wwPDB X-ray Structure Validation Summary Report (i)

Apr 16, 2024 – 10:00 AM EDT

PDB ID	:	3HAT
Title	:	ACTIVE SITE MIMETIC INHIBITION OF THROMBIN
Authors	:	Tulinsky, A.; Mathews, I.I.
Deposited on	:	1994-10-16
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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

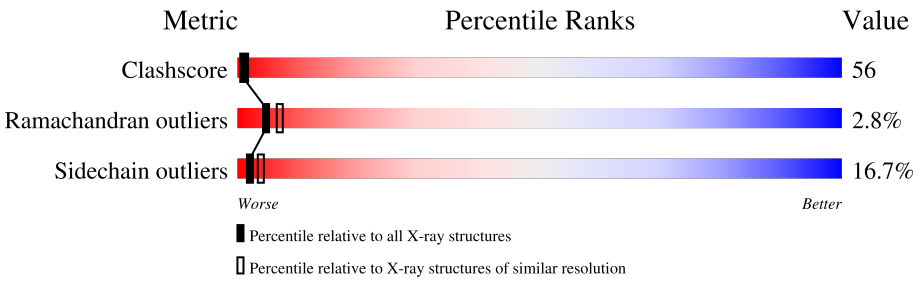
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.1

1 Overall quality at a glance (i)

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 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 <=5%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	L	36	17% 42% 31% 11%
2	H	259	25% 46% 22% . .
3	I	12	25% 33% 25% 17%
4	T	4	75% 25%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 2575 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 Thrombin light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	L	36	281	174	45	61	1	0	0	0

- Molecule 2 is a protein called Thrombin heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	H	251	1996	1271	355	356	14	0	0	0

- Molecule 3 is a protein called Hirudin variant-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	I	10	87	56	10	20	1	0	0	0

- Molecule 4 is a protein called FPAM (FIBRINOPEPTIDE A MIMIC).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	T	4	32	19	8	5	0	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	L	23	23	23	0	0
5	H	146	146	146	0	0
5	I	8	8	8	0	0
5	T	2	2	2	0	0

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: Thrombin light chain

Chain L: 17% 42% 31% 11%

T1H F1G G1F S1E G1D E1C A1B D1A C1 G2 L3 R4 P5 L6 F7 E8 K9 K10 D14 K14A T14B E14C R14D E14E L14F L14G E14H S14I Y14J I14K D14L G14M R15

• Molecule 2: Thrombin heavy chain

Chain H: 25% 46% 22%

I16 V17 E18 G19 S20 D21 I24 S27 P28 W29 Q30 V31 M32 L33 F34 R35 K36 S36A P37 Q38 E39 L40 L41 C42 L46 I47 S48 D49 R50 W51 V52 L53 T54 A55 A56 H57 C58 L59 L60 Y60A P60B P60C W60D D60E K60F N60G F60H T60I E61 N62 D63 L64 L65 V66 R67 K70

H71 S72 R73 T74 R75 Y76 E77 R77A N78 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 V112 A113 F114 Y117 I118 H119 P120 V121 C122 L123 P124 D125 R126 E127 T128 A129

L129C A132 G133 Y134 K135 G136 R137 W141 G142 N143 L144 K145 E146 THR TRP 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 R175 I176 T177 D178 F181 C182 A183 G184 Y184A K185 P186

D186A E186B G186C K186D R187 G188 D189 A190 C191 E192 G193 D194 S195 P198 F199 V200 M201 K202 S203 P204 F204A N204B N205 R206 W207 Y208 Q209 M210 G211 I212 V213 S214 W215 G216 E217 G219 C220 D221 R221A D222 G223 K224 Y225 G226 F227 Y228 T229 H230 V231 F232 R233 L234 K235 K236 W237 I238 Q239 K240 V241

I242 D243 Q244 F245 G246 E247

• Molecule 3: Hirudin variant-2

Chain I: 25% 33% 25% 17%

ASN GLY D55 E58 I59 P60 E61 E62 Y63 L64

• Molecule 4: FPAM (FIBRINOPEPTIDE A MIMIC)

Chain T: 75% 25%

RN G303
G304
V305
R306

4 Data and refinement statistics (i)

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, α, β, γ	71.10Å 72.40Å 73.00Å 90.00° 101.00° 90.00°	Depositor
Resolution (Å)	(Not available) – 2.50	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-2.50)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	PROLSQ	Depositor
R, R_{free}	0.140 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2575	wwPDB-VP
Average B, all atoms ($Å^2$)	28.0	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: TYS, RNG

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	1.31	1/284 (0.4%)	2.76	21/377 (5.6%)
2	H	1.19	2/2047 (0.1%)	2.65	128/2764 (4.6%)
3	I	1.15	0/71	2.15	1/93 (1.1%)
4	T	1.53	0/22	3.22	3/26 (11.5%)
All	All	1.21	3/2424 (0.1%)	2.65	153/3260 (4.7%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	27	SER	CB-OG	-5.84	1.34	1.42
2	H	58	CYS	CB-SG	-5.53	1.72	1.81
1	L	14(C)	GLU	CD-OE2	-5.44	1.19	1.25

The worst 5 of 153 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	H	75	ARG	NE-CZ-NH1	24.86	132.73	120.30
2	H	77(A)	ARG	NE-CZ-NH2	-23.46	108.57	120.30
2	H	173	ARG	NE-CZ-NH2	21.25	130.92	120.30
1	L	4	ARG	NE-CZ-NH1	19.51	130.05	120.30
2	H	126	ARG	CD-NE-CZ	19.50	150.90	123.60

There are no chirality outliers.

There are no planarity outliers.

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	281	0	267	68	3
2	H	1996	0	1955	223	3
3	I	87	0	68	9	0
4	T	32	0	34	16	0
5	H	146	0	0	19	0
5	I	8	0	0	1	0
5	L	23	0	0	2	0
5	T	2	0	0	0	0
All	All	2575	0	2324	263	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 56.

The worst 5 of 263 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)
2:H:224:LYS:HE2	5:H:465:HOH:O	1.40	1.19
1:L:1(F):GLY:O	1:L:1(E):SER:HB3	1.46	1.14
1:L:15:ARG:NH1	2:H:204(A):PHE:O	1.83	1.08
1:L:1(H):THR:HG22	1:L:1(D):GLY:O	1.51	1.08
2:H:201:MET:CE	2:H:210:MET:HG3	1.82	1.08

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:NH1[4_556]	2.07	0.13
1:L:14(M):GLY:O	2:H:173:ARG:NH1[4_556]	2.09	0.11
1:L:15:ARG:CZ	2:H:172:THR: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%)	21 (62%)	7 (21%)	6 (18%)	0	0
2	H	247/259 (95%)	228 (92%)	17 (7%)	2 (1%)	19	35
3	I	7/12 (58%)	7 (100%)	0	0	100	100
4	T	2/4 (50%)	1 (50%)	1 (50%)	0	100	100
All	All	290/311 (93%)	257 (89%)	25 (9%)	8 (3%)	5	7

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	L	1(F)	GLY
1	L	1(E)	SER
1	L	1(D)	GLY
1	L	1(C)	GLU
1	L	14(M)	GLY

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	30/31 (97%)	23 (77%)	7 (23%)	1	1
2	H	212/225 (94%)	178 (84%)	34 (16%)	2	4
3	I	7/10 (70%)	6 (86%)	1 (14%)	3	6
4	T	2/2 (100%)	2 (100%)	0	100	100
All	All	251/268 (94%)	209 (83%)	42 (17%)	2	4

5 of 42 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	112	VAL
2	H	202	LYS
2	H	127	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	H	182	CYS
2	H	233	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
2	H	156	GLN
2	H	204(B)	ASN
2	H	239	GLN
2	H	60(G)	ASN
2	H	38	GLN

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	2.40	2 (13%)	18,22,24	2.00	5 (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
3	TYS	I	63	3	-	2/10/11/13	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	7.75	1.70	1.58
3	I	63	TYS	OH-CZ	-4.25	1.35	1.42

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	I	63	TYS	CG-CB-CA	-4.49	105.01	114.10
3	I	63	TYS	O2-S-O1	4.12	128.75	112.22
3	I	63	TYS	OH-S-O2	-3.15	98.52	107.71
3	I	63	TYS	CB-CG-CD1	-2.30	116.35	120.91
3	I	63	TYS	CD2-CG-CD1	2.13	121.52	118.17

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	I	63	TYS	CA-CB-CG-CD1
3	I	63	TYS	CA-CB-CG-CD2

There are no ring outliers.

1 monomer is involved in 3 short contacts:

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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.