Full wwPDB X-ray Structure Validation Report (i)

Oct 7, 2023 – 03:35 PM EDT

PDB ID	:	6DKP
Title	:	The complex among DMF5(alpha-D26Y, alpha-Y50A,beta-L98W) TCR, human Class I MHC HLA-A2 and MART-1(26-35)(A27L) peptide
Authors	:	Hellman, L.M.; Singh, N.K.
Deposited on	:	2018-05-30
Resolution	:	2.97 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 (i) 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 (1)) were used in the production of this report:

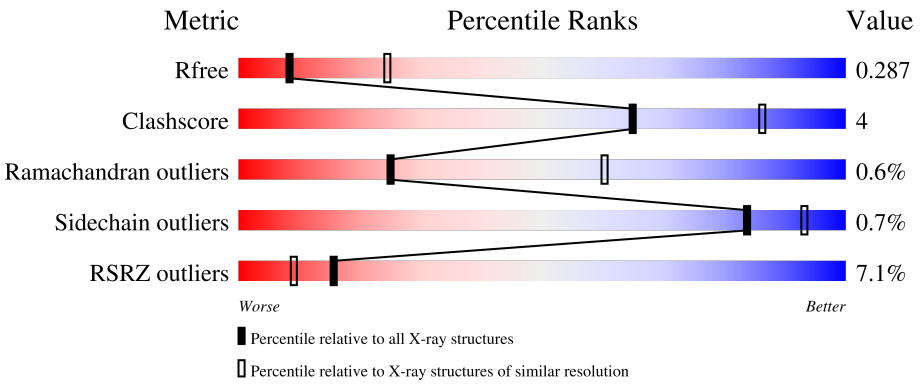
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.97 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	3104 (3.00-2.92)
Clashscore	141614	3462 (3.00-2.92)
Ramachandran outliers	138981	3340 (3.00-2.92)
Sidechain outliers	138945	3343 (3.00-2.92)
RSRZ outliers	127900	2986 (3.00-2.92)

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	8%, 88%, 12%
2	B	100	90%, 9%, ·
3	C	10	90%, 10%
4	D	200	8%, 87%, 8%, ·
5	E	243	7%, 86%, 13%, ·

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 6539 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 HLA class I histocompatibility antigen, A-2 alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	275	2247	1403	409	426	9	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP P01892

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	99	828	528	140	157	3	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	MET	-	initiating methionine	UNP P61769

- Molecule 3 is a protein called Melanoma antigen recognized by T-cells 1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	C	10	69	45	10	14	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	2	LEU	ALA	engineered mutation	UNP Q16655

- Molecule 4 is a protein called DMF5 T-cell Receptor Alpha Chain fusion.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	192	1492	933	247	304	8	0	0	0

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	0	MET	-	initiating methionine	UNP A0A075B6T6
D	26	TYR	ASP	engineered mutation	UNP A0A075B6T6
D	50	ALA	TYR	engineered mutation	UNP A0A075B6T6
D	92	PHE	-	linker	UNP A0A075B6T6
D	93	GLY	-	linker	UNP A0A075B6T6
D	94	GLY	-	linker	UNP A0A075B6T6
D	95	GLY	-	linker	UNP A0A075B6T6
D	96	LYS	-	linker	UNP A0A075B6T6
D	97	LEU	-	linker	UNP A0A075B6T6
D	98	ILE	-	linker	UNP A0A075B6T6
D	99	PHE	-	linker	UNP A0A075B6T6
D	100	GLY	-	linker	UNP A0A075B6T6
D	101	GLN	-	linker	UNP A0A075B6T6
D	102	GLY	-	linker	UNP A0A075B6T6
D	103	THR	-	linker	UNP A0A075B6T6
D	104	GLU	-	linker	UNP A0A075B6T6
D	105	LEU	-	linker	UNP A0A075B6T6
D	106	SER	-	linker	UNP A0A075B6T6
D	107	VAL	-	linker	UNP A0A075B6T6
D	108	LYS	-	linker	UNP A0A075B6T6
D	157	CYS	THR	conflict	UNP P01848

- Molecule 5 is a protein called DMF5 T-cell Receptor Beta Chain fusion.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	E	242	1903	1196	334	365	8	0	0	0

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	3	MET	-	initiating methionine	UNP A0A1B0GXE1
E	98	TRP	-	linker	UNP A0A1B0GXE1
E	99	SER	-	linker	UNP A0A1B0GXE1
E	100	PHE	-	linker	UNP A0A1B0GXE1
E	101	GLY	-	linker	UNP A0A1B0GXE1
E	102	THR	-	linker	UNP A0A1B0GXE1

Continued on next page...

Continued from previous page...

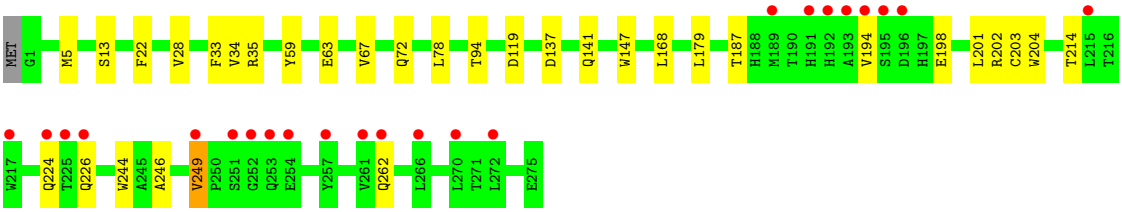
Chain	Residue	Modelled	Actual	Comment	Reference
E	103	GLU	-	linker	UNP A0A1B0GXE1
E	104	ALA	-	linker	UNP A0A1B0GXE1
E	105	PHE	-	linker	UNP A0A1B0GXE1
E	106	PHE	-	linker	UNP A0A1B0GXE1
E	107	GLY	-	linker	UNP A0A1B0GXE1
E	108	GLN	-	linker	UNP A0A1B0GXE1
E	109	GLY	-	linker	UNP A0A1B0GXE1
E	110	THR	-	linker	UNP A0A1B0GXE1
E	111	ARG	-	linker	UNP A0A1B0GXE1
E	112	LEU	-	linker	UNP A0A1B0GXE1
E	113	THR	-	linker	UNP A0A1B0GXE1
E	114	VAL	-	linker	UNP A0A1B0GXE1
E	115	VAL	-	linker	UNP A0A1B0GXE1
E	119	ASN	LYS	conflict	UNP K7N5M4
E	120	LYS	ASN	conflict	UNP K7N5M4
E	204	ASP	ASN	conflict	UNP K7N5M4

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.

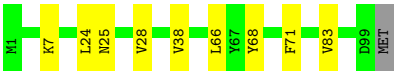
• Molecule 1: HLA class I histocompatibility antigen, A-2 alpha chain

Chain A:



• Molecule 2: Beta-2-microglobulin

Chain B:

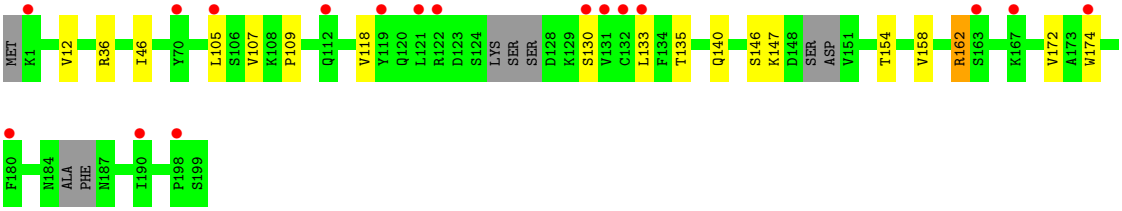


• Molecule 3: Melanoma antigen recognized by T-cells 1

Chain C:

• Molecule 4: DMF5 T-cell Receptor Alpha Chain fusion

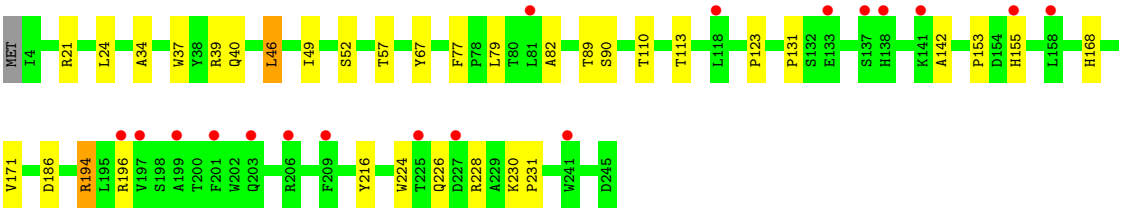
Chain D:



• Molecule 5: DMF5 T-cell Receptor Beta Chain fusion

Chain E:

7%
86%
13%


MET I4 R21 L24 A34 W37 Y38 R39 Q40 L46 I49 S52 T67 Y67 F77 P78 L79 T80 L81 A82 T89 S90 T110 T113 L118 P123 P131 S132 E133 S137 H138 K141 A142 P153 D154 H155 L158 H168
V171 D186 R194 L195 R196 V197 S198 A199 T200 F201 W202 Q203 R206 F209 Y216 W224 T225 Q226 D227 R228 A229 K230 P231 W241 D245

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	226.49Å 49.06Å 92.50Å 90.00° 94.98° 90.00°	Depositor
Resolution (Å)	47.94 – 2.97 47.94 – 2.97	Depositor EDS
% Data completeness (in resolution range)	94.0 (47.94-2.97) 94.1 (47.94-2.97)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.95 (at 2.96Å)	Xtriage
Refinement program	PHENIX (1.11.1_2575)	Depositor
R, R_{free}	0.229 , 0.286 0.229 , 0.287	Depositor DCC
R_{free} test set	1033 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å^2)	87.0	Xtriage
Anisotropy	0.534	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 69.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6539	wwPDB-VP
Average B, all atoms (Å^2)	118.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.06% 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 (i)

5.1 Standard geometry (i)

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	0.24	0/2312	0.43	0/3137
2	B	0.24	0/851	0.45	0/1152
3	C	0.24	0/68	0.50	0/90
4	D	0.25	0/1521	0.44	0/2053
5	E	0.25	0/1954	0.45	0/2663
All	All	0.24	0/6706	0.44	0/9095

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 (i)

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	2247	0	2096	21	0
2	B	828	0	794	4	0
3	C	69	0	79	1	0
4	D	1492	0	1414	10	0
5	E	1903	0	1807	20	0
All	All	6539	0	6190	52	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 (52) 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:194:VAL:CG1	1:A:198:GLU:HB3	1.86	1.04
1:A:194:VAL:HG12	1:A:198:GLU:HB3	1.40	1.01
1:A:194:VAL:CG1	1:A:198:GLU:CB	2.67	0.72
1:A:5:MET:HB2	1:A:168:LEU:HD13	1.72	0.71
2:B:7:LYS:O	2:B:28:VAL:HA	1.90	0.70
1:A:194:VAL:HG13	1:A:198:GLU:CB	2.24	0.67
5:E:40:GLN:HB2	5:E:46:LEU:HB3	1.83	0.61
5:E:123:PRO:HD3	5:E:231:PRO:HB3	1.84	0.60
5:E:39:ARG:HB2	5:E:49:ILE:HD11	1.85	0.59
1:A:194:VAL:HG13	1:A:198:GLU:HB3	1.76	0.58
4:D:36:ARG:HB2	4:D:46:ILE:HD11	1.86	0.57
4:D:109:PRO:HG3	4:D:158:VAL:HG11	1.87	0.56
2:B:38:VAL:HG22	2:B:83:VAL:HG22	1.89	0.55
1:A:13:SER:HB3	1:A:78:LEU:HD13	1.89	0.55
4:D:12:VAL:HG12	4:D:105:LEU:HD11	1.90	0.54
5:E:224:TRP:CG	5:E:230:LYS:HB3	2.45	0.52
1:A:187:THR:HA	1:A:204:TRP:O	2.11	0.51
4:D:135:THR:OG1	5:E:196:ARG:NH2	2.44	0.51
1:A:28:VAL:HG11	1:A:179:LEU:HD13	1.93	0.50
1:A:202:ARG:HG2	1:A:246:ALA:HB2	1.93	0.50
5:E:21:ARG:HG3	5:E:82:ALA:HA	1.93	0.49
5:E:224:TRP:CD1	5:E:230:LYS:HB3	2.47	0.49
5:E:186:ASP:OD1	5:E:186:ASP:N	2.46	0.49
5:E:89:THR:HG23	5:E:113:THR:HA	1.94	0.48
4:D:140:GLN:N	4:D:140:GLN:OE1	2.47	0.48
5:E:37:TRP:NE1	5:E:77:PHE:O	2.42	0.47
1:A:203:CYS:O	1:A:244:TRP:HA	2.15	0.47
5:E:34:ALA:HA	5:E:52:SER:O	2.16	0.46
1:A:224:GLN:O	1:A:226:GLN:N	2.44	0.45
1:A:214:THR:HB	1:A:262:GLN:HB2	1.99	0.44
5:E:155:HIS:HB3	5:E:216:TYR:HB2	1.98	0.44
4:D:130:SER:O	4:D:174:TRP:HA	2.18	0.44
1:A:59:TYR:O	1:A:63:GLU:HG2	2.17	0.43
1:A:201:LEU:HD12	1:A:249:VAL:HG11	1.99	0.43
4:D:118:VAL:HA	4:D:133:LEU:O	2.18	0.43
4:D:12:VAL:O	4:D:107:VAL:HA	2.18	0.43
5:E:24:LEU:HD22	5:E:110:THR:HG21	2.01	0.43
2:B:25:ASN:HB3	2:B:66:LEU:HD11	2.01	0.43
5:E:40:GLN:O	5:E:90:SER:OG	2.35	0.42
1:A:147:TRP:CD1	3:C:8:LEU:HD23	2.54	0.42
5:E:67:TYR:HB3	5:E:79:LEU:HD11	2.01	0.42
1:A:137:ASP:O	1:A:141:GLN:HG2	2.19	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:E:224:TRP:NE1	5:E:226:GLN:HB2	2.35	0.42
1:A:72:GLN:HG2	5:E:57:THR:HG21	2.02	0.41
4:D:162:ARG:NH2	5:E:168:HIS:O	2.54	0.41
5:E:131:PRO:HG2	5:E:142:ALA:HB1	2.03	0.41
5:E:171:VAL:HA	5:E:194:ARG:O	2.20	0.41
1:A:94:THR:O	1:A:119:ASP:N	2.42	0.41
2:B:24:LEU:O	2:B:68:TYR:HA	2.21	0.41
4:D:154:THR:HG22	4:D:172:VAL:O	2.21	0.40
1:A:33:PHE:CD2	1:A:34:VAL:HG13	2.57	0.40
1:A:22:PHE:HE2	1:A:67:VAL:HG22	1.86	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	273/276 (99%)	262 (96%)	11 (4%)	0	100	100
2	B	97/100 (97%)	95 (98%)	2 (2%)	0	100	100
3	C	8/10 (80%)	8 (100%)	0	0	100	100
4	D	184/200 (92%)	170 (92%)	11 (6%)	3 (2%)	9	36
5	E	240/243 (99%)	230 (96%)	8 (3%)	2 (1%)	19	53
All	All	802/829 (97%)	765 (95%)	32 (4%)	5 (1%)	25	60

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	D	146	SER
4	D	147	LYS
4	D	162	ARG
5	E	228	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
5	E	153	PRO

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	231/232 (100%)	229 (99%)	2 (1%)	78	91
2	B	94/95 (99%)	93 (99%)	1 (1%)	73	89
3	C	7/7 (100%)	7 (100%)	0	100	100
4	D	169/176 (96%)	169 (100%)	0	100	100
5	E	204/205 (100%)	202 (99%)	2 (1%)	76	90
All	All	705/715 (99%)	700 (99%)	5 (1%)	84	93

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

Mol	Chain	Res	Type
1	A	35	ARG
1	A	249	VAL
2	B	71	PHE
5	E	46	LEU
5	E	194	ARG

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 (i)

6.1 Protein, DNA and RNA chains (i)

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(Å^2)	Q<0.9
1	A	275/276 (99%)	0.47	23 (8%)	11	6	62, 102, 200, 227	0
2	B	99/100 (99%)	0.08	0	100	100	67, 87, 131, 151	0
3	C	10/10 (100%)	0.18	0	100	100	70, 74, 89, 89	0
4	D	192/200 (96%)	0.50	17 (8%)	9	5	78, 129, 187, 200	0
5	E	242/243 (99%)	0.43	18 (7%)	14	8	61, 122, 169, 194	0
All	All	818/829 (98%)	0.41	58 (7%)	16	9	61, 113, 188, 227	0

All (58) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	257	TYR	5.2
1	A	251	SER	4.7
1	A	252	GLY	4.5
1	A	254	GLU	4.3
4	D	119	TYR	4.2
5	E	155	HIS	4.2
4	D	70	TYR	4.2
4	D	131	VAL	4.0
1	A	215	LEU	3.9
1	A	193	ALA	3.4
4	D	198	PRO	3.3
1	A	196	ASP	3.2
1	A	195	SER	3.2
5	E	137	SER	3.2
4	D	174	TRP	3.1
5	E	225	THR	3.1
5	E	118	LEU	3.0
1	A	270	LEU	3.0
1	A	194	VAL	2.9
4	D	122	ARG	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
5	E	201	PHE	2.9
1	A	272	LEU	2.8
1	A	224	GLN	2.8
1	A	217	TRP	2.7
5	E	203	GLN	2.7
1	A	226	GLN	2.7
4	D	121	LEU	2.6
5	E	241	TRP	2.6
5	E	138	HIS	2.6
4	D	180	PHE	2.5
5	E	196	ARG	2.5
4	D	163	SER	2.5
5	E	199	ALA	2.4
4	D	112	GLN	2.4
4	D	133	LEU	2.4
5	E	206	ARG	2.4
1	A	225	THR	2.4
5	E	209	PHE	2.4
1	A	262	GLN	2.4
4	D	190	ILE	2.3
1	A	191	HIS	2.3
5	E	133	GLU	2.3
4	D	1	LYS	2.3
1	A	261	VAL	2.2
5	E	158	LEU	2.2
1	A	189	MET	2.2
1	A	192	HIS	2.2
5	E	227	ASP	2.2
1	A	266	LEU	2.2
5	E	141	LYS	2.1
1	A	253	GLN	2.1
5	E	81	LEU	2.1
4	D	105	LEU	2.1
5	E	197	VAL	2.1
4	D	130	SER	2.0
4	D	132	CYS	2.0
1	A	249	VAL	2.0
4	D	167	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.