



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 6, 2025 – 12:23 pm GMT

PDB ID : 9HE8  
Title : The molecular structure of a beta-1,4-D-xylosidase from the probiotic bacterium *Levilactobacillus brevis*  
Authors : Linares-Pasten, J.; Logan, D.T.; Nordberg Karlsson, E.  
Deposited on : 2024-11-13  
Resolution : 1.90 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	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

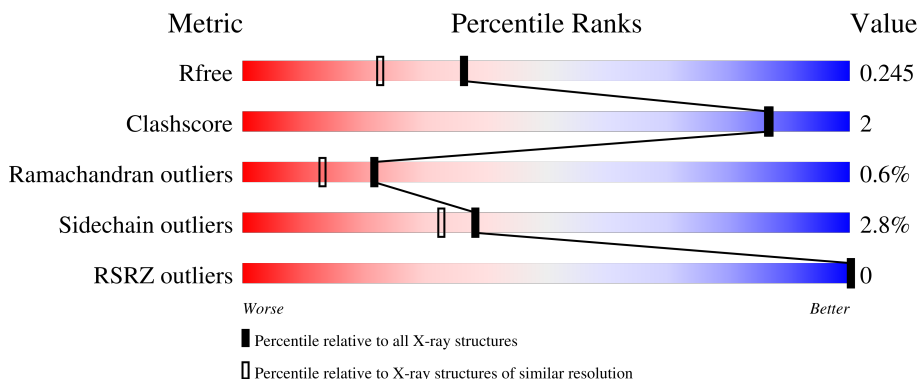
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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))
$R_{free}$	164625	7293 (1.90-1.90)
Clashscore	180529	8090 (1.90-1.90)
Ramachandran outliers	177936	8022 (1.90-1.90)
Sidechain outliers	177891	8022 (1.90-1.90)
RSRZ outliers	164620	7292 (1.90-1.90)

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	560	
1	B	560	
1	C	560	
1	D	560	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 18316 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-xylosidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	540	Total	C	N	O	S	0	1	0
			4354	2783	717	845	9			
1	B	548	Total	C	N	O	S	0	2	0
			4421	2825	731	854	11			
1	C	540	Total	C	N	O	S	0	2	0
			4359	2787	717	845	10			
1	D	540	Total	C	N	O	S	0	1	0
			4355	2785	717	843	10			

There are 80 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	initiating methionine	UNP A0A1L7HA01
A	-18	GLY	-	expression tag	UNP A0A1L7HA01
A	-17	SER	-	expression tag	UNP A0A1L7HA01
A	-16	SER	-	expression tag	UNP A0A1L7HA01
A	-15	HIS	-	expression tag	UNP A0A1L7HA01
A	-14	HIS	-	expression tag	UNP A0A1L7HA01
A	-13	HIS	-	expression tag	UNP A0A1L7HA01
A	-12	HIS	-	expression tag	UNP A0A1L7HA01
A	-11	HIS	-	expression tag	UNP A0A1L7HA01
A	-10	HIS	-	expression tag	UNP A0A1L7HA01
A	-9	SER	-	expression tag	UNP A0A1L7HA01
A	-8	SER	-	expression tag	UNP A0A1L7HA01
A	-7	GLY	-	expression tag	UNP A0A1L7HA01
A	-6	LEU	-	expression tag	UNP A0A1L7HA01
A	-5	VAL	-	expression tag	UNP A0A1L7HA01
A	-4	PRO	-	expression tag	UNP A0A1L7HA01
A	-3	ARG	-	expression tag	UNP A0A1L7HA01
A	-2	GLY	-	expression tag	UNP A0A1L7HA01
A	-1	SER	-	expression tag	UNP A0A1L7HA01
A	0	HIS	-	expression tag	UNP A0A1L7HA01
B	-19	MET	-	initiating methionine	UNP A0A1L7HA01

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-18	GLY	-	expression tag	UNP A0A1L7HA01
B	-17	SER	-	expression tag	UNP A0A1L7HA01
B	-16	SER	-	expression tag	UNP A0A1L7HA01
B	-15	HIS	-	expression tag	UNP A0A1L7HA01
B	-14	HIS	-	expression tag	UNP A0A1L7HA01
B	-13	HIS	-	expression tag	UNP A0A1L7HA01
B	-12	HIS	-	expression tag	UNP A0A1L7HA01
B	-11	HIS	-	expression tag	UNP A0A1L7HA01
B	-10	HIS	-	expression tag	UNP A0A1L7HA01
B	-9	SER	-	expression tag	UNP A0A1L7HA01
B	-8	SER	-	expression tag	UNP A0A1L7HA01
B	-7	GLY	-	expression tag	UNP A0A1L7HA01
B	-6	LEU	-	expression tag	UNP A0A1L7HA01
B	-5	VAL	-	expression tag	UNP A0A1L7HA01
B	-4	PRO	-	expression tag	UNP A0A1L7HA01
B	-3	ARG	-	expression tag	UNP A0A1L7HA01
B	-2	GLY	-	expression tag	UNP A0A1L7HA01
B	-1	SER	-	expression tag	UNP A0A1L7HA01
B	0	HIS	-	expression tag	UNP A0A1L7HA01
C	-19	MET	-	initiating methionine	UNP A0A1L7HA01
C	-18	GLY	-	expression tag	UNP A0A1L7HA01
C	-17	SER	-	expression tag	UNP A0A1L7HA01
C	-16	SER	-	expression tag	UNP A0A1L7HA01
C	-15	HIS	-	expression tag	UNP A0A1L7HA01
C	-14	HIS	-	expression tag	UNP A0A1L7HA01
C	-13	HIS	-	expression tag	UNP A0A1L7HA01
C	-12	HIS	-	expression tag	UNP A0A1L7HA01
C	-11	HIS	-	expression tag	UNP A0A1L7HA01
C	-10	HIS	-	expression tag	UNP A0A1L7HA01
C	-9	SER	-	expression tag	UNP A0A1L7HA01
C	-8	SER	-	expression tag	UNP A0A1L7HA01
C	-7	GLY	-	expression tag	UNP A0A1L7HA01
C	-6	LEU	-	expression tag	UNP A0A1L7HA01
C	-5	VAL	-	expression tag	UNP A0A1L7HA01
C	-4	PRO	-	expression tag	UNP A0A1L7HA01
C	-3	ARG	-	expression tag	UNP A0A1L7HA01
C	-2	GLY	-	expression tag	UNP A0A1L7HA01
C	-1	SER	-	expression tag	UNP A0A1L7HA01
C	0	HIS	-	expression tag	UNP A0A1L7HA01
D	-19	MET	-	initiating methionine	UNP A0A1L7HA01
D	-18	GLY	-	expression tag	UNP A0A1L7HA01
D	-17	SER	-	expression tag	UNP A0A1L7HA01

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	-16	SER	-	expression tag	UNP A0A1L7HA01
D	-15	HIS	-	expression tag	UNP A0A1L7HA01
D	-14	HIS	-	expression tag	UNP A0A1L7HA01
D	-13	HIS	-	expression tag	UNP A0A1L7HA01
D	-12	HIS	-	expression tag	UNP A0A1L7HA01
D	-11	HIS	-	expression tag	UNP A0A1L7HA01
D	-10	HIS	-	expression tag	UNP A0A1L7HA01
D	-9	SER	-	expression tag	UNP A0A1L7HA01
D	-8	SER	-	expression tag	UNP A0A1L7HA01
D	-7	GLY	-	expression tag	UNP A0A1L7HA01
D	-6	LEU	-	expression tag	UNP A0A1L7HA01
D	-5	VAL	-	expression tag	UNP A0A1L7HA01
D	-4	PRO	-	expression tag	UNP A0A1L7HA01
D	-3	ARG	-	expression tag	UNP A0A1L7HA01
D	-2	GLY	-	expression tag	UNP A0A1L7HA01
D	-1	SER	-	expression tag	UNP A0A1L7HA01
D	0	HIS	-	expression tag	UNP A0A1L7HA01

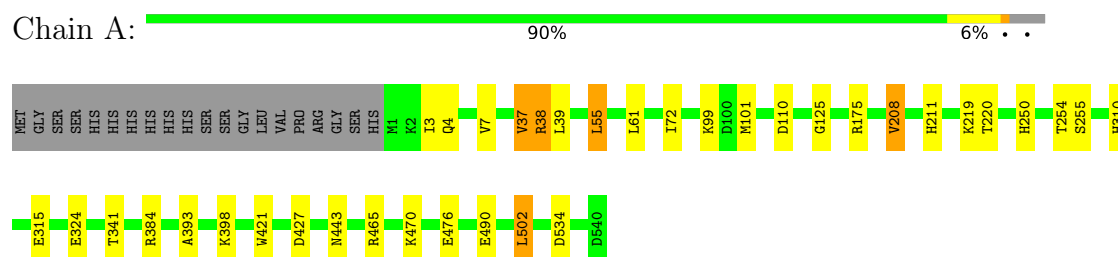
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	181	Total O 181 181	0	0
2	B	238	Total O 238 238	0	0
2	C	247	Total O 247 247	0	0
2	D	161	Total O 161 161	0	0

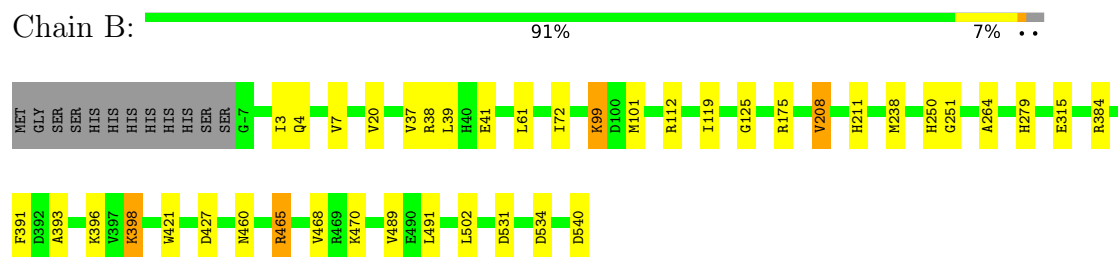
### 3 Residue-property plots

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.

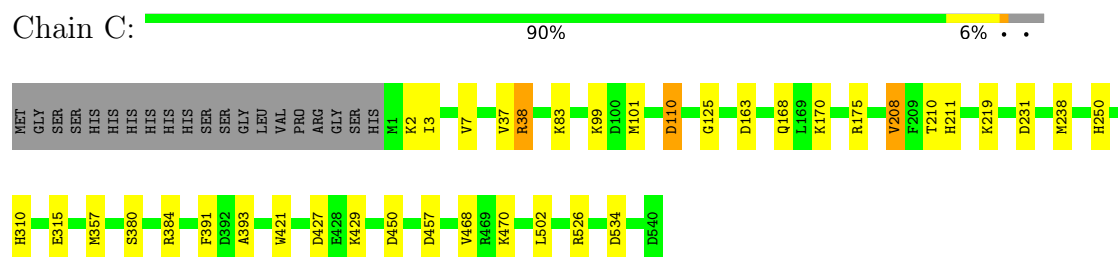
#### • Molecule 1: Beta-xylosidase



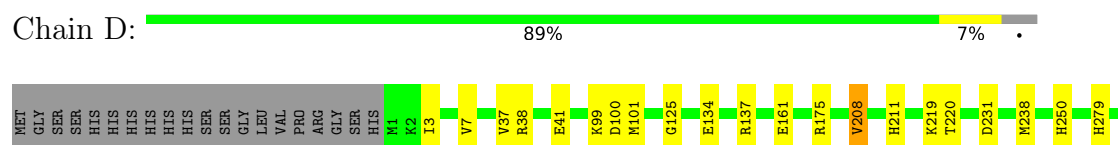
#### • Molecule 1: Beta-xylosidase



#### • Molecule 1: Beta-xylosidase



#### • Molecule 1: Beta-xylosidase





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	78.43Å 177.18Å 78.75Å 90.00° 98.84° 90.00°	Depositor
Resolution (Å)	47.05 – 1.90 47.05 – 1.90	Depositor EDS
% Data completeness (in resolution range)	97.8 (47.05-1.90) 97.0 (47.05-1.90)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.47 (at 1.90Å)	Xtriage
Refinement program	REFMAC 5.8.0425	Depositor
R, $R_{free}$	0.214 , 0.241 0.218 , 0.245	Depositor DCC
$R_{free}$ test set	8167 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	12.8	Xtriage
Anisotropy	1.269	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 15.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.43$ , $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.126 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	18316	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.87% 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

### 5.1 Standard geometry

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	0.43	0/4494	0.79	1/6132 (0.0%)
1	B	0.45	0/4563	0.80	2/6224 (0.0%)
1	C	0.46	0/4502	0.83	3/6142 (0.0%)
1	D	0.44	0/4495	0.80	3/6134 (0.0%)
All	All	0.44	0/18054	0.81	9/24632 (0.0%)

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	1
1	B	0	1
1	D	0	1
All	All	0	3

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	38	ARG	CG-CD-NE	-9.55	91.75	111.80
1	D	38	ARG	NE-CZ-NH2	-8.19	116.21	120.30
1	C	526	ARG	NE-CZ-NH1	6.93	123.77	120.30
1	A	38	ARG	NE-CZ-NH2	-5.87	117.36	120.30
1	B	465	ARG	NE-CZ-NH1	5.85	123.22	120.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	465	ARG	Sidechain

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Group
1	B	112	ARG	Sidechain
1	D	465	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	4354	0	4058	14	0
1	B	4421	0	4128	16	0
1	C	4359	0	4067	19	0
1	D	4355	0	4062	21	0
2	A	181	0	0	3	0
2	B	238	0	0	3	0
2	C	247	0	0	5	0
2	D	161	0	0	6	0
All	All	18316	0	16315	68	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 2.

The worst 5 of 68 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:C:219:LYS:HB2	2:C:832:HOH:O	1.53	1.08
1:A:476:GLU:HG3	2:A:753:HOH:O	1.56	1.04
1:A:443:ASN:HB3	2:A:772:HOH:O	1.61	1.00
1:C:450:ASP:HB3	2:C:830:HOH:O	1.66	0.93
1:D:454:LYS:HE3	2:D:737:HOH:O	1.68	0.93

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	539/560 (96%)	514 (95%)	22 (4%)	3 (1%)	22	13
1	B	548/560 (98%)	522 (95%)	23 (4%)	3 (0%)	25	17
1	C	540/560 (96%)	514 (95%)	23 (4%)	3 (1%)	22	13
1	D	539/560 (96%)	513 (95%)	23 (4%)	3 (1%)	22	13
All	All	2166/2240 (97%)	2063 (95%)	91 (4%)	12 (1%)	22	13

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	470	LYS
1	B	470	LYS
1	C	470	LYS
1	D	470	LYS
1	A	37	VAL

### 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	464/481 (96%)	449 (97%)	15 (3%)	34	27
1	B	471/481 (98%)	458 (97%)	13 (3%)	38	33
1	C	465/481 (97%)	453 (97%)	12 (3%)	41	36
1	D	464/481 (96%)	452 (97%)	12 (3%)	41	36
All	All	1864/1924 (97%)	1812 (97%)	52 (3%)	38	33

5 of 52 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	38	ARG
1	C	384	ARG
1	D	427	ASP
1	C	83	LYS
1	C	208	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	4	GLN
1	B	432	HIS
1	C	168	GLN
1	D	371	GLN

### 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](#)

There are no ligands in this entry.

## 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	540/560 (96%)	-1.49	0 100 100	13, 19, 33, 52	1 (0%)
1	B	548/560 (97%)	-1.62	0 100 100	8, 15, 26, 50	2 (0%)
1	C	540/560 (96%)	-1.62	0 100 100	11, 15, 26, 46	2 (0%)
1	D	540/560 (96%)	-1.51	0 100 100	10, 18, 34, 51	1 (0%)
All	All	2168/2240 (96%)	-1.56	0 100 100	8, 16, 30, 52	6 (0%)

There are no RSRZ outliers to report.

### 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 monosaccharides in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.