wwPDB X-ray Structure Validation Summary Report (i)

Jun 23, 2024 – 08:34 AM EDT

PDB ID	:	5D1Z
Title	:	IsdB NEAT1 bound by clone D4-10
Authors	:	Deng, X.
Deposited on	:	2015-08-04
Resolution	:	3.17 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:

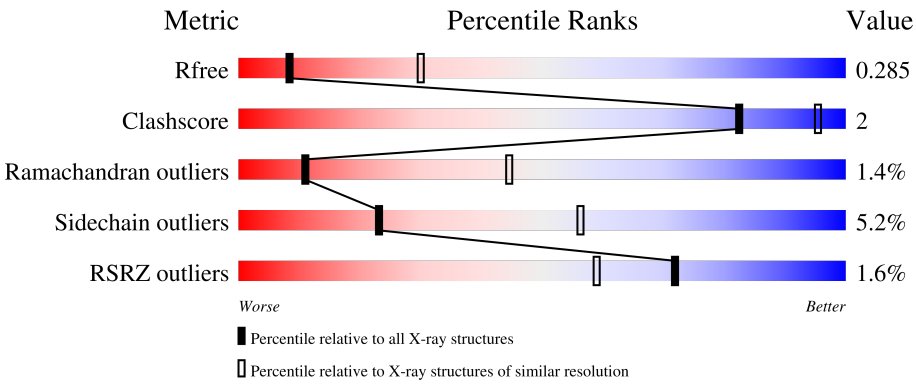
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 3.17 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	1467 (3.20-3.16)
Clashscore	141614	1599 (3.20-3.16)
Ramachandran outliers	138981	1574 (3.20-3.16)
Sidechain outliers	138945	1573 (3.20-3.16)
RSRZ outliers	127900	1423 (3.20-3.16)

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	214	92% 7%
1	H	214	% 93% 7%
2	B	264	% 75% 7% • 17%
2	G	264	3% 73% 10% 16%
3	C	214	% 81% 14% • •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	E	214	2%, 88%, 11%
4	D	270	3%, 72%, 13%, 15%
4	F	270	%, 67%, 13%, 17%
5	I	157	2%, 91%, 5%
5	J	157	%, 86%, 8%, 5%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 15704 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 Y10 Light Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	213	1644	1033	275	331	5	0	0	0
1	H	214	1650	1036	276	332	6	0	0	0

- Molecule 2 is a protein called Y10 Heavy Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	218	1645	1041	279	319	6	0	0	0
2	G	222	1668	1053	283	325	7	0	0	0

- Molecule 3 is a protein called D4-10 Light Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	208	1601	1004	268	325	4	0	0	0
3	E	213	1635	1025	274	332	4	0	0	0

- Molecule 4 is a protein called D4-10 Heavy Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	230	1729	1090	289	343	7	0	0	0
4	F	223	1680	1062	280	331	7	0	0	0

- Molecule 5 is a protein called Iron-regulated surface determinant protein B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	I	149	1226	787	204	233	2	0	0	0
			Total	C	N	O	S			
5	J	149	1226	787	204	233	2	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

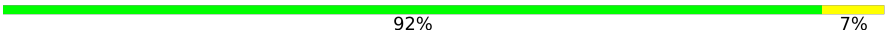
Chain	Residue	Modelled	Actual	Comment	Reference
I	113	GLY	-	expression tag	UNP Q6GA86
I	114	PRO	-	expression tag	UNP Q6GA86
J	113	GLY	-	expression tag	UNP Q6GA86
J	114	PRO	-	expression tag	UNP Q6GA86

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.

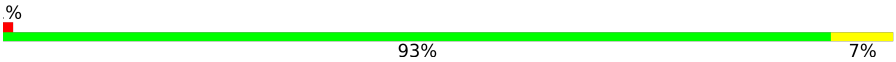
• Molecule 1: Y10 Light Chain

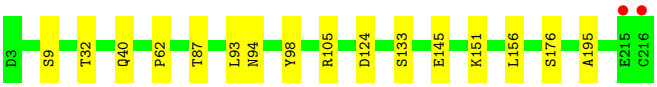
Chain A:



• Molecule 1: Y10 Light Chain

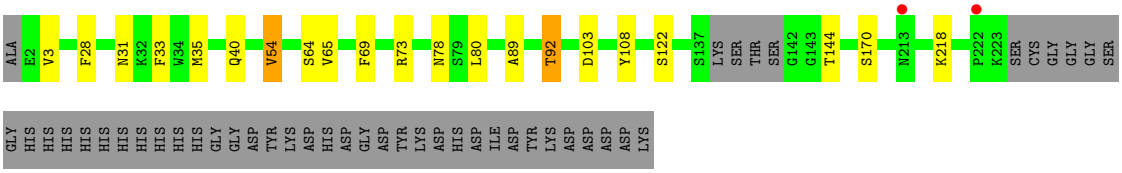
Chain H:





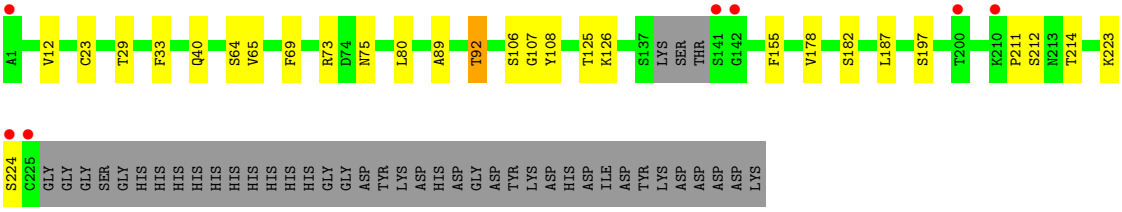
• Molecule 2: Y10 Heavy Chain

Chain B:



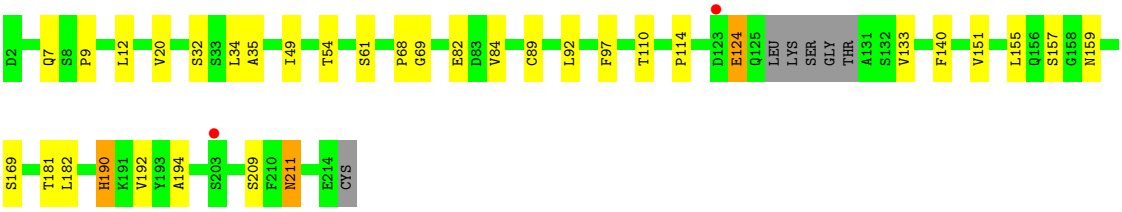
• Molecule 2: Y10 Heavy Chain

Chain G:

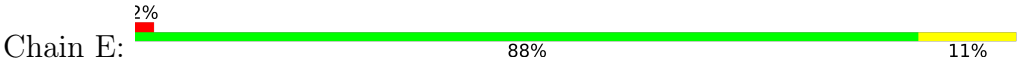


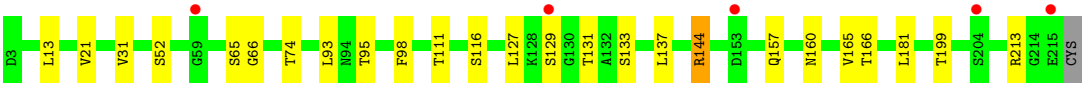
• Molecule 3: D4-10 Light Chain

Chain C:

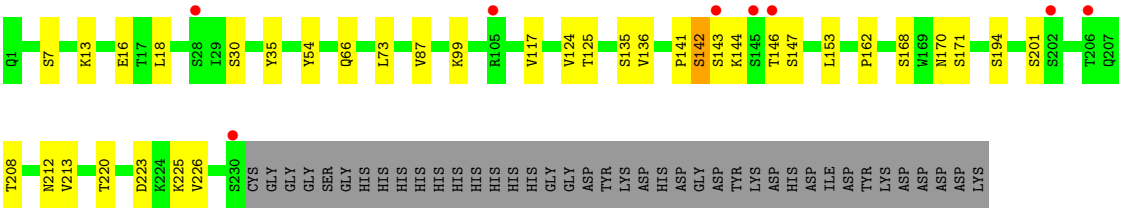


• Molecule 3: D4-10 Light Chain

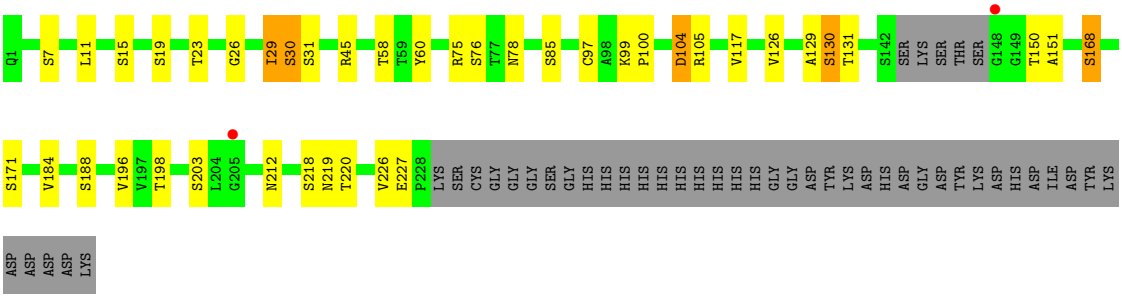




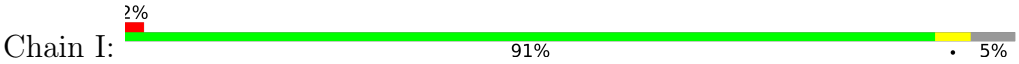
• Molecule 4: D4-10 Heavy Chain

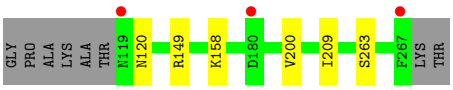


• Molecule 4: D4-10 Heavy Chain

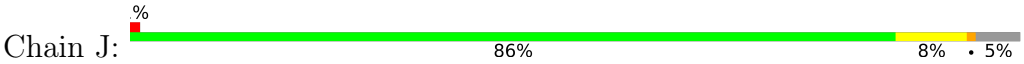


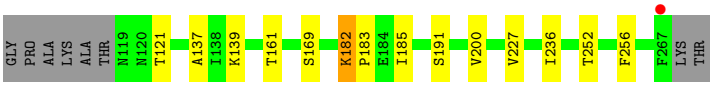
• Molecule 5: Iron-regulated surface determinant protein B





• Molecule 5: Iron-regulated surface determinant protein B




GLY
PRO
ALA
LYS
ALA
THR
N119
N120
T121
A137
I138
K139
T161
S169
K182
P183
E184
I185
S191
V200
V227
I236
T252
F256
F267
LYS
THR

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 2 21 21	Depositor
Cell constants a, b, c, α, β, γ	115.29Å 147.00Å 165.04Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	24.96 – 3.17 24.96 – 3.17	Depositor EDS
% Data completeness (in resolution range)	98.0 (24.96-3.17) 98.2 (24.96-3.17)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.73 (at 3.17Å)	Xtriage
Refinement program	REFMAC 5.8.0123	Depositor
R, R_{free}	0.233 , 0.287 0.233 , 0.285	Depositor DCC
R_{free} test set	2344 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å^2)	49.7	Xtriage
Anisotropy	0.101	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 6.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	15704	wwPDB-VP
Average B, all atoms (Å^2)	49.0	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 20.53 % 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 8.5101e-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 (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	0.77	0/1682	0.84	1/2288 (0.0%)
1	H	0.80	0/1688	0.82	1/2296 (0.0%)
2	B	0.82	0/1686	0.88	0/2294
2	G	0.79	0/1709	0.87	0/2325
3	C	0.79	0/1635	0.84	0/2223
3	E	0.79	0/1670	0.86	2/2271 (0.1%)
4	D	0.82	0/1772	0.91	1/2415 (0.0%)
4	F	0.80	0/1722	0.92	1/2348 (0.0%)
5	I	0.80	0/1257	0.85	1/1695 (0.1%)
5	J	0.78	0/1257	0.81	0/1695
All	All	0.80	0/16078	0.86	7/21850 (0.0%)

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	D	73	LEU	CA-CB-CG	6.67	130.65	115.30
3	E	127	LEU	CA-CB-CG	6.50	130.24	115.30
3	E	144	ARG	NE-CZ-NH1	5.96	123.28	120.30
5	I	149	ARG	NE-CZ-NH1	5.41	123.00	120.30
4	F	45	ARG	NE-CZ-NH1	5.21	122.90	120.30

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	1644	0	1592	5	0
1	H	1650	0	1596	10	0
2	B	1645	0	1599	9	0
2	G	1668	0	1621	13	0
3	C	1601	0	1553	14	0
3	E	1635	0	1593	7	0
4	D	1729	0	1700	12	0
4	F	1680	0	1646	11	0
5	I	1226	0	1204	1	0
5	J	1226	0	1204	8	0
All	All	15704	0	15308	76	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 2.

The worst 5 of 76 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:H:87:THR:HG22	1:H:105:ARG:HG3	1.82	0.61
4:F:58:THR:HG21	5:J:169:SER:HB2	1.83	0.60
4:D:170:ASN:O	4:D:212:ASN:ND2	2.36	0.59
4:F:99:LYS:HE3	4:F:117:VAL:HG22	1.85	0.59
1:A:165:VAL:HG22	1:A:177:LEU:HD12	1.84	0.59

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	211/214 (99%)	199 (94%)	12 (6%)	0	100 100
1	H	212/214 (99%)	199 (94%)	12 (6%)	1 (0%)	29 66

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	214/264 (81%)	202 (94%)	10 (5%)	2 (1%)	17	54
2	G	218/264 (83%)	205 (94%)	10 (5%)	3 (1%)	11	44
3	C	204/214 (95%)	186 (91%)	15 (7%)	3 (2%)	10	43
3	E	211/214 (99%)	196 (93%)	13 (6%)	2 (1%)	17	54
4	D	228/270 (84%)	198 (87%)	24 (10%)	6 (3%)	5	29
4	F	219/270 (81%)	190 (87%)	18 (8%)	11 (5%)	2	14
5	I	147/157 (94%)	134 (91%)	13 (9%)	0	100	100
5	J	147/157 (94%)	139 (95%)	8 (5%)	0	100	100
All	All	2011/2238 (90%)	1848 (92%)	135 (7%)	28 (1%)	11	44

5 of 28 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	144	THR
4	F	30	SER
4	F	130	SER
2	B	54	VAL
4	D	142	SER

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	186/187 (100%)	177 (95%)	9 (5%)	25	60
1	H	187/187 (100%)	182 (97%)	5 (3%)	44	74
2	B	182/220 (83%)	176 (97%)	6 (3%)	38	70
2	G	185/220 (84%)	175 (95%)	10 (5%)	22	55
3	C	183/188 (97%)	172 (94%)	11 (6%)	19	51
3	E	187/188 (100%)	176 (94%)	11 (6%)	19	52
4	D	199/232 (86%)	188 (94%)	11 (6%)	21	55
4	F	192/232 (83%)	171 (89%)	21 (11%)	6	24

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
5	I	135/140 (96%)	132 (98%)	3 (2%)	52	78
5	J	135/140 (96%)	130 (96%)	5 (4%)	34	67
All	All	1771/1934 (92%)	1679 (95%)	92 (5%)	23	56

5 of 92 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	F	78	ASN
4	F	220	THR
4	F	97	CYS
4	F	171	SER
2	G	126	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	190	HIS
5	I	262	ASN
1	H	162	GLN
2	G	53	GLN
1	H	40	GLN

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	213/214 (99%)	0.00	1 (0%)	91	86	34, 53, 66, 72	0
1	H	214/214 (100%)	-0.01	2 (0%)	84	75	35, 51, 67, 75	0
2	B	218/264 (82%)	-0.12	2 (0%)	84	75	30, 43, 62, 89	0
2	G	222/264 (84%)	-0.05	7 (3%)	47	31	32, 42, 60, 84	2 (0%)
3	C	208/214 (97%)	-0.10	2 (0%)	82	72	26, 40, 93, 108	0
3	E	213/214 (99%)	0.07	5 (2%)	60	46	28, 49, 92, 102	0
4	D	230/270 (85%)	0.16	8 (3%)	44	28	27, 50, 87, 96	6 (2%)
4	F	223/270 (82%)	0.10	2 (0%)	84	75	30, 55, 83, 95	0
5	I	149/157 (94%)	-0.23	3 (2%)	65	50	26, 38, 61, 78	0
5	J	149/157 (94%)	-0.16	1 (0%)	87	81	29, 40, 66, 80	0
All	All	2039/2238 (91%)	-0.02	33 (1%)	72	59	26, 47, 82, 108	8 (0%)

The worst 5 of 33 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	216	CYS	5.1
2	G	1	ALA	4.8
2	G	225	CYS	3.9
4	D	143	SER	3.6
2	G	142	GLY	3.3

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.