



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 19, 2016 – 10:19 PM GMT

PDB ID : 5CLR  
Title : Crystal structure of LegK4\_APO Kinase  
Authors : Flayhan, A.; Terradot, L.  
Deposited on : 2015-07-16  
Resolution : 3.71 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  
Mogul : unknown  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026982  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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-20026982

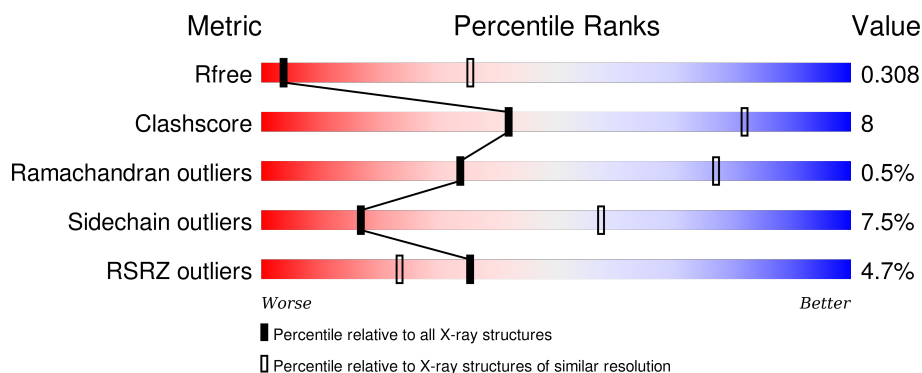
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.71 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))
$R_{free}$	91344	1101 (3.90-3.50)
Clashscore	102246	1224 (3.90-3.50)
Ramachandran outliers	100387	1172 (3.90-3.50)
Sidechain outliers	100360	1170 (3.90-3.50)
RSRZ outliers	91569	1108 (3.90-3.50)

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	451	<div> <div>2%</div> <div>64%</div> <div>18%</div> <div>•</div> <div>17%</div> </div>
1	B	451	<div> <div>5%</div> <div>51%</div> <div>16%</div> <div>•</div> <div>29%</div> </div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 5596 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 LegK4.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	376	Total	C	N	O	S	Se	0	0	0
			3017	1955	496	553	5	8			
1	B	321	Total	C	N	O	S	Se	0	0	0
			2579	1685	424	460	3	7			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	GLY	-	expression tag	UNP Q5WZW9
A	-4	ILE	-	expression tag	UNP Q5WZW9
A	-3	ASP	-	expression tag	UNP Q5WZW9
A	-2	PRO	-	expression tag	UNP Q5WZW9
A	-1	PHE	-	expression tag	UNP Q5WZW9
A	0	THR	-	expression tag	UNP Q5WZW9
A	1	MSE	-	expression tag	UNP Q5WZW9
B	-5	GLY	-	expression tag	UNP Q5WZW9
B	-4	ILE	-	expression tag	UNP Q5WZW9
B	-3	ASP	-	expression tag	UNP Q5WZW9
B	-2	PRO	-	expression tag	UNP Q5WZW9
B	-1	PHE	-	expression tag	UNP Q5WZW9
B	0	THR	-	expression tag	UNP Q5WZW9
B	1	MSE	-	expression tag	UNP Q5WZW9

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

Chain A:

2% 64% 18% 17%

GLY ILE ASP PRO PHE T0 M1 E8 LEU LYS LEU PRO GLY MSE D16 V26 L27 G28 N29 I32 I36 D40 G41 R46 Q58 S67 L68 K69 V73 P74 T75 E76 N77 K78 I79 SER GLY GLY LYS SER Y86 K91 F102 K107 I112

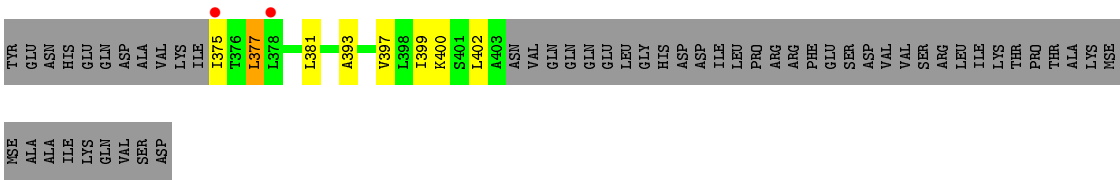
K125 V130 G131 T132 M135 E136 Q137 T138 Q139 H141 T145 I148 K149 S154 K157 N158 V164 L173 G174 I175 M176 L191 R194 K197 D204 A205 C209 V212 S218 S219 D220 T230 A235 L238 F247 V250 V264

L260 D261 E262 V270 K271 I277 L284 Y285 L286 T303 T306 S311 L314 Q315 V319 F320 GLU SER PHE LYS SER ARG F327 H328 A329 F330 L336 K341 SER SER CYS TYR ALA PRO GLU K349 E355 I356 K361 Y364 E365 ASN HIS ALA THR

D370 A371 V372 K373 I374 I375 T376 Q380 L381 S382 K383 ALA ASP GLY L387 L388 H391 K392 A393 L398 V405 Q406 GLN GLU LEU GLY HTS ASP ASP ILE LEU PRO ARG ARG PHE GLU SER SER ASP VAL SER ARG LEU ILE THR THR ALA LYS MSE MSE ALA THR

LYS GLN VAL SER ASP

Chain B:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	154.82Å 154.82Å 87.95Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	44.69 – 3.71 44.69 – 3.71	Depositor EDS
% Data completeness (in resolution range)	100.0 (44.69-3.71) 96.1 (44.69-3.71)	Depositor EDS
$R_{merge}$	0.24	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.78 (at 3.66Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, $R_{free}$	0.282 , 0.311 0.277 , 0.308	Depositor DCC
$R_{free}$ test set	620 reflections (4.89%)	DCC
Wilson B-factor (Å <sup>2</sup> )	105.5	Xtriage
Anisotropy	0.307	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 58.3	EDS
Estimated twinning fraction	0.046 for -h,-k,l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	0 of 13213 reflections	Xtriage
$F_o, F_c$ correlation	0.87	EDS
Total number of atoms	5596	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 4.32% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.21	0/3061	0.38	0/4111
1	B	0.21	0/2613	0.36	0/3502
All	All	0.21	0/5674	0.37	0/7613

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	3017	0	3108	52	0
1	B	2579	0	2695	47	0
All	All	5596	0	5803	96	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 (96) 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:A:194:ARG:NH1	1:A:218:SER:HB3	2.04	0.72
1:A:311:SER:O	1:A:315:GLN:NE2	2.23	0.72

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:250:VAL:HG12	1:A:306:ARG:HD2	1.75	0.68
1:B:136:GLU:OE2	1:B:136:GLU:N	2.29	0.65
1:A:135:ASN:ND2	1:A:137:GLN:O	2.31	0.63
1:B:228:TRP:CE3	1:B:228:TRP:HA	2.32	0.63
1:B:228:TRP:HE3	1:B:228:TRP:HA	1.65	0.61
1:B:135:ASN:HB2	1:B:140:VAL:HG23	1.83	0.59
1:B:257:ALA:HB2	1:B:298:LEU:HB3	1.84	0.58
1:A:286:LEU:HD12	1:A:393:ALA:HB2	1.87	0.57
1:A:197:LYS:HE3	1:A:230:THR:HG21	1.86	0.57
1:A:271:LYS:HA	1:B:299:MSE:HE1	1.87	0.57
1:B:132:THR:OG1	1:B:141:HIS:NE2	2.39	0.56
1:B:360:ILE:HD13	1:B:377:LEU:HD22	1.88	0.55
1:B:103:LEU:HD11	1:B:142:TYR:HB3	1.89	0.54
1:A:69:LYS:HB3	1:A:69:LYS:NZ	2.21	0.54
1:B:106:VAL:HG11	1:B:116:PHE:HB2	1.89	0.54
1:B:71:ILE:HG12	1:B:92:VAL:HB	1.89	0.54
1:A:250:VAL:HG11	1:A:303:THR:HG23	1.90	0.53
1:A:154:SER:O	1:A:158:ASN:ND2	2.40	0.53
1:A:235:ALA:HB3	1:A:238:LEU:HG	1.91	0.53
1:B:375:ILE:HG23	1:B:399:ILE:HG23	1.90	0.53
1:A:29:ASN:HB2	1:A:32:ILE:HG13	1.90	0.52
1:A:27:LEU:N	1:A:28:GLY:HA3	2.23	0.52
1:B:356:ILE:HD13	1:B:381:LEU:HB2	1.90	0.52
1:B:254:PHE:HB2	1:B:306:ARG:HH11	1.74	0.51
1:A:204:ASP:OD1	1:A:205:ALA:N	2.41	0.51
1:A:270:VAL:HG11	1:B:296:THR:HA	1.92	0.51
1:B:220:ASP:HB3	1:B:222:GLU:HG2	1.92	0.51
1:A:107:LYS:NZ	1:A:138:THR:HG22	2.25	0.50
1:B:125:LYS:NZ	1:B:191:LEU:HD11	2.27	0.50
1:A:270:VAL:HG12	1:B:299:MSE:HE2	1.93	0.50
1:A:26:VAL:C	1:A:28:GLY:HA3	2.33	0.49
1:B:284:LEU:HD23	1:B:397:VAL:HG22	1.94	0.49
1:A:77:ASN:ND2	1:A:91:LYS:O	2.46	0.49
1:B:281:ASN:HB3	1:B:283:HIS:CE1	2.48	0.48
1:B:341:LYS:NZ	1:B:341:LYS:HB3	2.28	0.48
1:A:194:ARG:NH2	1:A:220:ASP:O	2.41	0.48
1:A:1:MSE:HE1	1:A:36:ILE:HG22	1.96	0.48
1:A:376:THR:O	1:A:380:GLN:HG2	2.14	0.47
1:B:21:GLU:O	1:B:24:ILE:HG12	2.15	0.47
1:B:283:HIS:CE1	1:B:400:LYS:HG2	2.49	0.47
1:B:169:PHE:O	1:B:173:LEU:HB2	2.15	0.47

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:67:SER:HB2	1:A:68:LEU:HB3	1.97	0.47
1:A:132:THR:OG1	1:A:141:HIS:NE2	2.39	0.47
1:A:356:ILE:O	1:A:360:ILE:HG12	2.15	0.46
1:A:125:LYS:HD2	1:A:191:LEU:HD11	1.96	0.46
1:A:176:MSE:HE3	1:A:260:LEU:HG	1.97	0.46
1:A:336:LEU:HD23	1:A:398:LEU:HD11	1.98	0.46
1:A:138:THR:HB	1:A:140:VAL:HG23	1.98	0.46
1:A:40:ASP:HA	1:A:41:GLY:HA2	1.51	0.46
1:B:359:GLU:HG3	1:B:377:LEU:HD23	1.98	0.45
1:B:294:CYS:HB3	1:B:313:ILE:HB	1.97	0.45
1:B:119:GLU:HG3	1:B:214:PHE:HB2	1.98	0.45
1:A:102:PHE:HD2	1:A:145:THR:HB	1.81	0.45
1:A:77:ASN:OD1	1:A:77:ASN:N	2.48	0.45
1:B:175:ILE:HG23	1:B:209:CYS:SG	2.57	0.45
1:A:135:ASN:HD21	1:A:137:GLN:HB2	1.82	0.44
1:B:103:LEU:HD13	1:B:144:LEU:HD12	1.99	0.44
1:A:107:LYS:HZ2	1:A:138:THR:HG22	1.83	0.44
1:B:258:TYR:HD1	1:B:258:TYR:HA	1.73	0.44
1:A:77:ASN:HD22	1:A:91:LYS:HG3	1.82	0.44
1:A:67:SER:HA	1:A:68:LEU:HA	1.69	0.44
1:B:64:THR:HG21	1:B:69:LYS:HE3	2.00	0.44
1:B:63:THR:HA	1:B:68:LEU:HD22	2.00	0.44
1:B:298:LEU:HD21	1:B:310:PHE:HB2	2.00	0.44
1:A:327:PRO:HB2	1:A:328:GLU:H	1.64	0.44
1:B:94:ALA:H	1:B:100:THR:N	2.16	0.44
1:A:58:GLN:HB2	1:A:73:VAL:HG13	2.00	0.44
1:A:330:PHE:HE1	1:A:374:ILE:HG13	1.83	0.43
1:B:4:LEU:HB2	1:B:5:ARG:H	1.59	0.43
1:A:69:LYS:HB3	1:A:69:LYS:HZ3	1.83	0.42
1:A:194:ARG:HH12	1:A:218:SER:HB3	1.82	0.42
1:B:357:ILE:O	1:B:361:LYS:HG3	2.19	0.42
1:A:254:PHE:HB2	1:A:306:ARG:NH1	2.34	0.42
1:B:356:ILE:O	1:B:360:ILE:HG12	2.19	0.42
1:A:194:ARG:HG2	1:A:247:PHE:HB3	2.02	0.41
1:B:176:MSE:HE3	1:B:260:LEU:HG	2.02	0.41
1:A:330:PHE:HA	1:A:330:PHE:HD1	1.76	0.41
1:B:70:PRO:HG2	1:B:95:LEU:HB2	2.01	0.41
1:A:341:LYS:NZ	1:A:341:LYS:HB3	2.36	0.41
1:A:388:LEU:HD12	1:A:391:HIS:HE1	1.85	0.41
1:B:194:ARG:HG3	1:B:247:PHE:CG	2.56	0.41
1:B:393:ALA:O	1:B:397:VAL:HG23	2.20	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:218:SER:OG	1:A:219:SER:N	2.54	0.41
1:A:387:LEU:HD11	1:A:391:HIS:HB2	2.01	0.41
1:A:175:ILE:HG23	1:A:209:CYS:SG	2.61	0.41
1:B:196:LEU:HA	1:B:196:LEU:HD23	1.91	0.41
1:B:192:ILE:HG23	1:B:249:GLN:HA	2.03	0.40
1:A:46:ARG:HD2	1:A:75:THR:HG21	2.03	0.40
1:B:286:LEU:HD12	1:B:286:LEU:H	1.86	0.40
1:B:146:THR:HB	1:B:147:PHE:HD1	1.86	0.40
1:B:194:ARG:HB2	1:B:216:SER:HB3	2.04	0.40
1:B:224:LYS:HG2	1:B:225:PRO:HD2	2.02	0.40
1:A:157:LYS:HE2	1:A:262:GLU:O	2.21	0.40
1:A:277:ILE:HD12	1:A:284:LEU:HB2	2.04	0.40

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	362/451 (80%)	344 (95%)	17 (5%)	1 (0%)	46	83
1	B	297/451 (66%)	284 (96%)	11 (4%)	2 (1%)	26	72
All	All	659/902 (73%)	628 (95%)	28 (4%)	3 (0%)	34	77

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	147	PHE
1	B	148	ILE
1	A	130	VAL

### 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	338/390 (87%)	322 (95%)	16 (5%)	32	72
1	B	290/390 (74%)	259 (89%)	31 (11%)	8	42
All	All	628/780 (80%)	581 (92%)	47 (8%)	17	57

All (47) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	68	LEU
1	A	91	LYS
1	A	113	LEU
1	A	138	THR
1	A	148	ILE
1	A	149	LYS
1	A	164	VAL
1	A	173	LEU
1	A	212	VAL
1	A	286	LEU
1	A	314	LEU
1	A	315	GLN
1	A	319	VAL
1	A	364	TYR
1	A	365	GLU
1	A	372	VAL
1	B	4	LEU
1	B	5	ARG
1	B	22	LEU
1	B	68	LEU
1	B	71	ILE
1	B	92	VAL
1	B	100	THR
1	B	132	THR
1	B	133	PHE
1	B	140	VAL
1	B	143	ILE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	144	LEU
1	B	160	LEU
1	B	162	LEU
1	B	173	LEU
1	B	193	HIS
1	B	197	LYS
1	B	206	ASP
1	B	212	VAL
1	B	224	LYS
1	B	228	TRP
1	B	230	THR
1	B	241	GLN
1	B	242	GLU
1	B	258	TYR
1	B	298	LEU
1	B	314	LEU
1	B	333	LEU
1	B	357	ILE
1	B	377	LEU
1	B	402	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	135	ASN

### 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 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	368/451 (81%)	0.17	8 (2%) 65 50	76, 96, 154, 184	0
1	B	314/451 (69%)	0.48	24 (7%) 17 10	77, 116, 186, 233	0
All	All	682/902 (75%)	0.31	32 (4%) 35 23	76, 100, 176, 233	0

All (32) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	93	TYR	4.5
1	B	59	ILE	4.4
1	B	360	ILE	4.0
1	B	92	VAL	4.0
1	B	332	TYR	4.0
1	B	330	PHE	3.9
1	B	23	LEU	3.7
1	B	24	ILE	3.5
1	B	375	ILE	3.5
1	A	374	ILE	3.2
1	B	329	ALA	3.2
1	B	142	TYR	3.1
1	A	67	SER	2.9
1	B	359	GLU	2.8
1	B	74	PRO	2.7
1	B	112	SER	2.6
1	B	138	THR	2.5
1	B	378	LEU	2.5
1	B	71	ILE	2.5
1	B	72	ILE	2.4
1	A	381	LEU	2.3
1	B	135	ASN	2.3
1	A	371	ALA	2.3
1	B	137	GLN	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	361	LYS	2.2
1	B	116	PHE	2.2
1	A	370	ASP	2.1
1	B	190	GLY	2.1
1	A	405	VAL	2.0
1	A	355	GLU	2.0
1	B	333	LEU	2.0
1	B	68	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.