



# Full wwPDB X-ray Structure Validation Report ⓘ

May 15, 2020 – 08:51 pm BST

PDB ID : 2Q8B  
Title : Structure of the malaria antigen AMA1 in complex with a growth-inhibitory antibody  
Authors : Gupta, A.; Murphy, V.J.; Anders, R.F.; Batchelor, A.H.  
Deposited on : 2007-06-10  
Resolution : 2.30 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

<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
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

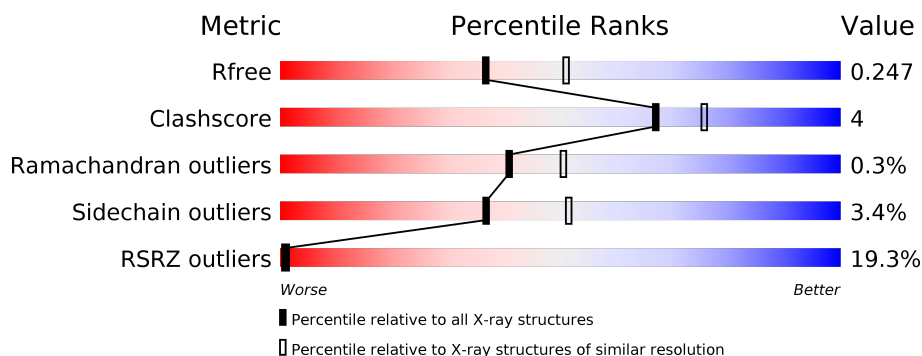
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	336	<div> <div>14%</div> <div> <div></div> <div>77%</div> <div>9%</div> <div>•</div> <div>13%</div> </div> </div>
2	L	214	<div> <div>17%</div> <div> <div></div> <div>87%</div> <div>13%</div> </div> </div>
3	H	210	<div> <div>25%</div> <div> <div></div> <div>86%</div> <div>8%</div> <div>•</div> <div>5%</div> </div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5598 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 Apical membrane antigen 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	292	Total	C	N	O	S	0	0	0
			2303	1458	382	447	16			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	103	GLY	-	EXPRESSION TAG	UNP Q7KQK5

- Molecule 2 is a protein called 1F9 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	213	Total	C	N	O	S	0	0	0
			1616	1019	264	327	6			

- Molecule 3 is a protein called 1F9 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	H	199	Total	C	N	O	S	0	0	0
			1459	924	236	293	6			

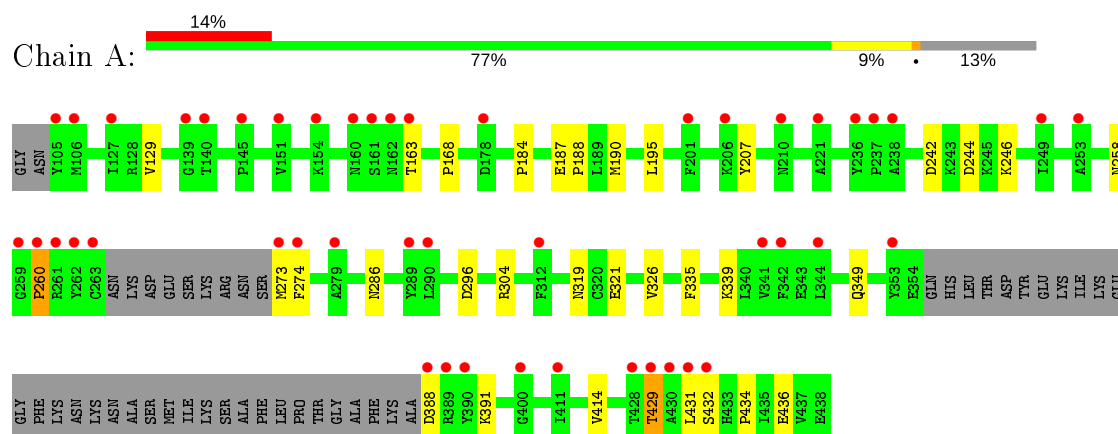
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	175	Total	O	0	0
			175	175		
4	L	11	Total	O	0	0
			11	11		
4	H	34	Total	O	0	0
			34	34		

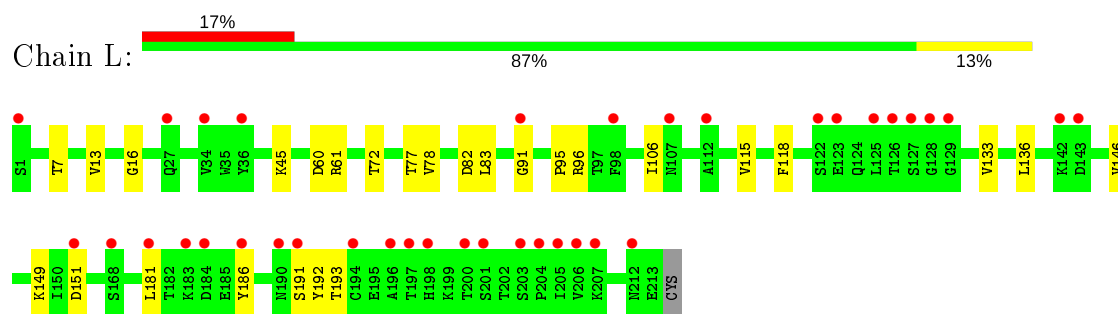
### 3 Residue-property plots [i](#)

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.

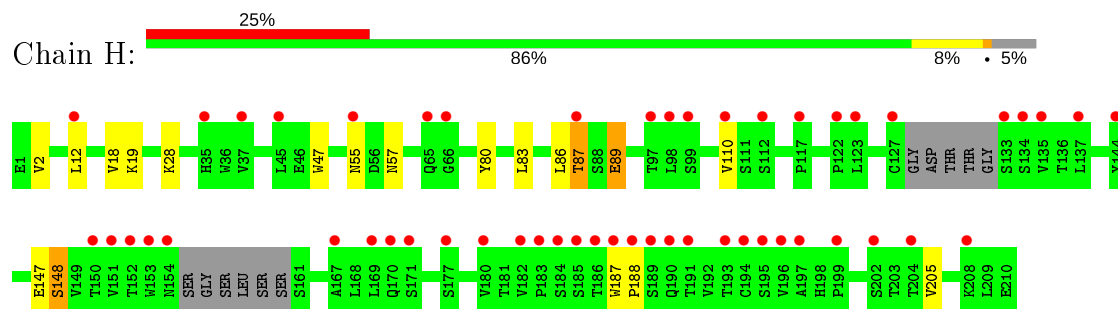
#### • Molecule 1: Apical membrane antigen 1



#### • Molecule 2: 1F9 light chain



#### • Molecule 3: 1F9 heavy chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	184.91Å 51.49Å 95.26Å 90.00° 108.49° 90.00°	Depositor
Resolution (Å)	27.95 – 2.30 27.95 – 2.30	Depositor EDS
% Data completeness (in resolution range)	100.0 (27.95-2.30) 99.4 (27.95-2.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.62 (at 2.31Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.209 , 0.256 0.202 , 0.247	Depositor DCC
$R_{free}$ test set	1899 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.8	Xtriage
Anisotropy	0.376	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 44.6	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	0.93	EDS
Total number of atoms	5598	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.05% 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	0.45	0/2363	0.58	0/3208
2	L	0.36	0/1655	0.50	0/2260
3	H	0.36	0/1495	0.52	0/2051
All	All	0.40	0/5513	0.54	0/7519

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	2303	0	2137	19	0
2	L	1616	0	1494	12	0
3	H	1459	0	1360	11	0
4	A	175	0	0	3	0
4	H	34	0	0	0	0
4	L	11	0	0	1	0
All	All	5598	0	4991	40	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 (40) 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:304:ARG:NH1	1:A:436:GLU:OE2	2.01	0.93
1:A:258:ASN:O	1:A:260:PRO:HD3	1.88	0.73
3:H:55:ASN:HB2	3:H:57:ASN:HB2	1.73	0.69
2:L:16:GLY:O	2:L:77:THR:HG22	1.93	0.69
1:A:242:ASP:HB2	1:A:286:ASN:HB3	1.83	0.61
2:L:149:LYS:HB2	2:L:193:THR:HB	1.84	0.59
1:A:429:THR:HG22	1:A:431:LEU:H	1.67	0.58
3:H:87:THR:HG22	3:H:89:GLU:H	1.73	0.53
1:A:335:PHE:HD1	1:A:434:PRO:HB2	1.75	0.52
1:A:429:THR:HB	1:A:432:SER:OG	2.10	0.52
2:L:115:VAL:HG22	2:L:136:LEU:HD22	1.93	0.50
1:A:187:GLU:HA	1:A:188:PRO:C	2.32	0.49
1:A:190:MET:HG3	3:H:2:VAL:HG21	1.94	0.49
3:H:147:GLU:O	3:H:148:SER:CB	2.61	0.48
1:A:429:THR:CG2	1:A:431:LEU:H	2.25	0.48
3:H:19:LYS:HD2	3:H:80:TYR:HB3	1.96	0.48
1:A:168:PRO:HB3	1:A:184:PRO:HA	1.95	0.47
1:A:335:PHE:CD1	1:A:434:PRO:HB2	2.49	0.47
1:A:429:THR:HG22	1:A:431:LEU:N	2.28	0.47
1:A:349:GLN:NE2	4:A:497:HOH:O	2.47	0.47
2:L:118:PHE:HB2	2:L:133:VAL:HG13	1.97	0.46
3:H:12:LEU:O	3:H:110:VAL:HA	2.17	0.45
1:A:244:ASP:HB3	1:A:246:LYS:HG3	1.99	0.44
2:L:83:LEU:HD13	2:L:106:ILE:HG12	1.99	0.44
1:A:129:VAL:HG11	1:A:349:GLN:HG2	2.00	0.43
2:L:61:ARG:NH1	2:L:82:ASP:OD1	2.52	0.42
1:A:391:LYS:NZ	4:A:591:HOH:O	2.52	0.42
2:L:136:LEU:HD11	2:L:146:VAL:HG22	2.02	0.42
2:L:151:ASP:HA	2:L:191:SER:HB3	2.02	0.41
1:A:388:ASP:N	1:A:391:LYS:HE2	2.36	0.41
4:A:440:HOH:O	3:H:28:LYS:HE3	2.19	0.41
2:L:91:GLY:HA2	2:L:96:ARG:HD3	2.01	0.41
2:L:95:PRO:HA	3:H:47:TRP:CZ3	2.56	0.41
2:L:186:TYR:HA	2:L:192:TYR:OH	2.21	0.41
3:H:187:TRP:HA	3:H:188:PRO:HA	1.87	0.41
3:H:87:THR:HG22	3:H:89:GLU:N	2.35	0.41
1:A:207:TYR:OH	1:A:296:ASP:OD1	2.32	0.40
1:A:335:PHE:CE2	1:A:339:LYS:HE2	2.57	0.40
2:L:45:LYS:HE2	4:L:215:HOH:O	2.20	0.40
3:H:12:LEU:HD13	3:H:18:VAL:HB	2.03	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	286/336 (85%)	274 (96%)	11 (4%)	1 (0%)	41	50
2	L	211/214 (99%)	205 (97%)	6 (3%)	0	100	100
3	H	193/210 (92%)	186 (96%)	6 (3%)	1 (0%)	29	35
All	All	690/760 (91%)	665 (96%)	23 (3%)	2 (0%)	41	50

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	H	148	SER
1	A	260	PRO

### 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	248/296 (84%)	239 (96%)	9 (4%)	35	49
2	L	175/191 (92%)	169 (97%)	6 (3%)	37	51
3	H	157/183 (86%)	152 (97%)	5 (3%)	39	54
All	All	580/670 (87%)	560 (97%)	20 (3%)	37	51

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	163	THR

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	195	LEU
1	A	273	MET
1	A	274	PHE
1	A	319	ASN
1	A	321	GLU
1	A	326	VAL
1	A	414	VAL
1	A	429	THR
2	L	7	THR
2	L	13	VAL
2	L	60	ASP
2	L	72	THR
2	L	78	VAL
2	L	181	LEU
3	H	83	LEU
3	H	86	LEU
3	H	87	THR
3	H	89	GLU
3	H	205	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	349	GLN
2	L	107	ASN
2	L	145	ASN
3	H	55	ASN
3	H	77	ASN

### 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

There are no ligands in this entry.

## 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	292/336 (86%)	0.96	47 (16%)  	30, 35, 48, 58	0
2	L	213/214 (99%)	1.02	37 (17%)  	33, 37, 40, 41	0
3	H	199/210 (94%)	1.36	52 (26%)  	32, 38, 43, 46	0
All	All	704/760 (92%)	1.09	136 (19%)  	30, 37, 44, 58	0

All (136) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	262	TYR	6.9
3	H	187	TRP	5.7
1	A	259	GLY	5.7
3	H	194	CYS	5.7
3	H	195	SER	5.5
1	A	105	TYR	5.4
1	A	261	ARG	5.3
3	H	193	THR	5.3
2	L	206	VAL	5.2
3	H	171	SER	5.1
1	A	390	TYR	5.0
1	A	353	TYR	4.9
3	H	65	GLN	4.9
3	H	191	THR	4.8
1	A	388	ASP	4.6
2	L	34	VAL	4.6
3	H	189	SER	4.2
3	H	153	TRP	4.1
3	H	66	GLY	4.0
2	L	200	THR	3.9
1	A	151	VAL	3.9
3	H	186	THR	3.8
1	A	289	TYR	3.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	L	198	HIS	3.8
1	A	127	ILE	3.7
3	H	152	THR	3.7
1	A	237	PRO	3.7
1	A	249	ILE	3.7
2	L	196	ALA	3.6
2	L	212	ASN	3.5
1	A	263	CYS	3.5
3	H	184	SER	3.5
3	H	185	SER	3.5
3	H	137	LEU	3.4
1	A	238	ALA	3.4
1	A	106	MET	3.4
2	L	128	GLY	3.4
1	A	221	ALA	3.4
3	H	196	VAL	3.4
3	H	110	VAL	3.4
2	L	204	PRO	3.4
1	A	430	ALA	3.3
3	H	182	VAL	3.3
3	H	117	PRO	3.3
3	H	197	ALA	3.3
2	L	205	ILE	3.3
2	L	190	ASN	3.3
2	L	203	SER	3.3
2	L	191	SER	3.2
1	A	411	ILE	3.2
1	A	160	ASN	3.2
1	A	341	VAL	3.2
3	H	37	VAL	3.2
2	L	98	PHE	3.2
3	H	144	TYR	3.2
2	L	194	CYS	3.2
2	L	197	THR	3.2
2	L	127	SER	3.2
2	L	125	LEU	3.2
3	H	169	LEU	3.1
1	A	279	ALA	3.1
3	H	167	ALA	3.1
1	A	274	PHE	3.0
3	H	150	THR	3.0
3	H	134	SER	2.9

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
3	H	202	SER	2.9
1	A	342	PHE	2.9
1	A	389	ARG	2.9
3	H	199	PRO	2.9
2	L	184	ASP	2.9
3	H	183	PRO	2.9
1	A	162	ASN	2.8
2	L	201	SER	2.8
2	L	112	ALA	2.8
1	A	428	THR	2.8
1	A	163	THR	2.7
1	A	206	LYS	2.7
3	H	133	SER	2.7
1	A	429	THR	2.6
3	H	123	LEU	2.6
2	L	151	ASP	2.6
1	A	290	LEU	2.6
3	H	204	THR	2.6
1	A	236	TYR	2.5
3	H	208	LYS	2.5
2	L	123	GLU	2.5
2	L	181	LEU	2.5
3	H	177	SER	2.5
2	L	143	ASP	2.5
1	A	260	PRO	2.4
3	H	180	VAL	2.4
3	H	190	GLN	2.4
2	L	129	GLY	2.4
3	H	99	SER	2.4
2	L	1	SER	2.4
2	L	186	TYR	2.4
1	A	431	LEU	2.4
1	A	210	ASN	2.4
1	A	139	GLY	2.4
1	A	344	LEU	2.3
3	H	55	ASN	2.3
1	A	432	SER	2.3
2	L	107	ASN	2.3
1	A	312	PHE	2.3
1	A	145	PRO	2.2
2	L	122	SER	2.2
3	H	87	THR	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	154	LYS	2.2
2	L	183	LYS	2.2
1	A	400	GLY	2.2
1	A	253	ALA	2.2
3	H	97	THR	2.2
3	H	98	LEU	2.2
2	L	142	LYS	2.2
2	L	27	GLN	2.1
1	A	201	PHE	2.1
2	L	126	THR	2.1
3	H	12	LEU	2.1
3	H	45	LEU	2.1
3	H	151	VAL	2.1
3	H	154	ASN	2.1
2	L	36	TYR	2.1
3	H	112	SER	2.1
1	A	273	MET	2.1
3	H	170	GLN	2.1
3	H	127	CYS	2.1
3	H	135	VAL	2.1
2	L	91	GLY	2.1
2	L	168	SER	2.1
2	L	207	LYS	2.1
3	H	188	PRO	2.1
3	H	122	PRO	2.1
1	A	140	THR	2.0
1	A	161	SER	2.0
1	A	178	ASP	2.0
3	H	35	HIS	2.0

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

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

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.