



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 26, 2020 – 10:19 pm BST

PDB ID : 6AMA  
Title : Structure of S. coelicolor/S. venezuelae BldC-smeA-ssfA complex to 3.09 Angstrom  
Authors : Schumacher, M.A.  
Deposited on : 2017-08-09  
Resolution : 3.09 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

<https://www.wwpdb.org/validation/2017/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)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

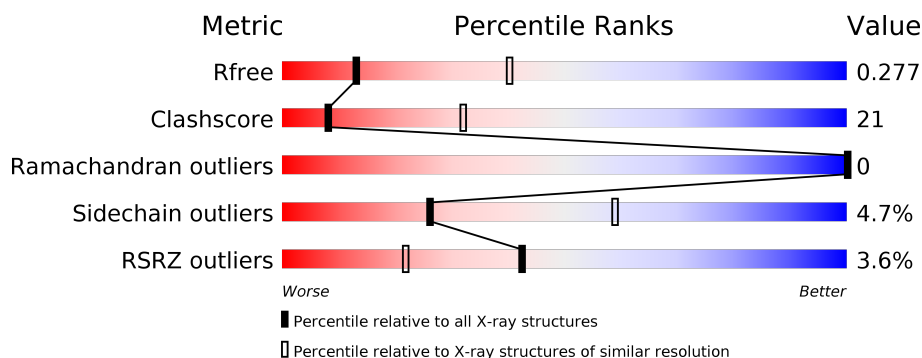
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.09 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))
$R_{free}$	130704	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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 respectively. 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	71	<div> <div>41%</div> <div>34%</div> <div>24%</div> </div>
1	B	71	<div> <div>4%</div> <div>51%</div> <div>25%</div> <div>23%</div> </div>
1	C	71	<div> <div>45%</div> <div>30%</div> <div>24%</div> </div>
1	D	71	<div> <div>51%</div> <div>24%</div> <div>24%</div> </div>
1	G	71	<div> <div>3%</div> <div>49%</div> <div>25%</div> <div>24%</div> </div>
1	H	71	<div> <div>4%</div> <div>45%</div> <div>28%</div> <div>23%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	K	71	<div><div><div></div><div></div><div></div><div></div></div><div>4%37%37%•24%</div></div>
1	L	71	<div><div><div></div><div></div><div></div><div></div></div><div>3%35%37%•24%</div></div>
1	O	71	<div><div><div></div><div></div><div></div><div></div></div><div>4%46%28%•23%</div></div>
1	P	71	<div><div><div></div><div></div><div></div><div></div></div><div>8%56%23%21%</div></div>
1	Y	71	<div><div><div></div><div></div><div></div><div></div></div><div>6%48%31%•20%</div></div>
2	N	99	<div><div><div></div><div></div><div></div><div></div></div><div>34%64%•</div></div>
3	R	99	<div><div><div></div><div></div><div></div><div></div></div><div>39%57%•</div></div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 8733 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 Putative DNA-binding protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	54	Total	C	N	O	S	0	0	0
			418	266	81	70	1			
1	B	55	Total	C	N	O	S	0	0	0
			427	271	82	73	1			
1	G	54	Total	C	N	O	S	0	0	0
			422	268	81	72	1			
1	H	55	Total	C	N	O	S	0	0	0
			431	273	83	74	1			
1	K	54	Total	C	N	O	S	0	0	0
			418	266	81	70	1			
1	L	54	Total	C	N	O	S	0	0	0
			418	266	81	70	1			
1	Y	57	Total	C	N	O	S	0	0	0
			442	280	84	77	1			
1	C	54	Total	C	N	O	S	0	0	0
			418	266	81	70	1			
1	D	54	Total	C	N	O	S	0	0	0
			418	266	81	70	1			
1	O	55	Total	C	N	O	S	0	0	0
			427	271	83	72	1			
1	P	56	Total	C	N	O	S	0	0	0
			435	275	83	76	1			

There are 33 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	expression tag	UNP A0A0M7QSG5
A	-1	SER	-	expression tag	UNP A0A0M7QSG5
A	0	HIS	-	expression tag	UNP A0A0M7QSG5
B	-2	GLY	-	expression tag	UNP A0A0M7QSG5
B	-1	SER	-	expression tag	UNP A0A0M7QSG5
B	0	HIS	-	expression tag	UNP A0A0M7QSG5
G	-2	GLY	-	expression tag	UNP A0A0M7QSG5

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
G	-1	SER	-	expression tag	UNP A0A0M7QSG5
G	0	HIS	-	expression tag	UNP A0A0M7QSG5
H	-2	GLY	-	expression tag	UNP A0A0M7QSG5
H	-1	SER	-	expression tag	UNP A0A0M7QSG5
H	0	HIS	-	expression tag	UNP A0A0M7QSG5
K	-2	GLY	-	expression tag	UNP A0A0M7QSG5
K	-1	SER	-	expression tag	UNP A0A0M7QSG5
K	0	HIS	-	expression tag	UNP A0A0M7QSG5
L	-2	GLY	-	expression tag	UNP A0A0M7QSG5
L	-1	SER	-	expression tag	UNP A0A0M7QSG5
L	0	HIS	-	expression tag	UNP A0A0M7QSG5
Y	-2	GLY	-	expression tag	UNP A0A0M7QSG5
Y	-1	SER	-	expression tag	UNP A0A0M7QSG5
Y	0	HIS	-	expression tag	UNP A0A0M7QSG5
C	-2	GLY	-	expression tag	UNP A0A0M7QSG5
C	-1	SER	-	expression tag	UNP A0A0M7QSG5
C	0	HIS	-	expression tag	UNP A0A0M7QSG5
D	-2	GLY	-	expression tag	UNP A0A0M7QSG5
D	-1	SER	-	expression tag	UNP A0A0M7QSG5
D	0	HIS	-	expression tag	UNP A0A0M7QSG5
O	-2	GLY	-	expression tag	UNP A0A0M7QSG5
O	-1	SER	-	expression tag	UNP A0A0M7QSG5
O	0	HIS	-	expression tag	UNP A0A0M7QSG5
P	-2	GLY	-	expression tag	UNP A0A0M7QSG5
P	-1	SER	-	expression tag	UNP A0A0M7QSG5
P	0	HIS	-	expression tag	UNP A0A0M7QSG5

- Molecule 2 is a DNA chain called DNA (99-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	N	99	Total	C	N	O	P	0	0	0
			2002	957	363	583	99			

- Molecule 3 is a DNA chain called DNA (99-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	R	99	Total	C	N	O	P	0	0	0
			2057	979	374	605	99			

### 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.

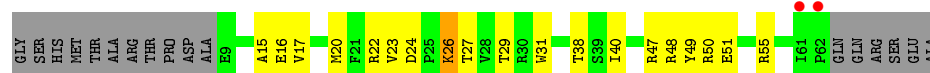
- Molecule 1: Putative DNA-binding protein



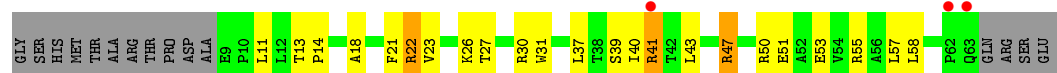
- Molecule 1: Putative DNA-binding protein



- Molecule 1: Putative DNA-binding protein



- Molecule 1: Putative DNA-binding protein



- Molecule 1: Putative DNA-binding protein



- Molecule 1: Putative DNA-binding protein

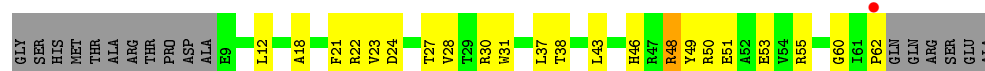
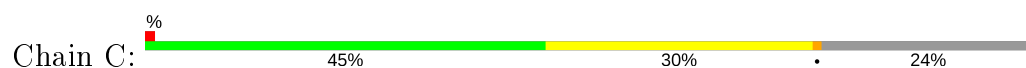


ALA

- Molecule 1: Putative DNA-binding protein



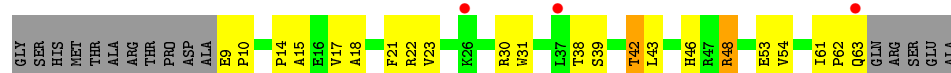
- Molecule 1: Putative DNA-binding protein



- Molecule 1: Putative DNA-binding protein



- Molecule 1: Putative DNA-binding protein

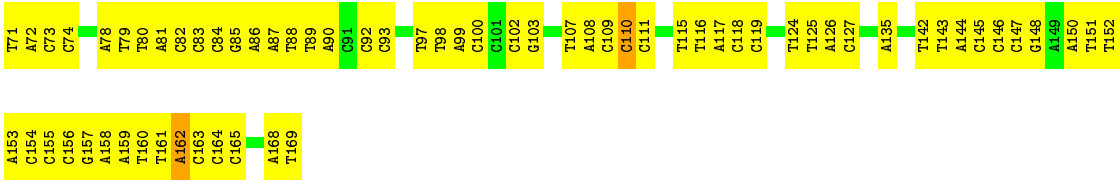


- Molecule 1: Putative DNA-binding protein

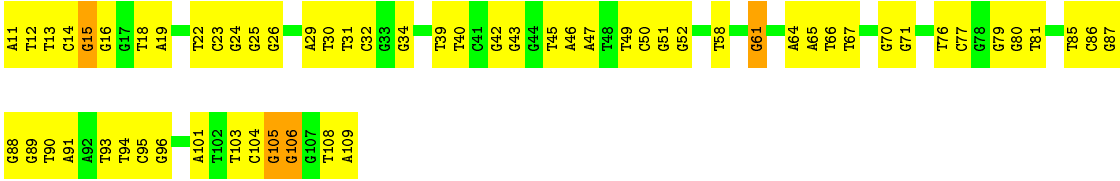


- Molecule 2: DNA (99-MER)





• Molecule 3: DNA (99-MER)





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	159.35Å 159.35Å 130.83Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	101.12 – 3.09 101.12 – 3.09	Depositor EDS
% Data completeness (in resolution range)	98.2 (101.12-3.09) 98.3 (101.12-3.09)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.12	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.04 (at 3.07Å)	Xtriage
Refinement program	PHENIX 1.6.4_486	Depositor
R, $R_{free}$	0.210 , 0.281 0.204 , 0.277	Depositor DCC
$R_{free}$ test set	1989 reflections (6.43%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	122.5	Xtriage
Anisotropy	0.037	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 94.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	8733	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	148.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.77% 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	0.23	0/426	0.45	0/577
1	B	0.22	0/435	0.43	0/589
1	C	0.23	0/426	0.44	0/577
1	D	0.24	0/426	0.45	0/577
1	G	0.23	0/430	0.45	0/582
1	H	0.23	0/439	0.44	0/594
1	K	0.23	0/426	0.43	0/577
1	L	0.23	0/426	0.48	0/577
1	O	0.24	0/435	0.43	0/589
1	P	0.23	0/443	0.44	0/600
1	Y	0.22	0/451	0.43	0/611
2	N	0.47	0/2243	1.09	3/3450 (0.1%)
3	R	0.49	0/2309	1.16	6/3571 (0.2%)
All	All	0.37	0/9315	0.87	9/13471 (0.1%)

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	R	106	DG	O4'-C1'-N9	6.00	112.20	108.00
2	N	162	DA	O4'-C1'-N9	-5.94	103.84	108.00
2	N	165	DC	C4'-C3'-C2'	-5.54	98.11	103.10
3	R	105	DG	O4'-C4'-C3'	-5.48	102.31	104.50
2	N	110	DC	O4'-C1'-N1	5.43	111.80	108.00

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	418	0	443	18	0
1	B	427	0	452	21	0
1	C	418	0	443	26	0
1	D	418	0	443	22	0
1	G	422	0	447	22	0
1	H	431	0	455	20	0
1	K	418	0	443	33	0
1	L	418	0	443	28	0
1	O	427	0	451	17	0
1	P	435	0	456	21	0
1	Y	442	0	464	23	0
2	N	2002	0	1112	90	0
3	R	2057	0	1123	92	0
All	All	8733	0	7175	335	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 21.

The worst 5 of 335 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)
3:R:12:DT:H2''	3:R:13:DT:H5''	1.47	0.95
2:N:93:DC:H42	3:R:87:DG:H1	1.10	0.95
3:R:29:DA:H2''	3:R:30:DT:H5''	1.47	0.94
1:C:38:THR:HG21	1:C:50:ARG:HE	1.35	0.89
2:N:90:DA:H5'	2:N:90:DA:C8	2.09	0.87

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	52/71 (73%)	45 (86%)	7 (14%)	0	100	100
1	B	53/71 (75%)	49 (92%)	4 (8%)	0	100	100
1	C	52/71 (73%)	48 (92%)	4 (8%)	0	100	100
1	D	52/71 (73%)	49 (94%)	3 (6%)	0	100	100
1	G	52/71 (73%)	45 (86%)	7 (14%)	0	100	100
1	H	53/71 (75%)	52 (98%)	1 (2%)	0	100	100
1	K	52/71 (73%)	46 (88%)	6 (12%)	0	100	100
1	L	52/71 (73%)	50 (96%)	2 (4%)	0	100	100
1	O	53/71 (75%)	48 (91%)	5 (9%)	0	100	100
1	P	54/71 (76%)	49 (91%)	5 (9%)	0	100	100
1	Y	55/71 (78%)	50 (91%)	5 (9%)	0	100	100
All	All	580/781 (74%)	531 (92%)	49 (8%)	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	42/56 (75%)	40 (95%)	2 (5%)	25	58
1	B	43/56 (77%)	42 (98%)	1 (2%)	50	77
1	C	42/56 (75%)	41 (98%)	1 (2%)	49	76
1	D	42/56 (75%)	41 (98%)	1 (2%)	49	76
1	G	43/56 (77%)	41 (95%)	2 (5%)	26	59
1	H	44/56 (79%)	41 (93%)	3 (7%)	16	45
1	K	42/56 (75%)	40 (95%)	2 (5%)	25	58
1	L	42/56 (75%)	38 (90%)	4 (10%)	8	31
1	O	43/56 (77%)	39 (91%)	4 (9%)	9	32
1	P	44/56 (79%)	43 (98%)	1 (2%)	50	77

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	Y	45/56 (80%)	44 (98%)	1 (2%)	52 78
All	All	472/616 (77%)	450 (95%)	22 (5%)	26 59

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	K	53	GLU
1	L	23	VAL
1	O	48	ARG
1	L	19	THR
1	L	22	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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	54/71 (76%)	0.42	1 (1%) 66 46	98, 131, 151, 193	0
1	B	55/71 (77%)	0.46	3 (5%) 25 11	90, 124, 174, 187	0
1	C	54/71 (76%)	0.25	1 (1%) 66 46	91, 122, 148, 207	0
1	D	54/71 (76%)	0.38	1 (1%) 66 46	92, 121, 137, 180	0
1	G	54/71 (76%)	0.25	2 (3%) 41 21	84, 121, 155, 211	0
1	H	55/71 (77%)	0.58	3 (5%) 25 11	97, 118, 146, 169	0
1	K	54/71 (76%)	0.40	3 (5%) 24 11	100, 137, 178, 203	0
1	L	54/71 (76%)	0.26	2 (3%) 41 21	85, 111, 132, 149	0
1	O	55/71 (77%)	0.38	3 (5%) 25 11	90, 113, 151, 178	0
1	P	56/71 (78%)	0.37	6 (10%) 6 2	98, 130, 178, 192	0
1	Y	57/71 (80%)	0.55	4 (7%) 16 7	94, 129, 193, 203	0
2	N	99/99 (100%)	-0.58	0 100 100	113, 164, 235, 298	0
3	R	99/99 (100%)	-0.55	0 100 100	112, 168, 249, 265	0
All	All	800/979 (81%)	0.16	29 (3%) 42 22	84, 131, 207, 298	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	63	GLN	5.8
1	B	61	ILE	5.0
1	Y	62	PRO	4.9
1	Y	61	ILE	4.8
1	G	62	PRO	3.8

### 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.