



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 10, 2020 – 04:44 AM BST

PDB ID : 3BE1  
Title : Dual specific bH1 Fab in complex with the extracellular domain of HER2/ErbB-2  
Authors : Bostrom, J.M.; Wiesmann, C.; Appleton, B.A.  
Deposited on : 2007-11-15  
Resolution : 2.90 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.13.1

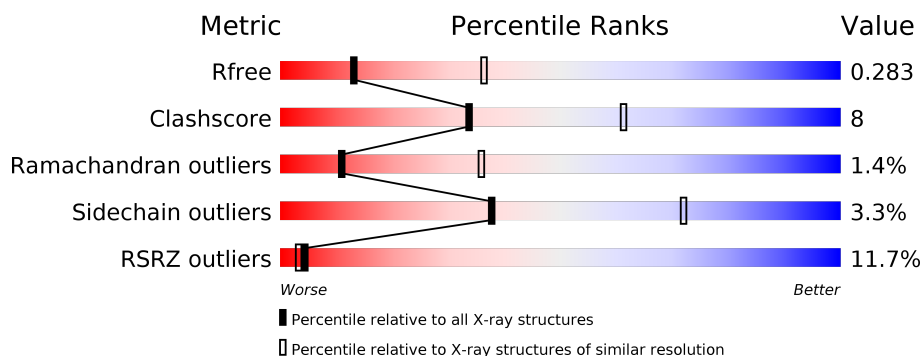
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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))
$R_{free}$	130704	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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 on 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	624	<div> <div>2%</div> <div> <div></div> <div>77%</div> <div>16%</div> <div>• 7%</div> </div> </div>
2	H	230	<div> <div>27%</div> <div> <div></div> <div>75%</div> <div>17%</div> <div>• 7%</div> </div> </div>
3	L	218	<div> <div>20%</div> <div> <div></div> <div>75%</div> <div>22%</div> <div>••</div> </div> </div>

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 7810 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 tyrosine-protein kinase erbB-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	583	Total	C	N	O	S	0	0	0
			4501	2800	807	842	52			

- Molecule 2 is a protein called Fab Fragment-Heavy Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	213	Total	C	N	O	S	0	0	0
			1605	1018	270	311	6			

- Molecule 3 is a protein called Fab Fragment-Light Chain.

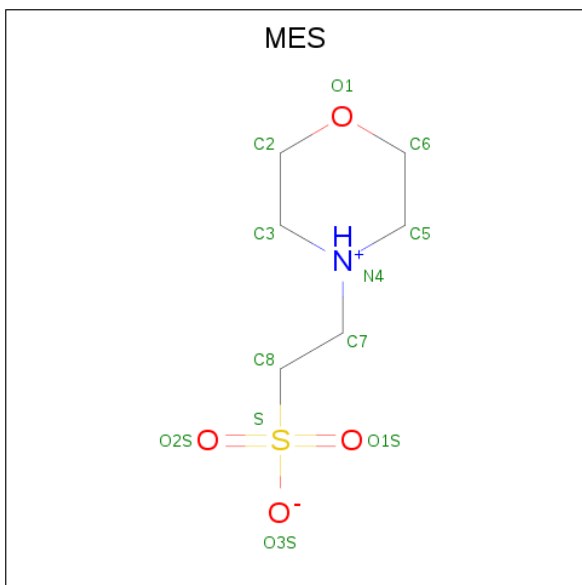
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	216	Total	C	N	O	S	0	0	0
			1664	1048	276	335	5			

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

- Molecule 5 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: C<sub>6</sub>H<sub>13</sub>NO<sub>4</sub>S).

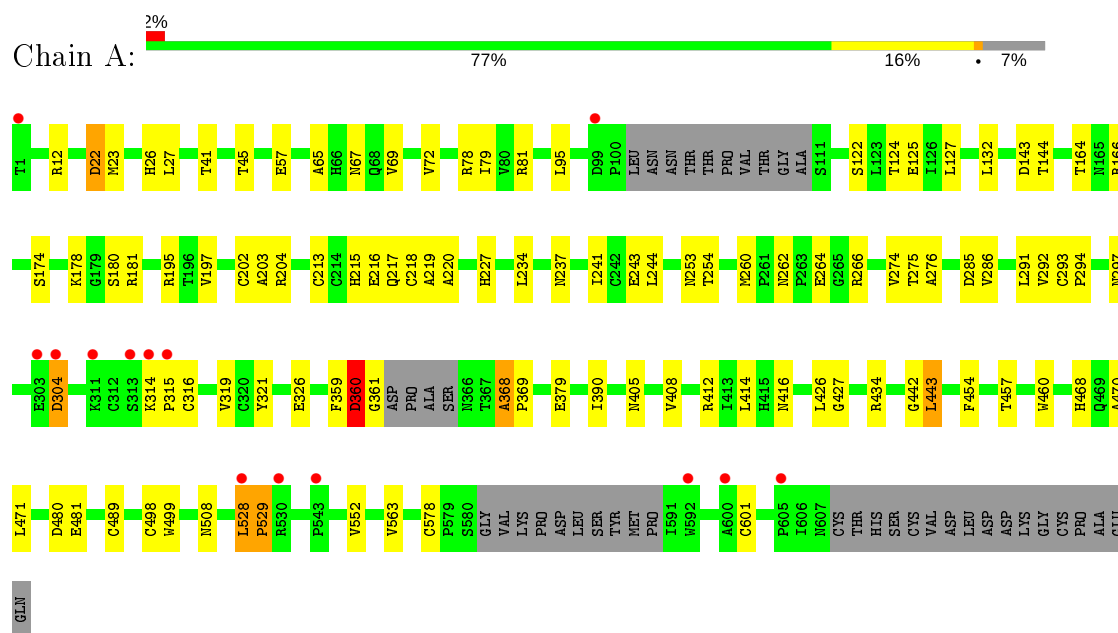


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	L	1	Total	C	N	O	S	0	0
			12	6	1	4	1		

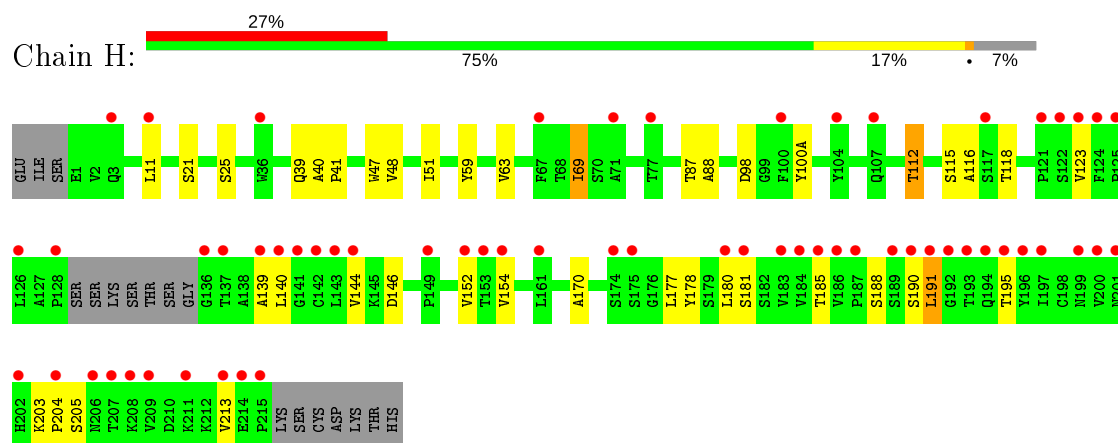
### 3 Residue-property plots [i](#)

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.

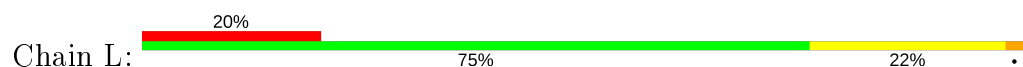
- Molecule 1: Receptor tyrosine-protein kinase erbB-2

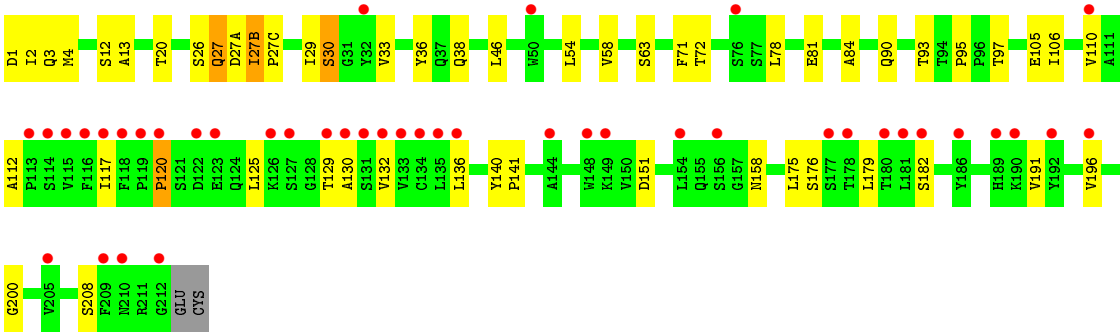


- Molecule 2: Fab Fragment-Heavy Chain



- Molecule 3: Fab Fragment-Light Chain





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	62.32Å 115.06Å 208.17Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.90 48.49 – 2.89	Depositor EDS
% Data completeness (in resolution range)	99.8 (30.00-2.90) 99.3 (48.49-2.89)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.36 (at 2.91Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.236 , 0.277 0.220 , 0.283	Depositor DCC
$R_{free}$ test set	1715 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	58.6	Xtriage
Anisotropy	0.701	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 66.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	7810	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.12% 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, MES

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	0.52	1/4607 (0.0%)	0.66	0/6266
2	H	0.35	0/1646	0.52	0/2245
3	L	0.36	0/1705	0.51	0/2320
All	All	0.46	1/7958 (0.0%)	0.60	0/10831

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

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	218	CYS	CB-SG	-5.09	1.73	1.81

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	427	GLY	Peptide

### 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	4501	0	4308	73	0
2	H	1605	0	1562	23	0
3	L	1664	0	1608	32	0
4	A	28	0	26	5	0
5	L	12	0	12	1	0
All	All	7810	0	7516	127	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 8.

The worst 5 of 127 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:237:ASN:HD21	4:A:1001:NAG:C1	1.21	1.53
1:A:508:ASN:HD21	4:A:1002:NAG:C1	1.34	1.38
1:A:124:THR:HG21	1:A:220:ALA:O	1.45	1.14
1:A:578:CYS:HG	1:A:601:CYS:HG	0.98	0.93
1:A:144:THR:HG22	1:A:181:ARG:HA	1.49	0.91

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	575/624 (92%)	523 (91%)	44 (8%)	8 (1%)	11	36
2	H	209/230 (91%)	185 (88%)	21 (10%)	3 (1%)	11	36
3	L	214/218 (98%)	191 (89%)	20 (9%)	3 (1%)	11	36
All	All	998/1072 (93%)	899 (90%)	85 (8%)	14 (1%)	11	36

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	178	LYS
1	A	213	CYS
1	A	215	HIS
1	A	360	ASP
2	H	191	LEU

### 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	503/538 (94%)	487 (97%)	16 (3%)	39	73
2	H	175/191 (92%)	170 (97%)	5 (3%)	42	76
3	L	189/192 (98%)	181 (96%)	8 (4%)	30	63
All	All	867/921 (94%)	838 (97%)	29 (3%)	38	72

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	408	VAL
2	H	21	SER
3	L	71	PHE
1	A	443	LEU
2	H	25	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	237	ASN
1	A	307	GLN
1	A	542	HIS
1	A	217	GLN
1	A	526	GLN

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

3 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$
5	MES	L	215	-	12,12,12	2.14	1 (8%)	14,16,16	2.56	7 (50%)
4	NAG	A	1001	1	14,14,15	0.74	1 (7%)	17,19,21	1.24	2 (11%)
4	NAG	A	1002	1	14,14,15	0.59	0	17,19,21	0.89	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
5	MES	L	215	-	-	4/6/14/14	0/1/1/1
4	NAG	A	1001	1	-	0/6/23/26	0/1/1/1
4	NAG	A	1002	1	-	2/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	L	215	MES	C8-S	-7.10	1.67	1.77
4	A	1001	NAG	O5-C1	-2.07	1.40	1.43

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	L	215	MES	C5-N4-C3	4.60	119.19	108.83
5	L	215	MES	C7-N4-C5	3.64	120.54	111.23
5	L	215	MES	C7-N4-C3	3.53	120.26	111.23
5	L	215	MES	O1S-S-C8	3.47	111.09	106.92
5	L	215	MES	C2-C3-N4	-3.04	105.50	110.10

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1002	NAG	C8-C7-N2-C2
4	A	1002	NAG	O7-C7-N2-C2
5	L	215	MES	C7-C8-S-O3S
5	L	215	MES	C8-C7-N4-C5
5	L	215	MES	C7-C8-S-O1S

There are no ring outliers.

3 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	L	215	MES	1	0
4	A	1001	NAG	3	0
4	A	1002	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 [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	583/624 (93%)	0.20	14 (2%) 59 56	19, 41, 79, 109	0
2	H	213/230 (92%)	1.51	61 (28%) 0 0	56, 86, 141, 149	0
3	L	216/218 (99%)	1.21	43 (19%) 1 0	49, 86, 143, 148	0
All	All	1012/1072 (94%)	0.69	118 (11%) 4 3	19, 59, 135, 149	0

The worst 5 of 118 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	140	LEU	11.9
2	H	184	VAL	9.2
2	H	125	PRO	8.5
3	L	181	LEU	8.4
1	A	528	LEU	8.3

### 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](#)

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
4	NAG	A	1002	14/15	0.73	0.36	107,113,123,125	0
4	NAG	A	1001	14/15	0.90	0.22	28,46,49,49	0
5	MES	L	215	12/12	0.93	0.21	76,91,99,100	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.