



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 09:19 AM GMT

PDB ID : 3HZ3  
Title : Lactobacillus reuteri N-terminally truncated glucansucrase GTF180(D1025N)  
-sucrose complex  
Authors : Vujicic-Zagar, A.; Pijning, T.; Kralj, S.; Eeuwema, W.; Dijkhuizen, L.; Dijkstra, B.W.  
Deposited on : 2009-06-23  
Resolution : 2.22 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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

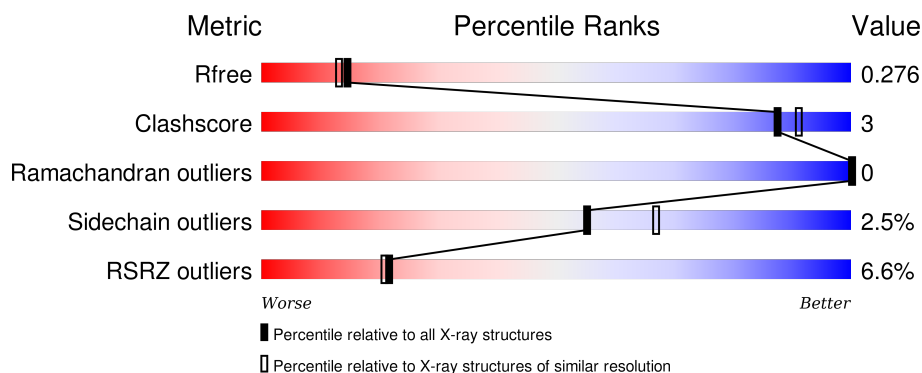
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.22 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))
$R_{free}$	91344	4405 (2.24-2.20)
Clashscore	102246	5146 (2.24-2.20)
Ramachandran outliers	100387	5065 (2.24-2.20)
Sidechain outliers	100360	5066 (2.24-2.20)
RSRZ outliers	91569	4414 (2.24-2.20)

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.

Mol	Chain	Length	Quality of chain
1	A	1039	<div> <div>6%</div> <div>88%</div> <div>8%</div> <div>.</div> </div>

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
3	SUC	A	1780	-	-	-	X

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 8251 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 Glucansucrase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1003	Total	C	N	O	S	0	0	0
			7960	4979	1349	1611	21			

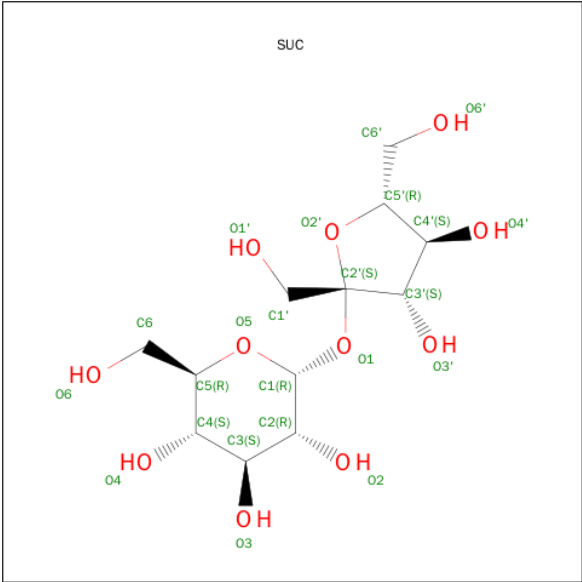
There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	740	MET	-	EXPRESSION TAG	UNP Q5SBN3
A	741	GLY	-	EXPRESSION TAG	UNP Q5SBN3
A	1025	ASN	ASP	ENGINEERED MUTATION	UNP Q5SBN3
A	1674	LEU	PHE	SEE REMARK 999	UNP Q5SBN3
A	1773	HIS	-	EXPRESSION TAG	UNP Q5SBN3
A	1774	HIS	-	EXPRESSION TAG	UNP Q5SBN3
A	1775	HIS	-	EXPRESSION TAG	UNP Q5SBN3
A	1776	HIS	-	EXPRESSION TAG	UNP Q5SBN3
A	1777	HIS	-	EXPRESSION TAG	UNP Q5SBN3
A	1778	HIS	-	EXPRESSION TAG	UNP Q5SBN3

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

- Molecule 3 is SUGAR (SUCROSE) (three-letter code: SUC) (formula: C<sub>12</sub>H<sub>22</sub>O<sub>11</sub>).



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

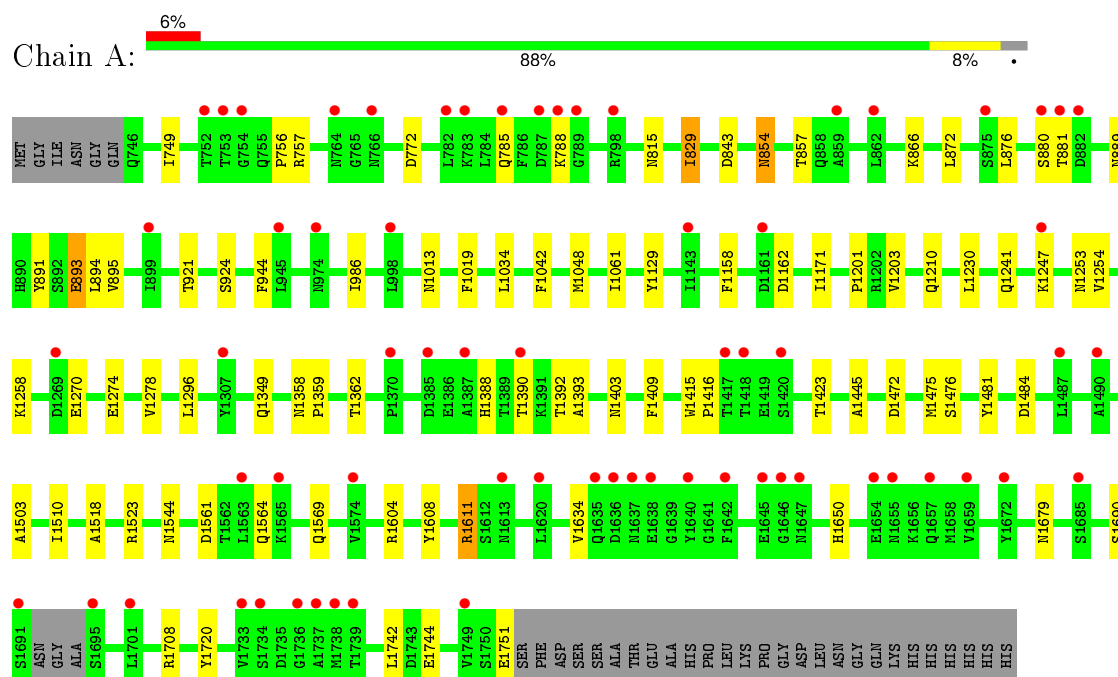
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	244	Total	O	0	0
			244	244		

### 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 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 ( $\text{RSRZ} > 2$ ). 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: Glucansucrase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	57.78Å 66.23Å 82.72Å 107.13° 101.37° 85.40°	Depositor
Resolution (Å)	20.00 – 2.22 20.00 – 2.22	Depositor EDS
% Data completeness (in resolution range)	96.3 (20.00-2.22) 87.9 (20.00-2.22)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.97 (at 2.21Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, $R_{free}$	0.193 , 0.243 0.234 , 0.276	Depositor DCC
$R_{free}$ test set	2721 reflections (4.98%)	DCC
Wilson B-factor (Å <sup>2</sup> )	31.4	Xtriage
Anisotropy	0.153	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 43.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 54678 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	8251	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.67% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

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

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	0.66	0/8136	0.69	1/11066 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	1296	LEU	CA-CB-CG	5.05	126.92	115.30

There are no chirality outliers.

There are no planarity outliers.

### 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	7960	0	7511	40	0
2	A	1	0	0	0	0
3	A	46	0	44	1	0
4	A	244	0	0	2	0
All	All	8251	0	7555	41	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 3.

All (41) 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:1390:THR:HG23	1:A:1392:THR:HG22	1.78	0.66
1:A:1544:ASN:HD21	1:A:1604:ARG:HH22	1.42	0.65
1:A:815:ASN:HD22	1:A:1518:ALA:H	1.53	0.54
1:A:866:LYS:HB2	1:A:872:LEU:HB2	1.92	0.52
1:A:854:ASN:HD22	1:A:857:THR:H	1.58	0.52
1:A:1061:ILE:HB	1:A:1129:TYR:CE2	2.45	0.51
1:A:815:ASN:ND2	1:A:1518:ALA:H	2.08	0.50
1:A:749:ILE:HA	1:A:756:PRO:HA	1.94	0.50
1:A:866:LYS:HE3	1:A:876:LEU:O	2.12	0.49
1:A:1409:PHE:CG	1:A:1445:ALA:HB2	2.47	0.49
1:A:1158:PHE:HB2	1:A:1162:ASP:HB2	1.94	0.49
1:A:1475:MET:HG3	1:A:1484:ASP:HB3	1.95	0.49
1:A:1472:ASP:OD2	1:A:1476:SER:OG	2.22	0.48
1:A:1258:LYS:HD3	1:A:1270:GLU:OE1	2.13	0.48
1:A:1203:VAL:HG21	1:A:1230:LEU:HD21	1.97	0.47
1:A:1358:ASN:HB2	1:A:1359:PRO:CD	2.44	0.47
1:A:944:PHE:O	1:A:1523:ARG:HD2	2.15	0.46
1:A:986:ILE:HD11	1:A:1510:ILE:HG13	1.97	0.46
1:A:854:ASN:ND2	1:A:857:THR:H	2.15	0.45
1:A:894:LEU:HA	1:A:894:LEU:HD12	1.80	0.45
1:A:891:TYR:O	1:A:895:VAL:HG23	2.17	0.45
1:A:1247:LYS:HG2	4:A:182:HOH:O	2.16	0.44
1:A:1611:ARG:HD2	4:A:225:HOH:O	2.18	0.44
1:A:1171:ILE:HD12	1:A:1210:GLN:HG2	2.00	0.44
1:A:829:ILE:HD13	1:A:843:ASP:HB3	2.00	0.44
1:A:1388:HIS:HE1	1:A:1393:ALA:O	2.01	0.43
1:A:1423:THR:HB	1:A:1481:TYR:HB3	2.00	0.43
3:A:1780:SUC:O6	3:A:1780:SUC:O4	2.36	0.43
1:A:1604:ARG:HD3	1:A:1608:TYR:CE2	2.54	0.42
1:A:1720:TYR:HB3	1:A:1742:LEU:HD11	2.00	0.42
1:A:757:ARG:NH2	1:A:1751:GLU:OE2	2.53	0.42
1:A:889:ASN:O	1:A:893:GLU:HG3	2.20	0.42
1:A:1201:PRO:HD2	1:A:1403:ASN:O	2.20	0.41
1:A:1241:GLN:HA	1:A:1254:VAL:O	2.20	0.41
1:A:1253:ASN:HB3	1:A:1278:VAL:HB	2.02	0.41
1:A:1019:PHE:CG	1:A:1503:ALA:HB2	2.55	0.41
1:A:1034:LEU:HA	1:A:1034:LEU:HD23	1.86	0.41
1:A:1650:HIS:HE1	1:A:1679:ASN:OD1	2.04	0.40
1:A:1415:TRP:HA	1:A:1416:PRO:HD3	1.99	0.40
1:A:1270:GLU:O	1:A:1274:GLU:HG3	2.21	0.40
1:A:1042:PHE:HB3	1:A:1048:MET:HG3	2.04	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	999/1039 (96%)	955 (96%)	44 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 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	855/883 (97%)	834 (98%)	21 (2%)	55	67

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	772	ASP
1	A	785	GLN
1	A	788	LYS
1	A	829	ILE
1	A	854	ASN
1	A	880	SER
1	A	881	THR
1	A	893	GLU
1	A	921	THR
1	A	924	SER

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	1013	ASN
1	A	1349	GLN
1	A	1362	THR
1	A	1561	ASP
1	A	1564	GLN
1	A	1569	GLN
1	A	1611	ARG
1	A	1634	VAL
1	A	1690	SER
1	A	1708	ARG
1	A	1744	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	755	GLN
1	A	815	ASN
1	A	854	ASN
1	A	1013	ASN
1	A	1349	GLN
1	A	1544	ASN
1	A	1650	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 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	SUC	A	1779	-	24,24,24	1.14	2 (8%)	36,36,36	1.27	5 (13%)
3	SUC	A	1780	-	24,24,24	1.44	4 (16%)	36,36,36	1.90	11 (30%)

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	SUC	A	1779	-	-	0/12/51/51	0/2/2/2
3	SUC	A	1780	-	-	0/12/51/51	0/2/2/2

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1779	SUC	C4-C3	2.09	1.57	1.52
3	A	1780	SUC	O3'-C3'	2.27	1.47	1.42
3	A	1780	SUC	O2'-C5'	2.35	1.49	1.43
3	A	1780	SUC	O5-C1	2.41	1.48	1.41
3	A	1779	SUC	O5-C1	3.21	1.50	1.41
3	A	1780	SUC	O2'-C2'	3.63	1.51	1.42

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1780	SUC	C1-O5-C5	-4.25	105.50	113.75
3	A	1780	SUC	C6'-C5'-C4'	-2.91	108.20	115.08
3	A	1780	SUC	C3-C4-C5	-2.78	105.35	110.20
3	A	1779	SUC	O5-C1-C2	-2.74	104.65	110.28
3	A	1780	SUC	O1-C1-C2	-2.17	101.12	108.36
3	A	1779	SUC	O4'-C4'-C3'	-2.08	105.41	112.01
3	A	1779	SUC	O2-C2-C1	2.12	114.66	110.02
3	A	1780	SUC	C6-C5-C4	2.25	118.56	113.02
3	A	1779	SUC	C2'-O1-C1	2.51	124.15	117.53

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1779	SUC	O5-C5-C4	2.63	114.62	109.68
3	A	1780	SUC	C2'-O1-C1	2.73	124.73	117.53
3	A	1780	SUC	O2'-C2'-C1'	2.81	115.64	107.98
3	A	1780	SUC	O5-C5-C6	2.92	113.73	106.36
3	A	1780	SUC	C1-C2-C3	3.01	115.90	109.97
3	A	1780	SUC	O2'-C5'-C6'	3.37	118.76	108.57
3	A	1780	SUC	O5-C1-C2	4.05	118.60	110.28

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1780	SUC	1	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

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	1003/1039 (96%)	0.55	66 (6%)	22 21	35, 56, 84, 110	0

All (66) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1691	SER	5.1
1	A	766	ASN	4.9
1	A	881	THR	4.3
1	A	882	ASP	4.2
1	A	1387	ALA	4.1
1	A	1636	ASP	4.1
1	A	1565	LYS	4.0
1	A	789	GLY	3.8
1	A	752	THR	3.8
1	A	1647	ASN	3.5
1	A	1695	SER	3.3
1	A	1385	ASP	3.3
1	A	1637	ASN	3.2
1	A	1417	THR	3.1
1	A	1635	GLN	3.1
1	A	783	LYS	3.0
1	A	1143	ILE	3.0
1	A	753	THR	3.0
1	A	862	LEU	3.0
1	A	788	LYS	2.9
1	A	1638	GLU	2.9
1	A	782	LEU	2.8
1	A	785	GLN	2.8
1	A	880	SER	2.8
1	A	764	ASN	2.7
1	A	998	LEU	2.7
1	A	1487	LEU	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	1738	MET	2.7
1	A	1307	TYR	2.6
1	A	1640	TYR	2.6
1	A	1734	SER	2.6
1	A	945	LEU	2.6
1	A	1490	ALA	2.6
1	A	1655	ASN	2.6
1	A	1574	VAL	2.5
1	A	1739	THR	2.5
1	A	1736	GLY	2.5
1	A	1642	PHE	2.5
1	A	1420	SER	2.5
1	A	1737	ALA	2.4
1	A	1418	THR	2.4
1	A	1563	LEU	2.4
1	A	754	GLY	2.4
1	A	875	SER	2.4
1	A	1645	GLU	2.4
1	A	1657	GLN	2.4
1	A	1390	THR	2.4
1	A	1247	LYS	2.4
1	A	1654	GLU	2.3
1	A	798	ARG	2.3
1	A	1370	PRO	2.3
1	A	899	ILE	2.2
1	A	787	ASP	2.2
1	A	1701	LEU	2.1
1	A	1646	GLY	2.1
1	A	1620	LEU	2.1
1	A	859	ALA	2.1
1	A	1672	TYR	2.1
1	A	1733	VAL	2.1
1	A	974	ASN	2.1
1	A	1685	SER	2.1
1	A	1269	ASP	2.1
1	A	1161	ASP	2.0
1	A	1613	ASN	2.0
1	A	1659	VAL	2.0
1	A	1749	VAL	2.0

## 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 carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	SUC	A	1780	23/23	0.63	0.35	6.19	49,60,70,72	0
3	SUC	A	1779	23/23	0.93	0.12	-0.71	36,40,43,44	0
2	CA	A	1	1/1	0.98	0.10	-3.53	36,36,36,36	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.