



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 25, 2017 – 02:12 PM EST

PDB ID : 5MJ4  
Title : INTERLEUKIN-23 COMPLEX WITH AN ANTAGONISTIC ALPHABODY,  
CRYSTAL FORM 2  
Authors : Desmet, J.; Verstraete, K.; Bloch, Y.; Lorent, E.; Wen, Y.; Devreese, B.; Van-  
denbroucke, K.; Loverix, S.; Hettmann, T.; Deroo, S.; Somers, K.; Henderikx,  
P.; Lasters, I.; Savvides, S.  
Deposited on : 2016-11-29  
Resolution : 3.40 Å(reported)

This is a Full wwPDB X-ray Structure 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/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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20028442  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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-20028442

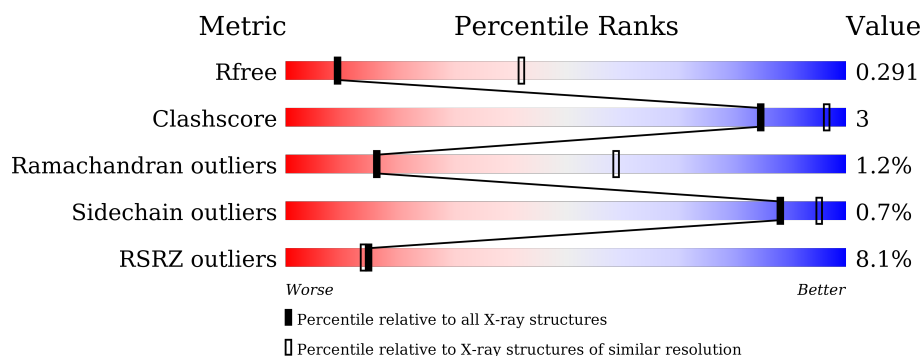
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.40 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))
$R_{free}$	91344	1476 (3.50-3.30)
Clashscore	102246	1611 (3.50-3.30)
Ramachandran outliers	100387	1571 (3.50-3.30)
Sidechain outliers	100360	1571 (3.50-3.30)
RSRZ outliers	91569	1485 (3.50-3.30)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	306	<div> <div>12%</div> <div>91%</div> <div>6%</div> </div>
2	B	179	<div> <div>2%</div> <div>80%</div> <div>9%</div> <div>9%</div> </div>
3	C	118	<div> <div>2%</div> <div>64%</div> <div>31%</div> </div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 8007 atoms, of which 3892 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 Interleukin-12 subunit beta.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	289	Total	C	H	N	O	S	0	0	0
			4233	1404	2023	351	443	12			

- Molecule 2 is a protein called Interleukin-23 subunit alpha.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	162	Total	C	H	N	O	S	0	0	0
			2383	776	1160	221	220	6			

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	190	GLY	-	expression tag	UNP Q9NPF7
B	191	THR	-	expression tag	UNP Q9NPF7
B	192	LYS	-	expression tag	UNP Q9NPF7
B	193	HIS	-	expression tag	UNP Q9NPF7
B	194	HIS	-	expression tag	UNP Q9NPF7
B	195	HIS	-	expression tag	UNP Q9NPF7
B	196	HIS	-	expression tag	UNP Q9NPF7
B	197	HIS	-	expression tag	UNP Q9NPF7
B	198	HIS	-	expression tag	UNP Q9NPF7

- Molecule 3 is a protein called ALPHABODY MA12.

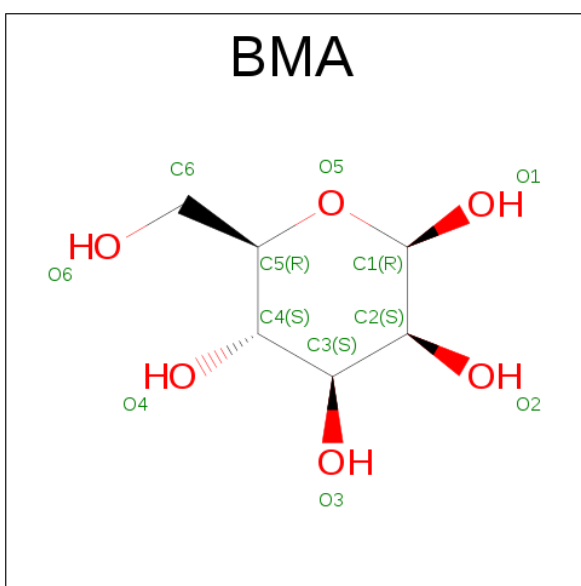
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	C	81	Total	C	H	N	O	S	0	0	0
			1275	402	654	100	113	6			

- Molecule 4 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



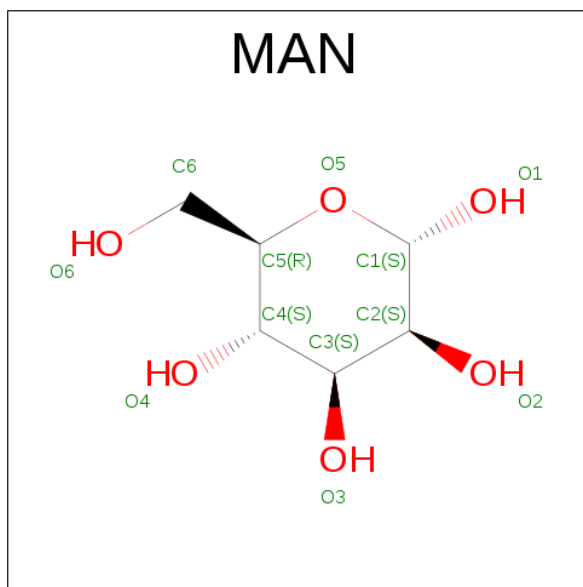
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	H	N	O	0	0
			27	8	13	1	5		
4	A	1	Total	C	H	N	O	0	0
			27	8	13	1	5		

- Molecule 5 is BETA-D-MANNOSE (three-letter code: BMA) (formula:  $C_6H_{12}O_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	H	O	0	0
			20	6	9	5		

- Molecule 6 is ALPHA-D-MANNOSE (three-letter code: MAN) (formula:  $C_6H_{12}O_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	H	O	0	0
			21	6	10	5		
6	A	1	Total	C	H	O	0	0
			21	6	10	5		



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	57.85Å 57.85Å 366.48Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	57.00 – 3.40 57.14 – 3.39	Depositor EDS
% Data completeness (in resolution range)	99.8 (57.00-3.40) 99.9 (57.14-3.39)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.98 (at 3.40Å)	Xtriage
Refinement program	PHENIX (1.11.1 _2575: ???)	Depositor
R, $R_{free}$	0.255 , 0.296 0.247 , 0.291	Depositor DCC
$R_{free}$ test set	933 reflections (9.96%)	DCC
Wilson B-factor (Å <sup>2</sup> )	109.9	Xtriage
Anisotropy	0.453	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 76.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	8007	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	122.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.00% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: BMA, 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	0.27	0/2266	0.51	0/3097
2	B	0.28	0/1256	0.52	0/1714
3	C	0.27	0/622	0.46	0/832
All	All	0.28	0/4144	0.51	0/5643

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

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	2210	2023	2023	9	0
2	B	1223	1160	1166	9	0
3	C	621	654	654	3	0
4	A	28	26	24	3	0
5	A	11	9	8	0	0
6	A	22	20	20	0	1
All	All	4115	3892	3895	23	1

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 3.



All (23) 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)
3:C:61:GLN:NE2	3:C:98:GLN:OE1	2.22	0.72
1:A:303:ASN:O	1:A:304:ALA:HB2	1.94	0.67
2:B:135:LEU:O	2:B:137:PRO:HD3	2.02	0.58
4:A:500:NAG:O7	4:A:500:NAG:C3	2.52	0.57
2:B:53:HIS:O	2:B:57:ARG:NH2	2.37	0.56
3:C:91:SER:N	3:C:94:GLU:OE1	2.39	0.56
4:A:500:NAG:O7	4:A:500:NAG:O3	2.25	0.55
1:A:303:ASN:O	1:A:304:ALA:CB	2.56	0.54
1:A:327:CYS:O	1:A:328:SER:CB	2.56	0.53
4:A:500:NAG:O3	4:A:501:NAG:O5	2.25	0.52
2:B:151:SER:HB2	2:B:152:PRO:HD2	1.92	0.51
1:A:304:ALA:O	1:A:327:CYS:SG	2.71	0.48
1:A:304:ALA:HB1	1:A:327:CYS:H	1.80	0.46
2:B:136:GLN:O	2:B:137:PRO:O	2.34	0.45
1:A:304:ALA:HA	1:A:327:CYS:HB2	1.97	0.45
1:A:38:TYR:HB2	1:A:39:PRO:HD2	1.98	0.44
3:C:117:VAL:O	3:C:117:VAL:HG12	2.18	0.43
2:B:91:GLN:HG2	2:B:142:TRP:CE2	2.55	0.41
2:B:143:GLU:OE1	2:B:144:THR:N	2.53	0.41
1:A:81:GLU:HG3	1:A:82:PHE:N	2.35	0.41
2:B:68:VAL:HG13	2:B:99:PHE:CE2	2.56	0.41
2:B:33:CYS:SG	2:B:134:LEU:HD23	2.61	0.40
1:A:314:TYR:CD2	2:B:178:ARG:HG2	2.56	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:A:504:MAN:O3	6:A:504:MAN:O3[7_645]	2.18	0.02

## 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	281/306 (92%)	273 (97%)	5 (2%)	3 (1%)	17	61
2	B	160/179 (89%)	153 (96%)	4 (2%)	3 (2%)	10	49
3	C	75/118 (64%)	75 (100%)	0	0	100	100
All	All	516/603 (86%)	501 (97%)	9 (2%)	6 (1%)	16	59

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	137	PRO
1	A	302	LYS
1	A	304	ALA
1	A	305	SER
2	B	28	PRO
2	B	136	GLN

### 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	239/277 (86%)	239 (100%)	0	100	100
2	B	127/152 (84%)	124 (98%)	3 (2%)	57	85
3	C	62/81 (76%)	62 (100%)	0	100	100
All	All	428/510 (84%)	425 (99%)	3 (1%)	88	95

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	67	ASP
2	B	140	HIS
2	B	143	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 ⓘ

5 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	$\# Z  > 2$
4	NAG	A	500	1,4	14,14,15	0.18	0	15,19,21	0.80	1 (6%)
4	NAG	A	501	5,4	14,14,15	0.17	0	15,19,21	0.27	0
5	BMA	A	502	4,6	11,11,12	0.92	0	15,15,17	1.06	1 (6%)
6	MAN	A	503	5	11,11,12	0.84	0	15,15,17	0.92	2 (13%)
6	MAN	A	504	5	11,11,12	1.42	2 (18%)	15,15,17	1.99	4 (26%)

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
4	NAG	A	500	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	501	5,4	-	0/6/23/26	0/1/1/1
5	BMA	A	502	4,6	-	0/2/19/22	0/1/1/1
6	MAN	A	503	5	-	0/2/19/22	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	MAN	A	504	5	-	0/2/19/22	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	504	MAN	O5-C1	2.67	1.48	1.43
6	A	504	MAN	C1-C2	3.48	1.60	1.52

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	504	MAN	O2-C2-C3	-2.45	105.24	110.19
6	A	503	MAN	O2-C2-C3	-2.07	106.01	110.19
6	A	503	MAN	C1-O5-C5	2.07	115.19	112.14
4	A	500	NAG	C2-N2-C7	2.32	126.12	123.11
6	A	504	MAN	O5-C1-C2	2.41	114.74	110.89
5	A	502	BMA	O3-C3-C2	2.72	114.99	110.01
6	A	504	MAN	C1-C2-C3	3.05	113.25	109.55
6	A	504	MAN	C1-O5-C5	5.94	120.87	112.14

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	500	NAG	3	0
4	A	501	NAG	1	0
6	A	504	MAN	0	1

## 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

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	289/306 (94%)	0.87	38 (13%) <b>5</b> <b>4</b>	78, 132, 175, 197	0
2	B	162/179 (90%)	0.49	3 (1%) <b>70</b> <b>64</b>	76, 113, 145, 162	0
3	C	81/118 (68%)	0.26	2 (2%) <b>61</b> <b>55</b>	100, 140, 160, 167	0
All	All	532/603 (88%)	0.66	43 (8%) <b>15</b> <b>14</b>	76, 125, 168, 197	0

All (43) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	23	ILE	7.1
1	A	76	THR	5.6
1	A	46	VAL	5.2
1	A	104	LEU	4.9
1	A	37	TRP	4.6
1	A	24	TRP	4.4
1	A	45	MET	3.9
1	A	35	LEU	3.8
1	A	105	HIS	3.8
1	A	54	GLU	3.6
1	A	36	ASP	3.6
1	A	220	TYR	3.5
1	A	44	GLU	3.4
1	A	77	ILE	3.4
1	A	48	LEU	3.3
1	A	95	GLU	3.0
1	A	144	TRP	3.0
1	A	79	VAL	3.0
1	A	218	LEU	3.0
2	B	188	SER	2.9
1	A	213	ASP	2.8
1	A	300	CYS	2.8
1	A	33	VAL	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	97	LEU	2.7
1	A	27	LYS	2.7
1	A	32	VAL	2.6
1	A	102	LEU	2.5
1	A	227	PHE	2.5
1	A	75	LEU	2.4
1	A	242	GLN	2.4
1	A	314	TYR	2.3
3	C	28	MET	2.3
2	B	180	PHE	2.3
1	A	208	ILE	2.2
1	A	247	LYS	2.2
1	A	298	VAL	2.2
3	C	21	ILE	2.2
2	B	101	GLU	2.2
1	A	143	TRP	2.2
1	A	40	ASP	2.2
1	A	270	SER	2.1
1	A	271	LEU	2.0
1	A	186	GLU	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

## 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
4	NAG	A	500	14/15	0.91	0.32	-1.06	103,108,114,116	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
5	BMA	A	502	11/12	0.82	0.15	-	107,137,141,142	0
4	NAG	A	501	14/15	0.92	0.24	-	105,109,118,119	0
6	MAN	A	504	11/12	0.64	0.31	-	107,111,146,148	0
6	MAN	A	503	11/12	0.68	0.26	-	106,111,150,152	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.