



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 31, 2016 – 07:35 PM GMT

PDB ID : 1GCS  
Title : STRUCTURE OF THE BOVINE GAMMA-B CRYSTALLIN AT 150K  
Authors : Najmudin, S.; Lindley, P.; Slingsby, C.; Bateman, O.; Myles, D.; Kumaraswamy, S.; Glover, I.  
Deposited on : 1994-01-27  
Resolution : 2.00 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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
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) : trunk26865

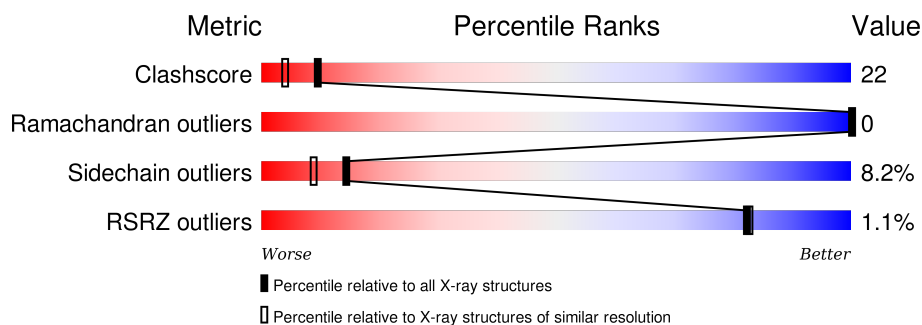
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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))
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)
RSRZ outliers	91569	6262 (2.00-2.00)

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	174	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 1729 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 GAMMA-B CRYSTALLIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	174	Total	C	N	O	S	0	0	0
			1474	926	266	268	14			

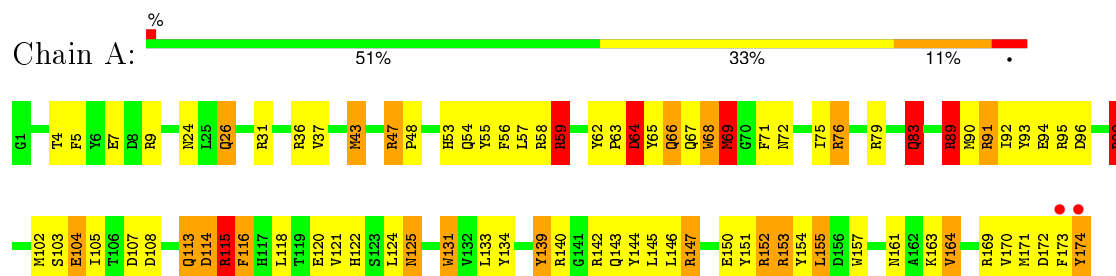
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	255	Total	O	0	0
			255	255		

### 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 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.

• Molecule 1: GAMMA-B CRYSTALLIN



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	56.41Å 56.41Å 97.11Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	8.00 – 2.00 7.91 – 1.95	Depositor EDS
% Data completeness (in resolution range)	96.0 (8.00-2.00) 94.1 (7.91-1.95)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	9.48 (at 1.95Å)	Xtriage
Refinement program	RESTRAIN	Depositor
R, $R_{free}$	0.170 , (Not available) 0.160 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	11.9	Xtriage
Anisotropy	0.173	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 74.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 11118 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	1729	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.47% 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	1.52	11/1518 (0.7%)	2.06	49/2046 (2.4%)

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	13

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	104	GLU	CD-OE1	-6.82	1.18	1.25
1	A	174	TYR	C-OXT	6.04	1.34	1.23
1	A	93	TYR	CE1-CZ	5.64	1.45	1.38
1	A	68	TRP	NE1-CE2	-5.59	1.30	1.37
1	A	54	GLN	CD-OE1	5.49	1.36	1.24
1	A	62	TYR	CA-CB	-5.43	1.42	1.53
1	A	64	ASP	CA-CB	-5.29	1.42	1.53
1	A	83	GLN	CD-OE1	5.17	1.35	1.24
1	A	125	ASN	N-CA	5.08	1.56	1.46
1	A	131	TRP	CD1-NE1	5.05	1.46	1.38
1	A	125	ASN	CG-OD1	5.03	1.35	1.24

All (49) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	142	ARG	NE-CZ-NH1	14.86	127.73	120.30
1	A	59	ARG	CD-NE-CZ	11.42	139.59	123.60
1	A	140	ARG	NE-CZ-NH1	11.41	126.00	120.30
1	A	89	ARG	CD-NE-CZ	-9.72	109.99	123.60
1	A	89	ARG	NE-CZ-NH1	-9.34	115.63	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	115	ARG	NE-CZ-NH1	9.05	124.83	120.30
1	A	76	ARG	NE-CZ-NH2	8.43	124.52	120.30
1	A	76	ARG	NE-CZ-NH1	-8.34	116.13	120.30
1	A	140	ARG	C-N-CA	8.24	139.60	122.30
1	A	142	ARG	NE-CZ-NH2	-8.19	116.21	120.30
1	A	151	TYR	CG-CD2-CE2	-7.91	114.97	121.30
1	A	164	VAL	CA-CB-CG1	7.73	122.50	110.90
1	A	36	ARG	NE-CZ-NH1	-7.64	116.48	120.30
1	A	142	ARG	CD-NE-CZ	7.54	134.15	123.60
1	A	37	VAL	CA-CB-CG1	-7.16	100.16	110.90
1	A	71	PHE	CB-CG-CD2	-7.15	115.80	120.80
1	A	154	TYR	CB-CG-CD2	-7.10	116.74	121.00
1	A	89	ARG	NE-CZ-NH2	7.05	123.82	120.30
1	A	116	PHE	CB-CG-CD2	6.81	125.57	120.80
1	A	4	THR	CA-CB-CG2	-6.53	103.26	112.40
1	A	144	TYR	CB-CG-CD1	-6.46	117.12	121.00
1	A	174	TYR	CB-CG-CD1	-6.42	117.15	121.00
1	A	154	TYR	CG-CD1-CE1	-6.39	116.19	121.30
1	A	170	VAL	CA-CB-CG2	6.12	120.08	110.90
1	A	114	ASP	CB-CG-OD1	6.06	123.76	118.30
1	A	147	ARG	CD-NE-CZ	5.99	131.98	123.60
1	A	55	TYR	CG-CD2-CE2	-5.83	116.64	121.30
1	A	79	ARG	CD-NE-CZ	-5.82	115.45	123.60
1	A	43	MET	CG-SD-CE	5.81	109.50	100.20
1	A	140	ARG	NE-CZ-NH2	-5.71	117.45	120.30
1	A	124	LEU	C-N-CA	-5.57	107.77	121.70
1	A	79	ARG	CG-CD-NE	-5.57	100.11	111.80
1	A	65	TYR	CB-CG-CD2	-5.52	117.69	121.00
1	A	68	TRP	CG-CD1-NE1	-5.51	104.59	110.10
1	A	69	MET	CG-SD-CE	5.50	109.00	100.20
1	A	5	PHE	CZ-CE2-CD2	5.44	126.63	120.10
1	A	94	GLU	CB-CG-CD	-5.31	99.86	114.20
1	A	64	ASP	CB-CA-C	5.29	120.98	110.40
1	A	115	ARG	N-CA-CB	-5.26	101.13	110.60
1	A	150	GLU	O-C-N	-5.25	114.30	122.70
1	A	151	TYR	CB-CG-CD1	-5.21	117.87	121.00
1	A	157	TRP	CA-CB-CG	5.17	123.53	113.70
1	A	93	TYR	CD1-CE1-CZ	-5.14	115.18	119.80
1	A	108	ASP	CB-CG-OD1	5.11	122.90	118.30
1	A	53	HIS	CB-CA-C	-5.10	100.19	110.40
1	A	75	ILE	O-C-N	5.07	130.82	122.70
1	A	7	GLU	OE1-CD-OE2	-5.03	117.26	123.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	57	LEU	CA-CB-CG	5.02	126.85	115.30
1	A	139	TYR	C-N-CA	5.00	134.21	121.70

There are no chirality outliers.

All (13) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	115	ARG	Sidechain
1	A	147	ARG	Sidechain
1	A	152	ARG	Sidechain
1	A	153	ARG	Sidechain
1	A	169	ARG	Sidechain
1	A	31	ARG	Sidechain
1	A	47	ARG	Sidechain
1	A	59	ARG	Sidechain
1	A	76	ARG	Sidechain
1	A	89	ARG	Sidechain
1	A	9	ARG	Sidechain
1	A	91	ARG	Sidechain
1	A	99	ARG	Sidechain

## 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	1474	0	1370	64	0
2	A	255	0	0	28	4
All	All	1729	0	1370	64	4

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 22.

All (64) 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:68:TRP:O	1:A:69:MET:HB2	1.49	1.10

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:67:GLN:HG3	2:A:402:HOH:O	1.53	1.09
1:A:96:ASP:HB3	2:A:287:HOH:O	1.65	0.97
1:A:173:PHE:HZ	2:A:441:HOH:O	1.50	0.92
1:A:58:ARG:NH1	2:A:434:HOH:O	2.08	0.86
1:A:69:MET:HB3	2:A:328:HOH:O	1.77	0.85
1:A:89:ARG:NH1	1:A:104:GLU:OE2	2.11	0.83
1:A:113:GLN:NE2	2:A:282:HOH:O	2.14	0.80
1:A:59:ARG:NH1	1:A:173:PHE:CD1	2.51	0.79
1:A:56:PHE:H	1:A:143:GLN:HE22	1.31	0.79
1:A:83:GLN:HG2	2:A:377:HOH:O	1.82	0.78
1:A:68:TRP:O	1:A:69:MET:CB	2.26	0.77
1:A:83:GLN:CG	2:A:377:HOH:O	2.34	0.73
1:A:89:ARG:HD2	1:A:104:GLU:CD	2.09	0.73
1:A:59:ARG:HH12	1:A:173:PHE:HB2	1.56	0.70
1:A:152:ARG:HG2	2:A:383:HOH:O	1.94	0.67
1:A:116:PHE:HB3	1:A:118:LEU:HD12	1.76	0.66
1:A:114:ASP:OD1	2:A:426:HOH:O	2.14	0.64
1:A:83:GLN:O	2:A:272:HOH:O	2.15	0.63
1:A:122:HIS:HB3	2:A:287:HOH:O	1.99	0.62
1:A:26:GLN:OE1	1:A:48:PRO:HG3	1.99	0.62
1:A:122:HIS:HD2	2:A:335:HOH:O	1.81	0.62
1:A:173:PHE:C	1:A:174:TYR:CD1	2.75	0.60
1:A:59:ARG:NH1	1:A:173:PHE:HD1	1.97	0.59
1:A:172:ASP:HB3	1:A:174:TYR:O	2.02	0.59
1:A:66:GLN:OE1	2:A:413:HOH:O	2.17	0.59
1:A:69:MET:HA	1:A:69:MET:CE	2.35	0.57
1:A:116:PHE:CB	1:A:118:LEU:HD12	2.34	0.56
1:A:24:ASN:OD1	1:A:26:GLN:HG3	2.05	0.55
1:A:171:MET:HG2	2:A:330:HOH:O	2.07	0.54
1:A:122:HIS:CD2	2:A:335:HOH:O	2.58	0.53
1:A:89:ARG:HG3	1:A:90:MET:N	2.24	0.52
1:A:155:LEU:HD12	2:A:386:HOH:O	2.09	0.51
1:A:134:TYR:CG	1:A:139:TYR:HA	2.46	0.51
1:A:43:MET:HG2	1:A:145:LEU:HB2	1.92	0.51
1:A:173:PHE:HA	2:A:419:HOH:O	2.11	0.51
1:A:120:GLU:OE1	1:A:122:HIS:HE1	1.93	0.50
1:A:72:ASN:HB2	2:A:279:HOH:O	2.11	0.50
1:A:173:PHE:HD2	2:A:439:HOH:O	1.94	0.49
1:A:91:ARG:HD3	1:A:102:MET:SD	2.53	0.49
1:A:58:ARG:NH1	2:A:417:HOH:O	2.45	0.49
1:A:96:ASP:OD2	1:A:153:ARG:NH2	2.47	0.48

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:173:PHE:C	1:A:174:TYR:HD1	2.16	0.48
1:A:174:TYR:CD1	1:A:174:TYR:N	2.82	0.47
1:A:99:ARG:HB2	1:A:99:ARG:HE	1.32	0.46
1:A:83:GLN:HG3	2:A:377:HOH:O	2.07	0.46
1:A:59:ARG:NH2	1:A:173:PHE:H	2.14	0.46
1:A:92:ILE:O	1:A:102:MET:HA	2.15	0.46
1:A:63:PRO:HD2	2:A:402:HOH:O	2.16	0.45
1:A:47:ARG:HD2	2:A:389:HOH:O	2.17	0.45
1:A:152:ARG:CG	2:A:383:HOH:O	2.60	0.43
1:A:173:PHE:CZ	2:A:441:HOH:O	2.41	0.43
1:A:59:ARG:HH12	1:A:173:PHE:CB	2.29	0.42
1:A:90:MET:HG2	1:A:131:TRP:CZ2	2.55	0.41
1:A:161:ASN:OD1	1:A:163:LYS:HB2	2.20	0.41
1:A:92:ILE:HG12	1:A:105:ILE:HD12	2.02	0.41
1:A:47:ARG:HG2	2:A:389:HOH:O	2.21	0.41
1:A:121:VAL:HB	1:A:164:VAL:HG21	2.02	0.41
1:A:64:ASP:OD1	1:A:66:GLN:HB2	2.21	0.41
1:A:59:ARG:HH22	1:A:173:PHE:H	1.68	0.40
1:A:59:ARG:CZ	1:A:173:PHE:HD1	2.33	0.40
1:A:95:ARG:HB2	1:A:99:ARG:HB3	2.02	0.40
1:A:115:ARG:NH2	2:A:423:HOH:O	2.50	0.40
1:A:133:LEU:HB2	1:A:146:LEU:HD11	2.02	0.40

All (4) 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:A:452:HOH:O	2:A:452:HOH:O[8_555]	1.61	0.59
2:A:255:HOH:O	2:A:255:HOH:O[7_556]	1.82	0.38
2:A:304:HOH:O	2:A:389:HOH:O[4_454]	1.87	0.33
2:A:303:HOH:O	2:A:375:HOH:O[4_454]	2.07	0.13

## 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	172/174 (99%)	167 (97%)	5 (3%)	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	158/158 (100%)	145 (92%)	13 (8%)	14	9

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	26	GLN
1	A	59	ARG
1	A	64	ASP
1	A	66	GLN
1	A	69	MET
1	A	83	GLN
1	A	99	ARG
1	A	103	SER
1	A	107	ASP
1	A	113	GLN
1	A	115	ARG
1	A	125	ASN
1	A	155	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	66	GLN
1	A	83	GLN
1	A	113	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	122	HIS
1	A	125	ASN
1	A	143	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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	174/174 (100%)	-0.31	2 (1%) 82 83	3, 10, 28, 41	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	173	PHE	5.3
1	A	174	TYR	2.9

### 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.