Full wwPDB X-ray Structure Validation Report ⓘ

Aug 4, 2025 – 05:14 PM JST

PDB ID	:	9J4U / pdb_00009j4u
Title	:	Structural basis for recognition of SARS-CoV-2 conserved nucleocapside epitopes by dominant T cell receptors
Authors	:	Yuan, P.; Wu, D.C.
Deposited on	:	2024-08-10
Resolution	:	2.17 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:

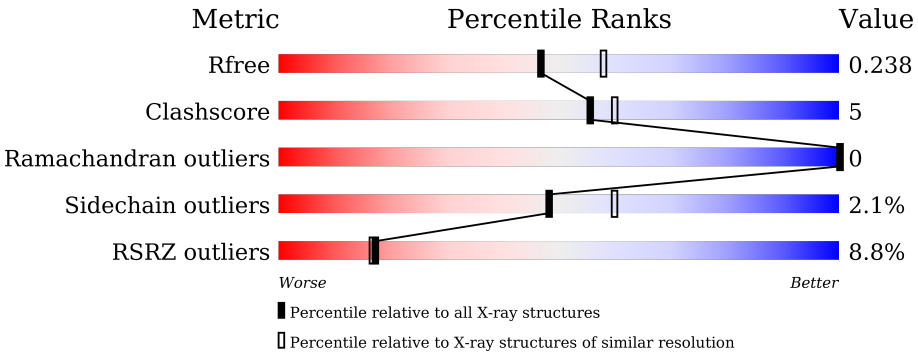
MolProbity	:	4-5-2 with Phenix2.0rc1
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.45.1

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 2.17 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}	164625	8336 (2.20-2.16)
Clashscore	180529	9404 (2.20-2.16)
Ramachandran outliers	177936	9297 (2.20-2.16)
Sidechain outliers	177891	9297 (2.20-2.16)
RSRZ outliers	164620	8337 (2.20-2.16)

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	276	3% 91% 8%
2	B	100	5% 87% 11%
3	D	204	17% 79% 15%
4	E	244	9% 88% 9%
5	C	9	78% 22%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 6777 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 MHC class I antigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	274	Total	C	N	O	S	0	1	0
			2230	1395	407	419	9			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP Q8WLS4

- 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
			814	520	139	152	3			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	MET	-	initiating methionine	UNP P61769

- Molecule 3 is a protein called LLL epitope specific TCR APHLA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	195	Total	C	N	O	S	0	0	0
			1492	932	251	300	9			

- Molecule 4 is a protein called LLL epitope specific TCR BETA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	240	Total	C	N	O	S	0	0	0
			1912	1209	329	369	5			

- Molecule 5 is a protein called Nucleoprotein.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	C	9	Total 77	C 49	N 14	O 14	0	0	0

- Molecule 6 is water.

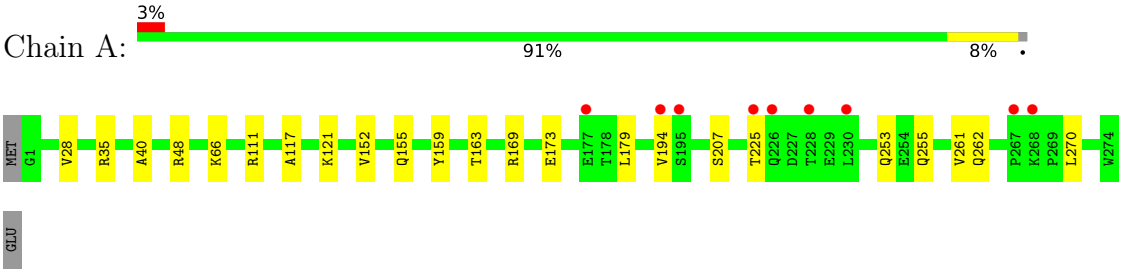
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	97	Total 97	O 97	0	0
6	B	32	Total 32	O 32	0	0
6	D	37	Total 37	O 37	0	0
6	E	79	Total 79	O 79	0	0
6	C	7	Total 7	O 7	0	0

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.

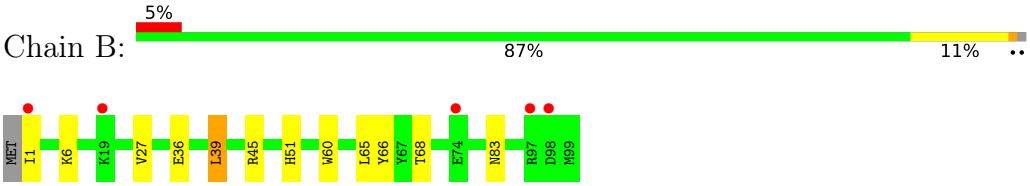
• Molecule 1: MHC class I antigen

Chain A:



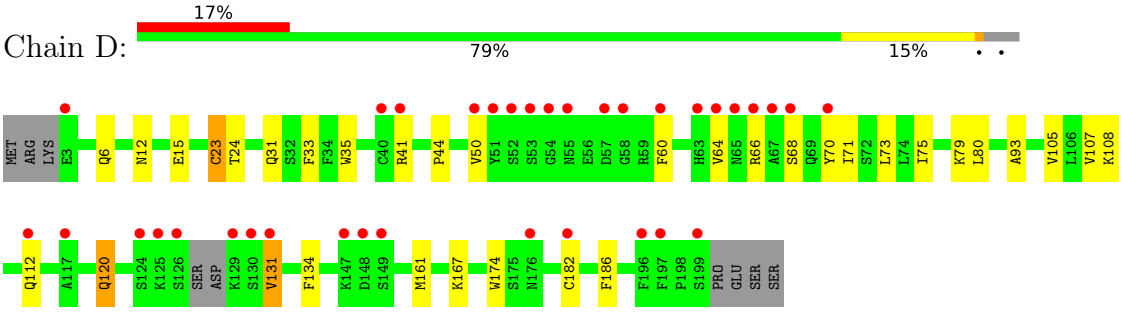
• Molecule 2: Beta-2-microglobulin

Chain B:



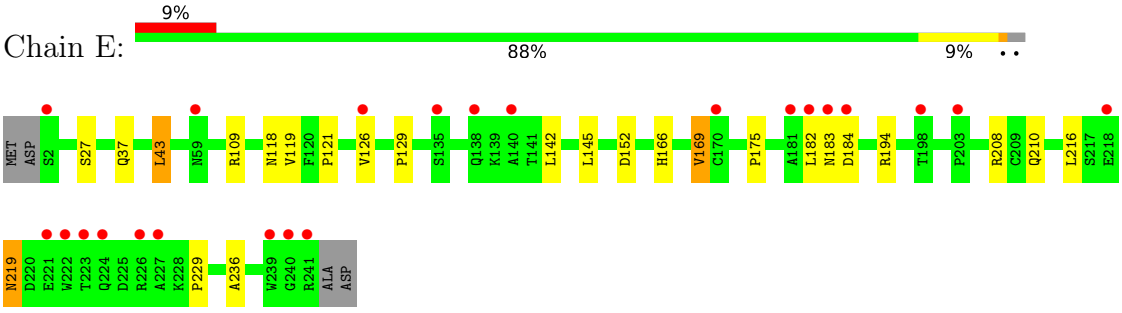
• Molecule 3: LLL epitope specific TCR APHLA

Chain D:



• Molecule 4: LLL epitope specific TCR BETA

Chain E:



• Molecule 5: Nucleoprotein

Chain C:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	54.76Å 148.27Å 64.21Å 90.00° 109.15° 90.00°	Depositor
Resolution (Å)	60.66 – 2.17 60.66 – 2.17	Depositor EDS
% Data completeness (in resolution range)	99.8 (60.66-2.17) 99.9 (60.66-2.17)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.75 (at 2.16Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R, R_{free}	0.193 , 0.239 0.193 , 0.238	Depositor DCC
R_{free} test set	2481 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å^2)	34.7	Xtriage
Anisotropy	0.599	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 47.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	6777	wwPDB-VP
Average B, all atoms (Å^2)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.54% 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	0.38	0/2295	0.53	0/3116
2	B	0.31	0/837	0.46	0/1134
3	D	0.31	0/1524	0.52	0/2068
4	E	0.37	0/1963	0.56	0/2675
5	C	0.45	0/76	0.65	0/100
All	All	0.36	0/6695	0.53	0/9093

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	2230	0	2081	18	0
2	B	814	0	773	8	0
3	D	1492	0	1390	22	0
4	E	1912	0	1811	17	0
5	C	77	0	88	2	0
6	A	97	0	0	2	0
6	B	32	0	0	1	0
6	C	7	0	0	0	0
6	D	37	0	0	0	0
6	E	79	0	0	0	0
All	All	6777	0	6143	58	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 5.

All (58) 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:35:ARG:NH1	1:A:48:ARG:HH21	1.71	0.88
3:D:15:GLU:OE1	3:D:167:LYS:NZ	2.16	0.79
1:A:35:ARG:HH11	1:A:48:ARG:HH21	1.35	0.70
3:D:120:GLN:HG2	3:D:182:CYS:SG	2.34	0.67
3:D:31:GLN:HA	3:D:66:ARG:HH12	1.60	0.65
4:E:37:GLN:HB2	4:E:43:LEU:HD13	1.79	0.63
1:A:169:ARG:O	1:A:173:GLU:HG3	2.00	0.62
1:A:66:LYS:HG3	3:D:93:ALA:HA	1.80	0.61
1:A:66:LYS:HE2	5:C:2:LEU:HB2	1.83	0.59
4:E:208:ARG:NH1	4:E:210:GLN:HB2	2.17	0.59
4:E:219:ASN:N	4:E:219:ASN:OD1	2.38	0.57
4:E:216:LEU:HD22	4:E:229:PRO:HD2	1.85	0.57
3:D:6:GLN:HB3	3:D:23:CYS:HB2	1.85	0.57
4:E:126:VAL:HG23	4:E:236:ALA:HB3	1.87	0.56
3:D:44:PRO:HG3	4:E:43:LEU:HD21	1.88	0.56
3:D:131:VAL:HG22	3:D:174:TRP:HB3	1.87	0.55
2:B:45:ARG:NH1	6:B:101:HOH:O	2.37	0.55
3:D:12:ASN:HB3	3:D:108:LYS:HD2	1.89	0.54
4:E:119:VAL:HG12	4:E:229:PRO:HB2	1.89	0.54
4:E:152:ASP:HB2	4:E:175:PRO:HG2	1.89	0.54
3:D:33:PHE:HB2	3:D:50:VAL:HG23	1.89	0.54
3:D:79:LYS:HD2	3:D:80:LEU:H	1.72	0.53
1:A:261:VAL:HG13	1:A:270:LEU:HB2	1.90	0.53
2:B:39:LEU:HD23	2:B:68:THR:HG22	1.90	0.53
1:A:261:VAL:CG1	1:A:270:LEU:HB2	2.39	0.53
3:D:60:PHE:CD2	3:D:75:ILE:HG12	2.44	0.51
3:D:80:LEU:HD12	3:D:107:VAL:HG12	1.93	0.50
1:A:117:ALA:HB2	2:B:60:TRP:CE2	2.48	0.49
3:D:161:MET:HE1	4:E:194:ARG:HG2	1.96	0.48
1:A:35:ARG:HH11	1:A:48:ARG:NH2	2.09	0.47
1:A:152:VAL:O	1:A:155:GLN:HG3	2.14	0.47
4:E:129:PRO:HD3	4:E:142:LEU:HG	1.95	0.47
4:E:109:ARG:NH2	4:E:152:ASP:OD1	2.33	0.47
3:D:64:VAL:HG22	3:D:71:ILE:HG13	1.96	0.47
3:D:112:GLN:H	3:D:112:GLN:CD	2.23	0.46
1:A:255:GLN:H	1:A:255:GLN:CD	2.23	0.46
4:E:182:LEU:HG	4:E:183:ASN:O	2.16	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:D:134:PHE:HB2	3:D:186:PHE:CE2	2.52	0.45
3:D:71:ILE:N	3:D:71:ILE:HD12	2.31	0.45
1:A:163:THR:OG1	3:D:66:ARG:NH2	2.50	0.45
2:B:51:HIS:HB3	2:B:66:TYR:CD2	2.52	0.44
3:D:66:ARG:H	3:D:66:ARG:HG2	1.64	0.44
4:E:183:ASN:O	4:E:184:ASP:HB2	2.16	0.44
4:E:118:ASN:ND2	4:E:184:ASP:O	2.51	0.44
3:D:24:THR:HG22	3:D:70:TYR:HB3	2.00	0.44
2:B:6:LYS:O	2:B:27:VAL:HA	2.18	0.43
2:B:36:GLU:OE1	2:B:83:ASN:HB3	2.20	0.42
1:A:40:ALA:HB3	6:A:346:HOH:O	2.19	0.42
3:D:41:ARG:NH2	4:E:152:ASP:OD2	2.52	0.42
4:E:166:HIS:O	4:E:169:VAL:HG13	2.20	0.42
4:E:121:PRO:HD3	4:E:229:PRO:HB3	2.02	0.42
1:A:111:ARG:HB3	6:A:333:HOH:O	2.19	0.41
2:B:51:HIS:HA	2:B:65:LEU:O	2.21	0.41
3:D:35:TRP:CE2	3:D:73:LEU:HB2	2.55	0.41
1:A:253:GLN:HA	1:A:255:GLN:HE22	1.86	0.41
1:A:159:TYR:HB2	5:C:3:LEU:HD13	2.03	0.40
1:A:28:VAL:HG11	1:A:179:LEU:HD13	2.03	0.40
1:A:121:LYS:HG2	2:B:1:ILE:HG13	2.03	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	272/276 (99%)	265 (97%)	7 (3%)	0	100	100
2	B	97/100 (97%)	95 (98%)	2 (2%)	0	100	100
3	D	191/204 (94%)	183 (96%)	8 (4%)	0	100	100
4	E	238/244 (98%)	230 (97%)	8 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
5	C	7/9 (78%)	7 (100%)	0	0	100 100
All	All	805/833 (97%)	780 (97%)	25 (3%)	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	A	228/232 (98%)	224 (98%)	4 (2%)	54 66
2	B	90/95 (95%)	89 (99%)	1 (1%)	70 80
3	D	166/181 (92%)	161 (97%)	5 (3%)	36 45
4	E	209/214 (98%)	204 (98%)	5 (2%)	44 54
5	C	9/9 (100%)	9 (100%)	0	100 100
All	All	702/731 (96%)	687 (98%)	15 (2%)	48 60

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	194	VAL
1	A	207	SER
1	A	225	THR
1	A	262	GLN
2	B	39	LEU
3	D	23	CYS
3	D	68	SER
3	D	105	VAL
3	D	120	GLN
3	D	131	VAL
4	E	27	SER
4	E	43	LEU
4	E	145	LEU
4	E	169	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	E	219	ASN

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	74	HIS
1	A	141	GLN
1	A	255	GLN
1	A	262	GLN
2	B	17	ASN
3	D	12	ASN
4	E	44	GLN
4	E	67	GLN
4	E	102	GLN

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 oligosaccharides 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^{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	274/276 (99%)	0.12	9 (3%) 49 49	24, 40, 61, 82	0
2	B	99/100 (99%)	0.46	5 (5%) 34 34	28, 50, 73, 78	0
3	D	195/204 (95%)	1.01	35 (17%) 4 4	31, 59, 93, 108	0
4	E	240/244 (98%)	0.57	23 (9%) 15 14	24, 48, 84, 102	0
5	C	9/9 (100%)	-0.44	0 100 100	28, 29, 31, 35	0
All	All	817/833 (98%)	0.50	72 (8%) 17 17	24, 47, 82, 108	0

All (72) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	E	182	LEU	5.1
3	D	199	SER	4.4
3	D	182	CYS	3.8
4	E	184	ASP	3.7
3	D	147	LYS	3.6
3	D	126	SER	3.6
3	D	197	PHE	3.5
3	D	51	TYR	3.4
3	D	130	SER	3.4
3	D	50	VAL	3.4
3	D	66	ARG	3.3
1	A	267	PRO	3.2
3	D	53	SER	3.2
4	E	203	PRO	3.2
3	D	129	LYS	3.1
2	B	98	ASP	3.1
4	E	222	TRP	3.1
3	D	54	GLY	3.1
3	D	68	SER	2.9
1	A	225	THR	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	226	GLN	2.8
4	E	241	ARG	2.8
3	D	67	ALA	2.8
4	E	227	ALA	2.7
3	D	52	SER	2.7
4	E	183	ASN	2.6
4	E	2	SER	2.6
2	B	19	LYS	2.6
3	D	149	SER	2.6
3	D	70	TYR	2.6
4	E	218	GLU	2.6
3	D	125	LYS	2.5
4	E	223	THR	2.5
4	E	126	VAL	2.5
3	D	196	PHE	2.5
1	A	230	LEU	2.5
4	E	181	ALA	2.5
3	D	63	HIS	2.4
4	E	170	CYS	2.4
4	E	239	TRP	2.4
3	D	55	ASN	2.3
3	D	124	SER	2.3
4	E	224	GLN	2.3
2	B	1	ILE	2.3
2	B	74	GLU	2.3
3	D	60	PHE	2.3
4	E	140	ALA	2.3
1	A	177	GLU	2.3
3	D	131	VAL	2.3
4	E	138	GLN	2.3
4	E	135	SER	2.3
3	D	3	GLU	2.2
3	D	64	VAL	2.2
3	D	58	GLY	2.2
3	D	57	ASP	2.2
2	B	97	ARG	2.2
3	D	65	ASN	2.2
3	D	117	ALA	2.2
4	E	240	GLY	2.2
3	D	148	ASP	2.1
3	D	41	ARG	2.1
3	D	40	CYS	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
4	E	198	THR	2.1
3	D	112	GLN	2.1
1	A	194	VAL	2.1
3	D	176	ASN	2.0
4	E	221	GLU	2.0
1	A	195	SER	2.0
1	A	228	THR	2.0
4	E	59	ASN	2.0
1	A	268	LYS	2.0
4	E	226	ARG	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 oligosaccharides 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.