



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 07:02 AM GMT

PDB ID : 2ZB5  
Title : Crystal structure of the measles virus hemagglutinin (complex-sugar-type)  
Authors : Hashiguchi, T.; Kajikawa, M.; Maita, N.; Takeda, M.; Kuroki, K.; Sasaki, K.; Kohda, D.; Yanagi, Y.; Maenaka, K.  
Deposited on : 2007-10-16  
Resolution : 3.00 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

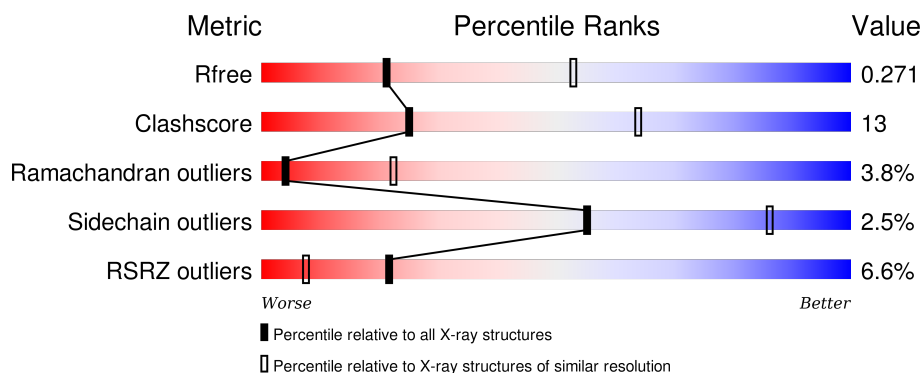
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (3.00-3.00)

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	481	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3269 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 Hemagglutinin protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	424	3241	2075	539	602	25	0	0	0

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	146	GLU	-	EXPRESSION TAG	UNP Q83625
A	147	THR	-	EXPRESSION TAG	UNP Q83625
A	148	GLY	-	EXPRESSION TAG	UNP Q83625
A	484	THR	ASN	ENGINEERED	UNP Q83625
A	492	GLY	GLU	ENGINEERED	UNP Q83625
A	618	GLY	-	EXPRESSION TAG	UNP Q83625
A	619	THR	-	EXPRESSION TAG	UNP Q83625
A	620	LYS	-	EXPRESSION TAG	UNP Q83625
A	621	HIS	-	EXPRESSION TAG	UNP Q83625
A	622	HIS	-	EXPRESSION TAG	UNP Q83625
A	623	HIS	-	EXPRESSION TAG	UNP Q83625
A	624	HIS	-	EXPRESSION TAG	UNP Q83625
A	625	HIS	-	EXPRESSION TAG	UNP Q83625
A	626	HIS	-	EXPRESSION TAG	UNP Q83625

- Molecule 2 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).

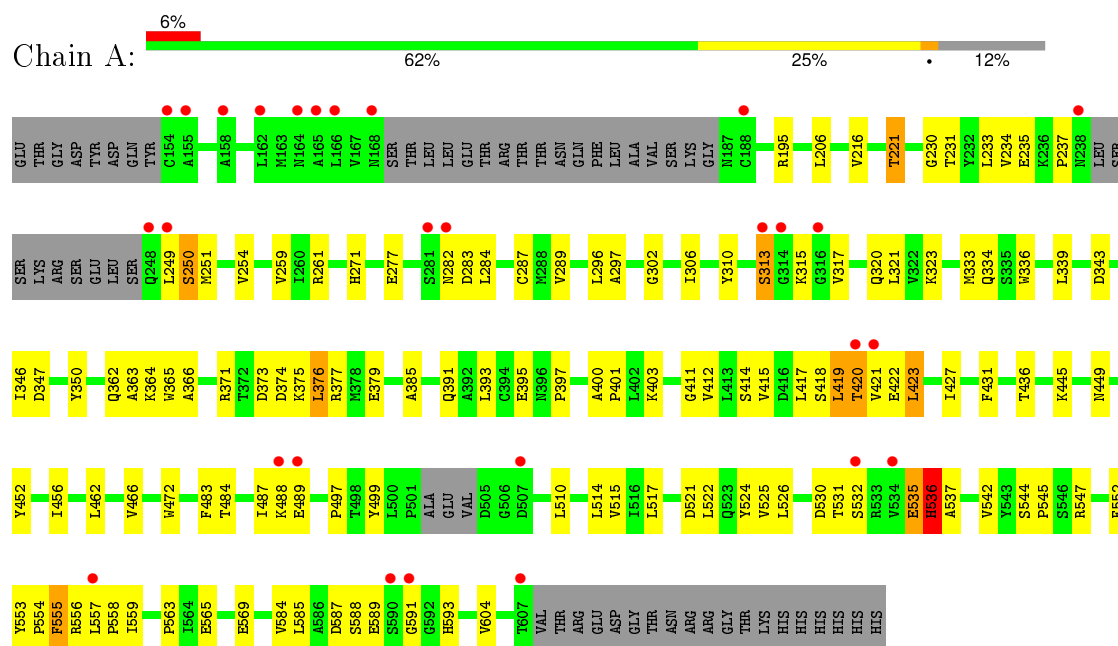


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		

### 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 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 ( $RSRZ > 2$ ). 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: Hemagglutinin protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	134.52Å 134.52Å 99.25Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 3.00 29.82 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.7 (30.00-3.00) 99.7 (29.82-3.00)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.61 (at 3.00Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.227 , 0.271 0.228 , 0.271	Depositor DCC
$R_{free}$ test set	1283 reflections (6.87%)	DCC
Wilson B-factor (Å <sup>2</sup> )	51.4	Xtriage
Anisotropy	0.903	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 41.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	1 of 18688 reflections (0.005%)	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	3269	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.68% 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: NAG

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	0.41	0/3322	0.68	0/4524

There are no bond length outliers.

There are no bond angle outliers.

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	3241	0	3150	81	0
2	A	28	0	26	3	0
All	All	3269	0	3176	82	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 13.

All (82) 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:230:GLY:HA2	1:A:289:VAL:HG11	1.58	0.85
1:A:593:HIS:H	2:A:11:NAG:H81	1.50	0.76

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:393:LEU:H	1:A:393:LEU:HD12	1.57	0.69
1:A:397:PRO:HG2	1:A:403:LYS:HB3	1.75	0.69
1:A:544:SER:HB3	1:A:547:ARG:O	1.94	0.67
1:A:391:GLN:O	1:A:395:GLU:HG2	1.94	0.66
1:A:537:ALA:HA	1:A:556:ARG:CG	2.25	0.66
1:A:235:GLU:HG2	1:A:249:LEU:HD23	1.79	0.65
1:A:362:GLN:HE21	1:A:364:LYS:HE2	1.61	0.65
1:A:565:GLU:HB3	1:A:584:VAL:HB	1.83	0.60
1:A:231:THR:HG21	1:A:287:CYS:HB2	1.83	0.60
1:A:364:LYS:HE3	1:A:472:TRP:CE2	2.36	0.60
1:A:249:LEU:O	1:A:250:SER:HB2	2.02	0.60
1:A:557:LEU:N	1:A:558:PRO:HD2	2.16	0.59
2:A:11:NAG:H82	2:A:11:NAG:H3	1.84	0.59
1:A:593:HIS:N	2:A:11:NAG:H81	2.18	0.59
1:A:343:ASP:HB3	1:A:346:ILE:HG12	1.85	0.58
1:A:259:VAL:HB	1:A:261:ARG:NH1	2.19	0.58
1:A:514:LEU:HD11	1:A:526:LEU:HD13	1.86	0.58
1:A:261:ARG:HD2	1:A:271:HIS:ND1	2.20	0.57
1:A:456:ILE:HB	1:A:466:VAL:HB	1.87	0.56
1:A:362:GLN:NE2	1:A:364:LYS:HE2	2.21	0.56
1:A:289:VAL:HG13	1:A:296:LEU:HD11	1.88	0.56
1:A:517:LEU:HD11	1:A:525:VAL:CG2	2.35	0.56
1:A:310:TYR:H	1:A:313:SER:HB2	1.71	0.55
1:A:385:ALA:HB2	1:A:487:ILE:HG13	1.89	0.55
1:A:535:GLU:HG3	1:A:552:PHE:HZ	1.72	0.55
1:A:530:ASP:OD2	1:A:532:SER:HB2	2.07	0.54
1:A:522:LEU:O	1:A:545:PRO:HD3	2.09	0.53
1:A:254:VAL:HG22	1:A:277:GLU:HG3	1.89	0.53
1:A:339:LEU:HD11	1:A:427:ILE:HD11	1.91	0.52
1:A:283:ASP:C	1:A:284:LEU:HD22	2.31	0.51
1:A:604:VAL:O	1:A:604:VAL:HG13	2.12	0.50
1:A:419:LEU:HA	1:A:423:LEU:HG	1.92	0.50
1:A:306:ILE:HG21	1:A:317:VAL:HG11	1.93	0.49
1:A:334:GLN:O	1:A:334:GLN:HG3	2.12	0.49
1:A:517:LEU:HD11	1:A:525:VAL:HG22	1.94	0.49
1:A:283:ASP:HB2	1:A:336:TRP:HH2	1.77	0.49
1:A:536:HIS:CD2	1:A:559:ILE:HA	2.48	0.49
1:A:321:LEU:HD22	1:A:415:VAL:HG11	1.94	0.49
1:A:554:PRO:HG2	1:A:555:PHE:CD1	2.47	0.48
1:A:563:PRO:HA	1:A:585:LEU:HD23	1.96	0.48
1:A:556:ARG:O	1:A:556:ARG:CG	2.60	0.48

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:587:ASP:O	1:A:589:GLU:N	2.46	0.48
1:A:553:TYR:HE1	1:A:555:PHE:HB2	1.78	0.47
1:A:483:PHE:N	1:A:483:PHE:CD2	2.82	0.47
1:A:233:LEU:HD13	1:A:251:MET:CE	2.44	0.47
1:A:365:TRP:O	1:A:412:VAL:HA	2.15	0.47
1:A:553:TYR:CE1	1:A:555:PHE:HB2	2.50	0.46
1:A:445:LYS:HA	1:A:452:TYR:HD1	1.81	0.46
1:A:284:LEU:HD11	1:A:320:GLN:OE1	2.16	0.46
1:A:364:LYS:HE3	1:A:472:TRP:CZ2	2.51	0.46
1:A:363:ALA:O	1:A:414:SER:HA	2.17	0.45
1:A:436:THR:HG23	1:A:497:PRO:HG2	1.99	0.44
1:A:555:PHE:O	1:A:556:ARG:CG	2.66	0.44
1:A:488:LYS:O	1:A:489:GLU:HB2	2.18	0.44
1:A:483:PHE:HZ	1:A:524:TYR:HH	1.63	0.44
1:A:216:VAL:HG22	1:A:234:VAL:HG12	1.99	0.44
1:A:296:LEU:HD12	1:A:297:ALA:H	1.84	0.43
1:A:376:LEU:HD23	1:A:377:ARG:HG3	2.00	0.43
1:A:557:LEU:N	1:A:558:PRO:CD	2.81	0.43
1:A:488:LYS:N	1:A:488:LYS:HD2	2.32	0.43
1:A:233:LEU:HD13	1:A:251:MET:HE3	2.00	0.43
1:A:452:TYR:CE2	1:A:472:TRP:CZ3	3.07	0.42
1:A:554:PRO:O	1:A:555:PHE:O	2.38	0.42
1:A:366:ALA:HA	1:A:411:GLY:O	2.19	0.42
1:A:393:LEU:HD11	1:A:499:TYR:CE1	2.54	0.42
1:A:525:VAL:HG13	1:A:542:VAL:HG22	2.01	0.41
1:A:484:THR:HG22	1:A:497:PRO:HA	2.02	0.41
1:A:350:TYR:CE1	1:A:401:PRO:HB3	2.55	0.41
1:A:221:THR:HB	1:A:569:GLU:OE1	2.20	0.41
1:A:531:THR:HG22	1:A:536:HIS:CE1	2.55	0.41
1:A:323:LYS:HD3	1:A:417:LEU:HD22	2.02	0.41
1:A:418:SER:C	1:A:420:THR:H	2.23	0.41
1:A:462:LEU:HD21	1:A:499:TYR:CE1	2.55	0.41
1:A:556:ARG:C	1:A:558:PRO:HD2	2.40	0.41
1:A:195:ARG:HH11	1:A:195:ARG:HG2	1.85	0.41
1:A:237:PRO:HA	1:A:249:LEU:HD11	2.02	0.41
1:A:371:ARG:HB3	1:A:373:ASP:OD1	2.20	0.41
1:A:515:VAL:HG23	1:A:515:VAL:O	2.21	0.41
1:A:375:LYS:HA	1:A:379:GLU:HB2	2.03	0.40
1:A:374:ASP:HA	1:A:431:PHE:CZ	2.56	0.40

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	416/481 (86%)	354 (85%)	46 (11%)	16 (4%)	<b>4</b> <b>22</b>

All (16) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	420	THR
1	A	555	PHE
1	A	282	ASN
1	A	302	GLY
1	A	449	ASN
1	A	250	SER
1	A	419	LEU
1	A	421	VAL
1	A	536	HIS
1	A	313	SER
1	A	422	GLU
1	A	423	LEU
1	A	535	GLU
1	A	588	SER
1	A	591	GLY
1	A	400	ALA

### 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.

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
-----	-------	----------	-----------	----------	-------------

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	354/420 (84%)	345 (98%)	9 (2%)	55 86

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	206	LEU
1	A	221	THR
1	A	315	LYS
1	A	333	MET
1	A	347	ASP
1	A	376	LEU
1	A	510	LEU
1	A	521	ASP
1	A	536	HIS

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	225	GLN
1	A	362	GLN
1	A	384	GLN
1	A	449	ASN
1	A	450	ASN
1	A	575	GLN
1	A	593	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

2 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	1	1	14,14,15	0.60	0	15,19,21	0.75	0
2	NAG	A	11	1	14,14,15	0.70	0	15,19,21	0.79	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
2	NAG	A	1	1	-	2/6/23/26	0/1/1/1
2	NAG	A	11	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1	NAG	C8-C7-N2-C2
2	A	1	NAG	O7-C7-N2-C2

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	11	NAG	3	0

## 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 ⓘ

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	424/481 (88%)	-0.04	28 (6%) 22 7	30, 51, 107, 133	0

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	590	SER	6.8
1	A	421	VAL	4.8
1	A	154	CYS	4.0
1	A	282	ASN	3.6
1	A	420	THR	3.6
1	A	165	ALA	3.4
1	A	313	SER	3.4
1	A	557	LEU	3.3
1	A	164	ASN	3.1
1	A	155	ALA	2.9
1	A	168	ASN	2.9
1	A	158	ALA	2.9
1	A	607	THR	2.8
1	A	238	ASN	2.7
1	A	489	GLU	2.7
1	A	507	ASP	2.6
1	A	314	GLY	2.5
1	A	591	GLY	2.5
1	A	534	VAL	2.5
1	A	316	GLY	2.4
1	A	188	CYS	2.4
1	A	248	GLN	2.4
1	A	488	LYS	2.3
1	A	532	SER	2.3
1	A	249	LEU	2.3
1	A	162	LEU	2.1
1	A	281	SER	2.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	166	LEU	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(Å <sup>2</sup> )	Q<0.9
2	NAG	A	11	14/15	0.91	0.36	0.98	79,83,85,85	0
2	NAG	A	1	14/15	0.64	0.40	-	88,94,96,96	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.