



# Full wwPDB X-ray Structure Validation Report ⓘ

Dec 2, 2023 – 09:48 pm GMT

PDB ID : 2BJJ  
Title : Structure of recombinant human lactoferrin produced in the milk of transgenic cows  
Authors : Thomassen, E.A.J.; Van Veen, H.A.; Van Berkel, P.H.C.; Nuijens, J.H.; Abrahams, J.P.  
Deposited on : 2005-02-03  
Resolution : 2.40 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

<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 : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtrriage (Phenix) : 1.13  
EDS : 2.36  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

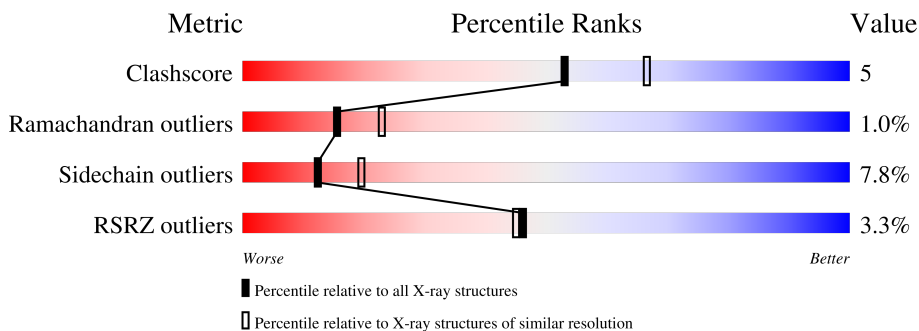
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

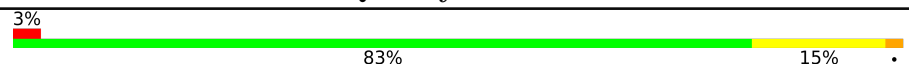
The reported resolution of this entry is 2.40 Å.

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))
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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	X	692	

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
4	NAG	X	901	X	-	-	-

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 5457 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
			Total	C	N	O	S			
1	X	692	5365	3348	964	1015	38	0	1	0

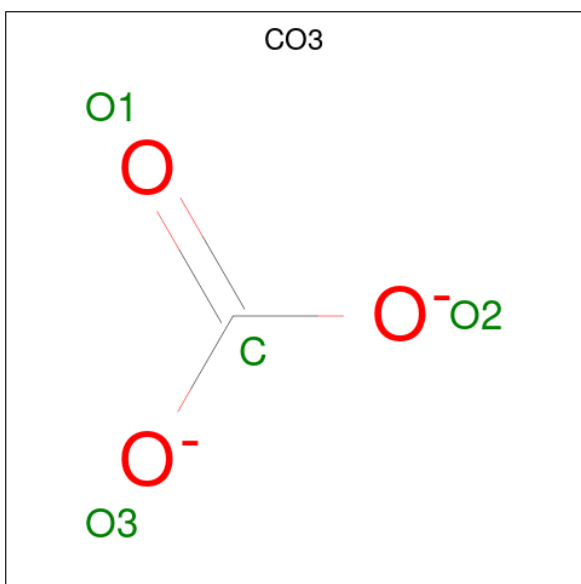
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
X	29	ARG	LYS	conflict	UNP P02788
X	561	ASP	GLU	conflict	UNP P02788

- Molecule 2 is FE (III) ION (three-letter code: FE) (formula: Fe).

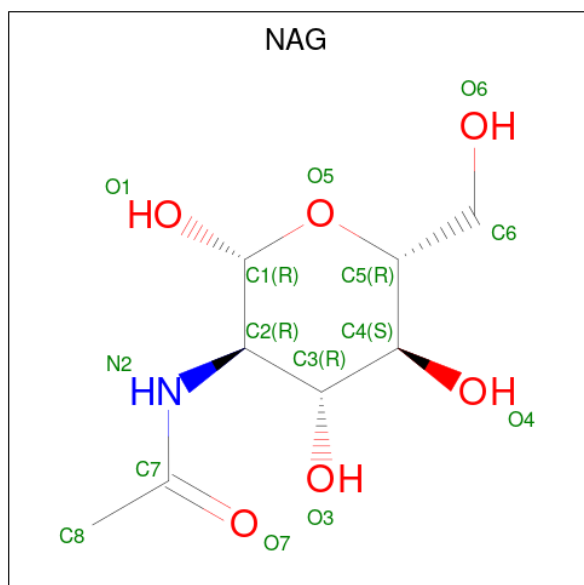
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	X	2	Total Fe 2 2	0	0

- Molecule 3 is CARBONATE ION (three-letter code: CO3) (formula: CO<sub>3</sub>).



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

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	X	1	Total C N O 14 8 1 5	0	0
4	X	1	Total C N O 14 8 1 5	0	0

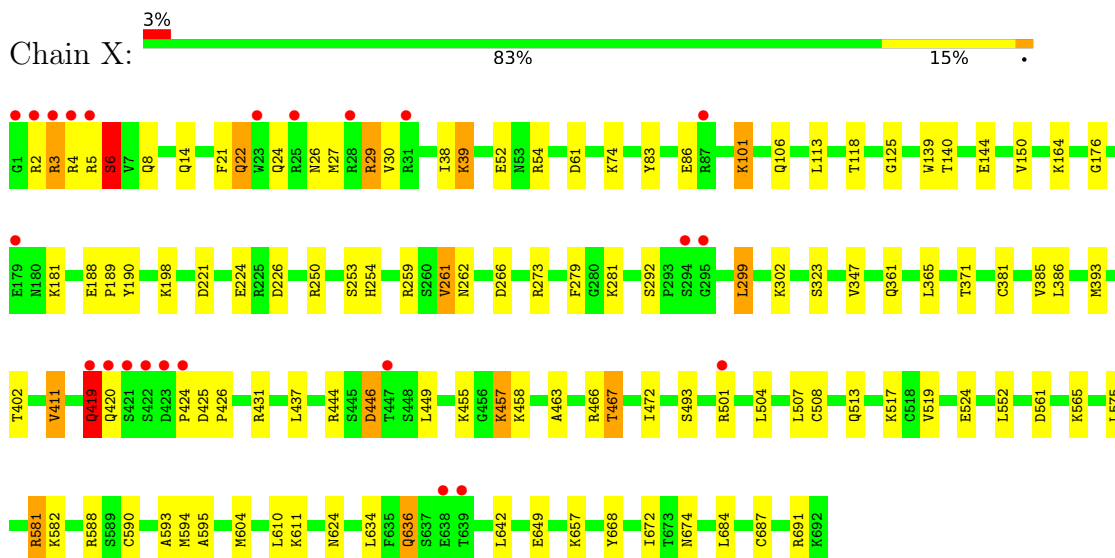
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	X	54	Total O 54 54	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



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	55.94Å 97.38Å 156.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.63 – 2.40 19.63 – 2.40	Depositor EDS
% Data completeness (in resolution range)	95.1 (19.63-2.40) 95.1 (19.63-2.40)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.41 (at 2.41Å)	Xtrriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.181 , 0.231 0.182 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	46.9	Xtrriage
Anisotropy	0.436	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 43.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	5457	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	50.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.01% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, CO3, FE

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	X	0.63	0/5479	0.68	0/7408

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 [i](#)

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	X	5365	0	5223	58	0
2	X	2	0	0	0	0
3	X	8	0	0	0	0
4	X	28	0	26	0	0
5	X	54	0	0	1	0
All	All	5457	0	5249	58	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 5.

All (58) 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:X:261:VAL:HG13	1:X:262:ASN:H	1.13	1.06
1:X:261:VAL:HG13	1:X:262:ASN:N	1.92	0.84
1:X:261:VAL:CG1	1:X:262:ASN:H	1.92	0.82
1:X:2:ARG:HG2	1:X:3:ARG:H	1.55	0.71
1:X:221:ASP:O	1:X:224:GLU:HG2	1.93	0.69
1:X:687:CYS:O	1:X:691:ARG:HG2	1.92	0.69
1:X:61:ASP:HA	1:X:254:HIS:CD2	2.31	0.66
1:X:259:ARG:NH1	1:X:262:ASN:O	2.30	0.65
1:X:52:GLU:HG3	1:X:54:ARG:HD2	1.78	0.65
1:X:347:VAL:HG22	1:X:371:THR:HG22	1.78	0.65
1:X:347:VAL:HG22	1:X:371:THR:CG2	2.28	0.64
1:X:323:SER:HB3	1:X:386:LEU:O	1.99	0.63
1:X:419:GLN:HG2	1:X:419:GLN:O	2.00	0.61
1:X:27:MET:HE2	1:X:279:PHE:HE2	1.66	0.60
1:X:27:MET:CE	1:X:279:PHE:HE2	2.16	0.58
1:X:467:THR:HG21	1:X:595:ALA:HB1	1.90	0.54
1:X:27:MET:CE	1:X:279:PHE:CE2	2.91	0.53
1:X:2:ARG:HG2	1:X:3:ARG:N	2.25	0.51
1:X:144:GLU:O	1:X:144:GLU:HG3	2.10	0.51
1:X:463:ALA:HB3	1:X:466:ARG:HD3	1.93	0.50
1:X:261:VAL:CG1	1:X:262:ASN:N	2.61	0.49
1:X:14:GLN:HE21	1:X:39:LYS:HE2	1.78	0.49
1:X:27:MET:HE2	1:X:279:PHE:CE2	2.45	0.49
1:X:106:GLN:NE2	5:X:2006:HOH:O	2.44	0.49
1:X:190:TYR:CE1	1:X:198:LYS:HG2	2.49	0.48
1:X:446:ASP:O	1:X:581:ARG:NH1	2.47	0.47
1:X:508:CYS:HB3	1:X:524:GLU:OE1	2.15	0.47
1:X:5:ARG:O	1:X:6:SER:HB2	2.13	0.47
1:X:472:ILE:HD13	1:X:593:ALA:HB3	1.97	0.47
1:X:449:LEU:CD1	1:X:457:LYS:HD3	2.45	0.47
1:X:30:VAL:O	1:X:30:VAL:HG12	2.15	0.46
1:X:431:ARG:HH22	1:X:649:GLU:CD	2.19	0.46
1:X:190:TYR:HE1	1:X:198:LYS:HG2	1.81	0.46
1:X:29:ARG:HD3	1:X:30:VAL:HG23	1.98	0.46
1:X:176:GLY:HA3	1:X:181:LYS:HA	1.99	0.45
1:X:668:TYR:CZ	1:X:672:ILE:HD11	2.51	0.45
1:X:493:SER:HB2	1:X:504:LEU:HA	1.98	0.45
1:X:164:LYS:HB2	1:X:164:LYS:HE3	1.86	0.45
1:X:83:TYR:CE2	1:X:253:SER:HB2	2.53	0.44
1:X:101:LYS:HD2	1:X:226:ASP:O	2.17	0.44
1:X:467:THR:HG21	1:X:595:ALA:CB	2.46	0.44
1:X:22:GLN:O	1:X:26:ASN:ND2	2.51	0.43

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:X:3:ARG:HG3	1:X:4:ARG:N	2.32	0.43
1:X:8:GLN:HB3	1:X:38:ILE:HD13	1.99	0.43
1:X:292:SER:HB3	1:X:299:LEU:HD22	2.01	0.43
1:X:411:VAL:HG11	1:X:610:LEU:HD23	2.02	0.42
1:X:458:LYS:HB3	1:X:507:LEU:HD11	2.01	0.42
1:X:21:PHE:O	1:X:24:GLN:HB3	2.20	0.42
1:X:582:LYS:HD2	1:X:590:CYS:HB2	2.00	0.41
1:X:425:ASP:HA	1:X:426:PRO:HD2	1.88	0.41
1:X:381:CYS:HB3	1:X:393:MET:CE	2.50	0.41
1:X:419:GLN:O	1:X:419:GLN:CG	2.68	0.41
1:X:402:THR:HG23	1:X:684:LEU:HD13	2.02	0.41
1:X:636:GLN:HE21	1:X:636:GLN:HA	1.86	0.41
1:X:684:LEU:HD12	1:X:684:LEU:HA	1.90	0.41
1:X:139:TRP:HE1	1:X:144:GLU:HG3	1.85	0.41
1:X:118:THR:OG1	1:X:125:GLY:HA3	2.21	0.40
1:X:188:GLU:HA	1:X:189:PRO:HD3	1.96	0.40

There are no symmetry-related clashes.

## 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	X	691/692 (100%)	653 (94%)	31 (4%)	7 (1%)	<b>15</b> <b>23</b>

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	X	261	VAL
1	X	281	LYS
1	X	419	GLN
1	X	424	PRO
1	X	6	SER

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
1	X	420	GLN
1	X	467	THR

### 5.3.2 Protein sidechains [i](#)

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	X	576/575 (100%)	530 (92%)	46 (8%)	<b>12</b> <b>18</b>

All (46) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	X	3	ARG
1	X	6	SER
1	X	22	GLN
1	X	29	ARG
1	X	39	LYS
1	X	74	LYS
1	X	86	GLU
1	X	101	LYS
1	X	113	LEU
1	X	140	THR
1	X	150	VAL
1	X	250	ARG
1	X	266	ASP
1	X	273	ARG
1	X	299	LEU
1	X	302	LYS
1	X	361	GLN
1	X	365	LEU
1	X	385	VAL
1	X	411	VAL
1	X	419	GLN
1	X	437	LEU
1	X	444	ARG
1	X	446	ASP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	X	455	LYS
1	X	457	LYS
1	X	501	ARG
1	X	513	GLN
1	X	517	LYS
1	X	519	VAL
1	X	552	LEU
1	X	561	ASP
1	X	565	LYS
1	X	575	LEU
1	X	581	ARG
1	X	588	ARG
1	X	594	MET
1	X	604[A]	MET
1	X	604[B]	MET
1	X	611	LYS
1	X	624	ASN
1	X	634	LEU
1	X	636	GLN
1	X	642	LEU
1	X	657	LYS
1	X	674	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
1	X	14	GLN
1	X	22	GLN
1	X	26	ASN
1	X	106	GLN
1	X	235	ASN
1	X	330	GLN
1	X	331	ASN
1	X	361	GLN
1	X	471	ASN
1	X	513	GLN
1	X	636	GLN

### 5.3.3 RNA

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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 2 are monoatomic - leaving 4 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
3	CO3	X	695	2	2,3,3	0.54	0	2,3,3	0.50	0
4	NAG	X	901	1	14,14,15	0.62	0	17,19,21	1.70	6 (35%)
4	NAG	X	941	1	14,14,15	0.53	0	17,19,21	1.30	1 (5%)
3	CO3	X	696	2	2,3,3	0.52	0	2,3,3	0.80	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
4	NAG	X	901	1	1/1/5/7	6/6/23/26	0/1/1/1
4	NAG	X	941	1	-	4/6/23/26	0/1/1/1

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	X	941	NAG	C1-O5-C5	3.87	117.44	112.19
4	X	901	NAG	C2-N2-C7	3.03	127.22	122.90

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	X	901	NAG	C3-C4-C5	2.92	115.45	110.24
4	X	901	NAG	C1-O5-C5	2.62	115.74	112.19
4	X	901	NAG	C4-C3-C2	2.34	114.44	111.02
4	X	901	NAG	O5-C1-C2	-2.24	107.76	111.29
4	X	901	NAG	C1-C2-N2	2.04	113.97	110.49

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	X	901	NAG	C1

All (10) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	X	901	NAG	C4-C5-C6-O6
4	X	941	NAG	C1-C2-N2-C7
4	X	941	NAG	C8-C7-N2-C2
4	X	901	NAG	O5-C5-C6-O6
4	X	941	NAG	O7-C7-N2-C2
4	X	901	NAG	C8-C7-N2-C2
4	X	941	NAG	C3-C2-N2-C7
4	X	901	NAG	O7-C7-N2-C2
4	X	901	NAG	C1-C2-N2-C7
4	X	901	NAG	C3-C2-N2-C7

There are no ring outliers.

No monomer is involved in short contacts.

## 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	X	692/692 (100%)	-0.25	23 (3%) 46 45	29, 49, 72, 108	0

All (23) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	X	1	GLY	6.6
1	X	3	ARG	6.3
1	X	422	SER	6.1
1	X	2	ARG	5.9
1	X	5	ARG	5.5
1	X	423	ASP	4.2
1	X	4	ARG	4.1
1	X	420	GLN	3.8
1	X	419	GLN	3.5
1	X	501	ARG	3.4
1	X	295	GLY	3.0
1	X	421	SER	2.8
1	X	638	GLU	2.8
1	X	87	ARG	2.6
1	X	28	ARG	2.5
1	X	639	THR	2.5
1	X	31	ARG	2.4
1	X	179	GLU	2.3
1	X	23	TRP	2.3
1	X	424	PRO	2.2
1	X	294	SER	2.2
1	X	25	ARG	2.1
1	X	447	THR	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 monosaccharides 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
4	NAG	X	941	14/15	0.81	0.42	80,85,89,89	0
4	NAG	X	901	14/15	0.86	0.41	81,86,87,88	0
3	CO3	X	695	4/4	0.99	0.07	28,29,30,31	0
3	CO3	X	696	4/4	0.99	0.09	22,22,23,23	0
2	FE	X	693	1/1	1.00	0.06	33,33,33,33	0
2	FE	X	694	1/1	1.00	0.07	30,30,30,30	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.