



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 10:04 AM UTC

PDB ID : 4QAE / pdb\_00004qae  
Title : Crystal structure of an engineered lipocalin (Anticalin) in complex with human hepcidin  
Authors : Giese, T.; Skerra, A.  
Deposited on : 2014-05-04  
Resolution : 2.10 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](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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4-5-2 with Phenix2.0  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 2.0  
EDS : 3.0  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.49

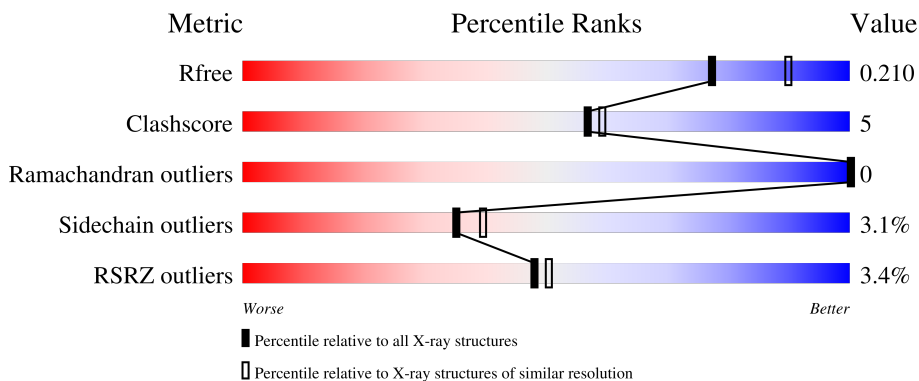
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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))
$R_{free}$	180053	6658 (2.10-2.10)
Clashscore	190562	7164 (2.10-2.10)
Ramachandran outliers	187476	7099 (2.10-2.10)
Sidechain outliers	187428	7100 (2.10-2.10)
RSRZ outliers	180081	6662 (2.10-2.10)

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 of 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	P	25	 4% 76% 24%
1	Q	25	 12% 80% 16% .
1	R	25	 8% 88% 8% .
1	S	25	 24% 84% 8% .
1	T	25	 12% 72% 20% 8%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	U	25	<p>12% 76% 20%</p>
2	A	188	<p>4% 84% 7% 9%</p>
2	B	188	<p>4% 81% 11% 7%</p>
2	C	188	<p>2% 77% 10% 13%</p>
2	D	188	<p>% 77% 10% 13%</p>
2	E	188	<p>% 81% 7% 12%</p>
2	F	188	<p>% 80% 7% 13%</p>

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 9938 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 Heparin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	P	25	187	113	34	31	9	0	0	0
1	Q	24	179	109	33	28	9	0	0	0
1	R	24	179	109	33	28	9	0	0	0
1	S	24	179	109	33	28	9	0	0	0
1	T	23	172	105	32	26	9	0	0	0
1	U	25	187	113	34	31	9	0	0	0

- Molecule 2 is a protein called Neutrophil gelatinase-associated lipocalin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	172	1391	902	229	254	6	0	0	0
2	B	175	1408	911	232	259	6	0	0	0
2	C	163	1331	866	216	243	6	0	0	0
2	D	164	1338	870	217	245	6	0	0	0
2	E	166	1351	877	219	249	6	0	0	0
2	F	164	1338	870	217	245	6	0	0	0

There are 216 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	28	HIS	GLN	engineered mutation	UNP P80188

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	40	GLU	ALA	engineered mutation	UNP P80188
A	41	VAL	ILE	engineered mutation	UNP P80188
A	49	MET	GLN	engineered mutation	UNP P80188
A	52	TRP	TYR	engineered mutation	UNP P80188
A	59	GLU	LYS	engineered mutation	UNP P80188
A	68	ILE	SER	engineered mutation	UNP P80188
A	70	MET	LEU	engineered mutation	UNP P80188
A	71	PRO	PHE	engineered mutation	UNP P80188
A	72	LEU	ARG	engineered mutation	UNP P80188
A	73	ALA	LYS	engineered mutation	UNP P80188
A	74	GLU	LYS	engineered mutation	UNP P80188
A	77	GLU	ASP	engineered mutation	UNP P80188
A	79	LEU	TRP	engineered mutation	UNP P80188
A	80	PHE	ILE	engineered mutation	UNP P80188
A	81	GLN	ARG	engineered mutation	UNP P80188
A	87	SER	CYS	engineered mutation	UNP P80188
A	96	GLY	ASN	engineered mutation	UNP P80188
A	100	GLY	TYR	engineered mutation	UNP P80188
A	103	ARG	LEU	engineered mutation	UNP P80188
A	106	GLY	TYR	engineered mutation	UNP P80188
A	125	VAL	LYS	engineered mutation	UNP P80188
A	127	TRP	SER	engineered mutation	UNP P80188
A	132	VAL	TYR	engineered mutation	UNP P80188
A	134	TRP	LYS	engineered mutation	UNP P80188
A	135	VAL	ILE	engineered mutation	UNP P80188
A	179	SER	-	expression tag	UNP P80188
A	180	ALA	-	expression tag	UNP P80188
A	181	TRP	-	expression tag	UNP P80188
A	182	SER	-	expression tag	UNP P80188
A	183	HIS	-	expression tag	UNP P80188
A	184	PRO	-	expression tag	UNP P80188
A	185	GLN	-	expression tag	UNP P80188
A	186	PHE	-	expression tag	UNP P80188
A	187	GLU	-	expression tag	UNP P80188
A	188	LYS	-	expression tag	UNP P80188
B	28	HIS	GLN	engineered mutation	UNP P80188
B	40	GLU	ALA	engineered mutation	UNP P80188
B	41	VAL	ILE	engineered mutation	UNP P80188
B	49	MET	GLN	engineered mutation	UNP P80188
B	52	TRP	TYR	engineered mutation	UNP P80188
B	59	GLU	LYS	engineered mutation	UNP P80188
B	68	ILE	SER	engineered mutation	UNP P80188

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	70	MET	LEU	engineered mutation	UNP P80188
B	71	PRO	PHE	engineered mutation	UNP P80188
B	72	LEU	ARG	engineered mutation	UNP P80188
B	73	ALA	LYS	engineered mutation	UNP P80188
B	74	GLU	LYS	engineered mutation	UNP P80188
B	77	GLU	ASP	engineered mutation	UNP P80188
B	79	LEU	TRP	engineered mutation	UNP P80188
B	80	PHE	ILE	engineered mutation	UNP P80188
B	81	GLN	ARG	engineered mutation	UNP P80188
B	87	SER	CYS	engineered mutation	UNP P80188
B	96	GLY	ASN	engineered mutation	UNP P80188
B	100	GLY	TYR	engineered mutation	UNP P80188
B	103	ARG	LEU	engineered mutation	UNP P80188
B	106	GLY	TYR	engineered mutation	UNP P80188
B	125	VAL	LYS	engineered mutation	UNP P80188
B	127	TRP	SER	engineered mutation	UNP P80188
B	132	VAL	TYR	engineered mutation	UNP P80188
B	134	TRP	LYS	engineered mutation	UNP P80188
B	135	VAL	ILE	engineered mutation	UNP P80188
B	179	SER	-	expression tag	UNP P80188
B	180	ALA	-	expression tag	UNP P80188
B	181	TRP	-	expression tag	UNP P80188
B	182	SER	-	expression tag	UNP P80188
B	183	HIS	-	expression tag	UNP P80188
B	184	PRO	-	expression tag	UNP P80188
B	185	GLN	-	expression tag	UNP P80188
B	186	PHE	-	expression tag	UNP P80188
B	187	GLU	-	expression tag	UNP P80188
B	188	LYS	-	expression tag	UNP P80188
C	28	HIS	GLN	engineered mutation	UNP P80188
C	40	GLU	ALA	engineered mutation	UNP P80188
C	41	VAL	ILE	engineered mutation	UNP P80188
C	49	MET	GLN	engineered mutation	UNP P80188
C	52	TRP	TYR	engineered mutation	UNP P80188
C	59	GLU	LYS	engineered mutation	UNP P80188
C	68	ILE	SER	engineered mutation	UNP P80188
C	70	MET	LEU	engineered mutation	UNP P80188
C	71	PRO	PHE	engineered mutation	UNP P80188
C	72	LEU	ARG	engineered mutation	UNP P80188
C	73	ALA	LYS	engineered mutation	UNP P80188
C	74	GLU	LYS	engineered mutation	UNP P80188
C	77	GLU	ASP	engineered mutation	UNP P80188

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	79	LEU	TRP	engineered mutation	UNP P80188
C	80	PHE	ILE	engineered mutation	UNP P80188
C	81	GLN	ARG	engineered mutation	UNP P80188
C	87	SER	CYS	engineered mutation	UNP P80188
C	96	GLY	ASN	engineered mutation	UNP P80188
C	100	GLY	TYR	engineered mutation	UNP P80188
C	103	ARG	LEU	engineered mutation	UNP P80188
C	106	GLY	TYR	engineered mutation	UNP P80188
C	125	VAL	LYS	engineered mutation	UNP P80188
C	127	TRP	SER	engineered mutation	UNP P80188
C	132	VAL	TYR	engineered mutation	UNP P80188
C	134	TRP	LYS	engineered mutation	UNP P80188
C	135	VAL	ILE	engineered mutation	UNP P80188
C	179	SER	-	expression tag	UNP P80188
C	180	ALA	-	expression tag	UNP P80188
C	181	TRP	-	expression tag	UNP P80188
C	182	SER	-	expression tag	UNP P80188
C	183	HIS	-	expression tag	UNP P80188
C	184	PRO	-	expression tag	UNP P80188
C	185	GLN	-	expression tag	UNP P80188
C	186	PHE	-	expression tag	UNP P80188
C	187	GLU	-	expression tag	UNP P80188
C	188	LYS	-	expression tag	UNP P80188
D	28	HIS	GLN	engineered mutation	UNP P80188
D	40	GLU	ALA	engineered mutation	UNP P80188
D	41	VAL	ILE	engineered mutation	UNP P80188
D	49	MET	GLN	engineered mutation	UNP P80188
D	52	TRP	TYR	engineered mutation	UNP P80188
D	59	GLU	LYS	engineered mutation	UNP P80188
D	68	ILE	SER	engineered mutation	UNP P80188
D	70	MET	LEU	engineered mutation	UNP P80188
D	71	PRO	PHE	engineered mutation	UNP P80188
D	72	LEU	ARG	engineered mutation	UNP P80188
D	73	ALA	LYS	engineered mutation	UNP P80188
D	74	GLU	LYS	engineered mutation	UNP P80188
D	77	GLU	ASP	engineered mutation	UNP P80188
D	79	LEU	TRP	engineered mutation	UNP P80188
D	80	PHE	ILE	engineered mutation	UNP P80188
D	81	GLN	ARG	engineered mutation	UNP P80188
D	87	SER	CYS	engineered mutation	UNP P80188
D	96	GLY	ASN	engineered mutation	UNP P80188
D	100	GLY	TYR	engineered mutation	UNP P80188

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	103	ARG	LEU	engineered mutation	UNP P80188
D	106	GLY	TYR	engineered mutation	UNP P80188
D	125	VAL	LYS	engineered mutation	UNP P80188
D	127	TRP	SER	engineered mutation	UNP P80188
D	132	VAL	TYR	engineered mutation	UNP P80188
D	134	TRP	LYS	engineered mutation	UNP P80188
D	135	VAL	ILE	engineered mutation	UNP P80188
D	179	SER	-	expression tag	UNP P80188
D	180	ALA	-	expression tag	UNP P80188
D	181	TRP	-	expression tag	UNP P80188
D	182	SER	-	expression tag	UNP P80188
D	183	HIS	-	expression tag	UNP P80188
D	184	PRO	-	expression tag	UNP P80188
D	185	GLN	-	expression tag	UNP P80188
D	186	PHE	-	expression tag	UNP P80188
D	187	GLU	-	expression tag	UNP P80188
D	188	LYS	-	expression tag	UNP P80188
E	28	HIS	GLN	engineered mutation	UNP P80188
E	40	GLU	ALA	engineered mutation	UNP P80188
E	41	VAL	ILE	engineered mutation	UNP P80188
E	49	MET	GLN	engineered mutation	UNP P80188
E	52	TRP	TYR	engineered mutation	UNP P80188
E	59	GLU	LYS	engineered mutation	UNP P80188
E	68	ILE	SER	engineered mutation	UNP P80188
E	70	MET	LEU	engineered mutation	UNP P80188
E	71	PRO	PHE	engineered mutation	UNP P80188
E	72	LEU	ARG	engineered mutation	UNP P80188
E	73	ALA	LYS	engineered mutation	UNP P80188
E	74	GLU	LYS	engineered mutation	UNP P80188
E	77	GLU	ASP	engineered mutation	UNP P80188
E	79	LEU	TRP	engineered mutation	UNP P80188
E	80	PHE	ILE	engineered mutation	UNP P80188
E	81	GLN	ARG	engineered mutation	UNP P80188
E	87	SER	CYS	engineered mutation	UNP P80188
E	96	GLY	ASN	engineered mutation	UNP P80188
E	100	GLY	TYR	engineered mutation	UNP P80188
E	103	ARG	LEU	engineered mutation	UNP P80188
E	106	GLY	TYR	engineered mutation	UNP P80188
E	125	VAL	LYS	engineered mutation	UNP P80188
E	127	TRP	SER	engineered mutation	UNP P80188
E	132	VAL	TYR	engineered mutation	UNP P80188
E	134	TRP	LYS	engineered mutation	UNP P80188

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	135	VAL	ILE	engineered mutation	UNP P80188
E	179	SER	-	expression tag	UNP P80188
E	180	ALA	-	expression tag	UNP P80188
E	181	TRP	-	expression tag	UNP P80188
E	182	SER	-	expression tag	UNP P80188
E	183	HIS	-	expression tag	UNP P80188
E	184	PRO	-	expression tag	UNP P80188
E	185	GLN	-	expression tag	UNP P80188
E	186	PHE	-	expression tag	UNP P80188
E	187	GLU	-	expression tag	UNP P80188
E	188	LYS	-	expression tag	UNP P80188
F	28	HIS	GLN	engineered mutation	UNP P80188
F	40	GLU	ALA	engineered mutation	UNP P80188
F	41	VAL	ILE	engineered mutation	UNP P80188
F	49	MET	GLN	engineered mutation	UNP P80188
F	52	TRP	TYR	engineered mutation	UNP P80188
F	59	GLU	LYS	engineered mutation	UNP P80188
F	68	ILE	SER	engineered mutation	UNP P80188
F	70	MET	LEU	engineered mutation	UNP P80188
F	71	PRO	PHE	engineered mutation	UNP P80188
F	72	LEU	ARG	engineered mutation	UNP P80188
F	73	ALA	LYS	engineered mutation	UNP P80188
F	74	GLU	LYS	engineered mutation	UNP P80188
F	77	GLU	ASP	engineered mutation	UNP P80188
F	79	LEU	TRP	engineered mutation	UNP P80188
F	80	PHE	ILE	engineered mutation	UNP P80188
F	81	GLN	ARG	engineered mutation	UNP P80188
F	87	SER	CYS	engineered mutation	UNP P80188
F	96	GLY	ASN	engineered mutation	UNP P80188
F	100	GLY	TYR	engineered mutation	UNP P80188
F	103	ARG	LEU	engineered mutation	UNP P80188
F	106	GLY	TYR	engineered mutation	UNP P80188
F	125	VAL	LYS	engineered mutation	UNP P80188
F	127	TRP	SER	engineered mutation	UNP P80188
F	132	VAL	TYR	engineered mutation	UNP P80188
F	134	TRP	LYS	engineered mutation	UNP P80188
F	135	VAL	ILE	engineered mutation	UNP P80188
F	179	SER	-	expression tag	UNP P80188
F	180	ALA	-	expression tag	UNP P80188
F	181	TRP	-	expression tag	UNP P80188
F	182	SER	-	expression tag	UNP P80188
F	183	HIS	-	expression tag	UNP P80188

*Continued on next page...*

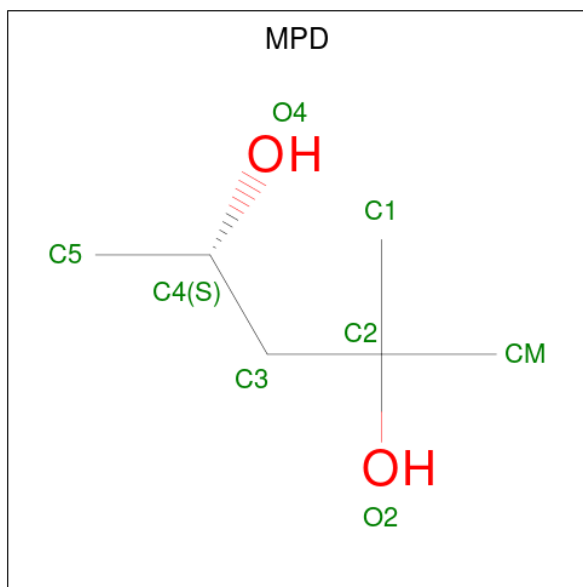
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	184	PRO	-	expression tag	UNP P80188
F	185	GLN	-	expression tag	UNP P80188
F	186	PHE	-	expression tag	UNP P80188
F	187	GLU	-	expression tag	UNP P80188
F	188	LYS	-	expression tag	UNP P80188

- Molecule 3 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	P	1	Total Cl 1 1	0	0
3	S	1	Total Cl 1 1	0	0
3	U	1	Total Cl 1 1	0	0
3	B	1	Total Cl 1 1	0	0
3	D	1	Total Cl 1 1	0	0

- Molecule 4 is (4S)-2-METHYL-2,4-PENTANEDIOL (CCD ID: MPD) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 8 6 2	0	0

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	C	O	0	0
			8	6	2		
4	D	1	Total	C	O	0	0
			8	6	2		
4	E	1	Total	C	O	0	0
			8	6	2		
4	E	1	Total	C	O	0	0
			8	6	2		
4	F	1	Total	C	O	0	0
			8	6	2		
4	F	1	Total	C	O	0	0
			8	6	2		
4	F	1	Total	C	O	0	0
			8	6	2		

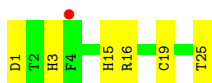
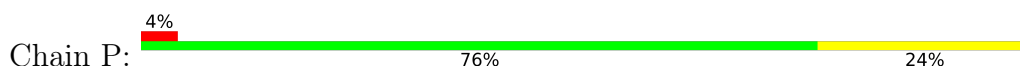
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	P	11	Total	O	0	0
			11	11		
5	Q	10	Total	O	0	0
			10	10		
5	R	11	Total	O	0	0
			11	11		
5	S	8	Total	O	0	0
			8	8		
5	T	8	Total	O	0	0
			8	8		
5	U	9	Total	O	0	0
			9	9		
5	A	126	Total	O	0	0
			126	126		
5	B	117	Total	O	0	0
			117	117		
5	C	66	Total	O	0	0
			66	66		
5	D	76	Total	O	0	0
			76	76		
5	E	107	Total	O	0	0
			107	107		
5	F	80	Total	O	0	0
			80	80		

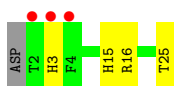
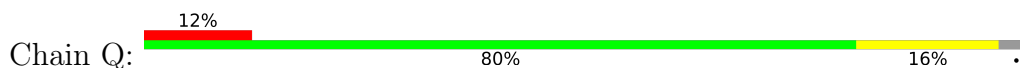
### 3 Residue-property plots [i](#)

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 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.

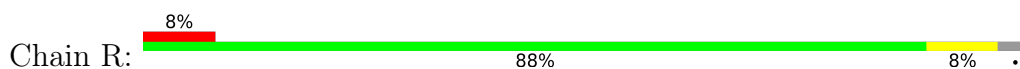
- Molecule 1: Hepcidin



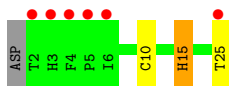
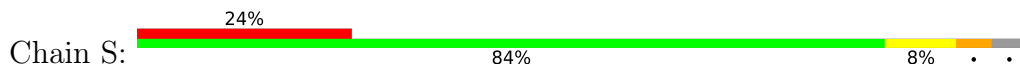
- Molecule 1: Hepcidin



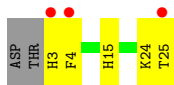
- Molecule 1: Hepcidin



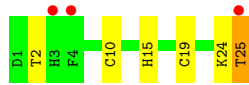
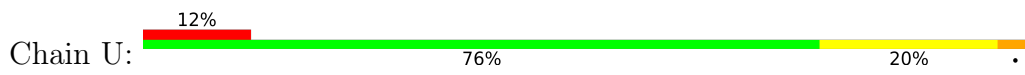
- Molecule 1: Hepcidin



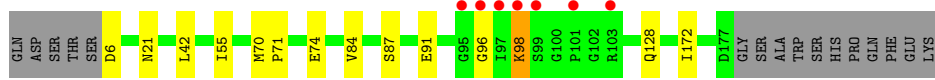
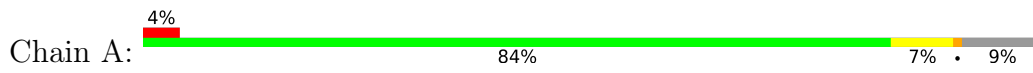
- Molecule 1: Hepcidin



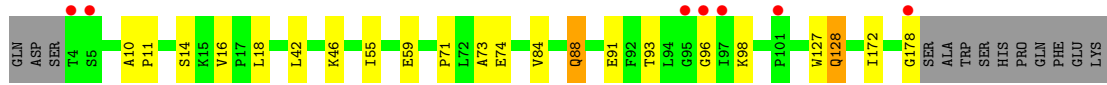
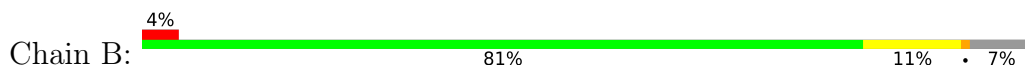
- Molecule 1: Hepcidin



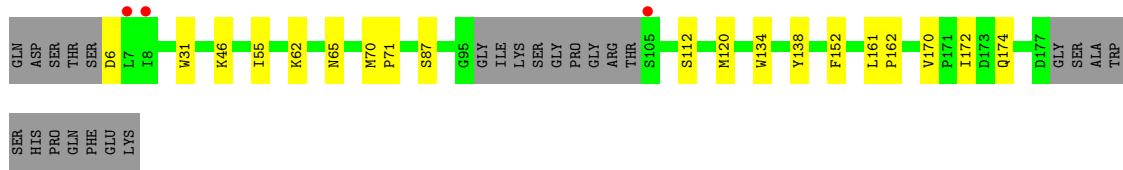
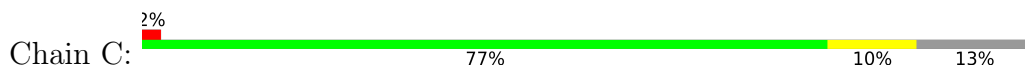
● Molecule 2: Neutrophil gelatinase-associated lipocalin



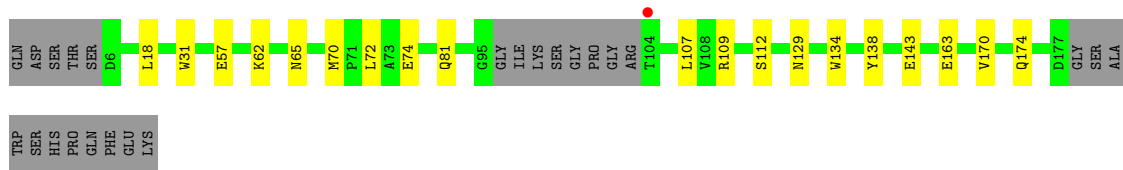
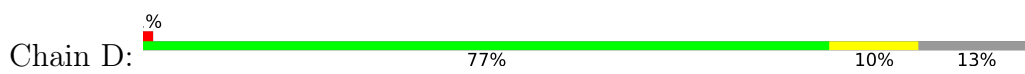
● Molecule 2: Neutrophil gelatinase-associated lipocalin



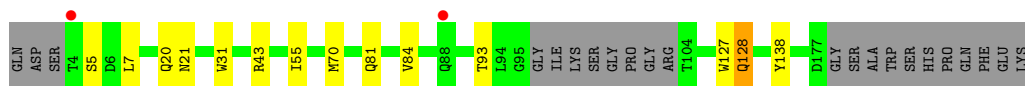
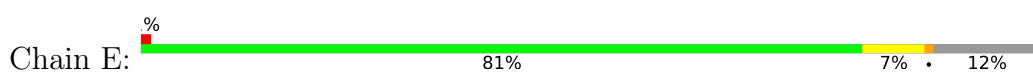
● Molecule 2: Neutrophil gelatinase-associated lipocalin




● Molecule 2: Neutrophil gelatinase-associated lipocalin

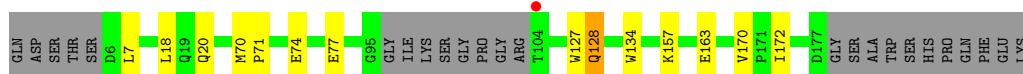


● Molecule 2: Neutrophil gelatinase-associated lipocalin



● Molecule 2: Neutrophil gelatinase-associated lipocalin

Chain F:  %



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	126.89Å 126.89Å 156.71Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	33.18 – 2.10 33.18 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.8 (33.18-2.10) 99.8 (33.18-2.10)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.28 (at 2.10Å)	Xtrriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.166 , 0.206 (Not available) , 0.210	Depositor DCC
$R_{free}$ test set	4267 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	29.0	Xtrriage
Anisotropy	0.054	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 48.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.000 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	9938	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 44.64 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.4820e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: CL, MPD

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	P	1.03	0/191	0.86	0/252
1	Q	1.00	0/183	0.96	0/241
1	R	0.77	0/183	0.84	0/241
1	S	0.91	0/183	0.85	0/241
1	T	0.98	0/176	1.22	2/231 (0.9%)
1	U	0.84	0/191	0.83	0/252
2	A	0.92	0/1430	0.91	0/1943
2	B	0.99	2/1447 (0.1%)	0.98	3/1966 (0.2%)
2	C	0.77	0/1368	0.90	2/1859 (0.1%)
2	D	0.81	0/1375	0.90	2/1869 (0.1%)
2	E	0.88	0/1388	0.92	0/1887
2	F	0.87	0/1375	0.92	1/1869 (0.1%)
All	All	0.88	2/9490 (0.0%)	0.92	10/12851 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	10	ALA	C-O	-6.08	1.16	1.24
2	B	73	ALA	C-O	-5.55	1.17	1.24

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	170	VAL	CB-CA-C	-6.98	103.36	110.89
2	F	170	VAL	CB-CA-C	-6.47	103.90	110.89
2	D	170	VAL	CB-CA-C	-5.79	104.64	110.89
2	B	16	VAL	N-CA-C	5.60	113.59	107.60
2	B	11	PRO	CA-C-N	-5.28	114.80	120.03

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	P	187	0	172	7	0
1	Q	179	0	165	5	0
1	R	179	0	165	1	0
1	S	179	0	165	2	0
1	T	172	0	158	2	0
1	U	187	0	172	5	0
2	A	1391	0	1369	15	0
2	B	1408	0	1384	14	0
2	C	1331	0	1303	15	0
2	D	1338	0	1310	12	0
2	E	1351	0	1322	10	0
2	F	1338	0	1310	8	0
3	B	1	0	0	0	0
3	D	1	0	0	0	0
3	P	1	0	0	0	0
3	S	1	0	0	0	0
3	U	1	0	0	0	0
4	A	8	0	14	0	0
4	B	8	0	14	0	0
4	D	8	0	14	4	0
4	E	16	0	28	3	0
4	F	24	0	42	3	0
5	A	126	0	0	5	0
5	B	117	0	0	5	0
5	C	66	0	0	3	0
5	D	76	0	0	4	0
5	E	107	0	0	9	0
5	F	80	0	0	3	0
5	P	11	0	0	0	0
5	Q	10	0	0	1	0
5	R	11	0	0	0	0
5	S	8	0	0	0	0
5	T	8	0	0	0	0
5	U	9	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	9938	0	9107	84	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 5.

The worst 5 of 84 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:A:84:VAL:HG21	2:A:96:GLY:HA2	1.51	0.91
2:E:43:ARG:HD2	5:E:362:HOH:O	1.73	0.88
2:A:21:ASN:HB2	5:A:318:HOH:O	1.86	0.76
2:D:70:MET:CE	5:D:341:HOH:O	2.35	0.74
4:D:202:MPD:H53	4:D:202:MPD:HM1	1.69	0.73

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	P	23/25 (92%)	22 (96%)	1 (4%)	0	100	100
1	Q	22/25 (88%)	21 (96%)	1 (4%)	0	100	100
1	R	22/25 (88%)	21 (96%)	1 (4%)	0	100	100
1	S	22/25 (88%)	21 (96%)	1 (4%)	0	100	100
1	T	21/25 (84%)	20 (95%)	1 (5%)	0	100	100
1	U	23/25 (92%)	22 (96%)	1 (4%)	0	100	100
2	A	170/188 (90%)	166 (98%)	4 (2%)	0	100	100
2	B	173/188 (92%)	167 (96%)	6 (4%)	0	100	100
2	C	159/188 (85%)	154 (97%)	5 (3%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	D	160/188 (85%)	156 (98%)	4 (2%)	0	100	100
2	E	162/188 (86%)	158 (98%)	4 (2%)	0	100	100
2	F	160/188 (85%)	155 (97%)	5 (3%)	0	100	100
All	All	1117/1278 (87%)	1083 (97%)	34 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	P	23/23 (100%)	22 (96%)	1 (4%)	26	27
1	Q	22/23 (96%)	21 (96%)	1 (4%)	24	25
1	R	22/23 (96%)	21 (96%)	1 (4%)	24	25
1	S	22/23 (96%)	20 (91%)	2 (9%)	9	6
1	T	21/23 (91%)	20 (95%)	1 (5%)	23	23
1	U	23/23 (100%)	20 (87%)	3 (13%)	4	2
2	A	155/169 (92%)	153 (99%)	2 (1%)	61	69
2	B	157/169 (93%)	153 (98%)	4 (2%)	42	48
2	C	149/169 (88%)	147 (99%)	2 (1%)	61	69
2	D	150/169 (89%)	144 (96%)	6 (4%)	28	29
2	E	152/169 (90%)	148 (97%)	4 (3%)	40	46
2	F	150/169 (89%)	145 (97%)	5 (3%)	33	37
All	All	1046/1152 (91%)	1014 (97%)	32 (3%)	35	39

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	F	18	LEU
2	F	20	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	B	88	GLN
2	B	14	SER
2	F	77	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	21	ASN
2	D	26	GLN
2	F	26	GLN
2	D	81	GLN
2	B	128	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 oligosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

Of 13 ligands modelled in this entry, 5 are monoatomic - leaving 8 for Mogul analysis.

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	MPD	F	201	-	7,7,7	0.41	0	9,10,10	1.07	1 (11%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	MPD	E	202	-	7,7,7	0.42	0	9,10,10	0.79	0
4	MPD	D	202	-	7,7,7	0.48	0	9,10,10	0.67	0
4	MPD	E	201	-	7,7,7	0.37	0	9,10,10	0.76	0
4	MPD	F	202	-	7,7,7	0.37	0	9,10,10	0.64	0
4	MPD	A	201	-	7,7,7	0.44	0	9,10,10	0.62	0
4	MPD	B	202	-	7,7,7	0.54	0	9,10,10	1.33	2 (22%)
4	MPD	F	203	-	7,7,7	0.37	0	9,10,10	0.69	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	MPD	F	201	-	-	3/5/5/5	-
4	MPD	E	202	-	-	4/5/5/5	-
4	MPD	D	202	-	-	0/5/5/5	-
4	MPD	E	201	-	-	5/5/5/5	-
4	MPD	F	202	-	-	2/5/5/5	-
4	MPD	A	201	-	-	4/5/5/5	-
4	MPD	B	202	-	-	4/5/5/5	-
4	MPD	F	203	-	-	3/5/5/5	-

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	F	201	MPD	O2-C2-C3	2.38	118.16	109.27
4	B	202	MPD	O2-C2-CM	-2.31	100.78	107.99
4	B	202	MPD	O2-C2-C3	2.15	117.30	109.27

There are no chirality outliers.

5 of 25 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	E	201	MPD	C2-C3-C4-O4
4	E	201	MPD	C2-C3-C4-C5
4	F	201	MPD	C2-C3-C4-C5
4	F	202	MPD	C2-C3-C4-C5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
4	F	203	MPD	C1-C2-C3-C4

There are no ring outliers.

4 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	F	201	MPD	1	0
4	D	202	MPD	4	0
4	E	201	MPD	3	0
4	F	203	MPD	2	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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	P	25/25 (100%)	-0.02	1 (4%) 42 44	17, 21, 77, 107	0
1	Q	24/25 (96%)	0.06	3 (12%) 8 8	17, 24, 77, 90	0
1	R	24/25 (96%)	0.12	2 (8%) 17 18	25, 32, 74, 78	0
1	S	24/25 (96%)	0.40	6 (25%) 2 2	25, 31, 75, 86	0
1	T	23/25 (92%)	0.24	3 (13%) 7 7	21, 29, 73, 94	0
1	U	25/25 (100%)	0.19	3 (12%) 9 9	22, 27, 88, 92	0
2	A	172/188 (91%)	-0.24	7 (4%) 41 43	17, 30, 58, 88	0
2	B	175/188 (93%)	-0.15	7 (4%) 42 44	14, 30, 74, 137	0
2	C	163/188 (86%)	0.08	3 (1%) 67 70	19, 47, 72, 82	0
2	D	164/188 (87%)	-0.00	1 (0%) 85 87	21, 43, 67, 80	0
2	E	166/188 (88%)	-0.24	2 (1%) 76 78	16, 35, 67, 90	0
2	F	164/188 (87%)	-0.16	1 (0%) 85 87	21, 38, 56, 72	1 (0%)
All	All	1149/1278 (89%)	-0.08	39 (3%) 48 50	14, 36, 72, 137	1 (0%)

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	A	97	ILE	5.3
2	B	4	THR	5.3
2	E	4	THR	4.4
1	T	4	PHE	4.3
2	B	97	ILE	4.3

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

### 6.3 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	MPD	E	202	8/8	0.71	0.28	72,76,81,94	0
4	MPD	A	201	8/8	0.75	0.24	56,57,62,69	0
4	MPD	F	203	8/8	0.75	0.26	62,67,69,71	0
4	MPD	F	201	8/8	0.81	0.22	55,59,66,68	0
4	MPD	D	202	8/8	0.81	0.19	47,58,61,63	0
4	MPD	F	202	8/8	0.82	0.22	58,72,76,82	0
4	MPD	B	202	8/8	0.82	0.17	41,45,50,64	0
4	MPD	E	201	8/8	0.85	0.16	48,54,56,57	0
3	CL	S	101	1/1	0.87	0.15	70,70,70,70	0
3	CL	D	201	1/1	0.88	0.15	80,80,80,80	0
3	CL	P	101	1/1	0.91	0.11	61,61,61,61	0
3	CL	B	201	1/1	0.92	0.14	76,76,76,76	0
3	CL	U	101	1/1	0.93	0.10	62,62,62,62	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.