



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 16, 2021 – 10:09 PM EDT

PDB ID : 1STX  
Title : Structure of the K38A mutant of EcoRV bound to cognate DNA and Mn<sup>2+</sup>  
Authors : Horton, N.C.; Perona, J.J.  
Deposited on : 2004-03-25  
Resolution : 2.10 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  
Xtrriage (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) : 2.23.2

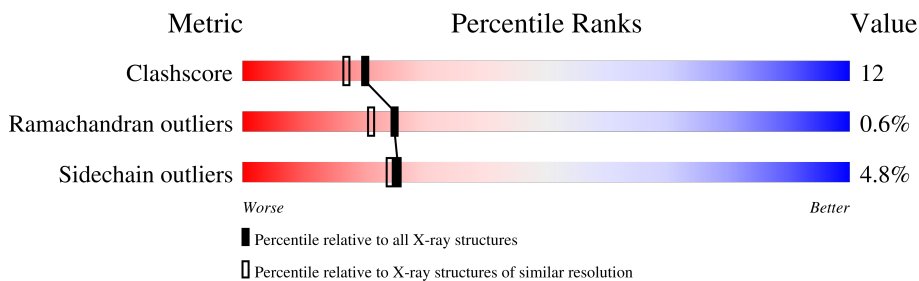
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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))
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)

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 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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	C	6	
1	E	6	
2	D	5	
2	F	5	
3	A	244	
3	B	244	

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 4806 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\*GP\*AP\*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	C	6	Total 123	C 60	N 27	O 31	P 5	0	0	0
1	E	6	Total 123	C 60	N 27	O 31	P 5	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	D	5	Total 101	C 49	N 14	O 33	P 5	0	0	0
2	F	5	Total 101	C 49	N 14	O 33	P 5	0	0	0

- Molecule 3 is a protein called Type II restriction enzyme EcoRV.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	244	Total 1997	C 1284	N 333	O 379	S 1	0	0	0
3	B	244	Total 1972	C 1272	N 323	O 376	S 1	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	38	ALA	LYS	engineered mutation	UNP P04390
B	38	ALA	LYS	engineered mutation	UNP P04390

- Molecule 4 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	4	Total 4	Mn 4	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	2	Total	Mn	0	0
			2	2		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	C	15	Total	O	0	0
			15	15		
5	D	9	Total	O	0	0
			9	9		
5	E	12	Total	O	0	0
			12	12		
5	F	14	Total	O	0	0
			14	14		
5	A	181	Total	O	0	0
			181	181		
5	B	152	Total	O	0	0
			152	152		

### 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. 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\*GP\*AP\*T)-3'

Chain C: 



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

Chain E: 



- Molecule 2: 5'-D(P\*AP\*TP\*CP\*TP\*T)-3'

Chain D: 



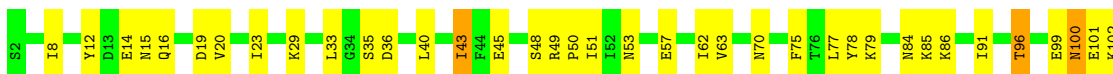
- Molecule 2: 5'-D(P\*AP\*TP\*CP\*TP\*T)-3'

Chain F: 


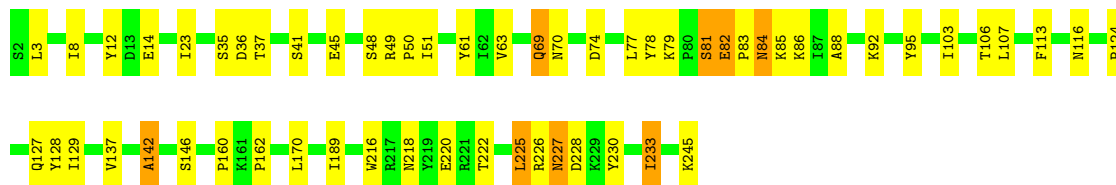


- Molecule 3: Type II restriction enzyme EcoRV

Chain A: 



## ● Molecule 3: Type II restriction enzyme EcoRV

Chain B:  76% 20%

## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	63.71Å 58.69Å 82.10Å 90.00° 107.70° 90.00°	Depositor
Resolution (Å)	20.00 – 2.10	Depositor
% Data completeness (in resolution range)	99.2 (20.00-2.10)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNS	Depositor
R, $R_{free}$	0.201 , 0.262	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4806	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	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: 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	C	1.35	0/139	1.25	1/213 (0.5%)
1	E	1.16	0/139	1.19	1/213 (0.5%)
2	D	1.28	1/111 (0.9%)	1.44	1/167 (0.6%)
2	F	1.63	3/111 (2.7%)	1.40	2/167 (1.2%)
3	A	1.00	0/2048	0.95	4/2776 (0.1%)
3	B	0.94	0/2023	0.95	0/2746
All	All	1.02	4/4571 (0.1%)	1.00	9/6282 (0.1%)

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
2	D	0	1
3	B	0	1
All	All	0	2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	8	DT	C4-C5	6.48	1.50	1.45
2	D	7	DA	OP3-P	-5.86	1.54	1.61
2	F	9	DC	C2-O2	5.42	1.29	1.24
2	F	8	DT	N1-C2	5.37	1.42	1.38

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	7	DA	OP1-P-OP2	-8.66	106.61	119.60

*Continued on next page...*



Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	7	DA	OP1-P-OP2	-8.03	107.56	119.60
3	A	137	VAL	CB-CA-C	-7.35	97.43	111.40
2	F	8	DT	OP2-P-O3'	5.65	117.64	105.20
1	E	2	DA	C5'-C4'-C3'	-5.62	103.99	114.10
1	C	5	DA	OP2-P-O3'	5.54	117.38	105.20
3	A	242	ARG	NE-CZ-NH1	5.46	123.03	120.30
3	A	179	ASP	CB-CG-OD1	5.33	123.10	118.30
3	A	35	SER	N-CA-C	5.00	124.50	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	B	95	TYR	Sidechain
2	D	8	DT	Sidechain

## 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	C	123	0	69	7	0
1	E	123	0	69	4	0
2	D	101	0	59	2	0
2	F	101	0	59	10	0
3	A	1997	0	1928	49	0
3	B	1972	0	1883	44	0
4	A	4	0	0	0	0
4	B	2	0	0	0	0
5	A	181	0	0	5	0
5	B	152	0	0	5	0
5	C	15	0	0	0	0
5	D	9	0	0	0	0
5	E	12	0	0	0	0
5	F	14	0	0	4	0
All	All	4806	0	4067	105	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 12.

All (105) 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)
3:A:242:ARG:HH11	3:A:242:ARG:HB2	1.20	1.05
2:F:7:DA:H2''	2:F:8:DT:H5''	1.53	0.90
3:A:14:GLU:HG2	3:A:51:ILE:HD11	1.61	0.83
1:E:6:DT:O4'	3:B:70:ASN:HB2	1.80	0.82
3:B:14:GLU:HG2	3:B:51:ILE:HD11	1.64	0.80
3:A:49:ARG:HG2	3:A:75:PHE:HZ	1.47	0.79
3:A:36:ASP:HB2	3:B:69:GLN:HE22	1.50	0.76
3:A:220:GLU:HB2	3:A:226:ARG:HG2	1.69	0.74
3:A:242:ARG:HB2	3:A:242:ARG:NH1	2.01	0.72
3:A:96:THR:HG22	3:A:139:THR:HA	1.71	0.72
2:F:7:DA:OP2	5:F:63:HOH:O	2.09	0.70
3:B:79:LYS:HB2	3:B:82:GLU:HG3	1.74	0.69
2:F:7:DA:H2''	2:F:8:DT:C5'	2.25	0.67
3:B:63:VAL:HG12	3:B:77:LEU:HD22	1.76	0.67
5:F:63:HOH:O	3:B:45:GLU:OE2	2.12	0.66
3:A:12:TYR:O	3:A:16:GLN:HG2	1.97	0.65
3:B:69:GLN:HE21	3:B:69:GLN:HA	1.60	0.65
3:A:100:ASN:O	3:A:101:GLU:HB3	1.97	0.64
2:F:9:DC:H1'	3:A:70:ASN:HD21	1.62	0.64
3:B:227:ASN:HA	5:B:1127:HOH:O	1.96	0.63
3:A:23:ILE:HG22	3:B:23:ILE:HG22	1.80	0.62
3:B:12:TYR:HB3	5:B:1098:HOH:O	1.98	0.62
5:F:63:HOH:O	3:B:74:ASP:OD1	2.16	0.60
3:A:147:SER:HB3	3:B:49:ARG:HH22	1.66	0.59
3:B:82:GLU:N	3:B:83:PRO:HD3	2.18	0.58
3:B:142:ALA:N	5:B:1157:HOH:O	2.28	0.58
1:C:1:DA:H2'	1:C:2:DA:N7	2.19	0.57
3:A:84:ASN:C	3:A:85:LYS:HD2	2.25	0.57
1:E:6:DT:H2''	2:F:7:DA:H5''	1.88	0.56
3:A:19:ASP:HB3	3:A:161:LYS:HE2	1.87	0.56
3:B:92:LYS:HE3	3:B:107:LEU:HA	1.89	0.54
3:B:81:SER:C	3:B:82:GLU:HG2	2.27	0.54
3:B:124:PRO:HD2	3:B:127:GLN:NE2	2.23	0.54
3:A:49:ARG:HB2	3:A:50:PRO:HD3	1.90	0.53
3:B:220:GLU:HB3	3:B:225:LEU:HB3	1.90	0.53
3:A:78:TYR:HB3	3:A:86:LYS:HG2	1.91	0.53
3:A:170:LEU:C	3:A:170:LEU:HD23	2.28	0.53
3:B:222:THR:O	3:B:226:ARG:HG3	2.10	0.52

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:170:LEU:HD23	3:A:171:GLN:N	2.24	0.52
3:A:99:GLU:O	3:A:100:ASN:HB2	2.09	0.51
3:B:82:GLU:HB2	5:B:1015:HOH:O	2.11	0.51
3:B:113:PHE:HA	3:B:116:ASN:O	2.11	0.51
1:E:5:DA:H2'	1:E:6:DT:H71	1.93	0.50
3:A:15:ASN:HD21	3:A:166:VAL:CG2	2.24	0.50
3:A:53:ASN:O	3:A:57:GLU:HG3	2.11	0.50
2:F:9:DC:H1'	3:A:70:ASN:ND2	2.25	0.49
3:A:20:VAL:HG11	3:A:43:ILE:HG23	1.94	0.49
3:A:36:ASP:CB	3:B:69:GLN:HE22	2.23	0.49
3:A:158:GLU:HG3	5:A:1185:HOH:O	2.11	0.49
1:C:1:DA:C2'	1:C:2:DA:C8	2.96	0.49
2:F:9:DC:H2'	2:F:10:DT:H72	1.93	0.49
3:A:62:ILE:O	3:A:77:LEU:HA	2.12	0.49
1:C:6:DT:H2'	5:A:1151:HOH:O	2.13	0.49
1:C:1:DA:H2'	1:C:2:DA:C8	2.47	0.48
3:A:152:ASN:OD1	3:A:154:ASN:HB2	2.14	0.48
1:C:2:DA:H5''	1:C:2:DA:H8	1.78	0.48
3:A:20:VAL:HG12	3:A:33:LEU:HD11	1.96	0.48
3:A:8:ILE:HG12	3:A:170:LEU:HB3	1.96	0.48
3:B:8:ILE:HG12	3:B:170:LEU:HB3	1.96	0.48
3:B:83:PRO:C	3:B:84:ASN:HD22	2.17	0.47
3:A:45:GLU:HG2	3:A:91:ILE:HB	1.97	0.47
3:B:3:LEU:HD22	3:B:61:TYR:CE2	2.50	0.47
3:B:106:THR:HG23	5:B:1076:HOH:O	2.14	0.47
2:D:7:DA:H2'	5:A:1151:HOH:O	2.16	0.46
2:D:7:DA:OP1	2:D:7:DA:H3'	2.16	0.46
3:A:84:ASN:O	3:A:85:LYS:HD2	2.16	0.46
3:A:53:ASN:HA	3:A:63:VAL:HG21	1.98	0.46
3:A:195:HIS:HD2	3:A:197:LYS:NZ	2.14	0.46
3:B:85:LYS:HG2	3:B:129:ILE:HG12	1.98	0.45
3:A:220:GLU:HB2	3:A:226:ARG:CG	2.45	0.45
3:A:169:PHE:C	3:A:169:PHE:CD1	2.90	0.45
3:B:222:THR:OG1	3:B:225:LEU:HB2	2.16	0.45
3:A:86:LYS:HE3	3:A:123:TYR:CG	2.52	0.45
3:A:79:LYS:HE2	5:A:1058:HOH:O	2.17	0.44
3:B:84:ASN:HD22	3:B:84:ASN:N	2.15	0.44
1:E:6:DT:H2''	2:F:7:DA:C5'	2.47	0.44
3:A:48:SER:HA	3:A:51:ILE:HD12	1.99	0.44
1:C:2:DA:H5'	3:B:226:ARG:NH2	2.33	0.43
3:B:49:ARG:HB2	3:B:50:PRO:HD3	1.99	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:B:48:SER:HA	3:B:51:ILE:HD12	2.00	0.43
3:B:37:THR:O	3:B:41:SER:HB3	2.18	0.43
3:B:88:ALA:HB2	3:B:128:TYR:CD2	2.53	0.43
1:C:1:DA:H2''	1:C:2:DA:C8	2.54	0.43
3:A:96:THR:HG21	5:A:1011:HOH:O	2.19	0.43
3:B:170:LEU:C	3:B:170:LEU:HD23	2.39	0.43
3:A:233:ILE:HG23	3:A:234:SER:N	2.34	0.43
3:B:78:TYR:HB3	3:B:86:LYS:HG2	2.02	0.42
3:A:144:ARG:O	3:A:147:SER:HB2	2.19	0.42
3:B:103:ILE:HD12	3:B:137:VAL:CG2	2.50	0.42
3:B:218:ASN:HB2	3:B:230:TYR:OH	2.19	0.42
3:A:40:LEU:HD12	3:A:138:TYR:CZ	2.55	0.42
3:A:103:ILE:HG21	3:A:137:VAL:HG11	2.00	0.42
3:A:29:LYS:HE3	3:A:150:THR:CB	2.50	0.42
3:A:29:LYS:HE3	3:A:150:THR:HG21	2.01	0.42
3:B:146:SER:OG	3:B:160:PRO:HG3	2.19	0.42
3:B:35:SER:HA	3:B:162:PRO:HB2	2.01	0.42
3:B:189:ILE:HG23	3:B:216:TRP:CE2	2.54	0.42
2:F:10:DT:H5'	3:A:70:ASN:HD22	1.83	0.41
3:B:82:GLU:N	3:B:83:PRO:CD	2.84	0.41
3:A:137:VAL:O	3:A:166:VAL:HA	2.21	0.41
3:B:36:ASP:OD1	3:B:36:ASP:N	2.49	0.41
2:F:11:DT:H6	5:F:374:HOH:O	2.04	0.40
3:A:220:GLU:HB3	3:A:225:LEU:HB3	2.03	0.40
3:B:233:ILE:O	3:B:233:ILE:HG13	2.20	0.40
3:A:109:GLY:HA2	3:A:187:THR:O	2.21	0.40

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	
3	A	242/244 (99%)	228 (94%)	13 (5%)	1 (0%)	34	32
3	B	242/244 (99%)	225 (93%)	15 (6%)	2 (1%)	19	15
All	All	484/488 (99%)	453 (94%)	28 (6%)	3 (1%)	25	21

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	B	142	ALA
3	A	100	ASN
3	B	227	ASN

### 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	
3	A	213/219 (97%)	201 (94%)	12 (6%)	21	18
3	B	207/219 (94%)	199 (96%)	8 (4%)	32	33
All	All	420/438 (96%)	400 (95%)	20 (5%)	25	24

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	43	ILE
3	A	96	THR
3	A	102	LYS
3	A	140	ARG
3	A	147	SER
3	A	148	LEU
3	A	149	LYS
3	A	153	ILE
3	A	197	LYS
3	A	223	SER
3	A	227	ASN
3	A	242	ARG
3	B	69	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	B	81	SER
3	B	82	GLU
3	B	84	ASN
3	B	225	LEU
3	B	228	ASP
3	B	233	ILE
3	B	245	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
3	A	15	ASN
3	A	70	ASN
3	A	195	HIS
3	A	227	ASN
3	B	59	HIS
3	B	69	GLN
3	B	70	ASN
3	B	84	ASN
3	B	227	ASN

### 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](#)

Of 6 ligands modelled in this entry, 6 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section is therefore empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.