wwPDB X-ray Structure Validation Summary Report (i)

Jun 12, 2024 – 01:40 AM EDT

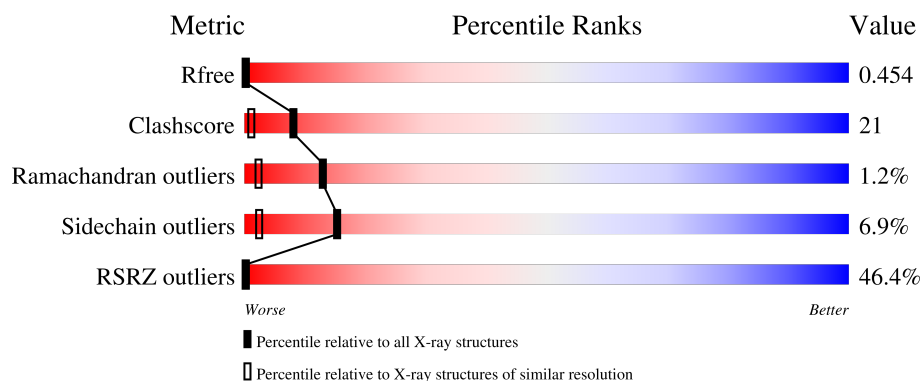|		
PDB ID	:	1AWR
Title	:	CYPA COMPLEXED WITH HAGPIA
Authors	:	Vajdos, F.F.
Deposited on	:	1997-10-04
Resolution	:	1.58 Å(reported)

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 1.58 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		0.454
Clashscore		21
Ramachandran outliers		1.2%
Sidechain outliers		6.9%
RSRZ outliers		46.4%

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	5534 (1.60-1.56)
Clashscore	141614	5861 (1.60-1.56)
Ramachandran outliers	138981	5708 (1.60-1.56)
Sidechain outliers	138945	5703 (1.60-1.56)
RSRZ outliers	127900	5431 (1.60-1.56)

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	164	24% / 65% 32% •
1	B	164	76% / 49% 45% 6%
1	C	164	51% / 62% 35% •
1	D	164	37% / 66% 30% •
1	E	164	24% / 70% 25% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	164	67% 57% 40%
2	G	6	50% 50% 33% 17%
2	H	6	50% 83% 17%
2	I	6	33% 83% 17%
2	J	6	33% 67% 33%
2	K	6	33% 100%
2	L	6	67% 67% 33%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 7963 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 CYCLOPHILIN A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	164	Total	C	N	O	S	0	0	0
			1258	797	217	236	8			
1	B	164	Total	C	N	O	S	0	0	0
			1258	797	217	236	8			
1	C	164	Total	C	N	O	S	0	0	0
			1258	797	217	236	8			
1	D	164	Total	C	N	O	S	0	0	0
			1258	797	217	236	8			
1	E	164	Total	C	N	O	S	0	0	0
			1258	797	217	236	8			
1	F	164	Total	C	N	O	S	0	0	0
			1258	797	217	236	8			

- Molecule 2 is a protein called PEPTIDE FROM THE HIV-1 CAPSID PROTEIN.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	G	6	Total	C	N	O	0	0	0
			40	25	8	7			
2	H	6	Total	C	N	O	0	0	0
			40	25	8	7			
2	I	6	Total	C	N	O	0	0	0
			40	25	8	7			
2	J	6	Total	C	N	O	0	0	0
			40	25	8	7			
2	K	6	Total	C	N	O	0	0	0
			40	25	8	7			
2	L	6	Total	C	N	O	0	0	0
			40	25	8	7			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	33	Total 33	O 33	0	0
3	B	22	Total 22	O 22	0	0
3	C	32	Total 32	O 32	0	0
3	D	20	Total 20	O 20	0	0
3	E	29	Total 29	O 29	0	0
3	F	30	Total 30	O 30	0	0
3	G	2	Total 2	O 2	0	0
3	H	1	Total 1	O 1	0	0
3	I	1	Total 1	O 1	0	0
3	J	2	Total 2	O 2	0	0
3	K	2	Total 2	O 2	0	0
3	L	1	Total 1	O 1	0	0

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: CYCLOPHILIN A
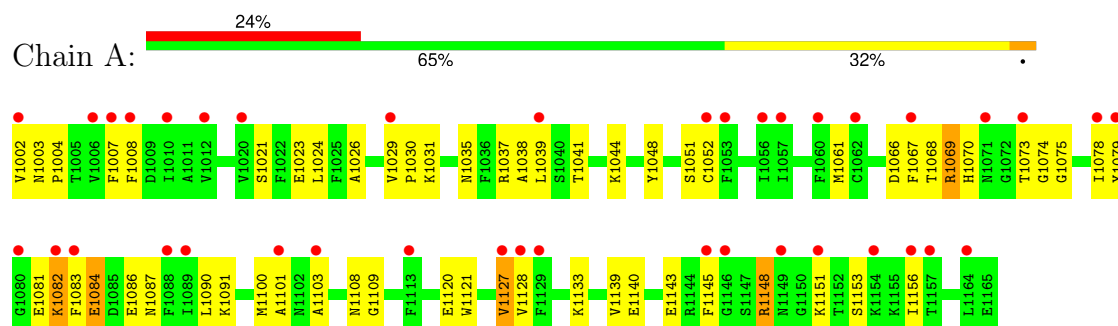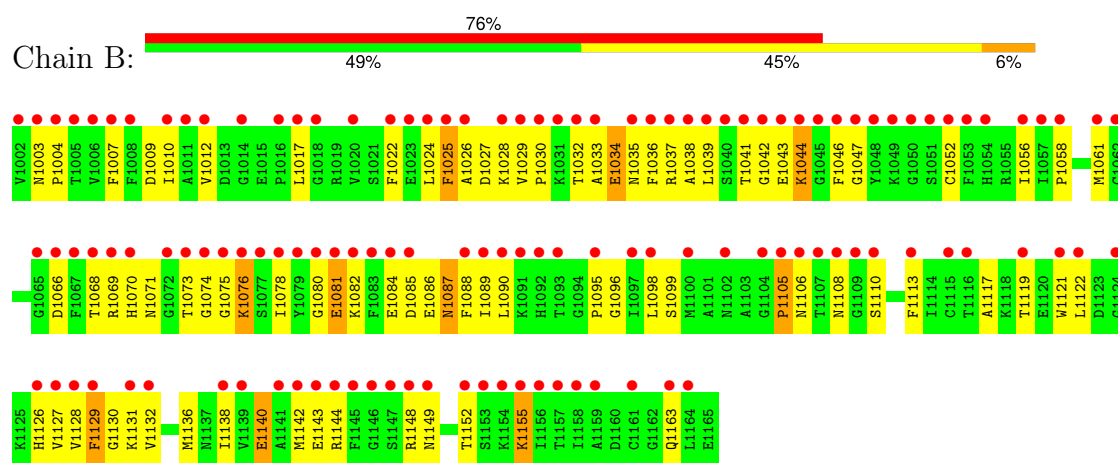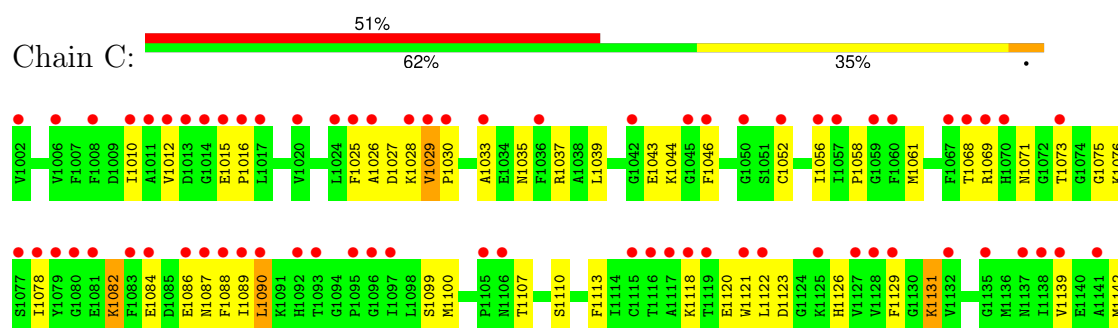
- Molecule 1: CYCLOPHILIN A

- Molecule 1: CYCLOPHILIN A

• Molecule 1: CYCLOPHILIN A

Chain D:

• Molecule 1: CYCLOPHILIN A

Chain E:

• Molecule 1: CYCLOPHILIN A

Chain F:

• Molecule 2: PEPTIDE FROM THE HIV-1 CAPSID PROTEIN

Chain G:

• Molecule 2: PEPTIDE FROM THE HIV-1 CAPSID PROTEIN

Chain H:

• Molecule 2: PEPTIDE FROM THE HIV-1 CAPSID PROTEIN

Chain I:

• Molecule 2: PEPTIDE FROM THE HIV-1 CAPSID PROTEIN

Chain J:

• Molecule 2: PEPTIDE FROM THE HIV-1 CAPSID PROTEIN

Chain K:

• Molecule 2: PEPTIDE FROM THE HIV-1 CAPSID PROTEIN

Chain L:
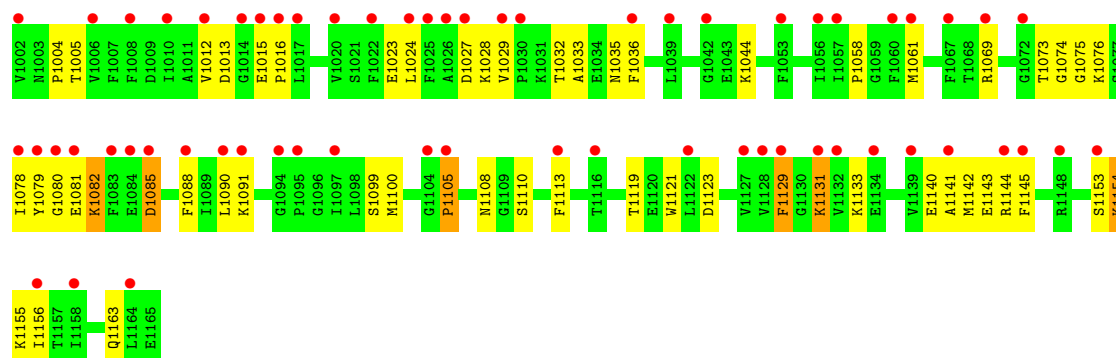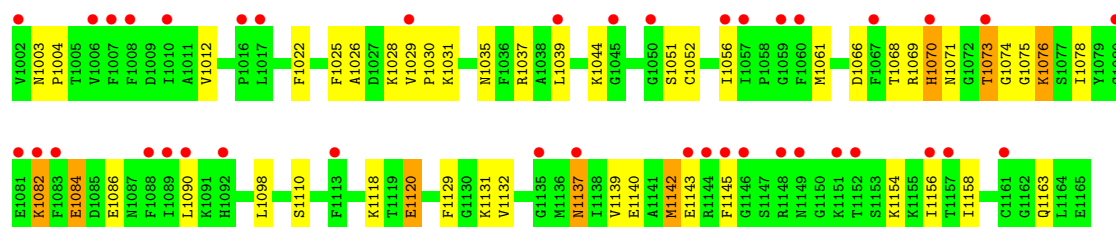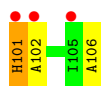
4 Data and refinement statistics ⓘ

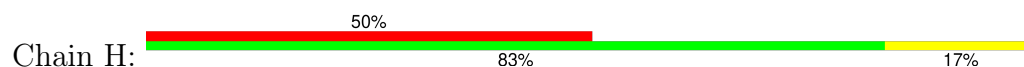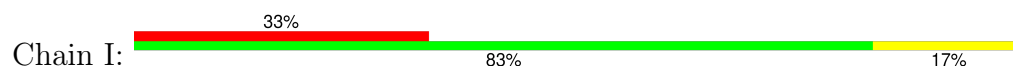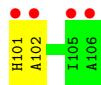| Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, α, β, γ	72.91Å 72.91Å 188.56Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	15.00 – 1.58 14.96 – 1.53	Depositor EDS
% Data completeness (in resolution range)	88.3 (15.00-1.58) 79.5 (14.96-1.53)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	1.00 (at 1.53Å)	Xtriage
Refinement program	X-PLOR 3.843	Depositor
R, R_{free}	0.394 , 0.461 0.391 , 0.454	Depositor DCC
R_{free} test set	6255 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	12.9	Xtriage
Anisotropy	0.289	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 46.1	EDS
L-test for twinning[2]	$< |L| > = 0.47, < L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.044 for h,-k,-l	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	7963	wwPDB-VP
Average B, all atoms (Å^2)	16.0	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 82.07 % 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 3.0506e-07. 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	0.62	0/1286	0.76	0/1723
1	B	0.63	0/1286	0.77	0/1723
1	C	0.64	0/1286	0.77	0/1723
1	D	0.58	0/1286	0.73	0/1723
1	E	0.64	0/1286	0.76	0/1723
1	F	0.60	0/1286	0.79	0/1723
2	G	0.63	0/41	0.89	0/54
2	H	0.65	0/41	0.72	0/54
2	I	0.74	0/41	0.90	0/54
2	J	0.78	0/41	0.91	0/54
2	K	0.76	0/41	0.92	0/54
2	L	0.57	0/41	0.62	0/54
All	All	0.62	0/7962	0.76	0/10662

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	1258	0	1225	41	0
1	B	1258	0	1225	73	0
1	C	1258	0	1225	62	0
1	D	1258	0	1225	47	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	1258	0	1225	39	0
1	F	1258	0	1225	65	0
2	G	40	0	37	3	0
2	H	40	0	37	1	0
2	I	40	0	37	1	0
2	J	40	0	37	2	0
2	K	40	0	37	0	0
2	L	40	0	37	1	0
3	A	33	0	0	7	0
3	B	22	0	0	6	0
3	C	32	0	0	4	0
3	D	20	0	0	3	0
3	E	29	0	0	0	0
3	F	30	0	0	12	0
3	G	2	0	0	0	0
3	H	1	0	0	0	0
3	I	1	0	0	0	0
3	J	2	0	0	0	0
3	K	2	0	0	0	0
3	L	1	0	0	0	0
All	All	7963	0	7572	327	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 327 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:F:1100:MET:HB2	1:F:1127:VAL:HG23	1.39	1.02
1:E:1069:ARG:HG2	1:E:1073:THR:HG22	1.48	0.93
1:F:1052:CYS:HB3	1:F:1155:LYS:HE2	1.51	0.92
1:C:1148:ARG:HH11	1:C:1148:ARG:HA	1.36	0.90
1:F:1048:TYR:HA	3:F:7016:HOH:O	1.70	0.90

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	162/164 (99%)	147 (91%)	14 (9%)	1 (1%)	25	8
1	B	162/164 (99%)	141 (87%)	18 (11%)	3 (2%)	8	1
1	C	162/164 (99%)	147 (91%)	12 (7%)	3 (2%)	8	1
1	D	162/164 (99%)	146 (90%)	13 (8%)	3 (2%)	8	1
1	E	162/164 (99%)	148 (91%)	13 (8%)	1 (1%)	25	8
1	F	162/164 (99%)	147 (91%)	14 (9%)	1 (1%)	25	8
2	G	4/6 (67%)	3 (75%)	1 (25%)	0	100	100
2	H	4/6 (67%)	3 (75%)	1 (25%)	0	100	100
2	I	4/6 (67%)	4 (100%)	0	0	100	100
2	J	4/6 (67%)	4 (100%)	0	0	100	100
2	K	4/6 (67%)	4 (100%)	0	0	100	100
2	L	4/6 (67%)	3 (75%)	1 (25%)	0	100	100
All	All	996/1020 (98%)	897 (90%)	87 (9%)	12 (1%)	13	2

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	1025	PHE
1	B	1105	PRO
1	D	1105	PRO
1	B	1071	ASN
1	D	1080	GLY

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	132/132 (100%)	123 (93%)	9 (7%)	16	2
1	B	132/132 (100%)	122 (92%)	10 (8%)	13	2
1	C	132/132 (100%)	125 (95%)	7 (5%)	22	4
1	D	132/132 (100%)	125 (95%)	7 (5%)	22	4
1	E	132/132 (100%)	122 (92%)	10 (8%)	13	2
1	F	132/132 (100%)	122 (92%)	10 (8%)	13	2
2	G	3/3 (100%)	2 (67%)	1 (33%)	0	0
2	H	3/3 (100%)	3 (100%)	0	100	100
2	I	3/3 (100%)	3 (100%)	0	100	100
2	J	3/3 (100%)	2 (67%)	1 (33%)	0	0
2	K	3/3 (100%)	3 (100%)	0	100	100
2	L	3/3 (100%)	2 (67%)	1 (33%)	0	0
All	All	810/810 (100%)	754 (93%)	56 (7%)	15	2

5 of 56 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	1085	ASP
2	L	101	HIS
1	E	1073	THR
2	J	101	HIS
1	F	1127	VAL

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	E	1137	ASN
1	F	1070	HIS
1	F	1137	ASN
1	F	1108	ASN
1	C	1149	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(Å²)	Q<0.9
1	A	164/164 (100%)	1.45	39 (23%)	0	0	5, 13, 20, 23	0
1	B	164/164 (100%)	3.23	124 (75%)	0	0	11, 20, 28, 34	0
1	C	164/164 (100%)	2.27	83 (50%)	0	0	5, 16, 25, 35	0
1	D	164/164 (100%)	1.94	61 (37%)	0	0	7, 18, 25, 33	0
1	E	164/164 (100%)	1.52	40 (24%)	0	0	3, 12, 21, 27	0
1	F	164/164 (100%)	2.90	110 (67%)	0	0	10, 18, 26, 30	0
2	G	6/6 (100%)	2.34	3 (50%)	0	0	13, 17, 21, 29	0
2	H	6/6 (100%)	2.43	3 (50%)	0	0	11, 14, 22, 24	0
2	I	6/6 (100%)	2.08	2 (33%)	0	0	6, 8, 11, 26	0
2	J	6/6 (100%)	1.71	2 (33%)	0	0	9, 10, 12, 18	0
2	K	6/6 (100%)	1.88	2 (33%)	0	0	7, 9, 20, 23	0
2	L	6/6 (100%)	2.94	4 (66%)	0	0	16, 19, 24, 24	0
All	All	1020/1020 (100%)	2.22	473 (46%)	0	0	3, 17, 26, 35	0

The worst 5 of 473 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	1080	GLY	12.4
1	B	1024	LEU	11.2
1	B	1025	PHE	10.2
1	F	1029	VAL	8.9
1	F	1078	ILE	8.3

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

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.