345:ILE:HD12	1:C:381:ILE:CG2	1.72	1.19
1:A:345:ILE:HD12	1:A:381:ILE:CG2	1.72	1.18
2:F:121:PHE:HD2	2:F:145:SER:HB3	1.10	1.09
1:C:403:MET:HE3	1:C:403:MET:HA	1.29	1.07

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	484/486 (100%)	463 (96%)	21 (4%)	0	<b>100</b> <b>100</b>

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	484/486 (100%)	466 (96%)	18 (4%)	0	100	100
1	C	484/486 (100%)	465 (96%)	19 (4%)	0	100	100
2	D	513/526 (98%)	498 (97%)	15 (3%)	0	100	100
2	E	513/526 (98%)	497 (97%)	16 (3%)	0	100	100
2	F	513/526 (98%)	498 (97%)	15 (3%)	0	100	100
All	All	2991/3036 (98%)	2887 (96%)	104 (4%)	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	420/420 (100%)	399 (95%)	21 (5%)	24	51
1	B	420/420 (100%)	394 (94%)	26 (6%)	18	46
1	C	420/420 (100%)	400 (95%)	20 (5%)	25	52
2	D	473/483 (98%)	442 (93%)	31 (7%)	16	43
2	E	473/483 (98%)	448 (95%)	25 (5%)	22	50
2	F	473/483 (98%)	450 (95%)	23 (5%)	25	52
All	All	2679/2709 (99%)	2533 (95%)	146 (5%)	22	49

5 of 146 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	479	PHE
2	F	520	GLN
2	E	531	ASN
2	F	286	CYS
1	C	316	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
2	E	52	ASN
2	E	57	ASN
1	B	334	GLN
1	C	418	ASN
2	D	52	ASN

### 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](#)

There are no monosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

Of 30 ligands modelled in this entry, 12 are monoatomic - leaving 18 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	F	603	2	14,14,15	0.51	0	17,19,21	0.71	1 (5%)
3	NAG	A	802	1	14,14,15	0.26	0	17,19,21	0.37	0
3	NAG	D	601	2	14,14,15	0.74	1 (7%)	17,19,21	0.77	0
3	NAG	C	803	1	14,14,15	0.31	0	17,19,21	0.47	0
3	NAG	D	603	2	14,14,15	0.39	0	17,19,21	0.79	1 (5%)
3	NAG	A	803	1	14,14,15	0.19	0	17,19,21	0.40	0
3	NAG	B	802	1	14,14,15	0.35	0	17,19,21	0.79	1 (5%)
3	NAG	E	601	2	14,14,15	0.36	0	17,19,21	0.43	0
3	NAG	E	603	2	14,14,15	0.25	0	17,19,21	0.57	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	NAG	F	601	2	14,14,15	0.35	0	17,19,21	0.74	0
3	NAG	B	803	1	14,14,15	0.52	0	17,19,21	1.03	1 (5%)
3	NAG	A	801	1	14,14,15	0.21	0	17,19,21	0.53	0
3	NAG	C	802	1	14,14,15	0.60	0	17,19,21	0.74	1 (5%)
3	NAG	C	801	1	14,14,15	0.63	0	17,19,21	0.44	0
3	NAG	D	602	2	14,14,15	0.24	0	17,19,21	0.40	0
3	NAG	F	602	2	14,14,15	0.26	0	17,19,21	0.39	0
3	NAG	E	602	2	14,14,15	0.34	0	17,19,21	0.63	1 (5%)
3	NAG	B	801	1	14,14,15	0.24	0	17,19,21	0.61	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	F	603	2	-	1/6/23/26	0/1/1/1
3	NAG	A	802	1	-	2/6/23/26	0/1/1/1
3	NAG	D	601	2	-	0/6/23/26	0/1/1/1
3	NAG	C	803	1	-	0/6/23/26	0/1/1/1
3	NAG	D	603	2	-	2/6/23/26	0/1/1/1
3	NAG	A	803	1	-	0/6/23/26	0/1/1/1
3	NAG	B	802	1	-	2/6/23/26	0/1/1/1
3	NAG	E	601	2	-	1/6/23/26	0/1/1/1
3	NAG	E	603	2	-	0/6/23/26	0/1/1/1
3	NAG	F	601	2	-	1/6/23/26	0/1/1/1
3	NAG	B	803	1	-	2/6/23/26	0/1/1/1
3	NAG	A	801	1	-	0/6/23/26	0/1/1/1
3	NAG	C	802	1	-	0/6/23/26	0/1/1/1
3	NAG	C	801	1	-	1/6/23/26	0/1/1/1
3	NAG	D	602	2	-	0/6/23/26	0/1/1/1
3	NAG	F	602	2	-	0/6/23/26	0/1/1/1
3	NAG	E	602	2	-	2/6/23/26	0/1/1/1
3	NAG	B	801	1	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	601	NAG	O5-C1	-2.44	1.39	1.43

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
3	B	803	NAG	C1-O5-C5	3.94	117.53	112.19
3	D	603	NAG	C1-O5-C5	2.89	116.11	112.19
3	F	603	NAG	C1-O5-C5	2.29	115.30	112.19
3	E	602	NAG	C1-O5-C5	2.19	115.15	112.19
3	B	802	NAG	C1-O5-C5	2.11	115.05	112.19

There are no chirality outliers.

5 of 14 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	D	603	NAG	O5-C5-C6-O6
3	D	603	NAG	C4-C5-C6-O6
3	B	803	NAG	O5-C5-C6-O6
3	A	802	NAG	O5-C5-C6-O6
3	B	802	NAG	C4-C5-C6-O6

There are no ring outliers.

4 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	803	NAG	1	0
3	F	601	NAG	1	0
3	B	803	NAG	1	0
3	A	801	NAG	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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.