



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 04:57 PM UTC

PDB ID : 2PMS / pdb\_00002pms  
Title : Crystal structure of the complex of human lactoferrin N-lobe and lactoferrin-binding domain of pneumococcal surface protein A  
Authors : Chattopadhyay, D.; Senkovich, O.; Cook, W.J.  
Deposited on : 2007-04-23  
Resolution : 2.91 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4-5-2 with Phenix2.0  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 2.0  
EDS : 3.0  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.49

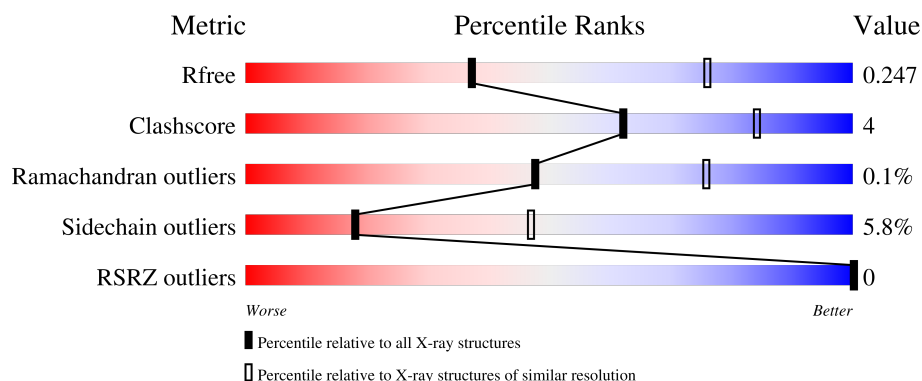
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*




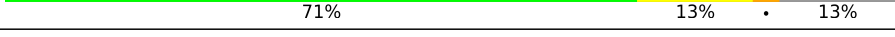
The reported resolution of this entry is 2.91 Å.

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))
$R_{free}$	180053	2995 (2.94-2.90)
Clashscore	190562	3213 (2.94-2.90)
Ramachandran outliers	187476	3128 (2.94-2.90)
Sidechain outliers	187428	3130 (2.94-2.90)
RSRZ outliers	180081	2995 (2.94-2.90)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. 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	344	
1	B	344	
2	C	125	
2	D	125	

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 6996 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 Lactotransferrin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	331	Total	C	N	O	S	0	0	0
			2591	1634	471	473	13			
1	B	331	Total	C	N	O	S	0	0	0
			2591	1634	471	473	13			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	123	THR	ASN	conflict	UNP Q5EK51
B	123	THR	ASN	conflict	UNP Q5EK51

- Molecule 2 is a protein called PNEUMOCOCCAL SURFACE PROTEIN A (PSPA).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	109	Total	C	N	O	S	0	0	0
			860	533	140	186	1			
2	D	109	Total	C	N	O	S	0	0	0
			860	533	140	186	1			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	164	GLY	-	expression tag	UNP Q8DRI0
C	165	SER	-	expression tag	UNP Q8DRI0
C	166	HIS	-	expression tag	UNP Q8DRI0
C	167	MET	-	expression tag	UNP Q8DRI0
D	164	GLY	-	expression tag	UNP Q8DRI0
D	165	SER	-	expression tag	UNP Q8DRI0
D	166	HIS	-	expression tag	UNP Q8DRI0
D	167	MET	-	expression tag	UNP Q8DRI0

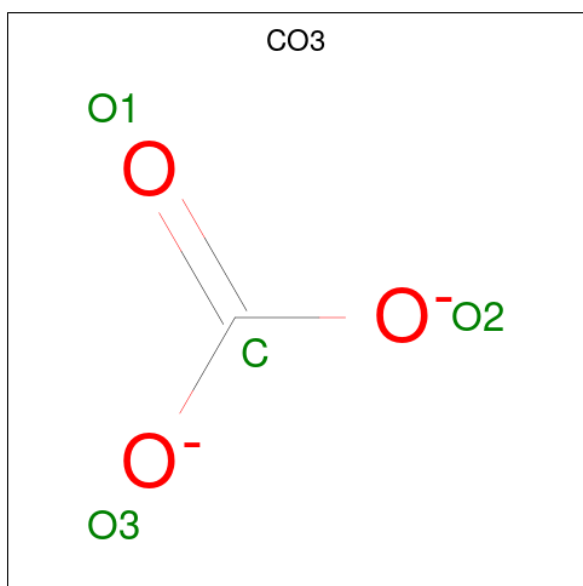
- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (CCD ID: NAG) (formula:

C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



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

- Molecule 4 is CARBONATE ION (CCD ID: CO3) (formula: CO<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			4	1	3		

*Continued on next page...*

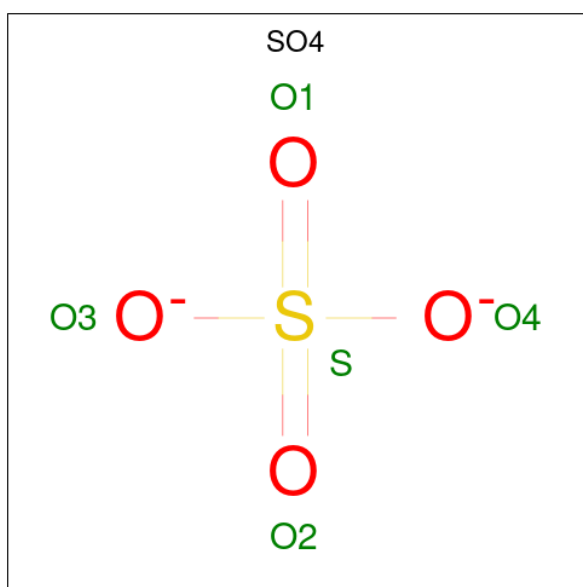
*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	C	O	0	0
			4	1	3		

- Molecule 5 is FE (III) ION (CCD ID: FE) (formula: Fe).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	Fe	0	0
			1	1		
5	B	1	Total	Fe	0	0
			1	1		

- Molecule 6 is SULFATE ION (CCD ID: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	O	S	0	0
			5	4	1		
6	A	1	Total	O	S	0	0
			5	4	1		
6	A	1	Total	O	S	0	0
			5	4	1		
6	B	1	Total	O	S	0	0
			5	4	1		
6	B	1	Total	O	S	0	0
			5	4	1		
6	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 7 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	C	1	Total 1	Zn 1	0	0
7	D	1	Total 1	Zn 1	0	0

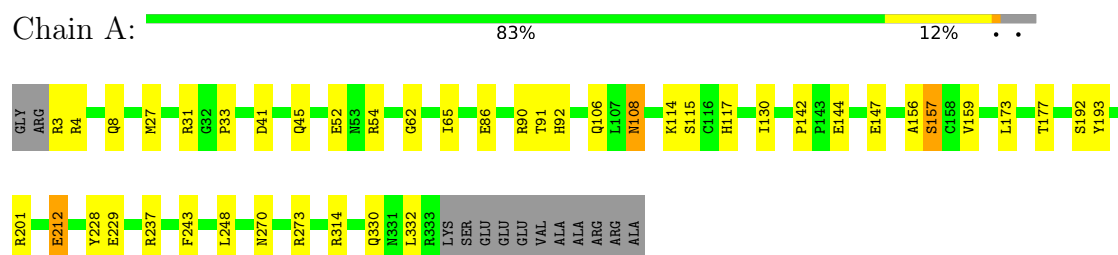
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	8	Total 8	O 8	0	0
8	B	10	Total 10	O 10	0	0
8	C	3	Total 3	O 3	0	0
8	D	3	Total 3	O 3	0	0

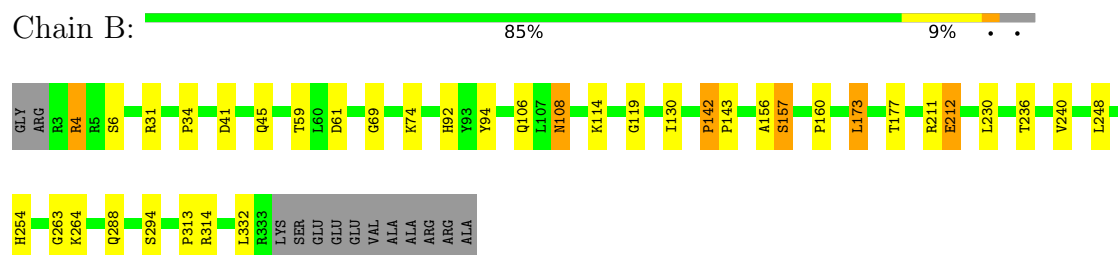
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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.

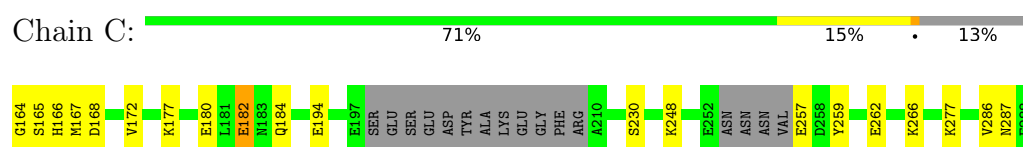
- Molecule 1: Lactotransferrin



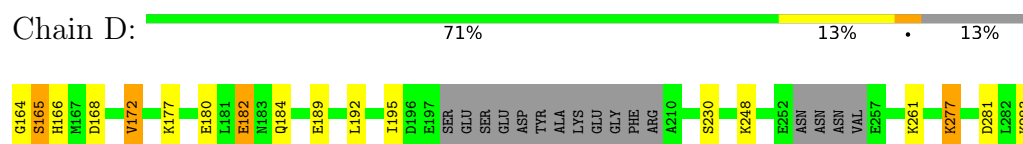
- Molecule 1: Lactotransferrin



- Molecule 2: PNEUMOCOCCAL SURFACE PROTEIN A (PSPA)



- Molecule 2: PNEUMOCOCCAL SURFACE PROTEIN A (PSPA)



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	130.18Å 130.18Å 80.80Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	15.00 – 2.91 15.00 – 2.91	Depositor EDS
% Data completeness (in resolution range)	95.2 (15.00-2.91) 95.2 (15.00-2.91)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.24 (at 2.91Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.203 , 0.249 0.201 , 0.247	Depositor DCC
$R_{free}$ test set	1694 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	39.5	Xtriage
Anisotropy	0.221	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 3.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.467 for -h,-k,l 0.043 for h,-h-k,-l 0.043 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	6996	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.21% 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: ZN, SO4, FE, NAG, CO3

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	0.58	0/2655	0.82	0/3591
1	B	0.58	1/2655 (0.0%)	0.80	0/3591
2	C	0.61	0/863	0.94	3/1148 (0.3%)
2	D	0.64	1/863 (0.1%)	0.97	3/1148 (0.3%)
All	All	0.59	2/7036 (0.0%)	0.85	6/9478 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	165	SER	N-CA	5.63	1.52	1.46
1	B	142	PRO	CA-C	5.01	1.54	1.51

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	165	SER	N-CA-C	7.91	121.25	108.76
2	C	165	SER	N-CA-C	7.11	120.15	108.99
2	D	166	HIS	N-CA-C	6.04	118.03	108.79
2	C	168	ASP	N-CA-C	5.69	117.48	111.28
2	D	168	ASP	N-CA-C	5.27	117.11	111.36

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	2591	0	2534	30	0
1	B	2591	0	2534	21	1
2	C	860	0	862	5	1
2	D	860	0	862	10	0
3	A	14	0	13	0	0
3	B	14	0	13	0	0
4	A	4	0	0	0	0
4	B	4	0	0	0	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
6	A	15	0	0	1	0
6	B	15	0	0	0	0
7	C	1	0	0	0	0
7	D	1	0	0	0	0
8	A	8	0	0	0	0
8	B	10	0	0	0	0
8	C	3	0	0	1	0
8	D	3	0	0	1	0
All	All	6996	0	6818	57	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 4.

The worst 5 of 57 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)
1:A:92:HIS:HE2	2:D:164:GLY:N	1.52	1.08
1:A:330:GLN:HE22	1:B:294:SER:H	1.03	0.99
1:A:92:HIS:NE2	2:D:164:GLY:N	2.30	0.78
1:A:330:GLN:NE2	1:B:294:SER:H	1.86	0.66
1:A:114:LYS:HB3	1:A:173:LEU:HD11	1.77	0.66

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)
1:B:92:HIS:NE2	2:C:164:GLY:N[1_556]	2.16	0.04

## 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	329/344 (96%)	312 (95%)	17 (5%)	0	100	100
1	B	329/344 (96%)	315 (96%)	14 (4%)	0	100	100
2	C	103/125 (82%)	100 (97%)	2 (2%)	1 (1%)	12	37
2	D	103/125 (82%)	97 (94%)	6 (6%)	0	100	100
All	All	864/938 (92%)	824 (95%)	39 (4%)	1 (0%)	48	76

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	194	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	A	273/282 (97%)	262 (96%)	11 (4%)	28	61
1	B	273/282 (97%)	261 (96%)	12 (4%)	25	57
2	C	91/105 (87%)	83 (91%)	8 (9%)	9	28
2	D	91/105 (87%)	80 (88%)	11 (12%)	5	15
All	All	728/774 (94%)	686 (94%)	42 (6%)	18	47

5 of 42 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	277	LYS
2	D	195	ILE
2	C	286	VAL
2	D	172	VAL
2	D	261	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	166	GLN
1	B	270	ASN
2	C	183	ASN
1	B	330	GLN
1	A	262	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 14 ligands modelled in this entry, 4 are monoatomic - leaving 10 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
6	SO4	A	350	-	4,4,4	0.23	0	6,6,6	0.12	0
3	NAG	B	345	1	14,14,15	0.69	0	17,19,21	1.26	2 (11%)
6	SO4	A	349	-	4,4,4	0.23	0	6,6,6	0.19	0
4	CO3	B	346	5	3,3,3	0.86	0	2,3,3	0.30	0
6	SO4	B	348	-	4,4,4	0.24	0	6,6,6	0.33	0
6	SO4	A	348	-	4,4,4	0.24	0	6,6,6	0.33	0
4	CO3	A	346	5	3,3,3	0.90	0	2,3,3	0.13	0
6	SO4	B	349	-	4,4,4	0.25	0	6,6,6	0.22	0
6	SO4	B	350	-	4,4,4	0.24	0	6,6,6	0.16	0
3	NAG	A	345	1	14,14,15	0.69	0	17,19,21	1.26	1 (5%)

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	B	345	1	-	2/6/23/26	0/1/1/1
3	NAG	A	345	1	-	3/6/23/26	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(°)
3	A	345	NAG	C1-O5-C5	3.14	116.39	112.19
3	B	345	NAG	C1-O5-C5	2.80	115.93	112.19
3	B	345	NAG	C4-C3-C2	2.46	114.62	111.02

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	345	NAG	C8-C7-N2-C2
3	A	345	NAG	O7-C7-N2-C2
3	B	345	NAG	C8-C7-N2-C2
3	B	345	NAG	O7-C7-N2-C2
3	A	345	NAG	C4-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	349	SO4	1	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, 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	331/344 (96%)	-1.84	0 100 100	17, 28, 43, 51	0
1	B	331/344 (96%)	-1.85	0 100 100	17, 28, 43, 51	0
2	C	109/125 (87%)	-1.77	0 100 100	11, 37, 55, 61	0
2	D	109/125 (87%)	-1.80	0 100 100	11, 37, 55, 61	0
All	All	880/938 (93%)	-1.83	0 100 100	11, 30, 47, 61	0

There are no RSRZ outliers to report.

### 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 oligosaccharides 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. 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	B-factors(Å <sup>2</sup> )	Q<0.9
3	NAG	B	345	14/15	0.98	0.04	57,60,63,64	0
3	NAG	A	345	14/15	0.99	0.04	57,59,62,62	0
6	SO4	A	349	5/5	0.99	0.05	73,73,73,73	0
6	SO4	A	350	5/5	0.99	0.03	71,72,72,72	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
6	SO4	B	349	5/5	0.99	0.06	65,65,65,65	0
5	FE	B	347	1/1	1.00	0.01	18,18,18,18	0
6	SO4	A	348	5/5	1.00	0.02	25,26,27,27	0
4	CO3	A	346	4/4	1.00	0.02	21,21,21,21	0
4	CO3	B	346	4/4	1.00	0.02	21,21,21,22	0
6	SO4	B	348	5/5	1.00	0.02	26,26,26,27	0
5	FE	A	347	1/1	1.00	0.01	18,18,18,18	0
6	SO4	B	350	5/5	1.00	0.05	62,63,63,63	0
7	ZN	C	502	1/1	1.00	0.01	20,20,20,20	0
7	ZN	D	501	1/1	1.00	0.01	16,16,16,16	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.