



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 09:25 PM GMT

PDB ID : 1OYH
Title : Crystal Structure of P13 Alanine Variant of Antithrombin
Authors : Johnson, D.J.D.; Huntington, J.A.
Deposited on : 2003-04-04
Resolution : 2.62 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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
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) : trunk26865

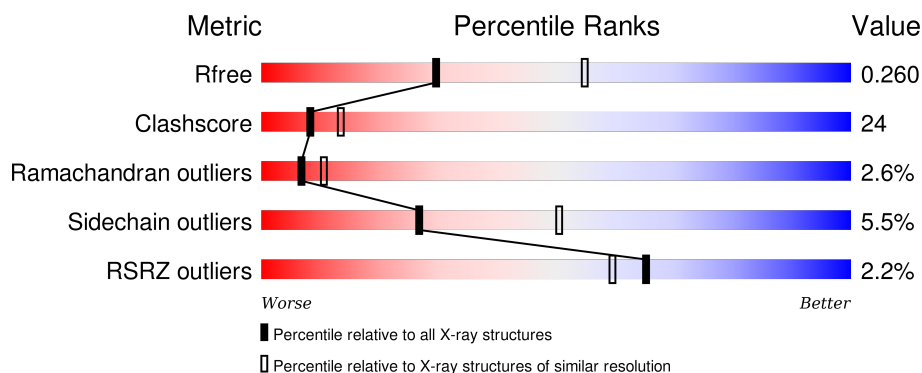
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.62 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	2700 (2.64-2.60)
Clashscore	102246	3065 (2.64-2.60)
Ramachandran outliers	100387	3015 (2.64-2.60)
Sidechain outliers	100360	3015 (2.64-2.60)
RSRZ outliers	91569	2706 (2.64-2.60)

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	I	432	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, green 61%, yellow 31%, orange 5%, grey 3%);"></div> <div style="display: flex; justify-content: space-between; padding: 0 5px;"> % 61% 31% 5% • </div> </div>
2	L	432	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 3%, green 53%, yellow 37%, orange 5%, grey 5%);"></div> <div style="display: flex; justify-content: space-between; padding: 0 5px;"> 3% 53% 37% 5% 5% </div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	NAG	L	831	-	-	-	X
4	NAG	I	841	-	-	-	X

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 6661 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 Antithrombin-III.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	I	416	Total	C	N	O	S	0	0	0
			3259	2083	545	615	16			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	137	ALA	SER	ENGINEERED	UNP P01008
I	381	ALA	GLU	ENGINEERED	UNP P01008

- Molecule 2 is a protein called Antithrombin-III.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	410	Total	C	N	O	S	0	0	0
			3175	2028	518	611	18			

- Molecule 3 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	I	1	Total	C	N	O	0	0
			14	8	1	5		
3	I	1	Total	C	N	O	0	0
			14	8	1	5		
3	L	1	Total	C	N	O	0	0
			14	8	1	5		
3	L	1	Total	C	N	O	0	0
			14	8	1	5		

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	I	2	Total	C	N	O	0	0
			28	16	2	10		
4	L	2	Total	C	N	O	0	0
			28	16	2	10		

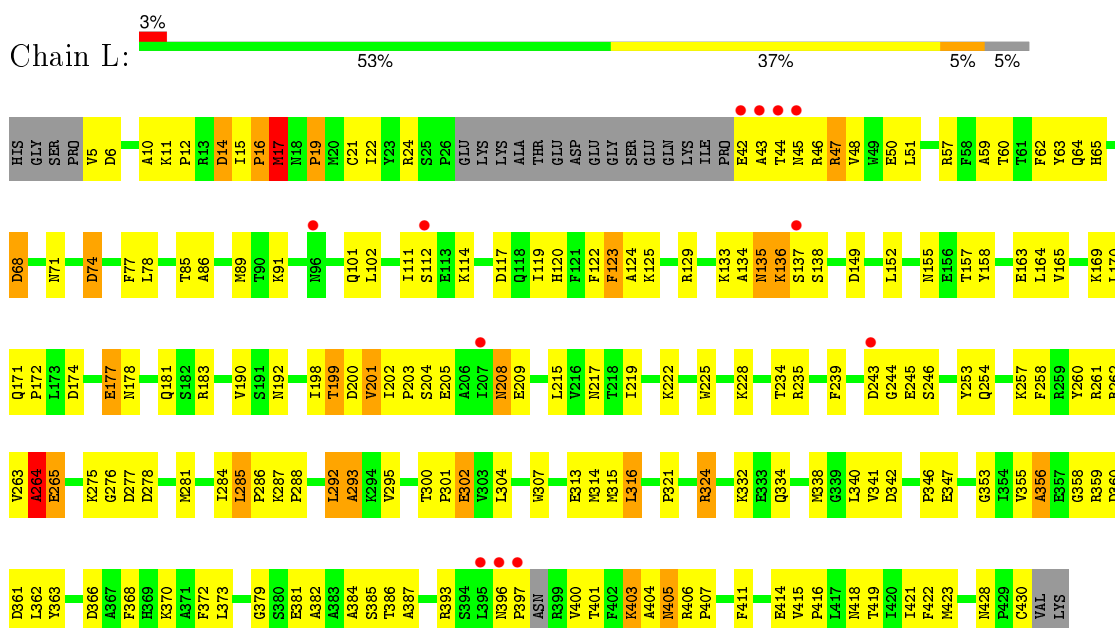
- Molecule 5 is a polymer of unknown type called SUGAR (5-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	L	5	Total	C	N	O	0	0
			61	34	2	25		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	I	33	Total	O	0	0
			33	33		
6	L	21	Total	O	0	0
			21	21		

- Molecule 1: Antithrombin-III



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	69.00Å 98.57Å 89.35Å 90.00° 105.53° 90.00°	Depositor
Resolution (Å)	39.00 – 2.62 39.45 – 2.62	Depositor EDS
% Data completeness (in resolution range)	99.7 (39.00-2.62) 99.8 (39.45-2.62)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	0.07	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.91 (at 2.61Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.221 , 0.258 0.223 , 0.260	Depositor DCC
R_{free} test set	1721 reflections (4.95%)	DCC
Wilson B-factor (Å ²)	55.3	Xtriage
Anisotropy	0.472	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 49.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 34900 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	6661	wwPDB-VP
Average B, all atoms (Å ²)	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.19% 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, 0.375 respectively for untwinned datasets, and 0.333, 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, NDG, 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	I	0.42	0/3325	0.69	1/4503 (0.0%)
2	L	0.39	0/3238	0.66	2/4392 (0.0%)
All	All	0.41	0/6563	0.67	3/8895 (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
2	L	0	1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	I	134	ALA	N-CA-C	6.12	127.53	111.00
2	L	414	GLU	N-CA-C	-5.44	96.31	111.00
2	L	264	ALA	N-CA-C	5.02	124.56	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	L	363	TYR	Sidechain

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	I	3259	0	3200	147	0
2	L	3175	0	3050	168	0
3	I	28	0	26	5	0
3	L	28	0	26	2	0
4	I	28	0	25	1	0
4	L	28	0	25	1	0
5	L	61	0	52	7	0
6	I	33	0	0	2	0
6	L	21	0	0	2	0
All	All	6661	0	6404	316	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 24.

The worst 5 of 316 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)
2:L:86:ALA:HA	2:L:89:MET:HE3	1.39	1.04
2:L:292:LEU:HD23	2:L:407:PRO:HG2	1.46	0.96
2:L:208:ASN:HB3	2:L:393:ARG:HH12	1.28	0.95
2:L:63:TYR:HB2	2:L:423:MET:HE1	1.50	0.91
2:L:42:GLU:HG2	2:L:45:ASN:HD22	1.35	0.91

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	I	412/432 (95%)	378 (92%)	28 (7%)	6 (2%)	13	24
2	L	404/432 (94%)	355 (88%)	34 (8%)	15 (4%)	4	5
All	All	816/864 (94%)	733 (90%)	62 (8%)	21 (3%)	7	10

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	I	112	SER
1	I	135	ASN
1	I	138	SER
2	L	10	ALA
2	L	17	MET

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	I	351/381 (92%)	332 (95%)	19 (5%)	27	51
2	L	338/383 (88%)	319 (94%)	19 (6%)	26	49
All	All	689/764 (90%)	651 (94%)	38 (6%)	27	50

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	I	366	ASP
2	L	68	ASP
2	L	324	ARG
2	L	17	MET
2	L	74	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
2	L	55	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	L	65	HIS
2	L	217	ASN
1	I	233	ASN
2	L	254	GLN

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 ⓘ

9 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	$\# Z > 2$
4	NAG	I	841	1,4	14,14,15	0.59	0	15,19,21	0.86	1 (6%)
4	NAG	I	842	4	14,14,15	0.62	0	15,19,21	0.77	1 (6%)
4	NAG	L	841	2,4	14,14,15	0.44	0	15,19,21	0.81	0
4	NAG	L	842	4	14,14,15	0.52	0	15,19,21	1.07	2 (13%)
5	NDG	L	861	2,5	14,14,15	0.69	0	15,19,21	0.90	1 (6%)
5	NAG	L	862	5	14,14,15	0.67	0	15,19,21	0.81	0
5	BMA	L	863	5	11,11,12	0.76	0	14,15,17	1.03	1 (7%)
5	MAN	L	864	5	11,11,12	0.53	0	14,15,17	0.73	1 (7%)
5	MAN	L	868	5	11,11,12	0.62	0	14,15,17	0.66	1 (7%)

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	I	841	1,4	-	0/6/23/26	0/1/1/1
4	NAG	I	842	4	-	0/6/23/26	0/1/1/1
4	NAG	L	841	2,4	-	0/6/23/26	0/1/1/1
4	NAG	L	842	4	-	0/6/23/26	0/1/1/1
5	NDG	L	861	2,5	-	1/6/23/26	0/1/1/1
5	NAG	L	862	5	-	0/6/23/26	0/1/1/1
5	BMA	L	863	5	-	0/2/19/22	0/1/1/1
5	MAN	L	864	5	-	0/2/19/22	1/1/1/1
5	MAN	L	868	5	-	0/2/19/22	1/1/1/1

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	L	842	NAG	C2-N2-C7	-3.08	119.08	123.04
5	L	861	NDG	C2-N2-C7	-2.80	119.44	123.04
4	I	841	NAG	C2-N2-C7	-2.53	119.79	123.04
4	I	842	NAG	C2-N2-C7	-2.30	120.09	123.04
4	L	842	NAG	C4-C3-C2	-2.07	108.02	111.23

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	L	861	NDG	O7-C7-N2-C2

All (2) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	L	864	MAN	C1-C2-C3-C4-C5-O5
5	L	868	MAN	C1-C2-C3-C4-C5-O5

6 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	I	841	NAG	1	0
4	I	842	NAG	1	0
4	L	841	NAG	1	0
5	L	861	NDG	5	0
5	L	862	NAG	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	L	863	BMA	1	0

5.6 Ligand geometry

4 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
3	NAG	I	801	1	14,14,15	0.59	0	15,19,21	0.67	0
3	NAG	I	861	1	14,14,15	0.67	0	15,19,21	0.72	1 (6%)
3	NAG	L	801	2	14,14,15	0.56	0	15,19,21	0.64	0
3	NAG	L	831	2	14,14,15	0.57	0	15,19,21	0.64	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	I	801	1	-	0/6/23/26	0/1/1/1
3	NAG	I	861	1	-	0/6/23/26	0/1/1/1
3	NAG	L	801	2	-	0/6/23/26	0/1/1/1
3	NAG	L	831	2	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	I	861	NAG	C2-N2-C7	-2.11	120.33	123.04

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	I	801	NAG	4	0
3	I	861	NAG	1	0
3	L	831	NAG	2	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	I	416/432 (96%)	0.06	6 (1%) 78 73	29, 51, 79, 103	0
2	L	410/432 (94%)	0.18	12 (2%) 55 48	32, 56, 99, 108	0
All	All	826/864 (95%)	0.12	18 (2%) 65 59	29, 53, 90, 108	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	I	207	ILE	5.4
2	L	112	SER	3.4
1	I	360	ASP	3.1
2	L	43	ALA	3.0
2	L	44	THR	2.9

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](#)

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(Å ²)	Q<0.9
4	NAG	I	841	14/15	0.87	0.28	2.97	64,69,75,82	0
4	NAG	L	841	14/15	0.90	0.14	-0.99	69,74,79,86	0
4	NAG	I	842	14/15	0.86	0.35	-	88,92,93,94	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
5	NDG	L	861	14/15	0.74	0.29	-	98,104,108,115	0
4	NAG	L	842	14/15	0.81	0.41	-	92,96,98,99	0
5	MAN	L	868	11/12	0.58	0.49	-	139,140,140,141	0
5	BMA	L	863	11/12	0.82	0.35	-	134,136,138,140	0
5	NAG	L	862	14/15	0.64	0.45	-	122,126,128,131	0
5	MAN	L	864	11/12	0.83	0.55	-	141,141,142,142	0

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
3	NAG	L	831	14/15	0.76	0.47	3.58	118,120,124,125	0
3	NAG	I	861	14/15	0.68	0.41	-	92,97,100,101	0
3	NAG	L	801	14/15	0.87	0.31	-	84,85,86,87	0
3	NAG	I	801	14/15	0.84	0.27	-	77,80,82,83	0

6.5 Other polymers [i](#)

There are no such residues in this entry.