



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 7, 2018 – 03:58 pm GMT

PDB ID : 1M6B  
Title : Structure of the HER3 (ERBB3) Extracellular Domain  
Authors : Leahy, D.J.; Cho, H.-S.  
Deposited on : 2002-07-15  
Resolution : 2.60 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) : 1.13  
EDS : trunk30967  
Percentile statistics : (not set)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30967

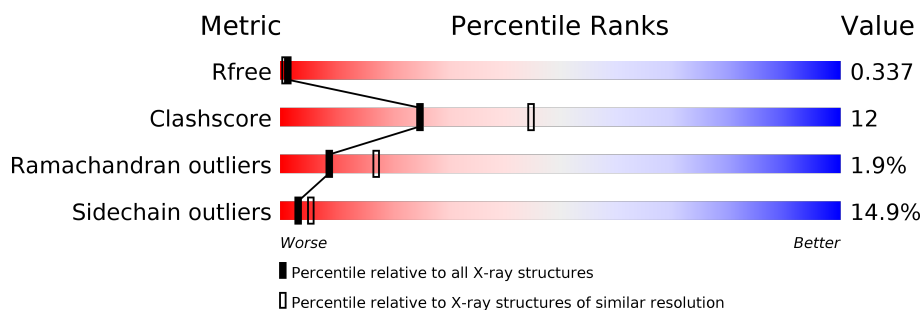
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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))
$R_{free}$	111664	2767 (2.60-2.60)
Clashscore	122126	3110 (2.60-2.60)
Ramachandran outliers	120053	3062 (2.60-2.60)
Sidechain outliers	120020	3062 (2.60-2.60)

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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	621	58% 25% 5% 12%
1	B	621	64% 24% 5% • 6%

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
2	NAG	A	626	X	-	-	-
2	NAG	A	628	X	-	-	-
2	NAG	B	627	X	-	-	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	NAG	B	628	X	-	-	-
4	SO4	B	5001	-	-	X	-

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 9012 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 Receptor protein-tyrosine kinase erbB-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	549	Total	C	N	O	S	0	0	0
			4206	2603	765	784	54			
1	B	584	Total	C	N	O	S	0	0	0
			4478	2766	813	840	59			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	61	VAL	ILE	ENGINEERED	UNP P21860
B	61	VAL	ILE	ENGINEERED	UNP P21860

- Molecule 2 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



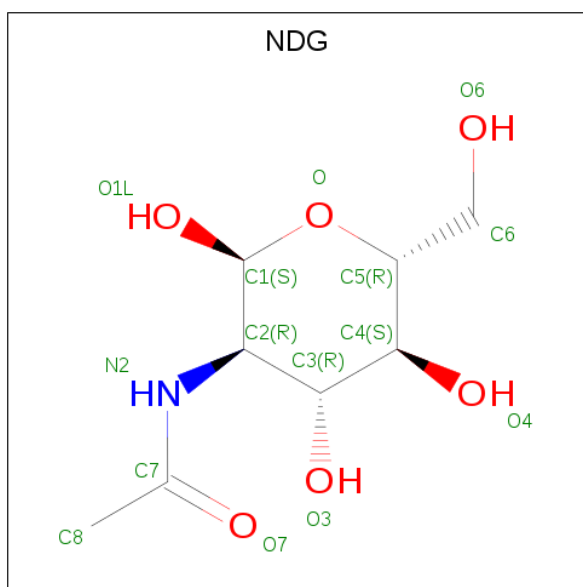
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		

*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		
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		
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		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	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<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



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

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



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

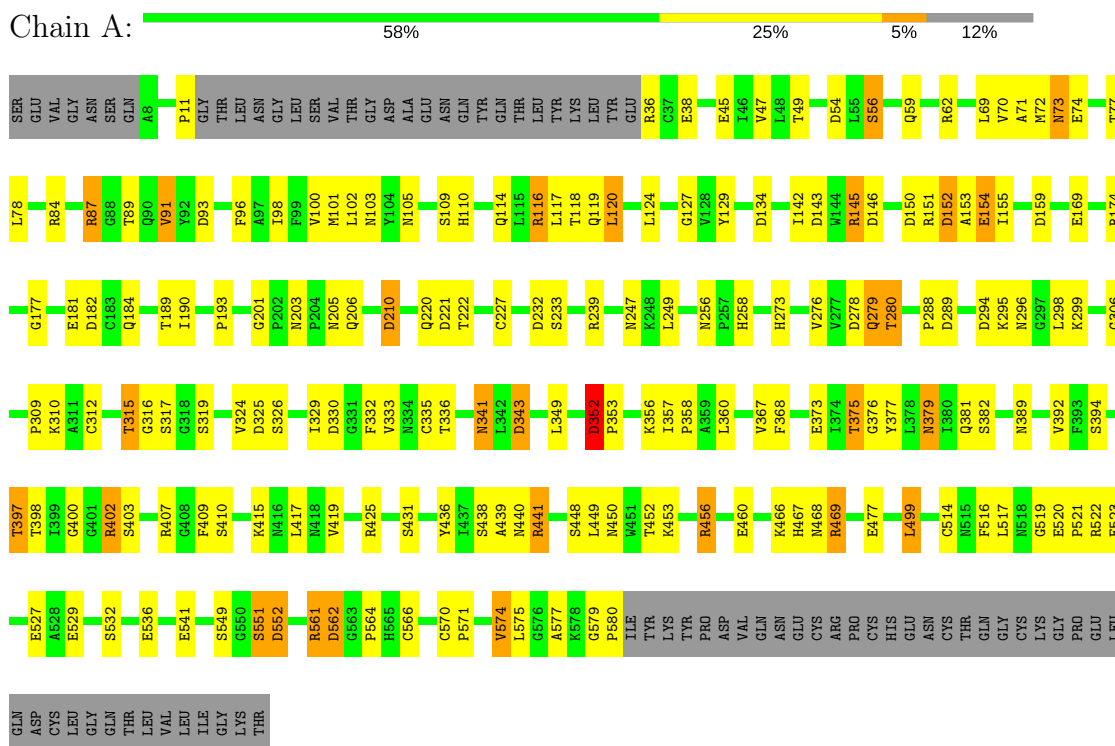
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	20	Total 20	O 20	0	0
5	B	65	Total 65	O 65	0	0

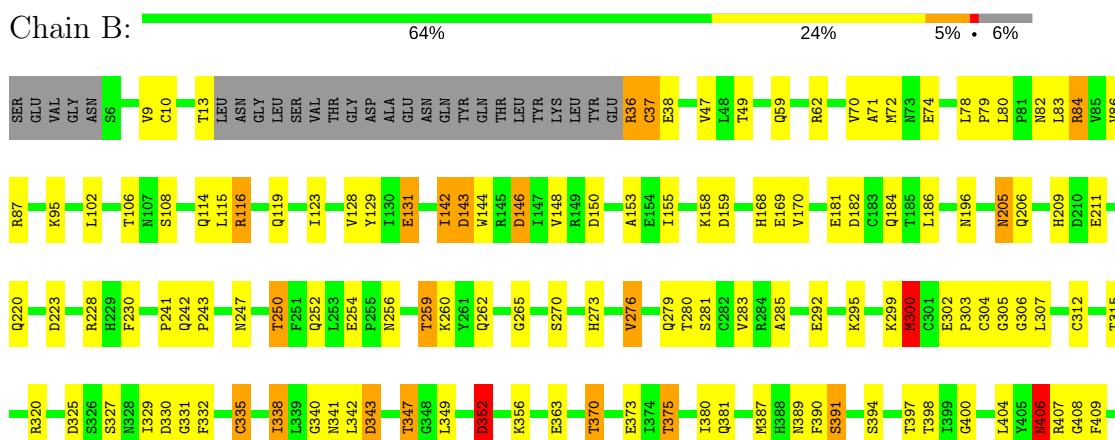
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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. 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: Receptor protein-tyrosine kinase erbB-3



- Molecule 1: Receptor protein-tyrosine kinase erbB-3





M414	K415	T420	R425	S426	S431	R434	I435	Y436	I437	S438	R441	H446	L449	T452	K453	T459	R462	K466	E477	G478	K479	W480	C481	C490	L499	S500	C501	R502	M503	Y504	S505	R506	N515	F516	L517	N518	G519	E520	P521	R522	E523	H534	P535
E536	C537	Q538	G542	T543	Q556	H559	F560	R561	D562	H565	G576	A577	K578	I581	Y582	K583	D586	N589	E590	H595	E596	N597	Q600	K603	E606	D609	C610	L611	GLY	GLN	THR	LEU	VAL	LEU	ILE	GLY	LYS	THR					

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	236.26Å 49.62Å 190.86Å 90.00° 125.56° 90.00°	Depositor
Resolution (Å)	20.00 – 2.60 28.74 – 2.60	Depositor EDS
% Data completeness (in resolution range)	94.5 (20.00-2.60) 90.7 (28.74-2.60)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.97 (at 2.61Å)	Xtriage
Refinement program	REFMAC 5.1.19	Depositor
R, $R_{free}$	0.235 , 0.294 0.299 , 0.337	Depositor DCC
$R_{free}$ test set	2754 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	48.1	Xtriage
Anisotropy	0.756	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 30.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.000 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.87	EDS
Total number of atoms	9012	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.52% 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, NDG, SO4

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	0.56	0/4308	0.84	15/5845 (0.3%)
1	B	0.62	1/4586 (0.0%)	0.86	8/6222 (0.1%)
All	All	0.59	1/8894 (0.0%)	0.85	23/12067 (0.2%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	300	MET	SD-CE	5.39	2.08	1.77

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	352	ASP	CB-CG-OD2	8.78	126.20	118.30
1	A	221	ASP	CB-CG-OD2	6.58	124.22	118.30
1	B	146	ASP	CB-CG-OD2	6.36	124.02	118.30
1	B	143	ASP	CB-CG-OD2	6.24	123.91	118.30
1	B	325	ASP	CB-CG-OD2	6.20	123.88	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	4206	0	3996	79	0
1	B	4478	0	4242	123	0
2	A	126	0	116	1	0
2	B	98	0	90	3	0
3	B	14	0	13	0	0
4	B	5	0	0	4	0
5	A	20	0	0	3	0
5	B	65	0	0	7	0
All	All	9012	0	8457	202	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 12.

The worst 5 of 202 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:B:300:MET:SD	1:B:300:MET:CE	2.08	1.42
1:B:116:ARG:HG3	1:B:116:ARG:HH11	1.11	1.11
1:B:335:CYS:O	1:B:370:THR:HG22	1.56	1.03
1:B:452:THR:HG22	1:B:459:THR:HG21	1.42	1.00
1:B:578:LYS:H	1:B:578:LYS:HE2	1.31	0.96

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	545/621 (88%)	482 (88%)	52 (10%)	11 (2%)	8	16
1	B	580/621 (93%)	525 (90%)	45 (8%)	10 (2%)	10	20
All	All	1125/1242 (91%)	1007 (90%)	97 (9%)	21 (2%)	9	17

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	317	SER
1	A	460	GLU
1	B	520	GLU
1	B	586	ASP
1	A	296	ASN

### 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	474/537 (88%)	398 (84%)	76 (16%)	2	4
1	B	506/537 (94%)	436 (86%)	70 (14%)	4	6
All	All	980/1074 (91%)	834 (85%)	146 (15%)	3	5

5 of 146 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	529	GLU
1	B	38	GLU
1	B	520	GLU
1	A	532	SER
1	A	570	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	468	ASN
1	A	565	HIS
1	B	229	HIS
1	A	467	HIS
1	B	467	HIS

### 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](#)

18 ligands 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$
2	NAG	A	622	1	14,14,15	0.42	0	17,19,21	2.21	2 (11%)
2	NAG	A	623	1	14,14,15	0.69	0	17,19,21	1.37	3 (17%)
2	NAG	A	624	1,2	14,14,15	0.82	0	17,19,21	1.66	3 (17%)
2	NAG	A	625	2	14,14,15	0.42	0	17,19,21	1.60	4 (23%)
2	NAG	A	626	1	14,14,15	0.61	0	17,19,21	1.22	2 (11%)
2	NAG	A	627	1	14,14,15	0.62	0	17,19,21	1.59	4 (23%)
2	NAG	A	628	1	14,14,15	0.73	0	17,19,21	1.92	4 (23%)
2	NAG	A	629	1	14,14,15	0.81	1 (7%)	17,19,21	1.61	4 (23%)
2	NAG	A	630	1	14,14,15	0.67	0	17,19,21	1.66	3 (17%)
4	SO4	B	5001	-	4,4,4	0.23	0	6,6,6	0.29	0
2	NAG	B	622	1	14,14,15	0.67	0	17,19,21	2.99	6 (35%)
2	NAG	B	623	1	14,14,15	0.70	0	17,19,21	1.64	4 (23%)
2	NAG	B	624	1,2	14,14,15	0.44	0	17,19,21	1.28	2 (11%)
2	NAG	B	625	2	14,14,15	0.56	0	17,19,21	0.90	1 (5%)
3	NDG	B	626	1	14,14,15	0.69	0	17,19,21	1.92	6 (35%)
2	NAG	B	627	1	14,14,15	0.56	0	17,19,21	1.63	2 (11%)
2	NAG	B	628	1	14,14,15	0.68	0	17,19,21	1.63	2 (11%)
2	NAG	B	629	1	14,14,15	1.05	1 (7%)	17,19,21	1.54	3 (17%)

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	622	1	-	0/6/23/26	0/1/1/1
2	NAG	A	623	1	-	0/6/23/26	0/1/1/1
2	NAG	A	624	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	625	2	-	0/6/23/26	0/1/1/1
2	NAG	A	626	1	1/1/5/7	0/6/23/26	0/1/1/1
2	NAG	A	627	1	-	0/6/23/26	0/1/1/1
2	NAG	A	628	1	1/1/5/7	0/6/23/26	0/1/1/1
2	NAG	A	629	1	-	0/6/23/26	0/1/1/1
2	NAG	A	630	1	-	0/6/23/26	0/1/1/1
4	SO4	B	5001	-	-	0/0/0/0	0/0/0/0
2	NAG	B	622	1	-	0/6/23/26	0/1/1/1
2	NAG	B	623	1	-	0/6/23/26	0/1/1/1
2	NAG	B	624	1,2	-	0/6/23/26	0/1/1/1
2	NAG	B	625	2	-	0/6/23/26	0/1/1/1
3	NDG	B	626	1	-	0/6/23/26	0/1/1/1
2	NAG	B	627	1	1/1/5/7	0/6/23/26	0/1/1/1
2	NAG	B	628	1	1/1/5/7	0/6/23/26	0/1/1/1
2	NAG	B	629	1	-	1/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	629	NAG	C1-C2	2.56	1.56	1.52
2	B	629	NAG	C1-C2	3.16	1.56	1.52

The worst 5 of 55 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	622	NAG	C2-N2-C7	-8.07	111.17	122.94
2	A	628	NAG	O5-C1-C2	-5.87	103.42	111.52
2	B	628	NAG	C2-N2-C7	-4.82	115.91	122.94
2	B	627	NAG	O5-C1-C2	-4.55	105.23	111.52
2	B	622	NAG	O7-C7-N2	-3.33	115.66	121.94

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	A	626	NAG	C1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atom
2	B	627	NAG	C1
2	B	628	NAG	C1
2	A	628	NAG	C1

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	629	NAG	O7-C7-N2-C2

There are no ring outliers.

4 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	624	NAG	1	0
4	B	5001	SO4	4	0
2	B	624	NAG	1	0
2	B	627	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

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.