



# wwPDB X-ray Structure Validation Summary Report

Feb 26, 2014 – 03:41 PM GMT

PDB ID : 2Q7N  
Title : Crystal structure of Leukemia inhibitory factor in complex with LIF receptor (domains 1-5)  
Authors : Huyton, T.; Zhang, J.G.; Nicola, N.A.; Garrett, T.P.J.  
Deposited on : 2007-06-07  
Resolution : 4.00 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

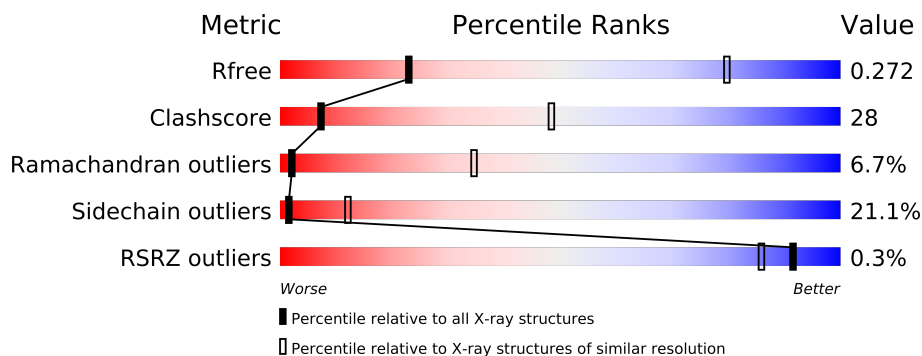
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 4.00 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))
$R_{free}$	66092	1035 (4.52-3.46)
Clashscore	79885	1235 (4.50-3.50)
Ramachandran outliers	78287	1170 (4.50-3.50)
Sidechain outliers	78261	1156 (4.50-3.50)
RSRZ outliers	66119	1035 (4.52-3.46)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	488	
1	C	488	
2	B	180	
2	D	180	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
5	NAG	A	601	-	X
5	NAG	C	501	-	X

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 10839 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Leukemia inhibitory factor receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	480	Total	C	N	O	S	0	0	0
			3739	2369	642	711	17			
1	C	480	Total	C	N	O	S	0	0	0
			3739	2369	642	711	17			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	ASP	-	CLONING ARTIFACT	UNP P42703
A	0	TYR	-	CLONING ARTIFACT	UNP P42703
A	1	LYS	-	CLONING ARTIFACT	UNP P42703
A	2	ASP	-	CLONING ARTIFACT	UNP P42703
A	3	ASP	-	CLONING ARTIFACT	UNP P42703
A	4	ASP	-	CLONING ARTIFACT	UNP P42703
A	5	ASP	-	CLONING ARTIFACT	UNP P42703
A	6	LYS	-	CLONING ARTIFACT	UNP P42703
C	-1	ASP	-	CLONING ARTIFACT	UNP P42703
C	0	TYR	-	CLONING ARTIFACT	UNP P42703
C	1	LYS	-	CLONING ARTIFACT	UNP P42703
C	2	ASP	-	CLONING ARTIFACT	UNP P42703
C	3	ASP	-	CLONING ARTIFACT	UNP P42703
C	4	ASP	-	CLONING ARTIFACT	UNP P42703
C	5	ASP	-	CLONING ARTIFACT	UNP P42703
C	6	LYS	-	CLONING ARTIFACT	UNP P42703

- Molecule 2 is a protein called Leukemia inhibitory factor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	180	Total	C	N	O	S	0	0	0
			1387	884	245	251	7			
2	D	180	Total	C	N	O	S	0	0	0
			1387	884	245	251	7			

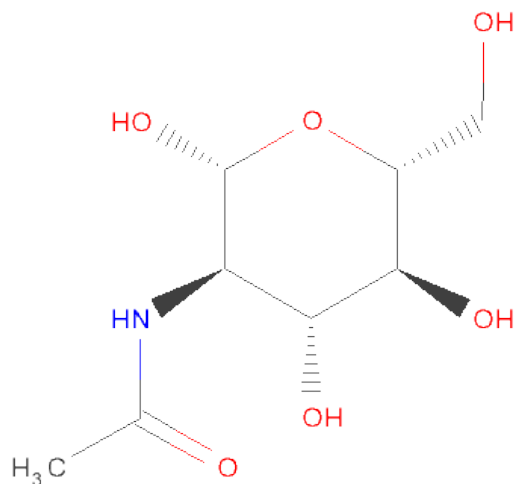
- Molecule 3 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	E	3	Total	C	N	O	0	0
			38	22	2	14		
3	H	3	Total	C	N	O	0	0
			38	22	2	14		
3	I	3	Total	C	N	O	0	0
			38	22	2	14		
3	N	3	Total	C	N	O	0	0
			38	22	2	14		
3	O	3	Total	C	N	O	0	0
			38	22	2	14		

- Molecule 4 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	F	2	Total	C	N	O	0	0
			28	16	2	10		
4	G	2	Total	C	N	O	0	0
			28	16	2	10		
4	L	2	Total	C	N	O	0	0
			28	16	2	10		
4	M	2	Total	C	N	O	0	0
			28	16	2	10		
4	P	2	Total	C	N	O	0	0
			28	16	2	10		
4	Q	2	Total	C	N	O	0	0
			28	16	2	10		
4	S	2	Total	C	N	O	0	0
			28	16	2	10		

- Molecule 5 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

- Molecule 6 is a polymer of unknown type called SUGAR (4-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	J	4	Total	C	N	O	0	0
			50	28	2	20		

- Molecule 7 is a polymer of unknown type called SUGAR (4-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	K	4	Total	C	N	O	0	0
			49	28	2	19		

- Molecule 8 is a polymer of unknown type called SUGAR (5-MER).

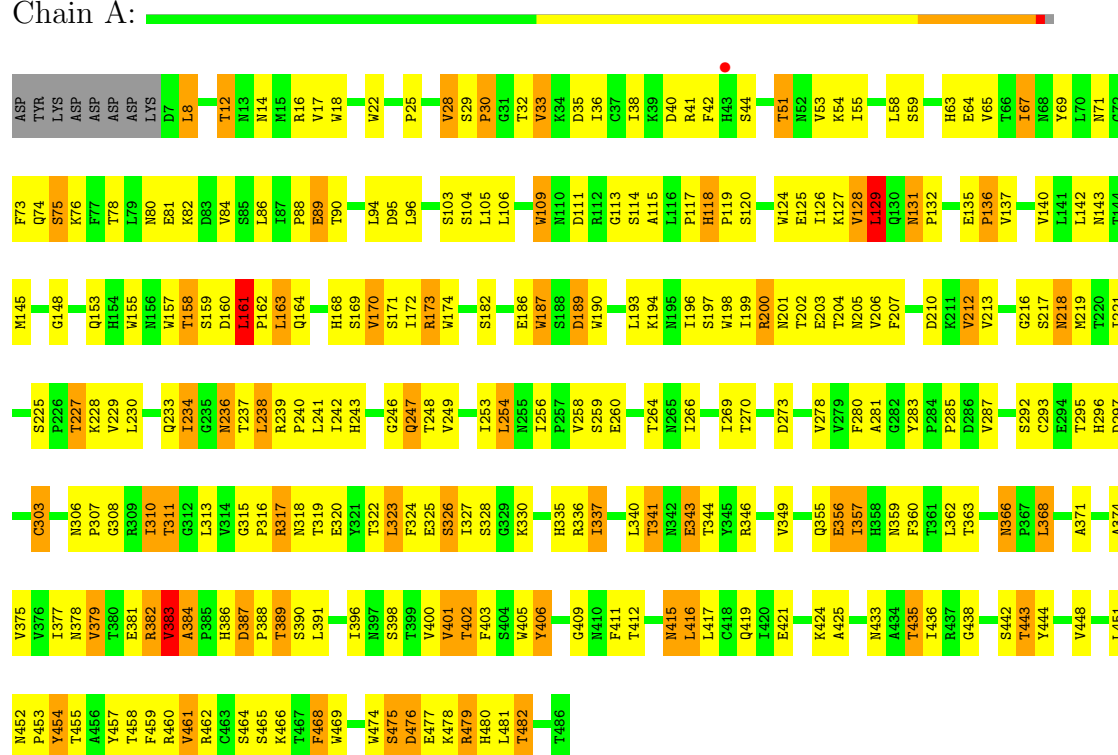
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	R	5	Total	C	N	O	0	0
			60	34	2	24		

### 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 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.

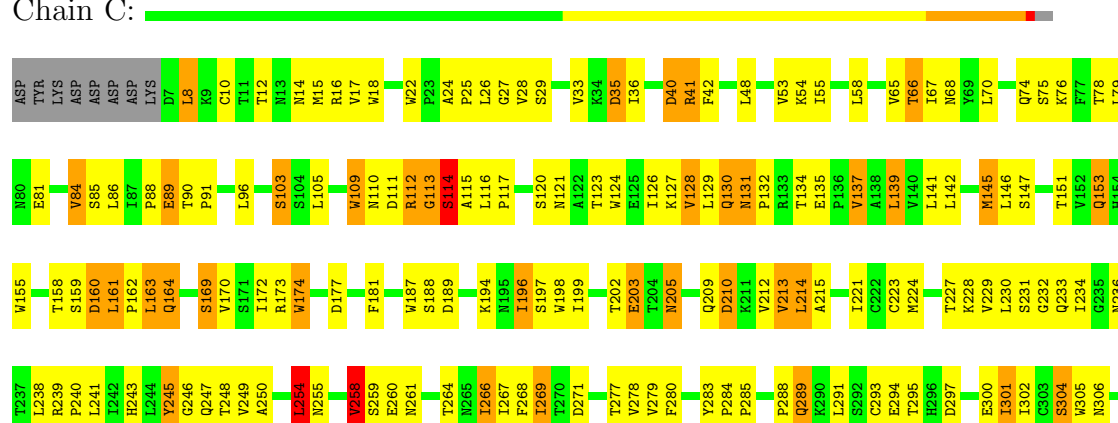
- Molecule 1: Leukemia inhibitory factor receptor

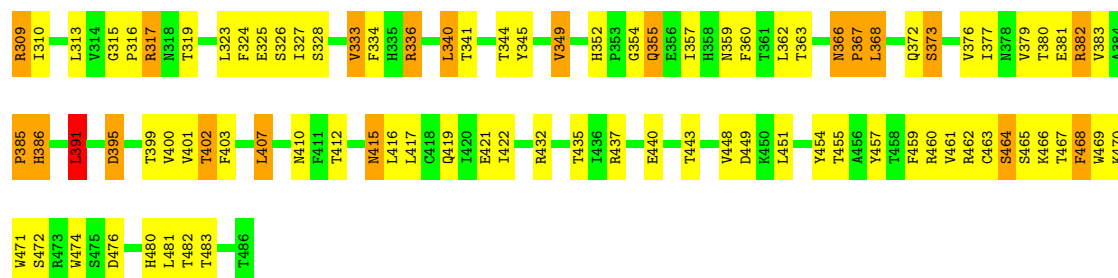
Chain A:



- Molecule 1: Leukemia inhibitory factor receptor

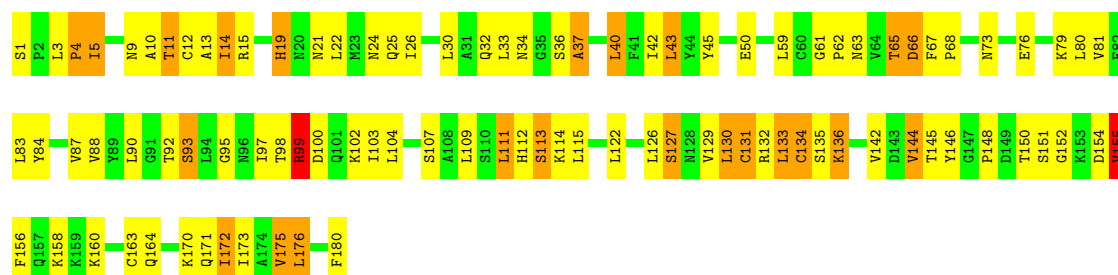
Chain C:





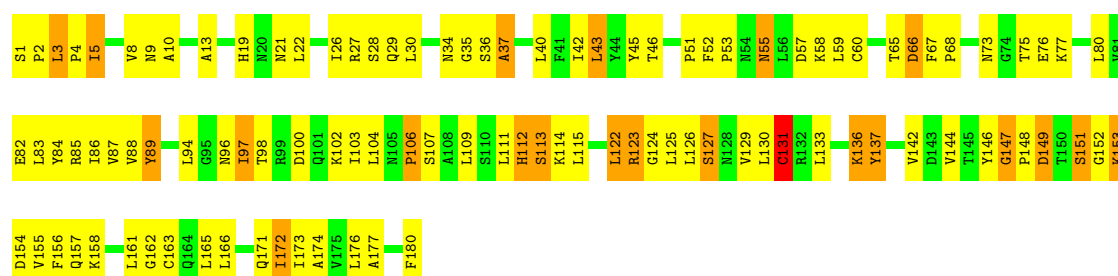
- Molecule 2: Leukemia inhibitory factor

Chain B:



- Molecule 2: Leukemia inhibitory factor

Chain D:



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	193.48Å 240.13Å 202.63Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 4.00 19.98 – 4.01	Depositor EDS
% Data completeness (in resolution range)	93.6 (20.00-4.00) 93.6 (19.98-4.01)	Depositor EDS
$R_{merge}$	0.23	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.84 (at 4.07Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.237 , 0.287 0.231 , 0.272	Depositor DCC
$R_{free}$ test set	1852 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	119.0	Xtriage
Anisotropy	0.286	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.20 , 95.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.28$	Xtriage
Outliers	0 of 37091 reflections	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	10839	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	102.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.76% of the height of the origin peak. No significant pseudotranslation is detected.*

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



## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: FUC, NAG, 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	0.76	2/3842 (0.1%)	0.82	2/5267 (0.0%)
1	C	0.73	0/3842	0.82	2/5267 (0.0%)
2	B	0.73	0/1416	0.77	0/1923
2	D	0.73	0/1416	0.81	0/1923
All	All	0.74	2/10516 (0.0%)	0.81	4/14380 (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
1	C	0	2
3	O	1	0
6	J	1	0
7	K	1	0
8	R	1	0
All	All	4	4

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	28	VAL	CA-CB	5.65	1.66	1.54
1	A	303	CYS	CB-SG	-5.65	1.72	1.81

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	254	LEU	CA-CB-CG	7.97	133.63	115.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	161	LEU	CA-CB-CG	6.46	130.15	115.30
1	A	254	LEU	CA-CB-CG	5.31	127.52	115.30
1	C	391	LEU	CA-CB-CG	5.14	127.12	115.30

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
6	J	703	MAN	C1
7	K	804	MAN	C1
3	O	501	NAG	C1
8	R	804	MAN	C1

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	239	ARG	Peptide
1	A	384	ALA	Peptide
1	C	28	VAL	Peptide
1	C	40	ASP	Peptide

## 5.2 Close contacts ⓘ

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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3739	0	3568	210	0
1	C	3739	0	3568	205	0
2	B	1387	0	1417	78	0
2	D	1387	0	1417	83	0
3	E	38	0	34	0	0
3	H	38	0	34	0	0
3	I	38	0	34	4	0
3	N	38	0	34	8	0
3	O	38	0	34	0	0
4	F	28	0	25	9	0
4	G	28	0	25	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	L	28	0	25	0	0
4	M	28	0	25	0	0
4	P	28	0	25	1	0
4	Q	28	0	25	0	0
4	S	28	0	25	7	0
5	A	14	0	13	0	0
5	C	28	0	26	0	0
6	J	50	0	43	0	0
7	K	49	0	43	5	0
8	R	60	0	52	0	0
All	All	10839	0	10492	596	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 28.

The worst 5 of 596 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
4:F:201:NAG:C6	4:F:203:NAG:H82	1.41	1.46
3:N:401:NAG:C6	3:N:403:NAG:H82	1.54	1.36
4:F:201:NAG:C6	4:F:203:NAG:C8	2.13	1.23
3:N:401:NAG:H61	3:N:403:NAG:C8	1.73	1.19
4:F:201:NAG:H62	4:F:203:NAG:C8	1.72	1.14

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	478/488 (98%)	374 (78%)	74 (16%)	30 (6%)	2	37
1	C	478/488 (98%)	383 (80%)	70 (15%)	25 (5%)	3	42
2	B	178/180 (99%)	125 (70%)	38 (21%)	15 (8%)	1	26
2	D	178/180 (99%)	115 (65%)	45 (25%)	18 (10%)	1	20

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	1312/1336 (98%)	997 (76%)	227 (17%)	88 (7%)	2	34

5 of 88 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	28	VAL
1	A	30	PRO
1	A	71	ASN
1	A	73	PHE
1	A	118	HIS

### 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	419/443 (95%)	328 (78%)	91 (22%)	1	11
1	C	419/443 (95%)	329 (78%)	90 (22%)	1	11
2	B	154/154 (100%)	123 (80%)	31 (20%)	2	14
2	D	154/154 (100%)	124 (80%)	30 (20%)	2	15
All	All	1146/1194 (96%)	904 (79%)	242 (21%)	1	12

5 of 242 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	130	LEU
1	C	103	SER
2	D	76	GLU
2	B	144	VAL
1	C	8	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	34	ASN
2	B	178	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	D	71	HIS
2	B	38	ASN
2	B	54	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

42 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	$\# Z  > 2$
3	NAG	E	101	1,3	12,14,15	0.65	0	15,19,21	1.11	1 (6%)
3	FUC	E	102	3	9,10,11	0.96	1 (11%)	10,14,16	1.36	1 (10%)
3	NAG	E	103	3	12,14,15	0.53	0	15,19,21	1.41	1 (6%)
4	NAG	F	201	1,4	12,14,15	0.73	1 (8%)	15,19,21	1.27	3 (20%)
4	NAG	F	203	4	12,14,15	0.66	0	15,19,21	1.69	2 (13%)
4	NAG	G	301	1,4	12,14,15	0.67	0	15,19,21	1.51	2 (13%)
4	NAG	G	302	4	12,14,15	0.76	1 (8%)	15,19,21	1.16	2 (13%)
3	NAG	H	401	1,3	12,14,15	0.66	0	15,19,21	0.82	1 (6%)
3	FUC	H	402	3	9,10,11	0.81	0	10,14,16	1.03	2 (20%)
3	NAG	H	403	3	12,14,15	0.61	0	15,19,21	0.93	0
3	NAG	I	501	1,3	12,14,15	0.62	0	15,19,21	1.15	2 (13%)
3	FUC	I	502	3	9,10,11	0.94	0	10,14,16	1.04	2 (20%)
3	NAG	I	503	3	12,14,15	0.64	0	15,19,21	0.87	1 (6%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	NAG	J	701	1,6	12,14,15	0.88	1 (8%)	15,19,21	1.64	3 (20%)
6	NAG	J	702	6	12,14,15	0.71	1 (8%)	15,19,21	0.98	0
6	MAN	J	703	6	10,11,12	0.69	0	11,15,17	2.10	3 (27%)
6	MAN	J	704	6	10,11,12	0.69	0	11,15,17	0.96	0
7	NAG	K	801	1,7	12,14,15	0.94	1 (8%)	15,19,21	1.57	2 (13%)
7	FUC	K	802	7	9,10,11	0.93	0	10,14,16	1.35	1 (10%)
7	NAG	K	803	7	12,14,15	0.66	0	15,19,21	0.77	0
7	MAN	K	804	7	10,11,12	0.79	0	11,15,17	1.36	2 (18%)
4	NAG	L	901	1,4	12,14,15	0.68	0	15,19,21	1.07	2 (13%)
4	NAG	L	902	4	12,14,15	0.60	0	15,19,21	0.95	1 (6%)
4	NAG	M	101	1,4	12,14,15	0.46	0	15,19,21	1.50	3 (20%)
4	NAG	M	102	4	12,14,15	0.64	0	15,19,21	1.04	1 (6%)
3	NAG	N	401	1,3	12,14,15	0.69	0	15,19,21	1.28	2 (13%)
3	FUC	N	402	3	9,10,11	0.82	0	10,14,16	1.30	2 (20%)
3	NAG	N	403	3	12,14,15	0.66	0	15,19,21	0.90	1 (6%)
3	NAG	O	501	1,3	12,14,15	0.58	0	15,19,21	1.69	1 (6%)
3	FUC	O	502	3	9,10,11	0.79	0	10,14,16	1.16	0
3	NAG	O	503	3	12,14,15	0.63	0	15,19,21	0.93	1 (6%)
4	NAG	P	601	1,4	12,14,15	0.62	0	15,19,21	1.60	1 (6%)
4	NAG	P	603	4	12,14,15	0.62	0	15,19,21	1.21	2 (13%)
4	NAG	Q	701	1,4	12,14,15	0.62	0	15,19,21	1.60	3 (20%)
4	NAG	Q	703	4	12,14,15	0.70	0	15,19,21	1.23	2 (13%)
8	NAG	R	801	1,8	12,14,15	1.01	1 (8%)	15,19,21	1.76	4 (26%)
8	FUC	R	802	8	9,10,11	0.78	0	10,14,16	0.59	0
8	NAG	R	803	8	12,14,15	0.82	1 (8%)	15,19,21	1.19	1 (6%)
8	MAN	R	804	8	10,11,12	0.99	1 (10%)	11,15,17	2.58	4 (36%)
8	MAN	R	805	8	10,11,12	0.53	0	11,15,17	2.19	3 (27%)
4	NAG	S	901	1,4	12,14,15	0.59	0	15,19,21	1.33	2 (13%)
4	NAG	S	902	4	12,14,15	0.54	0	15,19,21	1.34	2 (13%)

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	E	101	1,3	-	0/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	FUC	E	102	3	-	0/0/17/20	0/1/1/1
3	NAG	E	103	3	-	0/6/23/26	0/1/1/1
4	NAG	F	201	1,4	-	0/6/23/26	0/1/1/1
4	NAG	F	203	4	-	0/6/23/26	0/1/1/1
4	NAG	G	301	1,4	-	0/6/23/26	0/1/1/1
4	NAG	G	302	4	-	0/6/23/26	0/1/1/1
3	NAG	H	401	1,3	-	0/6/23/26	0/1/1/1
3	FUC	H	402	3	-	0/0/17/20	0/1/1/1
3	NAG	H	403	3	-	0/6/23/26	0/1/1/1
3	NAG	I	501	1,3	-	0/6/23/26	0/1/1/1
3	FUC	I	502	3	-	0/0/17/20	0/1/1/1
3	NAG	I	503	3	-	0/6/23/26	0/1/1/1
6	NAG	J	701	1,6	-	0/6/23/26	0/1/1/1
6	NAG	J	702	6	-	0/6/23/26	0/1/1/1
6	MAN	J	703	6	1/1/4/5	0/2/19/22	0/1/1/1
6	MAN	J	704	6	-	0/2/19/22	0/1/1/1
7	NAG	K	801	1,7	-	0/6/23/26	0/1/1/1
7	FUC	K	802	7	-	0/0/17/20	0/1/1/1
7	NAG	K	803	7	-	0/6/23/26	0/1/1/1
7	MAN	K	804	7	1/1/4/5	0/2/19/22	0/1/1/1
4	NAG	L	901	1,4	-	0/6/23/26	0/1/1/1
4	NAG	L	902	4	-	0/6/23/26	0/1/1/1
4	NAG	M	101	1,4	-	0/6/23/26	0/1/1/1
4	NAG	M	102	4	-	0/6/23/26	0/1/1/1
3	NAG	N	401	1,3	-	0/6/23/26	0/1/1/1
3	FUC	N	402	3	-	0/0/17/20	0/1/1/1
3	NAG	N	403	3	-	0/6/23/26	0/1/1/1
3	NAG	O	501	1,3	1/1/5/7	0/6/23/26	0/1/1/1
3	FUC	O	502	3	-	0/0/17/20	0/1/1/1
3	NAG	O	503	3	-	0/6/23/26	0/1/1/1
4	NAG	P	601	1,4	-	0/6/23/26	0/1/1/1
4	NAG	P	603	4	-	0/6/23/26	0/1/1/1
4	NAG	Q	701	1,4	-	0/6/23/26	0/1/1/1
4	NAG	Q	703	4	-	1/6/23/26	0/1/1/1
8	NAG	R	801	1,8	-	0/6/23/26	0/1/1/1
8	FUC	R	802	8	-	0/0/17/20	0/1/1/1
8	NAG	R	803	8	-	0/6/23/26	0/1/1/1
8	MAN	R	804	8	1/1/4/5	0/2/19/22	0/1/1/1
8	MAN	R	805	8	-	0/2/19/22	0/1/1/1
4	NAG	S	901	1,4	-	0/6/23/26	0/1/1/1
4	NAG	S	902	4	-	0/6/23/26	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	R	801	NAG	C2-N2	-2.63	1.43	1.46
8	R	803	NAG	O5-C5	-2.42	1.40	1.45
7	K	801	NAG	O5-C5	-2.29	1.41	1.45
6	J	701	NAG	O5-C5	-2.20	1.41	1.45
3	E	102	FUC	O5-C5	-2.18	1.41	1.45

The worst 5 of 69 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	R	804	MAN	O5-C5-C6	5.64	112.90	106.98
3	O	501	NAG	O5-C5-C6	5.58	112.84	106.98
4	F	203	NAG	O5-C5-C6	5.36	112.61	106.98
8	R	805	MAN	O5-C5-C4	5.31	117.39	110.65
4	P	601	NAG	O5-C5-C6	5.03	112.26	106.98

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
8	R	804	MAN	C1
6	J	703	MAN	C1
3	O	501	NAG	C1
7	K	804	MAN	C1

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	Q	703	NAG	O7-C7-N2-C2

There are no ring outliers.

## 5.6 Ligand geometry ⓘ

3 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
5	NAG	A	601	1	12,14,15	0.58	0	15,19,21	1.15	2 (13%)
5	NAG	C	501	1	12,14,15	0.59	0	15,19,21	0.90	1 (6%)
5	NAG	C	601	1	12,14,15	0.55	0	15,19,21	0.85	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
5	NAG	A	601	1	-	0/6/23/26	0/1/1/1
5	NAG	C	501	1	-	0/6/23/26	0/1/1/1
5	NAG	C	601	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	C	501	NAG	O5-C5-C6	2.44	109.54	106.98
5	A	601	NAG	O5-C5-C6	2.37	109.47	106.98
5	A	601	NAG	C3-C2-N2	-2.14	108.50	111.76

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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	480/488 (98%)	-0.30	1 (0%) 93 88	78, 100, 119, 144	0
1	C	480/488 (98%)	-0.31	0 100 100	80, 100, 119, 141	0
2	B	180/180 (100%)	-0.43	0 100 100	83, 100, 122, 138	0
2	D	180/180 (100%)	-0.41	0 100 100	85, 101, 122, 137	0
All	All	1320/1336 (98%)	-0.33	1 (0%) 91 90	78, 100, 120, 144	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	43	HIS	2.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates ⓘ

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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
4	NAG	F	203	14/15	0.56	15.93	87,89,90,90	14
4	NAG	P	601	14/15	0.23	7.60	101,102,105,108	14
4	NAG	L	902	14/15	0.29	3.28	87,87,87,87	14
4	NAG	G	302	14/15	0.31	3.00	100,101,103,103	14

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
7	NAG	K	801	14/15	0.23	2.84	56,60,62,64	14
4	NAG	S	902	14/15	0.33	2.49	96,96,97,97	14
4	NAG	M	101	14/15	0.28	2.40	103,107,108,108	14
3	FUC	E	102	10/11	0.27	2.35	68,69,69,70	10
7	NAG	K	803	14/15	0.20	1.99	59,60,62,63	14
3	FUC	I	502	10/11	0.41	1.94	104,105,106,106	10
3	NAG	N	401	14/15	0.18	1.84	96,98,102,102	14
4	NAG	Q	701	14/15	0.24	1.79	102,102,104,105	14
8	NAG	R	803	14/15	0.22	1.52	73,78,80,81	14
3	NAG	H	401	14/15	0.19	1.18	96,98,100,102	14
3	NAG	I	501	14/15	0.18	1.01	107,112,112,113	14
4	NAG	L	901	14/15	0.22	0.75	85,88,88,89	14
3	NAG	O	501	14/15	0.21	0.63	123,126,130,133	14
6	NAG	J	701	14/15	0.17	0.49	81,88,90,90	14
4	NAG	F	201	14/15	0.30	0.47	90,91,93,95	14
4	NAG	S	901	14/15	0.19	0.43	92,93,94,95	14
8	NAG	R	801	14/15	0.16	0.14	80,81,83,83	14
4	NAG	G	301	14/15	0.17	-0.60	92,94,96,98	14
3	NAG	E	101	14/15	0.17	-0.86	71,74,76,77	14
7	FUC	K	802	10/11	0.28	-	57,59,59,59	10
6	MAN	J	703	11/12	0.29	-	100,101,102,104	11
3	NAG	H	403	14/15	0.54	-	105,107,108,109	14
8	MAN	R	805	11/12	0.24	-	67,68,69,69	11
7	MAN	K	804	11/12	0.22	-	63,64,64,64	11
3	NAG	O	503	14/15	0.62	-	128,129,129,130	14
8	FUC	R	802	10/11	0.24	-	83,84,84,84	10
6	MAN	J	704	11/12	0.33	-	105,105,106,106	11
3	NAG	I	503	14/15	0.48	-	114,114,115,115	14
8	MAN	R	804	11/12	0.21	-	68,69,70,71	11
3	NAG	N	403	14/15	0.39	-	103,105,105,105	14
6	NAG	J	702	14/15	0.30	-	92,93,95,98	14
4	NAG	P	603	14/15	0.62	-	110,112,113,113	14
3	FUC	N	402	10/11	0.23	-	102,103,103,103	10
3	FUC	O	502	10/11	0.47	-	135,136,137,137	10
4	NAG	M	102	14/15	0.42	-	109,109,109,109	14
3	NAG	E	103	14/15	0.33	-	75,76,76,76	14
3	FUC	H	402	10/11	0.33	-	100,101,101,101	10
4	NAG	Q	703	14/15	0.42	-	105,107,108,109	14

## 6.4 Ligands

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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
5	NAG	C	501	14/15	0.34	3.56	89,92,93,94	14
5	NAG	A	601	14/15	0.23	3.03	77,86,87,88	14
5	NAG	C	601	14/15	0.16	-	94,97,99,99	14

## 6.5 Other polymers

There are no such residues in this entry.