



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 6, 2026 – 12:26 PM UTC

PDB ID : 2EXI / pdb\_00002exi  
Title : Structure of the family43 beta-Xylosidase D15G mutant from geobacillus stearothermophilus  
Authors : Brux, C.; Niefind, K.; Shallom-Shezifi, D.; Shoham, Y.; Schomburg, D.  
Deposited on : 2005-11-08  
Resolution : 2.15 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 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 : 4-5-2 with Phenix2.0  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 2.0  
EDS : 3.0  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.49

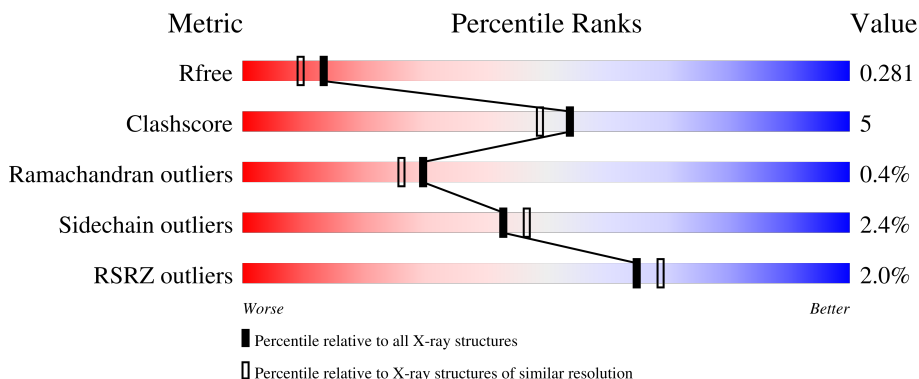
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.15 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))
$R_{free}$	180053	2057 (2.16-2.16)
Clashscore	190562	2159 (2.16-2.16)
Ramachandran outliers	187476	2134 (2.16-2.16)
Sidechain outliers	187428	2133 (2.16-2.16)
RSRZ outliers	180081	2059 (2.16-2.16)

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	A	535	
1	B	535	
1	C	535	
1	D	535	

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 19639 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 beta-D-xylosidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	533	4372	2812	748	802	10	0	0	0
1	B	533	4372	2812	748	802	10	0	0	0
1	C	533	4372	2812	748	802	10	0	0	0
1	D	533	4372	2812	748	802	10	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	15	GLY	ASP	engineered mutation	UNP Q68HB3
B	15	GLY	ASP	engineered mutation	UNP Q68HB3
C	15	GLY	ASP	engineered mutation	UNP Q68HB3
D	15	GLY	ASP	engineered mutation	UNP Q68HB3

- Molecule 2 is CALCIUM ION (CCD ID: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	Ca 1	0	0
2	B	1	Total 1	Ca 1	0	0
2	C	1	Total 1	Ca 1	0	0
2	D	1	Total 1	Ca 1	0	0

- Molecule 3 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (CCD ID: MES) (formula: C<sub>6</sub>H<sub>13</sub>NO<sub>4</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
3	A	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
3	B	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
3	D	1	Total	C	N	O	S	0	0
			12	6	1	4	1		

- Molecule 4 is GLYCEROL (CCD ID: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



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

*Continued on next page...*

*Continued from previous page...*

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

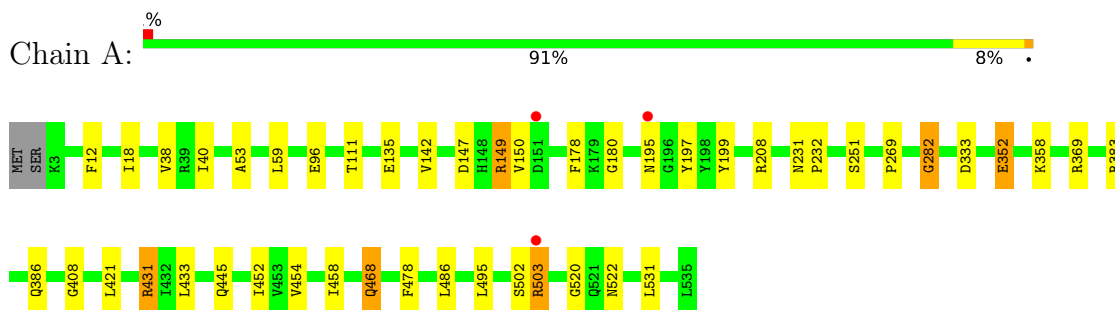
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	557	Total	O	0	0
			557	557		
5	B	584	Total	O	0	0
			584	584		
5	C	403	Total	O	0	0
			403	403		
5	D	543	Total	O	0	0
			543	543		

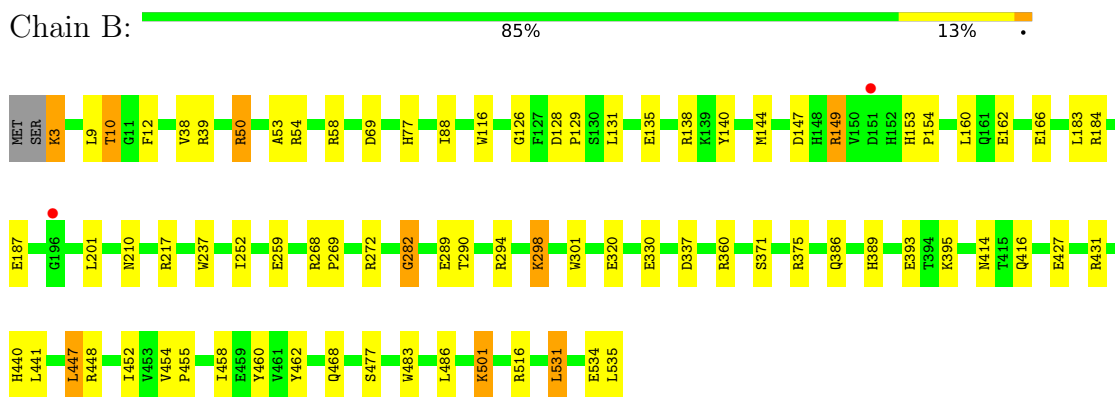
### 3 Residue-property plots [i](#)

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.

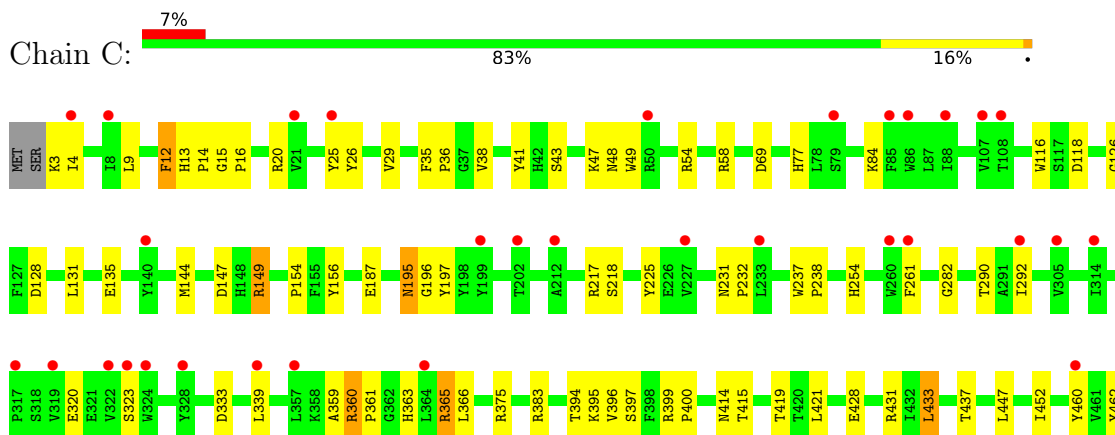
- Molecule 1: beta-D-xylosidase



- Molecule 1: beta-D-xylosidase



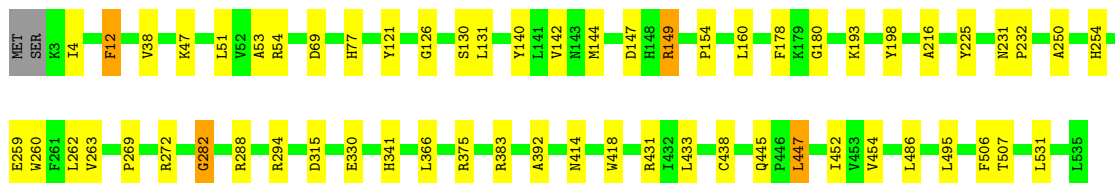
- Molecule 1: beta-D-xylosidase





- Molecule 1: beta-D-xylosidase

Chain D: 88% 10%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	140.18Å 140.18Å 232.32Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.15 20.00 – 2.15	Depositor EDS
% Data completeness (in resolution range)	99.5 (20.00-2.15) 99.7 (20.00-2.15)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.73 (at 2.15Å)	Xtrriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.199 , 0.276 0.200 , 0.281	Depositor DCC
$R_{free}$ test set	6323 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	28.0	Xtrriage
Anisotropy	0.274	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 50.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	19639	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.36% 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: CA, MES, GOL

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	0.53	0/4516	0.75	4/6154 (0.1%)
1	B	0.55	0/4516	0.78	1/6154 (0.0%)
1	C	0.52	0/4516	0.76	4/6154 (0.1%)
1	D	0.55	0/4516	0.76	1/6154 (0.0%)
All	All	0.54	0/18064	0.76	10/24616 (0.0%)

There are no bond length outliers.

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	282	GLY	N-CA-C	7.43	120.76	111.85
1	A	282	GLY	N-CA-C	6.69	118.94	111.85
1	B	282	GLY	N-CA-C	5.76	117.96	111.85
1	C	54	ARG	CA-C-N	5.41	124.60	118.97
1	C	54	ARG	C-N-CA	5.41	124.60	118.97

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	4372	0	4162	24	0
1	B	4372	0	4162	54	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	4372	0	4162	55	0
1	D	4372	0	4162	37	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	12	0	12	0	0
3	B	12	0	12	3	0
3	D	12	0	12	4	0
4	A	6	0	8	0	0
4	B	6	0	8	0	0
4	C	6	0	8	0	0
4	D	6	0	8	0	0
5	A	557	0	0	5	0
5	B	584	0	0	6	0
5	C	403	0	0	13	0
5	D	543	0	0	2	0
All	All	19639	0	16716	167	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.

The worst 5 of 167 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:50:ARG:HG2	1:B:50:ARG:HH11	1.22	1.02
1:B:50:ARG:HH11	1:B:50:ARG:CG	1.87	0.87
1:B:54:ARG:HE	3:B:3006:MES:H82	1.43	0.84
1:C:197:TYR:HB3	5:C:3298:HOH:O	1.79	0.81
1:B:3:LYS:HE3	5:B:3073:HOH:O	1.81	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	531/535 (99%)	501 (94%)	28 (5%)	2 (0%)	30	26
1	B	531/535 (99%)	505 (95%)	24 (4%)	2 (0%)	30	26
1	C	531/535 (99%)	492 (93%)	37 (7%)	2 (0%)	30	26
1	D	531/535 (99%)	504 (95%)	25 (5%)	2 (0%)	30	26
All	All	2124/2140 (99%)	2002 (94%)	114 (5%)	8 (0%)	30	26

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	12	PHE
1	A	38	VAL
1	B	12	PHE
1	B	38	VAL
1	C	38	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	468/470 (100%)	457 (98%)	11 (2%)	43	47
1	B	468/470 (100%)	454 (97%)	14 (3%)	36	38
1	C	468/470 (100%)	454 (97%)	14 (3%)	36	38
1	D	468/470 (100%)	462 (99%)	6 (1%)	61	68
All	All	1872/1880 (100%)	1827 (98%)	45 (2%)	43	47

5 of 45 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	195	ASN
1	C	428	GLU
1	C	320	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	365	ARG
1	C	447	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	389	HIS
1	D	210	ASN
1	B	195	ASN
1	B	210	ASN
1	B	482	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](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 11 ligands modelled in this entry, 4 are monoatomic - leaving 7 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	MES	B	3006	-	12,12,12	2.04	3 (25%)	15,16,16	6.08	8 (53%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	MES	A	3005	-	12,12,12	1.74	3 (25%)	15,16,16	5.85	9 (60%)
4	GOL	C	3010	-	5,5,5	0.41	0	5,5,5	0.21	0
3	MES	D	3007	-	12,12,12	5.94	4 (33%)	15,16,16	6.62	8 (53%)
4	GOL	B	3009	-	5,5,5	0.40	0	5,5,5	0.47	0
4	GOL	A	3008	-	5,5,5	0.41	0	5,5,5	0.70	0
4	GOL	D	3011	-	5,5,5	0.40	0	5,5,5	0.39	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
3	MES	B	3006	-	-	3/6/14/14	0/1/1/1
3	MES	A	3005	-	-	3/6/14/14	0/1/1/1
4	GOL	C	3010	-	-	0/4/4/4	-
3	MES	D	3007	-	-	4/6/14/14	0/1/1/1
4	GOL	B	3009	-	-	3/4/4/4	-
4	GOL	A	3008	-	-	3/4/4/4	-
4	GOL	D	3011	-	-	0/4/4/4	-

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	3007	MES	O2S-S	11.35	1.77	1.45
3	D	3007	MES	O1S-S	10.98	1.76	1.45
3	D	3007	MES	O3S-S	10.19	1.85	1.47
3	D	3007	MES	C8-S	7.73	1.88	1.77
3	B	3006	MES	C8-S	5.60	1.85	1.77

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	3007	MES	O3S-S-O2S	-18.38	65.43	111.40
3	B	3006	MES	O3S-S-O2S	-11.43	82.79	111.40
3	D	3007	MES	O3S-S-O1S	-11.20	83.39	111.40
3	A	3005	MES	O3S-S-O2S	-11.19	83.40	111.40
3	B	3006	MES	O3S-S-C8	-10.81	84.86	106.00

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	3005	MES	C7-C8-S-O1S
3	A	3005	MES	C7-C8-S-O3S
3	D	3007	MES	C8-C7-N4-C5
3	D	3007	MES	C7-C8-S-O2S
4	B	3009	GOL	O1-C1-C2-C3

There are no ring outliers.

2 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	3006	MES	3	0
3	D	3007	MES	4	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](#)

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	A	533/535 (99%)	-0.12	3 (0%) 85 87	13, 27, 41, 50	0
1	B	533/535 (99%)	-0.21	2 (0%) 88 90	13, 26, 40, 46	0
1	C	533/535 (99%)	0.84	38 (7%) 22 25	22, 42, 55, 64	0
1	D	533/535 (99%)	-0.16	0 100 100	14, 27, 41, 46	0
All	All	2132/2140 (99%)	0.09	43 (2%) 65 69	13, 30, 50, 64	0

The worst 5 of 43 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	151	ASP	3.7
1	B	151	ASP	3.6
1	C	483	TRP	3.2
1	C	21	VAL	2.9
1	C	261	PHE	2.9

### 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 oligosaccharides 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( $\text{\AA}^2$ )	Q<0.9
3	MES	D	3007	12/12	0.56	0.26	109,110,112,112	0
3	MES	B	3006	12/12	0.77	0.20	74,76,77,77	0
2	CA	C	3003	1/1	0.88	0.08	57,57,57,57	0
4	GOL	A	3008	6/6	0.91	0.09	24,25,26,26	0
4	GOL	C	3010	6/6	0.92	0.09	31,31,31,33	0
4	GOL	D	3011	6/6	0.92	0.07	22,25,25,28	0
4	GOL	B	3009	6/6	0.93	0.07	37,39,40,41	0
3	MES	A	3005	12/12	0.95	0.07	29,34,37,37	0
2	CA	D	3004	1/1	0.97	0.04	30,30,30,30	0
2	CA	B	3002	1/1	0.98	0.03	27,27,27,27	0
2	CA	A	3001	1/1	0.99	0.02	25,25,25,25	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.