



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 4, 2021 – 10:03 AM GMT

PDB ID : 6TQ0
Title : Crystal structure of the human Arc N-lobe bound to repeat peptide 5 from GKAP
Authors : Hallin, E.I.; Bramham, C.R.; Kursula, P.
Deposited on : 2019-12-15
Resolution : 1.95 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 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 : 2.16
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) : 2.16

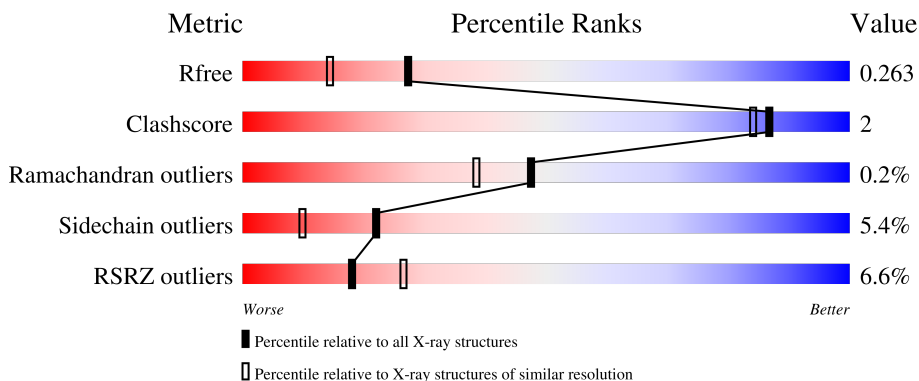
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.95 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	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)
RSRZ outliers	127900	2539 (1.96-1.96)

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 $\leq 5\%$. 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	75	 85% 11%
1	C	75	 80% 8% 11%
1	E	75	 73% 17% 9%
1	G	75	 80% 8% 12%
1	I	75	 81% 8% 11%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	K	75	
1	M	75	
1	O	75	
2	B	8	
2	D	8	
2	F	8	
2	H	8	
2	J	8	
2	L	8	
2	N	8	
2	P	8	

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 9873 atoms, of which 4741 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 Activity-regulated cytoskeleton-associated protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	A	67	1119	376	538	96	108	1	0	0	0
1	C	67	1119	376	538	96	108	1	0	0	0
1	E	68	1131	380	542	97	111	1	0	0	0
1	G	66	1105	372	531	95	106	1	0	0	0
1	I	67	1119	376	538	96	108	1	0	0	0
1	K	66	1105	372	531	95	106	1	0	0	0
1	M	67	1119	376	538	96	108	1	0	0	0
1	O	65	1097	370	528	94	104	1	0	0	0

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	203	GLY	-	expression tag	UNP Q7LC44
A	204	ALA	-	expression tag	UNP Q7LC44
A	205	MET	-	expression tag	UNP Q7LC44
A	206	GLY	-	expression tag	UNP Q7LC44
C	203	GLY	-	expression tag	UNP Q7LC44
C	204	ALA	-	expression tag	UNP Q7LC44
C	205	MET	-	expression tag	UNP Q7LC44
C	206	GLY	-	expression tag	UNP Q7LC44
E	203	GLY	-	expression tag	UNP Q7LC44
E	204	ALA	-	expression tag	UNP Q7LC44
E	205	MET	-	expression tag	UNP Q7LC44
E	206	GLY	-	expression tag	UNP Q7LC44
G	203	GLY	-	expression tag	UNP Q7LC44

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	204	ALA	-	expression tag	UNP Q7LC44
G	205	MET	-	expression tag	UNP Q7LC44
G	206	GLY	-	expression tag	UNP Q7LC44
I	203	GLY	-	expression tag	UNP Q7LC44
I	204	ALA	-	expression tag	UNP Q7LC44
I	205	MET	-	expression tag	UNP Q7LC44
I	206	GLY	-	expression tag	UNP Q7LC44
K	203	GLY	-	expression tag	UNP Q7LC44
K	204	ALA	-	expression tag	UNP Q7LC44
K	205	MET	-	expression tag	UNP Q7LC44
K	206	GLY	-	expression tag	UNP Q7LC44
M	203	GLY	-	expression tag	UNP Q7LC44
M	204	ALA	-	expression tag	UNP Q7LC44
M	205	MET	-	expression tag	UNP Q7LC44
M	206	GLY	-	expression tag	UNP Q7LC44
O	203	GLY	-	expression tag	UNP Q7LC44
O	204	ALA	-	expression tag	UNP Q7LC44
O	205	MET	-	expression tag	UNP Q7LC44
O	206	GLY	-	expression tag	UNP Q7LC44

- Molecule 2 is a protein called repeat peptide 5 from GKAP.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
2	B	8	Total	C	H	N	O	S	0	0	0
			134	41	67	14	9	3			
2	D	6	Total	C	H	N	O	S	0	0	0
			93	30	46	9	6	2			
2	F	7	Total	C	H	N	O	S	0	0	0
			118	36	59	13	8	2			
2	H	7	Total	C	H	N	O	S	0	0	0
			118	36	59	13	8	2			
2	J	8	Total	C	H	N	O	S	0	0	0
			134	41	67	14	9	3			
2	L	7	Total	C	H	N	O	S	0	0	0
			118	36	59	13	8	2			
2	N	6	Total	C	H	N	O	S	0	0	0
			93	30	46	9	6	2			
2	P	7	Total	C	H	N	O	S	0	0	0
			109	35	54	10	7	3			

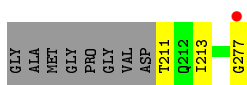
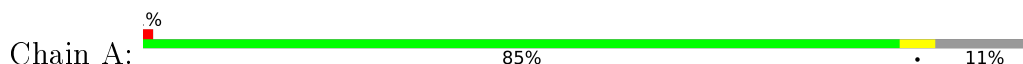
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	12	Total O 12 12	0	0
3	C	7	Total O 7 7	0	0
3	D	1	Total O 1 1	0	0
3	E	2	Total O 2 2	0	0
3	G	5	Total O 5 5	0	0
3	I	6	Total O 6 6	0	0
3	K	5	Total O 5 5	0	0
3	M	2	Total O 2 2	0	0
3	O	2	Total O 2 2	0	0

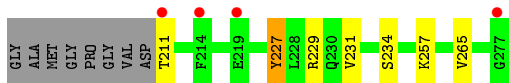
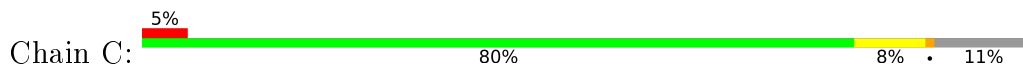
3 Residue-property plots [i](#)

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.

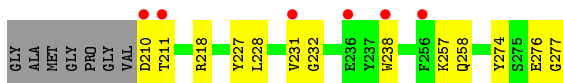
- Molecule 1: Activity-regulated cytoskeleton-associated protein



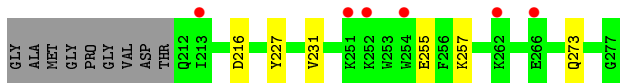
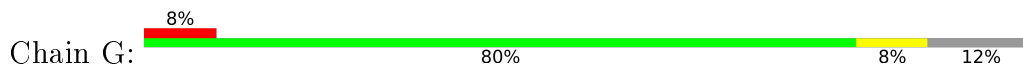
- Molecule 1: Activity-regulated cytoskeleton-associated protein



- Molecule 1: Activity-regulated cytoskeleton-associated protein



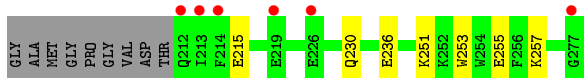
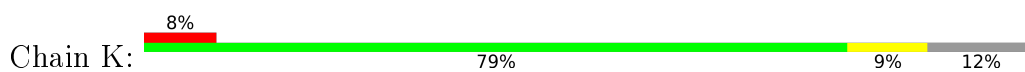
- Molecule 1: Activity-regulated cytoskeleton-associated protein



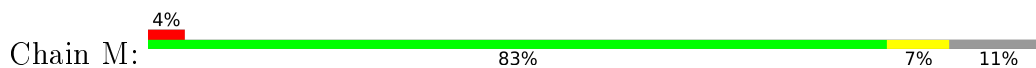
- Molecule 1: Activity-regulated cytoskeleton-associated protein



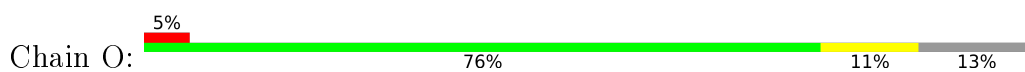
- Molecule 1: Activity-regulated cytoskeleton-associated protein



- Molecule 1: Activity-regulated cytoskeleton-associated protein



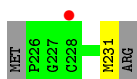
- Molecule 1: Activity-regulated cytoskeleton-associated protein



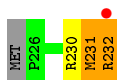
- Molecule 2: repeat peptide 5 from GKAP



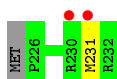
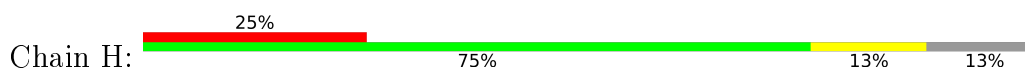
- Molecule 2: repeat peptide 5 from GKAP



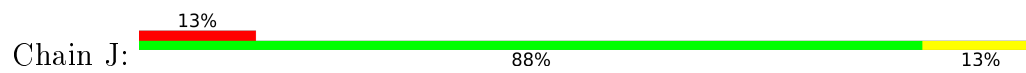
- Molecule 2: repeat peptide 5 from GKAP



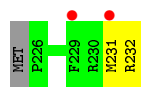
- Molecule 2: repeat peptide 5 from GKAP



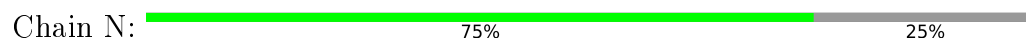
- Molecule 2: repeat peptide 5 from GKAP



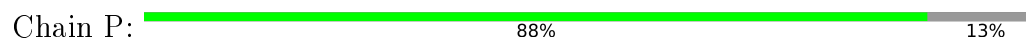
- Molecule 2: repeat peptide 5 from GKAP



- Molecule 2: repeat peptide 5 from GKAP



- Molecule 2: repeat peptide 5 from GKAP



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	39.31Å 192.30Å 39.56Å 90.00° 94.80° 90.00°	Depositor
Resolution (Å)	100.00 – 1.95 39.42 – 1.95	Depositor EDS
% Data completeness (in resolution range)	98.2 (100.00-1.95) 98.3 (39.42-1.95)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.04	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.24 (at 1.95Å)	Xtrriage
Refinement program	PHENIX 1.16-3549	Depositor
R, R_{free}	0.221 , 0.263 0.221 , 0.263	Depositor DCC
R_{free} test set	1998 reflections (4.79%)	wwPDB-VP
Wilson B-factor (Å ²)	45.0	Xtrriage
Anisotropy	0.291	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.41 , 55.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.088 for l,-k,h	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	9873	wwPDB-VP
Average B, all atoms (Å ²)	72.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.82% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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	0.81	0/600	0.58	0/806
1	C	0.69	0/600	0.54	0/806
1	E	0.66	0/608	0.48	0/817
1	G	0.58	0/593	0.50	0/796
1	I	0.78	2/600 (0.3%)	0.55	0/806
1	K	0.61	0/593	0.48	0/796
1	M	0.59	0/600	0.50	0/806
1	O	0.58	0/588	0.48	0/791
2	B	0.50	0/68	0.48	0/86
2	D	0.44	0/48	0.47	0/61
2	F	0.73	0/60	0.62	0/75
2	H	0.43	0/60	0.57	0/75
2	J	0.51	0/68	0.58	0/86
2	L	0.31	0/60	0.43	0/75
2	N	0.46	0/48	0.44	0/61
2	P	0.45	0/56	0.50	0/72
All	All	0.65	2/5250 (0.0%)	0.52	0/7015

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	I	235	GLU	CG-CD	6.08	1.61	1.51
1	I	235	GLU	CD-OE1	5.72	1.31	1.25

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	581	538	538	1	0
1	C	581	538	538	5	0
1	E	589	542	542	7	0
1	G	574	531	531	2	0
1	I	581	538	538	3	0
1	K	574	531	531	2	1
1	M	581	538	538	2	0
1	O	569	528	528	2	1
2	B	67	67	67	1	0
2	D	47	46	46	0	0
2	F	59	59	59	3	0
2	H	59	59	59	0	0
2	J	67	67	67	0	0
2	L	59	59	59	0	0
2	N	47	46	46	0	0
2	P	55	54	54	0	0
3	A	12	0	0	0	0
3	C	7	0	0	1	0
3	D	1	0	0	0	0
3	E	2	0	0	0	0
3	G	5	0	0	0	0
3	I	6	0	0	0	0
3	K	5	0	0	0	0
3	M	2	0	0	0	0
3	O	2	0	0	0	0
All	All	5132	4741	4741	21	1

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.

The worst 5 of 21 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:C:229:ARG:NH1	1:G:216:ASP:OD2	2.20	0.75
1:M:218:ARG:NE	1:M:276:GLU:OE2	2.31	0.64
1:K:251:LYS:NZ	1:K:255:GLU:OE1	2.32	0.62
1:A:277:GLY:O	1:O:234:SER:OG	2.15	0.61
1:E:227:TYR:O	1:E:231:VAL:HG13	2.08	0.54

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:K:215:GLU:OE2	1:O:252:LYS:NZ[1_454]	2.13	0.07

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	65/75 (87%)	65 (100%)	0	0	100	100
1	C	65/75 (87%)	65 (100%)	0	0	100	100
1	E	66/75 (88%)	65 (98%)	0	1 (2%)	10	3
1	G	64/75 (85%)	64 (100%)	0	0	100	100
1	I	65/75 (87%)	65 (100%)	0	0	100	100
1	K	64/75 (85%)	64 (100%)	0	0	100	100
1	M	65/75 (87%)	65 (100%)	0	0	100	100
1	O	63/75 (84%)	63 (100%)	0	0	100	100
2	B	6/8 (75%)	6 (100%)	0	0	100	100
2	D	4/8 (50%)	4 (100%)	0	0	100	100
2	F	5/8 (62%)	5 (100%)	0	0	100	100
2	H	5/8 (62%)	4 (80%)	1 (20%)	0	100	100
2	J	6/8 (75%)	5 (83%)	1 (17%)	0	100	100
2	L	5/8 (62%)	5 (100%)	0	0	100	100
2	N	4/8 (50%)	4 (100%)	0	0	100	100
2	P	5/8 (62%)	5 (100%)	0	0	100	100
All	All	557/664 (84%)	554 (100%)	2 (0%)	1 (0%)	47	38

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	211	THR

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	61/65 (94%)	59 (97%)	2 (3%)	38	26
1	C	61/65 (94%)	59 (97%)	2 (3%)	38	26
1	E	62/65 (95%)	59 (95%)	3 (5%)	25	12
1	G	60/65 (92%)	56 (93%)	4 (7%)	16	5
1	I	61/65 (94%)	60 (98%)	1 (2%)	62	58
1	K	60/65 (92%)	57 (95%)	3 (5%)	24	11
1	M	61/65 (94%)	60 (98%)	1 (2%)	62	58
1	O	60/65 (92%)	56 (93%)	4 (7%)	16	5
2	B	7/7 (100%)	6 (86%)	1 (14%)	3	0
2	D	5/7 (71%)	4 (80%)	1 (20%)	1	0
2	F	6/7 (86%)	3 (50%)	3 (50%)	0	0
2	H	6/7 (86%)	5 (83%)	1 (17%)	2	0
2	J	7/7 (100%)	6 (86%)	1 (14%)	3	0
2	L	6/7 (86%)	4 (67%)	2 (33%)	0	0
2	N	5/7 (71%)	5 (100%)	0	100	100
2	P	6/7 (86%)	6 (100%)	0	100	100
All	All	534/576 (93%)	505 (95%)	29 (5%)	22	10

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	255	GLU
2	H	231	MET
1	O	227	TYR
1	G	257	LYS
1	I	227	TYR

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	67/75 (89%)	0.59	1 (1%) 73 81	39, 54, 77, 94	0
1	C	67/75 (89%)	0.68	4 (5%) 21 30	40, 59, 80, 95	0
1	E	68/75 (90%)	0.83	6 (8%) 10 16	45, 63, 84, 91	0
1	G	66/75 (88%)	0.64	6 (9%) 9 15	44, 63, 92, 108	0
1	I	67/75 (89%)	0.60	1 (1%) 73 81	40, 57, 81, 115	0
1	K	66/75 (88%)	0.74	6 (9%) 9 15	46, 68, 94, 132	0
1	M	67/75 (89%)	0.53	3 (4%) 33 43	46, 62, 79, 95	0
1	O	65/75 (86%)	0.68	4 (6%) 20 29	47, 67, 100, 103	0
2	B	8/8 (100%)	1.32	1 (12%) 3 6	56, 73, 92, 106	0
2	D	6/8 (75%)	1.84	1 (16%) 1 2	70, 82, 92, 96	0
2	F	7/8 (87%)	1.11	1 (14%) 2 4	57, 69, 85, 103	0
2	H	7/8 (87%)	1.41	2 (28%) 0 0	85, 94, 114, 137	0
2	J	8/8 (100%)	0.99	1 (12%) 3 6	56, 83, 88, 101	0
2	L	7/8 (87%)	1.79	2 (28%) 0 0	91, 101, 126, 168	0
2	N	6/8 (75%)	0.80	0 100 100	64, 75, 90, 108	0
2	P	7/8 (87%)	0.81	0 100 100	76, 79, 119, 119	0
All	All	589/664 (88%)	0.72	39 (6%) 18 26	39, 63, 95, 168	0

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	277	GLY	6.5
1	I	277	GLY	6.3
1	O	256	PHE	5.3
2	B	232	ARG	5.0
1	M	256	PHE	4.7

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.