



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 12:48 PM GMT

PDB ID : 3S5K  
Title : Crystal structures of falcilysin, a M16 metalloprotease from the malaria parasite *Plasmodium falciparum*  
Authors : Morgunova, E.; Ponpuak, M.; Istvan, E.; Popov, A.; Goldberg, D.; Eneqvist, T.  
Deposited on : 2011-05-23  
Resolution : 3.20 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  
<http://wwpdb.org/validation/2016/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  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

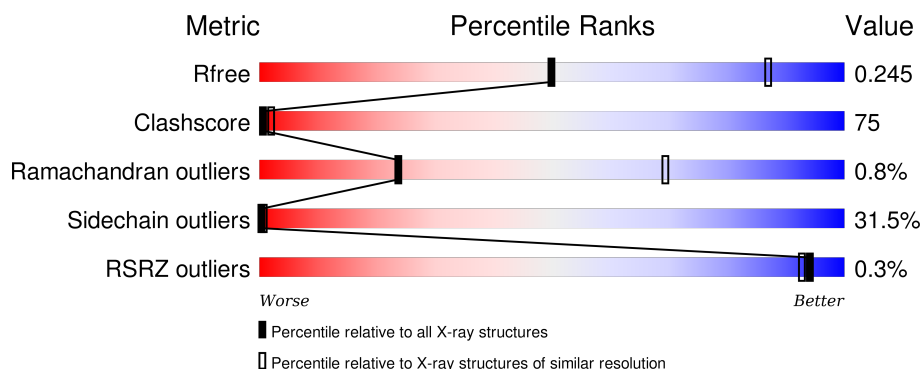
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*


The reported resolution of this entry is 3.20 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))
$R_{free}$	91344	1124 (3.24-3.16)
Clashscore	102246	1024 (3.22-3.18)
Ramachandran outliers	100387	1004 (3.22-3.18)
Sidechain outliers	100360	1003 (3.22-3.18)
RSRZ outliers	91569	1129 (3.24-3.16)

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. 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	1193	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 8741 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 Falcilysin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1053	Total	C	N	O	S	0	0	0
			8674	5577	1418	1653	26			

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Zn	0	0
			1	1		

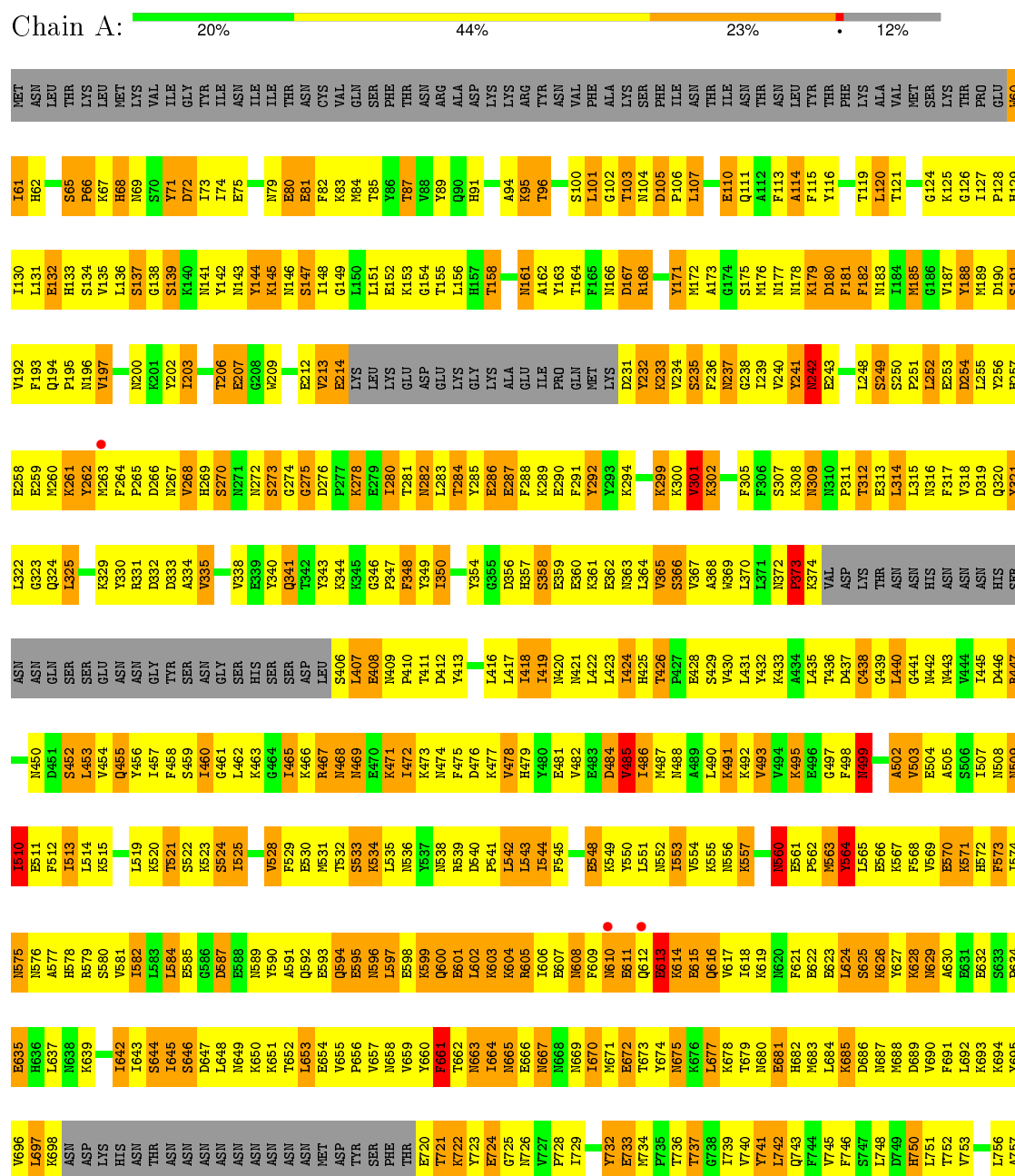
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	66	Total	O	0	0
			66	66		

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: Falcilysin



K1151	L1085	G1019	S956	N894	C826	Y758
F1152	R1086	E1020	Y957	T895	N827	L759
A1153	K1087	Y1021	F958	L896	D828	N760
D1154	M1088	L1022	E959	K397		
L1155		D1023	E960	L898		F762
L1156	T1091	P1024	N961	L899	I832	K763
E1157	M1092	S1025	D962	E900	A833	T764
S1158	T1093	F1026	K963	Q901	L834	L765
K1159	E1094	T1027	Y964	L902	E835	L766
V1160	N1095	V1028	N965	E903	A836	L767
N1161	D1096	I1029	ASN	L904	V837	E768
E1162	L1097	V1030	ASP	L905	K838	N769
F1163	L1098	A1031	MET	E906	E839	K770
E1164	R1099	A1032	GLN	N907	S840	T771
K1165	Y1100	L1033	ASN	D908	D841	
	I1101	K1034	VAL	F909	F842	R774
V1168	I1102	N1035	LVS	K910	S843	S775
I1169	N1103	S1036	ASN	T911		S776
I1170	T1104	Y1037	ASP	L912	K846	
T1171	I1105	L1038	PRQ	L915	K347	F779
T1172	G1106	W1039	THR	L916	V848	V780
K1173	T1107	D1040	VAL	L917	I849	I781
E1174	I1108	T1041	MET	N917	D850	L782
K1175	D1109	V1042	G979	R918	I851	R783
A1176	K1110	R1043	N980	L919	L852	
N1177	P1111	N981	N981	R920	K853	N786
E1178	R1112	G1047	Y984	N921	R854	L787
Y1179	R1113	A1048	K985	R922	K855	G788
I1180	G1114	Y1049	N985	L923	I856	S789
A1181	I1115	G1050	S986	F924	N857	N790
N1182	E1116	V1051	K987	N925	G858	S791
V1183	L1117	F1052	K988	K926	N859	A792
D1184	S1118	A1053	L989	K927	K860	
G1185	K1119	D1054	F990	N928	T861	Y797
E1186	L1120	I1055	D991	L929	T862	S798
F1187	S1121	E1056	E992	N930	F863	K799
K1188	F1122	Y1057	E993		S864	D800
K1189		D1058	K994	V933	E865	D801
V1190	I1126	G1059	Y995	T934	Y868	H802
L1191	S1127	S1060	K996	S935	L803	L303
I1192	M1128	Y1061	K997	D936	A869	N804
		V1062	E998	Y937	I870	V805
Q1132		F1063	F999	G938	L871	T806
D1133	R1134	L1064	A1066	L940	M872	D807
		S1065	V1001	L941	K873	K808
F1137		A1066	L1002	R942	Y874	Y809
R1138		R1067	P1003	L943	N875	N810
K1139		D1068	T1004	F944	K876	A811
R1140		P1069	F1005	N945		Q812
L1141		N1070	V1006	Y946	K882	A813
M1142		L1071	M1007	N947		L814
N1143		E1072		S947	A895	F815
T1144		K1073		N948	H886	N816
K1145		T1077		E949	N887	L817
K1146		F1078		S950	I888	E818
E1147		R1079	I1014	L951	N889	N819
D1148			L1015	K952	Y890	H820
F1149			F1016	N953	G891	V821
Y1150			K1017	L954	Y892	S822
			P1018	V055	E903	

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	93.78Å 105.29Å 114.76Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.83 – 3.20 42.83 – 2.84	Depositor EDS
% Data completeness (in resolution range)	99.9 (42.83-3.20) 95.4 (42.83-2.84)	Depositor EDS
$R_{merge}$	0.26	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.24 (at 2.86Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE: 1.6_289)	Depositor
R, $R_{free}$	0.210 , 0.316 0.220 , 0.245	Depositor DCC
$R_{free}$ test set	800 reflections (4.32%)	DCC
Wilson B-factor (Å <sup>2</sup> )	50.1	Xtriage
Anisotropy	0.258	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 78.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.42$ , $\langle L^2 \rangle = 0.24$	Xtriage
Outliers	0 of 26160 reflections	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	8741	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.34% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: ZN

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	1.79	58/8849 (0.7%)	1.11	36/11918 (0.3%)

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

The worst 5 of 58 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	438	CYS	CB-SG	-10.15	1.65	1.82
1	A	826	CYS	CB-SG	-7.60	1.69	1.82
1	A	818	GLU	CG-CD	-7.31	1.41	1.51
1	A	188	TYR	CE1-CZ	-6.19	1.30	1.38
1	A	821	VAL	CA-CB	-6.18	1.41	1.54

The worst 5 of 36 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	915	ILE	CB-CA-C	-7.40	96.80	111.60
1	A	365	VAL	CB-CA-C	-6.51	99.02	111.40
1	A	945	VAL	CB-CA-C	-6.49	99.07	111.40
1	A	837	VAL	CB-CA-C	-6.44	99.16	111.40
1	A	856	ILE	CB-CA-C	-6.38	98.83	111.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	560	ASN	Peptide

## 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	8674	0	8613	1304	2
2	A	1	0	0	0	0
3	A	66	0	0	6	0
All	All	8741	0	8613	1304	2

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 75.

The worst 5 of 1304 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:263:MET:CE	1:A:457:ILE:HD12	1.14	1.54
1:A:610:ASN:C	1:A:613:GLU:CG	1.75	1.50
1:A:232:TYR:HB2	1:A:614:LYS:NZ	1.30	1.45
1:A:151:LEU:O	1:A:155:THR:CG2	1.67	1.42
1:A:612:GLN:CB	1:A:616:GLN:HB3	1.49	1.42

All (2) 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)
1:A:608:ASN:ND2	1:A:988:LYS:CG[2_554]	1.98	0.22
1:A:538:ASN:O	1:A:679:THR:OG1[2_654]	2.10	0.10



## 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	1043/1193 (87%)	998 (96%)	37 (4%)	8 (1%)	24 69

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	213	VAL
1	A	610	ASN
1	A	613	GLU
1	A	1128	ASN
1	A	373	PRO

### 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	972/1104 (88%)	666 (68%)	306 (32%)	0 1

5 of 306 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	564	TYR
1	A	645	ILE
1	A	1127	SER
1	A	571	LYS
1	A	604	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such

sidechains are listed below:

Mol	Chain	Res	Type
1	A	538	ASN
1	A	600	GLN
1	A	953	ASN
1	A	560	ASN
1	A	572	HIS

### 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](#)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

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

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	1053/1193 (88%)	-0.50	3 (0%) 94 93	9, 32, 70, 128	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	610	ASN	4.0
1	A	612	GLN	2.8
1	A	263	MET	2.0

### 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

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	ZN	A	1194	1/1	0.35	0.20	-	38,38,38,38	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.