wwPDB X-ray Structure Validation Summary Report (i)

Aug 8, 2023 – 06:28 AM EDT

PDB ID	:	1NY2
Title	:	Human alpha thrombin inhibited by RPPGF and hirugen
Authors	:	Krishnan, R.; Tulinsky, A.; Schmaier, A.H.; Hasan, A.A.; Warnock, M.; Srikanth, S.; Mahdi, F.
Deposited on	:	2003-02-11
Resolution	:	2.30 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:

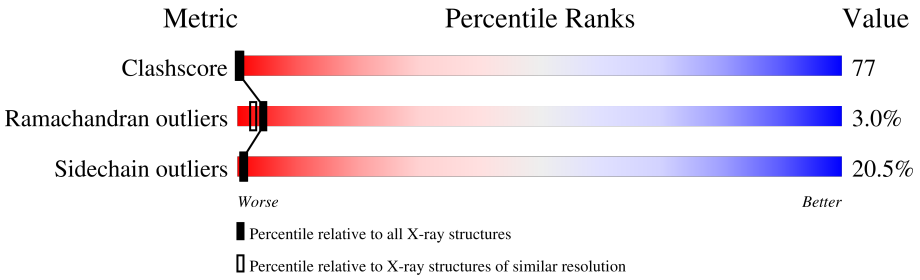
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 2.30 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	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)

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	1	36	8% 33% 44% 14%
2	2	259	19% 50% 25% 6%
3	3	10	10% 20% 60% 10%
4	4	5	20% 20% 20% 20% 20%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 2554 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
1	1	36	Total	C	N	O	S	0	0	0
			287	177	48	61	1			

- Molecule 2 is a protein called thrombin Heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	2	259	Total	C	N	O	S	0	0	0
			2093	1334	370	375	14			

- Molecule 3 is a protein called Hirugen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	3	10	Total	C	N	O	S	0	0	0
			94	59	10	24	1			

- Molecule 4 is a protein called Inhibitor peptide RPPGF.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	4	4	Total	C	N	O	0	0	0
			29	18	7	4			

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	1	3	Total	O	0	0
			3	3		
5	2	44	Total	O	0	0
			44	44		
5	3	4	Total	O	0	0
			4	4		

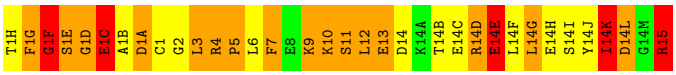
3 Residue-property plots (i)

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.

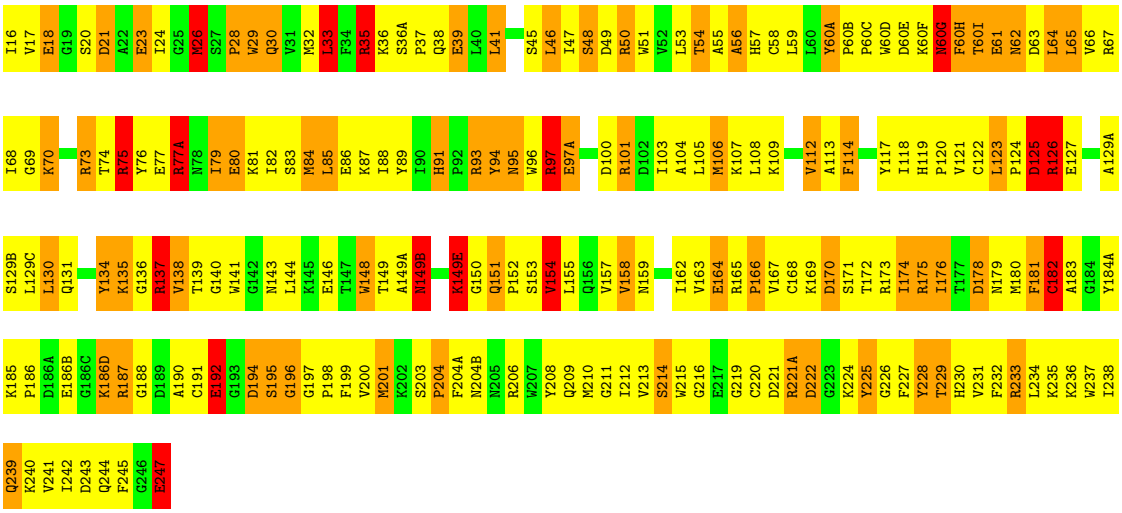
- Molecule 1: thrombin light chain

Chain 1:



- Molecule 2: thrombin Heavy chain

Chain 2:

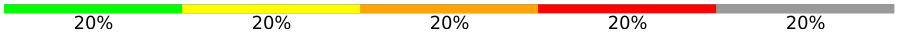


- Molecule 3: Hirugen

Chain 3:

- Molecule 4: Inhibitor peptide RPPGF

Chain 4:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	79.15Å 104.97Å 45.18Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	8.00 – 2.30	Depositor
% Data completeness (in resolution range)	(Not available) (8.00-2.30)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	PROLSQ	Depositor
R, R_{free}	0.202 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2554	wwPDB-VP
Average B, all atoms (Å^2)	24.0	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: 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	1	1.07	0/290	2.46	18/384 (4.7%)
2	2	1.18	2/2148 (0.1%)	2.58	123/2903 (4.2%)
3	3	1.14	0/78	2.45	6/103 (5.8%)
4	4	1.52	0/30	3.43	3/40 (7.5%)
All	All	1.17	2/2546 (0.1%)	2.57	150/3430 (4.4%)

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	2	0	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	2	75	ARG	NE-CZ	6.18	1.41	1.33
2	2	75	ARG	CD-NE	-5.66	1.36	1.46

The worst 5 of 150 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	2	75	ARG	CD-NE-CZ	30.67	166.53	123.60
2	2	206	ARG	CD-NE-CZ	25.83	159.76	123.60
2	2	73	ARG	NE-CZ-NH1	-23.82	108.39	120.30
2	2	175	ARG	CD-NE-CZ	20.54	152.35	123.60
2	2	175	ARG	NE-CZ-NH1	18.88	129.74	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	2	192	GLU	Mainchain
2	2	233	ARG	Sidechain

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	1	287	0	278	66	0
2	2	2093	0	2064	336	0
3	3	94	0	73	4	0
4	4	29	0	29	5	0
5	1	3	0	0	1	0
5	2	44	0	0	10	0
5	3	4	0	0	0	0
All	All	2554	0	2444	380	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 77.

The worst 5 of 380 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)
2:2:195:SER:HG	4:4:380:ARG:N	1.26	1.30
2:2:165:ARG:HB2	2:2:166:PRO:HD3	1.27	1.17
2:2:81:LYS:HD2	2:2:118:ILE:CD1	1.77	1.15
2:2:81:LYS:HD2	2:2:118:ILE:HD12	1.20	1.09
2:2:149(E):LYS:HE3	2:2:150:GLY:H	1.11	1.07

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
1	1	34/36 (94%)	22 (65%)	6 (18%)	6 (18%)	0 0
2	2	257/259 (99%)	235 (91%)	19 (7%)	3 (1%)	13 14
3	3	7/10 (70%)	5 (71%)	2 (29%)	0	100 100
4	4	2/5 (40%)	2 (100%)	0	0	100 100
All	All	300/310 (97%)	264 (88%)	27 (9%)	9 (3%)	4 2

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	1	1(D)	GLY
1	1	1(C)	GLU
1	1	14(K)	ILE
2	2	97	ARG
1	1	1(B)	ALA

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	1	31/31 (100%)	24 (77%)	7 (23%)	1 0
2	2	225/225 (100%)	183 (81%)	42 (19%)	1 1
3	3	9/9 (100%)	5 (56%)	4 (44%)	0 0
4	4	3/4 (75%)	1 (33%)	2 (67%)	0 0
All	All	268/269 (100%)	213 (80%)	55 (20%)	1 1

5 of 55 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	2	130	LEU
2	2	151	GLN
4	4	382	PRO
3	3	55	ASP
2	2	134	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
2	2	60(G)	ASN
2	2	131	GLN
2	2	204(B)	ASN
2	2	239	GLN
2	2	244	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	3	63	3	15,16,17	1.63	2 (13%)	18,22,24	1.83	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	3	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	3	63	TYS	OH-S	4.62	1.65	1.58
3	3	63	TYS	OH-CZ	-3.20	1.37	1.42

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	3	63	TYS	CB-CA-C	-3.24	105.40	111.47
3	3	63	TYS	CD1-CE1-CZ	-3.17	115.86	119.73
3	3	63	TYS	CG-CB-CA	-3.10	107.82	114.10
3	3	63	TYS	CB-CG-CD1	-2.72	115.52	120.91
3	3	63	TYS	CE2-CZ-CE1	2.27	123.68	120.18

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	3	63	TYS	1	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

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.