



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 12, 2024 – 08:56 PM EDT

PDB ID : 1HUT  
Title : THE STRUCTURE OF ALPHA-THROMBIN INHIBITED BY A 15-MER SINGLE-STRANDED DNA APTAMER  
Authors : Padmanabhan, K.; Padmanabhan, K.P.; Ferrara, J.D.; Sadler, J.E.; Tulinsky, A.  
Deposited on : 1993-05-27  
Resolution : 2.90 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 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	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
buster-report	:	1.1.7 (2018)
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)	:	2.36.2

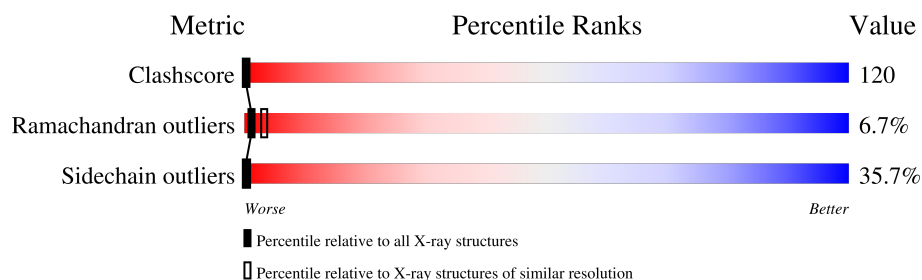
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)

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	D	15	100%
2	L	36	14% 39% 33% 14%
3	H	259	8% 47% 31% 12% .

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	OG7	H	1	-	-	X	-

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 2815 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			

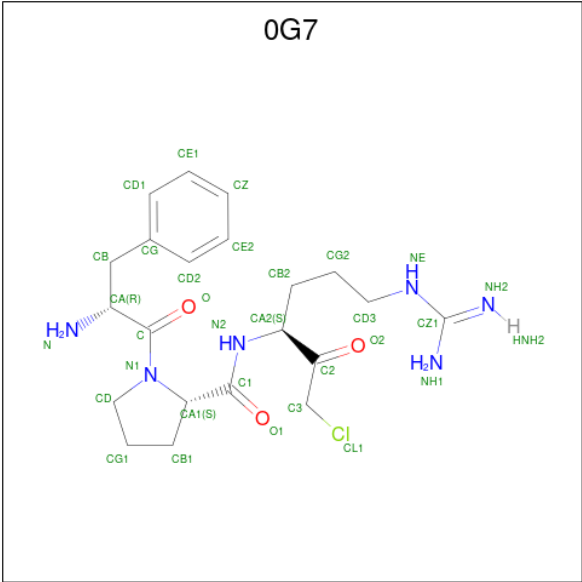
- Molecule 2 is a protein called ALPHA-Thrombin light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	36	Total	C	N	O	S	0	0	0
			287	177	48	61	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-[(3S)-6-carbamimidamido-1-chloro-2-oxohexan-3-yl]-L-prolinamide (three-letter code: 0G7) (formula: C<sub>21</sub>H<sub>31</sub>ClN<sub>6</sub>O<sub>3</sub>).



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	25	Total	O	0	0
			25	25		
5	L	15	Total	O	0	0
			15	15		
5	H	90	Total	O	0	0
			90	90		

### 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:  100%

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

- Molecule 2: ALPHA-Thrombin light chain

Chain L:  14% 39% 33% 14%

T1H  
F1G  
G1F  
S1E  
G1D  
E1C  
A1B  
D1A  
C1  
L3  
R4  
P5  
L6  
F7  
E8  
K9  
K10  
S11  
L12  
E13  
D14  
K14A  
T14B  
E14C  
R14D  
E14E  
L14F  
E14H  
S14I  
Y14J  
T14K  
D14L  
G14M  
R15

- Molecule 3: ALPHA-Thrombin heavy chain

Chain H:  8% 47% 31% 12%

I16  
V17  
E18  
G19  
S20  
D21  
A22  
E23  
I24  
G25  
R26  
S27  
P28  
V29  
Q30  
V31  
R32  
L33  
F34  
R35  
K36  
S36A  
P37  
Q38  
E39  
L40  
L41  
C42  
L43  
G43  
A44  
S45  
L46  
I47  
S48  
D49  
R50  
W51  
V52  
L53  
T54  
A55  
A56  
H57  
C58  
L59  
L60  
Y60A  
P60B  
P60C  
W60D  
D60E  
K60F  
N60G  
F60H  
Y17  
I18  
T60I  
E61  
N62  
D63  
L64  
L65

V66  
R67  
I68  
G69  
K70  
H71  
S72  
R73  
T74  
R75  
Y76  
E77  
R77A  
W78  
I79  
E80  
K81  
I82  
S83  
R84  
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  
E127  
T128  
A129  
A129A  
S129B  
L129C  
L130  
R131  
Q131  
Y134  
K135  
G136  
R137  
V138  
T139  
A190  
G140  
W141  
G142  
N143  
L144  
I145  
E146  
T147  
W148  
THR  
ALA  
ASN  
VAL  
GLY  
LYS  
G150  
Q151  
P152  
S153  
V154  
L155  
Q156  
V157  
V158  
N159  
L160  
P161  
I162  
V163  
E164  
R165  
P166  
V167  
C168  
K169  
D170  
S171  
T172  
R173  
I174  
T175  
L176

T177  
D178  
M179  
L234  
K235  
C236  
A237  
I238  
Q239  
K240  
V241  
I242  
G186C  
K186D  
R187  
G188  
D189  
A190  
C191  
E192  
G193  
D194  
S195  
G196  
G197  
F198  
F199  
V200  
M201  
K202  
S203  
P204  
F204A  
N205  
N206  
R206  
W207  
Y208  
Q209  
M210  
G211  
T212  
V213  
S214  
W215  
G216  
E217  
G219  
C220  
D221  
R221A  
D222  
G223  
K224  
Y225  
G226  
F227  
Y228  
T229  
H230

V231  
F232  
R233  
L234  
K235  
C236  
A237  
I238  
Q239  
K240  
V241  
I242  
G243  
Q244  
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, $\alpha$ , $\beta$ , $\gamma$	56.52Å 77.44Å 99.52Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	(Not available) – 2.90	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-2.90)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	X-PLOR	Depositor
R, $R_{free}$	0.159 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2815	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	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: 0G7

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	2.32	17/353 (4.8%)	5.35	121/547 (22.1%)
2	L	1.37	1/290 (0.3%)	2.37	14/384 (3.6%)
3	H	1.13	1/2107 (0.0%)	2.23	91/2846 (3.2%)
All	All	1.36	19/2750 (0.7%)	2.91	226/3777 (6.0%)

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
2	L	0	1
3	H	0	6
All	All	0	7

The worst 5 of 19 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	407	DT	C3'-C2'	9.18	1.63	1.52
1	D	404	DT	C4'-O4'	7.73	1.52	1.45
2	L	1(F)	GLY	C-N	7.53	1.51	1.34
1	D	401	DG	O4'-C1'	6.92	1.50	1.42
1	D	413	DT	O4'-C1'	6.77	1.50	1.42

The worst 5 of 226 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	406	DG	O4'-C1'-N9	33.65	131.56	108.00
3	H	101	ARG	NE-CZ-NH1	29.38	134.99	120.30
1	D	412	DT	O4'-C1'-N1	26.15	126.31	108.00

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	412	DT	P-O3'-C3'	24.52	149.12	119.70
1	D	404	DT	O4'-C1'-N1	22.93	124.05	108.00

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	H	36	LYS	Mainchain
3	H	63	ASP	Mainchain
3	H	64	LEU	Mainchain
3	H	81	LYS	Mainchain
2	L	14(K)	ILE	Mainchain

## 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	D	315	0	173	107	0
2	L	287	0	277	76	0
3	H	2053	0	2016	517	0
4	H	30	0	27	21	0
5	D	25	0	0	2	0
5	H	90	0	0	6	0
5	L	15	0	0	2	0
All	All	2815	0	2493	620	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 120.

The worst 5 of 620 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:D:408:DG:C8	3:H:75:ARG:HD2	1.43	1.53
1:D:407:DT:H3'	1:D:407:DT:C6	1.41	1.44
1:D:408:DG:H8	3:H:75:ARG:CD	1.34	1.41

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:407:DT:C6	1:D:407:DT:C3'	2.09	1.34
3:H:60(I):THR:O	3:H:63:ASP:HB2	1.22	1.28

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	34/36 (94%)	16 (47%)	12 (35%)	6 (18%)	0	0
3	H	249/259 (96%)	182 (73%)	54 (22%)	13 (5%)	2	6
All	All	283/295 (96%)	198 (70%)	66 (23%)	19 (7%)	1	3

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	L	1(G)	PHE
2	L	14(K)	ILE
3	H	116	ASP
3	H	186(C)	GLY
3	H	233	ARG

### 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	31/31 (100%)	21 (68%)	10 (32%)	0	0
3	H	221/225 (98%)	141 (64%)	80 (36%)	0	0
All	All	252/256 (98%)	162 (64%)	90 (36%)	0	0

5 of 90 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	H	147	THR
3	H	185	LYS
3	H	155	LEU
3	H	173	ARG
3	H	206	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
3	H	71	HIS
3	H	143	ASN
3	H	151	GLN
3	H	159	ASN

### 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 monosaccharides 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	0G7	H	1	-	31,31,32	2.25	3 (9%)	36,41,42	2.13	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	0G7	H	1	-	-	5/31/41/43	0/2/2/2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	H	1	0G7	C3-C2	-10.80	1.22	1.49
4	H	1	0G7	O1-C1	4.19	1.31	1.23
4	H	1	0G7	C1-N2	-2.83	1.27	1.34

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	H	1	0G7	CA1-C1-N2	7.32	132.90	116.58
4	H	1	0G7	O1-C1-N2	-4.37	114.84	122.93
4	H	1	0G7	CG-CB-CA	4.24	122.94	114.13
4	H	1	0G7	C1-CA1-N1	3.41	121.94	112.56
4	H	1	0G7	O1-C1-CA1	-3.37	112.23	120.63

There are no chirality outliers.

All (5) torsion outliers are listed below:

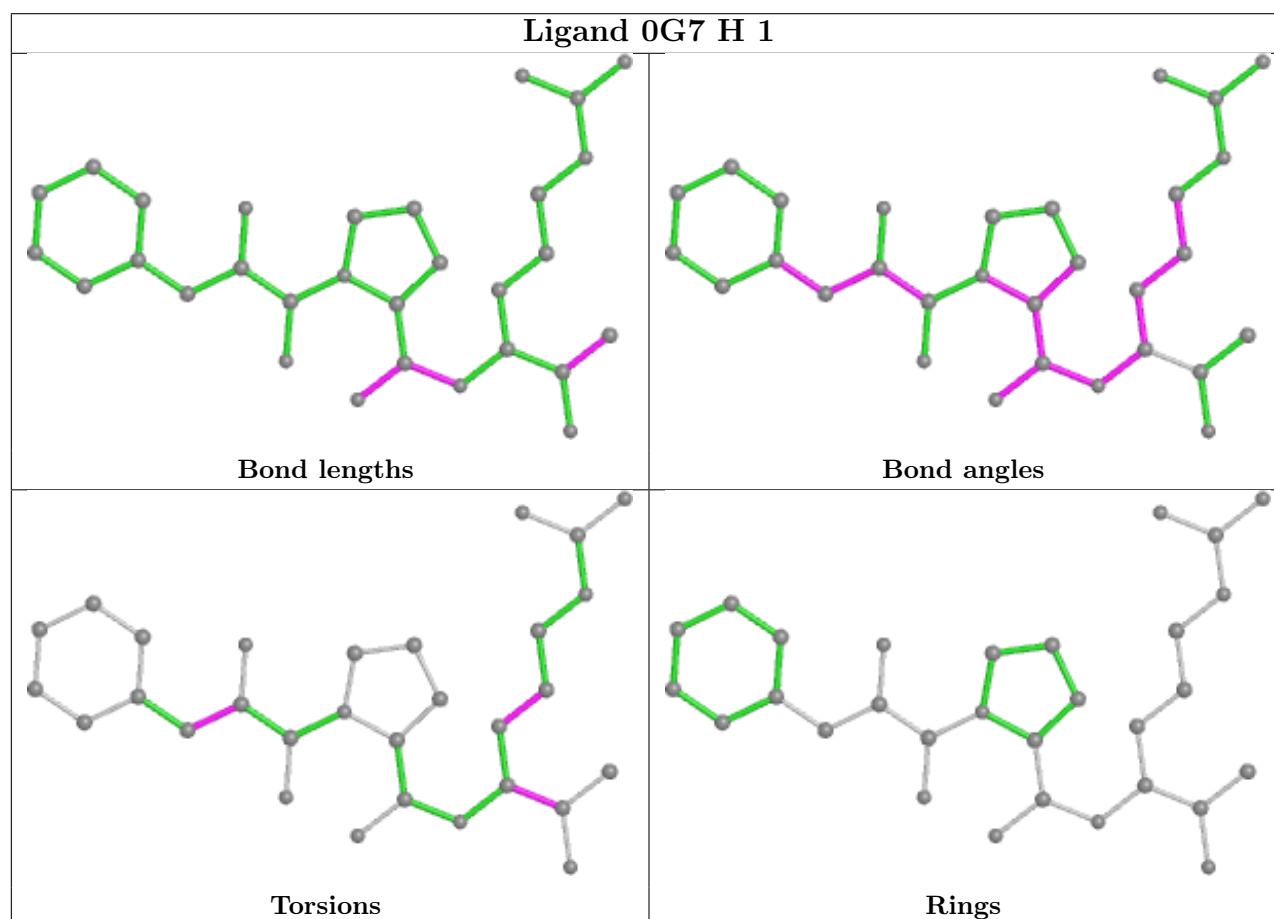
Mol	Chain	Res	Type	Atoms
4	H	1	0G7	C-CA-CB-CG
4	H	1	0G7	CA2-CB2-CG2-CD3
4	H	1	0G7	N-CA-CB-CG
4	H	1	0G7	C3-C2-CA2-CB2
4	H	1	0G7	O2-C2-CA2-CB2

There are no ring outliers.

1 monomer is involved in 21 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	H	1	0G7	21	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](#)

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.