



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 7, 2020 – 06:23 AM BST

PDB ID : 4THN  
Title : THE CRYSTAL STRUCTURE OF ALPHA-THROMBIN-HIRUNORM IV COMPLEX REVEALS A NOVEL SPECIFICITY SITE RECOGNITION MODE.  
Authors : Lombardi, A.; De Simone, G.; Nastri, F.; Galdiero, S.; Della Morte, R.; Stasiano, N.; Pedone, C.; Bolognesi, M.; Pavone, V.  
Deposited on : 1998-09-18  
Resolution : 2.50 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) : 2.13.1

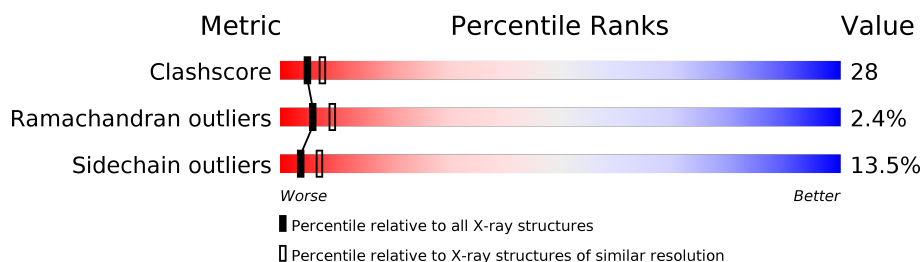
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)

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	L	36	
2	H	259	
3	I	26	

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	NAG	H	400	X	-	-	-

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 2566 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
1	L	36	Total	C	N	O	S	86	0	0
			287	177	48	61	1			

- Molecule 2 is a protein called ALPHA-THROMBIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	253	Total	C	N	O	S	120	0	0
			2053	1310	362	367	14			

- Molecule 3 is a protein called HIRUNORM IV.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	I	14	Total	C	N	O	43	0	0
			134	92	17	25			

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	L	9	Total	O	0	0
			9	9		
5	H	67	Total	O	0	0
			67	67		
5	I	2	Total	O	0	0
			2	2		

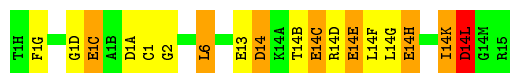
### 3 Residue-property plots [i](#)

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.

Note EDS was not executed.

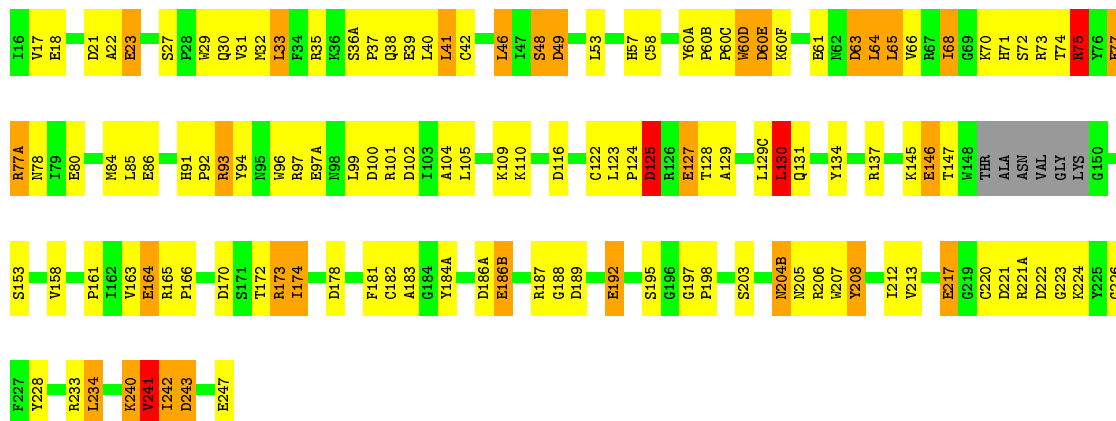
#### • Molecule 1: ALPHA-THROMBIN

Chain L: 




#### • Molecule 2: ALPHA-THROMBIN

Chain H: 



#### • Molecule 3: HIRUNORM IV

Chain I: 



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	72.30 Å   72.70 Å   73.60 Å 90.00°   100.00°   90.00°	Depositor
Resolution (Å)	20.00 – 2.50	Depositor
% Data completeness (in resolution range)	96.0 (20.00-2.50)	Depositor
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	TNT 5E	Depositor
R, $R_{free}$	0.173 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2566	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	47.0	wwPDB-VP

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, CHG, ALC, NAL, DGL, AIB

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	1.24	5/290 (1.7%)	1.45	5/384 (1.3%)
2	H	1.09	13/2107 (0.6%)	1.47	38/2846 (1.3%)
3	I	1.44	2/77 (2.6%)	1.84	3/102 (2.9%)
All	All	1.12	20/2474 (0.8%)	1.48	46/3332 (1.4%)

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 20 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	18	GLU	CD-OE2	9.60	1.36	1.25
2	H	217	GLU	CD-OE2	9.47	1.36	1.25
1	L	14(H)	GLU	CD-OE2	8.23	1.34	1.25
2	H	192	GLU	CD-OE2	7.81	1.34	1.25
2	H	77	GLU	CD-OE2	7.57	1.33	1.25

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	116	ASP	CB-CG-OD1	8.74	126.17	118.30
2	H	221	ASP	CB-CG-OD2	-8.73	110.44	118.30
2	H	137	ARG	NE-CZ-NH1	8.47	124.54	120.30
3	I	2	ARG	NE-CZ-NH1	8.30	124.45	120.30
2	H	49	ASP	CB-CG-OD2	-8.07	111.03	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	H	75	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	L	287	0	278	13	3
2	H	2053	0	2019	107	3
3	I	134	0	123	8	0
4	H	14	0	11	0	0
5	H	67	0	0	5	0
5	I	2	0	0	0	0
5	L	9	0	0	0	0
All	All	2566	0	2431	120	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 28.

The worst 5 of 120 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(D):ARG:O	1:L:14(H):GLU:HG3	1.69	0.92
2:H:32:MET:HE3	2:H:70:LYS:HD3	1.52	0.91
2:H:32:MET:HG3	2:H:40:LEU:HD13	1.61	0.82
2:H:64:LEU:HD12	2:H:85:LEU:HD11	1.62	0.81
2:H:233:ARG:C	2:H:234:LEU:HD23	2.01	0.81

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:14(L):ASP:O	2:H:173:ARG:NH2[4_556]	1.22	0.98
1:L:14(L):ASP:C	2:H:173:ARG:NH2[4_556]	2.05	0.15

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:L:14(L):ASP:O	2:H:173:ARG:CZ[4_556]	2.13	0.07

## 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%)	27 (79%)	3 (9%)	4 (12%)	0	0
2	H	249/259 (96%)	226 (91%)	20 (8%)	3 (1%)	13	24
3	I	7/26 (27%)	7 (100%)	0	0	100	100
All	All	290/321 (90%)	260 (90%)	23 (8%)	7 (2%)	6	9

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	H	240	LYS
1	L	1(D)	GLY
1	L	1(G)	PHE
1	L	14(K)	ILE
1	L	14(L)	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	
1	L	31/31 (100%)	26 (84%)	5 (16%)	2	4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	H	221/225 (98%)	193 (87%)	28 (13%)	4	8
3	I	8/14 (57%)	6 (75%)	2 (25%)	0	1
All	All	260/270 (96%)	225 (86%)	35 (14%)	4	7

5 of 35 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	75	ARG
2	H	130	LEU
2	H	242	ILE
2	H	110	LYS
2	H	125	ASP

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

Mol	Chain	Res	Type
2	H	71	HIS
2	H	78	ASN
2	H	143	ASN
2	H	204(B)	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

6 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	$\# Z  > 2$
3	AIB	I	23	3	1,5,6	1.10	0	2,7,9	0.09	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	CHG	I	1	3	9,10,11	1.18	1 (11%)	7,12,14	1.31	1 (14%)
3	NAL	I	3	3	15,16,17	1.21	1 (6%)	18,21,23	1.02	1 (5%)
3	ALC	I	25	3	9,11,12	0.51	0	10,13,15	1.25	1 (10%)
3	AIB	I	22	3	1,5,6	1.69	0	2,7,9	0.04	0

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	AIB	I	23	3	-	2/2/3/6	-
3	CHG	I	1	3	-	4/5/14/16	0/1/1/1
3	NAL	I	3	3	-	0/5/6/8	0/2/2/2
3	ALC	I	25	3	-	1/5/14/16	0/1/1/1
3	AIB	I	22	3	-	0/2/3/6	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	I	1	CHG	C1-CA	3.21	1.59	1.53
3	I	3	NAL	C9-CA	-2.51	1.48	1.53

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	I	25	ALC	CG-CB-CA	-2.36	111.35	114.52
3	I	1	CHG	C5-C6-C1	-2.18	108.21	111.93
3	I	3	NAL	C9-CA-C	-2.01	107.71	111.47

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	I	1	CHG	C2-C1-CA-N
3	I	1	CHG	C6-C1-CA-N
3	I	1	CHG	C6-C1-CA-C
3	I	23	AIB	O-C-CA-CB1
3	I	1	CHG	C2-C1-CA-C

There are no ring outliers.

4 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	I	23	AIB	1	0
3	I	1	CHG	4	0
3	I	3	NAL	1	0
3	I	22	AIB	1	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

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	NAG	H	400	2	14,14,15	1.08	1 (7%)	17,19,21	4.42	12 (70%)

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	2/2/5/7	4/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	H	400	NAG	C2-N2	2.68	1.50	1.46

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	H	400	NAG	O5-C5-C6	9.24	121.69	107.20
4	H	400	NAG	O7-C7-N2	-6.39	110.20	121.95
4	H	400	NAG	C3-C4-C5	5.98	120.90	110.24
4	H	400	NAG	O3-C3-C2	5.87	121.62	109.47
4	H	400	NAG	O7-C7-C8	-5.65	111.56	122.06

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	H	400	NAG	C4
4	H	400	NAG	C2

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	H	400	NAG	C3-C2-N2-C7
4	H	400	NAG	C8-C7-N2-C2
4	H	400	NAG	O7-C7-N2-C2
4	H	400	NAG	O5-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

## 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

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.