



## wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Aug 20, 2017 – 07:30 PM EDT

PDB ID : 3ZBJ  
EMDB ID: : EMD-2233  
Title : Fitting results in the I-layer of the subnanometer structure of the bacterial pKM101 type IV secretion system core complex digested with elastase  
Authors : Rivera-Calzada, A.; Fronzes, R.; Savva, C.G.; Chandran, V.; Lian, P.W.; Laeremans, T.; Pardon, E.; Steyaert, J.; Remaut, H.; Waksman, G.; Orlova, E.V.  
Deposited on : unknown  
Resolution : 8.50 Å(reported)

This is a wwPDB/EMDataBank EM Map/Model Validation Summary Report  
for a publicly released PDB/EMDB 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/EMValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

MolProbity : 4.02b-467  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20029824

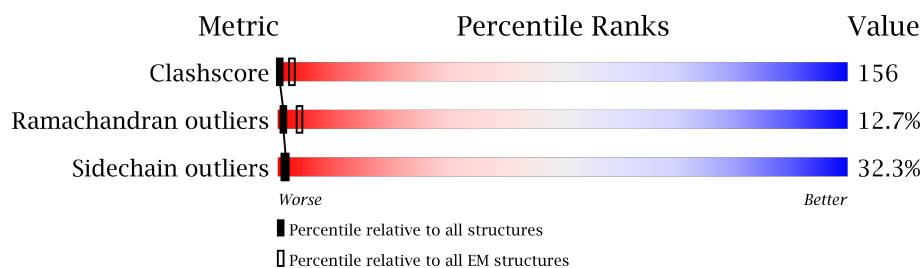
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 8.50 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.




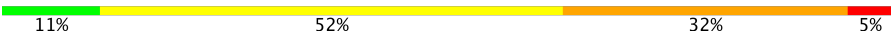
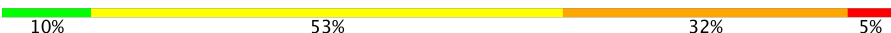


Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	112	11% 52% 32% 5%
1	B	112	10% 53% 32% 5%
1	C	112	10% 53% 32% 5%
1	D	112	10% 53% 32% 5%
1	E	112	10% 52% 33% 5%
1	F	112	11% 52% 32% 5%
1	G	112	11% 52% 32% 5%
1	H	112	10% 52% 33% 5%
1	I	112	11% 52% 32% 5%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	J	112	
1	K	112	
1	L	112	
1	M	112	
1	N	112	

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 12404 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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.

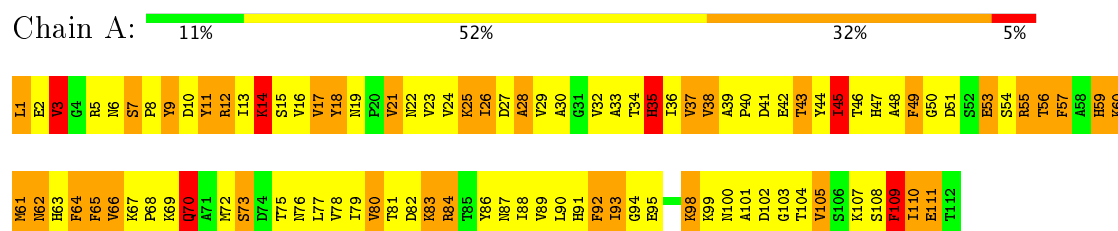
- Molecule 1 is a protein called TRA0 PROTEIN.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	112	Total	C	N	O	S	0	0
			886	564	151	169	2		
1	B	112	Total	C	N	O	S	0	0
			886	564	151	169	2		
1	C	112	Total	C	N	O	S	0	0
			886	564	151	169	2		
1	D	112	Total	C	N	O	S	0	0
			886	564	151	169	2		
1	E	112	Total	C	N	O	S	0	0
			886	564	151	169	2		
1	F	112	Total	C	N	O	S	0	0
			886	564	151	169	2		
1	G	112	Total	C	N	O	S	0	0
			886	564	151	169	2		
1	H	112	Total	C	N	O	S	0	0
			886	564	151	169	2		
1	I	112	Total	C	N	O	S	0	0
			886	564	151	169	2		
1	J	112	Total	C	N	O	S	0	0
			886	564	151	169	2		
1	K	112	Total	C	N	O	S	0	0
			886	564	151	169	2		
1	L	112	Total	C	N	O	S	0	0
			886	564	151	169	2		
1	M	112	Total	C	N	O	S	0	0
			886	564	151	169	2		
1	N	112	Total	C	N	O	S	0	0
			886	564	151	169	2		

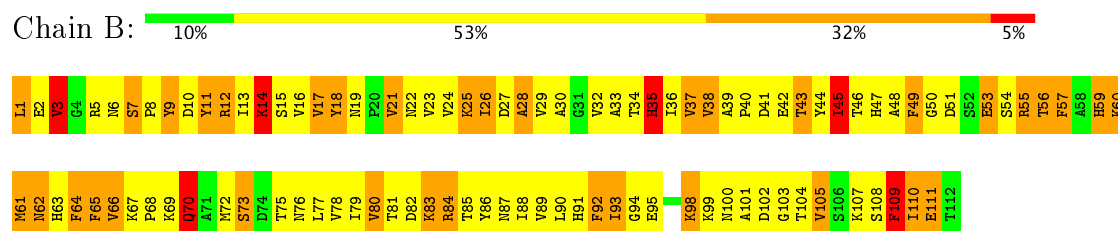
### 3 Residue-property plots [i](#)

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. 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. 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.

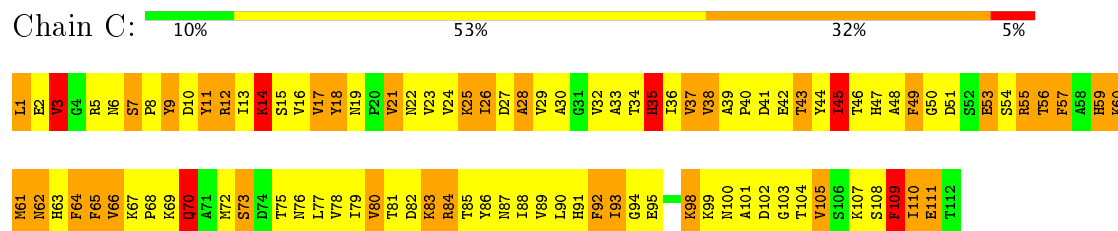
#### • Molecule 1: TRAO PROTEIN



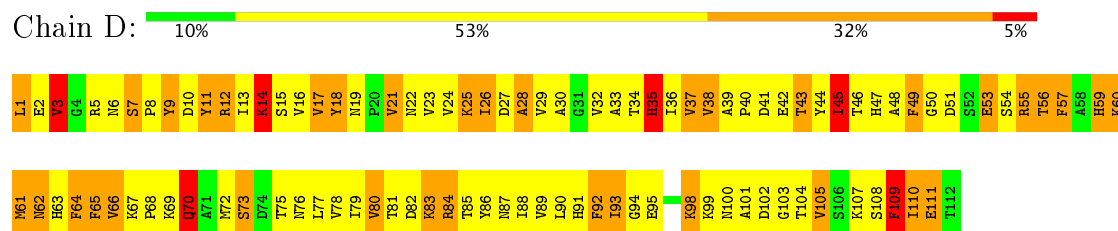
#### • Molecule 1: TRAO PROTEIN



#### • Molecule 1: TRAO PROTEIN

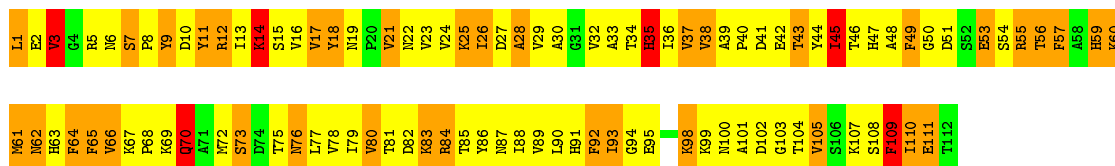


#### • Molecule 1: TRAO PROTEIN



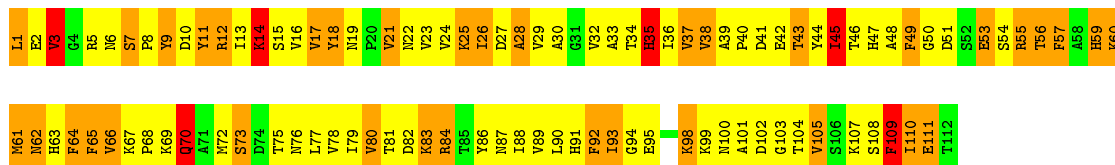
#### • Molecule 1: TRAO PROTEIN

Chain E: 10% 52% 33% 5%



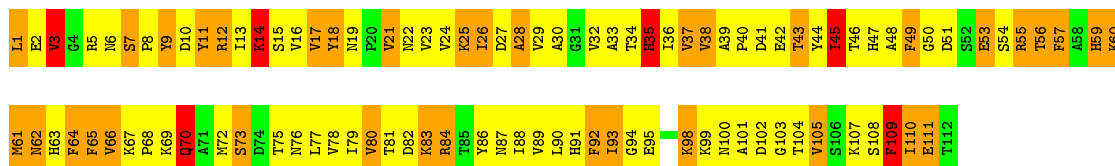
• Molecule 1: TRAO PROTEIN

Chain F: 11% 52% 32% 5%



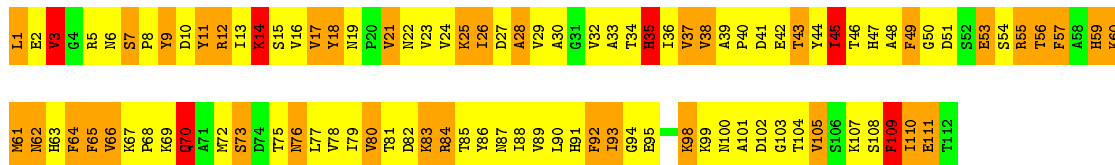
• Molecule 1: TRAO PROTEIN

Chain G: 11% 52% 32% 5%



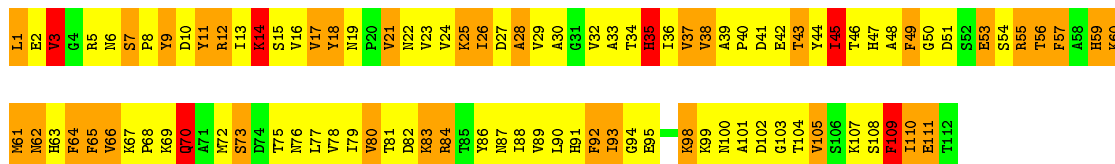
• Molecule 1: TRAO PROTEIN

Chain H: 10% 52% 33% 5%



• Molecule 1: TRAO PROTEIN

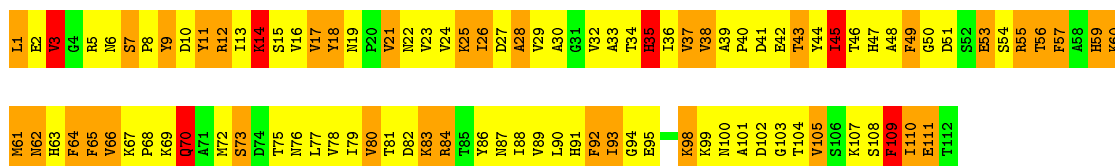
Chain I: 11% 52% 32% 5%



• Molecule 1: TRAO PROTEIN

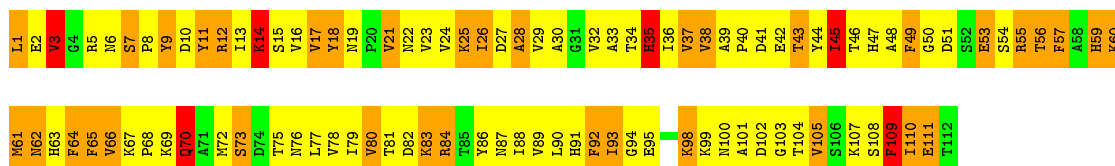
Chain J: 11% 52% 32% 5%





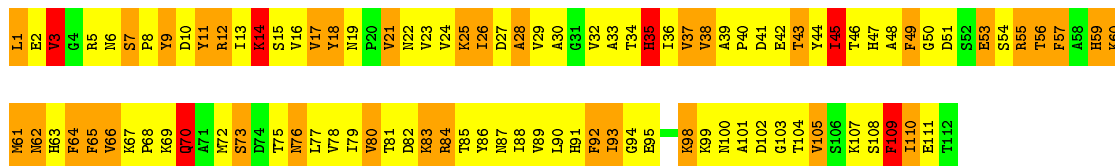
- Molecule 1: TRAO PROTEIN

Chain K: 11% 52% 32% 5%



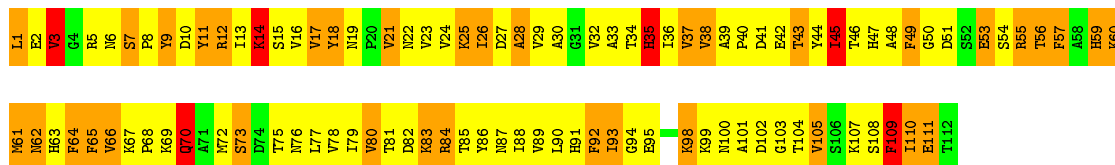
- Molecule 1: TRAO PROTEIN

Chain L: 10% 53% 32% 5%



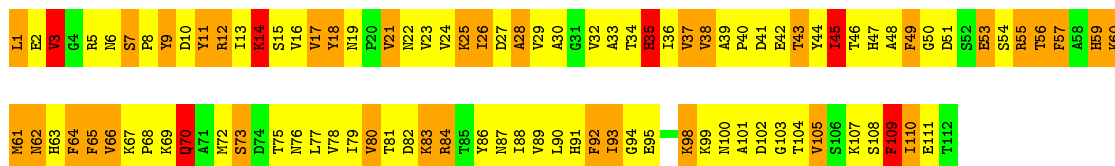
- Molecule 1: TRAO PROTEIN

Chain M: 10% 53% 32% 5%



- Molecule 1: TRAO PROTEIN

Chain N: 11% 53% 31% 5%



## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C14	Depositor
Number of particles used	5430	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	PHASE FLIPPING, EACH CCD IMAGE	Depositor
Microscope	FEI TECNAI 12	Depositor
Voltage (kV)	200	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	20	Depositor
Minimum defocus (nm)	1500	Depositor
Maximum defocus (nm)	3500	Depositor
Magnification	68100	Depositor
Image detector	GENERIC GATAN	Depositor



## 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  >2	RMSZ	# Z  >2
1	A	1.02	0/905	1.33	0/1227
1	B	1.02	0/905	1.33	0/1227
1	C	1.02	0/905	1.33	0/1227
1	D	1.02	0/905	1.33	0/1227
1	E	1.02	0/905	1.33	0/1227
1	F	1.02	0/905	1.33	0/1227
1	G	1.02	0/905	1.33	0/1227
1	H	1.02	0/905	1.33	0/1227
1	I	1.02	0/905	1.33	0/1227
1	J	1.02	0/905	1.33	0/1227
1	K	1.02	0/905	1.33	0/1227
1	L	1.02	0/905	1.33	0/1227
1	M	1.02	0/905	1.33	0/1227
1	N	1.02	0/905	1.33	0/1227
All	All	1.02	0/12670	1.33	0/17178

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
1	A	0	7
1	B	0	7
1	C	0	7
1	D	0	7
1	E	0	7
1	F	0	7
1	G	0	7
1	H	0	7
1	I	0	7
1	J	0	7
1	K	0	7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	#Chirality outliers	#Planarity outliers
1	L	0	7
1	M	0	7
1	N	0	7
All	All	0	98

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 98 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	12	ARG	Sidechain
1	A	18	TYR	Peptide
1	A	45	ILE	Peptide
1	A	56	THR	Peptide
1	A	7	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	886	0	881	302	0
1	B	886	0	881	307	0
1	C	886	0	881	307	0
1	D	886	0	881	304	0
1	E	886	0	881	305	0
1	F	886	0	881	305	0
1	G	886	0	881	306	0
1	H	886	0	881	308	0
1	I	886	0	881	308	0
1	J	886	0	881	301	0
1	K	886	0	881	302	0
1	L	886	0	881	309	0
1	M	886	0	881	308	0
1	N	886	0	881	306	0
All	All	12404	0	12334	3864	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 156.

The worst 5 of 3864 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:M:91:HIS:HA	1:M:104:THR:HB	1.24	1.18
1:F:1:LEU:HD12	1:F:16:VAL:HG23	1.23	1.18
1:A:91:HIS:HA	1:A:104:THR:HB	1.24	1.17
1:E:1:LEU:HD12	1:E:16:VAL:HG23	1.24	1.16
1:G:1:LEU:HD12	1:G:16:VAL:HG23	1.24	1.16

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 PDB entries followed by that with respect to all EM entries.

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	110/112 (98%)	70 (64%)	26 (24%)	14 (13%)	0	7
1	B	110/112 (98%)	70 (64%)	26 (24%)	14 (13%)	0	7
1	C	110/112 (98%)	70 (64%)	26 (24%)	14 (13%)	0	7
1	D	110/112 (98%)	70 (64%)	26 (24%)	14 (13%)	0	7
1	E	110/112 (98%)	70 (64%)	26 (24%)	14 (13%)	0	7
1	F	110/112 (98%)	70 (64%)	26 (24%)	14 (13%)	0	7
1	G	110/112 (98%)	70 (64%)	26 (24%)	14 (13%)	0	7
1	H	110/112 (98%)	70 (64%)	26 (24%)	14 (13%)	0	7
1	I	110/112 (98%)	70 (64%)	26 (24%)	14 (13%)	0	7
1	J	110/112 (98%)	70 (64%)	26 (24%)	14 (13%)	0	7
1	K	110/112 (98%)	70 (64%)	26 (24%)	14 (13%)	0	7
1	L	110/112 (98%)	70 (64%)	26 (24%)	14 (13%)	0	7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	M	110/112 (98%)	70 (64%)	26 (24%)	14 (13%)	0	7
1	N	110/112 (98%)	70 (64%)	26 (24%)	14 (13%)	0	7
All	All	1540/1568 (98%)	980 (64%)	364 (24%)	196 (13%)	1	7

5 of 196 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	3	VAL
1	A	14	LYS
1	A	21	VAL
1	A	28	ALA
1	A	33	ALA

### 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 PDB entries followed by that with respect to all EM entries.

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	99/99 (100%)	67 (68%)	32 (32%)	0	2
1	B	99/99 (100%)	67 (68%)	32 (32%)	0	2
1	C	99/99 (100%)	67 (68%)	32 (32%)	0	2
1	D	99/99 (100%)	67 (68%)	32 (32%)	0	2
1	E	99/99 (100%)	67 (68%)	32 (32%)	0	2
1	F	99/99 (100%)	67 (68%)	32 (32%)	0	2
1	G	99/99 (100%)	67 (68%)	32 (32%)	0	2
1	H	99/99 (100%)	67 (68%)	32 (32%)	0	2
1	I	99/99 (100%)	67 (68%)	32 (32%)	0	2
1	J	99/99 (100%)	67 (68%)	32 (32%)	0	2
1	K	99/99 (100%)	67 (68%)	32 (32%)	0	2
1	L	99/99 (100%)	67 (68%)	32 (32%)	0	2
1	M	99/99 (100%)	67 (68%)	32 (32%)	0	2
1	N	99/99 (100%)	67 (68%)	32 (32%)	0	2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1386/1386 (100%)	938 (68%)	448 (32%)	<b>1</b> <b>2</b>

5 of 448 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	45	ILE
1	H	92	PHE
1	M	110	ILE
1	G	60	LYS
1	H	11	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 73 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	91	HIS
1	I	70	GLN
1	M	100	ASN
1	H	87	ASN
1	I	87	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 ⓘ

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.