



# wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 12:06 PM GMT

PDB ID : 3EVJ  
Title : Intermediate structure of antithrombin bound to the natural pentasaccharide  
Authors : Huntington, J.A.; Belzar, K.J.  
Deposited on : 2008-10-13  
Resolution : 3.00 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>

---

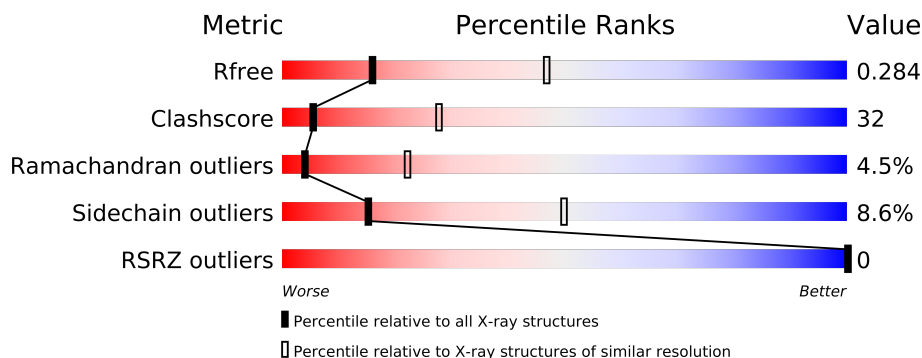
The following versions of software and data (see [references](#)) were used in the production of this report:

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 3.00 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	1216 (3.00-3.00)
Clashscore	79885	1594 (3.00-3.00)
Ramachandran outliers	78287	1537 (3.00-3.00)
Sidechain outliers	78261	1540 (3.00-3.00)
RSRZ outliers	66119	1217 (3.00-3.00)

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	I	432	
1	L	432	

## 2 Entry composition i

There are 7 unique types of molecules in this entry. The entry contains 6574 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 Antithrombin-III.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	I	404	Total	C	N	O	S	0	0	0
			3016	1922	498	579	17			
1	L	417	Total	C	N	O	S	0	0	0
			3148	2009	517	605	17			

- Molecule 2 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

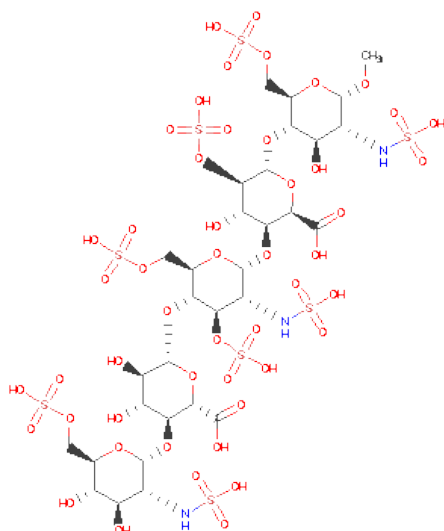
- Molecule 3 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	I	3	Total	C	N	O	0	0
			38	21	2	15		

- 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 TRISULFOAMINO HEPARIN PENTASACCHARIDE (three-letter code: NTO) (formula:  $C_{31}H_{53}N_3O_{49}S_8$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	I	1	Total	C	N	O	S	0	0
			91	31	3	49	8		
5	L	1	Total	C	N	O	S	0	0
			91	31	3	49	8		

- Molecule 6 is a polymer of unknown type called SUGAR (5-MER).

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

- Molecule 7 is water.

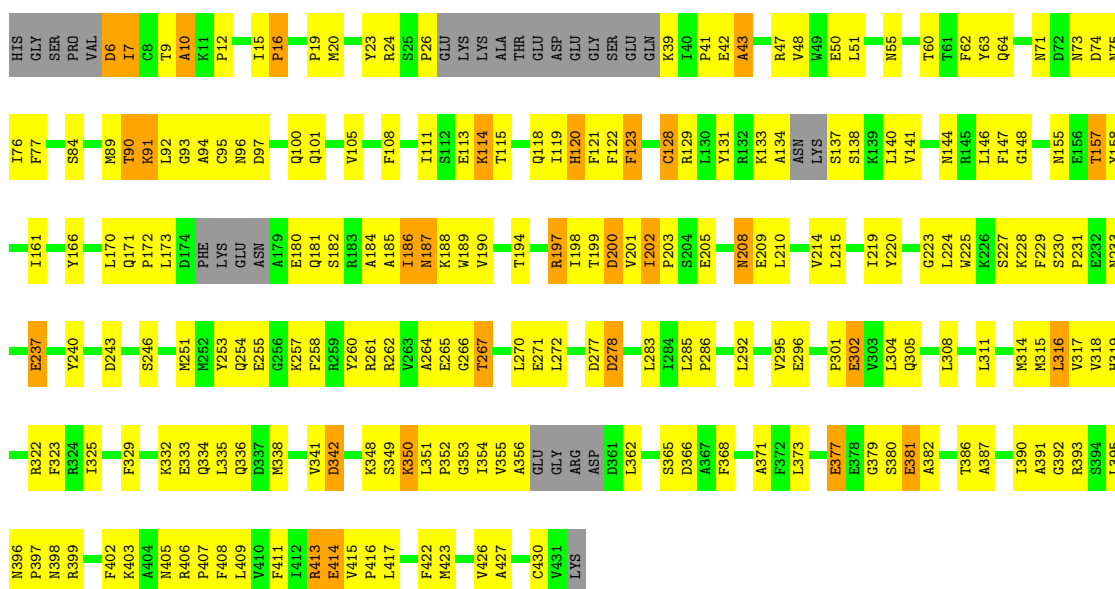
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	I	24	Total 24	O 24	0	0
7	L	21	Total 21	O 21	0	0

### 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.

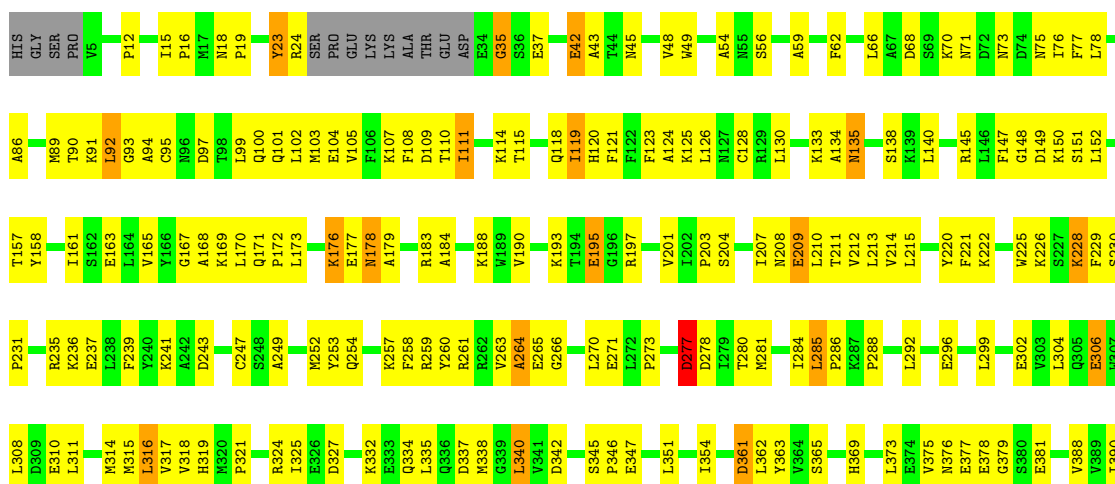
#### • Molecule 1: Antithrombin-III

Chain I: 



#### • Molecule 1: Antithrombin-III

Chain L: 



A391	G392	R393	S394	L395	ASN	P397	T401	A404	N405	R406	P407	F408	F411	I412	R413	E414	L417	N423	A427	N428	P429	C430	V431	LYS
------	------	------	------	------	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	65.85Å 87.06Å 92.38Å 90.00° 106.18° 90.00°	Depositor
Resolution (Å)	63.24 – 3.00 63.24 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.1 (63.24-3.00) 99.1 (63.24-3.00)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	0.12	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.29 (at 3.01Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, $R_{free}$	0.231 , 0.289 0.227 , 0.284	Depositor DCC
$R_{free}$ test set	1016 reflections (5.34%)	DCC
Wilson B-factor (Å <sup>2</sup> )	70.3	Xtriage
Anisotropy	0.162	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 58.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	3 of 20055 reflections (0.015%)	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	6574	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	63.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.70% 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, NTO, 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	0.43	0/3076	0.67	0/4187
1	L	0.46	0/3211	0.74	2/4365 (0.0%)
All	All	0.44	0/6287	0.70	2/8552 (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	I	0	1
3	I	1	0
6	L	1	0
All	All	2	1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	L	35	GLY	N-CA-C	6.30	128.86	113.10
1	L	95	CYS	N-CA-C	5.19	125.02	111.00

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	I	843	MAN	C1
6	L	863	MAN	C1

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	I	63	TYR	Sidechain

## 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	I	3016	0	2762	223	0
1	L	3148	0	2933	170	0
2	I	14	0	13	0	0
2	L	14	0	13	0	0
3	I	38	0	30	8	0
4	I	28	0	25	2	0
4	L	28	0	25	2	0
5	I	91	0	51	3	0
5	L	91	0	51	5	0
6	L	61	0	52	4	0
7	I	24	0	0	2	0
7	L	21	0	0	0	0
All	All	6574	0	5955	401	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 32.

The worst 5 of 401 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:I:332:LYS:O	1:I:336:GLN:HG3	1.71	0.89
1:L:258:PHE:HB2	1:L:316:LEU:HD21	1.55	0.88
1:L:178:ASN:N	1:L:178:ASN:HD22	1.68	0.88
6:L:861:NAG:H62	6:L:862:NAG:HN2	1.39	0.88
1:L:102:LEU:HD23	1:L:340:LEU:HD21	1.56	0.87

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	394/432 (91%)	315 (80%)	63 (16%)	16 (4%)	4	24
1	L	411/432 (95%)	336 (82%)	55 (13%)	20 (5%)	3	20
All	All	805/864 (93%)	651 (81%)	118 (15%)	36 (4%)	4	22

5 of 36 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	I	10	ALA
1	I	186	ILE
1	I	348	LYS
1	I	430	CYS
1	L	42	GLU

### 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	299/383 (78%)	270 (90%)	29 (10%)	12	42
1	L	318/383 (83%)	294 (92%)	24 (8%)	19	57
All	All	617/766 (80%)	564 (91%)	53 (9%)	15	50

5 of 53 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	I	334	GLN
1	L	23	TYR
1	L	361	ASP
1	I	342	ASP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	I	413	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	I	254	GLN
1	I	336	GLN
1	L	178	ASN
1	I	181	GLN
1	L	171	GLN

### 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 ⓘ

12 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	I	841	1,3	12,14,15	0.58	0	15,19,21	1.22	3 (20%)
3	NAG	I	842	3	8,11,15	5.33	1 (12%)	6,12,21	0.73	0
3	MAN	I	843	3	10,11,12	0.52	0	11,15,17	0.49	0
4	NAG	I	861	1,4	12,14,15	0.73	0	15,19,21	0.76	0
4	NAG	I	862	4	12,14,15	0.44	0	15,19,21	0.53	0
4	NAG	L	841	1,4	12,14,15	0.79	0	15,19,21	1.13	1 (6%)
4	NAG	L	842	4	12,14,15	0.43	0	15,19,21	0.76	1 (6%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	NAG	L	861	1,6	12,14,15	0.53	0	15,19,21	0.61	0
6	NAG	L	862	6	12,14,15	0.54	0	15,19,21	0.60	0
6	MAN	L	863	6	10,11,12	0.60	0	11,15,17	0.96	1 (9%)
6	MAN	L	864	6	10,11,12	0.58	0	11,15,17	0.80	0
6	MAN	L	868	6	10,11,12	0.51	0	11,15,17	0.43	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	841	1,3	-	0/6/23/26	0/1/1/1
3	NAG	I	842	3	-	0/8/10/26	0/0/0/1
3	MAN	I	843	3	1/1/4/5	0/2/19/22	0/1/1/1
4	NAG	I	861	1,4	-	0/6/23/26	0/1/1/1
4	NAG	I	862	4	-	0/6/23/26	0/1/1/1
4	NAG	L	841	1,4	-	0/6/23/26	0/1/1/1
4	NAG	L	842	4	-	0/6/23/26	0/1/1/1
6	NAG	L	861	1,6	-	0/6/23/26	0/1/1/1
6	NAG	L	862	6	-	0/6/23/26	0/1/1/1
6	MAN	L	863	6	1/1/4/5	0/2/19/22	0/1/1/1
6	MAN	L	864	6	-	0/2/19/22	0/1/1/1
6	MAN	L	868	6	-	0/2/19/22	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	I	842	NAG	O3-C3	15.05	1.42	1.25

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	I	841	NAG	O5-C5-C6	-3.04	103.79	106.98
4	L	841	NAG	C3-C4-C5	2.54	114.74	110.20
6	L	863	MAN	O5-C5-C6	2.28	109.38	106.98
3	I	841	NAG	C3-C2-N2	-2.28	108.29	111.76
4	L	842	NAG	C3-C2-N2	-2.06	108.62	111.76

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
6	L	863	MAN	C1
3	I	843	MAN	C1

There are no torsion outliers.

There are no ring outliers.

## 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	$\# Z  > 2$
2	NAG	I	801	1	12,14,15	0.36	0	15,19,21	0.76	0
5	NTO	I	863	-	95,95,95	2.78	15 (15%)	147,150,150	1.50	26 (17%)
2	NAG	L	801	1	12,14,15	0.53	0	15,19,21	0.91	0
5	NTO	L	869	-	95,95,95	2.78	15 (15%)	147,150,150	1.54	25 (17%)

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	NAG	I	801	1	-	0/6/23/26	0/1/1/1
5	NTO	I	863	-	-	0/69/169/169	0/5/5/5
2	NAG	L	801	1	-	0/6/23/26	0/1/1/1
5	NTO	L	869	-	-	0/69/169/169	0/5/5/5

The worst 5 of 30 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	L	869	NTO	O4A-S4B	-10.68	1.43	1.60
5	I	863	NTO	O4A-S4B	-10.64	1.43	1.60
5	I	863	NTO	O3A-S3B	-10.62	1.43	1.60

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	L	869	NTO	O3A-S3B	-10.61	1.43	1.60
5	L	869	NTO	O1I-S1G	8.09	1.50	1.42

The worst 5 of 51 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	L	869	NTO	C49-O4A-S4B	-4.81	110.52	118.37
5	L	869	NTO	C39-O3A-S3B	-4.57	110.92	118.37
5	I	863	NTO	C49-O4A-S4B	-4.24	111.45	118.37
5	I	863	NTO	C3F-N3G-S3H	-4.20	108.96	120.83
5	L	869	NTO	C1E-N1F-S1G	-4.14	109.11	120.83

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	I	404/432 (93%)	-0.29	0 100 100	37, 65, 90, 100	0
1	L	417/432 (96%)	-0.31	0 100 100	33, 59, 90, 106	0
All	All	821/864 (95%)	-0.30	0 100 100	33, 62, 90, 106	0

There are no RSRZ outliers to report.

### 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	I	862	14/15	0.38	6.96	128,132,134,134	0
4	NAG	L	842	14/15	0.39	3.70	87,90,92,92	0
3	NAG	I	842	13/15	0.30	1.92	63,71,74,79	0
4	NAG	I	861	14/15	0.21	1.40	110,114,118,123	0
4	NAG	L	841	14/15	0.23	1.24	63,71,75,82	0
6	NAG	L	862	14/15	0.16	0.58	68,74,77,79	0
6	MAN	L	868	11/12	0.19	-0.05	76,81,83,84	0
3	NAG	I	841	14/15	0.20	-0.12	69,72,75,76	0
6	MAN	L	863	11/12	0.17	-0.35	79,81,85,89	0
6	NAG	L	861	14/15	0.17	-0.47	64,67,68,71	0

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
6	MAN	L	864	11/12	0.17	-	90,91,92,92	0
3	MAN	I	843	11/12	0.20	-	82,85,87,87	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
2	NAG	L	801	14/15	0.24	0.34	80,81,83,83	0
5	NTO	I	863	91/91	0.20	0.07	76,97,110,116	0
5	NTO	L	869	91/91	0.22	-0.03	86,100,111,113	0
2	NAG	I	801	14/15	0.21	-0.42	105,107,108,109	6

## 6.5 Other polymers

There are no such residues in this entry.