Full wwPDB X-ray Structure Validation Report ⓘ

May 20, 2024 – 06:32 PM JST

PDB ID	:	8JHV
Title	:	The first crystal structure of a H-2Kb-restricted decapeptide from Cryptosporidium parvum
Authors	:	Fan, S.; Wang, Y.; Zhao, J.
Deposited on	:	2023-05-25
Resolution	:	3.47 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:

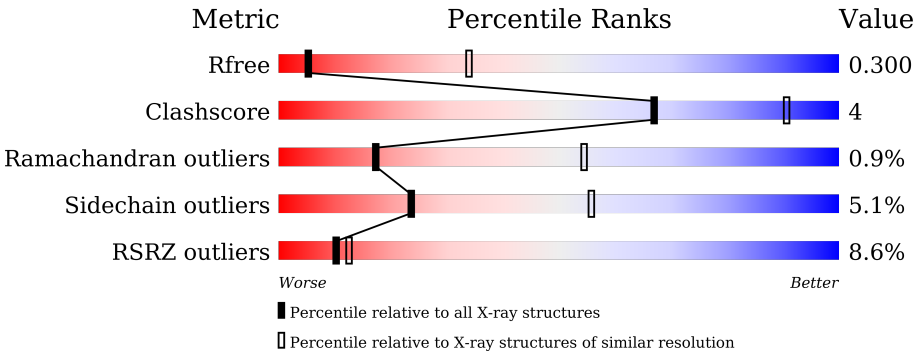
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 3.47 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	1379 (3.56-3.40)
Clashscore	141614	1461 (3.56-3.40)
Ramachandran outliers	138981	1424 (3.56-3.40)
Sidechain outliers	138945	1425 (3.56-3.40)
RSRZ outliers	127900	1289 (3.56-3.40)

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	275	11% 86% 13%
1	D	275	11% 85% 14%
2	B	100	86% 12%
2	E	100	2% 86% 13%
3	C	10	20% 70% 20% 10%
3	F	10	40% 50% 10%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 6282 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 H-2 class I histocompatibility antigen, K-B alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	275	Total	C	N	O	S	0	0	0
			2240	1413	394	424	9			
1	D	275	Total	C	N	O	S	0	0	0
			2240	1413	394	424	9			

- Molecule 2 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	99	Total	C	N	O	S	0	0	0
			818	523	138	150	7			
2	E	99	Total	C	N	O	S	0	0	0
			818	523	138	150	7			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	MET	-	initiating methionine	UNP P01887
E	0	MET	-	initiating methionine	UNP P01887

- Molecule 3 is a protein called VAL-THR-PHE-GLU-LYS-SER-TYR-ASN-THR-VAL.

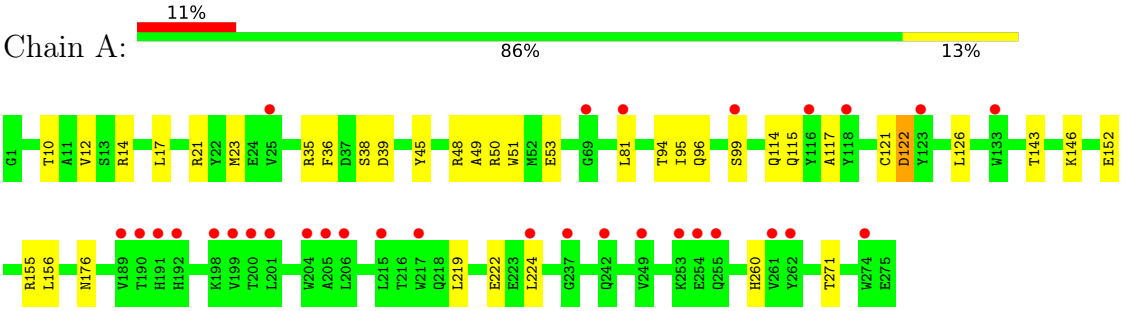
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	10	Total	C	N	O	0	0	0
			83	54	12	17			
3	F	10	Total	C	N	O	0	0	0
			83	54	12	17			

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.

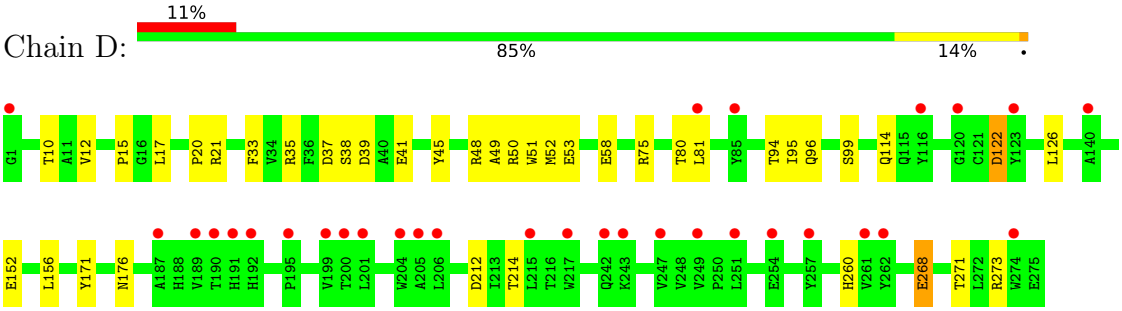
• Molecule 1: H-2 class I histocompatibility antigen, K-B alpha chain

Chain A:



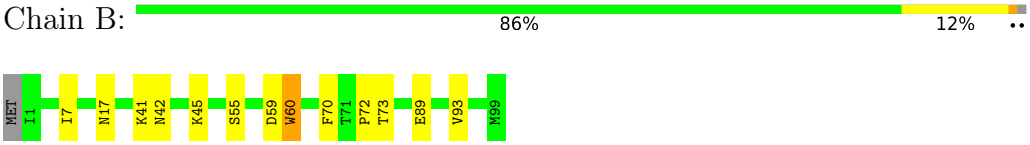
• Molecule 1: H-2 class I histocompatibility antigen, K-B alpha chain

Chain D:



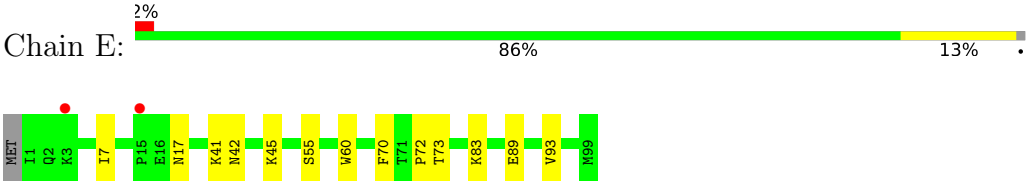
• Molecule 2: Beta-2-microglobulin

Chain B:



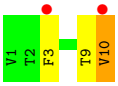
• Molecule 2: Beta-2-microglobulin

Chain E:



• Molecule 3: VAL-THR-PHE-GLU-LYS-SER-TYR-ASN-THR-VAL

Chain C:



• Molecule 3: VAL-THR-PHE-GLU-LYS-SER-TYR-ASN-THR-VAL

Chain F:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 1 2	Depositor
Cell constants a, b, c, α, β, γ	89.38Å 89.38Å 296.37Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.10 – 3.47 47.06 – 3.47	Depositor EDS
% Data completeness (in resolution range)	99.9 (47.10-3.47) 99.8 (47.06-3.47)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.22 (at 3.48Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.230 , 0.297 0.246 , 0.300	Depositor DCC
R_{free} test set	894 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	125.4	Xtriage
Anisotropy	0.030	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 93.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.298 for -h,-k,l	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	6282	wwPDB-VP
Average B, all atoms (Å^2)	151.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.04% 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	0.64	0/2301	0.74	0/3125
1	D	0.65	0/2301	0.75	0/3125
2	B	0.62	0/844	0.71	0/1144
2	E	0.62	0/844	0.71	0/1144
3	C	0.63	0/84	0.80	0/113
3	F	0.64	0/84	0.78	0/113
All	All	0.64	0/6458	0.74	0/8764

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	2240	0	2129	17	0
1	D	2240	0	2129	21	0
2	B	818	0	797	5	0
2	E	818	0	797	3	0
3	C	83	0	82	4	0
3	F	83	0	82	5	0
All	All	6282	0	6016	48	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 (48) 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:D:80:THR:HG21	3:F:10:VAL:HG12	1.55	0.89
3:F:9:THR:HG23	3:F:10:VAL:HG13	1.77	0.65
3:C:9:THR:O	3:C:10:VAL:HG13	2.02	0.60
1:D:41:GLU:N	1:D:41:GLU:OE1	2.38	0.56
1:D:50:ARG:NH2	1:D:53:GLU:OE1	2.40	0.55
1:A:143:THR:HG23	3:C:10:VAL:HG23	1.89	0.54
1:A:50:ARG:NH2	1:A:53:GLU:OE1	2.40	0.54
1:A:81:LEU:HD12	1:A:95:ILE:HD11	1.89	0.53
1:A:143:THR:HG23	3:C:10:VAL:CG2	2.38	0.53
1:D:81:LEU:HD12	1:D:95:ILE:HD11	1.89	0.53
1:D:80:THR:HG21	3:F:10:VAL:CG1	2.37	0.51
1:A:10:THR:HG22	1:A:96:GLN:HG2	1.93	0.50
1:D:152:GLU:OE2	3:F:8:ASN:ND2	2.45	0.50
2:B:17:ASN:HA	2:B:72:PRO:O	2.12	0.49
2:E:17:ASN:HA	2:E:72:PRO:O	2.12	0.49
1:D:33:PHE:CD2	1:D:52:MET:HE3	2.48	0.49
1:D:10:THR:HG22	1:D:96:GLN:HG2	1.96	0.48
1:D:12:VAL:HG22	1:D:94:THR:HG23	1.96	0.48
1:D:33:PHE:CD2	1:D:52:MET:CE	2.97	0.48
1:D:20:PRO:HD2	1:D:75:ARG:HD2	1.96	0.47
1:A:12:VAL:HG22	1:A:94:THR:HG23	1.97	0.46
1:A:122:ASP:OD1	1:A:122:ASP:N	2.50	0.45
1:D:35:ARG:HB3	1:D:48:ARG:HD3	1.99	0.45
1:D:126:LEU:HD22	1:D:156:LEU:HD11	1.99	0.45
1:A:146:LYS:HD2	3:C:10:VAL:HG11	1.98	0.44
1:A:35:ARG:HB3	1:A:48:ARG:HD3	1.99	0.44
1:A:152:GLU:HG3	1:A:156:LEU:HD23	2.00	0.44
1:A:126:LEU:HD22	1:A:156:LEU:HD11	2.00	0.43
1:A:219:LEU:HB2	1:A:224:LEU:HD11	2.01	0.43
1:D:15:PRO:O	1:D:17:LEU:HD12	2.17	0.43
2:B:7:ILE:HB	2:B:93:VAL:HG21	2.01	0.43
1:D:122:ASP:OD1	1:D:122:ASP:N	2.51	0.43
1:D:33:PHE:HD2	1:D:52:MET:CE	2.32	0.42
2:E:7:ILE:HB	2:E:93:VAL:HG21	2.02	0.42
1:D:260:HIS:ND1	1:D:271:THR:HG22	2.35	0.42
1:A:49:ALA:HB1	1:A:51:TRP:NE1	2.35	0.42
1:A:14:ARG:NH1	1:A:21:ARG:HB2	2.35	0.41
2:B:59:ASP:O	2:B:60:TRP:HB2	2.20	0.41
1:D:268:GLU:OE1	1:D:268:GLU:N	2.53	0.41
1:A:260:HIS:ND1	1:A:271:THR:HG22	2.34	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:E:41:LYS:HG2	2:E:42:ASN:HD22	1.86	0.41
2:B:41:LYS:HG2	2:B:42:ASN:HD22	1.86	0.41
1:D:49:ALA:HB1	1:D:51:TRP:NE1	2.36	0.41
1:A:23:MET:HA	1:A:36:PHE:O	2.21	0.40
1:D:21:ARG:NH2	1:D:37:ASP:OD2	2.55	0.40
1:D:52:MET:HE1	1:D:171:TYR:CD1	2.57	0.40
3:F:4:GLU:O	3:F:6:SER:N	2.54	0.40
1:A:117:ALA:HB2	2:B:60:TRP:CZ2	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	A	273/275 (99%)	256 (94%)	15 (6%)	2 (1%)	22	60
1	D	273/275 (99%)	255 (93%)	16 (6%)	2 (1%)	22	60
2	B	97/100 (97%)	87 (90%)	9 (9%)	1 (1%)	15	52
2	E	97/100 (97%)	87 (90%)	9 (9%)	1 (1%)	15	52
3	C	8/10 (80%)	6 (75%)	2 (25%)	0	100	100
3	F	8/10 (80%)	4 (50%)	3 (38%)	1 (12%)	0	4
All	All	756/770 (98%)	695 (92%)	54 (7%)	7 (1%)	17	54

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	F	5	LYS
1	A	176	ASN
1	D	176	ASN
1	A	114	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	114	GLN
2	B	60	TRP
2	E	60	TRP

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	233/233 (100%)	223 (96%)	10 (4%)	29 61
1	D	233/233 (100%)	223 (96%)	10 (4%)	29 61
2	B	93/94 (99%)	88 (95%)	5 (5%)	22 53
2	E	93/94 (99%)	87 (94%)	6 (6%)	17 49
3	C	10/10 (100%)	8 (80%)	2 (20%)	1 5
3	F	10/10 (100%)	9 (90%)	1 (10%)	7 30
All	All	672/674 (100%)	638 (95%)	34 (5%)	24 56

All (34) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	17	LEU
1	A	38	SER
1	A	39	ASP
1	A	45	TYR
1	A	99	SER
1	A	115	GLN
1	A	121	CYS
1	A	122	ASP
1	A	155	ARG
1	A	222	GLU
2	B	45	LYS
2	B	55	SER
2	B	70	PHE
2	B	73	THR
2	B	89	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	C	3	PHE
3	C	10	VAL
1	D	38	SER
1	D	39	ASP
1	D	45	TYR
1	D	58	GLU
1	D	99	SER
1	D	122	ASP
1	D	212	ASP
1	D	214	THR
1	D	268	GLU
1	D	273	ARG
2	E	45	LYS
2	E	55	SER
2	E	70	PHE
2	E	73	THR
2	E	83	LYS
2	E	89	GLU
3	F	10	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	127	ASN
1	A	174	ASN
1	A	242	GLN
2	B	29	GLN
2	B	42	ASN
1	D	127	ASN
1	D	174	ASN
2	E	29	GLN
2	E	42	ASN

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	A	275/275 (100%)	0.70	31 (11%) 5 6	113, 150, 192, 226	0
1	D	275/275 (100%)	0.69	31 (11%) 5 6	115, 150, 195, 207	0
2	B	99/100 (99%)	0.44	0 100 100	106, 153, 189, 204	0
2	E	99/100 (99%)	0.45	2 (2%) 65 61	120, 147, 168, 181	0
3	C	10/10 (100%)	1.19	2 (20%) 1 1	135, 139, 156, 161	0
3	F	10/10 (100%)	1.03	0 100 100	127, 136, 139, 142	0
All	All	768/770 (99%)	0.64	66 (8%) 10 12	106, 149, 190, 226	0

All (66) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	201	LEU	6.0
1	A	249	VAL	5.1
1	D	191	HIS	4.6
1	D	217	TRP	4.6
1	A	274	TRP	4.2
1	A	217	TRP	3.9
1	A	261	VAL	3.9
1	D	201	LEU	3.8
1	A	254	GLU	3.6
1	D	257	TYR	3.6
1	A	199	VAL	3.5
1	A	118	TYR	3.4
1	D	249	VAL	3.4
1	A	189	VAL	3.3
1	D	200	THR	3.3
1	A	123	TYR	3.3
1	D	195	PRO	3.2
1	D	215	LEU	3.1
1	D	204	TRP	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	215	LEU	3.1
1	D	247	VAL	3.0
1	D	243	LYS	3.0
1	D	123	TYR	3.0
1	A	191	HIS	3.0
1	A	206	LEU	2.9
1	A	204	TRP	2.9
1	A	192	HIS	2.8
1	A	262	TYR	2.8
1	D	254	GLU	2.8
1	D	120	GLY	2.8
1	A	198	LYS	2.7
1	D	187	ALA	2.7
1	A	99	SER	2.6
1	D	242	GLN	2.6
1	D	116	TYR	2.6
1	A	242	GLN	2.5
1	A	190	THR	2.5
1	A	200	THR	2.5
1	D	251	LEU	2.5
1	D	85	TYR	2.5
1	D	261	VAL	2.4
1	D	262	TYR	2.4
1	A	224	LEU	2.4
1	A	133	TRP	2.3
3	C	3	PHE	2.3
1	A	205	ALA	2.3
1	D	205	ALA	2.3
1	A	69	GLY	2.3
1	D	206	LEU	2.3
1	A	255	GLN	2.2
1	D	190	THR	2.2
2	E	15	PRO	2.2
1	A	25	VAL	2.2
2	E	3	LYS	2.2
1	D	274	TRP	2.2
1	D	81	LEU	2.1
1	A	253	LYS	2.1
1	D	140	ALA	2.1
1	A	237	GLY	2.1
1	D	189	VAL	2.1
1	D	192	HIS	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	1	GLY	2.1
3	C	10	VAL	2.1
1	D	199	VAL	2.1
1	A	116	TYR	2.0
1	A	81	LEU	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.