



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 12, 2024 – 04:31 PM EDT

PDB ID : 1HAO
Title : COMPLEX OF HUMAN ALPHA-THROMBIN WITH A 15MER OLIGONUCLEOTIDE GGTTGGTGTGGTTGG (BASED ON NMR MODEL OF DNA)
Authors : Tulinsky, A.; Padmanabhan, K.
Deposited on : 1995-10-03
Resolution : 2.80 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 ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

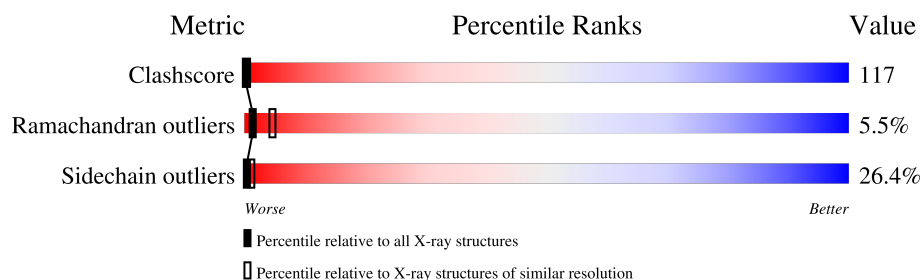
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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	180529	4123 (2.80-2.80)
Ramachandran outliers	177936	4071 (2.80-2.80)
Sidechain outliers	177891	4073 (2.80-2.80)

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 $\leq 5\%$.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	D	15	87% 13%
2	L	36	6% 25% 28% 17% 25%
3	H	259	10% 46% 32% 9% .

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
4	OG6	H	297	X	-	-	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 2769 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 DNA 5'-D(*GP*GP*TP*TP*GP*GP*TP*GP*TP*GP*GP*TP*TP*GP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	D	15	Total	C	N	O	P	0	0	0
			315	150	57	94	14			

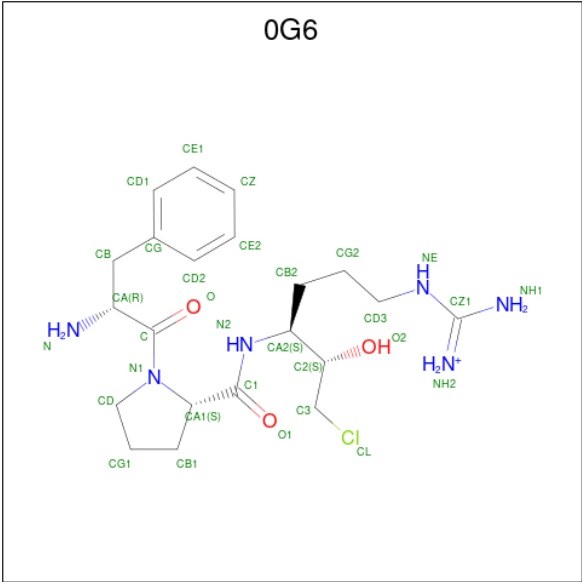
- Molecule 2 is a protein called ALPHA-THROMBIN light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	27	Total	C	N	O	S	0	0	0
			222	140	36	45	1			

- Molecule 3 is a protein called ALPHA-THROMBIN heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	H	253	Total	C	N	O	S	0	0	0
			2053	1310	362	367	14			

- Molecule 4 is D-phenylalanyl-N-[(2S,3S)-6-{[amino(iminio)methyl]amino}-1-chloro-2-hydroxyhexan-3-yl]-L-prolinamide (three-letter code: 0G6) (formula: C₂₁H₃₄ClN₆O₃).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	H	1	Total	C	N	O	0	0
			30	21	6	3		

- Molecule 5 is water.


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	D	19	Total	O	0	0
			19	19		
5	L	11	Total	O	0	0
			11	11		
5	H	119	Total	O	0	0
			119	119		

3 Residue-property plots

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: DNA 5'-D(*GP*GP*TP*TP*GP*GP*TP*GP*TP*GP*GP*TP*TP*GP*G)-3',

Chain D: 

G401
G402
T403
T404
G405
G406
T407
G408
T409
G410
G411
T412
T413
G414
G415

- Molecule 2: ALPHA-THROMBIN light chain

Chain L: 

THR
PHE
GLY
SER
GLY
GLY
A1B
D1A
C1
G2
L3
G2
R4
P5
L6
F7
E8
K9
K10
S11
L12
E13
D14
K14A
T14B
E14C
R14D
E14E
L14F
E14G
E14H
S14I
Y14J
I14K
ASP
GLY
ARG

- Molecule 3: ALPHA-THROMBIN heavy chain

Chain H: 

I16
V17
E18
G19
S20
D21
A22
E23
T24
G25
M26
S27
P28
K29
Q30
M32
L33
K34
R35
K36
S36A
P37
Q38
E39
L40
L41
C42
G43
A44
S45
L46
I47
S48
D49
R50
W51
V52
L53
T54
A55
A56
L105
H57
C58
L59
L60
Y60A
P60B
P60C
W60D
D60E
K60F
N60G
F60H
T60I
E61
N62
D63
L64
L65

V66
R67
G68
G69
H70
K71
S72
R73
T74
R75
Y76
E77
R77A
W78
I79
E80
K81
I82
S83
M84
L85
E86
K87
I88
Y89
I90
H91
P92
R93
Y94
N95
W96
R97
E97A
N98
L99
D100
R101
D102
I103
A104
L105
M106
K107
L108
K109
K110
P111
V112
A113
F114
S115
D116
Y117
I118
H119
P120
V121
C122
L123

P124
D125
R126
M127
T128
A129
A129A
S129B
L129C
L130
Y131
G132
Y134
K135
G136
R137
T138
V139
G140
W141
G142
N143
L144
K145
E146
T147
W148
THR
P199
V200
M201
K202
S203
P204
P204A
N205
P152
S153
V154
L155
Q209
Q156
M210
G211
V157
N159
L160
P161
I162
V163
E164
P165
P166
V167
C168
K169
D170
S171
T172
R173
I174
R175
I176

T177
D178
M179
M180
F181
C182
A183
G184
Y184A
K185
P186
D186A
E186B
G186C
R187
G188
D189
A190
C191
E192
G193
D194
S195
G196
G197
P198
F199
V200
M201
K202
S203
P204
P204A
N205
P152
S153
V154
L155
Q209
Q156
M210
G211
V157
N159
L160
P161
I162
V163
E164
P165
P166
V167
C168
K169
D170
S171
T172
R173
I174
R175
I176

F232
R233
L234
K235
K236
K237
I238
Q239
V240
Y241
D242
D243
F245
G246
E247

4 Data and refinement statistics

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

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	58.28Å 77.61Å 100.02Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	10.00 – 2.80	Depositor
% Data completeness (in resolution range)	68.0 (10.00-2.80)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	NUCLIN, PROLSQ	Depositor
R, R_{free}	0.155 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2769	wwPDB-VP
Average B, all atoms (Å ²)	17.0	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: 0G6

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	D	1.77	4/353 (1.1%)	3.93	95/547 (17.4%)
2	L	0.97	0/224	2.60	24/298 (8.1%)
3	H	1.06	3/2107 (0.1%)	2.27	111/2846 (3.9%)
All	All	1.17	7/2684 (0.3%)	2.61	230/3691 (6.2%)

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
1	D	0	2
3	H	0	3
All	All	0	5

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	H	44	ALA	C-N	-7.03	1.17	1.34
3	H	60(D)	TRP	C-N	-6.71	1.18	1.34
1	D	415	DG	O4'-C1'	6.26	1.49	1.42
1	D	405	DG	O4'-C1'	6.09	1.49	1.42
1	D	403	DT	C4-O4	5.33	1.28	1.23

The worst 5 of 230 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	410	DG	P-O3'-C3'	24.52	149.13	119.70
1	D	409	DT	P-O3'-C3'	17.95	141.24	119.70
3	H	187	ARG	NE-CZ-NH2	17.57	129.08	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	408	DG	P-O3'-C3'	-13.46	103.55	119.70
1	D	409	DT	O4'-C1'-N1	12.92	117.04	108.00

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	402	DG	Sidechain
1	D	405	DG	Sidechain
3	H	181	PHE	Mainchain
3	H	247	GLU	Sidechain
3	H	77(A)	ARG	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	D	315	0	173	59	0
2	L	222	0	224	57	0
3	H	2053	0	2017	492	0
4	H	30	0	30	19	0
5	D	19	0	0	0	0
5	H	119	0	0	23	0
5	L	11	0	0	2	0
All	All	2769	0	2444	588	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 117.

The worst 5 of 588 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:H:144:LEU:HD21	3:H:152:PRO:CD	1.58	1.31
3:H:195:SER:CB	4:H:297:OG6:C3	2.09	1.29
3:H:84:MET:SD	3:H:109:LYS:HE3	1.79	1.20

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:H:17:VAL:O	3:H:188:GLY:HA2	1.40	1.20
3:H:144:LEU:CD2	3:H:152:PRO:HD3	1.71	1.20

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	
2	L	25/36 (69%)	19 (76%)	3 (12%)	3 (12%)	0	1
3	H	249/259 (96%)	192 (77%)	45 (18%)	12 (5%)	2	6
All	All	274/295 (93%)	211 (77%)	48 (18%)	15 (6%)	1	4

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	L	2	GLY
3	H	38	GLN
3	H	77	GLU
2	L	14(D)	ARG
3	H	49	ASP

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	
2	L	25/31 (81%)	14 (56%)	11 (44%)	0	0
3	H	221/225 (98%)	167 (76%)	54 (24%)	0	2
All	All	246/256 (96%)	181 (74%)	65 (26%)	0	1

5 of 65 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	H	224	LYS
3	H	234	LEU
3	H	61	GLU
3	H	60(E)	ASP
3	H	239	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
3	H	179	ASN
3	H	204(B)	ASN
3	H	71	HIS
3	H	78	ASN
3	H	91	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

1 ligand is 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
4	0G6	H	297	3	30,31,32	2.25	2 (6%)	37,41,42	1.87	10 (27%)

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
4	0G6	H	297	3	1/1/8/10	7/31/41/43	0/2/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	H	297	0G6	C3-C2	-9.72	1.24	1.51
4	H	297	0G6	O2-C2	-6.35	1.26	1.43

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	H	297	0G6	NE-CZ1-NH2	-4.19	113.49	120.67
4	H	297	0G6	C1-CA1-N1	-3.87	101.93	112.50
4	H	297	0G6	O2-C2-C3	3.83	121.14	109.68
4	H	297	0G6	CA2-N2-C1	-3.77	116.69	123.25
4	H	297	0G6	CG2-CB2-CA2	-3.29	107.67	113.94

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	H	297	0G6	C2

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	H	297	0G6	C3-C2-CA2-N2

Continued on next page...

Continued from previous page...

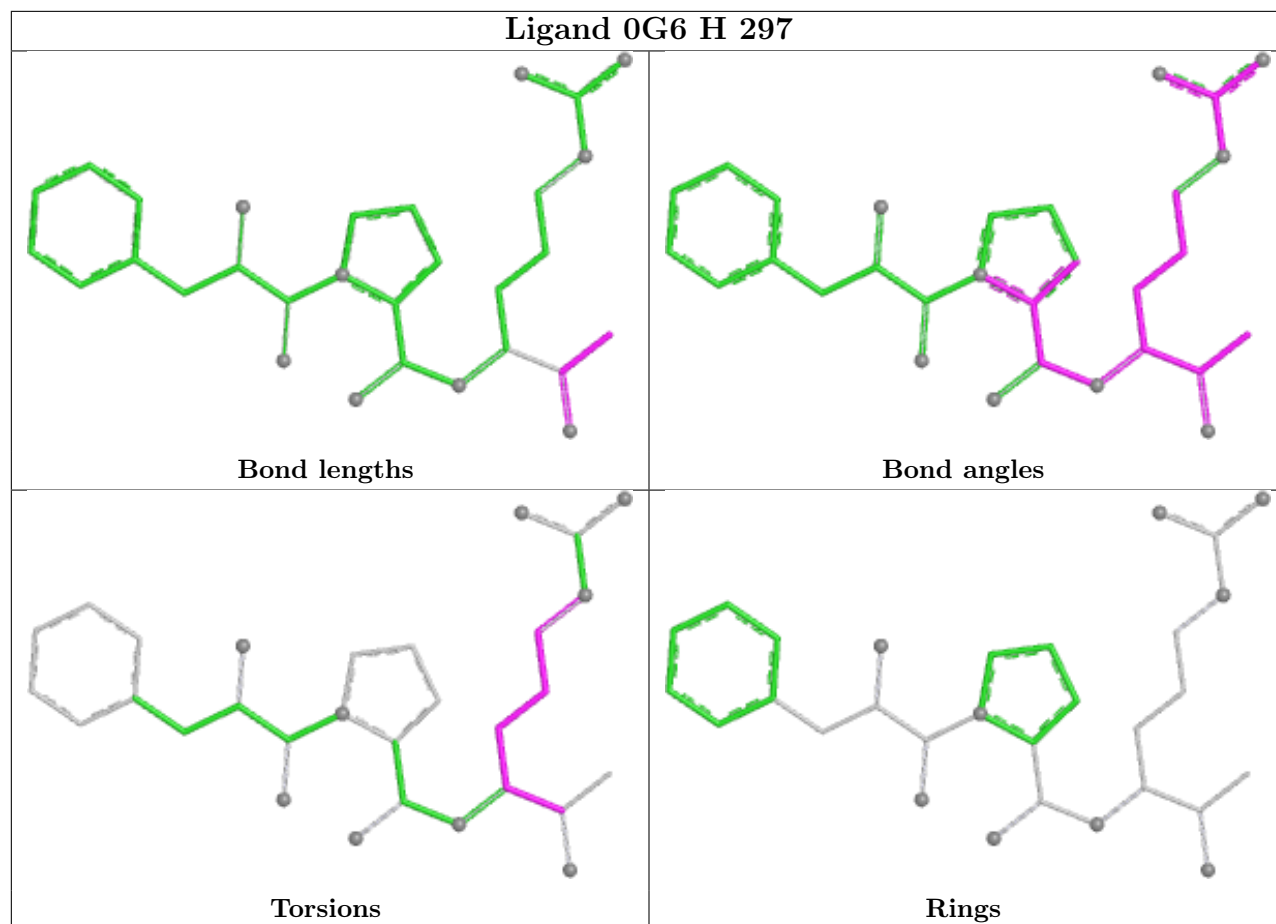
Mol	Chain	Res	Type	Atoms
4	H	297	0G6	C3-C2-CA2-CB2
4	H	297	0G6	N2-CA2-CB2-CG2
4	H	297	0G6	C2-CA2-CB2-CG2
4	H	297	0G6	NE-CD3-CG2-CB2

There are no ring outliers.

1 monomer is involved in 19 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	H	297	0G6	19	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
3	H	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	H	60(D):TRP	C	60(E):ASP	N	1.18
1	H	44:ALA	C	45:SER	N	1.17

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.