wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 01:40 AM GMT

PDB ID	:	2DXY
Title	:	Structure of the complex of C-terminal lobe of bovine lactoferrin with trehalose at 2.0 A resolution
Authors	:	Mir, R.; Singh, N.; Sinha, M.; Sharma, S.; Bhushan, A.; Singh, T.P.
Deposited on	:	2006-09-03
Resolution	:	2.03 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 (i) symbol.

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

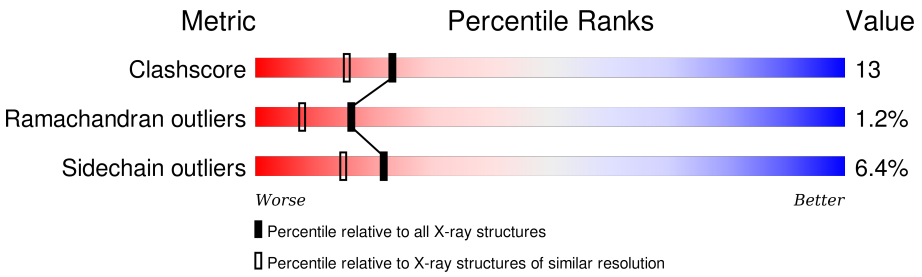
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.03 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	102246	9060 (2.04-2.00)
Ramachandran outliers	100387	8952 (2.04-2.00)
Sidechain outliers	100360	8951 (2.04-2.00)

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 >=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 <=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	77% 19% • •

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
9	TRE	A	1151	X	X	-	-

2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 3034 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.

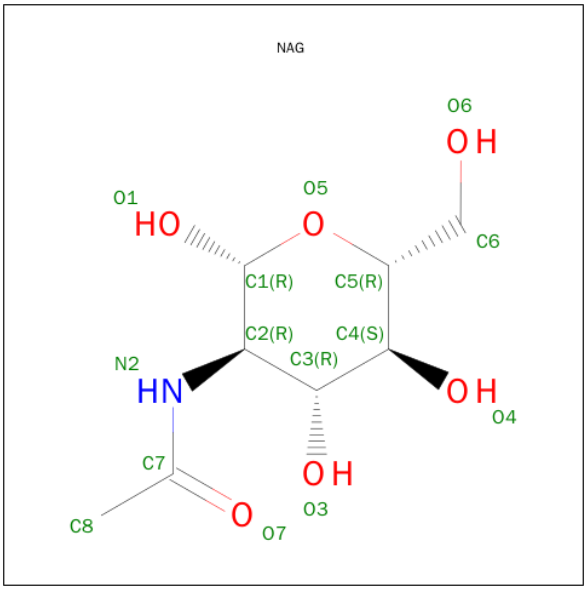
- Molecule 1 is a protein called Lactotransferrin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	341	2605	1622	454	508	21	0	0	0

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



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

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	3	39	22	2	15	0	0

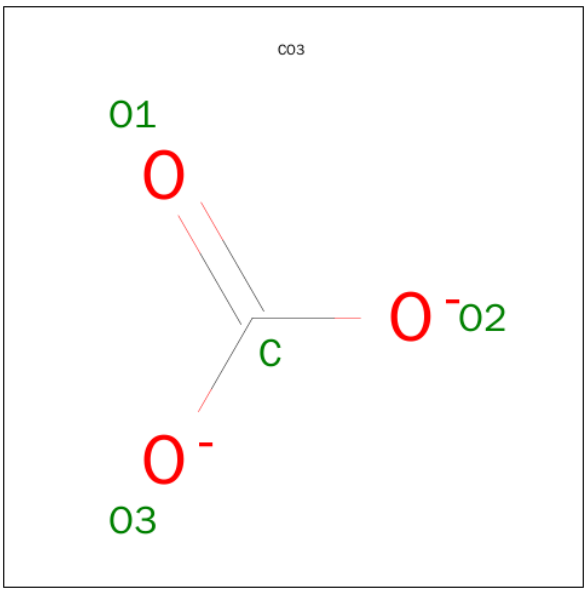
- Molecule 4 is a polymer of unknown type called SUGAR (4-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	A	4	50	28	2	20	0	0

- Molecule 5 is FE (III) ION (three-letter code: FE) (formula: Fe).

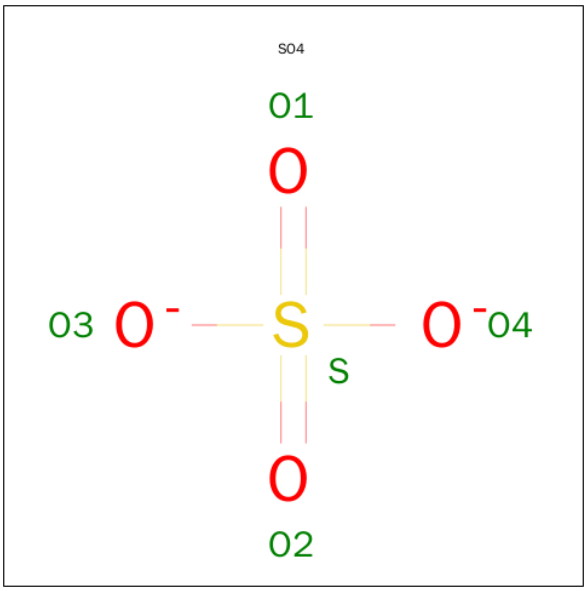
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Fe		
5	A	1	1	1	0	0

- Molecule 6 is CARBONATE ION (three-letter code: CO3) (formula: CO_3).



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

- Molecule 7 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

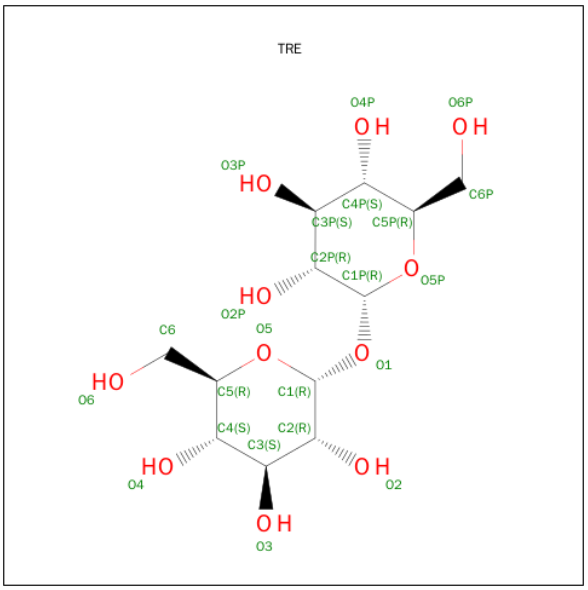


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total O S 5 4 1	0	0

- Molecule 8 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 9 is TREHALOSE (three-letter code: TRE) (formula: $C_{12}H_{22}O_{11}$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
9	A	1	23	12	11	0	0

- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
10	A	291	291	291	0	0

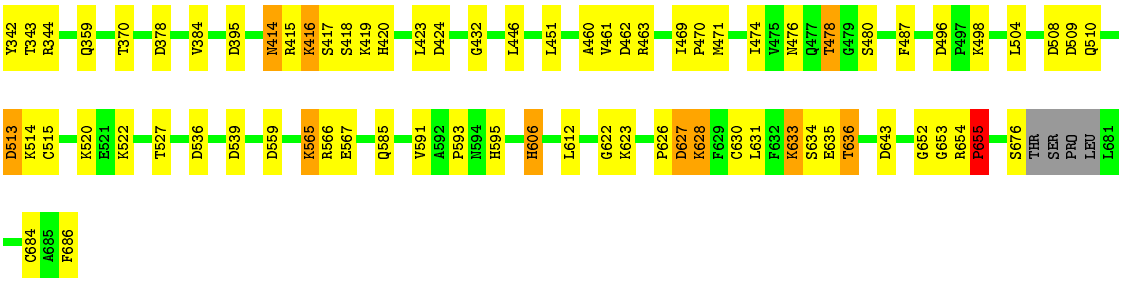
3 Residue-property plots (i)

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.

Note EDS was not executed.

- Molecule 1: Lactotransferrin

Chain A:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	63.37Å 50.41Å 65.91Å 90.00° 107.81° 90.00°	Depositor
Resolution (Å)	63.25 – 2.03	Depositor
% Data completeness (in resolution range)	97.5 (63.25-2.03)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	REFMAC 5	Depositor
R, R_{free}	0.169 , 0.211	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3034	wwPDB-VP
Average B, all atoms ($Å^2$)	37.0	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: CO3, BMA, NAG, TRE, ZN, 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	0.69	1/2653 (0.0%)	0.91	7/3591 (0.2%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	384	VAL	CB-CG1	5.06	1.63	1.52

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	655	PRO	CA-N-CD	-10.05	97.43	111.50
1	A	378	ASP	CB-CG-OD2	6.73	124.36	118.30
1	A	536	ASP	CB-CG-OD2	6.15	123.83	118.30
1	A	424	ASP	CB-CG-OD2	6.12	123.81	118.30
1	A	513	ASP	CB-CG-OD2	5.80	123.52	118.30

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	2605	0	2519	70	0
2	A	14	0	13	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	39	0	34	1	0
4	A	50	0	43	1	0
5	A	1	0	0	0	0
6	A	4	0	0	0	0
7	A	5	0	0	0	0
8	A	2	0	0	0	0
9	A	23	0	6	6	0
10	A	291	0	0	23	0
All	All	3034	0	2615	72	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 13.

The worst 5 of 72 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:416:LYS:HD3	10:A:2136:HOH:O	1.37	1.22
1:A:514:LYS:HD3	10:A:2034:HOH:O	1.44	1.14
1:A:359:GLN:HG2	10:A:2130:HOH:O	1.68	0.92
1:A:416:LYS:HE2	10:A:2065:HOH:O	1.71	0.89
1:A:478:THR:HG21	10:A:2040:HOH:O	1.83	0.79

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%)	319 (95%)	14 (4%)	4 (1%)	16 8

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	622	GLY
1	A	634	SER
1	A	509	ASP
1	A	684	CYS

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	A	282/286 (99%)	264 (94%)	18 (6%)	22 15

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	585	GLN
1	A	606	HIS
1	A	636	THR
1	A	559	ASP
1	A	565	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	414	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)

7 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	A	2	1,3	14,14,15	0.75	1 (7%)	15,19,21	1.61	2 (13%)
3	NAG	A	3	3	14,14,15	0.59	0	15,19,21	1.30	2 (13%)
3	BMA	A	4	3	11,11,12	0.61	0	14,15,17	1.37	3 (21%)
4	NAG	A	5	1,4	14,14,15	1.08	1 (7%)	15,19,21	1.53	2 (13%)
4	NAG	A	6	4	14,14,15	0.68	0	15,19,21	1.36	2 (13%)
4	MAN	A	7	4	11,11,12	0.75	0	14,15,17	2.43	6 (42%)
4	BMA	A	8	4	11,11,12	0.61	0	14,15,17	1.84	3 (21%)

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	A	2	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	3	3	-	0/6/23/26	0/1/1/1
3	BMA	A	4	3	-	0/2/19/22	0/1/1/1
4	NAG	A	5	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	6	4	-	0/6/23/26	0/1/1/1
4	MAN	A	7	4	-	0/2/19/22	0/1/1/1
4	BMA	A	8	4	-	0/2/19/22	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	2	NAG	C1-C2	2.11	1.55	1.52
4	A	5	NAG	C1-C2	3.40	1.57	1.52

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	6	NAG	C2-N2-C7	-3.38	118.70	123.04
3	A	2	NAG	C3-C4-C5	-3.36	104.34	110.20
4	A	5	NAG	C2-N2-C7	-3.32	118.78	123.04
4	A	8	BMA	O5-C1-C2	-2.93	106.11	110.86
4	A	6	NAG	C3-C4-C5	-2.89	105.16	110.20

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	2	NAG	1	0
3	A	3	NAG	1	0
4	A	8	BMA	1	0

5.6 Ligand geometry (i)

Of 7 ligands modelled in this entry, 3 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	$\#\|Z\| > 2$
2	NAG	A	1001	1	14,14,15	0.84	1 (7%)	15,19,21	1.00	0
9	TRE	A	1151	-	24,24,24	6.13	14 (58%)	35,35,35	4.31	30 (85%)
6	CO3	A	1999	5	0,3,3	0.00	-	0,3,3	0.00	-
7	SO4	A	2000	-	4,4,4	0.40	0	6,6,6	0.19	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
2	NAG	A	1001	1	-	0/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
9	TRE	A	1151	-	10/10/10/10	0/8/48/48	0/2/2/2
6	CO3	A	1999	5	-	0/0/0/0	0/0/0/0
7	SO4	A	2000	-	-	0/0/0/0	0/0/0/0

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	A	1151	TRE	C4P-C5P	-9.31	1.33	1.53
9	A	1151	TRE	C4-C5	-9.27	1.33	1.53
9	A	1151	TRE	O4P-C4P	-8.71	1.22	1.43
9	A	1151	TRE	O3-C3	-8.57	1.22	1.43
9	A	1151	TRE	O2-C2	-8.56	1.22	1.43

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
9	A	1151	TRE	C1P-O1-C1	2.39	118.14	114.39
9	A	1151	TRE	C6-C5-C4	2.49	119.15	113.02
9	A	1151	TRE	O4-C4-C5	3.52	118.56	109.24
9	A	1151	TRE	C6P-C5P-C4P	3.52	121.70	113.02
9	A	1151	TRE	C1P-C2P-C3P	3.61	117.09	109.97

5 of 10 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
9	A	1151	TRE	C5P
9	A	1151	TRE	C3P
9	A	1151	TRE	C4
9	A	1151	TRE	C5
9	A	1151	TRE	C2

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	A	1151	TRE	6	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)

EDS was not executed - this section will therefore be empty.

6.2 Non-standard residues in protein, DNA, RNA chains (i)

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates (i)

EDS was not executed - this section will therefore be empty.

6.4 Ligands (i)

EDS was not executed - this section will therefore be empty.

6.5 Other polymers (i)

EDS was not executed - this section will therefore be empty.