



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 10:26 am BST

PDB ID : 1K4S  
Title : HUMAN DNA TOPOISOMERASE I IN COVALENT COMPLEX WITH A  
22 BASE PAIR DNA DUPLEX  
Authors : Staker, B.L.; Hjerrild, K.; Feese, M.D.; Behnke, C.A.; Burgin Jr., A.B.; Stewart, L.J.  
Deposited on : 2001-10-08  
Resolution : 3.20 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
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

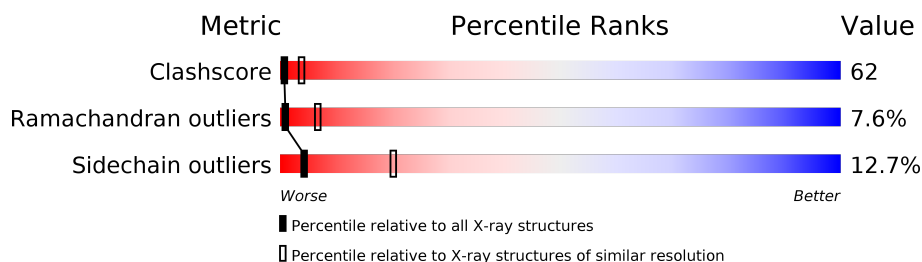
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.20 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))
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)

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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	B	10	90% 10%
2	C	12	8% 92%
3	D	22	86% 14%
4	A	592	24% 47% 10% 17%

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
1	5IU	B	9	-	-	X	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	5IU	C	18	-	-	X	-
2	5IU	C	19	-	-	X	-
2	5IU	C	20	-	-	X	-
3	5IU	D	107	-	-	X	-
3	5IU	D	109	-	-	X	-
3	5IU	D	110	-	-	X	-
3	5IU	D	116	-	-	X	-
3	5IU	D	118	-	-	X	-
3	5IU	D	119	-	-	X	-

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 4858 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\*AP\*AP\*GP\*AP\*CP\*(5IU)P\*(5IU))-3'.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	B	10	Total	C	I	N	O	P	0	0	0
			203	97	2	42	53	9			

- Molecule 2 is a DNA chain called 5'-D(\*(SPT)P\*GP\*AP\*AP\*AP\*AP\*AP\*(5IU)P\*(5IU)P\*(5IU)P\*(5IU)P\*T)-3'.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
2	C	12	Total	C	I	N	O	P	S	0	0	0
			244	116	4	42	70	11	1			

- Molecule 3 is a DNA chain called 5'-D(\*AP\*AP\*AP\*AP\*AP\*TP\*(IDO)UP\*(IDO)UP\*(IDO)UP\*(IDO)UP\*CP\*AP\*AP\*AP\*GP\*(IDO)UP\*CP\*(IDO)UP\*(IDO)UP\*(IDO)UP\*(IDO)UP\*T)-3'.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	D	22	Total	C	I	N	O	P			
			445	209	9	73	133	21	0	0	0

- Molecule 4 is a protein called DNA topoisomerase I.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
4	A	490	Total	C	N	O	P	S	0	0	0
			3966	2540	689	715	1	21			

There is a discrepancy between the modelled and reference sequences:


Chain	Residue	Modelled	Actual	Comment	Reference
A	723	PTR	TYR	MODIFIED RESIDUE	UNP P11387

### 3 Residue-property plots

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.

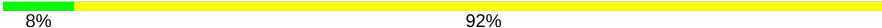
Note EDS was not executed.

- Molecule 1: 5'-D(\*AP\*AP\*AP\*AP\*AP\*GP\*AP\*CP\*(5IU)P\*(5IU))-3'

Chain B: 


A1 A2 A3 A4 A5 A6 A7 A8 A9 A10

- Molecule 2: 5'-D(\*(SPT)P\*GP\*AP\*AP\*AP\*AP\*AP\*(5IU)P\*(5IU)P\*(5IU)P\*(5IU)P\*T)-3',

Chain C: 

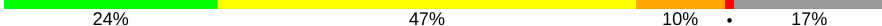
T11 G12 A13 A14 A15 A16 A17 A18 A19 A20 A21 A22

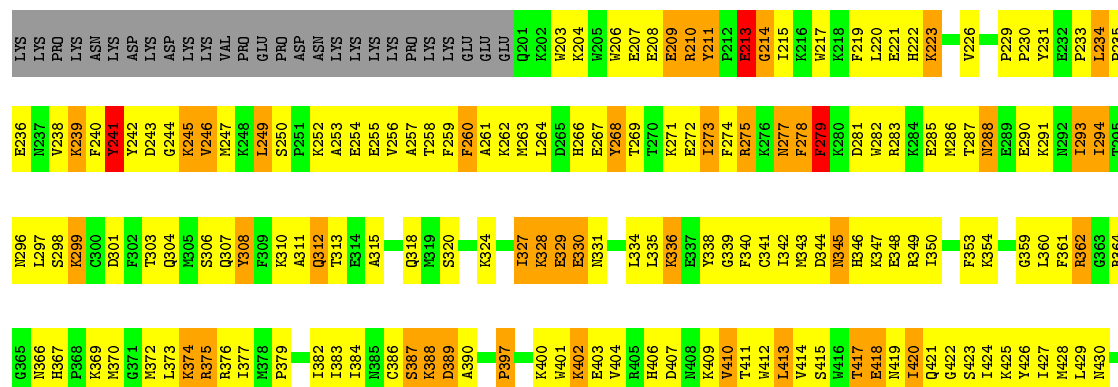
- Molecule 3: 5'-D(\*AP\*AP\*AP\*AP\*AP\*TP\*(IDO)UP\*(IDO)UP\*(IDO)UP\*(IDO)UP\*CP\*AP\*AP\*AP\*GP\*(IDO)UP\*CP\*(IDO)UP\*(IDO)UP\*(IDO)UP\*(IDO)UP\*T)-3'

Chain D: 

A101 A102 A103 A104 A105 A106 A107 A108 A109 A110 A111 A112 A113 A114 A115 A116 A117 A118 A119 A120 A121 A122

- Molecule 4: DNA topoisomerase I

Chain A: 



A753	G1N	H632	L588	D500	S433
A754	ARG	GLN	L569		R434
A755	LEU	ARG	T570	G503	I435
I756	GLU	ALA	G571	C504	K436
D757	GLU	PRO		C505	K437
M758	GLN	PRO	H574	S506	E438
A759	LEU	LYS	H575		K439
D760	MET	THR	H576	V509	D440
E761	LVS	PRE	L577	H510	H441
D762	LEU	GLU	D578	H511	Q442
V763	GLU	LYS	H579	L512	K443
E764	VAL	SER	L580	H513	H444
F765	GLN	MET	H581	L514	E445
	ALA	MET	E582	H515	T446
	THR	ASN	G583	P516	A447
	ASP	LEU	L584	E517	R448
	R708	GLN	T585	L518	R449
	E709	THR			L450
	H710	LVS	H588	Q521	
	N711	LVS	F589	V524	K456
	K712	ASP	H590	V525	I457
	Q713	ALA	T591	V526	K458
	I714	LVS	H592	F527	M459
	A715	LVS	H593	F527	Q460
	L716	GLU	A594	D528	
	G717	LEU	S595	F529	Y461
	T718	GLN	L596	L530	K462
	S719	ALA	T597	G531	E463
	K720	ASP	L598	K532	D464
	L721	ALA	Q599	D533	M465
	T722	ARG	H599	S534	K466
	T723	ARG	Q601	S534	S467
	L724	ASP	L600	R536	
	D725	LEU	K603	Y537	M470
	P726	LVS	G604	Y538	
	T727	SER	L605	M539	R473
	T728	ALA		R540	Q474
	T729	LVS	P608	V541	
	V730	ALA	D609		Y477
	A731	ASP	E510	E544	A478
	T732	ALA	H611		L479
	C733	LVS	L612	K549	Y480
	K734	VAL	H613	H550	F481
	K735	MET	A614	L551	I482
	W736	LVS	R615	Q552	D483
		ASP	L616	L553	K484
		ALA	L617	F554	L485
	P739	LVS	S618	H555	A486
	H740	LVS	S618	E556	L487
	E741	THR	P619	M557	R488
	K742	LVS	H620	K558	A489
	T743	LVS	R621		Q490
	Y744	VAL		Q559	M491
	N745	VAL	R624	P560	E492
	K746	GLU	A625	E561	K493
	T747	SER	V626	D562	E494
	Q748	LVS	A627	D563	E495
	R749	LVS	L628	L564	Q496
	T749	LVS	L629	F565	E497
	E750	LVS	P630	D566	T498
	K751	ALA	H631	P567	K499
	H752	VAL			

## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	73.23 Å 73.23 Å 186.63 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 3.20	Depositor
% Data completeness (in resolution range)	96.8 (50.00-3.20)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.13	Depositor
Refinement program	CNX	Depositor
R, $R_{free}$	0.217 , 0.222	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4858	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	41.0	wwPDB-VP

## 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: 5IU, SPT, PTR

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	0.56	0/186	0.90	1/285 (0.4%)
2	C	0.44	0/165	0.69	0/250
3	D	0.54	0/296	0.81	0/447
4	A	0.49	0/4046	0.69	0/5464
All	All	0.49	0/4693	0.71	1/6446 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
3	D	0	3

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	8	DC	C1'-O4'-C4'	-5.57	104.53	110.10

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	D	111	DC	Sidechain
3	D	114	DA	Sidechain
3	D	117	DC	Sidechain



## 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	203	0	106	45	0
2	C	244	0	127	28	0
3	D	445	0	228	95	0
4	A	3966	0	3841	434	0
All	All	4858	0	4302	559	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 62.

The worst 5 of 559 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:B:8:DC:H2''	1:B:9:5IU:I5	1.95	1.35
3:D:115:DG:H2'	3:D:116:5IU:I5	2.10	1.20
3:D:118:5IU:H2''	3:D:119:5IU:H5'	1.21	1.11
3:D:118:5IU:C2'	3:D:119:5IU:H5'	1.88	1.03
3:D:119:5IU:H2'	3:D:120:5IU:I5	2.28	1.03

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
4	A	485/592 (82%)	367 (76%)	81 (17%)	37 (8%)	<b>1</b> <b>7</b>

5 of 37 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	344	ASP
4	A	387	SER
4	A	388	LYS
4	A	389	ASP
4	A	495	GLU

### 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
4	A	408/535 (76%)	356 (87%)	52 (13%)	4 20

5 of 52 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	A	362	ARG
4	A	417	THR
4	A	709	GLU
4	A	397	PRO
4	A	402	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
4	A	421	GLN
4	A	474	GLN
4	A	632	HIS
4	A	419	ASN
4	A	722	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

17 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$
2	5IU	C	18	3,2	14,21,22	2.13	3 (21%)	16,30,33	3.95	3 (18%)
3	5IU	D	107	3,2	14,21,22	1.93	3 (21%)	16,30,33	4.01	4 (25%)
3	5IU	D	119	1,3	14,21,22	2.37	3 (21%)	16,30,33	3.91	3 (18%)
3	5IU	D	116	1,3	14,21,22	2.28	3 (21%)	16,30,33	3.91	3 (18%)
3	5IU	D	108	3,2	14,21,22	2.07	3 (21%)	16,30,33	3.92	3 (18%)
2	SPT	C	11	3,2	15,18,22	1.18	2 (13%)	16,26,33	3.64	2 (12%)
2	5IU	C	19	3,2	14,21,22	1.99	3 (21%)	16,30,33	3.99	4 (25%)
3	5IU	D	109	3,2	14,21,22	2.04	3 (21%)	16,30,33	3.97	4 (25%)
2	5IU	C	21	3,2	14,21,22	2.10	3 (21%)	16,30,33	3.92	3 (18%)
3	5IU	D	120	1,3	14,21,22	1.87	3 (21%)	16,30,33	3.96	4 (25%)
1	5IU	B	10	1,3,4	14,20,22	2.08	2 (14%)	13,28,33	4.58	5 (38%)
3	5IU	D	118	1,3	14,21,22	2.28	3 (21%)	16,30,33	3.88	3 (18%)
2	5IU	C	20	3,2	14,21,22	2.66	3 (21%)	16,30,33	3.98	3 (18%)
4	PTR	A	723	1,4	15,16,17	0.80	0	19,22,24	0.84	1 (5%)
1	5IU	B	9	1,3	14,21,22	2.24	3 (21%)	16,30,33	3.87	4 (25%)
3	5IU	D	110	3,2	14,21,22	2.06	3 (21%)	16,30,33	3.95	4 (25%)
3	5IU	D	121	1,3	14,21,22	2.11	3 (21%)	16,30,33	3.94	3 (18%)

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
2	5IU	C	18	3,2	-	3/4/21/22	0/2/2/2
3	5IU	D	107	3,2	-	2/4/21/22	0/2/2/2
3	5IU	D	119	1,3	-	3/4/21/22	0/2/2/2

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	5IU	D	116	1,3	-	1/4/21/22	0/2/2/2
3	5IU	D	108	3,2	-	1/4/21/22	0/2/2/2
2	SPT	C	11	3,2	-	2/3/18/22	0/2/2/2
2	5IU	C	19	3,2	-	1/4/21/22	0/2/2/2
3	5IU	D	109	3,2	-	2/4/21/22	0/2/2/2
2	5IU	C	21	3,2	-	1/4/21/22	0/2/2/2
3	5IU	D	120	1,3	-	1/4/21/22	0/2/2/2
1	5IU	B	10	1,3,4	-	1/4/18/22	0/2/2/2
3	5IU	D	118	1,3	-	0/4/21/22	0/2/2/2
2	5IU	C	20	3,2	-	2/4/21/22	0/2/2/2
4	PTR	A	723	1,4	-	2/10/11/13	0/1/1/1
1	5IU	B	9	1,3	-	0/4/21/22	0/2/2/2
3	5IU	D	110	3,2	-	2/4/21/22	0/2/2/2
3	5IU	D	121	1,3	-	1/4/21/22	0/2/2/2

The worst 5 of 46 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	20	5IU	C5-I5	-7.54	1.96	2.10
3	D	116	5IU	C5-I5	-6.53	1.98	2.10
3	D	119	5IU	C5-I5	-6.51	1.98	2.10
1	B	10	5IU	C5-I5	-6.38	1.98	2.10
1	B	9	5IU	C5-I5	-6.21	1.98	2.10

The worst 5 of 56 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	107	5IU	C4-N3-C2	14.16	127.10	115.14
2	C	11	SPT	C4-N3-C2	14.13	127.07	115.14
1	B	10	5IU	C4-N3-C2	14.04	127.00	115.14
2	C	19	5IU	C4-N3-C2	14.01	126.97	115.14
3	D	121	5IU	C4-N3-C2	13.94	126.91	115.14

There are no chirality outliers.

5 of 25 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	18	5IU	O4'-C1'-N1-C6
3	D	107	5IU	O4'-C1'-N1-C6
3	D	119	5IU	O4'-C4'-C5'-O5'

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
3	D	116	5IU	O4'-C1'-N1-C6
3	D	108	5IU	O4'-C1'-N1-C6

There are no ring outliers.

17 monomers are involved in 94 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	18	5IU	8	0
3	D	107	5IU	13	0
3	D	119	5IU	13	0
3	D	116	5IU	16	0
3	D	108	5IU	4	0
2	C	11	SPT	5	0
2	C	19	5IU	7	0
3	D	109	5IU	9	0
2	C	21	5IU	5	0
3	D	120	5IU	3	0
1	B	10	5IU	4	0
3	D	118	5IU	8	0
2	C	20	5IU	10	0
4	A	723	PTR	3	0
1	B	9	5IU	9	0
3	D	110	5IU	8	0
3	D	121	5IU	1	0

## 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 ⓘ

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

### 6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.