



# Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 10:47 AM GMT

PDB ID : 4H5O  
Title : Crystal Structure of Rift Valley Fever Virus Nucleocapsid Protein Pentamer  
Bound to Single-stranded RNA  
Authors : Raymond, D.D.; Smith, J.L.  
Deposited on : 2012-09-18  
Resolution : 3.90 Å(reported)

This is a full wwPDB 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/ValidationPDFNotes.html>

---

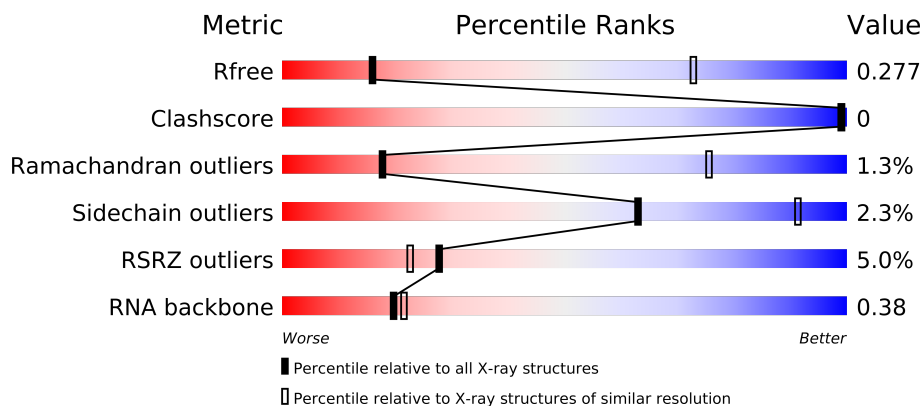
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 3.90 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))
$R_{free}$	66092	1022 (4.38-3.42)
Clashscore	79885	1173 (4.30-3.50)
Ramachandran outliers	78287	1118 (4.30-3.50)
Sidechain outliers	78261	1107 (4.30-3.50)
RSRZ outliers	66119	1000 (4.36-3.44)
RNA backbone	1838	1018 (5.00-2.80)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	245	
1	B	245	
1	C	245	
1	D	245	
1	E	245	
1	F	245	
1	G	245	
1	H	245	
1	I	245	
1	J	245	
2	K	35	
2	L	35	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 20484 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Nucleocapsid protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	242	Total	C	N	O	S	0	0	0
			1894	1197	340	345	12			
1	B	244	Total	C	N	O	S	0	0	0
			1910	1205	343	350	12			
1	C	244	Total	C	N	O	S	0	0	0
			1910	1205	343	350	12			
1	D	244	Total	C	N	O	S	0	0	0
			1910	1205	343	350	12			
1	E	244	Total	C	N	O	S	0	0	0
			1910	1205	343	350	12			
1	F	244	Total	C	N	O	S	0	0	0
			1910	1205	343	350	12			
1	G	244	Total	C	N	O	S	0	0	0
			1910	1205	343	350	12			
1	H	244	Total	C	N	O	S	0	0	0
			1910	1205	343	350	12			
1	I	244	Total	C	N	O	S	0	0	0
			1910	1205	343	350	12			
1	J	244	Total	C	N	O	S	0	0	0
			1910	1205	343	350	12			

- Molecule 2 is a RNA chain called 35-mer poly(U) RNA.

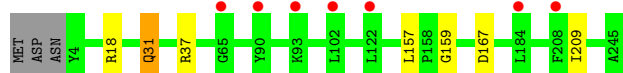
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	K	35	Total	C	N	O	P	0	0	0
			700	315	70	280	35			
2	L	35	Total	C	N	O	P	0	0	0
			700	315	70	280	35			

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

Chain A: 



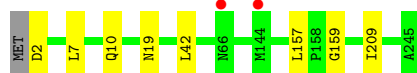
- Molecule 1: Nucleocapsid protein

Chain B: 



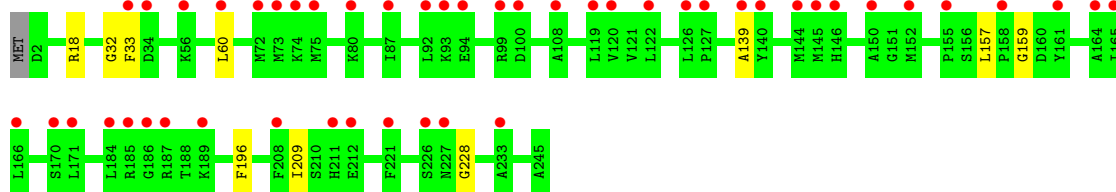
- Molecule 1: Nucleocapsid protein

Chain C: 



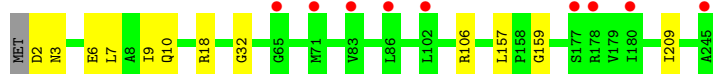
- Molecule 1: Nucleocapsid protein

Chain D: 



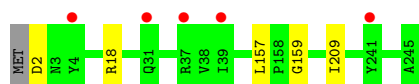
- Molecule 1: Nucleocapsid protein

Chain E: 



- Molecule 1: Nucleocapsid protein

Chain F: 



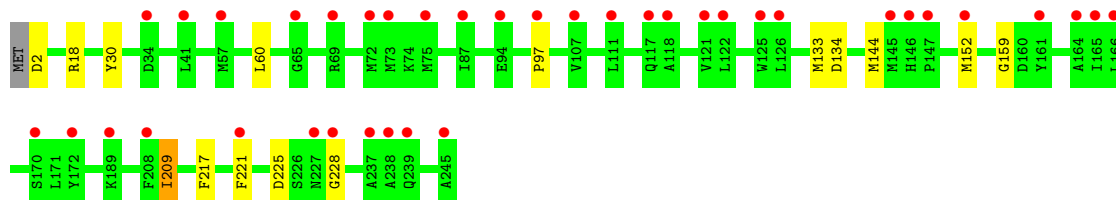
- Molecule 1: Nucleocapsid protein

Chain G:



- Molecule 1: Nucleocapsid protein

Chain H:



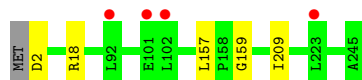
- Molecule 1: Nucleocapsid protein

Chain I:



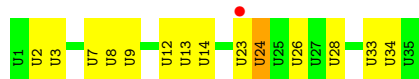
- Molecule 1: Nucleocapsid protein

Chain J:



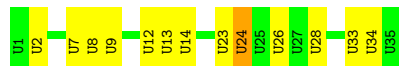
- Molecule 2: 35-mer poly(U) RNA

Chain K:



- Molecule 2: 35-mer poly(U) RNA

Chain L:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	79.81Å 93.59Å 124.78Å 101.70° 90.27° 114.18°	Depositor
Resolution (Å)	41.60 – 3.90 41.60 – 3.90	Depositor EDS
% Data completeness (in resolution range)	99.2 (41.60-3.90) 94.0 (41.60-3.90)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.21	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.85 (at 3.88Å)	Xtriage
Refinement program	BUSTER 2.10	Depositor
R, $R_{free}$	0.228 , 0.248 0.260 , 0.277	Depositor DCC
$R_{free}$ test set	1459 reflections (5.31%)	DCC
Wilson B-factor (Å <sup>2</sup> )	125.0	Xtriage
Anisotropy	0.251	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 34.4	EDS
Estimated twinning fraction	0.280 for h,-h-k,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.39$ , $\langle L^2 \rangle = 0.21$	Xtriage
Outliers	0 of 28959 reflections	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	20484	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.73% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 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	0.39	0/1931	0.55	0/2604
1	B	0.38	0/1947	0.54	0/2626
1	C	0.38	0/1947	0.55	0/2626
1	D	0.41	0/1947	0.56	0/2626
1	E	0.38	0/1947	0.54	0/2626
1	F	0.37	0/1947	0.53	0/2626
1	G	0.38	0/1947	0.54	0/2626
1	H	0.42	0/1947	0.55	0/2626
1	I	0.39	0/1947	0.54	0/2626
1	J	0.38	0/1947	0.53	0/2626
2	K	1.31	3/769 (0.4%)	0.95	0/1186
2	L	1.30	3/769 (0.4%)	0.96	0/1186
All	All	0.52	6/20992 (0.0%)	0.59	0/28610

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	K	9	U	C1'-N1	6.02	1.57	1.48
2	K	2	U	C1'-N1	5.98	1.57	1.48
2	L	2	U	C1'-N1	5.80	1.57	1.48
2	L	24	U	C1'-N1	5.80	1.57	1.48
2	K	24	U	C1'-N1	5.74	1.57	1.48
2	L	9	U	C1'-N1	5.54	1.57	1.48

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens

added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1894	0	8	0	0
1	B	1910	0	18	1	0
1	C	1910	0	18	0	0
1	D	1910	0	18	0	0
1	E	1910	0	18	2	0
1	F	1910	0	18	0	0
1	G	1910	0	18	0	0
1	H	1910	0	18	0	0
1	I	1910	0	18	1	0
1	J	1910	0	18	0	0
2	K	700	0	351	1	0
2	L	700	0	351	1	0
All	All	20484	0	872	4	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 0.

All (4) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:I:199:PRO:CB	2:L:23:U:H5'	2.36	0.55
1:B:199:PRO:CB	2:K:23:U:H5'	2.42	0.49
1:E:32:GLY:O	1:E:106:ARG:NH2	2.48	0.47
1:E:3:ASN:HD22	1:E:6:GLU:CD	2.22	0.43

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	240/245 (98%)	222 (92%)	15 (6%)	3 (1%)	18 75

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	242/245 (99%)	225 (93%)	13 (5%)	4 (2%)	14	71
1	C	242/245 (99%)	226 (93%)	14 (6%)	2 (1%)	27	83
1	D	242/245 (99%)	220 (91%)	16 (7%)	6 (2%)	9	62
1	E	242/245 (99%)	226 (93%)	14 (6%)	2 (1%)	27	83
1	F	242/245 (99%)	227 (94%)	13 (5%)	2 (1%)	27	83
1	G	242/245 (99%)	227 (94%)	13 (5%)	2 (1%)	27	83
1	H	242/245 (99%)	219 (90%)	17 (7%)	6 (2%)	9	62
1	I	242/245 (99%)	226 (93%)	14 (6%)	2 (1%)	27	83
1	J	242/245 (99%)	227 (94%)	13 (5%)	2 (1%)	27	83
All	All	2418/2450 (99%)	2245 (93%)	142 (6%)	31 (1%)	18	75

All (31) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	H	30	TYR
1	A	31	GLN
1	A	159	GLY
1	B	159	GLY
1	C	159	GLY
1	D	159	GLY
1	E	159	GLY
1	F	159	GLY
1	G	159	GLY
1	I	159	GLY
1	J	159	GLY
1	H	228	GLY
1	A	209	ILE
1	B	33	PHE
1	B	34	ASP
1	B	209	ILE
1	C	209	ILE
1	E	209	ILE
1	F	209	ILE
1	G	209	ILE
1	I	209	ILE
1	J	209	ILE
1	D	33	PHE
1	D	209	ILE
1	D	139	ALA

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	228	GLY
1	H	209	ILE
1	H	225	ASP
1	H	159	GLY
1	D	32	GLY
1	H	97	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	195/198 (98%)	190 (97%)	5 (3%)	59	90
1	B	197/198 (100%)	195 (99%)	2 (1%)	85	96
1	C	197/198 (100%)	191 (97%)	6 (3%)	53	89
1	D	197/198 (100%)	193 (98%)	4 (2%)	68	92
1	E	197/198 (100%)	191 (97%)	6 (3%)	53	89
1	F	197/198 (100%)	194 (98%)	3 (2%)	76	95
1	G	197/198 (100%)	193 (98%)	4 (2%)	68	92
1	H	197/198 (100%)	187 (95%)	10 (5%)	33	80
1	I	197/198 (100%)	195 (99%)	2 (1%)	85	96
1	J	197/198 (100%)	194 (98%)	3 (2%)	76	95
All	All	1968/1980 (99%)	1923 (98%)	45 (2%)	63	91

All (45) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	18	ARG
1	A	31	GLN
1	A	37	ARG
1	A	157	LEU
1	A	167	ASP
1	B	18	ARG
1	B	157	LEU
1	C	2	ASP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	7	LEU
1	C	10	GLN
1	C	19	ASN
1	C	42	LEU
1	C	157	LEU
1	D	18	ARG
1	D	60	LEU
1	D	157	LEU
1	D	196	PHE
1	E	2	ASP
1	E	7	LEU
1	E	9	ILE
1	E	10	GLN
1	E	18	ARG
1	E	157	LEU
1	F	2	ASP
1	F	18	ARG
1	F	157	LEU
1	G	2	ASP
1	G	18	ARG
1	G	157	LEU
1	G	231	SER
1	H	2	ASP
1	H	18	ARG
1	H	60	LEU
1	H	133	MET
1	H	134	ASP
1	H	144	MET
1	H	152	MET
1	H	209	ILE
1	H	217	PHE
1	H	221	PHE
1	I	18	ARG
1	I	157	LEU
1	J	2	ASP
1	J	18	ARG
1	J	157	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	K	34/35 (97%)	9 (26%)	3 (8%)
2	L	34/35 (97%)	8 (23%)	3 (8%)
All	All	68/70 (97%)	17 (25%)	6 (8%)

All (17) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	K	3	U
2	K	7	U
2	K	8	U
2	K	13	U
2	K	14	U
2	K	24	U
2	K	28	U
2	K	33	U
2	K	34	U
2	L	7	U
2	L	8	U
2	L	13	U
2	L	14	U
2	L	24	U
2	L	28	U
2	L	33	U
2	L	34	U

All (6) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	K	12	U
2	K	26	U
2	K	33	U
2	L	12	U
2	L	26	U
2	L	33	U

## 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	242/245 (98%)	0.50	7 (2%) 49 37	10, 53, 124, 161	0
1	B	244/245 (99%)	0.38	3 (1%) 75 59	4, 22, 67, 91	0
1	C	244/245 (99%)	0.37	2 (0%) 83 69	6, 22, 70, 114	0
1	D	244/245 (99%)	1.05	48 (19%) 2 3	23, 120, 216, 266	0
1	E	244/245 (99%)	0.45	9 (3%) 39 31	18, 50, 122, 203	0
1	F	244/245 (99%)	0.42	5 (2%) 62 46	6, 36, 104, 149	0
1	G	244/245 (99%)	0.40	2 (0%) 83 69	4, 17, 77, 129	0
1	H	244/245 (99%)	0.92	38 (15%) 3 4	23, 113, 187, 216	0
1	I	244/245 (99%)	0.40	6 (2%) 54 41	4, 19, 76, 121	0
1	J	244/245 (99%)	0.42	4 (1%) 68 53	4, 36, 110, 185	0
2	K	35/35 (100%)	0.36	1 (2%) 49 37	31, 37, 44, 44	0
2	L	35/35 (100%)	0.29	0 100 100	31, 37, 44, 44	0
All	All	2508/2520 (99%)	0.52	125 (4%) 28 23	4, 40, 150, 266	0

All (125) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	165	ILE	6.3
1	D	164	ALA	6.0
1	D	165	ILE	5.6
1	H	34	ASP	5.2
1	D	34	ASP	5.1
1	H	227	ASN	4.9
1	D	187	ARG	4.8
1	H	164	ALA	4.6
1	D	185	ARG	4.5
1	D	33	PHE	4.3
1	D	212	GLU	4.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	H	245	ALA	4.1
1	H	221	PHE	3.8
1	J	102	LEU	3.8
1	H	238	ALA	3.7
1	D	166	LEU	3.6
1	H	73	MET	3.6
1	D	227	ASN	3.6
1	H	172	TYR	3.6
1	D	211	HIS	3.5
1	D	75	MET	3.5
1	A	102	LEU	3.4
1	D	146	HIS	3.4
1	D	161	TYR	3.3
1	H	161	TYR	3.2
1	D	184	LEU	3.2
1	H	145	MET	3.2
1	D	122	LEU	3.2
1	H	75	MET	3.2
1	D	73	MET	3.1
1	D	152	MET	3.1
1	D	56	LYS	3.1
1	H	117	GLN	3.1
1	G	133	MET	3.1
1	D	155	PRO	3.0
1	D	100	ASP	3.0
1	D	126	LEU	3.0
1	E	71	MET	3.0
1	H	152	MET	3.0
1	H	189	LYS	2.9
1	H	121	VAL	2.9
1	D	99	ARG	2.9
1	I	99	ARG	2.9
1	B	99	ARG	2.9
1	A	90	TYR	2.9
1	H	122	LEU	2.9
1	D	189	LYS	2.9
1	D	74	LYS	2.9
1	D	93	LYS	2.9
2	K	23	U	2.8
1	D	72	MET	2.8
1	H	118	ALA	2.8
1	D	158	PRO	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	180	ILE	2.8
1	D	127	PRO	2.8
1	H	208	PHE	2.7
1	D	171	LEU	2.7
1	E	245	ALA	2.7
1	H	228	GLY	2.7
1	C	144	MET	2.7
1	D	140	TYR	2.7
1	J	92	LEU	2.7
1	H	147	PRO	2.6
1	D	92	LEU	2.6
1	D	226	SER	2.6
1	H	166	LEU	2.6
1	E	65	GLY	2.5
1	B	100	ASP	2.5
1	D	208	PHE	2.5
1	A	93	LYS	2.5
1	E	83	VAL	2.5
1	H	41	LEU	2.5
1	I	28	PHE	2.5
1	H	72	MET	2.5
1	H	126	LEU	2.5
1	I	144	MET	2.4
1	J	223	LEU	2.4
1	A	122	LEU	2.4
1	I	180	ILE	2.4
1	D	139	ALA	2.4
1	H	111	LEU	2.4
1	C	66	ASN	2.4
1	D	233	ALA	2.4
1	F	37	ARG	2.4
1	D	221	PHE	2.4
1	A	65	GLY	2.3
1	H	94	GLU	2.3
1	A	184	LEU	2.3
1	D	144	MET	2.3
1	D	108	ALA	2.3
1	H	237	ALA	2.3
1	D	87	ILE	2.2
1	H	239	GLN	2.2
1	H	107	VAL	2.2
1	H	125	TRP	2.2

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	I	150	ALA	2.2
1	D	120	VAL	2.2
1	H	146	HIS	2.2
1	F	4	TYR	2.2
1	H	87	ILE	2.2
1	E	178	ARG	2.2
1	J	101	GLU	2.2
1	D	80	LYS	2.2
1	D	145	MET	2.2
1	D	119	LEU	2.2
1	D	150	ALA	2.1
1	H	57	MET	2.1
1	E	86	LEU	2.1
1	F	241	TYR	2.1
1	H	69	ARG	2.1
1	H	170	SER	2.1
1	E	180	ILE	2.1
1	H	65	GLY	2.1
1	F	31	GLN	2.1
1	G	144	MET	2.1
1	E	102	LEU	2.1
1	D	170	SER	2.1
1	D	94	GLU	2.1
1	H	97	PRO	2.1
1	D	60	LEU	2.1
1	D	186	GLY	2.1
1	F	39	ILE	2.0
1	A	208	PHE	2.0
1	E	177	SER	2.0
1	I	31	GLN	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.