Full wwPDB X-ray Structure Validation Report ⓘ

Nov 20, 2023 – 06:55 PM JST

PDB ID	:	7D7S
Title	:	HIV-1 SF2 Nef in complex with the Fyn SH3 R96I mutant
Authors	:	Aldehaiman, A.; Shahul Hamed, U.F.; Arold, S.T.
Deposited on	:	2020-10-05
Resolution	:	3.32 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

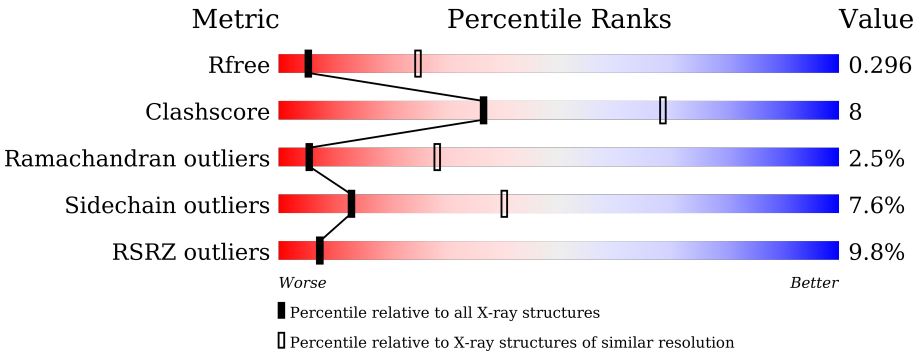
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 3.32 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	1089 (3.36-3.28)
Clashscore	141614	1137 (3.36-3.28)
Ramachandran outliers	138981	1115 (3.36-3.28)
Sidechain outliers	138945	1114 (3.36-3.28)
RSRZ outliers	127900	1059 (3.36-3.28)

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	223	2% 36% 10% • 53%
1	B	223	5% 33% 12% • 54%
2	C	72	10% 61% 19% • 18%
2	D	72	12% 68% 15% • 15%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2737 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 Protein Nef.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	105	Total	C	N	O	S	0	0	0
			897	595	152	147	3			
1	B	103	Total	C	N	O	S	0	0	0
			879	585	150	141	3			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	210	SER	-	expression tag	UNP P03407
A	211	LYS	-	expression tag	UNP P03407
A	212	LEU	-	expression tag	UNP P03407
A	213	ALA	-	expression tag	UNP P03407
A	214	ALA	-	expression tag	UNP P03407
A	215	ALA	-	expression tag	UNP P03407
A	216	LEU	-	expression tag	UNP P03407
A	217	GLU	-	expression tag	UNP P03407
A	218	HIS	-	expression tag	UNP P03407
A	219	HIS	-	expression tag	UNP P03407
A	220	HIS	-	expression tag	UNP P03407
A	221	HIS	-	expression tag	UNP P03407
A	222	HIS	-	expression tag	UNP P03407
A	223	HIS	-	expression tag	UNP P03407
B	210	SER	-	expression tag	UNP P03407
B	211	LYS	-	expression tag	UNP P03407
B	212	LEU	-	expression tag	UNP P03407
B	213	ALA	-	expression tag	UNP P03407
B	214	ALA	-	expression tag	UNP P03407
B	215	ALA	-	expression tag	UNP P03407
B	216	LEU	-	expression tag	UNP P03407
B	217	GLU	-	expression tag	UNP P03407
B	218	HIS	-	expression tag	UNP P03407
B	219	HIS	-	expression tag	UNP P03407
B	220	HIS	-	expression tag	UNP P03407

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	221	HIS	-	expression tag	UNP P03407
B	222	HIS	-	expression tag	UNP P03407
B	223	HIS	-	expression tag	UNP P03407

- Molecule 2 is a protein called Tyrosine-protein kinase Fyn.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	59	Total	C	N	O	0	0	0
			474	305	71	98			
2	D	61	Total	C	N	O	0	0	0
			487	314	72	101			

There are 20 discrepancies between the modelled and reference sequences:

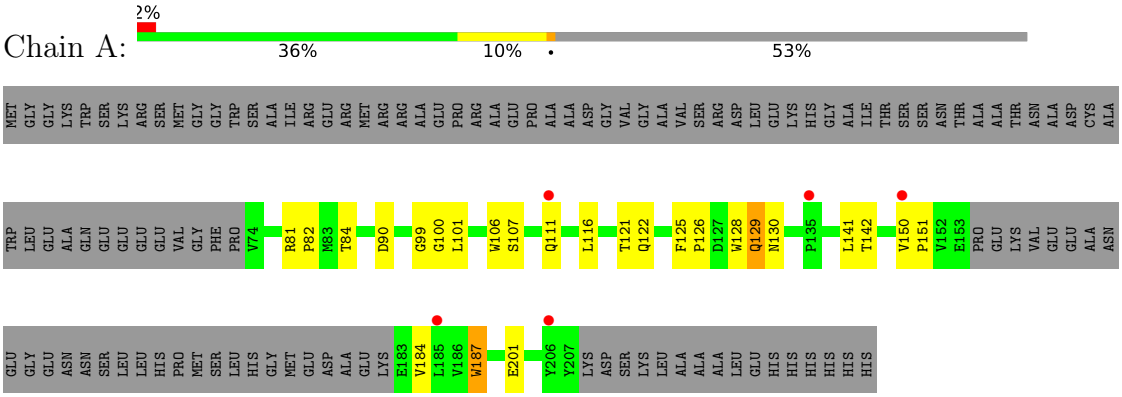
Chain	Residue	Modelled	Actual	Comment	Reference
C	81	MET	-	initiating methionine	UNP E5RFS5
C	96	ILE	ARG	engineered mutation	UNP E5RFS5
C	145	LEU	-	expression tag	UNP E5RFS5
C	146	GLU	-	expression tag	UNP E5RFS5
C	147	HIS	-	expression tag	UNP E5RFS5
C	148	HIS	-	expression tag	UNP E5RFS5
C	149	HIS	-	expression tag	UNP E5RFS5
C	150	HIS	-	expression tag	UNP E5RFS5
C	151	HIS	-	expression tag	UNP E5RFS5
C	152	HIS	-	expression tag	UNP E5RFS5
D	81	MET	-	initiating methionine	UNP E5RFS5
D	96	ILE	ARG	engineered mutation	UNP E5RFS5
D	145	LEU	-	expression tag	UNP E5RFS5
D	146	GLU	-	expression tag	UNP E5RFS5
D	147	HIS	-	expression tag	UNP E5RFS5
D	148	HIS	-	expression tag	UNP E5RFS5
D	149	HIS	-	expression tag	UNP E5RFS5
D	150	HIS	-	expression tag	UNP E5RFS5
D	151	HIS	-	expression tag	UNP E5RFS5
D	152	HIS	-	expression tag	UNP E5RFS5

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.

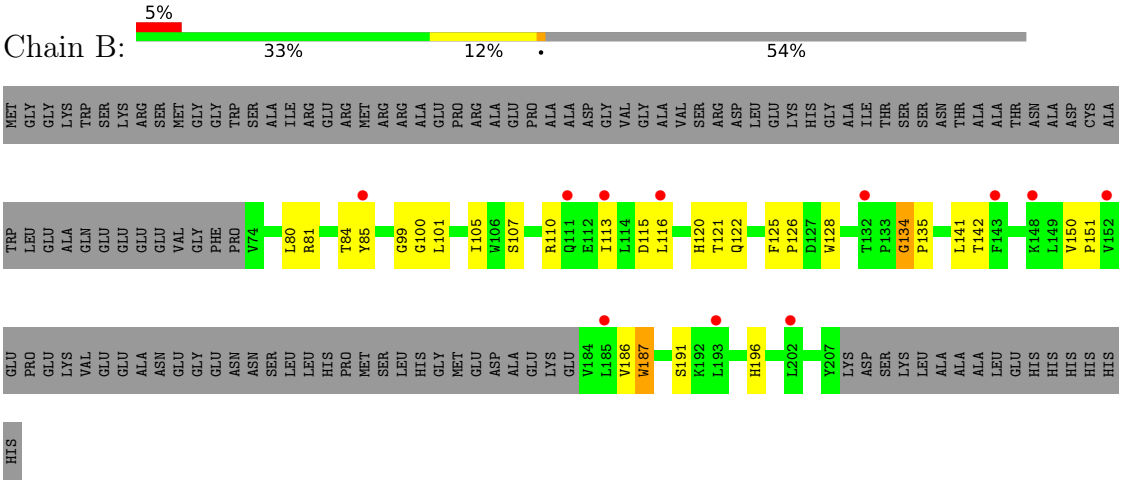
• Molecule 1: Protein Nef

Chain A:



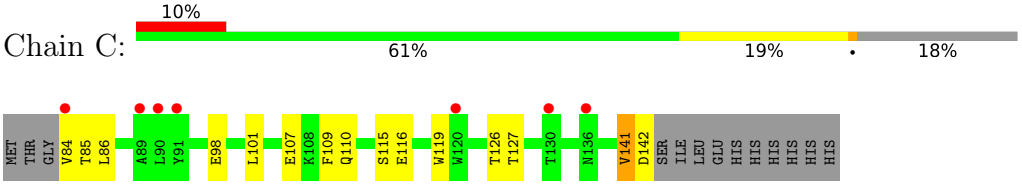
• Molecule 1: Protein Nef

Chain B:



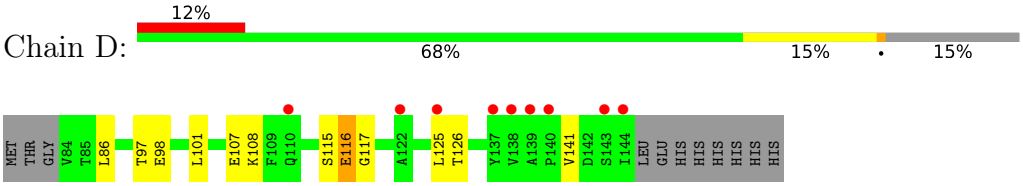
• Molecule 2: Tyrosine-protein kinase Fyn

Chain C:



• Molecule 2: Tyrosine-protein kinase Fyn

Chain D:


12%
68%
15%
15%
MET
THR
GLY
V84
T85
L86
T97
E98
L101
E107
K108
F109
Q110
S115
E116
G117
A122
L125
T126
Y137
V138
A139
P140
V141
D142
S143
I144
LEU
GLU
HIS
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 64 2 2	Depositor
Cell constants a, b, c, α, β, γ	122.11Å 122.11Å 145.89Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.68 – 3.32 48.63 – 3.32	Depositor EDS
% Data completeness (in resolution range)	92.2 (48.68-3.32) 92.3 (48.63-3.32)	Depositor EDS
R_{merge}	0.22	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.65 (at 3.33Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.286 , 0.300 0.277 , 0.296	Depositor DCC
R_{free} test set	923 reflections (10.03%)	wwPDB-VP
Wilson B-factor (Å^2)	142.8	Xtriage
Anisotropy	0.096	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 104.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	2737	wwPDB-VP
Average B, all atoms (Å^2)	150.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.20% 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	0.43	0/932	0.63	0/1268
1	B	0.39	0/914	0.64	1/1244 (0.1%)
2	C	0.51	0/487	0.71	0/665
2	D	0.51	0/500	0.63	0/682
All	All	0.45	0/2833	0.65	1/3859 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	134	GLY	N-CA-C	5.25	126.22	113.10

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	897	0	871	15	0
1	B	879	0	859	16	0
2	C	474	0	434	6	0
2	D	487	0	448	7	0
All	All	2737	0	2612	42	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 8.

All (42) 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:B:121:THR:HG22	1:B:122:GLN:HE21	1.25	0.98
1:A:121:THR:HG22	1:A:122:GLN:HE21	1.23	0.98
1:B:105:ILE:HD13	1:B:186:VAL:HG12	1.76	0.67
1:B:121:THR:HG22	1:B:122:GLN:NE2	2.05	0.66
1:A:121:THR:HG22	1:A:122:GLN:NE2	2.05	0.62
1:A:128:TRP:O	1:A:142:THR:OG1	2.18	0.61
2:D:141:VAL:HG12	2:D:141:VAL:O	2.01	0.60
1:A:106:TRP:NE1	1:A:111:GLN:OE1	2.29	0.59
2:C:141:VAL:HG12	2:C:141:VAL:O	2.02	0.59
1:B:128:TRP:O	1:B:142:THR:OG1	2.18	0.59
2:D:108:LYS:HB3	2:D:125:LEU:HD12	1.85	0.59
2:D:107:GLU:HG3	2:D:126:THR:OG1	2.07	0.54
1:A:130:ASN:HB3	1:A:141:LEU:HD12	1.90	0.54
1:B:101:LEU:HD11	1:B:187:TRP:CD2	2.44	0.53
1:A:101:LEU:HD11	1:A:187:TRP:CD2	2.45	0.52
2:C:85:THR:HG22	2:C:85:THR:O	2.11	0.50
1:A:81:ARG:HG3	1:A:82:PRO:HD2	1.92	0.50
2:C:110:GLN:HA	2:C:110:GLN:OE1	2.14	0.48
2:C:84:VAL:HG23	2:C:86:LEU:HD23	1.95	0.47
1:B:115:ASP:HB3	1:B:126:PRO:HB3	1.96	0.47
1:A:125:PHE:O	1:A:128:TRP:HD1	1.97	0.47
2:D:108:LYS:CB	2:D:125:LEU:HD12	2.43	0.47
1:B:125:PHE:O	1:B:128:TRP:HD1	1.99	0.46
1:A:125:PHE:CG	1:A:126:PRO:HD2	2.51	0.45
2:D:86:LEU:HD11	2:D:125:LEU:HD11	1.99	0.45
1:A:101:LEU:HD11	1:A:187:TRP:CE2	2.53	0.44
1:A:111:GLN:HG3	1:A:129:GLN:HE22	1.83	0.44
1:B:80:LEU:HD11	1:B:125:PHE:HB2	1.99	0.44
1:A:150:VAL:HG12	1:A:151:PRO:O	2.18	0.44
1:B:141:LEU:HD23	1:B:141:LEU:C	2.38	0.43
1:B:81:ARG:NH2	2:C:119:TRP:CH2	2.87	0.43
1:B:150:VAL:HG12	1:B:151:PRO:O	2.18	0.43
1:A:130:ASN:CB	1:A:141:LEU:HD12	2.50	0.42
2:D:116:GLU:O	2:D:117:GLY:C	2.56	0.42
1:A:90:ASP:CB	2:D:97:THR:HG22	2.50	0.42
1:A:99:GLY:O	1:A:101:LEU:N	2.53	0.42
1:B:105:ILE:HD13	1:B:186:VAL:CG1	2.46	0.41
1:B:110:ARG:O	1:B:113:ILE:HB	2.21	0.41
1:B:99:GLY:O	1:B:101:LEU:N	2.54	0.41
2:C:107:GLU:O	2:C:109:PHE:CE1	2.74	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:85:TYR:CD1	1:B:85:TYR:C	2.94	0.40
1:B:101:LEU:HD11	1:B:187:TRP:CE2	2.55	0.40

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	101/223 (45%)	91 (90%)	8 (8%)	2 (2%)	7 33
1	B	99/223 (44%)	92 (93%)	4 (4%)	3 (3%)	4 25
2	C	57/72 (79%)	53 (93%)	2 (4%)	2 (4%)	3 22
2	D	59/72 (82%)	52 (88%)	6 (10%)	1 (2%)	9 36
All	All	316/590 (54%)	288 (91%)	20 (6%)	8 (2%)	5 28

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	100	GLY
1	B	100	GLY
1	B	134	GLY
1	B	135	PRO
1	A	184	VAL
2	C	116	GLU
2	D	116	GLU
2	C	141	VAL

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	94/185 (51%)	88 (94%)	6 (6%)	17 48
1	B	92/185 (50%)	85 (92%)	7 (8%)	13 40
2	C	51/63 (81%)	45 (88%)	6 (12%)	5 21
2	D	53/63 (84%)	50 (94%)	3 (6%)	20 52
All	All	290/496 (58%)	268 (92%)	22 (8%)	13 40

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	84	THR
1	A	107	SER
1	A	116	LEU
1	A	129	GLN
1	A	187	TRP
1	A	201	GLU
1	B	84	THR
1	B	107	SER
1	B	116	LEU
1	B	120	HIS
1	B	187	TRP
1	B	191	SER
1	B	196	HIS
2	C	98	GLU
2	C	101	LEU
2	C	115	SER
2	C	126	THR
2	C	127	THR
2	C	142	ASP
2	D	98	GLU
2	D	101	LEU
2	D	115	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	120	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	122	GLN
1	A	129	GLN
1	B	122	GLN
1	B	129	GLN

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, 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	105/223 (47%)	0.52	5 (4%) 30 30	91, 115, 150, 192	0
1	B	103/223 (46%)	0.68	11 (10%) 6 5	143, 193, 265, 294	0
2	C	59/72 (81%)	0.68	7 (11%) 4 3	98, 122, 151, 167	0
2	D	61/72 (84%)	0.85	9 (14%) 2 1	112, 145, 208, 244	0
All	All	328/590 (55%)	0.66	32 (9%) 7 7	91, 141, 244, 294	0

All (32) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	143	SER	4.0
1	B	113	ILE	3.9
2	C	84	VAL	3.6
2	D	139	ALA	3.6
2	D	125	LEU	3.4
1	B	152	VAL	3.3
2	C	90	LEU	3.3
2	D	110	GLN	3.2
2	C	89	ALA	3.2
1	B	143	PHE	3.0
1	A	185	LEU	2.8
1	B	116	LEU	2.8
1	B	185	LEU	2.7
2	D	122	ALA	2.7
2	D	144	ILE	2.7
2	D	140	PRO	2.6
2	C	91	TYR	2.5
1	A	111	GLN	2.5
2	D	138	VAL	2.5
1	A	150	VAL	2.5
1	A	206	TYR	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	135	PRO	2.4
1	B	148	LYS	2.4
1	B	132	THR	2.3
2	C	130	THR	2.3
2	D	137	TYR	2.3
1	B	111	GLN	2.2
1	B	193	LEU	2.2
1	B	85	TYR	2.2
1	B	202	LEU	2.2
2	C	120	TRP	2.1
2	C	136	ASN	2.1

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.