



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 8, 2017 – 01:12 AM EDT

PDB ID : 2Z5H  
Title : Crystal structure of the head-to-tail junction of tropomyosin complexed with a fragment of TnT  
Authors : Murakami, K.; Nozawa, K.; Tomii, K.; Kudou, N.; Igarashi, N.; Shirakihara, Y.; Wakatsuki, S.; Stewart, M.; Yasunaga, T.; Wakabayashi, T.  
Deposited on : unknown  
Resolution : 2.89 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

<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-20029824
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-20029824

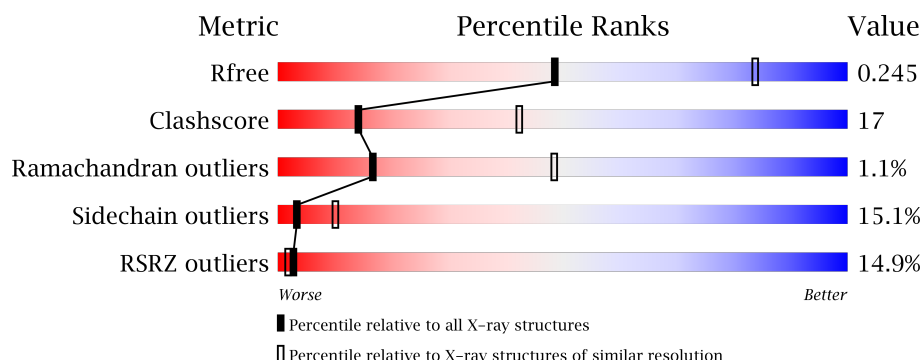
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.89 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	1586 (2.90-2.90)
Clashscore	112137	1807 (2.90-2.90)
Ramachandran outliers	110173	1768 (2.90-2.90)
Sidechain outliers	110143	1770 (2.90-2.90)
RSRZ outliers	101464	1596 (2.90-2.90)

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	52	<div> <div>2%</div> <div>65% 25% 6%</div> </div>
1	B	52	<div> <div>2%</div> <div>83% 12%</div> </div>
1	C	52	<div> <div>2%</div> <div>75% 21%</div> </div>
1	D	52	<div> <div>6%</div> <div>83% 12% 6%</div> </div>
1	E	52	<div> <div>83% 8% 6%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	52	<div><div>4%</div><div><div></div><div>81%</div><div>17%</div><div></div></div><div></div></div>
1	G	52	<div><div>4%</div><div><div></div><div>69%</div><div>27%</div><div></div></div><div></div></div>
1	H	52	<div><div>2%</div><div><div></div><div>81%</div><div>12%</div><div></div></div><div></div></div>
2	I	40	<div><div></div><div><div>90%</div><div>30%</div><div>48%</div><div>20%</div><div></div></div><div></div></div>
3	T	55	<div><div></div><div><div>44%</div><div>5%</div><div>33%</div><div>18%</div><div>5%</div><div>38%</div><div></div></div><div></div></div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 3940 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 General control protein GCN4 and Tropomyosin alpha-1 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	49	Total	C	N	O	S	0	0	0
			403	254	66	82	1			
1	B	51	Total	C	N	O	S	0	0	0
			416	261	68	86	1			
1	C	52	Total	C	N	O	S	0	0	0
			424	267	69	87	1			
1	D	49	Total	C	N	O	S	0	0	0
			403	254	66	82	1			
1	E	49	Total	C	N	O	S	0	0	0
			403	254	66	82	1			
1	F	51	Total	C	N	O	S	0	0	0
			416	261	68	86	1			
1	G	52	Total	C	N	O	S	0	0	0
			424	267	69	87	1			
1	H	49	Total	C	N	O	S	0	0	0
			403	254	66	82	1			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	233	GLY	-	EXPRESSION TAG	UNP P03069
B	233	GLY	-	EXPRESSION TAG	UNP P03069
C	233	GLY	-	EXPRESSION TAG	UNP P03069
D	233	GLY	-	EXPRESSION TAG	UNP P03069
E	233	GLY	-	EXPRESSION TAG	UNP P03069
F	233	GLY	-	EXPRESSION TAG	UNP P03069
G	233	GLY	-	EXPRESSION TAG	UNP P03069
H	233	GLY	-	EXPRESSION TAG	UNP P03069

- Molecule 2 is a protein called Tropomyosin alpha-1 chain and General control protein GCN4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	I	39	Total	C	N	O	S	0	0	0
			306	189	56	58	3			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	-3	GLY	-	EXPRESSION TAG	UNP P58772
I	-2	ALA	-	EXPRESSION TAG	UNP P58772
I	-1	ALA	-	EXPRESSION TAG	UNP P58772
I	0	SER	-	EXPRESSION TAG	UNP P58772

- Molecule 3 is a protein called Troponin T, fast skeletal muscle isoforms.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	T	34	Total	C	N	O	8	0	0
			287	179	53	55			

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	8	Total	O	0	0
			8	8		
4	B	7	Total	O	0	0
			7	7		
4	C	10	Total	O	0	0
			10	10		
4	D	3	Total	O	0	0
			3	3		
4	E	4	Total	O	0	0
			4	4		
4	F	12	Total	O	0	0
			12	12		
4	G	2	Total	O	0	0
			2	2		
4	H	2	Total	O	0	0
			2	2		
4	I	7	Total	O	0	0
			7	7		

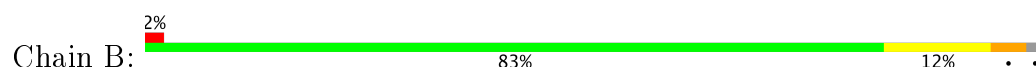
### 3 Residue-property plots [i](#)

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.

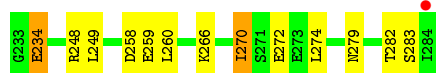
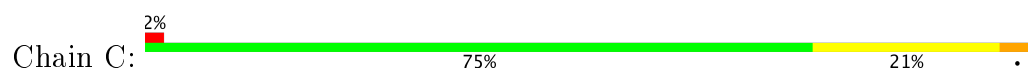
- Molecule 1: General control protein GCN4 and Tropomyosin alpha-1 chain



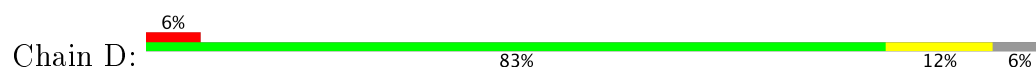
- Molecule 1: General control protein GCN4 and Tropomyosin alpha-1 chain



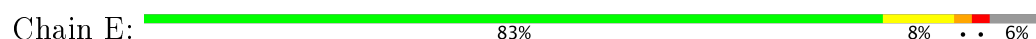
- Molecule 1: General control protein GCN4 and Tropomyosin alpha-1 chain



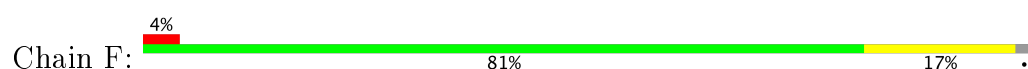
- Molecule 1: General control protein GCN4 and Tropomyosin alpha-1 chain



- Molecule 1: General control protein GCN4 and Tropomyosin alpha-1 chain

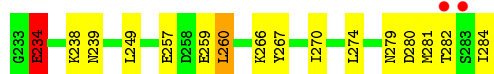


- Molecule 1: General control protein GCN4 and Tropomyosin alpha-1 chain

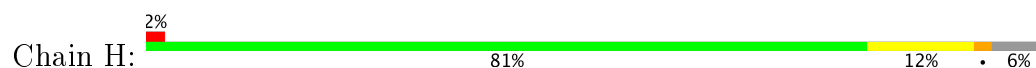




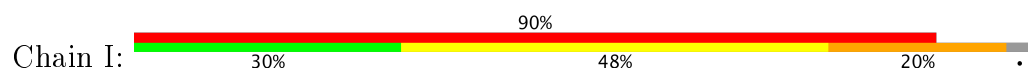
- Molecule 1: General control protein GCN4 and Tropomyosin alpha-1 chain



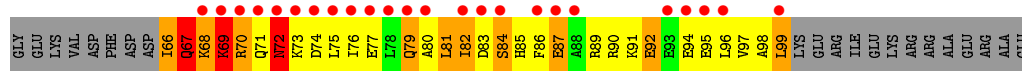
- Molecule 1: General control protein GCN4 and Tropomyosin alpha-1 chain



- Molecule 2: Tropomyosin alpha-1 chain and General control protein GCN4



- Molecule 3: Troponin T, fast skeletal muscle isoforms



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	86.81Å 158.31Å 163.57Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	113.96 – 2.89 40.55 – 2.89	Depositor EDS
% Data completeness (in resolution range)	95.9 (113.96-2.89) 95.9 (40.55-2.89)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.16 (at 2.90Å)	Xtriage
Refinement program	REFMAC 5.2.0019, CNS 1.1	Depositor
R, $R_{free}$	0.237 , 0.247 0.227 , 0.245	Depositor DCC
$R_{free}$ test set	1251 reflections (5.36%)	DCC
Wilson B-factor (Å <sup>2</sup> )	50.9	Xtriage
Anisotropy	0.730	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 46.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.064 for -h,-l,-k	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	3940	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.51% 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	0.55	0/407	0.98	2/545 (0.4%)
1	B	0.57	0/420	1.07	0/563
1	C	0.45	0/428	0.83	2/574 (0.3%)
1	D	0.47	0/407	0.81	0/545
1	E	0.51	0/407	0.94	3/545 (0.6%)
1	F	0.51	0/420	0.89	0/563
1	G	0.54	0/428	0.87	0/574
1	H	0.52	0/407	0.81	0/545
2	I	0.40	0/305	0.95	1/401 (0.2%)
3	T	0.76	2/288 (0.7%)	1.37	4/382 (1.0%)
All	All	0.53	2/3917 (0.1%)	0.95	12/5237 (0.2%)

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	C	0	1
1	G	0	1
All	All	0	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	T	69	LYS	CB-CG	-8.01	1.30	1.52
3	T	66	ILE	CB-CG2	-6.66	1.32	1.52

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	T	69	LYS	CA-CB-CG	7.17	129.17	113.40
1	E	255	ASP	CB-CG-OD1	6.12	123.81	118.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	T	82	ILE	CB-CA-C	-5.64	100.31	111.60
2	I	35	LEU	CA-CB-CG	5.60	128.17	115.30
1	C	248	ARG	NE-CZ-NH1	5.50	123.05	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	234	GLU	Peptide
1	G	234	GLU	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	403	0	400	24	0
1	B	416	0	412	10	0
1	C	424	0	423	3	0
1	D	403	0	400	2	0
1	E	403	0	400	3	0
1	F	416	0	412	3	0
1	G	424	0	423	10	0
1	H	403	0	400	7	0
2	I	306	0	334	29	2
3	T	287	0	297	75	2
4	A	8	0	0	0	0
4	B	7	0	0	1	0
4	C	10	0	0	0	0
4	D	3	0	0	0	0
4	E	4	0	0	0	0
4	F	12	0	0	0	0
4	G	2	0	0	0	0
4	H	2	0	0	0	0
4	I	7	0	0	1	0
All	All	3940	0	3901	134	2

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 17.

The worst 5 of 134 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:276:HIS:CE1	3:T:83:ASP:HB2	1.19	1.67
1:A:276:HIS:CE1	3:T:83:ASP:CB	1.97	1.47
1:A:276:HIS:ND1	3:T:83:ASP:HB2	1.22	1.41
3:T:86:PHE:CE1	3:T:87:GLU:HB2	1.61	1.33
3:T:84:SER:HA	3:T:86:PHE:CE2	1.69	1.25

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:I:9:GLN:NE2	3:T:84:SER:O[7_455]	1.87	0.33
2:I:9:GLN:CD	3:T:84:SER:O[7_455]	2.11	0.09

## 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	47/52 (90%)	46 (98%)	0	1 (2%)	8	30
1	B	49/52 (94%)	46 (94%)	3 (6%)	0	100	100
1	C	50/52 (96%)	48 (96%)	2 (4%)	0	100	100
1	D	47/52 (90%)	47 (100%)	0	0	100	100
1	E	47/52 (90%)	46 (98%)	1 (2%)	0	100	100
1	F	49/52 (94%)	46 (94%)	3 (6%)	0	100	100
1	G	50/52 (96%)	48 (96%)	2 (4%)	0	100	100
1	H	47/52 (90%)	47 (100%)	0	0	100	100
2	I	37/40 (92%)	32 (86%)	4 (11%)	1 (3%)	6	23

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	T	32/55 (58%)	26 (81%)	3 (9%)	3 (9%)	1	1
All	All	455/511 (89%)	432 (95%)	18 (4%)	5 (1%)	17	48

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	280	ASP
2	I	-1	ALA
3	T	70	ARG
3	T	67	GLN
3	T	72	ASN

### 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	44/47 (94%)	41 (93%)	3 (7%)	18	47
1	B	46/47 (98%)	40 (87%)	6 (13%)	5	15
1	C	47/47 (100%)	38 (81%)	9 (19%)	2	5
1	D	44/47 (94%)	40 (91%)	4 (9%)	11	32
1	E	44/47 (94%)	40 (91%)	4 (9%)	11	32
1	F	46/47 (98%)	41 (89%)	5 (11%)	7	22
1	G	47/47 (100%)	40 (85%)	7 (15%)	3	10
1	H	44/47 (94%)	40 (91%)	4 (9%)	11	32
2	I	32/33 (97%)	24 (75%)	8 (25%)	1	2
3	T	31/49 (63%)	17 (55%)	14 (45%)	0	0
All	All	425/458 (93%)	361 (85%)	64 (15%)	3	10

5 of 64 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	251	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	G	274	LEU
3	T	84	SER
1	F	252	LEU
1	G	249	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	239	ASN
3	T	67	GLN
2	I	17	ASN
1	C	276	HIS
1	H	244	ASN

### 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	49/52 (94%)	-0.25	1 (2%) 65 62	20, 37, 88, 102	0
1	B	51/52 (98%)	-0.20	1 (1%) 65 62	22, 32, 58, 119	0
1	C	52/52 (100%)	-0.08	1 (1%) 67 64	34, 46, 77, 98	0
1	D	49/52 (94%)	0.23	3 (6%) 22 17	37, 49, 80, 111	0
1	E	49/52 (94%)	-0.12	0 100 100	24, 38, 69, 95	0
1	F	51/52 (98%)	-0.03	2 (3%) 40 35	26, 35, 58, 109	0
1	G	52/52 (100%)	0.11	2 (3%) 41 35	32, 47, 75, 110	0
1	H	49/52 (94%)	-0.05	1 (2%) 65 62	37, 48, 76, 101	0
2	I	39/40 (97%)	4.26	36 (92%) 0 0	133, 167, 187, 189	2 (5%)
3	T	34/55 (61%)	3.33	24 (70%) 0 0	167, 199, 214, 216	10 (29%)
All	All	475/511 (92%)	0.55	71 (14%) 3 2	20, 45, 186, 216	12 (2%)

The worst 5 of 71 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	I	25	LEU	11.9
2	I	35	LEU	10.6
2	I	32	LEU	7.5
2	I	22	ALA	6.5
3	T	72	ASN	6.3

### 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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.