wwPDB X-ray Structure Validation Summary Report (i)

May 14, 2020 – 01:03 pm BST

PDB ID	:	1BIH
Title	:	CRYSTAL STRUCTURE OF THE INSECT IMMUNE PROTEIN HEMOLIN: A NEW DOMAIN ARRANGEMENT WITH IMPLICATIONS FOR HOMOPHILIC ADHESION
Authors	:	Su, X.-D.; Gastinel, L.N.; Vaughn, D.E.; Faye, I.; Poon, P.; Bjorkman, P.J.
Deposited on	:	1998-06-17
Resolution	:	3.10 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 following versions of software and data (see references (i)) were used in the production of this report:

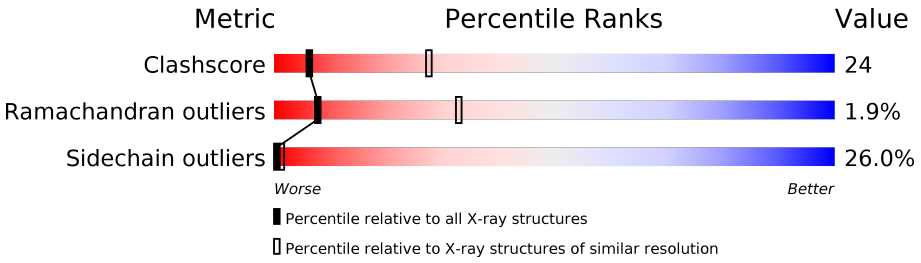
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 3.10 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	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)

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 >=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	A	395	49% 39% 11% .
1	B	395	49% 41% 9% .

2 Entry composition

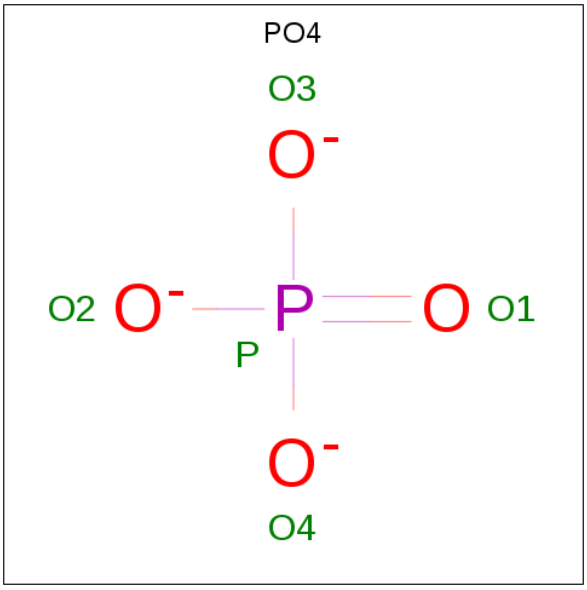
There are 2 unique types of molecules in this entry. The entry contains 6106 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 HEMOLIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	391	Total	C	N	O	S	0	0	0
			3048	1923	523	589	13			
1	B	391	Total	C	N	O	S	6	0	0
			3048	1923	523	589	13			

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	P	0	0
			5	4	1		
2	B	1	Total	O	P	0	0
			5	4	1		

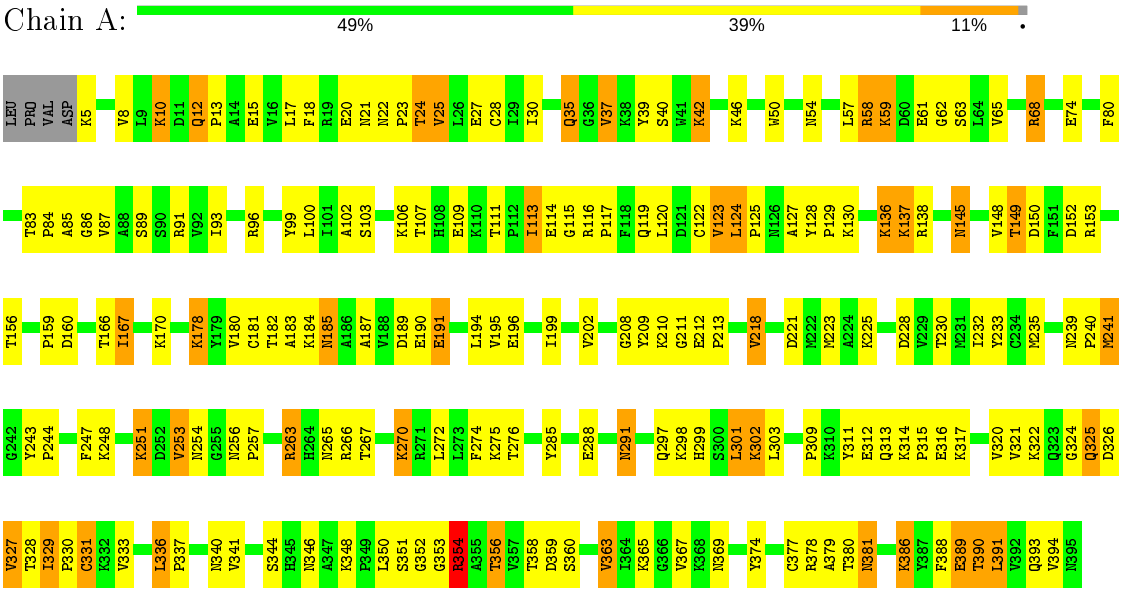
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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.

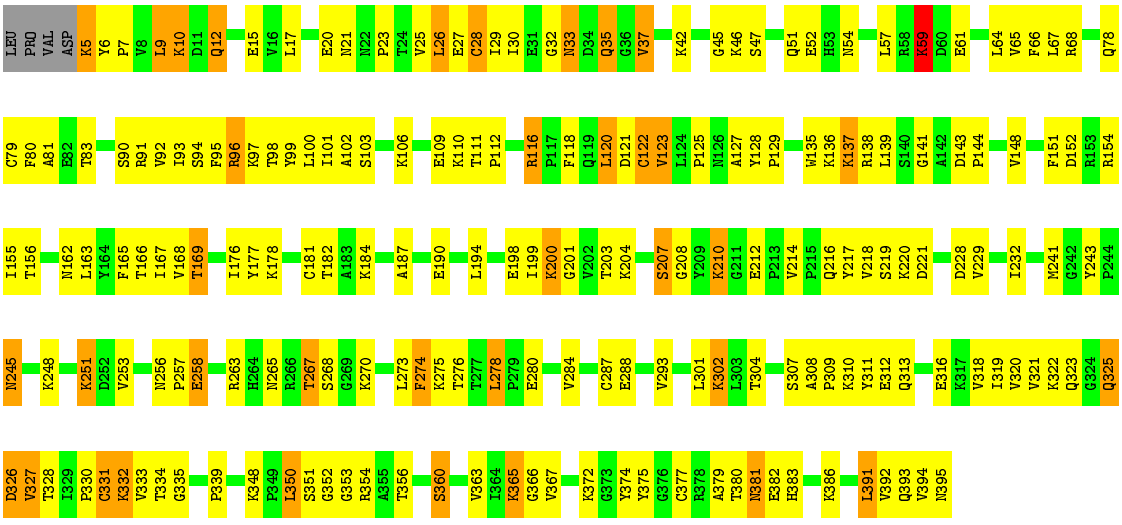
- Molecule 1: HEMOLIN

Chain A:



- Molecule 1: HEMOLIN

Chain B:



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 21	Depositor
Cell constants a, b, c, α, β, γ	85.00Å 90.30Å 143.10Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.10	Depositor
% Data completeness (in resolution range)	93.0 (20.00-3.10)	Depositor
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
Refinement program	X-PLOR 3.857	Depositor
R, R_{free}	0.218 , 0.264	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	6106	wwPDB-VP
Average B, all atoms (Å^2)	21.0	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: PO4

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	A	0.52	0/3119	0.68	0/4233
1	B	0.50	0/3119	0.68	0/4233
All	All	0.51	0/6238	0.68	0/8466

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	A	3048	0	2992	157	0
1	B	3048	0	2992	140	0
2	A	5	0	0	1	0
2	B	5	0	0	0	0
All	All	6106	0	5984	290	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 24.

The worst 5 of 290 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:B:96:ARG:HB3	1:B:96:ARG:HH11	1.28	0.95
1:B:122:CYS:SG	1:B:181:CYS:SG	2.64	0.92
1:A:309:PRO:HD3	1:A:381:ASN:HD21	1.31	0.91
1:B:122:CYS:HG	1:B:181:CYS:HG	0.89	0.88
1:A:288:GLU:HG3	1:A:298:LYS:HG2	1.56	0.87

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	A	389/395 (98%)	351 (90%)	31 (8%)	7 (2%)	8 34
1	B	389/395 (98%)	346 (89%)	35 (9%)	8 (2%)	7 30
All	All	778/790 (98%)	697 (90%)	66 (8%)	15 (2%)	8 33

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	59	LYS
1	A	208	GLY
1	B	59	LYS
1	B	141	GLY
1	B	208	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	A	333/340 (98%)	248 (74%)	85 (26%)	0 1
1	B	333/340 (98%)	245 (74%)	88 (26%)	0 1
All	All	666/680 (98%)	493 (74%)	173 (26%)	0 1

5 of 173 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	363	VAL
1	B	59	LYS
1	B	327	VAL
1	A	381	ASN
1	B	10	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	345	HIS
1	A	381	ASN
1	B	216	GLN
1	A	325	GLN
1	B	245	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)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

2 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	$\#\|Z\| > 2$
2	PO4	B	396	-	4,4,4	0.82	0	6,6,6	0.64	0
2	PO4	A	396	-	4,4,4	0.51	0	6,6,6	0.86	0

There are no bond length outliers.

There are no bond angle outliers.

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
2	A	396	PO4	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

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.