Full wwPDB X-ray Structure Validation Report (i)

Sep 24, 2023 – 05:21 PM EDT

PDB ID	:	5USH
Title	:	Structure of vaccinia virus D8 protein bound to human Fab vv66
Authors	:	Zajonc, D.M.
Deposited on	:	2017-02-13
Resolution	:	2.30 Å(reported)

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

|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
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)	:	2.35.1
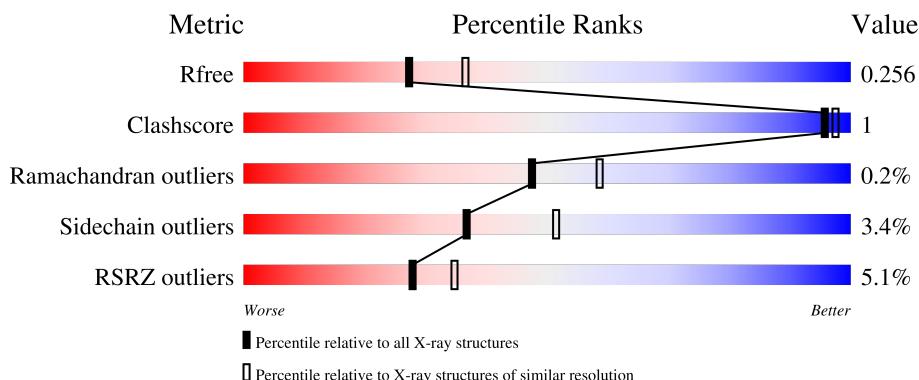
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION
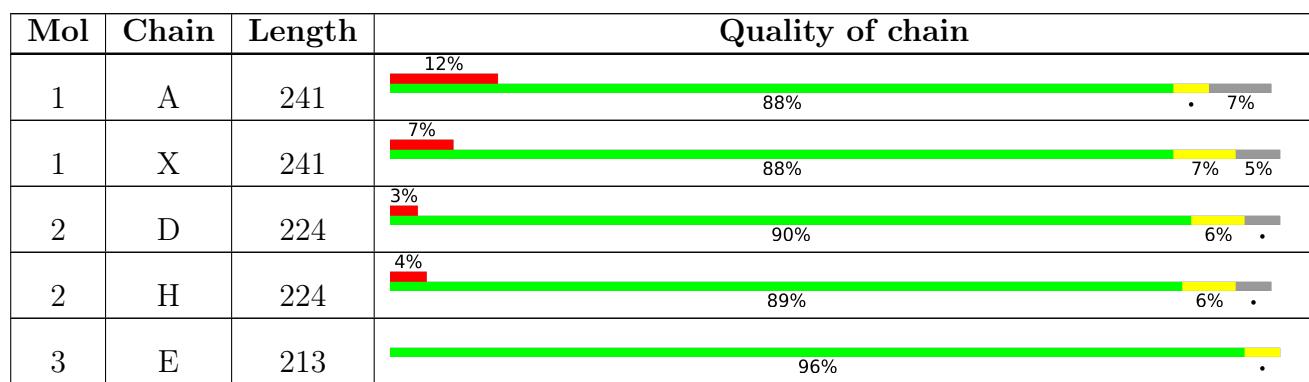
The reported resolution of this entry is 2.30 Å.

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	Percentile Ranks	Value
Rfree		0.256
Clashscore		1
Ramachandran outliers		0.2%
Sidechain outliers		3.4%
RSRZ outliers		5.1%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

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 of the lower bar indicate the fraction of residues that contain outliers for >=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 <=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	241	12% / 88% · 7%
1	X	241	7% / 88% 7% 5%
2	D	224	3% / 90% 6% ·
2	H	224	4% / 89% 6% ·
3	E	213	96% ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	L	213	

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 10202 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 IMV membrane protein.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	A	223	Total	C	N	O	0	0	0
			1792	1159	289	344			
1	X	230	Total	C	N	O	0	0	0
			1867	1206	306	355			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	235	GLY	-	expression tag	UNP Q1M1K6
A	236	HIS	-	expression tag	UNP Q1M1K6
A	237	HIS	-	expression tag	UNP Q1M1K6
A	238	HIS	-	expression tag	UNP Q1M1K6
A	239	HIS	-	expression tag	UNP Q1M1K6
A	240	HIS	-	expression tag	UNP Q1M1K6
A	241	HIS	-	expression tag	UNP Q1M1K6
X	235	GLY	-	expression tag	UNP Q1M1K6
X	236	HIS	-	expression tag	UNP Q1M1K6
X	237	HIS	-	expression tag	UNP Q1M1K6
X	238	HIS	-	expression tag	UNP Q1M1K6
X	239	HIS	-	expression tag	UNP Q1M1K6
X	240	HIS	-	expression tag	UNP Q1M1K6
X	241	HIS	-	expression tag	UNP Q1M1K6

- Molecule 2 is a protein called Fab vv66 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	216	Total	C	N	O	S	0	0	0
			1634	1044	268	318	4			
2	H	214	Total	C	N	O	S	0	0	0
			1630	1042	268	316	4			

- Molecule 3 is a protein called Fab vv66 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	212	Total C N O S					0	0	0
			1555 972 261 318 4							
3	L	212	Total C N O S					0	0	0
			1550 968 261 317 4							
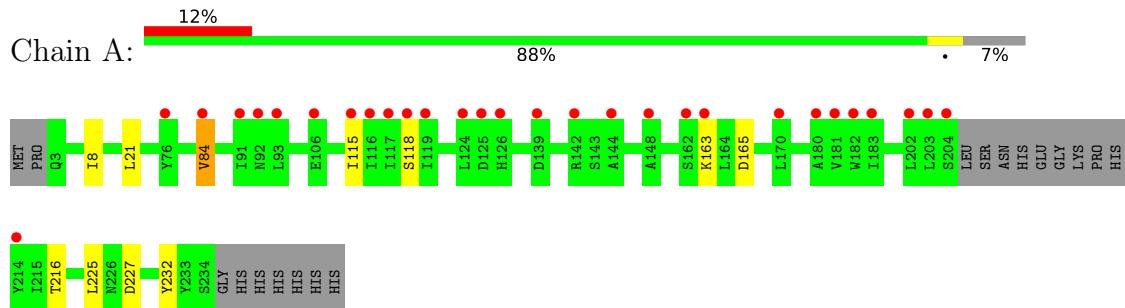
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	12	Total O		0	0
			12 12			
4	D	40	Total O		0	0
			40 40			
4	E	38	Total O		0	0
			38 38			
4	H	28	Total O		0	0
			28 28			
4	L	23	Total O		0	0
			23 23			
4	X	33	Total O		0	0
			33 33			
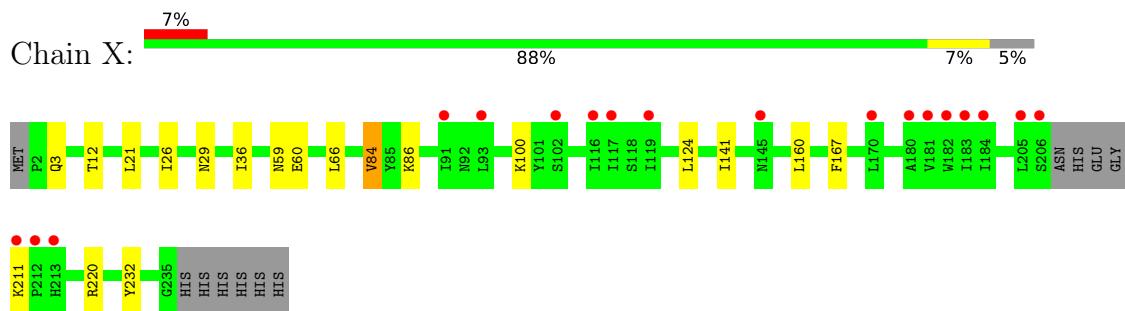
3 Residue-property plots ⓘ

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.
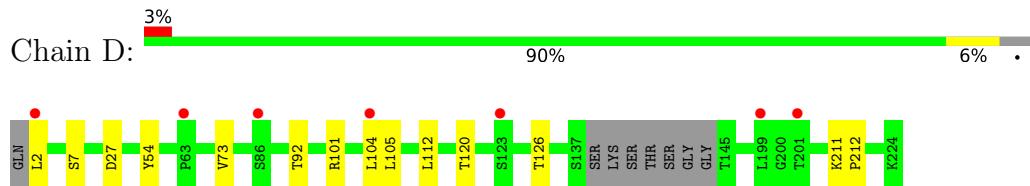
- Molecule 1: IMV membrane protein
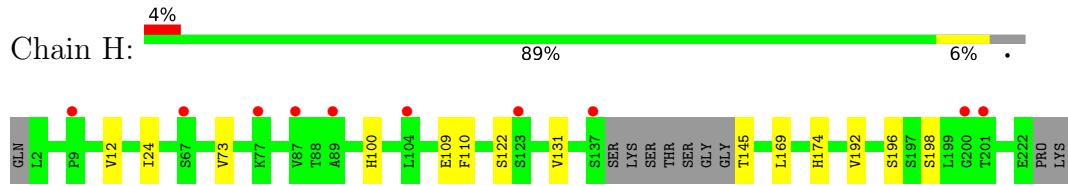
- Molecule 1: IMV membrane protein
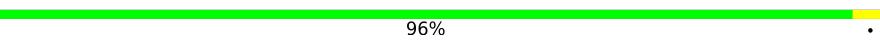
- Molecule 2: Fab vv66 heavy chain

- Molecule 2: Fab vv66 heavy chain

- Molecule 3: Fab vv66 light chain

Chain E:

96%

• Molecule 3: Fab vv66 light chain

Chain L:

%

96%

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	55.86Å 61.52Å 444.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.51 – 2.30 38.51 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.9 (38.51-2.30) 99.9 (38.51-2.30)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.41 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.218 , 0.257 0.220 , 0.256	Depositor DCC
R_{free} test set	3518 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	47.0	Xtriage
Anisotropy	0.095	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 25.6	EDS
L-test for twinning[2]	$< |L| > = 0.45, < L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	10202	wwPDB-VP
Average B, all atoms (Å^2)	54.0	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.*

---

[1] Intensities estimated from amplitudes.
[2] Theoretical values of $< |L| >$, $< L^2 >$ 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	0.41	0/1840	0.55	0/2510
1	X	0.40	0/1918	0.61	0/2610
2	D	0.39	0/1678	0.63	0/2297
2	H	0.40	0/1673	0.62	0/2286
3	E	0.40	0/1593	0.60	0/2177
3	L	0.39	0/1588	0.59	0/2172
All	All	0.40	0/10290	0.60	0/14052

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	1792	0	1701	4	0
1	X	1867	0	1805	5	0
2	D	1634	0	1594	5	0
2	H	1630	0	1607	5	0
3	E	1555	0	1497	2	0
3	L	1550	0	1486	1	0
4	A	12	0	0	0	0
4	D	40	0	0	0	0
4	E	38	0	0	0	0
4	H	28	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	L	23	0	0	0	0
4	X	33	0	0	0	0
All	All	10202	0	9690	21	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 1.

All (21) 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)
2:D:2:LEU:HD23	2:D:112:LEU:HD12	1.86	0.58
2:D:2:LEU:HD23	2:D:112:LEU:CD1	2.39	0.51
1:X:21:LEU:HD11	1:X:232:TYR:HB3	1.93	0.51
1:X:26:ILE:HG21	1:X:160:LEU:HD11	1.97	0.46
1:X:84:VAL:O	1:X:84:VAL:HG12	2.17	0.45
2:H:100:HIS:HB3	2:H:110:PHE:HA	2.00	0.44
2:D:211:LYS:N	2:D:212:PRO:CD	2.81	0.43
3:E:80:LEU:HD21	3:E:108:LEU:HD21	2.00	0.43
1:A:21:LEU:HD11	1:A:232:TYR:HB3	2.01	0.43
1:A:84:VAL:O	1:A:84:VAL:HG12	2.19	0.43
2:H:169:LEU:HD23	2:H:192:VAL:HG21	2.01	0.42
3:L:85:GLU:OE2	3:L:171:LYS:NZ	2.52	0.42
1:A:8:ILE:HD13	1:A:225:LEU:HD13	2.01	0.42
2:H:145:THR:HA	2:H:196:SER:HB2	2.02	0.42
1:X:12:THR:HG22	1:X:167:PHE:CZ	2.54	0.41
2:D:92:THR:HG23	2:D:120:THR:HA	2.02	0.41
2:H:24:ILE:HD12	2:H:24:ILE:N	2.35	0.41
1:A:115:ILE:HD11	2:D:105:LEU:HD11	2.01	0.41
1:X:36:ILE:HG12	1:X:141:ILE:HD13	2.03	0.41
3:E:168:LYS:HA	3:E:168:LYS:HD3	1.92	0.40
2:H:100:HIS:HB2	2:H:109:GLU:O	2.21	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	219/241 (91%)	209 (95%)	9 (4%)	1 (0%)	29	35
1	X	226/241 (94%)	217 (96%)	8 (4%)	1 (0%)	34	42
2	D	212/224 (95%)	209 (99%)	2 (1%)	1 (0%)	29	35
2	H	210/224 (94%)	205 (98%)	5 (2%)	0	100	100
3	E	210/213 (99%)	205 (98%)	5 (2%)	0	100	100
3	L	210/213 (99%)	207 (99%)	3 (1%)	0	100	100
All	All	1287/1356 (95%)	1252 (97%)	32 (2%)	3 (0%)	47	58

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	104	LEU
1	X	84	VAL
1	A	84	VAL

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	196/222 (88%)	191 (97%)	5 (3%)	46	63
1	X	208/222 (94%)	198 (95%)	10 (5%)	25	36
2	D	186/195 (95%)	180 (97%)	6 (3%)	39	54
2	H	187/195 (96%)	181 (97%)	6 (3%)	39	54
3	E	171/174 (98%)	166 (97%)	5 (3%)	42	58
3	L	170/174 (98%)	164 (96%)	6 (4%)	36	50
All	All	1118/1182 (95%)	1080 (97%)	38 (3%)	37	51

All (38) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	118	SER
1	A	163	LYS
1	A	165	ASP
1	A	216	THR
1	A	227	ASP
2	D	7	SER
2	D	27	ASP
2	D	54	TYR
2	D	73	VAL
2	D	101	ARG
2	D	126	THR
3	E	62	ASP
3	E	74	SER
3	E	184	SER
3	E	192	SER
3	E	205	SER
2	H	12	VAL
2	H	73	VAL
2	H	122	SER
2	H	131	VAL
2	H	174	HIS
2	H	198	SER
3	L	58	SER
3	L	62	ASP
3	L	65	SER
3	L	151	VAL
3	L	180	SER
3	L	192	SER
1	X	3	GLN
1	X	29	ASN
1	X	59	ASN
1	X	60	GLU
1	X	66	LEU
1	X	86	LYS
1	X	100	LYS
1	X	124	LEU
1	X	211	LYS
1	X	220	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
2	D	100	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	H	100	HIS
1	X	3	GLN
1	X	95	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 monosaccharides 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, 95th 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(Å²)	Q<0.9
1	A	223/241 (92%)	0.69	29 (13%)	3	4	51, 71, 93, 104	0
1	X	230/241 (95%)	0.35	18 (7%)	13	17	31, 45, 67, 94	0
2	D	216/224 (96%)	-0.00	7 (3%)	47	54	34, 51, 72, 85	0
2	H	214/224 (95%)	0.20	10 (4%)	31	38	36, 53, 82, 92	0
3	E	212/213 (99%)	-0.24	1 (0%)	91	94	34, 48, 61, 64	0
3	L	212/213 (99%)	-0.14	2 (0%)	84	88	31, 51, 76, 94	0
All	All	1307/1356 (96%)	0.15	67 (5%)	28	35	31, 52, 82, 104	0

All (67) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	203	LEU	6.3
1	X	205	LEU	5.0
1	A	116	ILE	4.0
1	X	211	LYS	4.0
1	A	181	VAL	4.0
1	X	212	PRO	3.8
1	A	117	ILE	3.8
1	X	206	SER	3.6
1	A	214	TYR	3.6
1	X	117	ILE	3.5
1	X	116	ILE	3.4
3	L	173	SER	3.3
2	H	77	LYS	3.2
2	D	104	LEU	3.2
1	A	124	LEU	3.2
1	A	125	ASP	3.2
1	A	93	LEU	3.0
2	D	201	THR	3.0
1	A	142	ARG	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	170	LEU	3.0
1	X	180	ALA	2.9
1	X	182	TRP	2.9
1	A	144	ALA	2.9
1	X	119	ILE	2.8
1	A	115	ILE	2.8
1	X	183	ILE	2.8
3	L	172	GLN	2.8
1	A	163	LYS	2.7
2	H	89	ALA	2.7
1	A	182	TRP	2.7
1	A	139	ASP	2.7
1	X	91	ILE	2.7
2	H	137	SER	2.7
2	H	104	LEU	2.7
1	A	91	ILE	2.6
1	X	145	ASN	2.6
1	A	118	SER	2.6
1	A	180	ALA	2.5
1	A	84	VAL	2.5
1	A	76	TYR	2.5
1	X	213	HIS	2.5
2	D	123	SER	2.4
1	A	183	ILE	2.4
1	X	170	LEU	2.4
2	H	87	VAL	2.4
1	A	126	HIS	2.4
1	A	162	SER	2.4
1	A	106	GLU	2.3
1	X	102	SER	2.3
1	A	92	ASN	2.3
1	X	181	VAL	2.3
2	D	63	PRO	2.3
2	H	9	PRO	2.3
1	A	119	ILE	2.3
3	E	59	GLY	2.2
1	A	148	ALA	2.2
1	A	202	LEU	2.1
1	A	204	SER	2.1
2	H	67	SER	2.1
2	H	123	SER	2.1
2	D	86	SER	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	X	93	LEU	2.1
2	D	2	LEU	2.1
2	H	200	GLY	2.1
1	X	184	ILE	2.0
2	D	199	LEU	2.0
2	H	201	THR	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 monosaccharides 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.