



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 17, 2018 – 01:38 am GMT

PDB ID : 2DSF  
Title : Structure of the complex of C-terminal lobe of bovine lactoferrin with xylose at 2.8Å resolution  
Authors : Mir, R.; Singh, N.; Sinha, M.; Sharma, S.; Bhushan, A.; Singh, T.P.  
Deposited on : 2006-06-29  
Resolution : 2.80 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : NOT EXECUTED  
EDS : NOT EXECUTED  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30686

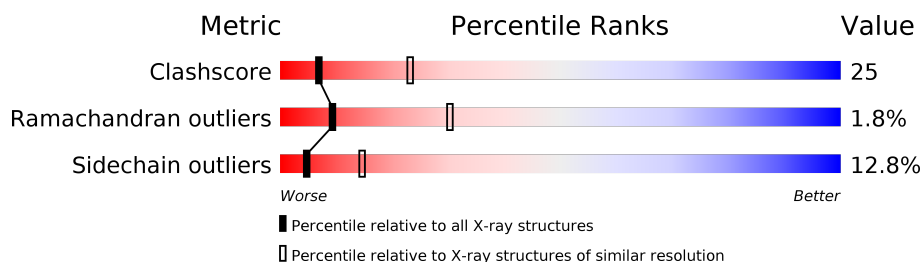
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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	122078	3207 (2.80-2.80)
Ramachandran outliers	120005	3156 (2.80-2.80)
Sidechain outliers	119972	3158 (2.80-2.80)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	345	

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
6	LXC	A	3937	X	-	-	-

## 2 Entry composition [i](#)

There are 11 unique types of molecules in this entry. The entry contains 2921 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	341	Total	C	N	O	S	0	0	0
			2604	1622	454	507	21			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	565	LYS	ASN	SEE REMARK 999	UNP P24627
A	608	GLU	LYS	SEE REMARK 999	UNP P24627

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



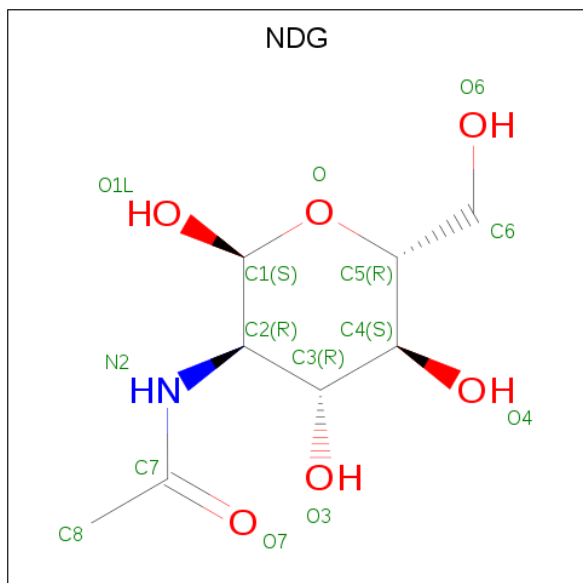
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		

*Continued on next page...*

Continued from previous page...

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

- Molecule 3 is 2-(ACETYLAMINO)-2-DEOXY-A-D-GLUCOPYRANOSE (three-letter code: NDG) (formula:  $C_8H_{15}NO_6$ ).



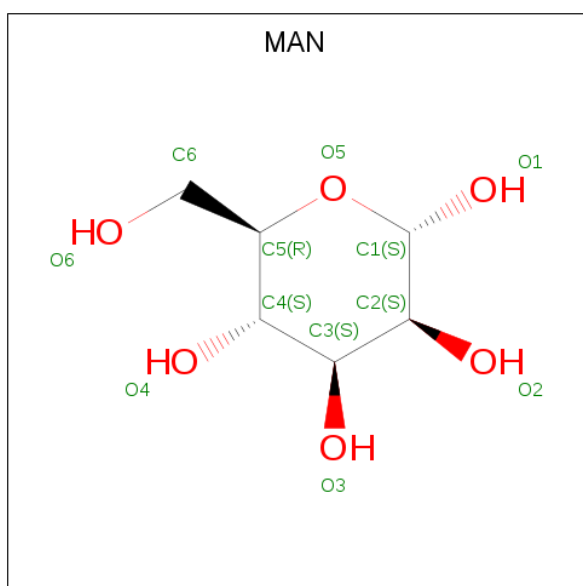
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 4 is BETA-D-MANNOSE (three-letter code: BMA) (formula:  $C_6H_{12}O_6$ ).



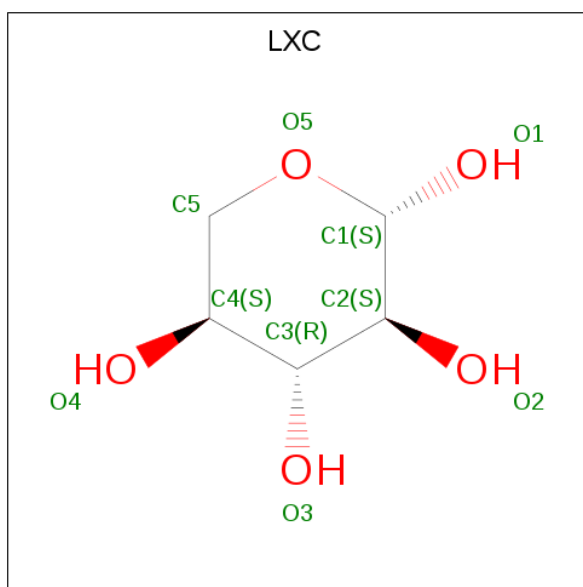
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			11	6	5		
4	A	1	Total	C	O	0	0
			11	6	5		

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



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			11	6	5		

- Molecule 6 is BETA-L-XYLOPYRANOSE (three-letter code: LXC) (formula:  $C_5H_{10}O_5$ ).



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

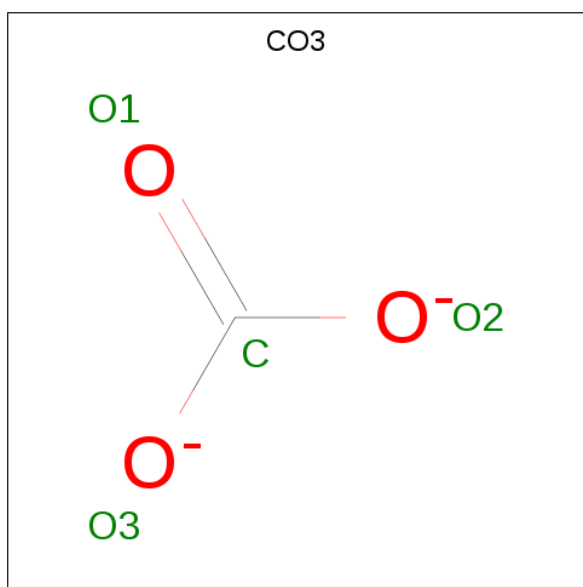
- Molecule 7 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	2	Total	Zn	0	0
			2	2		

- Molecule 8 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	1	Total	Fe	0	0
			1	1		

- Molecule 9 is CARBONATE ION (three-letter code: CO3) (formula: CO<sub>3</sub>).



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

- Molecule 10 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

- Molecule 11 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	192	Total 192	O 192	0	0

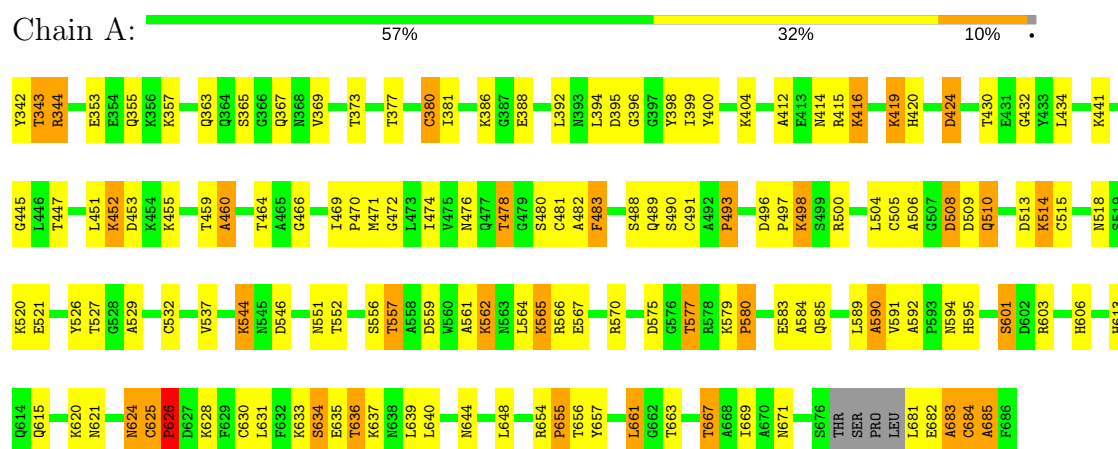


### 3 Residue-property plots

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

Note EDS was not executed.

#### • Molecule 1: Lactotransferrin



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	63.45Å 50.44Å 65.87Å 90.00° 107.80° 90.00°	Depositor
Resolution (Å)	19.94 – 2.80	Depositor
% Data completeness (in resolution range)	92.9 (19.94-2.80)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNS 0.9	Depositor
R, $R_{free}$	0.201 , 0.231	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2921	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

## 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: ZN, BMA, NAG, CO3, NDG, LXC, SO4, MAN, 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	A	0.57	2/2652 (0.1%)	1.07	14/3591 (0.4%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	683	ALA	CA-CB	-5.56	1.40	1.52
1	A	510	GLN	CA-CB	-5.01	1.43	1.53

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	625	CYS	C-N-CD	-22.02	72.15	120.60
1	A	683	ALA	N-CA-C	13.13	146.44	111.00
1	A	625	CYS	C-N-CA	8.49	157.66	122.00
1	A	656	THR	N-CA-C	-8.45	88.20	111.00
1	A	626	PRO	CA-N-CD	-6.66	102.18	111.50

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	A	2604	0	2518	131	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	56	0	49	2	0
3	A	14	0	12	1	0
4	A	22	0	20	1	0
5	A	11	0	9	0	0
6	A	10	0	7	0	0
7	A	2	0	0	0	0
8	A	1	0	0	0	0
9	A	4	0	0	0	0
10	A	5	0	0	0	0
11	A	192	0	0	22	0
All	All	2921	0	2615	132	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 25.

The worst 5 of 132 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:625:CYS:SG	1:A:626:PRO:HD3	1.71	1.28
1:A:395:ASP:HA	1:A:595:HIS:CD2	2.00	0.96
1:A:625:CYS:SG	1:A:626:PRO:CD	2.62	0.88
1:A:685:ALA:HB2	11:A:4026:HOH:O	1.75	0.85
1:A:474:ILE:O	1:A:478:THR:HB	1.76	0.85

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	A	337/345 (98%)	302 (90%)	29 (9%)	6 (2%)	<b>9</b> 30

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	508	ASP
1	A	626	PRO
1	A	685	ALA
1	A	590	ALA
1	A	636	THR

### 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	282/286 (99%)	246 (87%)	36 (13%)	5 14

5 of 36 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	515	CYS
1	A	562	LYS
1	A	661	LEU
1	A	544	LYS
1	A	565	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	458	HIS
1	A	671	ASN
1	A	585	GLN
1	A	414	ASN
1	A	510	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 14 ligands modelled in this entry, 3 are monoatomic - leaving 11 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	$\# Z  > 2$
2	NAG	A	1001	1	14,14,15	0.62	0	17,19,21	1.91	5 (29%)
9	CO3	A	1999	8	0,3,3	0.00	-	0,3,3	0.00	-
3	NDG	A	2	1,2	14,14,15	1.29	2 (14%)	17,19,21	1.37	4 (23%)
2	NAG	A	3	3,4	14,14,15	0.95	1 (7%)	17,19,21	1.55	5 (29%)
6	LXC	A	3937	-	10,10,10	1.22	1 (10%)	14,14,14	5.08	13 (92%)
10	SO4	A	3938	-	4,4,4	0.31	0	6,6,6	0.09	0
4	BMA	A	4	2	11,11,12	0.90	1 (9%)	15,15,17	0.87	0
2	NAG	A	5	1,2	14,14,15	0.89	0	17,19,21	0.72	0
2	NAG	A	6	2,5	14,14,15	0.97	0	17,19,21	1.05	1 (5%)
5	MAN	A	7	2,4	11,11,12	1.10	1 (9%)	15,15,17	1.83	3 (20%)
4	BMA	A	8	5	11,11,12	0.75	0	15,15,17	1.37	2 (13%)

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	A	1001	1	-	0/6/23/26	0/1/1/1
9	CO3	A	1999	8	-	0/0/0/0	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NDG	A	2	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	3	3,4	-	0/6/23/26	0/1/1/1
6	LXC	A	3937	-	3/3/4/4	0/0/17/17	0/1/1/1
10	SO4	A	3938	-	-	0/0/0/0	0/0/0/0
4	BMA	A	4	2	-	0/2/19/22	0/1/1/1
2	NAG	A	5	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	6	2,5	-	0/6/23/26	0/1/1/1
5	MAN	A	7	2,4	-	0/2/19/22	0/1/1/1
4	BMA	A	8	5	-	0/2/19/22	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	3937	LXC	O5-C1	-2.22	1.40	1.43
2	A	3	NAG	C1-C2	2.04	1.55	1.52
5	A	7	MAN	C2-C3	2.30	1.55	1.52
4	A	4	BMA	C2-C3	2.41	1.56	1.52
3	A	2	NDG	O-C5	2.56	1.48	1.43

The worst 5 of 33 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1001	NAG	C6-C5-C4	-3.40	104.95	112.99
2	A	1001	NAG	C2-N2-C7	-3.16	118.33	122.94
2	A	3	NAG	C4-C3-C2	-2.95	106.69	111.02
2	A	1001	NAG	O5-C1-C2	-2.87	107.56	111.52
2	A	3	NAG	C2-N2-C7	-2.69	119.02	122.94

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
6	A	3937	LXC	C3
6	A	3937	LXC	C2
6	A	3937	LXC	C1

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
3	A	2	NDG	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	4	BMA	1	0
2	A	5	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

### 6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.