



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 13, 2017 – 04:08 am GMT

PDB ID : 2SIM  
Title : THE STRUCTURES OF SALMONELLA TYPHIMURIUM LT2 NEURAMINIDASE AND ITS COMPLEX WITH A TRANSITION STATE ANALOGUE AT 1.6 ANGSTROMS RESOLUTION  
Authors : Taylor, G.L.; Crennell, S.J.; Garman, E.F.; Vimr, E.R.; Laver, W.G.  
Deposited on : 1994-07-15  
Resolution : 1.60 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

<http://wwpdb.org/validation/2016/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	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

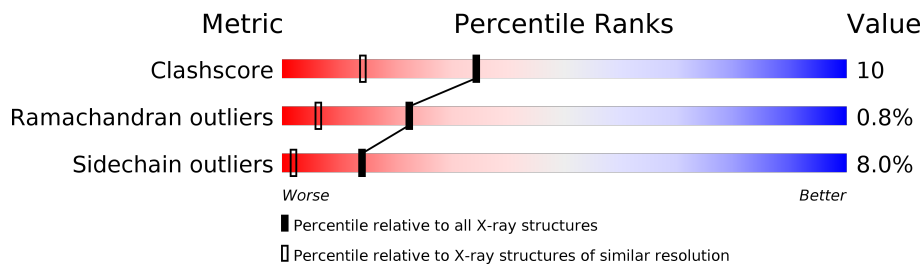
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.60 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	112137	2967 (1.60-1.60)
Ramachandran outliers	110173	2887 (1.60-1.60)
Sidechain outliers	110143	2886 (1.60-1.60)

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	381	 77% 17% . .

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 3216 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.

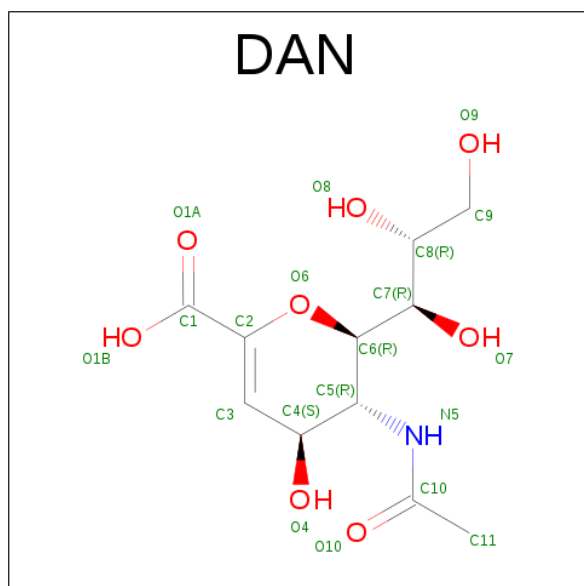
- Molecule 1 is a protein called SIALIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	381	Total	C	N	O	S	0	0	0
			2954	1847	513	584	10			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	329	ASP	ALA	CONFLICT	UNP P29768

- Molecule 2 is 2-DEOXY-2,3-DEHYDRO-N-ACETYL-NEURAMINIC ACID (three-letter code: DAN) (formula:  $C_{11}H_{17}NO_8$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			20	11	1	8		

- Molecule 3 is water.

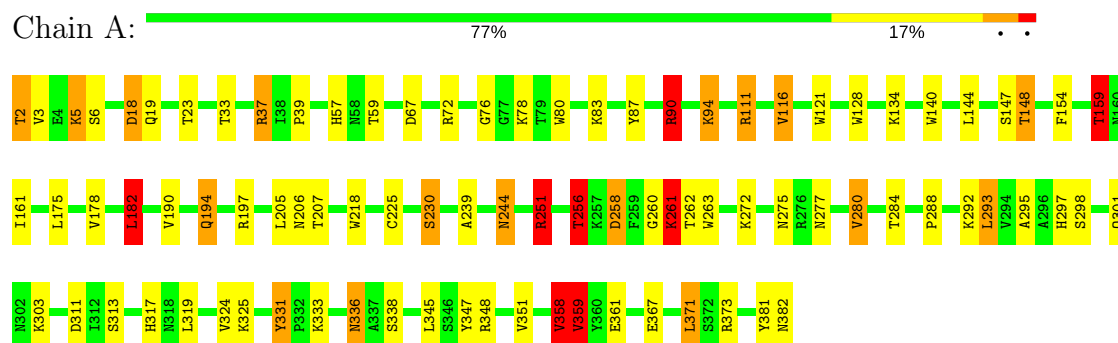
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	242	Total 242	O 242	0	0

### 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: SIALIDASE



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	47.60Å 82.50Å 91.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	6.00 – 1.60	Depositor
% Data completeness (in resolution range)	(Not available) (6.00-1.60)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	TNT	Depositor
R, $R_{free}$	0.185 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3216	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	13.0	wwPDB-VP

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: DAN

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	0.80	0/3012	1.84	57/4079 (1.4%)

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	A	0	3

There are no bond length outliers.

All (57) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	251	ARG	NE-CZ-NH2	-29.88	105.36	120.30
1	A	90	ARG	NE-CZ-NH1	29.40	135.00	120.30
1	A	251	ARG	NE-CZ-NH1	27.02	133.81	120.30
1	A	90	ARG	NE-CZ-NH2	-25.72	107.44	120.30
1	A	37	ARG	NE-CZ-NH2	-21.39	109.61	120.30
1	A	37	ARG	NE-CZ-NH1	20.65	130.62	120.30
1	A	72	ARG	NE-CZ-NH1	18.39	129.50	120.30
1	A	72	ARG	NE-CZ-NH2	-17.92	111.34	120.30
1	A	373	ARG	NE-CZ-NH1	10.24	125.42	120.30
1	A	373	ARG	NE-CZ-NH2	-9.99	115.30	120.30
1	A	111	ARG	NE-CZ-NH1	9.75	125.17	120.30
1	A	358	VAL	CB-CA-C	-9.20	93.93	111.40
1	A	159	THR	N-CA-CB	-9.10	93.02	110.30
1	A	90	ARG	CG-CD-NE	-8.79	93.34	111.80
1	A	280	VAL	CB-CA-C	-8.17	95.88	111.40
1	A	251	ARG	CD-NE-CZ	8.02	134.83	123.60

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	182	LEU	CA-CB-CG	7.97	133.62	115.30
1	A	218	TRP	CD1-CG-CD2	7.89	112.61	106.30
1	A	128	TRP	CD1-CG-CD2	7.88	112.61	106.30
1	A	263	TRP	CD1-CG-CD2	7.86	112.59	106.30
1	A	359	VAL	CG1-CB-CG2	7.84	123.45	110.90
1	A	111	ARG	NE-CZ-NH2	-7.61	116.50	120.30
1	A	218	TRP	CE2-CD2-CG	-7.59	101.22	107.30
1	A	263	TRP	CE2-CD2-CG	-7.47	101.33	107.30
1	A	121	TRP	CD1-CG-CD2	7.45	112.26	106.30
1	A	37	ARG	CD-NE-CZ	7.27	133.77	123.60
1	A	140	TRP	CD1-CG-CD2	7.27	112.11	106.30
1	A	331	TYR	CB-CG-CD2	-6.66	117.00	121.00
1	A	358	VAL	CG1-CB-CG2	6.62	121.50	110.90
1	A	128	TRP	CE2-CD2-CG	-6.53	102.08	107.30
1	A	80	TRP	CD1-CG-CD2	6.41	111.43	106.30
1	A	280	VAL	CG1-CB-CG2	6.38	121.12	110.90
1	A	358	VAL	N-CA-CB	6.34	125.44	111.50
1	A	80	TRP	CE2-CD2-CG	-6.24	102.31	107.30
1	A	348	ARG	NE-CZ-NH2	-6.24	117.18	120.30
1	A	121	TRP	CE2-CD2-CG	-6.23	102.31	107.30
1	A	347	TYR	CB-CG-CD2	-6.07	117.36	121.00
1	A	256	THR	CA-CB-CG2	-6.00	104.00	112.40
1	A	18	ASP	CB-CG-OD1	5.92	123.63	118.30
1	A	116	VAL	CB-CA-C	-5.91	100.17	111.40
1	A	256	THR	N-CA-CB	5.87	121.45	110.30
1	A	205	LEU	CA-CB-CG	5.85	128.75	115.30
1	A	2	THR	N-CA-CB	-5.82	99.24	110.30
1	A	251	ARG	CG-CD-NE	-5.81	99.59	111.80
1	A	90	ARG	CD-NE-CZ	5.81	131.73	123.60
1	A	371	LEU	CB-CG-CD1	-5.66	101.37	111.00
1	A	261	LYS	CB-CG-CD	5.60	126.15	111.60
1	A	256	THR	CB-CA-C	-5.36	97.14	111.60
1	A	280	VAL	N-CA-CB	5.36	123.28	111.50
1	A	140	TRP	CE2-CD2-CG	-5.35	103.02	107.30
1	A	218	TRP	CG-CD2-CE3	5.34	138.70	133.90
1	A	148	THR	N-CA-CB	-5.32	100.19	110.30
1	A	128	TRP	CG-CD1-NE1	-5.32	104.78	110.10
1	A	121	TRP	CB-CG-CD1	-5.15	120.31	127.00
1	A	116	VAL	CG1-CB-CG2	5.13	119.11	110.90
1	A	258	ASP	CB-CG-OD1	5.01	122.81	118.30
1	A	19	GLN	CA-CB-CG	-5.01	102.37	113.40

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	239	ALA	Peptide
1	A	251	ARG	Sidechain
1	A	90	ARG	Sidechain

## 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	2954	0	2883	56	0
2	A	20	0	16	0	0
3	A	242	0	0	8	0
All	All	3216	0	2899	56	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 10.

All (56) 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:256:THR:HG21	1:A:260:GLY:H	1.39	0.88
1:A:2:THR:HB	1:A:76:GLY:HA3	1.62	0.81
1:A:39:PRO:HB2	1:A:359:VAL:HG13	1.65	0.78
1:A:272:LYS:HB3	1:A:324:VAL:HG21	1.65	0.76
1:A:345:LEU:HD23	1:A:358:VAL:HG13	1.70	0.73
1:A:5:LYS:HZ1	1:A:367:GLU:HB3	1.55	0.70
1:A:159:THR:HG23	1:A:161:ILE:H	1.55	0.70
1:A:159:THR:HG21	3:A:626:HOH:O	1.91	0.70
1:A:256:THR:HG21	1:A:260:GLY:N	2.11	0.65
1:A:5:LYS:HE3	1:A:6:SER:N	2.12	0.65
1:A:37:ARG:HD2	1:A:361:GLU:OE2	1.99	0.63
1:A:23:THR:HG22	3:A:597:HOH:O	1.99	0.61
1:A:324:VAL:HG22	3:A:555:HOH:O	2.00	0.61
1:A:244:ASN:HD21	1:A:251:ARG:HD3	1.67	0.60
1:A:371:LEU:HA	3:A:654:HOH:O	2.02	0.58
1:A:256:THR:HG23	1:A:262:THR:O	2.06	0.56

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:197:ARG:HH11	1:A:206:ASN:ND2	2.05	0.55
1:A:2:THR:HG22	1:A:78:LYS:NZ	2.22	0.55
1:A:317:HIS:HD2	3:A:555:HOH:O	1.89	0.55
1:A:336:ASN:HD22	1:A:338:SER:H	1.55	0.54
1:A:3:VAL:HG11	1:A:6:SER:HB3	1.88	0.54
1:A:244:ASN:HD22	1:A:251:ARG:HH11	1.54	0.53
1:A:2:THR:HB	1:A:76:GLY:CA	2.34	0.53
1:A:111:ARG:HG3	1:A:111:ARG:HH11	1.73	0.53
1:A:207:THR:OG1	1:A:225:CYS:HB3	2.09	0.53
1:A:284:THR:OG1	1:A:297:HIS:HD2	1.92	0.52
1:A:244:ASN:ND2	1:A:251:ARG:HH11	2.09	0.51
1:A:2:THR:HG22	1:A:78:LYS:HZ1	1.75	0.51
1:A:295:ALA:HB3	1:A:317:HIS:HB2	1.93	0.50
1:A:331:TYR:CE1	1:A:333:LYS:HD2	2.46	0.50
1:A:159:THR:HG22	3:A:631:HOH:O	2.10	0.50
1:A:301:GLN:HB2	1:A:311:ASP:HB2	1.95	0.49
1:A:194:GLN:HB3	1:A:230:SER:HB2	1.93	0.49
1:A:358:VAL:HG22	1:A:371:LEU:HD11	1.94	0.48
1:A:33:THR:OG1	1:A:57:HIS:HD2	1.96	0.48
1:A:358:VAL:HG22	1:A:371:LEU:CD1	2.45	0.46
1:A:182:LEU:HD13	1:A:190:VAL:HB	1.98	0.45
1:A:57:HIS:HE1	1:A:67:ASP:OD2	2.00	0.45
1:A:288:PRO:HA	1:A:293:LEU:HD12	2.00	0.44
1:A:3:VAL:HB	1:A:78:LYS:HG2	2.00	0.43
1:A:18:ASP:HB2	3:A:596:HOH:O	2.19	0.43
1:A:39:PRO:HB2	1:A:359:VAL:CG1	2.40	0.43
1:A:94:LYS:HE3	3:A:668:HOH:O	2.17	0.43
1:A:298:SER:HA	1:A:313:SER:O	2.19	0.43
1:A:325:LYS:HA	1:A:325:LYS:HD2	1.82	0.43
1:A:5:LYS:NZ	1:A:367:GLU:HB3	2.28	0.42
1:A:345:LEU:CD2	1:A:358:VAL:HG13	2.43	0.42
1:A:336:ASN:ND2	1:A:338:SER:H	2.15	0.42
1:A:18:ASP:HA	1:A:87:TYR:CE2	2.55	0.42
1:A:258:ASP:O	1:A:261:LYS:HD2	2.20	0.42
1:A:301:GLN:OE1	1:A:303:LYS:HE3	2.20	0.41
1:A:251:ARG:HH21	1:A:275:ASN:ND2	2.18	0.41
1:A:5:LYS:HG2	1:A:5:LYS:HZ2	1.68	0.41
1:A:359:VAL:HA	1:A:367:GLU:O	2.21	0.41
1:A:251:ARG:HE	1:A:275:ASN:HD21	1.68	0.40
1:A:147:SER:HB2	1:A:154:PHE:CE1	2.56	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	A	379/381 (100%)	364 (96%)	12 (3%)	3 (1%)	22 6

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	230	SER
1	A	178	VAL
1	A	351	VAL

### 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	327/327 (100%)	301 (92%)	26 (8%)	14 2

All (26) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	5	LYS
1	A	59	THR
1	A	83	LYS
1	A	90	ARG
1	A	94	LYS
1	A	116	VAL
1	A	134	LYS
1	A	144	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	148	THR
1	A	159	THR
1	A	175	LEU
1	A	182	LEU
1	A	194	GLN
1	A	244	ASN
1	A	256	THR
1	A	261	LYS
1	A	277	ASN
1	A	280	VAL
1	A	292	LYS
1	A	293	LEU
1	A	319	LEU
1	A	336	ASN
1	A	358	VAL
1	A	359	VAL
1	A	381	TYR
1	A	382	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	57	HIS
1	A	63	GLN
1	A	185	ASN
1	A	194	GLN
1	A	206	ASN
1	A	232	ASN
1	A	244	ASN
1	A	275	ASN
1	A	278	HIS
1	A	297	HIS
1	A	317	HIS
1	A	336	ASN
1	A	382	ASN

### 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](#)

1 ligand is 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	DAN	A	800	-	17,20,20	1.81	1 (5%)	18,28,28	1.32	2 (11%)

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	DAN	A	800	-	-	0/14/34/34	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	800	DAN	C3-C2	6.42	1.40	1.32

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	800	DAN	O6-C2-C3	-2.88	120.28	124.31
2	A	800	DAN	C8-C7-C6	-2.48	108.22	113.04

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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

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.