



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 25, 2020 – 05:28 am BST

PDB ID : 3T2N  
Title : Human hepsin protease in complex with the Fab fragment of an inhibitory antibody  
Authors : Koschubs, T.; Dengl, S.; Duerr, H.; Kaluza, K.; Georges, G.; Hartl, C.; Jennewein, S.; Lanzendoerfer, M.; Auer, J.; Stern, A.; Huang, K.-S.; Kostrewa, D.; Ries, S.; Hansen, S.; Kohnert, U.; Cramer, P.; Mundigl, O.  
Deposited on : 2011-07-22  
Resolution : 2.55 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
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

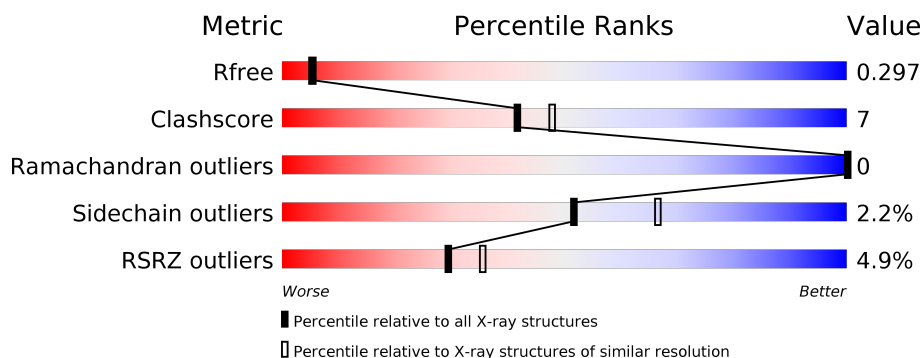
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.55 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))
$R_{free}$	130704	1284 (2.56-2.52)
Clashscore	141614	1332 (2.56-2.52)
Ramachandran outliers	138981	1315 (2.56-2.52)
Sidechain outliers	138945	1315 (2.56-2.52)
RSRZ outliers	127900	1272 (2.56-2.52)

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  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	372	<div> <div>7%</div> <div> <div></div> <div>75%</div> <div>16%</div> <div>9%</div> </div> </div>
1	B	372	<div> <div>5%</div> <div> <div></div> <div>75%</div> <div>15%</div> <div>10%</div> </div> </div>
2	H	225	<div> <div>3%</div> <div> <div></div> <div>75%</div> <div>15%</div> <div>10%</div> </div> </div>
2	I	225	<div> <div>3%</div> <div> <div></div> <div>79%</div> <div>12%</div> <div>9%</div> </div> </div>
3	L	215	<div> <div>2%</div> <div> <div></div> <div>81%</div> <div>16%</div> <div>3%</div> </div> </div>
3	M	215	<div> <div>4%</div> <div> <div></div> <div>81%</div> <div>16%</div> <div>3%</div> </div> </div>

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 11749 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 Serine protease hepsin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	339	Total	C	N	O	S	0	0	0
			2605	1638	470	478	19			
1	B	335	Total	C	N	O	S	0	0	0
			2576	1622	464	471	19			

- Molecule 2 is a protein called Antibody, Fab fragment, Heavy Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	203	Total	C	N	O	S	0	0	0
			1535	980	250	299	6			
2	I	205	Total	C	N	O	S	0	0	0
			1547	987	252	302	6			

- Molecule 3 is a protein called Antibody, Fab fragment, Light Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	211	Total	C	N	O	S	0	0	0
			1582	996	267	315	4			
3	M	212	Total	C	N	O	S	0	0	0
			1589	1000	268	317	4			

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	64	Total	O	0	0
			64	64		
4	B	74	Total	O	0	0
			74	74		
4	H	46	Total	O	0	0
			46	46		
4	I	36	Total	O	0	0
			36	36		

*Continued on next page...*

*Continued from previous page...*

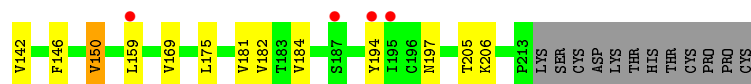
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	L	40	Total	O	0	0
			40	40		
4	M	55	Total	O	0	0
			55	55		

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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). 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.

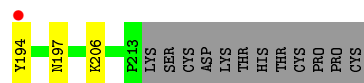
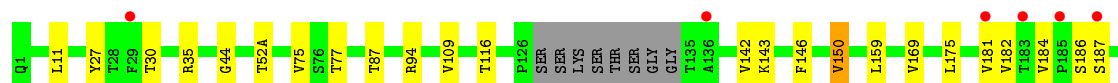
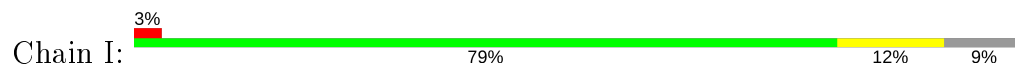
- Chain A:
- 
- 75% 16% 9% 7%
- SER ASP Q48 V66 S57 S58 A59 D60 A61 R62 D67 E70 D105 I106 R107 T108 A109 S115 L125 H126 H127 T128 Q129 V134 P141 R142 G143 R144 K157 L158 P159 VAL ASP ARG ILE VAL GLY GLY R167 D168 T169 S170 L171 C188 D196 L199 I212 V217 F218 V222 S226 P227 H228 G229 P245 F246 R247 D248 P249 N250 E253 M254 S255 N256 D257 L258 A259 H262 L263 P266 E271 Y272 L285 V286 T294 G295 N296 ASN THR GLN TTR TTR GLY GLN GLN ALA G307 V308 L309 Q310 I316 I317 S318 N319

- Chain B:
- 
- 5% 75% 15% 10%

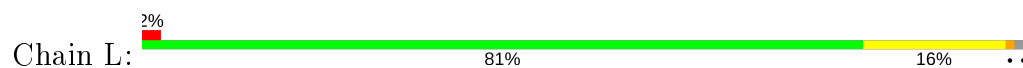
- Chain H:
- 
- | Amino Acid | Frequency (bits) |
|------------|------------------|
| Q1         | 0.02             |
| L11        | 0.01             |
| Y27        | 0.01             |
| T30        | 0.01             |
| R35        | 0.01             |
| A40        | 0.01             |
| F41        | 0.01             |
| G42        | 0.02             |
| Q43        | 0.01             |
| M48        | 0.01             |
| T52A       | 0.01             |
| S56        | 0.01             |
| P57        | 0.01             |
| D61        | 0.01             |
| S62        | 0.01             |
| F63        | 0.01             |
| M64        | 0.01             |
| F67        | 0.01             |
| S70        | 0.01             |
| V75        | 0.01             |
| S76        | 0.01             |
| T77        | 0.01             |
| A78        | 0.01             |
| Y79        | 0.01             |
| I82        | 0.02             |
| T87        | 0.01             |
| R94        | 0.01             |
| V109       | 0.01             |
| T116       | 0.01             |
| F126       | 0.01             |
| SER        | 0.01             |
| SER        | 0.01             |
| LYS        | 0.01             |
| SER        | 0.01             |
| SER        | 0.01             |
| THR        | 0.01             |
| SER        | 0.01             |
| SER        | 0.01             |
| GLY        | 0.01             |
| THR        | 0.01             |
| ALA        | 0.01             |
| A137       | 0.02             |



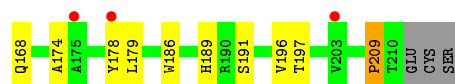
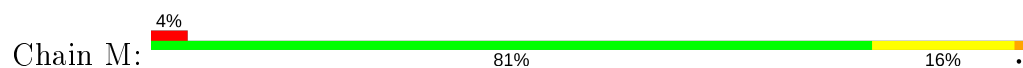
• Molecule 2: Antibody, Fab fragment, Heavy Chain



• Molecule 3: Antibody, Fab fragment, Light Chain



• Molecule 3: Antibody, Fab fragment, Light Chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	62.98Å 66.58Å 108.33Å 88.71° 94.30° 104.53°	Depositor
Resolution (Å)	47.33 – 2.55 47.33 – 2.55	Depositor EDS
% Data completeness (in resolution range)	98.7 (47.33-2.55) 79.2 (47.33-2.55)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.69 (at 2.54Å)	Xtriage
Refinement program	BUSTER 2.9.2	Depositor
R, $R_{free}$	0.242 , 0.275 0.257 , 0.297	Depositor DCC
$R_{free}$ test set	2772 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	43.1	Xtriage
Anisotropy	0.637	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 41.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	11749	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 17.14% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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	0/2665	0.75	3/3615 (0.1%)
1	B	0.54	1/2637 (0.0%)	0.74	0/3579
2	H	0.56	0/1575	0.78	0/2148
2	I	0.50	0/1587	0.77	0/2165
3	L	0.55	0/1623	0.74	0/2222
3	M	0.59	0/1630	0.75	1/2232 (0.0%)
All	All	0.54	1/11717 (0.0%)	0.75	4/15961 (0.0%)

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
1	A	0	1
1	B	0	1
All	All	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	206	PRO	N-CD	-5.43	1.40	1.47

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	107	ARG	NE-CZ-NH2	-7.07	116.76	120.30
1	A	107	ARG	NE-CZ-NH1	6.30	123.45	120.30
1	A	167	ARG	CD-NE-CZ	5.99	131.99	123.60
3	M	209	PRO	C-N-CA	5.32	135.00	121.70

There are no chirality outliers.



All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	167	ARG	Sidechain
1	B	257	ASP	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	A	2605	0	2536	43	0
1	B	2576	0	2500	37	0
2	H	1535	0	1502	22	0
2	I	1547	0	1514	16	0
3	L	1582	0	1541	25	0
3	M	1589	0	1548	27	0
4	A	64	0	0	0	0
4	B	74	0	0	0	0
4	H	46	0	0	0	0
4	I	36	0	0	0	0
4	L	40	0	0	1	0
4	M	55	0	0	1	0
All	All	11749	0	11141	158	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 7.

The worst 5 of 158 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:317:ILE:HD11	1:B:340:GLY:HA2	1.52	0.92
2:I:181:VAL:HG11	3:M:136:LEU:HD13	1.54	0.89
1:A:383:LEU:HG	1:A:383:LEU:O	1.76	0.85
2:H:75:VAL:HG13	2:H:77:THR:HG22	1.66	0.77
2:H:181:VAL:HG11	3:L:136:LEU:HD13	1.66	0.76

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	329/372 (88%)	319 (97%)	10 (3%)	0	100	100
1	B	325/372 (87%)	316 (97%)	9 (3%)	0	100	100
2	H	199/225 (88%)	193 (97%)	6 (3%)	0	100	100
2	I	201/225 (89%)	194 (96%)	7 (4%)	0	100	100
3	L	209/215 (97%)	203 (97%)	6 (3%)	0	100	100
3	M	210/215 (98%)	200 (95%)	10 (5%)	0	100	100
All	All	1473/1624 (91%)	1425 (97%)	48 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains ⓘ

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	280/302 (93%)	275 (98%)	5 (2%)	59	74
1	B	276/302 (91%)	268 (97%)	8 (3%)	42	57
2	H	172/191 (90%)	168 (98%)	4 (2%)	50	65
2	I	173/191 (91%)	169 (98%)	4 (2%)	50	65
3	L	175/179 (98%)	172 (98%)	3 (2%)	60	75
3	M	176/179 (98%)	172 (98%)	4 (2%)	50	65
All	All	1252/1344 (93%)	1224 (98%)	28 (2%)	52	66

5 of 28 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	377	TRP
2	H	169	VAL
3	M	34	ASN
2	H	67	PHE
2	H	150	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
2	I	43	GLN
2	I	199	ASN
3	L	189	HIS
2	H	199	ASN
3	M	79	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 carbohydrates 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 ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2		OWAB(Å <sup>2</sup> )	Q<0.9
1	A	339/372 (91%)	0.59	26 (7%)	13 17	25, 48, 72, 93	1 (0%)
1	B	335/372 (90%)	0.64	20 (5%)	21 25	28, 49, 72, 102	1 (0%)
2	H	203/225 (90%)	0.46	7 (3%)	45 52	25, 45, 68, 88	0
2	I	205/225 (91%)	0.43	7 (3%)	45 52	25, 44, 67, 88	0
3	L	211/215 (98%)	0.28	4 (1%)	66 73	24, 44, 61, 79	0
3	M	212/215 (98%)	0.34	9 (4%)	36 42	30, 43, 58, 64	0
All	All	1505/1624 (92%)	0.48	73 (4%)	29 35	24, 46, 69, 102	2 (0%)

The worst 5 of 73 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	222	VAL	6.1
1	B	123	GLY	5.2
1	A	383	LEU	4.7
1	B	239	TYR	4.0
2	H	137	ALA	3.7

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

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

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 6.4 Ligands ⓘ

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.