



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 01:53 am BST

PDB ID : 2ZVN  
Title : NEMO CoZi domain incomplex with diubiquitin in P212121 space group  
Authors : Rahighi, S.; Ikeda, F.; Kawasaki, M.; Akutsu, M.; Suzuki, N.; Kato, R.; Kensch, T.; Uejima, T.; Bloor, S.; Komander, D.; Randow, F.; Wakatsuki, S.; Dikic, I.  
Deposited on : 2008-11-12  
Resolution : 3.00 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](mailto: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	:	2.11
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)	:	2.11

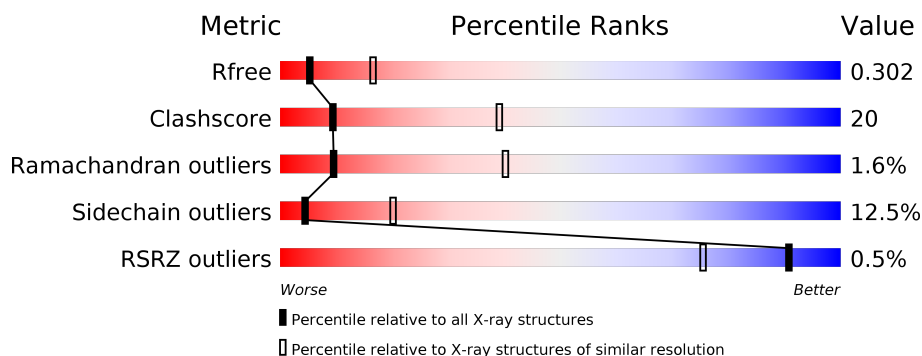
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	154	
1	C	154	
1	E	154	
1	G	154	
2	B	87	
2	D	87	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	87	 66%31%..
2	H	87	 52%34%13%.

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7598 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 UBC protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	148	Total	C	N	O	S	0	0	0
			1175	740	203	230	2			
1	G	146	Total	C	N	O	S	0	0	0
			1156	728	198	228	2			
1	C	149	Total	C	N	O	S	0	0	0
			1181	743	204	232	2			
1	E	146	Total	C	N	O	S	0	0	0
			1156	728	198	228	2			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	EXPRESSION TAG	UNP Q96C32
A	0	SER	-	EXPRESSION TAG	UNP Q96C32
G	-1	GLY	-	EXPRESSION TAG	UNP Q96C32
G	0	SER	-	EXPRESSION TAG	UNP Q96C32
C	-1	GLY	-	EXPRESSION TAG	UNP Q96C32
C	0	SER	-	EXPRESSION TAG	UNP Q96C32
E	-1	GLY	-	EXPRESSION TAG	UNP Q96C32
E	0	SER	-	EXPRESSION TAG	UNP Q96C32

- Molecule 2 is a protein called NF-kappa-B essential modulator.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	85	Total	C	N	O	S	0	0	0
			711	446	124	140	1			
2	D	85	Total	C	N	O	S	0	0	0
			714	449	125	139	1			
2	F	86	Total	C	N	O	S	0	0	0
			720	452	126	141	1			
2	H	86	Total	C	N	O	S	0	0	0
			720	452	126	141	1			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	251	GLY	-	EXPRESSION TAG	UNP O88522
B	252	SER	-	EXPRESSION TAG	UNP O88522
D	251	GLY	-	EXPRESSION TAG	UNP O88522
D	252	SER	-	EXPRESSION TAG	UNP O88522
F	251	GLY	-	EXPRESSION TAG	UNP O88522
F	252	SER	-	EXPRESSION TAG	UNP O88522
H	251	GLY	-	EXPRESSION TAG	UNP O88522
H	252	SER	-	EXPRESSION TAG	UNP O88522

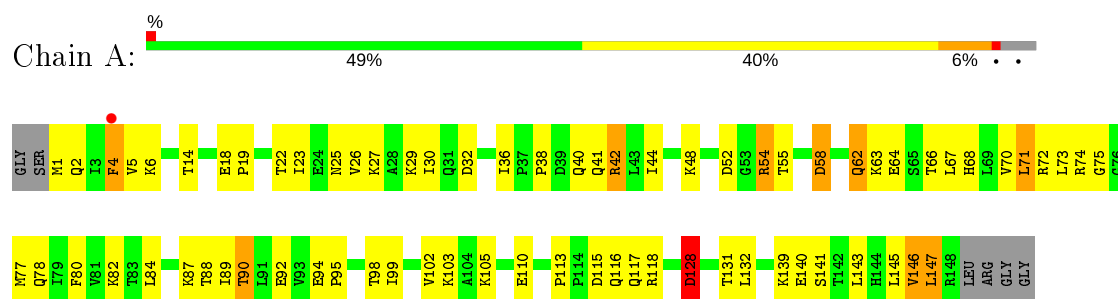
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	9	Total O 9 9	0	0
3	G	18	Total O 18 18	0	0
3	B	3	Total O 3 3	0	0
3	D	1	Total O 1 1	0	0
3	C	7	Total O 7 7	0	0
3	E	25	Total O 25 25	0	0
3	F	2	Total O 2 2	0	0

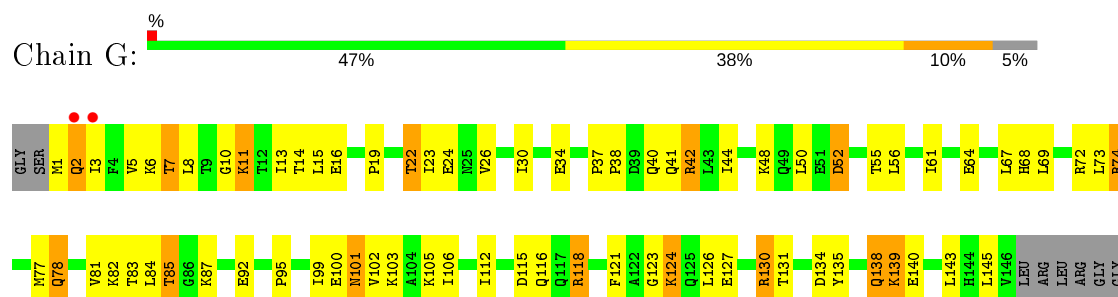
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

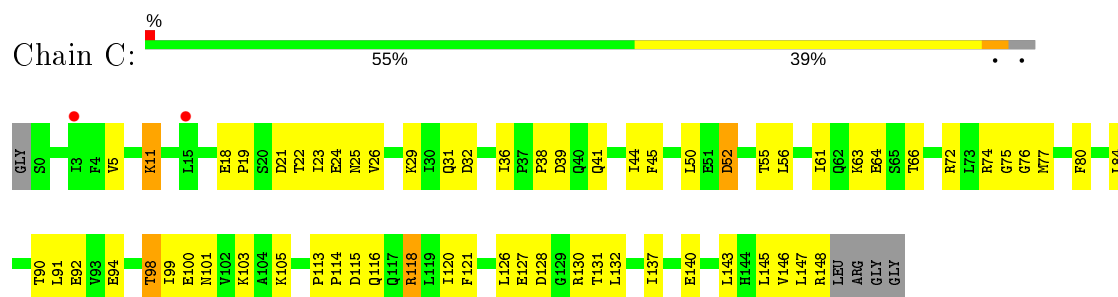
#### • Molecule 1: UBC protein



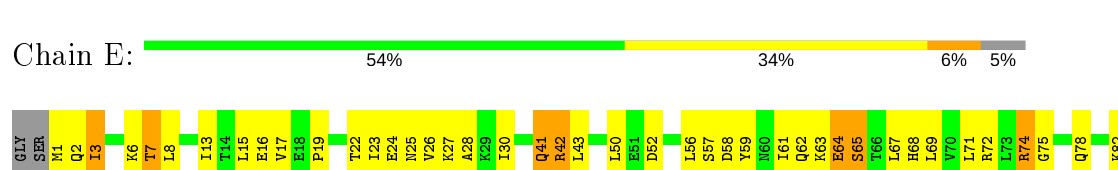
#### • Molecule 1: UBC protein



#### • Molecule 1: UBC protein

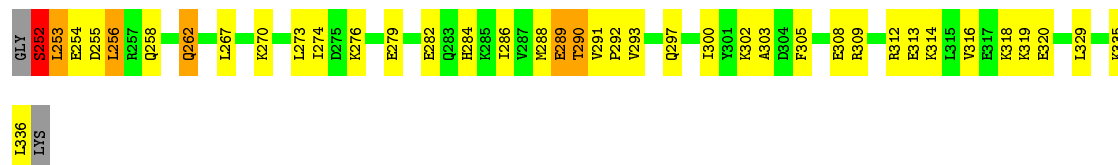


#### • Molecule 1: UBC protein





- Molecule 2: NF-kappa-B essential modulator



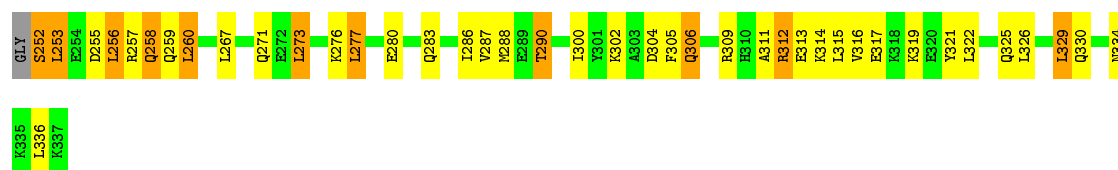
- Molecule 2: NF-kappa-B essential modulator



- Molecule 2: NF-kappa-B essential modulator



- Molecule 2: NF-kappa-B essential modulator



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	59.20 Å 141.29 Å 144.18 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.75 – 3.00 45.75 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.6 (45.75-3.00) 99.6 (45.75-3.00)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.95 (at 3.01 Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.261 , 0.303 0.256 , 0.302	Depositor DCC
$R_{free}$ test set	1267 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	71.2	Xtriage
Anisotropy	0.381	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 50.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.000 for -h,l,k	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	7598	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	77.0	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 46.02 % 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 1.2160e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup> Intensities estimated from amplitudes.

<sup>2</sup> Theoretical 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	0.83	4/1188 (0.3%)	0.76	4/1600 (0.2%)
1	C	0.75	0/1194	0.66	0/1608
1	E	0.79	0/1169	0.76	0/1575
1	G	0.75	1/1169 (0.1%)	0.71	1/1575 (0.1%)
2	B	0.69	0/717	0.70	0/958
2	D	0.75	1/720 (0.1%)	0.71	0/961
2	F	0.79	1/726 (0.1%)	0.73	0/969
2	H	0.75	0/726	0.68	0/969
All	All	0.77	7/7609 (0.1%)	0.72	5/10215 (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
2	B	0	1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	42	ARG	CZ-NH1	8.87	1.44	1.33
2	F	265	GLU	CB-CG	6.58	1.64	1.52
2	D	254	GLU	CD-OE1	6.03	1.32	1.25
1	A	48	LYS	CE-NZ	5.95	1.64	1.49
1	A	48	LYS	CD-CE	5.83	1.65	1.51

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	42	ARG	NE-CZ-NH2	-9.05	115.77	120.30
1	G	118	ARG	NE-CZ-NH2	-6.05	117.27	120.30
1	A	118	ARG	NE-CZ-NH1	5.83	123.22	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	147	LEU	CA-CB-CG	5.49	127.92	115.30
1	A	54	ARG	NE-CZ-NH2	-5.20	117.70	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	252	SER	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	1175	0	1226	47	0
1	C	1181	0	1231	48	0
1	E	1156	0	1202	50	0
1	G	1156	0	1202	50	0
2	B	711	0	721	48	0
2	D	714	0	730	49	0
2	F	720	0	735	40	0
2	H	720	0	734	47	0
3	A	9	0	0	1	0
3	B	3	0	0	0	0
3	C	7	0	0	0	0
3	D	1	0	0	0	0
3	E	25	0	0	3	0
3	F	2	0	0	0	0
3	G	18	0	0	0	0
All	All	7598	0	7781	301	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 20.

The worst 5 of 301 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)
2:B:252:SER:O	2:B:255:ASP:N	1.70	1.23
2:D:253:LEU:HD23	2:D:253:LEU:O	1.42	1.16
2:D:253:LEU:O	2:D:255:ASP:N	1.84	1.10
1:A:1:MET:HB3	1:A:19:PRO:HD3	1.33	1.10
1:G:74:ARG:HH11	1:G:74:ARG:HG3	1.15	1.06

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	146/154 (95%)	121 (83%)	24 (16%)	1 (1%)	22	60
1	C	147/154 (96%)	125 (85%)	19 (13%)	3 (2%)	7	34
1	E	144/154 (94%)	123 (85%)	19 (13%)	2 (1%)	11	43
1	G	144/154 (94%)	126 (88%)	15 (10%)	3 (2%)	7	33
2	B	83/87 (95%)	78 (94%)	3 (4%)	2 (2%)	6	29
2	D	83/87 (95%)	78 (94%)	2 (2%)	3 (4%)	3	19
2	F	84/87 (97%)	79 (94%)	5 (6%)	0	100	100
2	H	84/87 (97%)	78 (93%)	5 (6%)	1 (1%)	13	48
All	All	915/964 (95%)	808 (88%)	92 (10%)	15 (2%)	9	40

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	253	LEU
2	D	254	GLU
1	E	64	GLU
2	H	253	LEU
2	D	329	LEU

### 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	134/137 (98%)	119 (89%)	15 (11%)	6	24
1	C	135/137 (98%)	124 (92%)	11 (8%)	11	40
1	E	132/137 (96%)	116 (88%)	16 (12%)	5	21
1	G	132/137 (96%)	113 (86%)	19 (14%)	3	15
2	B	76/77 (99%)	65 (86%)	11 (14%)	3	15
2	D	76/77 (99%)	63 (83%)	13 (17%)	2	10
2	F	77/77 (100%)	71 (92%)	6 (8%)	12	42
2	H	77/77 (100%)	63 (82%)	14 (18%)	1	9
All	All	839/856 (98%)	734 (88%)	105 (12%)	4	20

5 of 105 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	258	GLN
1	C	5	VAL
2	H	277	LEU
2	D	273	LEU
2	D	288	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	310	HIS
1	C	40	GLN
2	H	283	GLN
2	D	323	GLN
2	D	330	GLN

### 5.3.3 RNA ⓘ

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 carbohydrates 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	148/154 (96%)	-0.02	1 (0%) 87 69	66, 101, 118, 121	0
1	C	149/154 (96%)	-0.01	2 (1%) 77 51	64, 100, 117, 119	0
1	E	146/154 (94%)	-0.23	0 100 100	44, 66, 92, 100	0
1	G	146/154 (94%)	-0.25	2 (1%) 75 49	46, 68, 96, 102	0
2	B	85/87 (97%)	-0.29	0 100 100	20, 63, 89, 98	0
2	D	85/87 (97%)	-0.31	0 100 100	31, 63, 85, 105	0
2	F	86/87 (98%)	-0.26	0 100 100	33, 66, 93, 107	0
2	H	86/87 (98%)	-0.24	0 100 100	36, 64, 92, 112	0
All	All	931/964 (96%)	-0.18	5 (0%) 91 75	20, 73, 115, 121	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	15	LEU	3.4
1	C	3	ILE	2.5
1	G	2	GLN	2.5
1	A	4	PHE	2.3
1	G	3	ILE	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 carbohydrates 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.