wwPDB X-ray Structure Validation Summary Report (i)

Mar 3, 2025 – 04:46 PM JST

PDB ID	:	8Z92
Title	:	Crystal structure of CrtAgo/TIR-APAZ in complex with guide DNA and 16-nt target DNA
Authors	:	Hu, R.; Chen, J.; Liu, L.
Deposited on	:	2024-04-22
Resolution	:	3.85 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 (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:

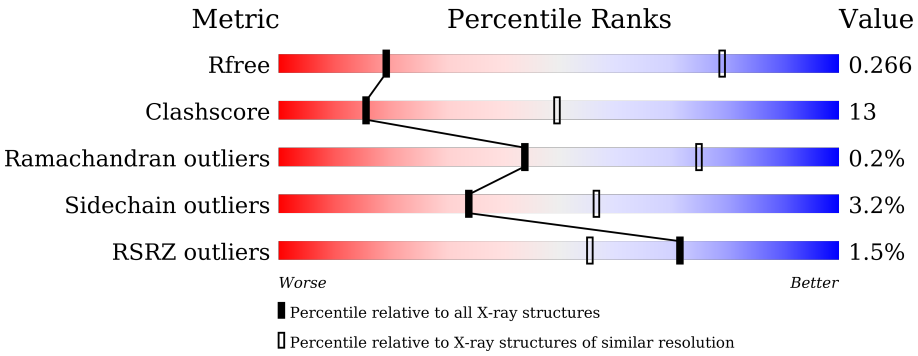
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.21
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.41.2

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 3.85 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.



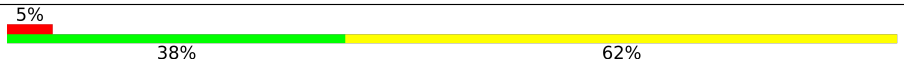
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	1056 (4.02-3.70)
Clashscore	180529	1117 (4.02-3.70)
Ramachandran outliers	177936	1077 (4.02-3.70)
Sidechain outliers	177891	1070 (4.02-3.70)
RSRZ outliers	164620	1056 (4.02-3.70)

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	507	62% 28% • 8%
1	B	507	4% 65% 27% • 6%
2	E	421	68% 30% •
2	G	421	% 69% 29% •
3	D	16	19% 19% 75% 6%
3	H	16	6% 38% 62%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	F	21	
4	I	21	

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 16050 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 Piwi domain-containing protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	476	Total	C	N	O	S	0	0	0
			3842	2488	639	703	12			
1	A	464	Total	C	N	O	S	0	0	0
			3749	2431	623	683	12			

- Molecule 2 is a protein called TIR domain-containing protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	419	Total	C	N	O	S	0	0	0
			3465	2245	586	623	11			
2	G	419	Total	C	N	O	S	0	0	0
			3480	2255	589	625	11			

- Molecule 3 is a DNA chain called DNA (5'-D(*CP*AP*AP*CP*CP*TP*AP*CP*TP*AP*CP*CP*TP*CP*AP*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	16	Total	C	N	O	P	0	0	0
			315	153	54	93	15			
3	H	16	Total	C	N	O	P	0	0	0
			315	153	54	93	15			

- Molecule 4 is a DNA chain called DNA (5'-D(P*TP*GP*AP*GP*GP*TP*AP*GP*TP*AP*GP*GP*TP*TP*GP*TP*AP*TP*AP*GP*T)-3').

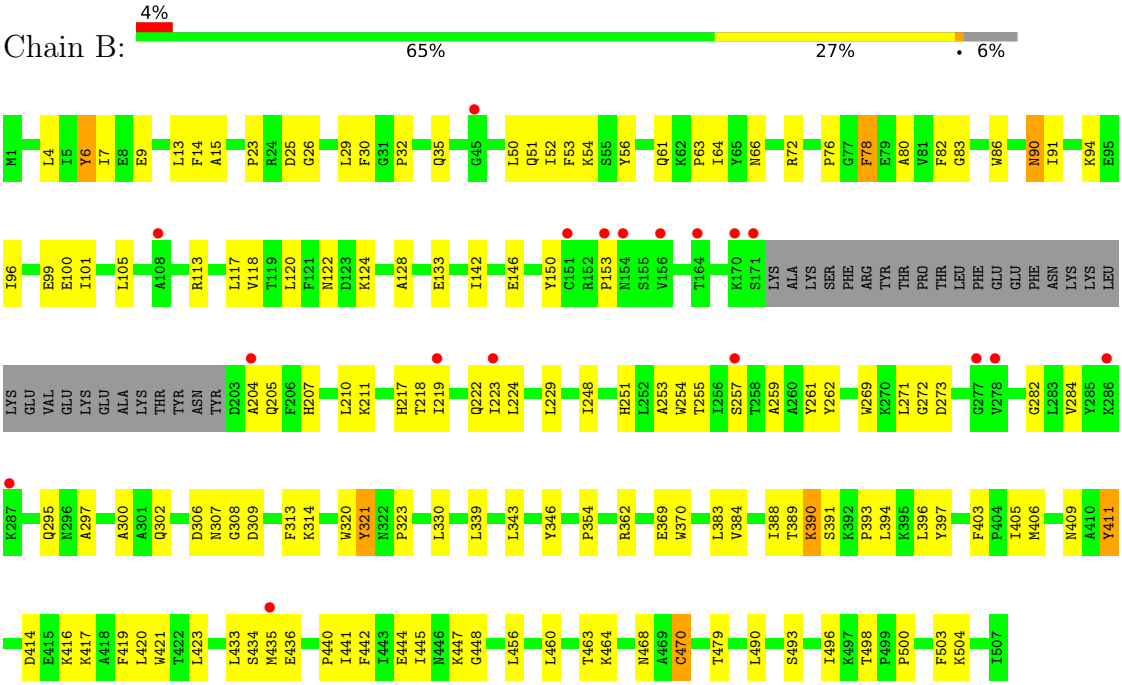
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	F	21	Total	C	N	O	P	0	0	0
			442	210	81	130	21			
4	I	21	Total	C	N	O	P	0	0	0
			442	210	81	130	21			

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.

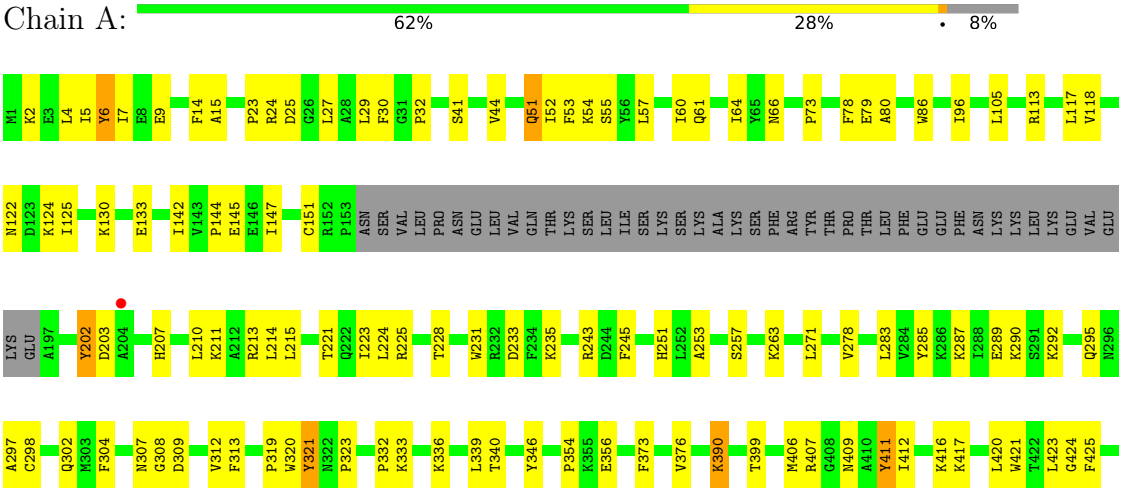
• Molecule 1: Piwi domain-containing protein

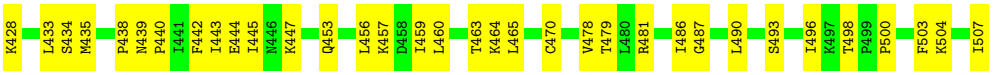
Chain B:



• Molecule 1: Piwi domain-containing protein

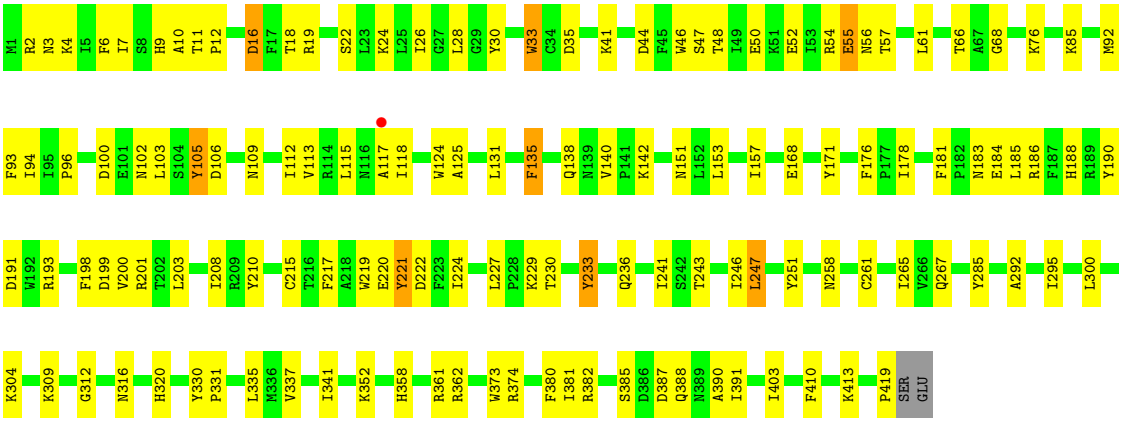
Chain A:





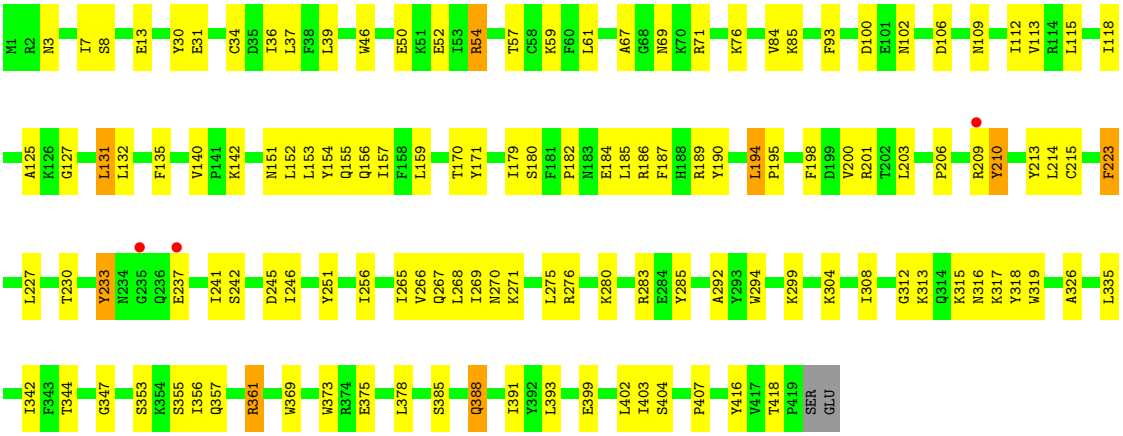
• Molecule 2: TIR domain-containing protein

Chain E:



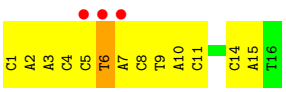
• Molecule 2: TIR domain-containing protein

Chain G:



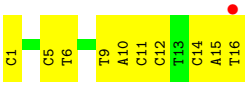
• Molecule 3: DNA (5’-D(*CP*AP*AP*CP*CP*TP*AP*CP*TP*AP*CP*CP*TP*CP*AP*T)-3’)

Chain D:



• Molecule 3: DNA (5’-D(*CP*AP*AP*CP*CP*TP*AP*CP*TP*AP*CP*CP*TP*CP*AP*T)-3’)

Chain H:



• Molecule 4: DNA (5'-D(P*TP*GP*AP*GP*GP*TP*AP*GP*TP*AP*GP*GP*TP*TP*GP*TP*AP*TP*AP*GP*T)-3')

Chain F:

• Molecule 4: DNA (5'-D(P*TP*GP*AP*GP*GP*TP*AP*GP*TP*AP*GP*GP*TP*TP*GP*TP*AP*TP*AP*GP*T)-3')

Chain I:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	201.18Å 287.27Å 111.77Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.25 – 3.85 46.25 – 3.85	Depositor EDS
% Data completeness (in resolution range)	64.8 (46.25-3.85) 77.5 (46.25-3.85)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.24 (at 3.88Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.247 , 0.265 0.250 , 0.266	Depositor DCC
R_{free} test set	29013 reflections (8.28%)	wwPDB-VP
Wilson B-factor (Å^2)	48.1	Xtriage
Anisotropy	0.234	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 13.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.39$, $< L^2 > = 0.22$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.79	EDS
Total number of atoms	16050	wwPDB-VP
Average B, all atoms (Å^2)	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.58% 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	0.30	0/3845	0.53	0/5206
1	B	0.28	0/3938	0.50	0/5332
2	E	0.28	0/3552	0.49	0/4794
2	G	0.29	0/3568	0.51	1/4814 (0.0%)
3	D	0.68	0/351	1.03	1/537 (0.2%)
3	H	0.60	0/351	1.01	0/537
4	F	0.83	1/496 (0.2%)	1.12	0/765
4	I	0.83	1/496 (0.2%)	1.06	0/765
All	All	0.36	2/16597 (0.0%)	0.60	2/22750 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	F	0	DT	OP3-P	-10.59	1.48	1.61
4	I	0	DT	OP3-P	-10.50	1.48	1.61

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	D	6	DT	O4'-C1'-N1	6.38	112.46	108.00
2	G	194	LEU	CA-CB-CG	6.14	129.43	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	203	ASP	Peptide

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	3749	0	3752	104	0
1	B	3842	0	3862	106	0
2	E	3465	0	3440	95	0
2	G	3480	0	3467	90	0
3	D	315	0	182	14	0
3	H	315	0	182	10	0
4	F	442	0	240	17	0
4	I	442	0	240	19	0
All	All	16050	0	15365	411	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 13.

The worst 5 of 411 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:41:SER:HG	1:A:86:TRP:HE1	0.91	0.89
1:B:210:LEU:HD22	1:B:223:ILE:HD11	1.57	0.86
1:A:9:GLU:OE1	1:A:464:LYS:NZ	2.11	0.83
2:G:30:TYR:HA	2:G:142:LYS:HD2	1.59	0.82
1:B:423:LEU:HD12	1:B:434:SER:HB2	1.59	0.82

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	460/507 (91%)	441 (96%)	18 (4%)	1 (0%)	44	75
1	B	472/507 (93%)	454 (96%)	17 (4%)	1 (0%)	44	75
2	E	417/421 (99%)	400 (96%)	16 (4%)	1 (0%)	44	75
2	G	417/421 (99%)	399 (96%)	17 (4%)	1 (0%)	44	75
All	All	1766/1856 (95%)	1694 (96%)	68 (4%)	4 (0%)	44	75

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	218	THR
2	E	191	ASP
2	G	388	GLN
1	A	202	TYR

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	401/446 (90%)	389 (97%)	12 (3%)	36	58
1	B	416/446 (93%)	405 (97%)	11 (3%)	41	62
2	E	374/387 (97%)	360 (96%)	14 (4%)	29	53
2	G	378/387 (98%)	365 (97%)	13 (3%)	32	55
All	All	1569/1666 (94%)	1519 (97%)	50 (3%)	34	57

5 of 50 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	G	106	ASP
2	G	315	LYS
1	A	442	PHE
2	G	131	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	G	223	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	G	270	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 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	464/507 (91%)	-0.08	1 (0%) 92 85	26, 33, 69, 114	0
1	B	476/507 (93%)	0.48	18 (3%) 44 36	31, 83, 132, 164	0
2	E	419/421 (99%)	0.13	1 (0%) 92 85	27, 55, 104, 126	0
2	G	419/421 (99%)	0.01	3 (0%) 84 71	26, 42, 86, 108	0
3	D	16/16 (100%)	1.35	3 (18%) 4 7	86, 122, 164, 172	0
3	H	16/16 (100%)	0.73	1 (6%) 27 25	31, 62, 171, 174	0
4	F	21/21 (100%)	0.84	0 100 100	44, 102, 135, 153	0
4	I	21/21 (100%)	0.76	1 (4%) 36 31	34, 71, 142, 155	0
All	All	1852/1930 (95%)	0.17	28 (1%) 71 55	26, 51, 117, 174	0

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	153	PRO	3.8
1	B	204	ALA	3.5
1	B	171	SER	3.3
1	B	287	LYS	3.2
3	D	5	DC	3.1

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.