wwPDB X-ray Structure Validation Summary Report (i)

Aug 2, 2023 – 10:43 PM EDT

PDB ID	:	1ICF
Title	:	CRYSTAL STRUCTURE OF MHC CLASS II ASSOCIATED P41 II FRAGMENT IN COMPLEX WITH CATHEPSIN L
Authors	:	Guncar, G.; Pungercic, G.; Klemencic, I.; Turk, V.; Turk, D.
Deposited on	:	1999-01-07
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*
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:

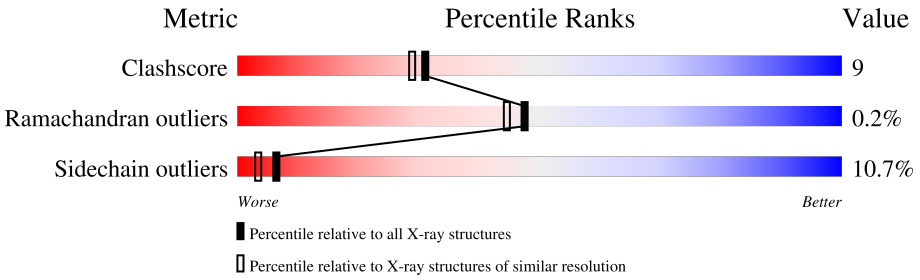
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.34

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 2.00 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 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	A	175	74% 23%
1	C	175	81% 15%
2	B	42	55% 40% 5%
2	D	42	76% 21%
3	I	65	75% 20% 5%
3	J	65	75% 20% 5%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 5060 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 PROTEIN (CATHEPSIN L: HEAVY CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	175	Total	C	N	O	S	35	0	0
			1339	837	217	275	10			
1	C	175	Total	C	N	O	S	36	0	0
			1339	837	217	275	10			

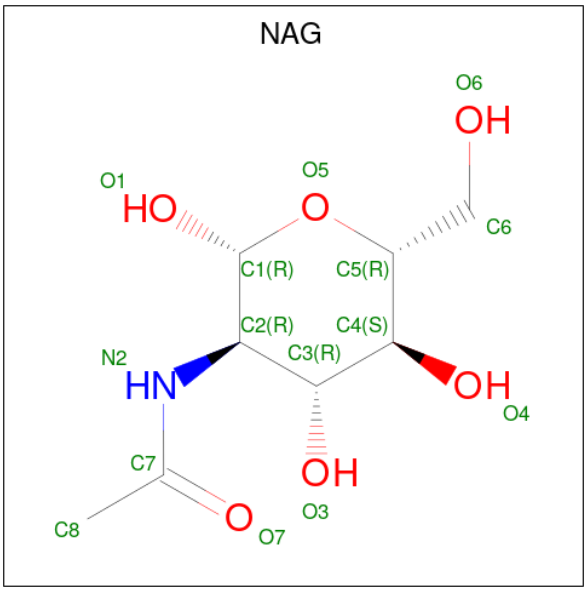
- Molecule 2 is a protein called PROTEIN (CATHEPSIN L: LIGHT CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	42	Total	C	N	O	S	15	0	0
			336	212	61	60	3			
2	D	42	Total	C	N	O	S	24	0	0
			336	212	61	60	3			

- Molecule 3 is a protein called PROTEIN (INVARIANT CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	I	65	Total	C	N	O	S	8	0	0
			507	312	92	97	6			
3	J	65	Total	C	N	O	S	17	0	0
			507	312	92	97	6			

- 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	I	1	Total C N O 14 8 1 5	0	0
4	J	1	Total C N O 14 8 1 5	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	196	Total O 196 196	0	0
5	B	49	Total O 49 49	0	0
5	C	190	Total O 190 190	1	0
5	D	36	Total O 36 36	0	0
5	I	110	Total O 110 110	0	0
5	J	87	Total O 87 87	0	0

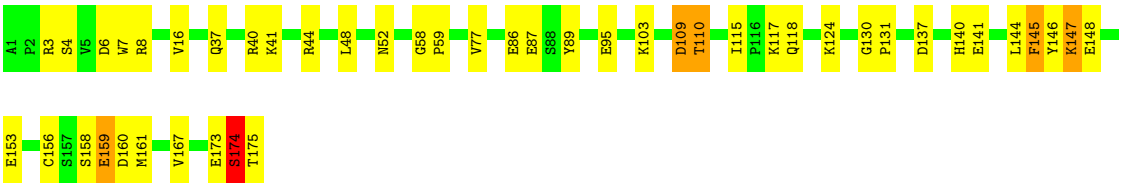
3 Residue-property plots (i)

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.

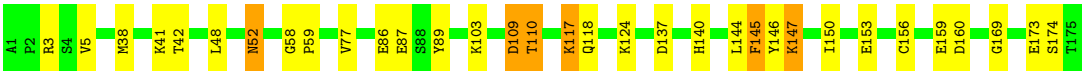
- Molecule 1: PROTEIN (CATHEPSIN L: HEAVY CHAIN)

Chain A:



- Molecule 1: PROTEIN (CATHEPSIN L: HEAVY CHAIN)

Chain C:



- Molecule 2: PROTEIN (CATHEPSIN L: LIGHT CHAIN)

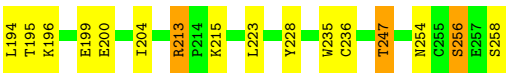
Chain B:

- Molecule 2: PROTEIN (CATHEPSIN L: LIGHT CHAIN)

Chain D:

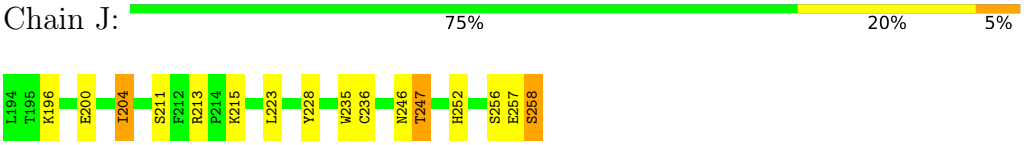
- Molecule 3: PROTEIN (INVARIANT CHAIN)

Chain I:



• Molecule 3: PROTEIN (INVARIANT CHAIN)

Chain J:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	62.59Å 80.59Å 64.25Å 90.00° 96.77° 90.00°	Depositor
Resolution (Å)	10.00 – 2.00	Depositor
% Data completeness (in resolution range)	97.0 (10.00-2.00)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
Refinement program	MAIN	Depositor
R, R_{free}	0.182 , 0.213	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	5060	wwPDB-VP
Average B, all atoms (Å^2)	32.0	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

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	0.55	0/1370	0.66	0/1852
1	C	0.52	0/1370	0.64	0/1852
2	B	0.54	0/346	0.67	0/465
2	D	0.58	0/346	0.66	0/465
3	I	0.50	0/523	0.63	0/710
3	J	0.51	0/523	0.68	0/710
All	All	0.53	0/4478	0.65	0/6054

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	1339	0	1239	24	2
1	C	1339	0	1239	21	2
2	B	336	0	316	16	6
2	D	336	0	316	9	0
3	I	507	0	457	12	1
3	J	507	0	457	10	7
4	I	14	0	13	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	J	14	0	13	0	0
5	A	196	0	0	1	0
5	B	49	0	0	0	1
5	C	190	0	0	1	1
5	D	36	0	0	0	0
5	I	110	0	0	3	0
5	J	87	0	0	3	0
All	All	5060	0	4050	74	10

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 9.

The worst 5 of 74 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:A:160:ASP:HB3	5:I:316:HOH:O	1.85	0.76
1:C:150:ILE:HD11	1:C:173:GLU:CG	2.23	0.68
1:C:48:LEU:O	1:C:86:GLU:HG2	1.93	0.68
3:I:215:LYS:HB3	3:I:223:LEU:HD12	1.77	0.66
1:A:6:ASP:OD1	1:A:8:ARG:HD3	1.97	0.65

The worst 5 of 10 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)
2:B:206:ARG:O	3:J:258:SER:OG[2_656]	0.88	1.32
2:B:207:ASN:CB	3:J:258:SER:OXT[2_656]	1.36	0.84
1:A:175:THR:OXT	1:C:38:MET:CE[1_656]	1.46	0.74
3:J:258:SER:CB	5:B:235:HOH:O[2_646]	1.57	0.63
2:B:207:ASN:N	3:J:258:SER:OXT[2_656]	1.77	0.43

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	173/175 (99%)	168 (97%)	4 (2%)	1 (1%)	25 19
1	C	173/175 (99%)	169 (98%)	4 (2%)	0	100 100
2	B	40/42 (95%)	37 (92%)	3 (8%)	0	100 100
2	D	40/42 (95%)	38 (95%)	2 (5%)	0	100 100
3	I	63/65 (97%)	61 (97%)	2 (3%)	0	100 100
3	J	63/65 (97%)	61 (97%)	2 (3%)	0	100 100
All	All	552/564 (98%)	534 (97%)	17 (3%)	1 (0%)	47 44

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	174	SER

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	143/143 (100%)	123 (86%)	20 (14%)	3 2
1	C	143/143 (100%)	129 (90%)	14 (10%)	8 4
2	B	33/33 (100%)	31 (94%)	2 (6%)	18 14
2	D	33/33 (100%)	32 (97%)	1 (3%)	41 41
3	I	58/58 (100%)	53 (91%)	5 (9%)	10 6
3	J	58/58 (100%)	50 (86%)	8 (14%)	3 2
All	All	468/468 (100%)	418 (89%)	50 (11%)	6 3

5 of 50 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	110	THR
1	C	174	SER
3	J	258	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	117	LYS
1	C	147	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	118	GLN
1	C	140	HIS
3	J	208	HIS
3	I	254	ASN
1	C	78	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)

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

5.5 Carbohydrates (i)

There are no monosaccharides 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	$\#\|Z\| > 2$
4	NAG	J	100	3	14,14,15	0.40	0	17,19,21	0.66	0
4	NAG	I	100	3	14,14,15	0.39	0	17,19,21	0.71	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
4	NAG	J	100	3	-	0/6/23/26	0/1/1/1
4	NAG	I	100	3	-	0/6/23/26	0/1/1/1

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.

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.