



# Full wwPDB X-ray Structure Validation Report ⓘ

May 14, 2020 – 09:55 pm BST

PDB ID : 6D6I  
Title : Ube2V1 in complex with ubiquitin variant Ubv.V1.1 and Ube2N/Ubc13  
Authors : Ceccarelli, D.F.; Garg, P.; Keszei, A.; Sidhu, S.; Sicheri, F.  
Deposited on : 2018-04-21  
Resolution : 2.55 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](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

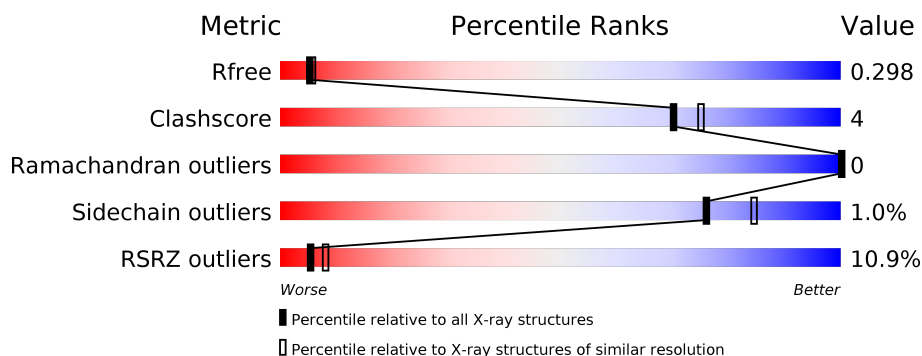
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.55 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	1279 (2.58-2.54)
Clashscore	141614	1327 (2.58-2.54)
Ramachandran outliers	138981	1312 (2.58-2.54)
Sidechain outliers	138945	1312 (2.58-2.54)
RSRZ outliers	127900	1269 (2.58-2.54)

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	143	<div> <div>3%</div> <div>91%</div> <div>5%</div> </div>
1	D	143	<div> <div>2%</div> <div>89%</div> <div>6%</div> <div>5%</div> </div>
2	B	155	<div> <div>19%</div> <div>81%</div> <div>8%</div> <div>10%</div> </div>
2	E	155	<div> <div>19%</div> <div>83%</div> <div>7%</div> <div>10%</div> </div>
3	C	88	<div> <div>8%</div> <div>70%</div> <div>13%</div> <div>17%</div> </div>
3	F	88	<div> <div>5%</div> <div>80%</div> <div>17%</div> </div>

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 9861 atoms, of which 4743 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 Ubiquitin-conjugating enzyme E2 variant 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	136	Total	C	H	N	O	S	0	0	0
			2040	657	1000	181	194	8			
1	D	136	Total	C	H	N	O	S	0	0	0
			2061	663	1014	182	195	7			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	5	GLY	-	expression tag	UNP Q13404
A	6	SER	-	expression tag	UNP Q13404
A	7	THR	-	expression tag	UNP Q13404
D	5	GLY	-	expression tag	UNP Q13404
D	6	SER	-	expression tag	UNP Q13404
D	7	THR	-	expression tag	UNP Q13404

- Molecule 2 is a protein called Ubiquitin-conjugating enzyme E2 N.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	139	Total	C	H	N	O	S	0	0	0
			1854	629	882	166	174	3			
2	E	140	Total	C	H	N	O	S	0	0	0
			1915	647	922	171	172	3			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	GLY	-	expression tag	UNP P61088
B	-1	ALA	-	expression tag	UNP P61088
B	0	MET	-	expression tag	UNP P61088
B	1	GLY	-	expression tag	UNP P61088
B	2	SER	-	expression tag	UNP P61088
E	-2	GLY	-	expression tag	UNP P61088

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	-1	ALA	-	expression tag	UNP P61088
E	0	MET	-	expression tag	UNP P61088
E	1	GLY	-	expression tag	UNP P61088
E	2	SER	-	expression tag	UNP P61088

- Molecule 3 is a protein called Ubv.V1.1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	C	73	Total	C	H	N	O	S	0	0	0
			1012	342	475	91	103	1			
3	F	73	Total	C	H	N	O	S	0	0	0
			979	337	450	92	99	1			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-11	GLY	-	expression tag	UNP Q96H31
C	-10	ALA	-	expression tag	UNP Q96H31
C	-9	GLY	-	expression tag	UNP Q96H31
C	-8	GLY	-	expression tag	UNP Q96H31
C	-7	ASP	-	expression tag	UNP Q96H31
C	-6	TYR	-	expression tag	UNP Q96H31
C	-5	LYS	-	expression tag	UNP Q96H31
C	-4	ASP	-	expression tag	UNP Q96H31
C	-3	ASP	-	expression tag	UNP Q96H31
C	-2	ASP	-	expression tag	UNP Q96H31
C	-1	ASP	-	expression tag	UNP Q96H31
C	0	LYS	-	expression tag	UNP Q96H31
C	62	HIS	GLN	engineered mutation	UNP Q96H31
C	63	TRP	LYS	engineered mutation	UNP Q96H31
C	68	LEU	HIS	engineered mutation	UNP Q96H31
C	70	TRP	VAL	engineered mutation	UNP Q96H31
C	71	TRP	LEU	engineered mutation	UNP Q96H31
C	74	LEU	ARG	engineered mutation	UNP Q96H31
C	75	ILE	GLY	engineered mutation	UNP Q96H31
C	76	ALA	GLY	engineered mutation	UNP Q96H31
F	-11	GLY	-	expression tag	UNP Q96H31
F	-10	ALA	-	expression tag	UNP Q96H31
F	-9	GLY	-	expression tag	UNP Q96H31
F	-8	GLY	-	expression tag	UNP Q96H31
F	-7	ASP	-	expression tag	UNP Q96H31
F	-6	TYR	-	expression tag	UNP Q96H31

*Continued on next page...*

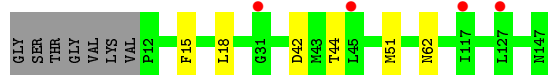
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
F	-5	LYS	-	expression tag	UNP Q96H31
F	-4	ASP	-	expression tag	UNP Q96H31
F	-3	ASP	-	expression tag	UNP Q96H31
F	-2	ASP	-	expression tag	UNP Q96H31
F	-1	ASP	-	expression tag	UNP Q96H31
F	0	LYS	-	expression tag	UNP Q96H31
F	62	HIS	GLN	engineered mutation	UNP Q96H31
F	63	TRP	LYS	engineered mutation	UNP Q96H31
F	68	LEU	HIS	engineered mutation	UNP Q96H31
F	70	TRP	VAL	engineered mutation	UNP Q96H31
F	71	TRP	LEU	engineered mutation	UNP Q96H31
F	74	LEU	ARG	engineered mutation	UNP Q96H31
F	75	ILE	GLY	engineered mutation	UNP Q96H31
F	76	ALA	GLY	engineered mutation	UNP Q96H31

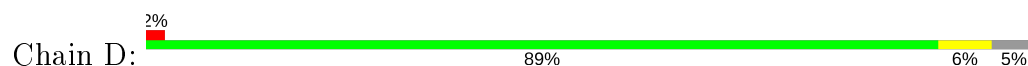
### 3 Residue-property plots [i](#)

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.

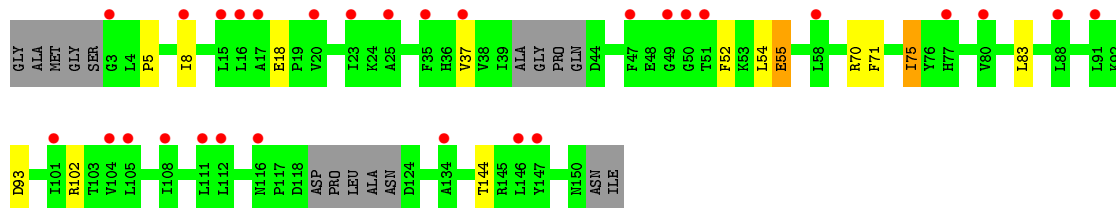
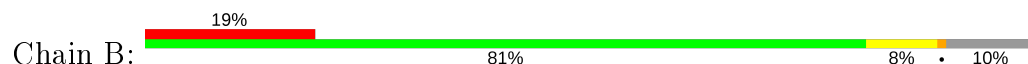
- Molecule 1: Ubiquitin-conjugating enzyme E2 variant 1



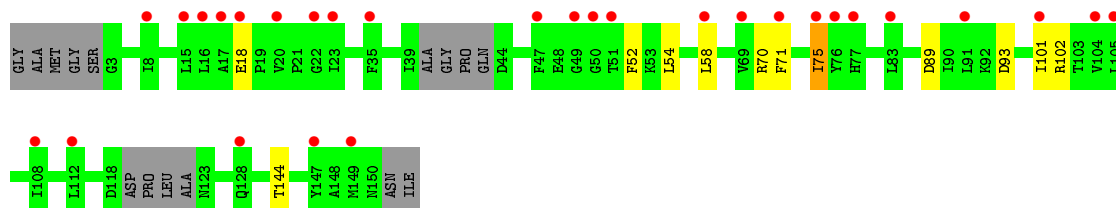
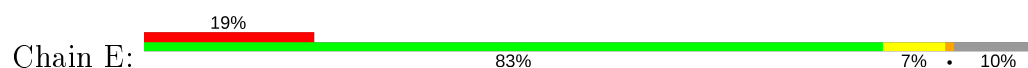
- Molecule 1: Ubiquitin-conjugating enzyme E2 variant 1



- Molecule 2: Ubiquitin-conjugating enzyme E2 N



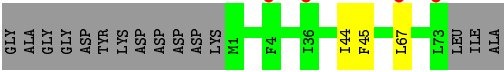
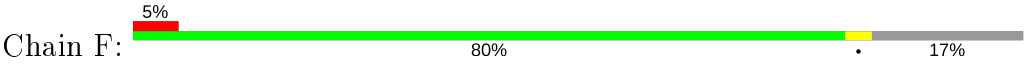
- Molecule 2: Ubiquitin-conjugating enzyme E2 N



- Molecule 3: Ubv.V1.1



● Molecule 3: Ubv.V1.1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	91.44Å 91.44Å 92.45Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	45.72 – 2.55 45.72 – 2.55	Depositor EDS
% Data completeness (in resolution range)	99.7 (45.72-2.55) 99.8 (45.72-2.55)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.43 (at 2.54Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.264 , 0.296 0.265 , 0.298	Depositor DCC
$R_{free}$ test set	1423 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	71.6	Xtriage
Anisotropy	0.170	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 74.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.029 for -h,-k,l 0.488 for h,-h-k,-l 0.034 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	9861	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	114.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.01% of the height of the origin peak. No significant pseudotranslation is detected.*

<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

### 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	0.24	0/1064	0.41	0/1448
1	D	0.24	0/1071	0.42	0/1457
2	B	0.24	0/994	0.41	0/1369
2	E	0.24	0/1016	0.41	0/1396
3	C	0.25	0/547	0.43	0/749
3	F	0.24	0/541	0.42	0/742
All	All	0.24	0/5233	0.42	0/7161

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	1040	1000	1000	5	0
1	D	1047	1014	1014	5	0
2	B	972	882	881	13	0
2	E	993	922	930	9	0
3	C	537	475	490	7	0
3	F	529	450	463	2	0
All	All	5118	4743	4778	36	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 4.

All (36) 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:E:18:GLU:O	2:E:102:ARG:NH2	2.03	0.91
2:B:18:GLU:O	2:B:102:ARG:NH2	2.20	0.75
1:D:42:ASP:OD1	1:D:44:THR:HG22	1.92	0.69
1:A:42:ASP:OD1	1:A:44:THR:HG22	1.94	0.68
2:E:71:PHE:CD1	2:E:75:ILE:HD11	2.32	0.65
2:B:70:ARG:NH1	2:B:83:LEU:O	2.32	0.62
2:B:71:PHE:CD1	2:B:75:ILE:HD11	2.34	0.62
3:C:23:ILE:HD13	3:C:50:LEU:HB3	1.81	0.61
2:B:55:GLU:OE2	2:B:70:ARG:NE	2.33	0.60
3:C:22:THR:HG23	3:C:25:ASN:H	1.70	0.56
2:B:54:LEU:HD21	2:B:71:PHE:CZ	2.41	0.55
1:D:89:ASN:OD1	1:D:136:LEU:HD21	2.06	0.54
1:A:18:LEU:HD11	2:B:70:ARG:HD2	1.89	0.53
1:A:18:LEU:HD11	2:B:70:ARG:CD	2.39	0.53
2:B:71:PHE:CG	2:B:75:ILE:HD11	2.44	0.53
3:C:8:LEU:HD23	3:C:8:LEU:O	2.09	0.52
2:E:54:LEU:HD21	2:E:71:PHE:CZ	2.43	0.52
3:C:36:ILE:O	3:C:41:GLN:NE2	2.45	0.50
2:B:52:PHE:CE2	2:B:75:ILE:HG23	2.47	0.49
3:F:45:PHE:HB2	3:F:67:LEU:HD13	1.92	0.49
2:E:89:ASP:OD1	2:E:89:ASP:N	2.47	0.47
1:D:18:LEU:HD11	2:E:70:ARG:HD2	1.96	0.47
3:C:23:ILE:HD11	3:C:59:TYR:CE2	2.50	0.47
2:B:37:VAL:O	2:B:54:LEU:N	2.46	0.47
2:E:58:LEU:HD11	2:E:101:ILE:HD11	1.97	0.46
1:A:15:PHE:HA	1:A:18:LEU:HD12	1.97	0.46
2:E:54:LEU:HD21	2:E:71:PHE:CE1	2.51	0.46
2:E:93:ASP:N	2:E:93:ASP:OD1	2.49	0.46
2:B:93:ASP:N	2:B:93:ASP:OD1	2.49	0.45
1:D:51:MET:HE1	3:F:44:ILE:HG23	1.98	0.44
3:C:45:PHE:HB2	3:C:67:LEU:HD13	1.98	0.44
2:B:5:PRO:HG2	2:B:8:ILE:HD12	2.00	0.43
2:B:52:PHE:HZ	2:B:144:THR:HG22	1.83	0.43
1:A:51:MET:HE1	3:C:44:ILE:HG23	2.02	0.41
2:E:52:PHE:HZ	2:E:144:THR:HG22	1.85	0.41
1:D:107:VAL:HG21	1:D:119:VAL:HG12	2.02	0.41

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	134/143 (94%)	132 (98%)	2 (2%)	0	100	100
1	D	134/143 (94%)	132 (98%)	2 (2%)	0	100	100
2	B	133/155 (86%)	130 (98%)	3 (2%)	0	100	100
2	E	134/155 (86%)	129 (96%)	5 (4%)	0	100	100
3	C	71/88 (81%)	69 (97%)	2 (3%)	0	100	100
3	F	71/88 (81%)	69 (97%)	2 (3%)	0	100	100
All	All	677/772 (88%)	661 (98%)	16 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	111/127 (87%)	110 (99%)	1 (1%)	78	87
1	D	112/127 (88%)	111 (99%)	1 (1%)	78	87
2	B	84/130 (65%)	82 (98%)	2 (2%)	49	63
2	E	88/130 (68%)	87 (99%)	1 (1%)	73	83
3	C	51/77 (66%)	51 (100%)	0	100	100
3	F	47/77 (61%)	47 (100%)	0	100	100
All	All	493/668 (74%)	488 (99%)	5 (1%)	76	84

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	62	ASN
2	B	55	GLU
2	B	75	ILE
1	D	62	ASN
2	E	75	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 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 ⓘ

### 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<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	136/143 (95%)	0.20	4 (2%) 51 61	45, 74, 118, 152	0
1	D	136/143 (95%)	0.22	3 (2%) 62 70	45, 74, 118, 147	0
2	B	139/155 (89%)	1.17	29 (20%) 1 1	81, 136, 198, 229	0
2	E	140/155 (90%)	1.14	29 (20%) 1 1	86, 133, 176, 264	0
3	C	73/88 (82%)	0.59	7 (9%) 8 11	56, 107, 145, 192	0
3	F	73/88 (82%)	0.57	4 (5%) 25 32	57, 110, 150, 166	0
All	All	697/772 (90%)	0.66	76 (10%) 5 8	45, 108, 166, 264	0

All (76) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	E	50	GLY	12.3
2	B	50	GLY	9.2
2	B	49	GLY	7.4
2	B	88	LEU	7.1
2	E	49	GLY	7.1
2	B	17	ALA	6.1
2	B	134	ALA	5.3
2	B	23	ILE	5.2
2	B	91	LEU	5.1
2	B	147	TYR	4.9
2	B	146	LEU	4.9
2	B	77	HIS	4.8
2	B	101	ILE	4.7
2	E	17	ALA	4.6
2	E	105	LEU	4.3
2	B	3	GLY	4.2
2	E	8	ILE	4.2
2	B	108	ILE	4.1
2	B	105	LEU	4.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	E	76	TYR	4.0
2	E	108	ILE	3.9
2	E	20	VAL	3.9
2	E	47	PHE	3.9
2	E	35	PHE	3.8
2	E	15	LEU	3.6
2	E	75	ILE	3.6
2	E	51	THR	3.5
3	C	68	LEU	3.5
3	F	67	LEU	3.4
2	B	51	THR	3.4
2	E	147	TYR	3.3
2	E	58	LEU	3.3
2	B	35	PHE	3.2
2	E	101	ILE	3.2
2	B	20	VAL	3.0
2	B	25	ALA	3.0
3	C	14	THR	3.0
2	B	58	LEU	3.0
2	B	8	ILE	2.9
2	E	104	VAL	2.9
2	E	128	GLN	2.9
1	D	31	GLY	2.8
2	B	80	VAL	2.8
1	A	117	ILE	2.8
2	B	112	LEU	2.7
3	C	8	LEU	2.7
1	A	31	GLY	2.7
2	E	91	LEU	2.7
2	E	71	PHE	2.7
1	A	127	LEU	2.6
3	F	73	LEU	2.6
2	B	116	ASN	2.6
2	E	23	ILE	2.6
2	E	149	MET	2.6
2	B	104	VAL	2.5
2	E	83	LEU	2.5
3	F	4	PHE	2.5
2	E	22	GLY	2.5
2	B	111	LEU	2.4
3	C	34	GLU	2.4
2	B	15	LEU	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	37	VAL	2.3
1	D	117	ILE	2.3
2	E	77	HIS	2.3
2	E	69	VAL	2.3
2	E	112	LEU	2.2
3	C	36	ILE	2.2
3	C	69	LEU	2.2
2	B	47	PHE	2.2
3	F	36	ILE	2.2
2	E	16	LEU	2.2
2	E	18	GLU	2.2
3	C	48	LYS	2.1
1	A	45	LEU	2.1
1	D	91	ASN	2.0
2	B	16	LEU	2.0

## 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.