WORLDWIDE PDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 26, 2023 – 06:54 AM EDT

PDB ID	:	6BPE
Title	:	Plasmodium vivax reticulocyte binding protein 2b (PvRBP2b) bound to monoclonal antibody 6H1
Authors	:	Gruszczyk, J.; Chan, L.J.; Tham, W.H.
Deposited on	:	2017-11-22
Resolution	:	3.34 Å(reported)

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

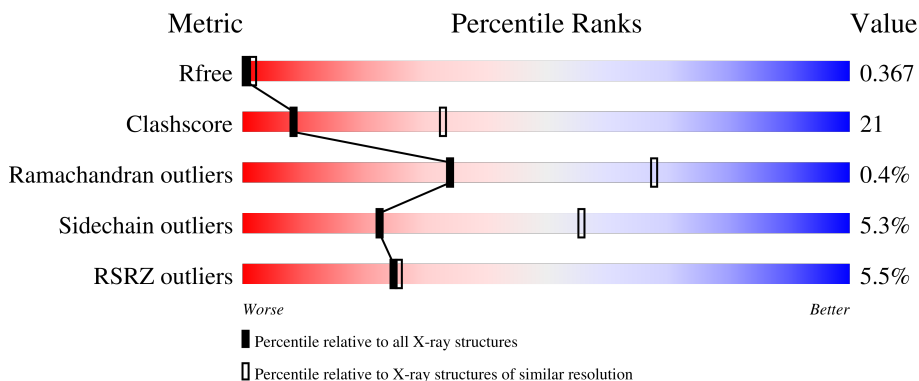|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 3.34 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		0.367
Clashscore		21
Ramachandran outliers		0.4%
Sidechain outliers		5.3%
RSRZ outliers		5.5%

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	1060 (3.38-3.30)
Clashscore	141614	1111 (3.38-3.30)
Ramachandran outliers	138981	1090 (3.38-3.30)
Sidechain outliers	138945	1089 (3.38-3.30)
RSRZ outliers	127900	1028 (3.38-3.30)

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	307	3% / 64% / 33% · ·
1	D	307	5% / 66% / 31% · ·
1	G	307	2% / 58% / 36% · ·
1	J	307	4% / 50% / 44% · ·
2	B	254	7% / 43% / 37% · / 18%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	E	254	7% 49% 31% 18%
2	H	254	6% 21% 22% 56%
3	C	238	6% 45% 37% 16%
3	F	238	3% 47% 34% 17%
3	I	238	4% 12% 15% 72%

WORLDWIDE
PDB
PROTEIN DATA BANK

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 17777 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 Reticulocyte binding protein 2, putative.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	299	Total	C	N	O	S	0	0	0
			2518	1612	430	467	9			
1	D	302	Total	C	N	O	S	0	0	0
			2544	1627	436	472	9			
1	G	302	Total	C	N	O	S	0	0	0
			2543	1627	436	471	9			
1	J	297	Total	C	N	O	S	0	0	0
			2500	1602	428	461	9			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	164	GLY	-	expression tag	UNP A5K736
A	165	ALA	-	expression tag	UNP A5K736
A	166	MET	-	expression tag	UNP A5K736
A	167	GLY	-	expression tag	UNP A5K736
A	168	SER	-	expression tag	UNP A5K736
D	164	GLY	-	expression tag	UNP A5K736
D	165	ALA	-	expression tag	UNP A5K736
D	166	MET	-	expression tag	UNP A5K736
D	167	GLY	-	expression tag	UNP A5K736
D	168	SER	-	expression tag	UNP A5K736
G	164	GLY	-	expression tag	UNP A5K736
G	165	ALA	-	expression tag	UNP A5K736
G	166	MET	-	expression tag	UNP A5K736
G	167	GLY	-	expression tag	UNP A5K736
G	168	SER	-	expression tag	UNP A5K736
J	164	GLY	-	expression tag	UNP A5K736
J	165	ALA	-	expression tag	UNP A5K736
J	166	MET	-	expression tag	UNP A5K736
J	167	GLY	-	expression tag	UNP A5K736
J	168	SER	-	expression tag	UNP A5K736

WORLDWIDE PDB
PROTEIN DATA BANK

• Molecule 2 is a protein called Monoclonal antibody 6H1 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	209	Total	C	N	O	S	0	0	0
			1594	1013	252	321	8			
2	E	208	Total	C	N	O	S	0	0	0
			1590	1011	254	318	7			
2	H	113	Total	C	N	O	S	0	0	0
			885	560	140	180	5			

• Molecule 3 is a protein called Monoclonal antibody 6H1 Fab light chain.

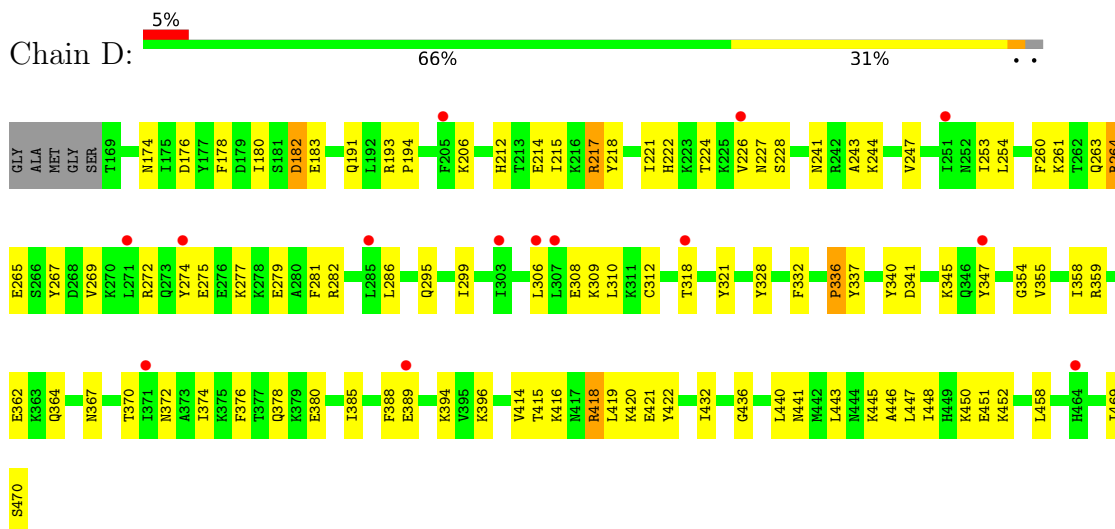| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	199	Total	C	N	O	S	0	0	0
			1539	960	258	314	7			
3	F	198	Total	C	N	O	S	0	0	0
			1535	962	255	311	7			
3	I	67	Total	C	N	O	S	0	0	0
			529	334	88	105	2			

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: Reticulocyte binding protein 2, putative

Chain A:

- Molecule 1: Reticulocyte binding protein 2, putative

Chain D:

- Molecule 1: Reticulocyte binding protein 2, putative

Chain G:

• Molecule 1: Reticulocyte binding protein 2, putative

Chain J:

• Molecule 2: Monoclonal antibody 6H1 Fab heavy chain

Chain B:

• Molecule 2: Monoclonal antibody 6H1 Fab heavy chain

Chain E:

• Molecule 2: Monoclonal antibody 6H1 Fab heavy chain

Chain H:

• Molecule 3: Monoclonal antibody 6H1 Fab light chain

Chain C:

• Molecule 3: Monoclonal antibody 6H1 Fab light chain

Chain F:

MET GLY ILE LYS MET GLU SER GLN THR GLN VAL PHE VAL TYR MET LEU LEU TRP LEU SER GLY VAL ASP GLY ASP ILE V27 M28 T29 K33 F34 T37 S38 I39 G40 V41 S44 C47 V53 G54 T55 N56 V57 A58 W59 Y60 P64 GLY GLN S67 A70 L71 I72

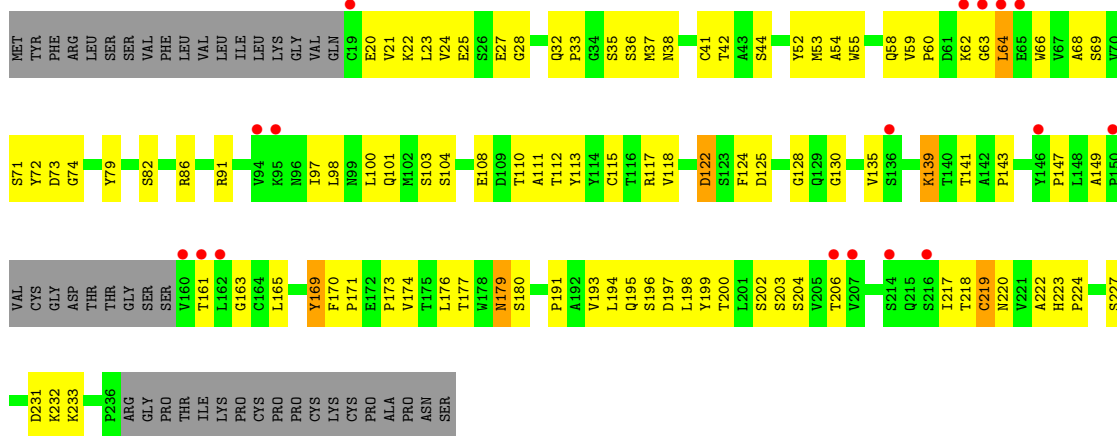R78 N79 D84 G92 L97 S100 ASN VAL GLN SER GLU ASP LEU A108 F111 C112 Q113 Q114 Y115 D116 S117 Y118 P119 Y120 P121 F122 K127 L128 E129 I130 K131 R132 A133 D134 P137 S140 I141 F142 P143 P144 S145 S146 E147 Q148 L149 T150 S151 G152 G153 A154 S155

V156 V157 C158 F159 F163 Y164 P165 K166 N169 V170 K171 W172 K173 I174 E178 V183 L184 N185 T188 D189 Q190 D191 D194 Y197 S198 M199 S200 S201 T202 L203 T204 L205 T206 K207 D208 E209 Y210 E211 R212 S215 Y216 E219 A220 T221 H222 K223 P228 K231 S232

F233 ASN ARG ASN GLU CYS

- Molecule 3: Monoclonal antibody 6H1 Fab light chain

4%

Chain I: 12%   15%   •   72%

MET GLY ILE LYS MET GLU SER GLN THR GLN VAL PHE VAL TYR MET LEU LEU TRP LEU SER GLY VAL ASP GLY ASP ILE VAL MET THR GLN SER GLN LYS PHE MET SER THR SER ILE GLY V41 R42 V43 S44 V45 T46 C47 K48 A49 S50 Q51 N52 V53 G54 T55 N56 V57 A58 W59 Y60

Q61 Q62 K63 PRO GLY GLN S67 P68 K69 A70 L71 S74 A75 R78 D84 R85 F86 T87 G88 S89 G90 S91 F95 T96 L97 T98 ILE SER ASN VAL GLN SER GLU ASP LEU ALA E109 Y110 F111 C112 Q113 Q114 Y115 D116 S117 Y118 P119 Y120 PRO PHE GLY GLY THR LYS LEU

GLU ILE LYS ARG ALA ASP SER ALA ALA PRO THR VAL SER ILE PHE PRO PRO SER SER GLU GLN LEU THR SER GLY GLY ALA SER VAL VAL CYS PHE LEU ASN ASN PHE TYR PRO LYS ASP ILE ASN VAL LYS TRP LYS ILE ASP GLY SER GLU ARG GLN ASN GLY VAL LEU ASN SER TRP THR

ASP GLN ASP SER LYS ASP SER THR TYR SER MET SER SER THR LEU THR LEU THR LYS ASP GLU TYR GLU ARG HIS ASN SER TYR THR CYS GLU ALA THR HIS LYS THR SER THR SER PRO ILE VAL LYS SER PHE ASN ARG ASN GLU CYS

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	106.60Å 123.48Å 143.28Å 90.00° 105.25° 90.00°	Depositor
Resolution (Å)	73.75 – 3.34 75.49 – 3.34	Depositor EDS
% Data completeness (in resolution range)	99.2 (73.75-3.34) 99.4 (75.49-3.34)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.82 (at 3.33Å)	Xtriage
Refinement program	PHENIX 1.12_2829	Depositor
R, R_{free}	0.318 , 0.367 0.318 , 0.367	Depositor DCC
R_{free} test set	1068 reflections (2.06%)	wwPDB-VP
Wilson B-factor (Å^2)	75.2	Xtriage
Anisotropy	0.681	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.25 , 45.4	EDS
L-test for twinning [2]	$< |L| > = 0.44, < L^2 > = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.81	EDS
Total number of atoms	17777	wwPDB-VP
Average B, all atoms (Å^2)	86.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 30.93 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.2039e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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	0.25	0/2567	0.36	0/3444
1	D	0.24	0/2593	0.38	0/3477
1	G	0.26	0/2592	0.43	0/3477
1	J	0.27	0/2549	0.46	0/3420
2	B	0.27	0/1632	0.53	0/2231
2	E	0.27	0/1628	0.51	0/2225
2	H	0.29	0/904	0.51	0/1229
3	C	0.26	0/1574	0.51	0/2136
3	F	0.29	0/1570	0.51	0/2128
3	I	0.35	0/541	0.62	0/732
All	All	0.27	0/18150	0.46	0/24499

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	2518	0	2523	72	0
1	D	2544	0	2552	68	0
1	G	2543	0	2552	104	0
1	J	2500	0	2511	107	0
2	B	1594	0	1555	80	0
2	E	1590	0	1557	75	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	H	885	0	843	60	0
3	C	1539	0	1465	70	0
3	F	1535	0	1466	69	0
3	I	529	0	491	68	0
All	All	17777	0	17515	736	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 21.

The worst 5 of 736 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)
3:I:57:VAL:CG2	3:I:59:TRP:CH2	1.76	1.62
3:I:57:VAL:HG21	3:I:59:TRP:CH2	1.37	1.49
3:I:57:VAL:HG23	3:I:59:TRP:CZ3	1.48	1.44
3:I:53:VAL:HG21	3:I:116:ASP:CG	1.48	1.33
2:E:56:VAL:HG23	2:E:65:GLU:O	1.24	1.30

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	297/307 (97%)	289 (97%)	8 (3%)	0	100	100
1	D	300/307 (98%)	295 (98%)	4 (1%)	1 (0%)	41	72
1	G	300/307 (98%)	289 (96%)	11 (4%)	0	100	100
1	J	295/307 (96%)	284 (96%)	10 (3%)	1 (0%)	41	72
2	B	205/254 (81%)	193 (94%)	10 (5%)	2 (1%)	15	49
2	E	204/254 (80%)	192 (94%)	10 (5%)	2 (1%)	15	49

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	H	111/254 (44%)	107 (96%)	4 (4%)	0	100	100
3	C	195/238 (82%)	183 (94%)	10 (5%)	2 (1%)	15	49
3	F	192/238 (81%)	180 (94%)	12 (6%)	0	100	100
3	I	61/238 (26%)	56 (92%)	5 (8%)	0	100	100
All	All	2160/2704 (80%)	2068 (96%)	84 (4%)	8 (0%)	34	68

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	E	179	ASN
3	C	52	ASN
3	C	228	PRO
1	D	336	PRO
2	E	224	PRO

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	281/286 (98%)	270 (96%)	11 (4%)	32	63
1	D	284/286 (99%)	277 (98%)	7 (2%)	47	74
1	G	284/286 (99%)	261 (92%)	23 (8%)	11	39
1	J	279/286 (98%)	260 (93%)	19 (7%)	16	47
2	B	187/227 (82%)	177 (95%)	10 (5%)	22	56
2	E	186/227 (82%)	176 (95%)	10 (5%)	22	55
2	H	101/227 (44%)	94 (93%)	7 (7%)	15	47
3	C	175/211 (83%)	173 (99%)	2 (1%)	73	86
3	F	175/211 (83%)	164 (94%)	11 (6%)	18	49
3	I	58/211 (28%)	51 (88%)	7 (12%)	5	21
All	All	2010/2458 (82%)	1903 (95%)	107 (5%)	22	56

5 of 107 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	261	LYS
1	G	457	LEU
1	J	322	PHE
1	G	308	GLU
1	G	357	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	292	ASN
1	G	372	ASN
1	J	455	HIS
3	I	113	GLN
2	E	179	ASN

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(Å^2)	Q<0.9
1	A	299/307 (97%)	0.24	10 (3%)	46	45	28, 67, 115, 149	0
1	D	302/307 (98%)	0.28	14 (4%)	32	33	36, 72, 116, 173	0
1	G	302/307 (98%)	0.30	7 (2%)	60	60	41, 85, 140, 181	0
1	J	297/307 (96%)	0.35	11 (3%)	41	40	69, 109, 149, 189	0
2	B	209/254 (82%)	0.52	17 (8%)	12	12	31, 72, 135, 165	0
2	E	208/254 (81%)	0.42	17 (8%)	11	12	34, 86, 130, 144	0
2	H	113/254 (44%)	0.70	14 (12%)	4	3	56, 86, 153, 180	0
3	C	199/238 (83%)	0.44	14 (7%)	16	17	34, 79, 130, 176	0
3	F	198/238 (83%)	0.38	8 (4%)	38	37	38, 74, 130, 170	0
3	I	67/238 (28%)	0.83	9 (13%)	3	3	81, 117, 163, 192	0
All	All	2194/2704 (81%)	0.38	121 (5%)	25	26	28, 82, 140, 192	0

The worst 5 of 121 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	64	LEU	6.8
3	I	60	TYR	6.5
1	J	213	THR	6.3
1	J	211	TYR	6.1
2	B	62	LYS	5.5

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.