Full wwPDB X-ray Structure Validation Report ⓘ

Mar 6, 2025 – 04:06 PM EST

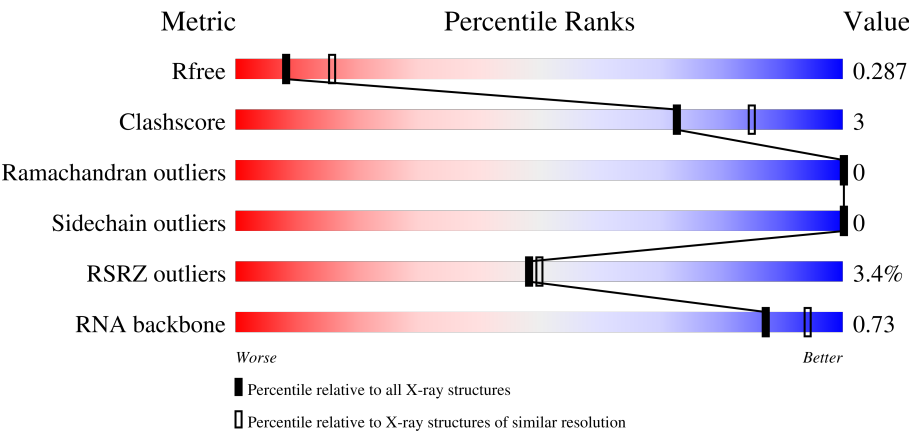|		
PDB ID	:	9DE5
Title	:	Structure of full-length HIV TAR RNA bound to HIV Tat RNA-binding domain
Authors	:	Bou-Nader, C.; Zhang, J.
Deposited on	:	2024-08-28
Resolution	:	2.75 Å(reported)

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 2.75 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.
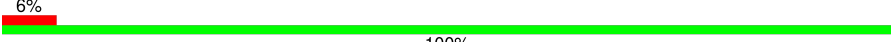
Metric	Percentile Ranks	Value
Rfree		0.287
Clashscore		3
Ramachandran outliers		0
Sidechain outliers		0
RSRZ outliers		3.4%
RNA backbone		0.73

Worse ← → *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	1606 (2.78-2.74)
Clashscore	180529	1689 (2.78-2.74)
Ramachandran outliers	177936	1665 (2.78-2.74)
Sidechain outliers	177891	1665 (2.78-2.74)
RSRZ outliers	164620	1606 (2.78-2.74)
RNA backbone	3690	1094 (3.00-2.52)

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	B	57	2% / 70% / 30%
1	C	57	65% / 30% / · ·
1	D	57	4% / 75% / 19% / 5%
1	E	57	7% / 68% / 30% / ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	A	17	

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 4902 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 RNA chain called RNA (55-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	55	Total	C	N	O	P	0	0	0
			1166	515	209	385	57			
1	E	57	Total	C	N	O	P	0	0	0
			1206	534	211	402	59			
1	B	57	Total	C	N	O	P	0	0	0
			1222	542	216	405	59			
1	D	54	Total	C	N	O	P	0	0	0
			1162	515	209	382	56			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	31	G	U	engineered mutation	GB 1945658931
C	32	A	G	engineered mutation	GB 1945658931
E	31	G	U	engineered mutation	GB 1945658931
E	32	A	G	engineered mutation	GB 1945658931
B	31	G	U	engineered mutation	GB 1945658931
B	32	A	G	engineered mutation	GB 1945658931
D	31	G	U	engineered mutation	GB 1945658931
D	32	A	G	engineered mutation	GB 1945658931

- Molecule 2 is a protein called Protein Tat.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	A	17	Total	C	N	O	0	0	0
			134	78	36	20			

- Molecule 3 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	1	Total	K	0	0
			1	1		

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	E	1	Total 1	K 1	0	0
3	D	1	Total 1	K 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	C	4	Total 4	O 4	0	0
4	E	2	Total 2	O 2	0	0
4	B	2	Total 2	O 2	0	0
4	D	1	Total 1	O 1	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.
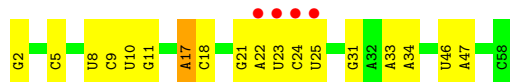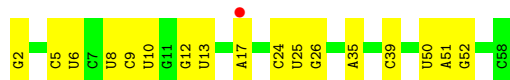
• Molecule 1: RNA (55-MER)

Chain C:
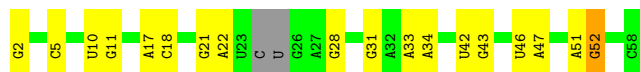
• Molecule 1: RNA (55-MER)

Chain E:

• Molecule 1: RNA (55-MER)

Chain B:

• Molecule 1: RNA (55-MER)

Chain D:

• Molecule 2: Protein Tat

Chain A:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	68.87Å 40.52Å 120.41Å 90.00° 94.61° 90.00°	Depositor
Resolution (Å)	47.09 – 2.75 47.09 – 2.75	Depositor EDS
% Data completeness (in resolution range)	97.8 (47.09-2.75) 97.9 (47.09-2.75)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.43 (at 2.77Å)	Xtriage
Refinement program	PHENIX 1.21.1_5286	Depositor
R, R_{free}	0.248 , 0.288 0.248 , 0.287	Depositor DCC
R_{free} test set	1024 reflections (4.78%)	wwPDB-VP
Wilson B-factor (Å^2)	45.3	Xtriage
Anisotropy	0.285	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 40.1	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	4902	wwPDB-VP
Average B, all atoms (Å^2)	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.55% 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 (i)

5.1 Standard geometry (i)

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

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	B	0.22	0/1329	0.70	0/2068
1	C	0.20	0/1266	0.67	0/1969
1	D	0.17	0/1262	0.66	0/1962
1	E	0.20	0/1311	0.68	0/2040
2	A	0.26	0/135	0.71	0/176
All	All	0.20	0/5303	0.68	0/8215

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 (i)

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	B	1222	0	615	5	0
1	C	1166	0	584	8	0
1	D	1162	0	585	2	0
1	E	1206	0	606	7	0
2	A	134	0	125	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	2	0	0	0	0
4	C	4	0	0	0	0
4	D	1	0	0	0	0
4	E	2	0	0	0	0
All	All	4902	0	2515	20	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 3.

All (20) 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:E:17:A:H8	1:B:52:G:H21	1.37	0.72
1:C:31:G:N2	1:C:34:A:OP2	2.26	0.62
1:C:46:U:H2'	1:C:47:A:H8	1.75	0.51
1:C:42:U:H2'	1:C:43:G:C8	2.46	0.50
1:E:17:A:O2'	1:B:10:U:O4'	2.26	0.49
1:B:8:U:H2'	1:B:9:C:C6	2.47	0.49
1:C:46:U:H2'	1:C:47:A:C8	2.47	0.48
1:C:10:U:H2'	1:C:11:G:H8	1.79	0.48
1:E:8:U:H2'	1:E:9:C:C6	2.49	0.47
1:C:51:A:H2'	1:C:52:G:O4'	2.14	0.47
1:B:50:U:H2'	1:B:51:A:C8	2.50	0.47
1:D:8:U:H2'	1:D:9:C:C6	2.51	0.46
1:C:10:U:H2'	1:C:11:G:C8	2.52	0.45
1:E:46:U:H2'	1:E:47:A:C8	2.52	0.44
1:E:33:A:H2'	1:E:34:A:C8	2.54	0.43
1:B:12:G:H2'	1:B:13:U:C6	2.55	0.42
1:E:31:G:N2	1:E:34:A:OP2	2.53	0.41
1:D:38:U:H2'	1:D:39:C:C6	2.56	0.41
1:C:33:A:H2'	1:C:34:A:C8	2.57	0.40
1:E:10:U:H2'	1:E:11:G:C8	2.56	0.40

There are no symmetry-related clashes.

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	
2	A	15/17 (88%)	15 (100%)	0	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	
2	A	10/14 (71%)	10 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	B	55/57 (96%)	8 (14%)	0
1	C	51/57 (89%)	7 (13%)	0
1	D	51/57 (89%)	6 (11%)	0
1	E	55/57 (96%)	7 (12%)	1 (1%)
All	All	212/228 (92%)	28 (13%)	1 (0%)

All (28) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	C	5	C
1	C	17	A
1	C	18	C
1	C	21	G
1	C	22	A
1	C	28	G
1	C	52	G
1	E	5	C
1	E	17	A
1	E	18	C
1	E	21	G
1	E	23	U
1	E	24	C
1	E	25	U
1	B	5	C
1	B	6	U
1	B	17	A
1	B	24	C
1	B	25	U
1	B	26	G
1	B	35	A
1	B	39	C
1	D	5	C
1	D	6	U
1	D	17	A
1	D	18	C
1	D	35	A
1	D	44	G

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	E	22	A

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

4 non-standard protein/DNA/RNA residues 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	$\#|Z| > 2$
1	GTP	D	2	1	29,34,34	1.30	4 (13%)	35,54,54	1.31	6 (17%)
1	GTP	B	2	1	29,34,34	1.26	3 (10%)	35,54,54	1.29	4 (11%)
1	GTP	E	2	1	29,34,34	1.32	4 (13%)	35,54,54	1.33	4 (11%)
1	GTP	C	2	1	29,34,34	1.36	5 (17%)	35,54,54	1.26	4 (11%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	GTP	D	2	1	-	5/18/38/38	0/3/3/3
1	GTP	B	2	1	-	5/18/38/38	0/3/3/3
1	GTP	E	2	1	-	8/18/38/38	0/3/3/3
1	GTP	C	2	1	-	4/18/38/38	0/3/3/3

All (16) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	2	GTP	C5-C6	-4.28	1.39	1.47
1	C	2	GTP	C5-C6	-4.27	1.39	1.47
1	D	2	GTP	C5-C6	-4.27	1.39	1.47
1	E	2	GTP	C5-C6	-4.17	1.39	1.47
1	C	2	GTP	PB-O3A	2.70	1.62	1.59
1	E	2	GTP	C2-N3	2.47	1.39	1.33
1	E	2	GTP	PB-O3B	2.44	1.62	1.59
1	C	2	GTP	C2-N3	2.41	1.39	1.33
1	B	2	GTP	C2-N3	2.35	1.38	1.33
1	D	2	GTP	PA-O3A	2.33	1.62	1.59
1	C	2	GTP	PA-O3A	2.27	1.61	1.59
1	E	2	GTP	PB-O3A	2.24	1.61	1.59
1	D	2	GTP	C2-N3	2.22	1.38	1.33
1	D	2	GTP	PB-O3A	2.22	1.61	1.59
1	C	2	GTP	PB-O3B	2.13	1.61	1.59
1	B	2	GTP	PB-O3B	2.10	1.61	1.59

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	2	GTP	C8-N7-C5	3.69	108.83	102.55
1	E	2	GTP	C8-N7-C5	3.67	108.79	102.55
1	B	2	GTP	C8-N7-C5	3.65	108.76	102.55
1	D	2	GTP	C8-N7-C5	3.62	108.72	102.55
1	B	2	GTP	C5-C6-N1	2.98	119.76	114.07
1	E	2	GTP	C2-N1-C6	-2.97	119.67	125.11
1	B	2	GTP	C2-N1-C6	-2.91	119.78	125.11
1	C	2	GTP	C2-N1-C6	-2.90	119.81	125.11
1	E	2	GTP	C5-C6-N1	2.87	119.55	114.07
1	D	2	GTP	C5-C6-N1	2.87	119.54	114.07
1	C	2	GTP	C5-C6-N1	2.86	119.53	114.07
1	D	2	GTP	C2-N1-C6	-2.85	119.90	125.11
1	E	2	GTP	O2A-PA-O3A	2.28	113.44	107.27
1	D	2	GTP	O6-C6-C5	-2.18	120.00	124.32
1	B	2	GTP	O6-C6-C5	-2.14	120.09	124.32
1	D	2	GTP	O2A-PA-O3A	2.06	112.84	107.27
1	C	2	GTP	O6-C6-C5	-2.05	120.25	124.32
1	D	2	GTP	O2B-PB-O3A	2.03	112.75	107.27

There are no chirality outliers.

All (22) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	C	2	GTP	C5'-O5'-PA-O3A
1	C	2	GTP	C5'-O5'-PA-O2A
1	C	2	GTP	O4'-C4'-C5'-O5'
1	E	2	GTP	PB-O3B-PG-O3G
1	E	2	GTP	O4'-C4'-C5'-O5'
1	B	2	GTP	C5'-O5'-PA-O1A
1	B	2	GTP	C5'-O5'-PA-O2A
1	C	2	GTP	C3'-C4'-C5'-O5'
1	E	2	GTP	C3'-C4'-C5'-O5'
1	B	2	GTP	C3'-C4'-C5'-O5'
1	D	2	GTP	C3'-C4'-C5'-O5'
1	E	2	GTP	C5'-O5'-PA-O1A
1	B	2	GTP	C5'-O5'-PA-O3A
1	E	2	GTP	PB-O3A-PA-O2A
1	D	2	GTP	PB-O3A-PA-O2A
1	B	2	GTP	O4'-C4'-C5'-O5'
1	E	2	GTP	PB-O3B-PG-O1G
1	D	2	GTP	PG-O3B-PB-O2B
1	E	2	GTP	C4'-C5'-O5'-PA
1	D	2	GTP	O4'-C4'-C5'-O5'

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
1	E	2	GTP	PB-O3A-PA-O1A
1	D	2	GTP	PB-O3A-PA-O1A

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

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 ⓘ

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, 95th 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	B	56/57 (98%)	-0.01	1 (1%)	67	68	36, 52, 67, 76	0
1	C	54/57 (94%)	0.34	0	100	100	32, 59, 147, 165	0
1	D	53/57 (92%)	0.58	2 (3%)	44	46	41, 80, 110, 119	0
1	E	56/57 (98%)	0.28	4 (7%)	23	27	42, 59, 107, 171	0
2	A	17/17 (100%)	0.38	1 (5%)	29	32	44, 60, 68, 72	0
All	All	236/245 (96%)	0.30	8 (3%)	48	50	32, 60, 142, 171	0

All (8) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	23	U	3.3
2	A	45	ILE	2.9
1	D	49	C	2.8
1	E	25	U	2.7
1	D	15	A	2.6
1	E	24	C	2.3
1	E	22	A	2.0
1	B	17	A	2.0

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

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, 95th 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
1	GTP	B	2	32/32	0.73	0.12	52,61,116,118	0
1	GTP	E	2	32/32	0.80	0.13	56,66,137,148	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å²)	Q<0.9
1	GTP	C	2	32/32	0.86	0.12	37,45,94,113	0
1	GTP	D	2	32/32	0.89	0.10	37,42,80,88	0

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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, 95th 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(Å²)	Q<0.9
3	K	D	101	1/1	0.80	0.11	61,61,61,61	0
3	K	E	101	1/1	0.89	0.09	65,65,65,65	0
3	K	C	101	1/1	0.91	0.09	46,46,46,46	0

6.5 Other polymers ⓘ

There are no such residues in this entry.