



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 4, 2017 – 10:44 PM EDT

PDB ID : 5MY6  
Title : Crystal structure of a HER2-Nb complex  
Authors : Sterckx, Y.G.-J.; D'Huyvetter, M.; Devoogdt, N.  
Deposited on : unknown  
Resolution : 2.25 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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030345  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : rb-20030345

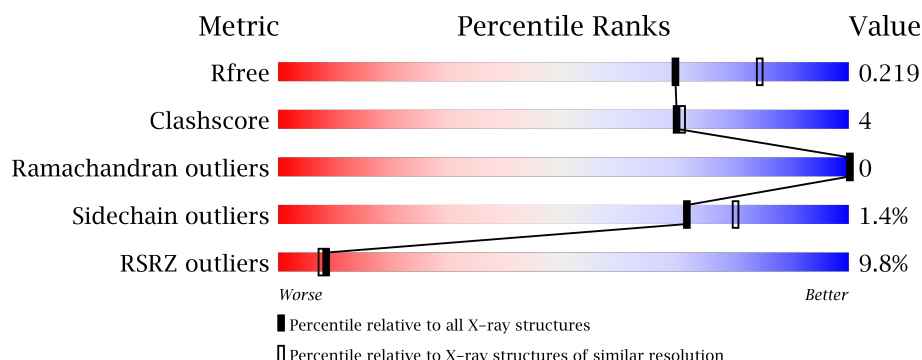
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.25 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))
$R_{free}$	100719	1804 (2.26-2.22)
Clashscore	112137	1957 (2.26-2.22)
Ramachandran outliers	110173	1916 (2.26-2.22)
Sidechain outliers	110143	1917 (2.26-2.22)
RSRZ outliers	101464	1809 (2.26-2.22)

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	622	<div> <div>10%</div> <div>78%</div> <div>11%</div> <div>11%</div> </div>
2	B	115	<div> <div>3%</div> <div>97%</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	NAG	A	701	-	-	-	X
4	GOL	A	704	-	-	-	X
4	GOL	A	705	-	-	-	X

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 5581 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.

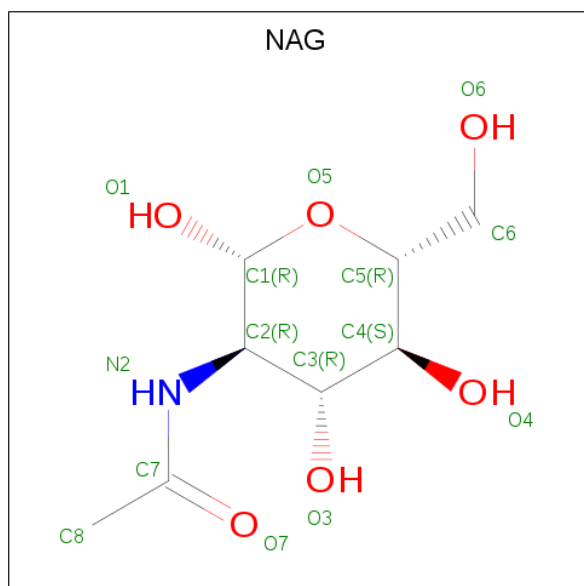
- Molecule 1 is a protein called Receptor tyrosine-protein kinase erbB-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	556	Total	C	N	O	S	0	5	0
			4282	2658	769	806	49			

- Molecule 2 is a protein called Nanobody 2Rs15d.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	115	Total	C	N	O	S	0	0	0
			875	545	150	174	6			

- Molecule 3 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

*Continued on next page...*

*Continued from previous page...*

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

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



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

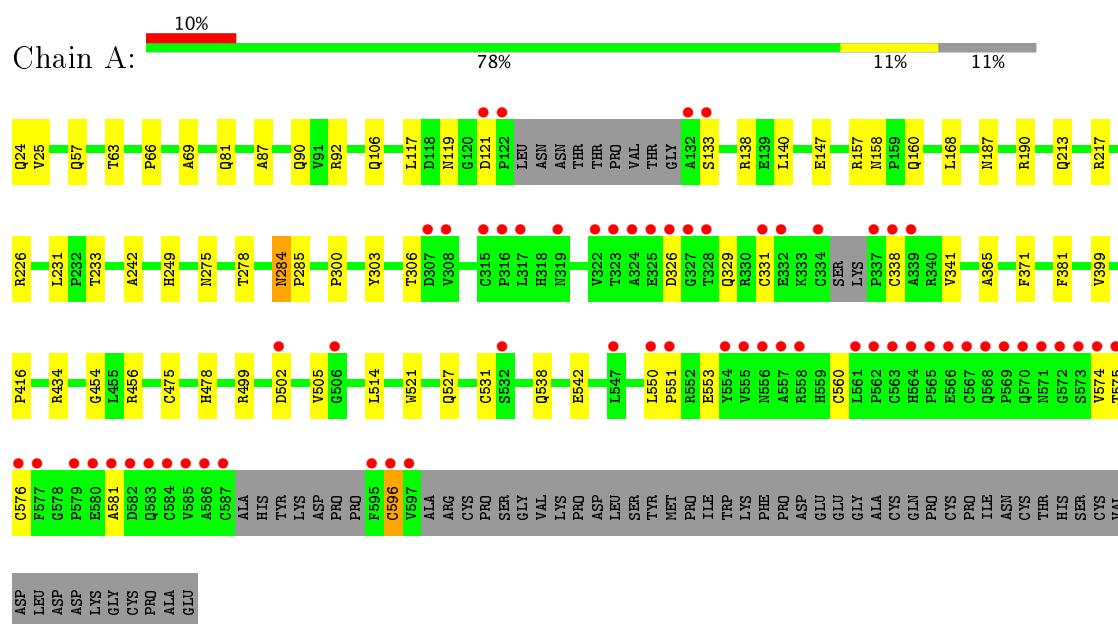
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	297	Total	O	0	0
			297	297		
5	B	73	Total	O	0	0
			73	73		

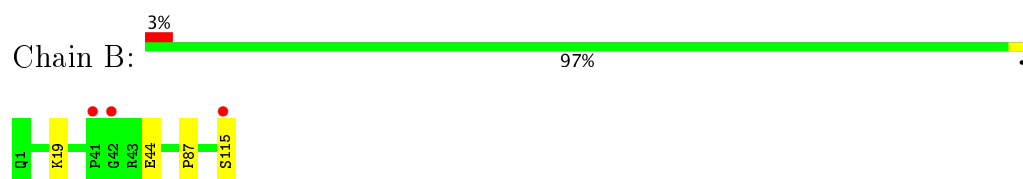
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: Nanobody 2Rs15d



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	66.52Å 113.52Å 137.43Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.79 – 2.25 47.79 – 2.25	Depositor EDS
% Data completeness (in resolution range)	98.1 (47.79-2.25) 98.1 (47.79-2.25)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.50 (at 2.24Å)	Xtriage
Refinement program	PHENIX 1.10 _2155	Depositor
R, $R_{free}$	0.185 , 0.220 0.182 , 0.219	Depositor DCC
$R_{free}$ test set	2475 reflections (5.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	39.5	Xtriage
Anisotropy	0.120	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 40.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	5581	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.69% 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: GOL, 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	0.31	0/4376	0.52	0/5954
2	B	0.31	0/892	0.55	0/1205
All	All	0.31	0/5268	0.53	0/7159

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 [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	4282	0	4059	41	0
2	B	875	0	832	3	0
3	A	42	0	38	0	0
4	A	12	0	16	0	0
5	A	297	0	0	2	0
5	B	73	0	0	1	0
All	All	5581	0	4945	43	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 4.

All (43) 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:275:ASN:HD22	1:A:278:THR:H	1.16	0.94
1:A:275:ASN:ND2	1:A:278:THR:H	1.73	0.85
1:A:581:ALA:HB1	1:A:596:CYS:H	1.52	0.74
1:A:168:LEU:H	1:A:213:GLN:HE22	1.37	0.73
1:A:57:GLN:OE1	1:A:81:GLN:NE2	2.30	0.64
1:A:106:GLN:NE2	5:A:802:HOH:O	2.31	0.64
1:A:92:ARG:HB3	1:A:138:ARG:HG3	1.79	0.63
1:A:574:VAL:HG12	1:A:575:THR:H	1.64	0.63
1:A:66:PRO:HD2	1:A:69:ALA:HB2	1.80	0.62
1:A:119:ASN:HB2	1:A:158:ASN:HD22	1.68	0.57
1:A:574:VAL:HG12	1:A:575:THR:N	2.19	0.57
1:A:514:LEU:HD21	1:A:542:GLU:HG3	1.87	0.56
1:A:284:ASN:HD22	1:A:285:PRO:HD2	1.71	0.54
1:A:306:THR:HG22	1:A:331:CYS:HB3	1.91	0.52
1:A:160:GLN:NE2	1:A:190:ARG:HD3	2.25	0.52
2:B:19:LYS:HE3	5:B:209:HOH:O	2.11	0.50
1:A:456:ARG:HA	1:A:521:TRP:CD1	2.47	0.50
1:A:502:ASP:HA	1:A:505:VAL:HG12	1.94	0.49
1:A:147:GLU:HB2	1:A:242:ALA:HA	1.95	0.49
1:A:92:ARG:NH2	1:A:133:SER:O	2.46	0.49
1:A:329:GLN:OE1	1:A:434:ARG:NH2	2.48	0.47
1:A:90:GLN:NE2	1:A:121:ASP:OD1	2.46	0.47
1:A:231:LEU:HB2	1:A:233:THR:HG22	1.96	0.46
1:A:160:GLN:HE21	1:A:190:ARG:HD3	1.79	0.46
1:A:365:ALA:HA	1:A:399:VAL:HG22	1.98	0.46
1:A:478:HIS:HD2	5:A:1061:HOH:O	1.99	0.46
1:A:475:CYS:SG	1:A:499:ARG:HD3	2.56	0.46
1:A:87:ALA:HA	1:A:117:LEU:O	2.16	0.45
1:A:550:LEU:HD23	1:A:551:PRO:HA	1.99	0.45
1:A:119:ASN:HD22	1:A:158:ASN:ND2	2.15	0.45
1:A:341:VAL:HG13	1:A:371:PHE:CD1	2.53	0.44
1:A:300:PRO:HG2	1:A:303:TYR:CD1	2.53	0.44
1:A:63:THR:HA	1:A:87:ALA:O	2.17	0.44
1:A:119:ASN:H	1:A:158:ASN:ND2	2.16	0.44
1:A:24:GLN:HG3	1:A:25:VAL:HG23	2.00	0.43
1:A:381:PHE:CZ	1:A:416:PRO:HD3	2.54	0.43
1:A:454:GLY:HA2	1:A:527:GLN:OE1	2.17	0.43
1:A:575:THR:HG22	1:A:576:CYS:SG	2.59	0.43
1:A:553:GLU:OE1	1:A:560:CYS:HB3	2.19	0.43
1:A:157:ARG:NH1	2:B:44:GLU:HB3	2.35	0.42
2:B:87:PRO:HG3	2:B:115:SER:HB3	2.02	0.41

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:140:LEU:HD23	1:A:140:LEU:HA	1.91	0.40
1:A:217:ARG:HB2	1:A:226:ARG:NH1	2.36	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	553/622 (89%)	530 (96%)	23 (4%)	0	100	100
2	B	113/115 (98%)	113 (100%)	0	0	100	100
All	All	666/737 (90%)	643 (96%)	23 (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	472/536 (88%)	464 (98%)	8 (2%)	66	74
2	B	94/96 (98%)	94 (100%)	0	100	100
All	All	566/632 (90%)	558 (99%)	8 (1%)	71	80

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	187	ASN
1	A	249	HIS
1	A	284	ASN
1	A	326	ASP
1	A	338	CYS
1	A	531	CYS
1	A	538	GLN
1	A	596	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	75	GLN
1	A	81	GLN
1	A	158	ASN
1	A	160	GLN
1	A	213	GLN
1	A	275	ASN
1	A	284	ASN
2	B	13	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

5 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
3	NAG	A	701	1,3	14,14,15	0.40	0	15,19,21	0.45	0
3	NAG	A	702	3	14,14,15	0.22	0	15,19,21	0.59	0
3	NAG	A	703	-	14,14,15	1.47	2 (14%)	15,19,21	0.91	1 (6%)
4	GOL	A	704	-	5,5,5	0.36	0	5,5,5	0.30	0
4	GOL	A	705	-	5,5,5	0.36	0	5,5,5	0.32	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	NAG	A	701	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	702	3	-	0/6/23/26	0/1/1/1
3	NAG	A	703	-	-	0/6/23/26	0/1/1/1
4	GOL	A	704	-	-	0/4/4/4	0/0/0/0
4	GOL	A	705	-	-	0/4/4/4	0/0/0/0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	703	NAG	C1-C2	2.92	1.56	1.52
3	A	703	NAG	O5-C1	4.08	1.50	1.43

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	703	NAG	C1-O5-C5	-2.05	109.34	112.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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	556/622 (89%)	0.58	63 (11%) 6 5	23, 42, 112, 158	0
2	B	115/115 (100%)	-0.04	3 (2%) 56 58	26, 36, 70, 107	0
All	All	671/737 (91%)	0.48	66 (9%) 8 7	23, 41, 108, 158	0

All (66) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	587	CYS	10.9
1	A	584	CYS	9.5
1	A	567	CYS	8.8
1	A	596	CYS	8.6
1	A	586	ALA	7.7
1	A	564	HIS	7.0
1	A	317	LEU	6.8
1	A	334	CYS	6.8
1	A	562	PRO	6.6
1	A	337	PRO	6.5
1	A	561	LEU	6.4
1	A	338	CYS	6.3
1	A	565	PRO	6.3
1	A	583	GLN	6.2
1	A	563	CYS	6.1
1	A	585	VAL	5.7
1	A	557	ALA	5.6
1	A	595	PHE	5.6
1	A	581	ALA	5.5
1	A	582	ASP	5.5
1	A	597	VAL	5.5
1	A	132	ALA	5.3
1	A	580	GLU	5.2
1	A	579	PRO	5.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	316	PRO	5.0
1	A	576	CYS	4.8
1	A	566	GLU	4.7
1	A	326	ASP	4.4
1	A	577	PHE	4.4
1	A	569	PRO	4.3
1	A	319	ASN	4.2
2	B	115	SER	4.2
1	A	121	ASP	4.2
1	A	307	ASP	4.1
2	B	42	GLY	3.9
1	A	556	ASN	3.9
2	B	41	PRO	3.8
1	A	568	GLN	3.8
1	A	572	GLY	3.6
1	A	554	TYR	3.6
1	A	502	ASP	3.6
1	A	327	GLY	3.4
1	A	555	VAL	3.4
1	A	573	SER	3.3
1	A	532	SER	3.3
1	A	339	ALA	3.3
1	A	325	GLU	3.0
1	A	322	VAL	3.0
1	A	547	LEU	3.0
1	A	308	VAL	2.9
1	A	122	PRO	2.7
1	A	550	LEU	2.7
1	A	574	VAL	2.6
1	A	575	THR	2.6
1	A	570	GLN	2.5
1	A	328	THR	2.5
1	A	506	GLY	2.5
1	A	331	CYS	2.5
1	A	133	SER	2.4
1	A	315	CYS	2.4
1	A	324	ALA	2.3
1	A	571	ASN	2.2
1	A	332	GLU	2.2
1	A	323	THR	2.1
1	A	558	ARG	2.1
1	A	551	PRO	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
4	GOL	A	705	6/6	0.66	0.30	5.16	81,86,93,104	0
4	GOL	A	704	6/6	0.80	0.26	4.55	54,62,69,74	0
3	NAG	A	701	14/15	0.82	0.16	2.23	46,53,75,79	0
3	NAG	A	703	14/15	0.54	0.40	-	81,98,110,113	0
3	NAG	A	702	14/15	0.73	0.34	-	73,90,97,97	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.