



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 2, 2017 – 09:53 AM EDT

PDB ID : 2D1L  
Title : Structure of F-actin binding domain IMD of MIM (Missing In Metastasis)  
Authors : Lee, S.H.; Kerff, F.; Chereau, D.; Ferron, F.; Dominguez, R.  
Deposited on : unknown  
Resolution : 1.85 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

<http://wwpdb.org/validation/2016/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) : 1.9-1692  
EDS : rb-20030345  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030345

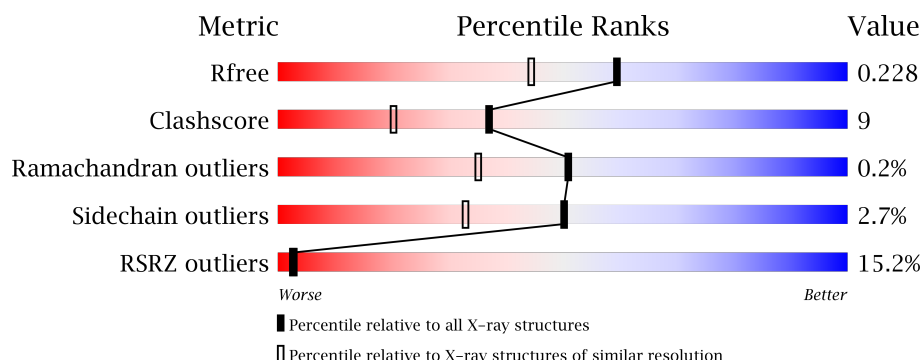
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.85 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))
$R_{free}$	100719	1923 (1.86-1.86)
Clashscore	112137	2083 (1.86-1.86)
Ramachandran outliers	110173	2060 (1.86-1.86)
Sidechain outliers	110143	2060 (1.86-1.86)
RSRZ outliers	101464	1932 (1.86-1.86)

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. 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	253	<div> <div>11%</div> <div> <div></div> <div>86%</div> <div>11%</div> <div>••</div> </div> </div>
1	B	253	<div> <div>17%</div> <div> <div></div> <div>73%</div> <div>16%</div> <div>• 9%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4390 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 Metastasis suppressor protein 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	249	Total	C	N	O	S	Se	0	19	0
			2058	1290	358	394	4	12			
1	B	231	Total	C	N	O	S	Se	0	20	0
			1940	1224	339	362	4	11			

There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	ALA	-	CLONING ARTIFACT	UNP Q8R1S4
A	-1	GLY	-	CLONING ARTIFACT	UNP Q8R1S4
A	0	HIS	-	CLONING ARTIFACT	UNP Q8R1S4
A	1	MSE	MET	MODIFIED RESIDUE	UNP Q8R1S4
A	23	MSE	MET	MODIFIED RESIDUE	UNP Q8R1S4
A	64	MSE	MET	MODIFIED RESIDUE	UNP Q8R1S4
A	82	MSE	MET	MODIFIED RESIDUE	UNP Q8R1S4
A	84	MSE	MET	MODIFIED RESIDUE	UNP Q8R1S4
A	112	MSE	MET	MODIFIED RESIDUE	UNP Q8R1S4
A	139	LYS	ASN	MODIFIED RESIDUE	UNP Q8R1S4
A	205	MSE	MET	MODIFIED RESIDUE	UNP Q8R1S4
A	216	MSE	MET	MODIFIED RESIDUE	UNP Q8R1S4
A	235	MSE	MET	MODIFIED RESIDUE	UNP Q8R1S4
B	-2	ALA	-	CLONING ARTIFACT	UNP Q8R1S4
B	-1	GLY	-	CLONING ARTIFACT	UNP Q8R1S4
B	0	HIS	-	CLONING ARTIFACT	UNP Q8R1S4
B	1	MSE	MET	MODIFIED RESIDUE	UNP Q8R1S4
B	23	MSE	MET	MODIFIED RESIDUE	UNP Q8R1S4
B	64	MSE	MET	MODIFIED RESIDUE	UNP Q8R1S4
B	82	MSE	MET	MODIFIED RESIDUE	UNP Q8R1S4
B	84	MSE	MET	MODIFIED RESIDUE	UNP Q8R1S4
B	112	MSE	MET	MODIFIED RESIDUE	UNP Q8R1S4
B	139	LYS	ASN	MODIFIED RESIDUE	UNP Q8R1S4
B	205	MSE	MET	MODIFIED RESIDUE	UNP Q8R1S4
B	216	MSE	MET	MODIFIED RESIDUE	UNP Q8R1S4

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	235	MSE	MET	MODIFIED RESIDUE	UNP Q8R1S4

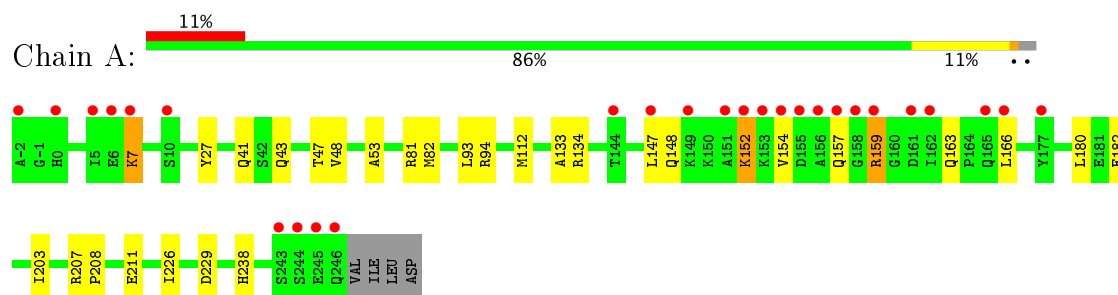
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	208	Total 208	O 208	0	0
2	B	184	Total 184	O 184	0	0

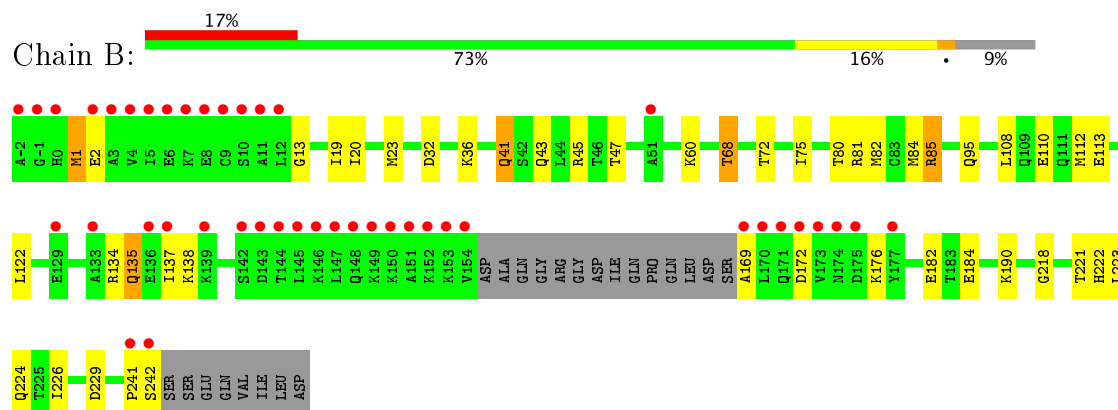
### 3 Residue-property plots [i](#)

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: Metastasis suppressor protein 1



#### • Molecule 1: Metastasis suppressor protein 1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	53.50 Å 37.32 Å 129.02 Å 90.00° 94.07° 90.00°	Depositor
Resolution (Å)	48.10 – 1.85 48.10 – 1.80	Depositor EDS
% Data completeness (in resolution range)	98.7 (48.10-1.85) 96.1 (48.10-1.80)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.65 (at 1.79 Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.183 , 0.228 0.184 , 0.228	Depositor DCC
$R_{free}$ test set	2197 reflections (5.32%)	DCC
Wilson B-factor (Å <sup>2</sup> )	24.9	Xtriage
Anisotropy	0.295	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 53.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	4390	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.41% 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 [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	0.52	0/2127	0.55	0/2825
1	B	0.56	1/2012 (0.0%)	0.57	0/2670
All	All	0.54	1/4139 (0.0%)	0.56	0/5495

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	1	MSE	SE-CE	8.71	2.46	1.95

There are no bond angle outliers.

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	2058	0	2157	31	0
1	B	1940	0	2058	47	0
2	A	208	0	0	1	0
2	B	184	0	0	9	0
All	All	4390	0	4215	73	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 9.

All (73) 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:224[B]:GLN:HG3	2:B:340:HOH:O	1.47	1.14
1:B:1:MSE:SE	1:B:1:MSE:CE	2.46	1.13
1:A:48:VAL:HG11	1:A:94:ARG:HG2	1.46	0.98
1:B:85[A]:ARG:HD3	1:B:222:HIS:HB3	1.45	0.95
1:B:85[A]:ARG:HD3	1:B:222:HIS:CB	1.98	0.93
1:B:84[A]:MSE:SE	2:B:420:HOH:O	2.40	0.89
1:A:152:LYS:HE3	1:A:152:LYS:HA	1.57	0.84
1:A:148:GLN:HG3	1:A:166:LEU:HD11	1.62	0.82
1:A:48:VAL:HG22	1:A:93:LEU:HD23	1.62	0.81
1:B:134:ARG:HD2	1:B:138:LYS:HE2	1.62	0.80
1:A:207[B]:ARG:HG3	1:A:208:PRO:HD3	1.65	0.77
1:B:81[A]:ARG:HA	1:B:84[A]:MSE:HE3	1.67	0.77
1:B:113[B]:GLU:HB2	2:B:316:HOH:O	1.86	0.75
1:A:154:VAL:HG13	1:A:159:ARG:HG3	1.72	0.72
1:B:95:GLN:NE2	2:B:432:HOH:O	2.23	0.72
1:A:48:VAL:HG11	1:A:94:ARG:CG	2.20	0.72
1:A:27:TYR:HD1	1:A:112[A]:MSE:HE1	1.57	0.70
1:B:80:THR:HG22	1:B:84[A]:MSE:HE2	1.74	0.69
1:B:81[A]:ARG:NH2	1:B:229[A]:ASP:OD2	2.20	0.68
1:A:27:TYR:HD1	1:A:112[A]:MSE:CE	2.08	0.66
1:B:122:LEU:HD13	1:B:190:LYS:HE2	1.76	0.66
1:A:112[B]:MSE:HE2	1:A:112[B]:MSE:HA	1.77	0.66
1:B:41:GLN:HE22	1:B:45[A]:ARG:HH11	1.45	0.63
1:B:68:THR:OG1	1:B:72:THR:HB	2.01	0.61
1:A:112[B]:MSE:CE	1:A:112[B]:MSE:HA	2.31	0.61
1:A:7:LYS:N	1:A:7:LYS:HD3	2.17	0.60
1:A:207[B]:ARG:HG3	1:A:208:PRO:CD	2.33	0.59
1:B:110[B]:GLU:OE2	2:B:292:HOH:O	2.17	0.59
1:A:133:ALA:HB3	1:A:180:LEU:HD13	1.83	0.59
1:B:19:ILE:HG22	1:B:23:MSE:HE2	1.85	0.58
1:B:80:THR:O	1:B:84[A]:MSE:HG3	2.04	0.58
1:A:43[B]:GLN:HG3	1:B:47:THR:OG1	2.03	0.58
1:B:85[A]:ARG:HD3	1:B:222:HIS:HB2	1.81	0.58
1:A:238:HIS:H	1:A:238:HIS:CD2	2.24	0.55
1:B:113[A]:GLU:HB2	2:B:316:HOH:O	2.07	0.55
1:B:134:ARG:O	1:B:137:ILE:HG12	2.07	0.54
1:B:218:GLY:O	1:B:221:THR:HG22	2.08	0.54
1:A:203:ILE:HD11	1:B:75:ILE:HG23	1.89	0.53
1:B:32:ASP:OD2	2:B:408:HOH:O	2.18	0.53
1:A:152:LYS:CE	1:A:152:LYS:HA	2.34	0.53
1:B:108:LEU:O	1:B:112:MSE:HG2	2.09	0.52
1:B:241:PRO:O	1:B:242:SER:C	2.47	0.52

Continued on next page...



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:43[B]:GLN:OE1	1:B:43[B]:GLN:NE2	2.40	0.52
1:B:13:GLY:HA3	1:B:184:GLU:OE2	2.10	0.50
1:A:207[B]:ARG:NH1	1:A:211:GLU:HB2	2.27	0.50
1:B:85[A]:ARG:HG3	1:B:226:ILE:CD1	2.41	0.50
1:B:113[A]:GLU:HG2	2:B:369:HOH:O	2.11	0.50
1:A:47:THR:HA	1:B:43[A]:GLN:HG2	1.94	0.50
1:A:133:ALA:CB	1:A:180:LEU:HD13	2.41	0.50
1:B:85[A]:ARG:HG3	1:B:226:ILE:HD12	1.95	0.48
1:B:137:ILE:HG22	1:B:176:LYS:CB	2.43	0.48
1:B:82[A]:MSE:SE	1:B:223:LEU:HD22	2.64	0.47
1:A:147:LEU:HB3	1:A:166:LEU:HD13	1.96	0.47
1:A:27:TYR:CD1	1:A:112[A]:MSE:CE	2.94	0.46
1:B:122:LEU:CD1	1:B:190:LYS:HE2	2.44	0.46
1:B:137:ILE:CG2	1:B:176:LYS:HB3	2.46	0.46
1:B:137:ILE:HG22	1:B:176:LYS:HB2	1.99	0.45
1:A:81[B]:ARG:NH1	1:A:229[B]:ASP:OD2	2.49	0.45
1:B:224[B]:GLN:CG	2:B:340:HOH:O	2.30	0.45
1:B:32:ASP:OD1	1:B:36[A]:LYS:NZ	2.47	0.44
1:B:169:ALA:HA	1:B:172:ASP:OD2	2.17	0.44
1:B:20:ILE:HD13	1:B:23:MSE:HE3	1.99	0.44
1:B:135:GLN:O	1:B:135:GLN:HG3	2.19	0.43
1:A:207[B]:ARG:NE	2:A:368:HOH:O	2.52	0.43
1:A:82[B]:MSE:HE2	1:A:226:ILE:HG22	2.00	0.43
1:B:134:ARG:HD2	1:B:138:LYS:CE	2.38	0.42
1:A:159:ARG:HH11	1:A:163:GLN:HG2	1.83	0.42
1:A:154:VAL:HG12	1:A:154:VAL:O	2.20	0.42
1:B:2:GLU:HG3	1:B:138:LYS:HD3	2.01	0.42
1:B:60:LYS:HE2	1:B:60:LYS:HB3	1.76	0.42
1:A:27:TYR:CD1	1:A:112[A]:MSE:HE3	2.55	0.41
1:A:53:ALA:HB1	1:B:36[B]:LYS:HD2	2.03	0.41
1:B:41:GLN:HE22	1:B:45[A]:ARG:NH1	2.14	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	266/253 (105%)	258 (97%)	8 (3%)	0	100	100
1	B	247/253 (98%)	242 (98%)	4 (2%)	1 (0%)	38	22
All	All	513/506 (101%)	500 (98%)	12 (2%)	1 (0%)	51	35

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	68	THR

### 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	234/210 (111%)	227 (97%)	7 (3%)	46	28
1	B	220/210 (105%)	215 (98%)	5 (2%)	56	39
All	All	454/420 (108%)	442 (97%)	12 (3%)	50	34

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	7	LYS
1	A	41	GLN
1	A	134	ARG
1	A	152	LYS
1	A	157	GLN
1	A	159	ARG
1	A	182	GLU
1	B	41	GLN
1	B	85[A]	ARG
1	B	85[B]	ARG
1	B	135	GLN
1	B	182	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	41	GLN
1	A	238	HIS
1	B	41	GLN
1	B	95	GLN
1	B	106	ASN
1	B	121	GLN
1	B	135	GLN
1	B	186	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 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	240/253 (94%)	0.76	27 (11%) 6 5	16, 27, 70, 84	0
1	B	222/253 (87%)	1.04	43 (19%) 1 1	14, 26, 64, 83	0
All	All	462/506 (91%)	0.89	70 (15%) 2 3	14, 26, 69, 84	0

All (70) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	158	GLY	16.2
1	A	156	ALA	8.7
1	B	151	ALA	8.4
1	B	-2	ALA	7.4
1	B	145	LEU	7.2
1	A	157	GLN	6.7
1	B	147	LEU	6.5
1	B	8	GLU	6.1
1	A	155	ASP	5.8
1	B	9	CYS	5.7
1	B	169	ALA	5.5
1	A	0	HIS	5.4
1	B	170	LEU	5.4
1	A	246	GLN	5.3
1	B	5	ILE	4.5
1	A	154	VAL	4.5
1	B	150	LYS	4.4
1	B	137	ILE	4.2
1	B	177	TYR	4.1
1	A	-2	ALA	4.0
1	B	6	GLU	4.0
1	A	161	ASP	3.9
1	B	148	GLN	3.7
1	A	147	LEU	3.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	142	SER	3.5
1	B	-1	GLY	3.5
1	B	12	LEU	3.5
1	B	136	GLU	3.5
1	A	151	ALA	3.4
1	B	172	ASP	3.4
1	A	6	GLU	3.3
1	A	244	SER	3.3
1	B	11	ALA	3.3
1	B	173	VAL	3.2
1	A	159	ARG	3.2
1	B	4	VAL	3.2
1	B	0	HIS	3.1
1	B	242	SER	3.1
1	B	154	VAL	3.0
1	A	7	LYS	2.9
1	B	10	SER	2.9
1	B	152	LYS	2.7
1	B	174	ASN	2.7
1	B	7	LYS	2.6
1	B	129	GLU	2.6
1	B	143	ASP	2.6
1	B	2	GLU	2.6
1	A	177	TYR	2.6
1	A	5	ILE	2.6
1	B	175	ASP	2.6
1	A	166	LEU	2.5
1	B	139	LYS	2.5
1	A	10	SER	2.5
1	B	241	PRO	2.4
1	A	144	THR	2.3
1	A	162	ILE	2.3
1	B	149	LYS	2.3
1	A	152	LYS	2.3
1	B	144	THR	2.2
1	B	171	GLN	2.2
1	A	243	SER	2.2
1	B	146	LYS	2.2
1	A	165	GLN	2.2
1	A	153	LYS	2.2
1	B	153	LYS	2.2
1	A	149	LYS	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	3	ALA	2.1
1	A	245	GLU	2.1
1	B	133	ALA	2.1
1	B	51	ALA	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.