



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 6, 2020 – 10:01 PM BST

PDB ID : 1UMA  
Title : ALPHA-THROMBIN (HIRUGEN) COMPLEXED WITH NA-(N,N-DIMETHYLCARBAMOYL)-ALPHA-AZALYSINE  
Authors : Nardini, M.; Pesce, A.; Rizzi, M.; Casale, E.; Ferraccioli, R.; Balliano, G.; Milla, P.; Ascenzi, P.; Bolognesi, M.  
Deposited on : 1996-03-26  
Resolution : 2.00 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) : 1.13  
EDS : 2.13.1  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.13.1

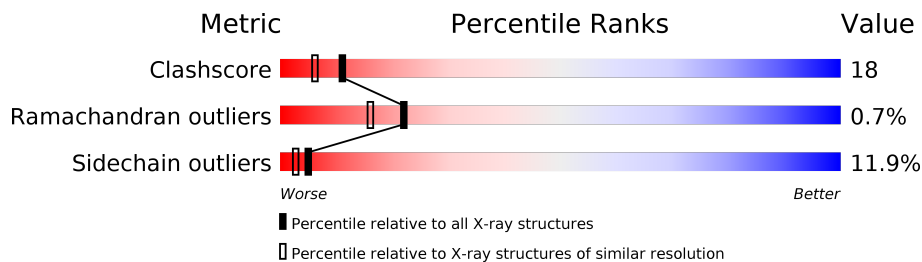
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)

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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	L	36	53% 33% 11% .
2	H	259	60% 29% 7% . .
3	I	10	40% 50% 10%

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 2658 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 protein called ALPHA-THROMBIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	L	36	287	177	48	61	1	57	0	0

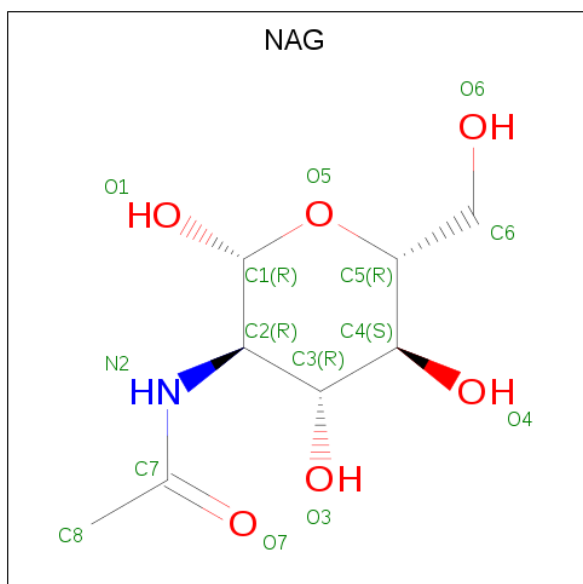
- Molecule 2 is a protein called ALPHA-THROMBIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	H	253	2053	1310	362	367	14	33	0	0

- Molecule 3 is a protein called HIRUDIN I.

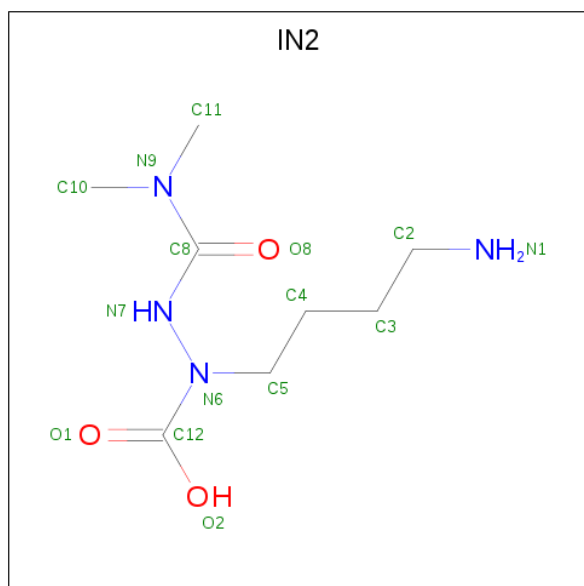
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	I	10	95	59	10	25	1	4	0	0

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	H	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 5 is N,N-DIMETHYLCARBAMOYL-ALPHA-AZALYSINE (three-letter code: IN2) (formula: C<sub>8</sub>H<sub>18</sub>N<sub>4</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	H	1	Total	C	N	O	0	0
			15	8	4	3		
5	H	1	Total	C	N	O	0	0
			14	8	4	2		

- Molecule 6 is water.

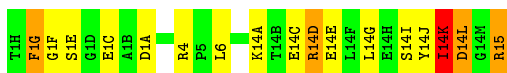
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	L	21	Total	O	1	0
			21	21		
6	H	157	Total	O	5	0
			157	157		
6	I	2	Total	O	0	0
			2	2		

### 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.

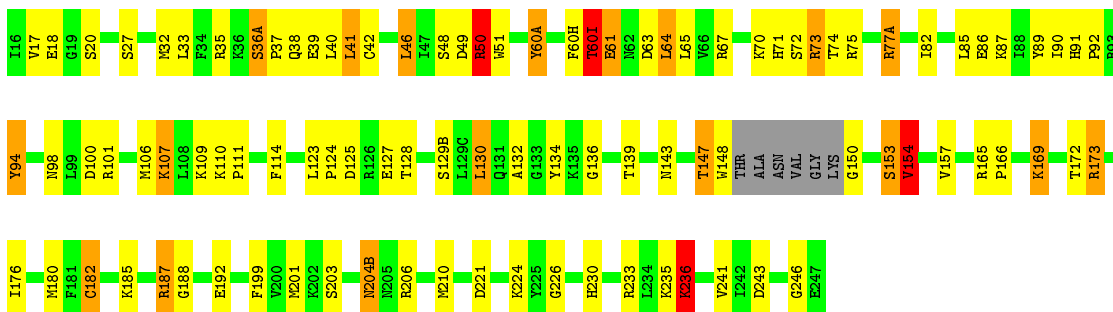
- Molecule 1: ALPHA-THROMBIN

Chain L: 



- Molecule 2: ALPHA-THROMBIN

Chain H: 



- Molecule 3: HIRUDIN I

Chain I: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	71.15Å 71.90Å 72.30Å 90.00° 100.49° 90.00°	Depositor
Resolution (Å)	20.00 – 2.00 20.48 – 2.00	Depositor EDS
% Data completeness (in resolution range)	(Not available) (20.00-2.00) 93.9 (20.48-2.00)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	6.85 (at 2.01Å)	Xtriage
Refinement program	TNT	Depositor
R, $R_{free}$	(Not available) , (Not available) 0.177 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	27.1	Xtriage
Anisotropy	0.246	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 97.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	2658	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.29% of the height of the origin peak. No significant pseudotranslation is detected.*

<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 i

### 5.1 Standard geometry i

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, IN2, TYS

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	L	6.28	2/290 (0.7%)	1.72	4/384 (1.0%)
2	H	1.25	6/2107 (0.3%)	1.61	37/2846 (1.3%)
3	I	9.00	1/79 (1.3%)	1.12	0/103
All	All	2.92	9/2476 (0.4%)	1.61	41/3333 (1.2%)

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
2	H	0	1

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	L	15	ARG	C-OXT	105.92	3.24	1.23
3	I	64	LEU	C-OXT	79.69	2.74	1.23
2	H	36(A)	SER	CB-OG	-35.86	0.95	1.42
2	H	153	SER	CB-OG	21.81	1.70	1.42
2	H	236	LYS	CE-NZ	-16.11	1.08	1.49

The worst 5 of 41 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	L	15	ARG	NE-CZ-NH1	14.04	127.32	120.30
2	H	101	ARG	NE-CZ-NH2	-12.00	114.30	120.30
2	H	77(A)	ARG	NE-CZ-NH2	-11.19	114.70	120.30
2	H	77(A)	ARG	NE-CZ-NH1	10.30	125.45	120.30
2	H	187	ARG	NE-CZ-NH1	-9.83	115.39	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	H	147	THR	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	L	287	0	278	16	3
2	H	2053	0	2018	72	3
3	I	95	0	73	5	0
4	H	14	0	13	0	0
5	H	29	0	34	1	0
6	H	157	0	0	7	0
6	I	2	0	0	1	0
6	L	21	0	0	0	0
All	All	2658	0	2416	85	3

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 18.

The worst 5 of 85 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:L:14(K):ILE:CD1	1:L:14(K):ILE:CG1	1.85	1.52
2:H:153:SER:CB	2:H:153:SER:OG	1.70	1.37
1:L:14(J):TYR:C	1:L:14(K):ILE:HG13	1.87	0.93
1:L:1(F):GLY:HA2	2:H:235:LYS:NZ	1.87	0.90
2:H:51:TRP:CZ2	2:H:107:LYS:HE3	2.12	0.84

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:L:15:ARG:OXT	2:H:173:ARG:O[4_556]	1.48	0.72

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:L:15:ARG:NH2	2:H:172:THR:O[4_556]	2.07	0.13
1:L:15:ARG:NH2	2:H:169:LYS:O[4_556]	2.18	0.02

### 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	
1	L	34/36 (94%)	25 (74%)	7 (21%)	2 (6%)	1	0
2	H	249/259 (96%)	239 (96%)	10 (4%)	0	100	100
3	I	7/10 (70%)	7 (100%)	0	0	100	100
All	All	290/305 (95%)	271 (93%)	17 (6%)	2 (1%)	22	16

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	L	14(L)	ASP
1	L	1(G)	PHE

#### 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	
1	L	31/31 (100%)	24 (77%)	7 (23%)	1	0
2	H	221/225 (98%)	197 (89%)	24 (11%)	6	3
3	I	9/9 (100%)	9 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	261/265 (98%)	230 (88%)	31 (12%)	<b>5</b> <b>3</b>

5 of 31 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	64	LEU
2	H	90	ILE
2	H	204(B)	ASN
2	H	65	LEU
2	H	94	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
2	H	78	ASN
2	H	156	GLN
2	H	204(B)	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](#)

1 non-standard protein/DNA/RNA residue 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
3	TYS	I	63	3	15,16,17	1.96	<b>2 (13%)</b>	18,22,24	<b>2.11</b>	<b>5 (27%)</b>

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	TYS	I	63	3	-	0/10/11/13	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	I	63	TYS	OH-S	-6.64	1.48	1.58
3	I	63	TYS	OH-CZ	-2.93	1.37	1.42

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
3	I	63	TYS	OH-CZ-CE1	4.78	128.03	118.64
3	I	63	TYS	CG-CB-CA	-4.52	104.94	114.10
3	I	63	TYS	OH-CZ-CE2	-3.93	110.92	118.64
3	I	63	TYS	OH-S-O2	2.44	114.83	107.71
3	I	63	TYS	CB-CA-C	-2.33	107.10	111.47

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	I	63	TYS	1	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

3 ligands 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$
4	NAG	H	400	2	14,14,15	0.83	0	17,19,21	2.15	3 (17%)
5	IN2	H	600	2	10,13,14	0.88	1 (10%)	13,15,17	0.93	0
5	IN2	H	500	-	9,14,14	0.99	1 (11%)	11,17,17	1.62	2 (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
4	NAG	H	400	2	-	2/6/23/26	0/1/1/1
5	IN2	H	600	2	-	1/10/15/17	-
5	IN2	H	500	-	-	3/11/17/17	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	H	600	IN2	C8-N9	-2.39	1.32	1.35
5	H	500	IN2	C8-N9	-2.17	1.32	1.35

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	H	400	NAG	O3-C3-C4	-5.53	97.56	110.35
4	H	400	NAG	C2-N2-C7	-5.14	115.59	122.90
5	H	500	IN2	C8-N7-N6	-3.46	112.58	119.93
5	H	500	IN2	C4-C5-N6	2.96	119.81	113.35
4	H	400	NAG	O7-C7-C8	2.13	126.02	122.06

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	H	400	NAG	C8-C7-N2-C2
5	H	500	IN2	C4-C5-N6-C12
5	H	500	IN2	O8-C8-N7-N6
4	H	400	NAG	O7-C7-N2-C2
5	H	600	IN2	C2-C3-C4-C5

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	H	500	IN2	1	0

## 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

### 6.1 Protein, DNA and RNA chains

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

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

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

### 6.3 Carbohydrates

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

### 6.4 Ligands

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

### 6.5 Other polymers

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