



# Full wwPDB X-ray Structure Validation Report

Oct 22, 2014 – 10:47 AM EDT

PDB ID : 4NYT  
Title : L-Ficolin Complexed to Phosphocholine  
Authors : Laffly, E.; Gaboriaud, C.; Martin, L.; Thielens, N.  
Deposited on : 2013-12-11  
Resolution : 2.25 Å(reported)

This is a full wwPDB 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/ValidationPDFNotes.html>

---

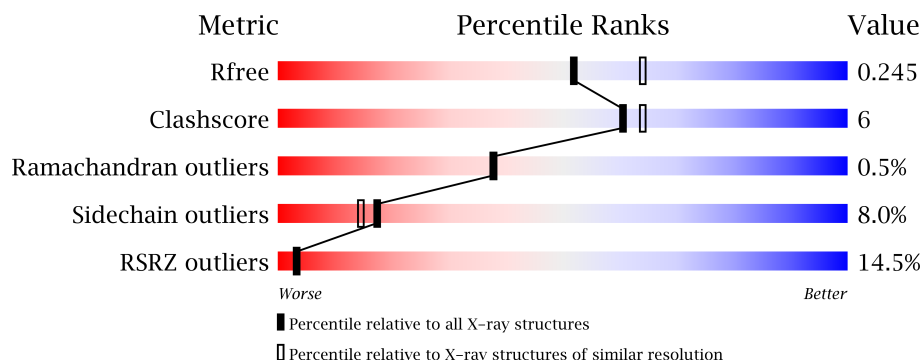
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.16 November 2013  
Xtriage (Phenix) : dev-1439  
EDS : stable24103  
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) : stable24103

# 1 Overall quality at a glance

The reported resolution of this entry is 2.25 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	1108 (2.28-2.24)
Clashscore	79885	1326 (2.28-2.24)
Ramachandran outliers	78287	1291 (2.28-2.24)
Sidechain outliers	78261	1291 (2.28-2.24)
RSRZ outliers	66119	1110 (2.28-2.24)

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	217	
1	B	217	
1	C	217	

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

Mol	Type	Chain	Res	Geometry	Electron density
4	CA	A	305	-	X
5	PO4	C	307	-	X
7	PC	A	303	-	X

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 5347 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.

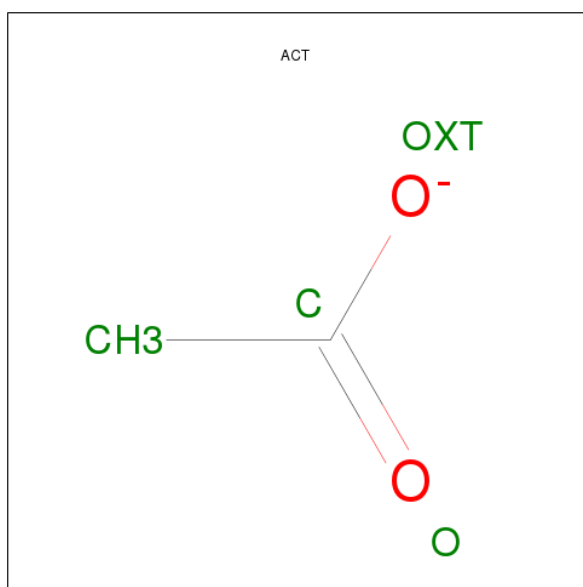
- Molecule 1 is a protein called Ficolin-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	216	Total	C	N	O	S	0	3	0
			1757	1106	310	331	10			
1	A	212	Total	C	N	O	S	0	3	0
			1727	1085	306	328	8			
1	B	212	Total	C	N	O	S	0	0	0
			1702	1071	300	323	8			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	168	THR	VAL	SEE REMARK 999	UNP Q15485
C	247	THR	VAL	SEE REMARK 999	UNP Q15485
A	168	THR	VAL	SEE REMARK 999	UNP Q15485
A	247	THR	VAL	SEE REMARK 999	UNP Q15485
B	168	THR	VAL	SEE REMARK 999	UNP Q15485
B	247	THR	VAL	SEE REMARK 999	UNP Q15485

- Molecule 2 is ACETATE ION (three-letter code: ACT) (formula: C<sub>2</sub>H<sub>3</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	C	1	Total	C	O	0	0
			4	2	2		
2	A	1	Total	C	O	0	0
			4	2	2		

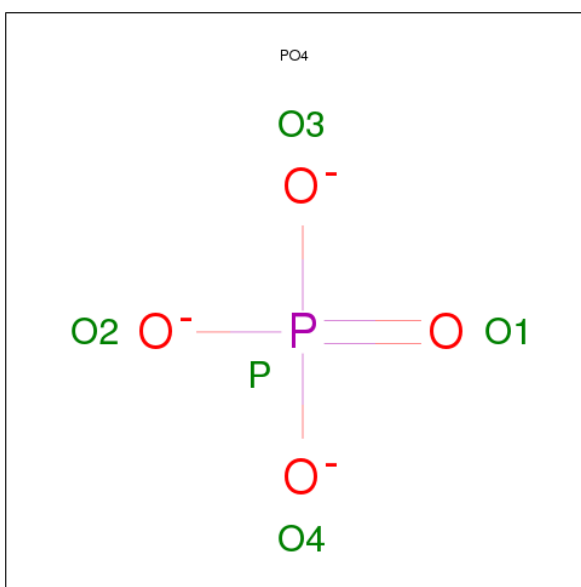
- Molecule 3 is a polymer of unknown type called SUGAR (4-MER).

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

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

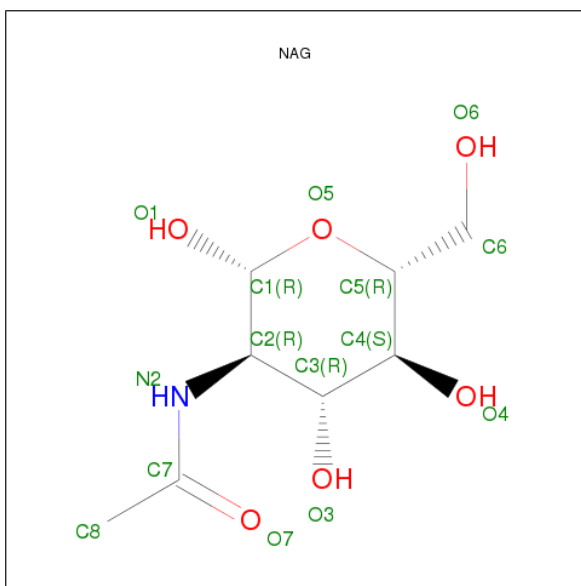
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Ca	0	0
			1	1		
4	A	2	Total	Ca	0	0
			2	2		
4	C	1	Total	Ca	0	0
			1	1		

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



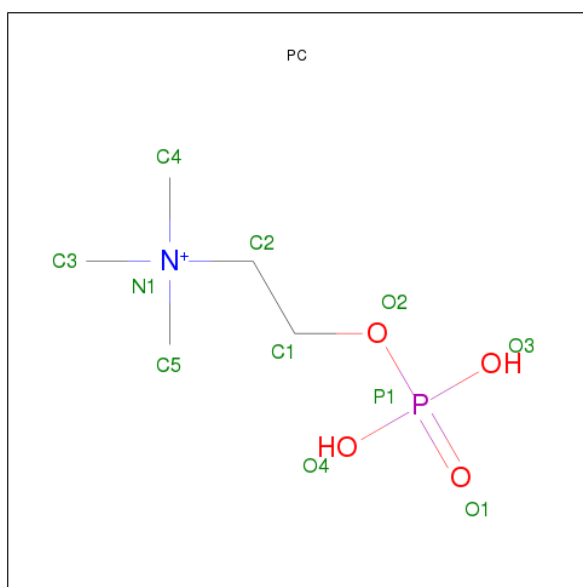
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	C	1	Total	O	P	0	0
			5	4	1		

- Molecule 6 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

- Molecule 7 is PHOSPHOCHOLINE (three-letter code: PC) (formula:  $C_5H_{15}NO_4P$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	A	1	Total	C	N	O	P	0	0
			11	5	1	4	1		

- Molecule 8 is water.

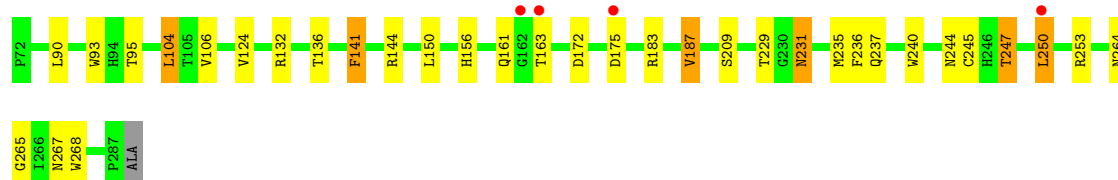
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	C	33	Total	O	0	0
			33	33		
8	A	34	Total	O	0	0
			34	34		
8	B	3	Total	O	0	0
			3	3		

### 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: Ficolin-2

Chain C: 



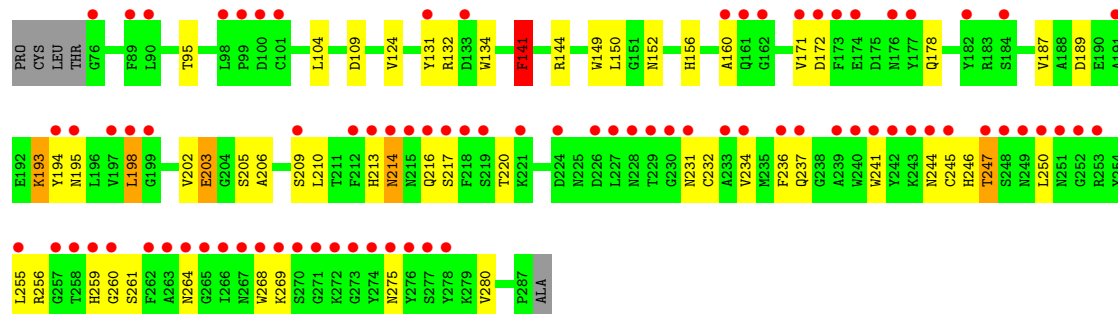
#### • Molecule 1: Ficolin-2

Chain A: 



#### • Molecule 1: Ficolin-2

Chain B: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	96.13Å 96.13Å 139.88Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.95 – 2.25 19.95 – 2.25	Depositor EDS
% Data completeness (in resolution range)	99.5 (19.95-2.25) 99.6 (19.95-2.25)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.62 (at 2.26Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, $R_{free}$	0.213 , 0.243 0.215 , 0.245	Depositor DCC
$R_{free}$ test set	1795 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	32.3	Xtriage
Anisotropy	0.424	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 28.1	EDS
Estimated twinning fraction	0.033 for -h,-k,l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 35895 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	5347	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.84% 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: NAG, PO4, PC, FUC, ACT, CA, 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	0.95	3/1774 (0.2%)	0.97	5/2400 (0.2%)
1	B	0.76	7/1749 (0.4%)	0.85	6/2366 (0.3%)
1	C	1.08	4/1805 (0.2%)	1.04	5/2441 (0.2%)
All	All	0.94	14/5328 (0.3%)	0.96	16/7207 (0.2%)

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
3	C	1	0

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	141	PHE	CG-CD2	-14.68	1.16	1.38
1	C	141	PHE	CG-CD1	-12.70	1.19	1.38
1	B	141	PHE	CG-CD2	-9.61	1.24	1.38
1	C	141	PHE	CE1-CZ	-9.48	1.19	1.37
1	A	141	PHE	CG-CD2	-9.02	1.25	1.38
1	C	141	PHE	CE2-CZ	-8.70	1.20	1.37
1	B	141	PHE	CG-CD1	-8.48	1.26	1.38
1	B	203	GLU	CD-OE1	7.54	1.33	1.25
1	B	203	GLU	CD-OE2	7.24	1.33	1.25
1	A	141	PHE	CG-CD1	-6.93	1.28	1.38
1	B	141	PHE	CE1-CZ	-6.46	1.25	1.37
1	A	141	PHE	CE1-CZ	-6.21	1.25	1.37
1	B	141	PHE	CE2-CZ	-5.33	1.27	1.37
1	B	206	ALA	C-O	5.10	1.33	1.23

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	132	ARG	NE-CZ-NH2	-12.29	114.15	120.30
1	C	132	ARG	NE-CZ-NH2	-12.05	114.28	120.30
1	A	132	ARG	NE-CZ-NH1	8.78	124.69	120.30
1	B	132	ARG	NE-CZ-NH2	-8.52	116.04	120.30
1	C	132	ARG	CG-CD-NE	-7.82	95.37	111.80
1	B	132	ARG	CG-CD-NE	-6.68	97.77	111.80
1	B	109	ASP	CB-CG-OD1	6.50	124.15	118.30
1	A	132	ARG	CG-CD-NE	-6.30	98.57	111.80
1	C	132	ARG	NE-CZ-NH1	5.94	123.27	120.30
1	B	132	ARG	NE-CZ-NH1	5.85	123.23	120.30
1	C	172	ASP	CB-CG-OD2	-5.61	113.26	118.30
1	A	109	ASP	CB-CG-OD1	5.49	123.24	118.30
1	C	141	PHE	CE1-CZ-CE2	-5.09	110.84	120.00
1	B	172	ASP	CB-CG-OD2	-5.04	113.76	118.30
1	A	113	ASP	CB-CG-OD1	5.04	122.83	118.30
1	B	144	ARG	NE-CZ-NH2	-5.01	117.80	120.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	C	304	MAN	C1

There are no planarity outliers.

## 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	1727	0	1596	15	0
1	B	1702	0	1573	28	0
1	C	1757	0	1631	17	0
2	A	4	0	3	0	0
2	C	4	0	3	0	0
3	C	49	0	43	0	0
4	A	2	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	1	0	0	0	0
4	C	1	0	0	0	0
5	C	5	0	0	0	0
6	A	14	0	13	0	0
7	A	11	0	13	0	0
8	A	34	0	0	0	0
8	B	3	0	0	0	0
8	C	33	0	0	0	0
All	All	5347	0	4875	60	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 6.

All (60) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:237:GLN:HE22	1:A:264:ASN:HD22	1.37	0.72
1:B:259:HIS:O	1:B:261:SER:N	2.23	0.72
1:A:237:GLN:NE2	1:A:264:ASN:HD22	1.88	0.72
1:C:237:GLN:NE2	1:C:264:ASN:HD22	1.94	0.65
1:B:141:PHE:CE1	1:B:149:TRP:HB3	2.35	0.62
1:C:235[B]:MET:HE2	1:C:236:PHE:CE2	2.34	0.62
1:B:134:TRP:HA	1:B:220:THR:HG21	1.83	0.61
1:C:237:GLN:HE21	1:C:264:ASN:HD22	1.49	0.59
1:C:265:GLY:H	1:C:267:ASN:HD21	1.49	0.58
1:B:198:LEU:N	1:B:214:ASN:HD21	2.01	0.58
1:B:156:HIS:CD2	1:B:187:VAL:O	2.57	0.57
1:B:198:LEU:H	1:B:214:ASN:HD21	1.50	0.57
1:A:247[A]:THR:HG23	1:A:269:LYS:HB3	1.88	0.56
1:A:265:GLY:H	1:A:267:ASN:HD21	1.54	0.55
1:B:244:ASN:N	1:B:245:CYS:HA	2.21	0.55
1:C:156:HIS:HD2	1:C:187:VAL:O	1.90	0.55
1:A:78:ARG:HB3	1:A:78:ARG:CZ	2.37	0.55
1:B:209:SER:HB2	1:B:268:TRP:CE2	2.42	0.54
1:B:171:VAL:HB	1:B:280:VAL:HB	1.89	0.54
1:B:156:HIS:HD2	1:B:187:VAL:O	1.90	0.53
1:B:259:HIS:O	1:B:259:HIS:ND1	2.42	0.52
1:A:209:SER:HB3	1:A:247[A]:THR:HG22	1.93	0.50
1:A:100:ASP:O	1:A:101:CYS:HB2	2.10	0.50
1:C:265:GLY:H	1:C:267:ASN:ND2	2.11	0.49
1:B:259:HIS:CE1	1:B:275:ASN:HA	2.49	0.48
1:C:209:SER:HB2	1:C:268:TRP:CE2	2.49	0.48
1:A:209:SER:HB2	1:A:268:TRP:CE2	2.49	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:131:TYR:CE1	1:B:237:GLN:HA	2.49	0.47
1:B:187:VAL:HA	1:B:195:ASN:O	2.15	0.47
1:B:217:SER:O	1:B:241:TRP:HA	2.15	0.46
1:C:104[A]:LEU:HD13	1:C:106:VAL:HG12	1.98	0.46
1:C:136:THR:HG22	1:C:141:PHE:HD2	1.81	0.46
1:C:104[A]:LEU:HD13	1:C:106:VAL:CG1	2.46	0.46
1:B:195:ASN:HA	1:B:216:GLN:O	2.15	0.46
1:C:253[B]:ARG:HH11	1:C:253[B]:ARG:HG2	1.82	0.45
1:B:189:ASP:O	1:B:193:LYS:O	2.35	0.44
1:A:237:GLN:HE21	1:A:253:ARG:HD3	1.83	0.44
1:A:231:ASN:C	1:A:231:ASN:HD22	2.19	0.44
1:A:265:GLY:H	1:A:267:ASN:ND2	2.14	0.44
1:B:198:LEU:HD13	1:B:210:LEU:HD23	2.00	0.43
1:B:231:ASN:OD1	1:B:234:VAL:HG23	2.17	0.43
1:B:210:LEU:O	1:B:213:HIS:N	2.48	0.43
1:C:209:SER:HB3	1:C:247:THR:HG23	2.00	0.43
1:B:237:GLN:NE2	1:B:264:ASN:HD22	2.15	0.43
1:B:247:THR:O	1:B:269:LYS:HB2	2.19	0.42
1:B:255:LEU:O	1:B:256:ARG:HB2	2.19	0.42
1:C:240:TRP:CH2	1:C:250:LEU:HB2	2.55	0.42
1:C:93:TRP:CZ2	1:C:144:ARG:HA	2.54	0.42
1:C:244:ASN:N	1:C:245:CYS:HA	2.33	0.42
1:A:198:LEU:H	1:A:214:ASN:ND2	2.17	0.42
1:A:237:GLN:HE21	1:A:253:ARG:HH11	1.67	0.42
1:B:178:GLN:HA	1:B:205:SER:OG	2.20	0.42
1:C:136:THR:HG22	1:C:141:PHE:CD2	2.54	0.42
1:A:228:ASN:ND2	1:A:244:ASN:OD1	2.51	0.42
1:B:194:TYR:O	1:B:195:ASN:C	2.58	0.42
1:C:231:ASN:C	1:C:231:ASN:HD22	2.23	0.42
1:B:232:CYS:O	1:B:236:PHE:N	2.54	0.41
1:B:213:HIS:ND1	1:B:241:TRP:O	2.50	0.41
1:A:93:TRP:CZ2	1:A:144:ARG:HA	2.56	0.40
1:B:152:ASN:HB3	1:B:194:TYR:CD1	2.56	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	213/217 (98%)	199 (93%)	13 (6%)	1 (0%)	38	38
1	B	210/217 (97%)	188 (90%)	20 (10%)	2 (1%)	22	17
1	C	217/217 (100%)	207 (95%)	10 (5%)	0	100	100
All	All	640/651 (98%)	594 (93%)	43 (7%)	3 (0%)	38	38

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	260	GLY
1	B	160	ALA
1	A	125	ASP

### 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	181/182 (100%)	164 (91%)	17 (9%)	13	9
1	B	178/182 (98%)	165 (93%)	13 (7%)	20	18
1	C	185/182 (102%)	170 (92%)	15 (8%)	17	14
All	All	544/546 (100%)	499 (92%)	45 (8%)	17	13

All (45) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	90	LEU
1	C	95	THR
1	C	104[A]	LEU
1	C	104[B]	LEU
1	C	124	VAL
1	C	150	LEU
1	C	161	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	163	THR
1	C	175	ASP
1	C	183	ARG
1	C	187	VAL
1	C	229	THR
1	C	231	ASN
1	C	247	THR
1	C	250	LEU
1	A	78	ARG
1	A	90	LEU
1	A	95	THR
1	A	104	LEU
1	A	122	ARG
1	A	124	VAL
1	A	150	LEU
1	A	163	THR
1	A	183	ARG
1	A	187	VAL
1	A	202	VAL
1	A	229	THR
1	A	231	ASN
1	A	237	GLN
1	A	250	LEU
1	A	258[A]	THR
1	A	258[B]	THR
1	B	95	THR
1	B	104	LEU
1	B	124	VAL
1	B	141	PHE
1	B	150	LEU
1	B	193	LYS
1	B	198	LEU
1	B	202	VAL
1	B	203	GLU
1	B	214	ASN
1	B	246	HIS
1	B	247	THR
1	B	250	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (21) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	139	GLN
1	C	156	HIS
1	C	161	GLN
1	C	214	ASN
1	C	216	GLN
1	C	231	ASN
1	C	237	GLN
1	C	267	ASN
1	A	139	GLN
1	A	156	HIS
1	A	195	ASN
1	A	214	ASN
1	A	231	ASN
1	A	237	GLN
1	A	267	ASN
1	B	139	GLN
1	B	156	HIS
1	B	214	ASN
1	B	225	ASN
1	B	237	GLN
1	B	267	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 ⓘ

4 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	C	302	1,3	12,14,15	1.06	1 (8%)	15,19,21	2.38	5 (33%)
3	NAG	C	303	3	12,14,15	0.91	1 (8%)	15,19,21	2.28	7 (46%)
3	MAN	C	304	3	10,11,12	0.66	0	11,15,17	1.76	3 (27%)
3	FUC	C	305	3	9,10,11	1.13	0	10,14,16	2.32	2 (20%)

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	C	302	1,3	-	0/6/23/26	0/1/1/1
3	NAG	C	303	3	-	0/6/23/26	0/1/1/1
3	MAN	C	304	3	1/1/4/5	0/2/19/22	0/1/1/1
3	FUC	C	305	3	-	0/0/17/20	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	302	NAG	O5-C5	-2.35	1.41	1.45
3	C	303	NAG	C2-N2	-2.20	1.44	1.46

All (17) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	302	NAG	O5-C5-C6	-4.98	101.75	106.98
3	C	305	FUC	C4-C3-C2	-4.75	105.31	110.61
3	C	303	NAG	O5-C5-C6	4.54	111.75	106.98
3	C	302	NAG	C3-C2-N2	-4.41	105.01	111.62
3	C	305	FUC	O3-C3-C2	3.99	116.98	109.74
3	C	304	MAN	O5-C5-C6	3.59	110.75	106.98
3	C	303	NAG	C3-C2-N2	-3.41	106.52	111.62
3	C	302	NAG	C3-C4-C5	-3.24	104.35	110.17
3	C	303	NAG	O5-C5-C4	2.95	114.40	110.65
3	C	304	MAN	O2-C2-C3	-2.89	103.98	110.10
3	C	303	NAG	O4-C4-C5	-2.87	101.67	109.25
3	C	302	NAG	C6-C5-C4	-2.86	106.05	113.04
3	C	303	NAG	C2-N2-C7	2.82	126.81	123.39
3	C	303	NAG	C6-C5-C4	-2.45	107.06	113.04
3	C	303	NAG	C4-C3-C2	-2.25	106.00	110.74
3	C	302	NAG	O5-C5-C4	-2.25	107.80	110.65

*Continued on next page...*



Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	304	MAN	O3-C3-C2	-2.22	105.70	109.74

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	C	304	MAN	C1

There are no torsion outliers.

There are no ring outliers.

## 5.6 Ligand geometry

Of 9 ligands modelled in this entry, 4 are monoatomic - leaving 5 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	ACT	A	301	-	1,3,3	1.92	0	0,3,3	0.00	-
6	NAG	A	302	1	12,14,15	1.08	1 (8%)	15,19,21	2.36	7 (46%)
7	PC	A	303	-	10,10,10	0.82	0	15,15,15	3.54	10 (66%)
2	ACT	C	301	-	1,3,3	1.10	0	0,3,3	0.00	-
5	PO4	C	307	-	4,4,4	0.88	0	6,6,6	0.34	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
2	ACT	A	301	-	-	0/0/0/0	0/0/0/0
6	NAG	A	302	1	-	0/6/23/26	0/1/1/1
7	PC	A	303	-	-	0/8/8/8	0/0/0/0
2	ACT	C	301	-	-	0/0/0/0	0/0/0/0
5	PO4	C	307	-	-	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)
6	A	302	NAG	C4-C5	2.21	1.57	1.53

All (17) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	A	303	PC	C5-N1-C3	-6.15	92.67	108.99
7	A	303	PC	C5-N1-C4	-6.05	92.93	108.99
6	A	302	NAG	C3-C2-N2	-5.98	102.65	111.62
7	A	303	PC	C5-N1-C2	-5.60	87.63	109.92
7	A	303	PC	O4-P1-O1	4.82	126.11	110.36
7	A	303	PC	C4-N1-C3	3.48	118.22	108.99
7	A	303	PC	C3-N1-C2	3.46	123.69	109.92
7	A	303	PC	P1-O2-C1	-3.24	110.35	118.63
6	A	302	NAG	O4-C4-C3	-2.90	103.87	110.36
7	A	303	PC	C1-C2-N1	-2.69	106.49	115.85
6	A	302	NAG	C4-C3-C2	2.66	116.32	110.74
6	A	302	NAG	O4-C4-C5	2.60	116.14	109.25
7	A	303	PC	O2-P1-O1	-2.59	98.80	106.72
6	A	302	NAG	C2-N2-C7	2.50	126.42	123.39
6	A	302	NAG	C6-C5-C4	2.15	118.29	113.04
6	A	302	NAG	O5-C5-C4	-2.11	107.97	110.65
7	A	303	PC	C4-N1-C2	2.03	117.99	109.92

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	212/217 (97%)	-0.18	6 (2%) 50 57	17, 34, 64, 89	5 (2%)
1	B	212/217 (97%)	1.87	83 (39%) 1 0	37, 109, 205, 253	4 (1%)
1	C	216/217 (99%)	-0.28	4 (1%) 64 70	16, 30, 56, 73	4 (1%)
All	All	640/651 (98%)	0.47	93 (14%) 3 3	16, 42, 179, 253	13 (2%)

All (93) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	229	THR	14.1
1	B	268	TRP	8.4
1	B	250	LEU	6.6
1	B	242	TYR	6.2
1	B	260	GLY	6.1
1	B	274	TYR	6.0
1	B	227	LEU	5.6
1	B	191	ALA	5.6
1	B	101	CYS	5.5
1	A	161	GLN	5.4
1	B	161	GLN	5.3
1	B	197	VAL	5.3
1	B	218	PHE	5.2
1	B	219	SER	4.9
1	B	174	GLU	4.9
1	B	248	SER	4.8
1	B	241	TRP	4.6
1	B	263	ALA	4.6
1	B	176	ASN	4.6
1	B	231	ASN	4.6
1	B	273	GLY	4.5
1	B	257	GLY	4.4
1	B	240	TRP	4.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	98	LEU	4.3
1	B	171	VAL	4.3
1	B	243	LYS	4.3
1	B	239	ALA	4.3
1	B	100	ASP	4.2
1	B	262	PHE	4.2
1	B	249	ASN	4.1
1	B	216	GLN	4.1
1	B	199	GLY	4.1
1	B	172	ASP	4.0
1	B	275	ASN	4.0
1	A	162	GLY	3.9
1	B	224	ASP	3.9
1	B	271	GLY	3.8
1	B	276	TYR	3.8
1	B	277	SER	3.7
1	B	173	PHE	3.7
1	B	236	PHE	3.6
1	B	264	ASN	3.6
1	B	258	THR	3.5
1	B	265	GLY	3.5
1	B	230	GLY	3.5
1	B	245	CYS	3.4
1	B	76	GLY	3.4
1	B	160	ALA	3.4
1	B	252	GLY	3.3
1	B	177	TYR	3.2
1	B	272	LYS	3.1
1	B	215	ASN	3.1
1	B	217	SER	3.1
1	B	89	PHE	3.0
1	B	267	ASN	3.0
1	B	212	PHE	3.0
1	B	182	TYR	3.0
1	B	90	LEU	2.9
1	B	226	ASP	2.9
1	A	229	THR	2.7
1	B	269	LYS	2.6
1	B	162	GLY	2.6
1	B	209	SER	2.6
1	A	212	PHE	2.5
1	B	184	SER	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	266	ILE	2.5
1	B	233	ALA	2.4
1	B	198	LEU	2.4
1	B	234	VAL	2.4
1	A	177	TYR	2.3
1	B	213	HIS	2.3
1	B	133	ASP	2.3
1	B	251	ASN	2.3
1	B	259	HIS	2.3
1	B	214	ASN	2.3
1	B	270	SER	2.2
1	B	99	PRO	2.2
1	B	221	LYS	2.2
1	C	162	GLY	2.2
1	B	194	TYR	2.2
1	B	244	ASN	2.2
1	B	237	GLN	2.2
1	B	253	ARG	2.2
1	C	175	ASP	2.2
1	C	163	THR	2.2
1	B	278	TYR	2.2
1	C	250	LEU	2.1
1	B	131	TYR	2.1
1	B	247	THR	2.1
1	B	255	LEU	2.0
1	B	195	ASN	2.0
1	B	228	ASN	2.0
1	A	78	ARG	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( $\text{\AA}^2$ )	Q<0.9
3	FUC	C	305	10/11	0.29	14.03	37,40,42,47	10
3	MAN	C	304	11/12	0.26	7.10	46,54,61,64	0
3	NAG	C	302	14/15	0.17	3.36	19,26,33,34	0
3	NAG	C	303	14/15	0.20	0.85	28,35,39,40	0

## 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( $\text{\AA}^2$ )	Q<0.9
5	PO4	C	307	5/5	0.24	14.58	30,37,45,46	0
7	PC	A	303	11/11	0.21	5.94	32,43,58,58	0
4	CA	A	305	1/1	0.18	2.04	46,46,46,46	0
2	ACT	A	301	4/4	0.13	0.29	37,38,38,39	0
2	ACT	C	301	4/4	0.12	0.01	29,31,32,33	0
4	CA	B	301	1/1	0.19	-1.26	67,67,67,67	0
4	CA	A	304	1/1	0.04	-1.71	38,38,38,38	0
4	CA	C	306	1/1	0.02	-1.85	20,20,20,20	0
6	NAG	A	302	14/15	0.42	-	39,49,57,59	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.