Full wwPDB X-ray Structure Validation Report ⓘ

Oct 15, 2023 – 04:29 AM EDT

PDB ID	:	7UFW
Title	:	Pfs230 D1D2 in complex with LMIV230-01
Authors	:	Tang, W.K.; Tolia, N.H.
Deposited on	:	2022-03-23
Resolution	:	2.10 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

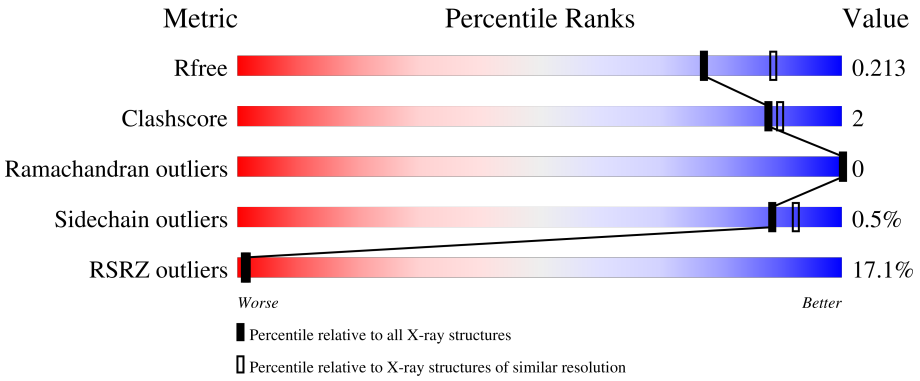
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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	H	261	9% 81% 7% 12%
1	I	261	8% 81% 6% 13%
2	A	355	18% 83% 5% 12%
2	B	355	22% 86% . 10%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 17571 atoms, of which 8579 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 LMIV230-01.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	H	229	Total	C	H	N	O	S	0	0	0
			3483	1110	1728	299	339	7			
1	I	228	Total	C	H	N	O	S	0	0	0
			3476	1108	1725	298	338	7			

- Molecule 2 is a protein called Gametocyte surface protein P230.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	A	314	Total	C	H	N	O	S	0	0	0
			5082	1633	2545	396	496	12			
2	B	320	Total	C	H	N	O	S	0	0	0
			5164	1659	2581	405	507	12			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	540	THR	-	expression tag	UNP P68874
A	541	GLY	-	expression tag	UNP P68874
A	887	GLY	-	expression tag	UNP P68874
A	888	THR	-	expression tag	UNP P68874
A	889	HIS	-	expression tag	UNP P68874
A	890	HIS	-	expression tag	UNP P68874
A	891	HIS	-	expression tag	UNP P68874
A	892	HIS	-	expression tag	UNP P68874
A	893	HIS	-	expression tag	UNP P68874
A	894	HIS	-	expression tag	UNP P68874
B	540	THR	-	expression tag	UNP P68874
B	541	GLY	-	expression tag	UNP P68874
B	887	GLY	-	expression tag	UNP P68874
B	888	THR	-	expression tag	UNP P68874
B	889	HIS	-	expression tag	UNP P68874
B	890	HIS	-	expression tag	UNP P68874

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	891	HIS	-	expression tag	UNP P68874
B	892	HIS	-	expression tag	UNP P68874
B	893	HIS	-	expression tag	UNP P68874
B	894	HIS	-	expression tag	UNP P68874

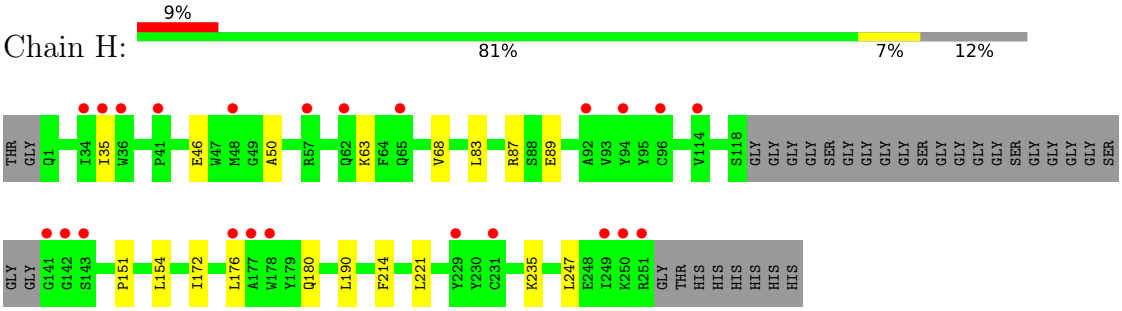
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	H	111	Total O 111 111	0	0
3	I	85	Total O 85 85	0	0
3	A	99	Total O 99 99	0	0
3	B	71	Total O 71 71	0	0

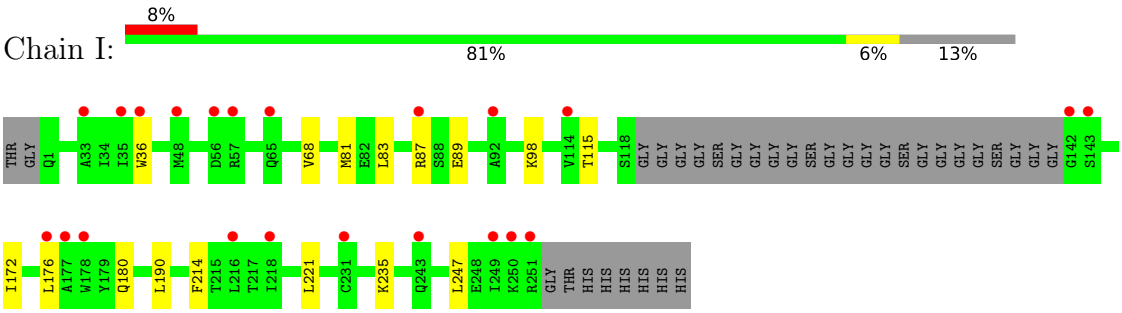
3 Residue-property plots (i)

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.

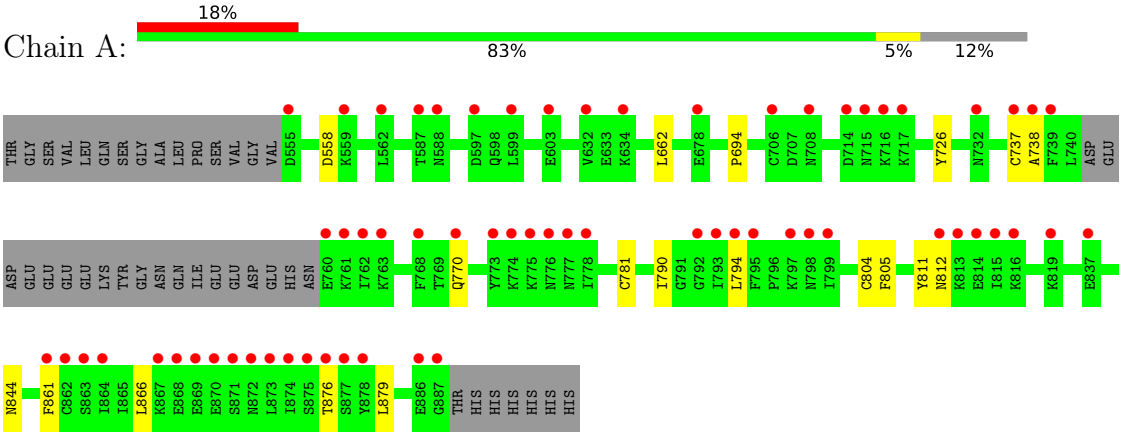
• Molecule 1: LMIV230-01



• Molecule 1: LMIV230-01

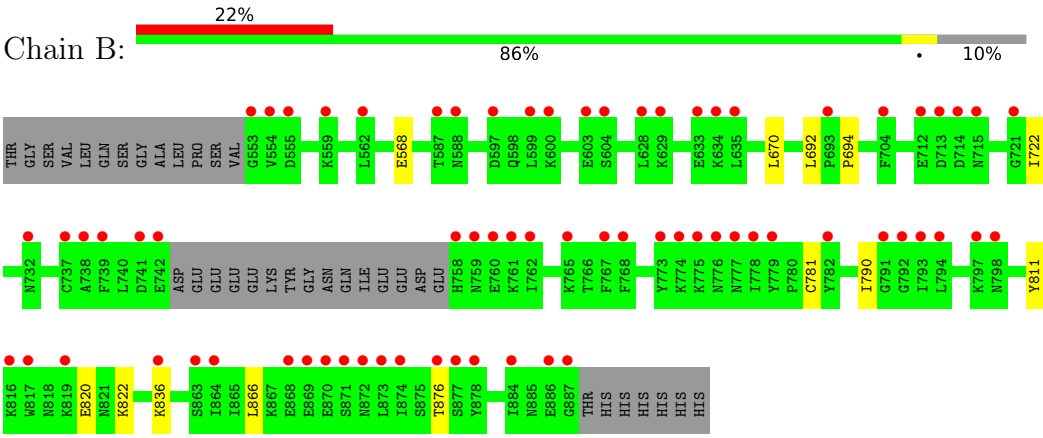


• Molecule 2: Gametocyte surface protein P230



• Molecule 2: Gametocyte surface protein P230

Chain B:


22%
86%
10%
THR GLY SER VAL LEU GLN SER GLY ALA LEU PRO SER VAL G553 V554 D555 K559 L562 E568 T587 N588 D597 Q598 L599 K600 E603 S604 L628 K629 E633 K634 L635 L670 L692 P693 P694 F704 E712 D713 D714 N715 G721 I722 V723 E724 V725 Y726
N732 C737 A738 F739 L740 D741 E742 ASP GLU GLU GLU GLU LYS TYR GLY ASN GLN ILE GLU GLU ASP GLU H758 N759 E760 K761 I762 K765 T766 F767 F768 Y773 K774 K775 N776 N777 I778 Y779 P780 C781 Y782 I790 G791 G792 I793 L794 K797 N798 Y811 N812 K813 E814 I815
K816 W817 N818 K819 E820 N821 K822 K836 S863 I864 I865 L866 K867 E868 E869 E870 S871 N872 L873 I874 S875 T876 S877 Y878 I884 N885 E886 G887 THR HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	44.44Å 151.37Å 99.04Å 90.00° 92.15° 90.00°	Depositor
Resolution (Å)	19.92 – 2.10 19.92 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.5 (19.92-2.10) 99.6 (19.92-2.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.08 (at 2.09Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.190 , 0.213 0.190 , 0.213	Depositor DCC
R_{free} test set	3783 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	39.5	Xtriage
Anisotropy	0.714	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 47.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.094 for h,-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	17571	wwPDB-VP
Average B, all atoms (Å^2)	63.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.06% 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	H	0.28	0/1794	0.55	0/2428
1	I	0.27	0/1790	0.53	0/2423
2	A	0.27	0/2587	0.51	0/3493
2	B	0.27	0/2634	0.52	0/3557
All	All	0.27	0/8805	0.52	0/11901

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	H	1755	1728	1730	9	0
1	I	1751	1725	1727	8	0
2	A	2537	2545	2545	10	0
2	B	2583	2581	2580	8	0
3	A	99	0	0	1	0
3	B	71	0	0	1	0
3	H	111	0	0	0	0
3	I	85	0	0	1	0
All	All	8992	8579	8582	35	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 2.

All (35) 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:B:866:LEU:HD12	2:B:876:THR:HB	1.60	0.83
1:I:221:LEU:HD21	1:I:247:LEU:HD21	1.74	0.69
2:A:844:ASN:ND2	3:A:901:HOH:O	2.26	0.68
1:H:176:LEU:HD22	1:H:214:PHE:CG	2.29	0.67
1:H:46:GLU:OE2	1:H:63:LYS:NZ	2.31	0.59
2:B:670:LEU:HD12	2:B:692:LEU:HD23	1.84	0.58
1:H:180:GLN:HB2	1:H:190:LEU:HD11	1.85	0.58
1:H:87:ARG:HG3	1:H:89:GLU:HG2	1.87	0.56
1:I:176:LEU:HD22	1:I:214:PHE:CG	2.42	0.55
2:A:558:ASP:HB2	2:A:662:LEU:HD21	1.88	0.54
2:A:861:PHE:HE1	2:A:879:LEU:HD22	1.73	0.54
1:I:180:GLN:HB2	1:I:190:LEU:HD11	1.91	0.53
2:A:861:PHE:CE1	2:A:879:LEU:HD22	2.43	0.52
1:I:115:THR:HG23	3:I:332:HOH:O	2.10	0.51
2:A:866:LEU:HD12	2:A:876:THR:HB	1.93	0.51
2:B:866:LEU:HD12	2:B:876:THR:CB	2.37	0.50
1:I:172:ILE:HA	1:I:235:LYS:HD2	1.95	0.47
2:A:804:CYS:HA	2:A:805:PHE:HA	1.73	0.47
1:I:68:VAL:HG22	1:I:83:LEU:HD13	1.97	0.46
2:A:811:TYR:O	2:A:812:ASN:HB2	2.15	0.46
2:B:811:TYR:O	2:B:812:ASN:HB2	2.15	0.46
2:B:568:GLU:O	2:B:722:ILE:HD11	2.17	0.45
2:B:820:GLU:HA	2:B:822:LYS:HE3	2.00	0.44
1:H:151:PRO:HG2	1:H:154:LEU:HG	2.00	0.43
1:H:172:ILE:HA	1:H:235:LYS:HD2	2.01	0.43
2:A:737:CYS:HA	2:A:770:GLN:O	2.20	0.42
2:B:694:PRO:HB3	2:B:790:ILE:HG21	2.00	0.42
2:A:694:PRO:HB3	2:A:790:ILE:HG21	2.00	0.42
1:H:221:LEU:HD21	1:H:247:LEU:HD21	2.01	0.42
1:H:68:VAL:HG22	1:H:83:LEU:HD13	2.02	0.41
2:B:836:LYS:NZ	3:B:905:HOH:O	2.52	0.41
1:I:87:ARG:HG3	1:I:89:GLU:HG2	2.02	0.41
2:A:738:ALA:HA	2:A:794:LEU:HB3	2.03	0.41
1:H:35:ILE:HG22	1:H:50:ALA:HB2	2.03	0.40
1:I:36:TRP:CE2	1:I:81:MET:HB2	2.57	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	H	225/261 (86%)	220 (98%)	5 (2%)	0	100	100
1	I	224/261 (86%)	218 (97%)	6 (3%)	0	100	100
2	A	310/355 (87%)	296 (96%)	14 (4%)	0	100	100
2	B	316/355 (89%)	300 (95%)	16 (5%)	0	100	100
All	All	1075/1232 (87%)	1034 (96%)	41 (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	H	188/200 (94%)	188 (100%)	0	100	100
1	I	188/200 (94%)	187 (100%)	1 (0%)	88	92
2	A	295/331 (89%)	293 (99%)	2 (1%)	84	88
2	B	300/331 (91%)	298 (99%)	2 (1%)	84	88
All	All	971/1062 (91%)	966 (100%)	5 (0%)	88	92

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	I	98	LYS
2	A	726	TYR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	A	781	CYS
2	B	726	TYR
2	B	781	CYS

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

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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

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^{th} 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^2)	Q<0.9
1	H	229/261 (87%)	0.63	23 (10%) 7 9	33, 43, 68, 112	0
1	I	228/261 (87%)	0.60	22 (9%) 8 10	35, 46, 72, 136	0
2	A	314/355 (88%)	1.15	65 (20%) 1 1	34, 54, 109, 174	0
2	B	320/355 (90%)	1.31	77 (24%) 0 0	36, 57, 122, 151	0
All	All	1091/1232 (88%)	0.97	187 (17%) 1 1	33, 50, 100, 174	0

All (187) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	873	LEU	12.3
2	A	871	SER	10.1
2	B	869	GLU	10.0
2	A	869	GLU	9.8
2	A	873	LEU	9.5
2	B	815	ILE	9.4
2	A	776	ASN	9.3
2	B	814	GLU	8.4
2	B	872	ASN	7.6
2	B	871	SER	7.6
2	B	761	LYS	7.4
2	A	761	LYS	7.3
2	A	814	GLU	7.2
2	B	759	ASN	7.0
2	A	762	ILE	7.0
2	A	815	ILE	7.0
2	B	715	ASN	6.8
2	B	813	LYS	6.8
2	B	777	ASN	6.6
1	I	251	ARG	6.6
1	H	251	ARG	6.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	H	141	GLY	6.3
2	A	872	ASN	6.2
1	I	142	GLY	6.0
2	A	870	GLU	6.0
2	B	870	GLU	6.0
2	A	874	ILE	5.9
2	B	554	VAL	5.5
2	A	813	LYS	5.5
1	I	48	MET	5.5
2	B	758	HIS	5.4
2	B	762	ILE	5.3
2	B	553	GLY	5.2
2	A	775	LYS	4.9
2	A	816	LYS	4.8
2	B	874	ILE	4.8
1	I	249	ILE	4.7
2	B	868	GLU	4.6
2	A	773	TYR	4.6
2	B	732	ASN	4.5
2	B	779	TYR	4.5
1	H	142	GLY	4.4
2	A	793	ILE	4.4
2	A	868	GLU	4.4
1	H	231	CYS	4.4
2	B	816	LYS	4.4
2	B	562	LEU	4.3
2	A	562	LEU	4.3
1	H	96	CYS	4.3
1	I	143	SER	4.3
2	A	714	ASP	4.3
2	B	774	LYS	4.2
2	A	760	GLU	4.2
1	H	143	SER	4.2
1	I	250	LYS	4.1
1	H	48	MET	4.0
1	I	65	GLN	4.0
2	A	715	ASN	4.0
2	A	798	ASN	3.9
1	I	36	TRP	3.9
1	I	231	CYS	3.8
2	A	887	GLY	3.8
2	A	732	ASN	3.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	A	588	ASN	3.7
1	H	178	TRP	3.7
2	B	713	ASP	3.7
2	B	819	LYS	3.7
2	B	887	GLY	3.7
1	H	36	TRP	3.6
1	I	178	TRP	3.6
2	A	597	ASP	3.6
2	A	877	SER	3.6
2	A	819	LYS	3.5
1	I	57	ARG	3.5
2	B	797	LYS	3.5
2	B	599	LEU	3.5
2	B	878	TYR	3.5
2	A	774	LYS	3.4
2	B	634	LYS	3.4
2	B	776	ASN	3.4
2	B	760	GLU	3.4
2	B	587	THR	3.4
2	B	603	GLU	3.4
2	A	864	ILE	3.4
1	H	176	LEU	3.3
2	A	777	ASN	3.3
2	A	603	GLU	3.3
1	I	33	ALA	3.3
2	B	782	TYR	3.2
2	A	794	LEU	3.2
2	B	798	ASN	3.2
2	B	792	GLY	3.2
2	A	875	SER	3.1
2	B	877	SER	3.1
2	B	793	ILE	3.1
2	B	597	ASP	3.1
2	B	737	CYS	3.1
2	B	738	ALA	3.1
1	H	249	ILE	3.1
1	I	56	ASP	3.1
2	B	778	ILE	3.0
2	A	739	PHE	3.0
2	B	864	ILE	3.0
2	B	741	ASP	3.0
2	B	836	LYS	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	H	41	PRO	3.0
2	B	555	ASP	2.9
2	B	886	GLU	2.9
1	H	94	TYR	2.9
2	A	792	GLY	2.9
2	B	600	LYS	2.9
2	A	632	VAL	2.9
1	H	65	GLN	2.9
2	A	634	LYS	2.9
2	B	775	LYS	2.9
2	A	555	ASP	2.8
2	B	742	GLU	2.8
1	H	57	ARG	2.8
2	A	763	LYS	2.8
2	A	878	TYR	2.8
2	B	884	ILE	2.8
2	B	721	GLY	2.8
2	A	861	PHE	2.8
2	B	588	ASN	2.8
2	A	737	CYS	2.8
2	B	773	TYR	2.7
1	H	92	ALA	2.7
2	A	886	GLU	2.7
2	A	599	LEU	2.7
2	B	633	GLU	2.6
2	B	635	LEU	2.6
2	A	678	GLU	2.6
2	B	725	VAL	2.6
2	B	812	ASN	2.6
2	B	714	ASP	2.6
1	H	177	ALA	2.5
1	H	62	GLN	2.5
2	B	768	PHE	2.5
2	A	837	GLU	2.5
1	H	229	TYR	2.5
2	A	867	LYS	2.5
1	I	35	ILE	2.5
2	A	738	ALA	2.5
2	A	768	PHE	2.5
1	I	177	ALA	2.4
2	B	712	GLU	2.4
2	A	716	LYS	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	H	35	ILE	2.4
2	B	704	PHE	2.4
1	I	114	VAL	2.4
2	B	767	PHE	2.4
2	B	794	LEU	2.4
2	A	770	GLN	2.4
2	A	863	SER	2.4
2	B	604	SER	2.4
2	A	559	LYS	2.3
2	B	817	TRP	2.3
2	A	795	PHE	2.3
2	A	876	THR	2.3
2	B	739	PHE	2.3
2	A	587	THR	2.3
2	A	862	CYS	2.3
2	B	723	VAL	2.3
2	B	628	LEU	2.3
1	I	92	ALA	2.3
1	I	87	ARG	2.3
2	A	706	CYS	2.2
2	A	778	ILE	2.2
1	I	176	LEU	2.2
1	I	218	ILE	2.2
2	A	799	ILE	2.2
2	B	876	THR	2.2
2	A	812	ASN	2.2
2	B	791	GLY	2.2
2	A	717	LYS	2.2
2	B	765	LYS	2.2
1	H	114	VAL	2.1
1	I	216	LEU	2.1
1	H	250	LYS	2.1
2	B	559	LYS	2.1
1	H	34	ILE	2.1
2	A	797	LYS	2.1
2	A	708	ASN	2.1
2	B	693	PRO	2.0
1	I	243	GLN	2.0
2	B	863	SER	2.0
2	B	629	LYS	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 monosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.