wwPDB X-ray Structure Validation Summary Report (i)

Oct 20, 2021 – 08:54 am BST

PDB ID	:	7OZQ
Title	:	Crystal structure of archaeal L7Ae bound to eukaryotic kink-loop
Authors	:	Hoefler, S.; Lukat, P.; Carlomagno, T.; Blankenfeldt, W.
Deposited on	:	2021-06-28
Resolution	:	1.91 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

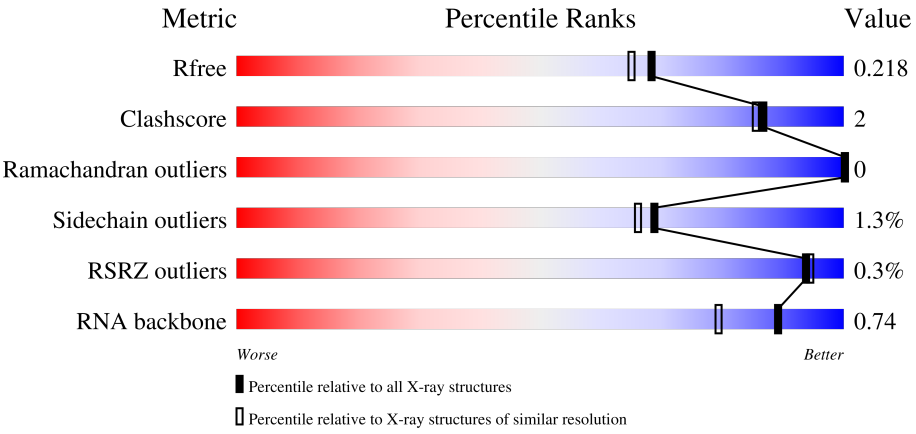
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.91 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.



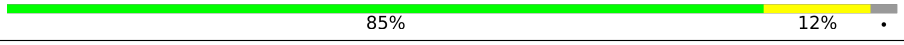
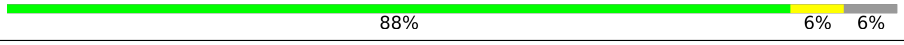
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	7937 (1.94-1.90)
Clashscore	141614	8644 (1.94-1.90)
Ramachandran outliers	138981	8530 (1.94-1.90)
Sidechain outliers	138945	8530 (1.94-1.90)
RSRZ outliers	127900	7793 (1.94-1.90)
RNA backbone	3102	1010 (2.44-1.40)

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 >=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 <=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	125	96%
1	B	125	92% 6%
1	C	125	98%
1	D	125	88% 9%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	E	33	
2	F	33	
2	G	33	
2	H	33	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	ACT	E	101	-	-	X	-

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 12412 atoms, of which 5300 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 50S ribosomal protein L7Ae.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	123	Total	C	H	N	O	S	0	3	0
			1931	603	988	153	184	3			
1	B	122	Total	C	H	N	O	S	0	2	0
			1904	596	972	150	183	3			
1	C	123	Total	C	H	N	O	S	0	3	0
			1914	600	974	155	181	4			
1	D	121	Total	C	H	N	O	S	0	3	0
			1933	601	991	155	183	3			

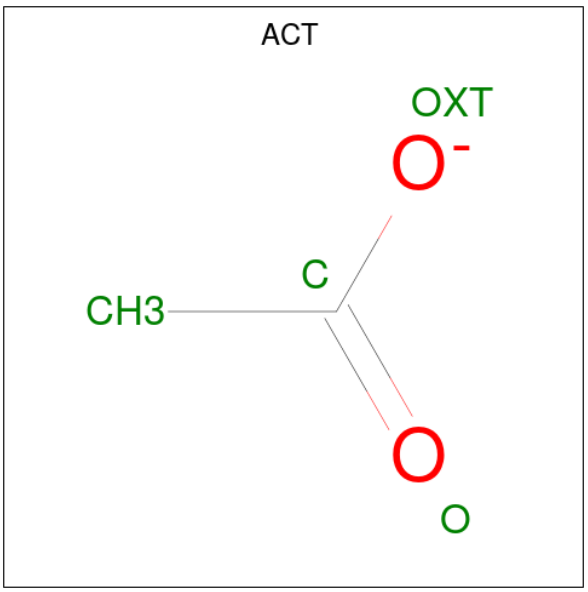
There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP Q8U160
A	0	ALA	-	expression tag	UNP Q8U160
B	-1	GLY	-	expression tag	UNP Q8U160
B	0	ALA	-	expression tag	UNP Q8U160
C	-1	GLY	-	expression tag	UNP Q8U160
C	0	ALA	-	expression tag	UNP Q8U160
D	-1	GLY	-	expression tag	UNP Q8U160
D	0	ALA	-	expression tag	UNP Q8U160

- Molecule 2 is a RNA chain called RNA.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	E	32	Total	C	H	N	O	P	0	0	0
			1030	307	343	125	223	32			
2	F	31	Total	C	H	N	O	P	0	0	0
			999	298	332	122	216	31			
2	G	32	Total	C	H	N	O	P	0	0	0
			1029	307	342	125	223	32			
2	H	32	Total	C	H	N	O	P	0	0	0
			1030	307	343	125	223	32			

- Molecule 3 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total C H O 7 2 3 2	0	0
3	E	1	Total C H O 7 2 3 2	0	0
3	F	1	Total C H O 7 2 3 2	0	0
3	G	1	Total C H O 7 2 3 2	0	0
3	H	1	Total C H O 7 2 3 2	0	0

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	D	1	Total Na 1 1	0	0
4	F	4	Total Na 4 4	0	0
4	G	1	Total Na 1 1	0	0

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	E	2	Total Ca 2 2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	F	1	Total Ca 1 1	0	0

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	H	1	Total Cl 1 1	0	0

- Molecule 7 is water.

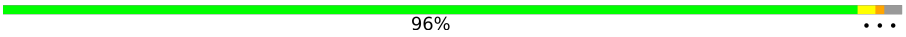
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	77	Total O 77 77	0	0
7	B	62	Total O 62 62	0	0
7	C	84	Total O 84 84	0	0
7	D	56	Total O 56 56	0	0
7	E	88	Total O 88 88	0	0
7	F	73	Total O 73 73	0	0
7	G	75	Total O 75 75	0	0
7	H	82	Total O 82 82	0	0

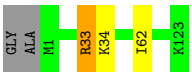
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: 50S ribosomal protein L7Ae

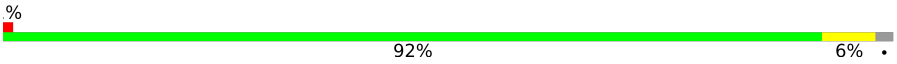
Chain A:

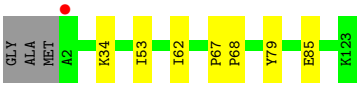




• Molecule 1: 50S ribosomal protein L7Ae

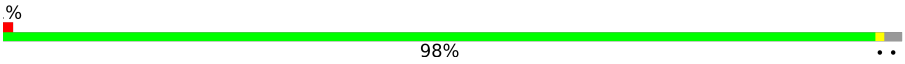
Chain B:

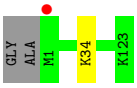




• Molecule 1: 50S ribosomal protein L7Ae

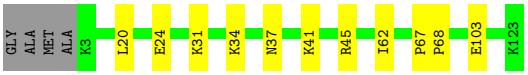
Chain C:





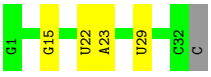
• Molecule 1: 50S ribosomal protein L7Ae

Chain D:



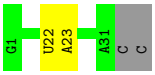
• Molecule 2: RNA

Chain E:



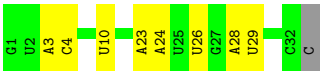
• Molecule 2: RNA

Chain F:



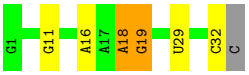
• Molecule 2: RNA

Chain G:



• Molecule 2: RNA

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	109.39Å 61.60Å 138.61Å 90.00° 108.36° 90.00°	Depositor
Resolution (Å)	54.50 – 1.91 54.50 – 1.91	Depositor EDS
% Data completeness (in resolution range)	99.9 (54.50-1.91) 95.7 (54.50-1.91)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.87 (at 1.91Å)	Xtriage
Refinement program	PHENIX 1.18rc4_3812	Depositor
R, R_{free}	0.173 , 0.218 0.173 , 0.218	Depositor DCC
R_{free} test set	3281 reflections (4.79%)	wwPDB-VP
Wilson B-factor (Å^2)	29.6	Xtriage
Anisotropy	0.462	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.009 for 1/2*h-3/2*k,-1/2*h-1/2*k,-1/2*h+1/2*k-l 0.016 for 1/2*h+3/2*k,1/2*h-1/2*k,-1/2*h-1/2*k-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	12412	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.89% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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, NA, CL, CA

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	0.54	0/964	0.56	0/1301
1	B	0.47	0/950	0.52	0/1284
1	C	0.51	0/962	0.58	0/1302
1	D	0.48	0/963	0.53	0/1299
2	E	0.75	0/769	0.89	0/1197
2	F	0.79	0/747	0.90	0/1163
2	G	0.78	0/769	0.90	0/1197
2	H	0.73	0/769	0.88	0/1197
All	All	0.63	0/6893	0.73	0/9940

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	943	988	990	3	0
1	B	932	972	972	4	0
1	C	940	974	976	0	0
1	D	942	991	991	7	0
2	E	687	343	344	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	F	667	332	333	1	0
2	G	687	342	343	6	1
2	H	687	343	344	5	1
3	B	4	3	3	0	0
3	E	4	3	3	2	0
3	F	4	3	3	0	0
3	G	4	3	3	1	0
3	H	4	3	3	1	0
4	D	1	0	0	0	0
4	F	4	0	0	0	0
4	G	1	0	0	0	0
5	E	2	0	0	0	0
5	F	1	0	0	0	0
6	H	1	0	0	1	0
7	A	77	0	0	1	0
7	B	62	0	0	1	0
7	C	84	0	0	0	1
7	D	56	0	0	0	0
7	E	88	0	0	1	0
7	F	73	0	0	0	1
7	G	75	0	0	1	0
7	H	82	0	0	0	0
All	All	7112	5300	5308	25	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 2.

The worst 5 of 25 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)
2:H:11:G:N7	6:H:102:CL:CL	2.75	0.56
2:G:26:U:O4	3:G:101:ACT:C	2.58	0.52
2:G:28:A:N7	7:G:204:HOH:O	2.33	0.51
2:G:3:A:H2'	2:G:4:C:O4'	2.11	0.51
1:B:67:PRO:HB2	1:B:68:PRO:HD3	1.95	0.48

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)
7:C:231:HOH:O	7:F:261:HOH:O[3_445]	2.05	0.15
2:G:10:U:OP1	2:H:16:A:O2'[3_555]	2.07	0.13

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	124/125 (99%)	124 (100%)	0	0	100	100
1	B	122/125 (98%)	122 (100%)	0	0	100	100
1	C	124/125 (99%)	123 (99%)	1 (1%)	0	100	100
1	D	122/125 (98%)	122 (100%)	0	0	100	100
All	All	492/500 (98%)	491 (100%)	1 (0%)	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	99/99 (100%)	97 (98%)	2 (2%)	55	49
1	B	98/99 (99%)	97 (99%)	1 (1%)	76	75
1	C	98/99 (99%)	97 (99%)	1 (1%)	76	75
1	D	100/99 (101%)	99 (99%)	1 (1%)	76	75
All	All	395/396 (100%)	390 (99%)	5 (1%)	69	66

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	33	ARG
1	A	34	LYS
1	B	34	LYS
1	C	34	LYS
1	D	34	LYS

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

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	E	31/33 (93%)	0	0
2	F	30/33 (90%)	0	0
2	G	31/33 (93%)	0	0
2	H	31/33 (93%)	2 (6%)	0
All	All	123/132 (93%)	2 (1%)	0

All (2) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	H	18	A
2	H	19	G

There are no RNA pucker outliers to report.

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)

Of 15 ligands modelled in this entry, 10 are monoatomic - leaving 5 for Mogul analysis.

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	$\#\|Z\| > 2$
3	ACT	B	201	-	1,3,3	5.91	1 (100%)	0,3,3	-	-
3	ACT	F	101	4	1,3,3	5.84	1 (100%)	0,3,3	-	-
3	ACT	G	101	-	1,3,3	4.78	1 (100%)	0,3,3	-	-
3	ACT	E	101	5	1,3,3	8.20	1 (100%)	0,3,3	-	-
3	ACT	H	101	-	1,3,3	6.01	1 (100%)	0,3,3	-	-

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	101	ACT	CH3-C	8.20	1.59	1.48
3	H	101	ACT	CH3-C	6.01	1.56	1.48
3	B	201	ACT	CH3-C	5.91	1.56	1.48
3	F	101	ACT	CH3-C	5.84	1.56	1.48
3	G	101	ACT	CH3-C	4.78	1.54	1.48

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	G	101	ACT	1	0
3	E	101	ACT	2	0
3	H	101	ACT	1	0

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^{th} 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^2)	Q<0.9
1	A	123/125 (98%)	-0.30	0 100 100	19, 29, 49, 76	0
1	B	122/125 (97%)	-0.13	1 (0%) 86 87	24, 33, 60, 105	0
1	C	123/125 (98%)	-0.30	1 (0%) 86 87	21, 30, 45, 97	0
1	D	121/125 (96%)	-0.24	0 100 100	24, 35, 55, 71	0
2	E	32/33 (96%)	-0.52	0 100 100	28, 35, 74, 77	0
2	F	31/33 (93%)	-0.62	0 100 100	30, 36, 58, 63	0
2	G	32/33 (96%)	-0.43	0 100 100	28, 44, 77, 78	0
2	H	32/33 (96%)	-0.56	0 100 100	27, 40, 58, 87	0
All	All	616/632 (97%)	-0.30	2 (0%) 94 94	19, 34, 60, 105	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	2	ALA	6.3
1	C	1	MET	3.6

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

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^{th} 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(Å^2)	Q<0.9
3	ACT	E	101	4/4	0.61	0.20	36,44,44,47	0
3	ACT	B	201	4/4	0.65	0.14	50,59,60,65	0
6	CL	H	102	1/1	0.70	0.34	77,77,77,77	0
3	ACT	G	101	4/4	0.76	0.11	45,54,64,70	0
4	NA	D	201	1/1	0.78	0.27	60,60,60,60	0
5	CA	E	103	1/1	0.85	0.29	77,77,77,77	0
3	ACT	H	101	4/4	0.88	0.10	42,57,69,69	0
4	NA	F	103	1/1	0.90	0.06	45,45,45,45	0
4	NA	F	105	1/1	0.90	0.12	57,57,57,57	0
3	ACT	F	101	4/4	0.90	0.10	38,56,57,57	0
4	NA	F	102	1/1	0.90	0.08	50,50,50,50	0
5	CA	E	102	1/1	0.96	0.12	41,41,41,41	1
4	NA	G	102	1/1	0.97	0.09	40,40,40,40	0
5	CA	F	106	1/1	0.97	0.21	45,45,45,45	1
4	NA	F	104	1/1	0.97	0.21	50,50,50,50	0

6.5 Other polymers (i)

There are no such residues in this entry.