



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 17, 2024 – 01:08 AM EDT

PDB ID : 5JGH  
Title : Crystal structure of the mitochondrial DNA packaging protein Abf2p in complex with DNA at 2.6 Angstrom resolution  
Authors : Chakraborty, A.; Lyonnais, S.; Sola, M.  
Deposited on : 2016-04-20  
Resolution : 2.60 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
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)	:	2.37.1

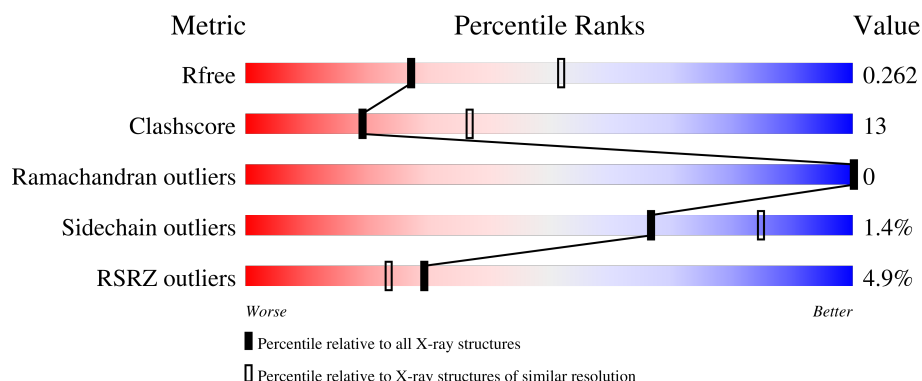
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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.





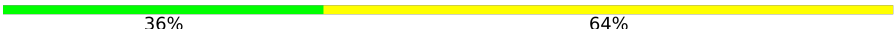




Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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 of 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	163	<div> <div>4%</div> <div>69%</div> <div>26%</div> <div>.</div> </div>
1	D	163	<div> <div>8%</div> <div>68%</div> <div>26%</div> <div>5%</div> </div>
1	G	163	<div> <div>7%</div> <div>67%</div> <div>28%</div> <div>.</div> </div>
1	J	163	<div> <div>6%</div> <div>65%</div> <div>27%</div> <div>7%</div> </div>
2	B	22	<div> <div>41%</div> <div>59%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	E	22	 59% 41%
2	H	22	 50% 50%
2	K	22	 36% 64%
3	C	22	 59% 41%
3	F	22	 59% 41%
3	I	22	 64% 36%
3	L	22	 50% 50%

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 8906 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 ARS-binding factor 2, mitochondrial.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	156	Total	C	N	O	S	0	0	0
			1309	840	229	239	1			
1	D	155	Total	C	N	O	S	0	0	0
			1301	834	228	238	1			
1	G	156	Total	C	N	O	S	0	0	0
			1306	837	229	239	1			
1	J	151	Total	C	N	O	S	0	0	0
			1272	816	222	233	1			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	21	HIS	-	expression tag	UNP Q02486
A	22	HIS	-	expression tag	UNP Q02486
A	23	HIS	-	expression tag	UNP Q02486
A	24	HIS	-	expression tag	UNP Q02486
A	25	HIS	-	expression tag	UNP Q02486
A	26	HIS	-	expression tag	UNP Q02486
D	21	HIS	-	expression tag	UNP Q02486
D	22	HIS	-	expression tag	UNP Q02486
D	23	HIS	-	expression tag	UNP Q02486
D	24	HIS	-	expression tag	UNP Q02486
D	25	HIS	-	expression tag	UNP Q02486
D	26	HIS	-	expression tag	UNP Q02486
G	21	HIS	-	expression tag	UNP Q02486
G	22	HIS	-	expression tag	UNP Q02486
G	23	HIS	-	expression tag	UNP Q02486
G	24	HIS	-	expression tag	UNP Q02486
G	25	HIS	-	expression tag	UNP Q02486
G	26	HIS	-	expression tag	UNP Q02486
J	21	HIS	-	expression tag	UNP Q02486
J	22	HIS	-	expression tag	UNP Q02486
J	23	HIS	-	expression tag	UNP Q02486

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
J	24	HIS	-	expression tag	UNP Q02486
J	25	HIS	-	expression tag	UNP Q02486
J	26	HIS	-	expression tag	UNP Q02486

- Molecule 2 is a DNA chain called DNA (5'-D(\*TP\*TP\*TP\*AP\*TP\*TP\*AP\*TP\*TP\*TP\*TP\*AP\*TP\*AP\*TP\*AP\*TP\*TP\*AP\*TP\*AP\*TP\*AP\*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	22	Total	C	N	O	P	0	0	0
			445	220	68	136	21			
2	E	22	Total	C	N	O	P	0	0	0
			445	220	68	136	21			
2	H	22	Total	C	N	O	P	0	0	0
			445	220	68	136	21			
2	K	22	Total	C	N	O	P	0	0	0
			445	220	68	136	21			

- Molecule 3 is a DNA chain called DNA (5'-D(\*TP\*TP\*AP\*TP\*AP\*TP\*AP\*AP\*TP\*AP\*TP\*AP\*AP\*AP\*TP\*AP\*AP\*TP\*AP\*AP\*TP\*AP\*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	22	Total	C	N	O	P	0	0	0
			451	220	86	124	21			
3	F	22	Total	C	N	O	P	0	0	0
			451	220	86	124	21			
3	I	22	Total	C	N	O	P	0	0	0
			451	220	86	124	21			
3	L	22	Total	C	N	O	P	0	0	0
			451	220	86	124	21			

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula: C<sub>2</sub>H<sub>3</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	G	1	Total	C	O	0	0
			4	2	2		
4	G	1	Total	C	O	0	0
			4	2	2		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	23	Total	O	0	0
			23	23		
5	B	4	Total	O	0	0
			4	4		
5	C	5	Total	O	0	0
			5	5		
5	D	10	Total	O	0	0
			10	10		
5	E	8	Total	O	0	0
			8	8		
5	F	3	Total	O	0	0
			3	3		
5	G	20	Total	O	0	0
			20	20		
5	H	11	Total	O	0	0
			11	11		
5	I	11	Total	O	0	0
			11	11		
5	J	18	Total	O	0	0
			18	18		

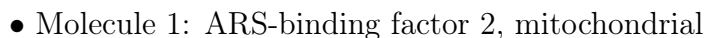
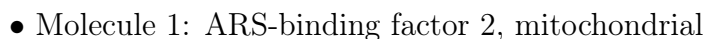
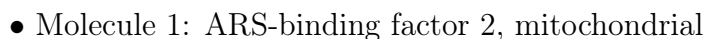
*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	K	9	Total	O	0	0
			9	9		
5	L	4	Total	O	0	0
			4	4		



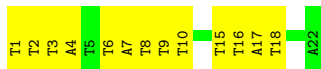
- Molecule 1: ARS-binding factor 2, mitochondrial







- Molecule 2: DNA (5'-D(\*TP\*TP\*TP\*AP\*TP\*TP\*AP\*TP\*TP\*TP\*TP\*AP\*TP\*AP\*TP\*TP\*AP\*TP\*AP\*TP\*AP\*A)-3')



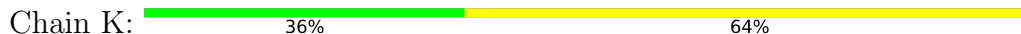
- Molecule 2: DNA (5'-D(\*TP\*TP\*TP\*AP\*TP\*TP\*AP\*TP\*TP\*TP\*TP\*AP\*TP\*AP\*TP\*TP\*AP\*TP\*AP\*TP\*AP\*A)-3')



- Molecule 2: DNA (5'-D(\*TP\*TP\*TP\*AP\*TP\*TP\*AP\*TP\*TP\*TP\*TP\*AP\*TP\*AP\*TP\*TP\*AP\*TP\*AP\*TP\*AP\*A)-3')



- Molecule 2: DNA (5'-D(\*TP\*TP\*TP\*AP\*TP\*TP\*AP\*TP\*TP\*TP\*TP\*AP\*TP\*AP\*TP\*TP\*AP\*TP\*AP\*TP\*AP\*A)-3')



- Molecule 3: DNA (5'-D(\*TP\*TP\*AP\*TP\*AP\*TP\*AP\*AP\*TP\*AP\*TP\*AP\*AP\*AP\*AP\*TP\*AP\*AP\*TP\*AP\*AP\*A)-3')

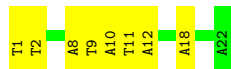


- Molecule 3: DNA (5'-D(\*TP\*TP\*AP\*TP\*AP\*TP\*AP\*AP\*TP\*AP\*TP\*AP\*AP\*AP\*AP\*TP\*AP\*AP\*TP\*AP\*AP\*A)-3')





- Molecule 3: DNA (5'-D(\*TP\*TP\*AP\*TP\*AP\*TP\*AP\*AP\*TP\*AP\*TP\*AP\*AP\*AP\*AP\*TP\*AP\*AP\*TP\*AP\*AP\*A)-3')



- Molecule 3: DNA (5'-D(\*TP\*TP\*AP\*TP\*AP\*TP\*AP\*AP\*TP\*AP\*TP\*AP\*AP\*AP\*AP\*TP\*AP\*AP\*TP\*AP\*AP\*A)-3')



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	71.99Å 131.99Å 71.99Å 90.00° 103.21° 90.00°	Depositor
Resolution (Å)	48.05 – 2.60 48.05 – 2.50	Depositor EDS
% Data completeness (in resolution range)	95.3 (48.05-2.60) 95.3 (48.05-2.50)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.28 (at 2.51Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, $R_{free}$	0.227 , 0.260 0.228 , 0.262	Depositor DCC
$R_{free}$ test set	2056 reflections (4.78%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	50.4	Xtriage
Anisotropy	0.289	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 32.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.207 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	8906	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 19.16% 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

Bond lengths and bond angles in the following residue types are not validated in this section: ACT

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	0.39	0/1339	0.50	0/1795
1	D	0.40	0/1331	0.52	0/1784
1	G	0.37	0/1336	0.51	0/1791
1	J	0.41	0/1302	0.51	0/1747
2	B	0.79	0/496	1.13	0/764
2	E	0.86	0/496	1.15	0/764
2	H	0.89	0/496	1.14	0/764
2	K	0.81	0/496	1.13	0/764
3	C	0.72	0/508	1.03	0/782
3	F	0.81	0/508	1.08	0/782
3	I	0.84	0/508	1.03	0/782
3	L	0.79	0/508	0.97	0/782
All	All	0.61	0/9324	0.83	0/13301

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	1309	0	1338	35	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	1301	0	1327	34	0
1	G	1306	0	1329	42	0
1	J	1272	0	1291	34	0
2	B	445	0	258	17	0
2	E	445	0	258	8	0
2	H	445	0	258	9	0
2	K	445	0	258	16	0
3	C	451	0	252	7	0
3	F	451	0	252	10	0
3	I	451	0	252	8	0
3	L	451	0	252	11	0
4	G	8	0	6	0	0
5	A	23	0	0	2	0
5	B	4	0	0	0	0
5	C	5	0	0	0	0
5	D	10	0	0	0	0
5	E	8	0	0	0	0
5	F	3	0	0	0	0
5	G	20	0	0	2	0
5	H	11	0	0	0	0
5	I	11	0	0	0	0
5	J	18	0	0	0	0
5	K	9	0	0	0	0
5	L	4	0	0	0	0
All	All	8906	0	7331	202	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 13.

The worst 5 of 202 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:29:SER:HB2	2:B:8:DT:H4'	1.30	1.05
1:A:31:ARG:NH2	2:B:8:DT:OP1	2.08	0.86
1:G:65:ASN:HB3	1:G:68:LEU:HD22	1.61	0.81
1:D:90:GLU:HA	1:D:93:ILE:HD12	1.69	0.75
1:G:74:SER:OG	2:K:14:DA:N3	2.20	0.74

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	154/163 (94%)	154 (100%)	0	0	100	100
1	D	153/163 (94%)	148 (97%)	5 (3%)	0	100	100
1	G	154/163 (94%)	151 (98%)	3 (2%)	0	100	100
1	J	149/163 (91%)	147 (99%)	2 (1%)	0	100	100
All	All	610/652 (94%)	600 (98%)	10 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains [i](#)

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	141/148 (95%)	140 (99%)	1 (1%)	84	94
1	D	140/148 (95%)	138 (99%)	2 (1%)	67	85
1	G	140/148 (95%)	139 (99%)	1 (1%)	84	94
1	J	137/148 (93%)	133 (97%)	4 (3%)	42	68
All	All	558/592 (94%)	550 (99%)	8 (1%)	67	85

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	J	179	ARG
1	J	143	GLN
1	J	51	PHE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	G	29	SER
1	J	57	HIS

Sometimes 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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
4	ACT	G	202	-	3,3,3	0.78	0	3,3,3	0.73	0
4	ACT	G	201	-	3,3,3	0.82	0	3,3,3	0.65	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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.



## 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	156/163 (95%)	0.10	6 (3%) 40 33	38, 60, 81, 92	1 (0%)
1	D	155/163 (95%)	0.45	13 (8%) 11 7	36, 70, 114, 135	0
1	G	156/163 (95%)	0.41	11 (7%) 16 11	34, 70, 104, 115	0
1	J	151/163 (92%)	0.09	9 (5%) 21 16	37, 58, 88, 105	2 (1%)
2	B	22/22 (100%)	-0.69	0 100 100	43, 62, 72, 79	0
2	E	22/22 (100%)	-0.73	0 100 100	34, 48, 52, 53	0
2	H	22/22 (100%)	-0.77	0 100 100	36, 46, 51, 55	0
2	K	22/22 (100%)	-0.58	0 100 100	41, 60, 77, 83	0
3	C	22/22 (100%)	-0.74	0 100 100	49, 65, 74, 75	0
3	F	22/22 (100%)	-0.85	0 100 100	42, 55, 60, 63	0
3	I	22/22 (100%)	-0.81	0 100 100	42, 54, 60, 64	0
3	L	22/22 (100%)	-0.66	0 100 100	50, 60, 80, 82	0
All	All	794/828 (95%)	0.04	39 (4%) 29 23	34, 60, 95, 135	3 (0%)

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	67	THR	8.9
1	D	61	PHE	5.7
1	J	33	GLN	4.9
1	G	67	THR	4.8
1	D	69	ARG	4.3

### 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 monosaccharides in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	ACT	G	202	4/4	0.64	0.31	74,91,92,92	0
4	ACT	G	201	4/4	0.93	0.18	44,53,55,61	1

### 6.5 Other polymers [i](#)

There are no such residues in this entry.