



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 8, 2018 – 07:07 pm GMT

PDB ID : 2A2T  
Title : crystal structure of d(AAATATTT)  
Authors : Valls, N.; Richter, M.; Subirana, J.A.  
Deposited on : 2005-06-23  
Resolution : 3.10 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

<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  
Mogul : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : 1.13  
EDS : trunk30967  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30967

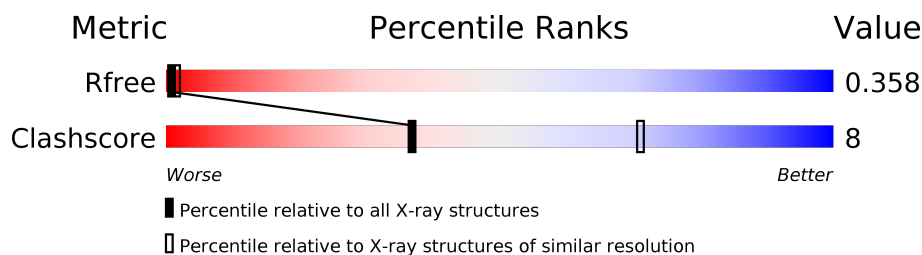
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*



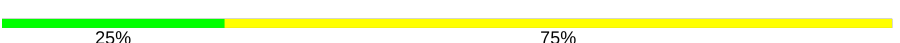



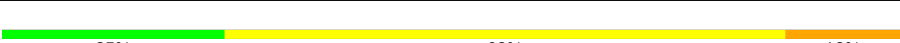
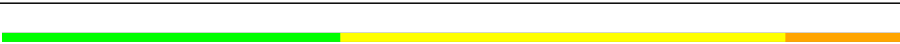
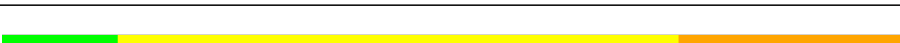
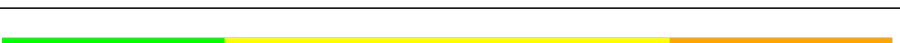
The reported resolution of this entry is 3.10 Å.

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 <sub>free</sub>	111664	1115 (3.12-3.08)
Clashscore	122126	1042 (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. 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	8	
1	B	8	
1	C	8	
1	D	8	
1	E	8	
1	F	8	
1	G	8	
1	H	8	
1	I	8	
1	J	8	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	K	8	 A horizontal bar chart showing the quality of chain 1. The bar is divided into three segments: green (50%), yellow (38%), and orange (13%).

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 1823 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 DNA chain called 5'-D(\*AP\*AP\*AP\*TP\*AP\*TP\*TP\*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	8	Total	C	N	O	P	0	0	0
			161	80	28	46	7			
1	G	8	Total	C	N	O	P	0	0	0
			161	80	28	46	7			
1	B	8	Total	C	N	O	P	0	0	0
			161	80	28	46	7			
1	H	8	Total	C	N	O	P	0	0	0
			161	80	28	46	7			
1	C	8	Total	C	N	O	P	0	0	0
			161	80	28	46	7			
1	I	8	Total	C	N	O	P	0	0	0
			161	80	28	46	7			
1	D	8	Total	C	N	O	P	0	0	0
			161	80	28	46	7			
1	J	8	Total	C	N	O	P	0	0	0
			161	80	28	46	7			
1	E	8	Total	C	N	O	P	0	0	0
			161	80	28	46	7			
1	K	8	Total	C	N	O	P	0	0	0
			161	80	28	46	7			
1	F	8	Total	C	N	O	P	0	0	0
			161	80	28	46	7			

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	D	1	Total	Mn	0	0
			1	1		
2	E	5	Total	Mn	0	0
			5	5		

- Molecule 3 is CACODYLATE ION (three-letter code: CAC) (formula: C<sub>2</sub>H<sub>6</sub>AsO<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	E	1	Total	As	C	O	0	0
			5	1	2	2		
3	E	1	Total	As	C	O	0	0
			5	1	2	2		
3	E	1	Total	As	C	O	0	0
			5	1	2	2		
3	E	1	Total	As	C	O	0	0
			5	1	2	2		
3	E	1	Total	As	C	O	0	0
			5	1	2	2		
3	E	1	Total	As	C	O	0	0
			5	1	2	2		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	O	0	0
			1	1		
4	G	1	Total	O	0	0
			1	1		
4	B	1	Total	O	0	0
			1	1		
4	H	1	Total	O	0	0
			1	1		
4	D	2	Total	O	0	0
			2	2		
4	J	1	Total	O	0	0
			1	1		

*Continued on next page...*

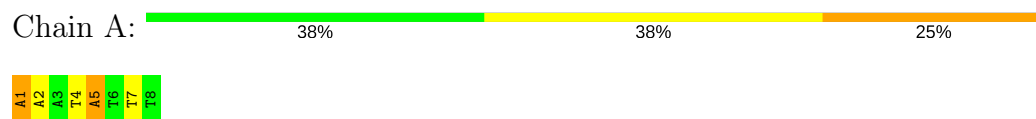
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	E	6	Total	O	0	0
			6	6		
4	K	3	Total	O	0	0
			3	3		

### 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. 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. 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: 5'-D(\*AP\*AP\*AP\*TP\*AP\*TP\*TP\*T)-3'



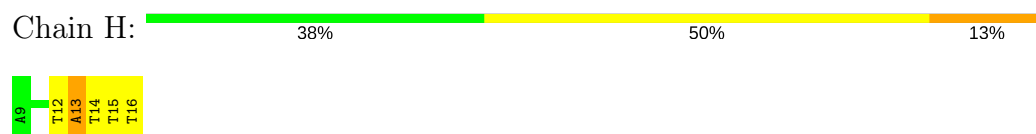
- Molecule 1: 5'-D(\*AP\*AP\*AP\*TP\*AP\*TP\*TP\*T)-3'



- Molecule 1: 5'-D(\*AP\*AP\*AP\*TP\*AP\*TP\*TP\*T)-3'



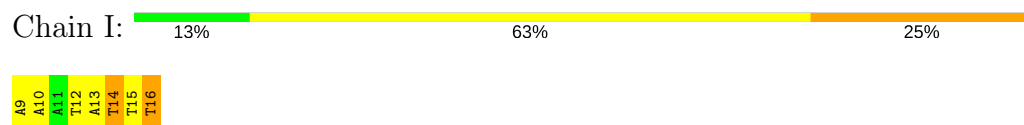
- Molecule 1: 5'-D(\*AP\*AP\*AP\*TP\*AP\*TP\*TP\*T)-3'



- Molecule 1: 5'-D(\*AP\*AP\*AP\*TP\*AP\*TP\*TP\*T)-3'

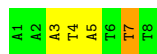


- Molecule 1: 5'-D(\*AP\*AP\*AP\*TP\*AP\*TP\*TP\*T)-3'



- Molecule 1: 5'-D(\*AP\*AP\*AP\*TP\*AP\*TP\*TP\*T)-3'

Chain D:  50% 38% 13%



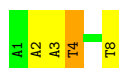
- Molecule 1: 5'-D(\*AP\*AP\*AP\*TP\*AP\*TP\*TP\*T)-3'

Chain J:  25% 50% 25%



- Molecule 1: 5'-D(\*AP\*AP\*AP\*TP\*AP\*TP\*TP\*T)-3'

Chain E:  50% 38% 13%



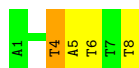
- Molecule 1: 5'-D(\*AP\*AP\*AP\*TP\*AP\*TP\*TP\*T)-3'

Chain K:  50% 38% 13%



- Molecule 1: 5'-D(\*AP\*AP\*AP\*TP\*AP\*TP\*TP\*T)-3'

Chain F:  50% 38% 13%





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	147.62Å 24.98Å 82.14Å 90.00° 90.51° 90.00°	Depositor
Resolution (Å)	15.00 – 3.10 36.90 – 2.78	Depositor EDS
% Data completeness (in resolution range)	99.4 (15.00-3.10) 98.8 (36.90-2.78)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.09 (at 2.77Å)	Xtriage
Refinement program	REFMAC 5.2.0016	Depositor
R, $R_{free}$	0.272 , 0.357 0.296 , 0.358	Depositor DCC
$R_{free}$ test set	362 reflections (4.59%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	49.4	Xtriage
Anisotropy	0.674	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 66.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.42$ , $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	0.000 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	1823	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 36.18 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 5.2042e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: CAC, MN

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	0.76	0/180	1.55	4/276 (1.4%)
1	B	0.77	0/180	1.50	4/276 (1.4%)
1	C	0.73	0/180	1.46	1/276 (0.4%)
1	D	0.73	0/180	1.67	3/276 (1.1%)
1	E	0.77	0/180	1.87	6/276 (2.2%)
1	F	0.61	0/180	1.39	2/276 (0.7%)
1	G	0.71	0/180	1.47	1/276 (0.4%)
1	H	0.70	0/180	1.92	7/276 (2.5%)
1	I	0.81	0/180	1.86	9/276 (3.3%)
1	J	0.78	0/180	1.56	2/276 (0.7%)
1	K	0.79	0/180	1.51	2/276 (0.7%)
All	All	0.75	0/1980	1.62	41/3036 (1.4%)

There are no bond length outliers.

All (41) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	H	13	DA	C1'-O4'-C4'	-10.00	100.10	110.10
1	E	3	DA	O4'-C1'-N9	-9.53	101.33	108.00
1	E	4	DT	O4'-C1'-N1	9.41	114.59	108.00
1	H	16	DT	C1'-O4'-C4'	-8.36	101.74	110.10
1	I	16	DT	C1'-O4'-C4'	-8.16	101.94	110.10
1	H	16	DT	O4'-C1'-N1	7.90	113.53	108.00
1	I	16	DT	O4'-C4'-C3'	-7.76	101.34	106.00
1	A	1	DA	O4'-C4'-C3'	-7.58	101.45	106.00
1	E	8	DT	C1'-O4'-C4'	-7.45	102.65	110.10
1	I	14	DT	O4'-C1'-N1	7.30	113.11	108.00
1	I	14	DT	C1'-O4'-C4'	-7.22	102.88	110.10
1	H	16	DT	O4'-C4'-C3'	-7.18	101.63	104.50
1	H	15	DT	O4'-C1'-N1	7.16	113.01	108.00

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	3	DA	O4'-C1'-N9	-7.15	102.99	108.00
1	H	12	DT	O4'-C1'-N1	6.61	112.62	108.00
1	A	4	DT	P-O3'-C3'	6.51	127.52	119.70
1	E	2	DA	O4'-C1'-N9	6.26	112.38	108.00
1	A	5	DA	P-O3'-C3'	6.23	127.18	119.70
1	B	4	DT	C1'-O4'-C4'	-6.17	103.93	110.10
1	D	7	DT	C6-C5-C7	-6.15	119.21	122.90
1	K	11	DA	O4'-C1'-N9	-6.10	103.73	108.00
1	I	10	DA	C1'-O4'-C4'	-6.04	104.06	110.10
1	J	16	DT	O4'-C4'-C3'	-6.04	102.08	104.50
1	F	8	DT	O4'-C1'-N1	6.01	112.20	108.00
1	K	12	DT	O4'-C1'-N1	6.00	112.20	108.00
1	E	8	DT	O4'-C4'-C3'	-5.90	102.14	104.50
1	F	4	DT	C4-C5-C7	5.77	122.46	119.00
1	G	12	DT	P-O3'-C3'	5.76	126.61	119.70
1	B	4	DT	O4'-C1'-N1	5.70	111.99	108.00
1	J	12	DT	O4'-C1'-N1	5.68	111.98	108.00
1	I	14	DT	O4'-C1'-C2'	-5.63	101.39	105.90
1	H	14	DT	O4'-C1'-N1	5.62	111.94	108.00
1	A	7	DT	O4'-C1'-N1	5.45	111.81	108.00
1	I	12	DT	N3-C4-O4	5.39	123.14	119.90
1	C	8	DT	N3-C4-O4	5.35	123.11	119.90
1	E	3	DA	P-O3'-C3'	5.28	126.04	119.70
1	D	7	DT	C4-C5-C7	5.18	122.11	119.00
1	I	10	DA	C3'-C2'-C1'	-5.18	96.28	102.50
1	B	8	DT	N3-C4-O4	5.12	122.97	119.90
1	I	12	DT	C5-C4-O4	-5.10	121.33	124.90
1	B	1	DA	P-O3'-C3'	5.09	125.80	119.70

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	161	0	94	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	161	0	94	2	0
1	C	161	0	94	3	0
1	D	161	0	94	4	0
1	E	161	0	94	1	0
1	F	161	0	94	2	0
1	G	161	0	94	5	0
1	H	161	0	94	1	0
1	I	161	0	94	4	0
1	J	161	0	94	6	0
1	K	161	0	94	3	0
2	D	1	0	0	0	0
2	E	5	0	0	0	0
3	E	30	0	0	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	D	2	0	0	0	0
4	E	6	0	0	0	0
4	G	1	0	0	0	0
4	H	1	0	0	0	0
4	J	1	0	0	0	0
4	K	3	0	0	0	0
All	All	1823	0	1034	24	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 8.

All (24) 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:J:9:DA:H5'	1:K:16:DT:OP2	1.99	0.62
1:B:5:DA:C2	1:H:13:DA:C2	2.93	0.56
1:I:15:DT:H2''	1:I:16:DT:O4'	2.08	0.53
1:E:4:DT:H3	1:K:13:DA:H2	1.57	0.53
1:D:4:DT:H2''	1:D:5:DA:O5'	2.10	0.52
1:D:7:DT:H3	1:J:10:DA:H61	1.59	0.51
1:J:12:DT:H1'	1:J:13:DA:H5'	1.93	0.51
1:C:3:DA:C2	1:C:4:DT:C2	2.99	0.51
1:I:13:DA:C2	1:I:14:DT:C2	3.00	0.49
1:I:9:DA:H1'	1:J:16:DT:H2'	1.94	0.48
1:F:5:DA:C2	1:F:6:DT:C2	3.02	0.47
1:G:9:DA:HO5'	1:G:9:DA:H8	1.63	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:G:13:DA:H1'	1:G:14:DT:H5'	1.97	0.45
1:D:7:DT:O2	1:J:11:DA:C2	2.69	0.45
1:K:12:DT:H2''	1:K:13:DA:C8	2.50	0.45
1:A:5:DA:C2	1:G:13:DA:C2	3.06	0.44
1:G:12:DT:H2''	1:G:13:DA:OP2	2.17	0.44
1:C:5:DA:C2	1:I:13:DA:C2	3.05	0.44
1:D:5:DA:C2	1:J:13:DA:C2	3.07	0.43
1:B:6:DT:H1'	1:B:7:DT:H5'	2.02	0.42
1:F:4:DT:H2''	1:F:5:DA:OP2	2.21	0.41
1:A:1:DA:H2''	1:A:2:DA:OP2	2.21	0.41
1:G:10:DA:C6	1:G:11:DA:C6	3.08	0.40
1:C:6:DT:H2''	1:C:7:DT:O5'	2.21	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

There are no protein molecules in this entry.

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

There are no protein molecules in this entry.

### 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

Of 12 ligands modelled in this entry, 6 are monoatomic - leaving 6 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	CAC	E	201	2	0,4,4	0.00	-	0,6,6	0.00	-
3	CAC	E	202	2	0,4,4	0.00	-	0,6,6	0.00	-
3	CAC	E	203	2	0,4,4	0.00	-	0,6,6	0.00	-
3	CAC	E	204	2	0,4,4	0.00	-	0,6,6	0.00	-
3	CAC	E	205	2	0,4,4	0.00	-	0,6,6	0.00	-
3	CAC	E	206	2	0,4,4	0.00	-	0,6,6	0.00	-

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
3	CAC	E	201	2	-	0/0/0/0	0/0/0/0
3	CAC	E	202	2	-	0/0/0/0	0/0/0/0
3	CAC	E	203	2	-	0/0/0/0	0/0/0/0
3	CAC	E	204	2	-	0/0/0/0	0/0/0/0
3	CAC	E	205	2	-	0/0/0/0	0/0/0/0
3	CAC	E	206	2	-	0/0/0/0	0/0/0/0

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 [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](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.