Molecular modeling experiments

Molecular modeling experiments were performed using the following molecular modeling softwares: Molecular Operating Environment (MOE). Both MOE 2018.1 and Yasara 18.12.27 [126–128] were used for homology modeling of the vicilin storage globulin. The general methodology for the modeling of the vicilin storage protein is summarized in Figure S1. Details are given in the following studies [51,129]. The used templates for the homology model are given in Table S1. The amino acid sequence of the storage protein vicilin from *Theobroma cacao* (Entry: Q43358; 60.8 kDa/525 amino acids reviewed version from the database UniProt - http://www.uniprot.org/; 28.01-2019) was used for the modeling experiments and determination of the possible phosphorylation and glycosylation (O-GLcNAc) sites.

Template	Total score	BLAST E- value	Align score	Cover	ID	Resolution	Header
1	380.14	5e-081	717.0	67%	5VF5- A	1.49 A	XRAY SM80.1 Vicilin [Solanum melongena] <na> (370 residues with quality score 0.793), released 2017-12-06</na>
2	352.24	3e-083	672.0	67%	5YJS- A	2.16 A	XRAY vicilin-like antimicrobial peptides 2-2 [Capsicum annuum] <na> (370 residues with quality score 0.784), released 2018-10-10</na>
3	346.19	1e-084	753.0	65%	5E1R- E	2.65 A	XRAY 7S vicilin [Carya illinoinensis] <b3stu4_caril(369-792)> (358 residues with quality score 0.712), released 2016-08-10</b3stu4_caril(369-792)>
4	320.10	9e-108	604.0	68%	2EA7- C	1.80 A	XRAY 7S globulin-1 (Fragment) [Phaseolus angularis] <a4pi98_phaan(1-434)> (379 residues with quality score 0.778), released 2008-02-05</a4pi98_phaan(1-434)>
5	319.40	2e-076	612.0	69%	5CAD- A	1.49 A	XRAY SM80.1 Vicilin [Solanum melongena] <a0a158rfr1_solme(1-392)> (379 residues with quality score 0.759), released 2016-04-06</a0a158rfr1_solme(1-392)>

Table S1 Templates utilized for the homology model for vicilin from Theobroma cacao

Figure S1 Chronological order of the steps taken to produce the final model of the vicilin storage protein.



Figure S2: Calculated and postulated phosphorylation sites (green) in the storage protein vicilin from *Theobroma cacao* (Entry: Q43358; 60.8 kDa/525 amino acids; reviewed version from the database UniProt - http://www.uniprot.org/; 28.01-2019). The data indicates that the reported modifications (orange) at positions 232 (Thr), 235 (Ser) and 240 (Ser), as reported in [50] are most probably occurring after the proteolytic processing of the 66-kDa precursor [9,22].

1: Vincillin_Monomer5vF5_a03	5	10	15	20	25	30	35	40	45	50	55	60	65
1: Vincillin_Monomer5vF5_a03	73	78	83	88	93	98	103	108	113	118	123	128 N F	133 R <mark>S</mark> EEEE
1: Vincillin_Monomer5vF5_a03	141 GQQRNNPY	146 YFPKR	151 RSFQT	156 RFRDE	161 EGNFK	166 I L Q R F	171 AENSF	176 PPLKGI	181 ND <mark>Y</mark> RL	186 . AMFE A	191	196 FILPHI	201 HCDAEA
1: Vincillin_Monomer5vF5_a03	209] I <mark>Y</mark> F V T N G K	214 GTITF	219 VT <mark>H</mark> EN	224 Kesyn	229 IVQRG <mark>T</mark>	234 VV <mark>S</mark> VF	239 PAG <mark>ST</mark> V	244 / <mark>Y</mark> VVSC	249 DNQEK	254	259	264 VN <mark>S</mark> PGI	269 XYELFF
1: Vincillin_Monomer5vF5_a03	277 PAGNNKPE	282 S <mark>YY</mark> GA	287 F <mark>SY</mark> EV	292	297 NTQRE	302 KLEEI	307 LEEQF	312 RGQKRC	317 QGQQQ	322 SMFRKA	327 KPEQ	332 I RA I S(337 QQATSP
1: Vincillin_Monomer5vF5_a03	345]R <mark>H</mark> RGGERL	350 AINLL	355 <mark>S</mark> Q <mark>S</mark> PV	360 <mark>YS</mark> NQN	365 IGRFFE	370 ACPED	375 DF <mark>S</mark> QFC	380 NMDVA	385 VSAFK	390	395	400 H <mark>Y</mark> NSK/	405 A <mark>T</mark> FVVF
1: Vincillin_Monomer5vF5_a03	413]VTDG <mark>Y</mark> G <mark>Y</mark> A	418 QMACP	423 <mark>H</mark> LSRQ	428 <mark>S</mark> QGSC	433 SGRQD	438 RREQE	443 EE <mark>S</mark> EE	448 E <mark>T</mark> FGE	453 FQQVK	458 CAPL <mark>S</mark> F	463 PGDVF	468 VAPAGI	473 HAVTFF
1: Vincillin_Monomer5vF5_a03	481]A <mark>S</mark> KDQPLN	486 AVAFG	491 LNAQN	496 NQR I F	501 LAGRP	506 FFLN <mark>H</mark>	511 <mark>I KQN T</mark> N	516 IVIKFT	521 VKA <mark>S</mark> A	526 <mark>Y</mark>	531	536	541

Figure S3: Calculated and postulated glycosylation sites (O-GLcNAc; pink) in the storage protein vicilin from *Theobroma cacao* (Entry: Q43358; 60.8 kDa/525 amino acids; reviewed version from the database UniProt - http://www.uniprot.org/; 28.01-2019). The data indicates that the reported modifications at positions 193 (Thr), 235 (Ser), 338 (Thr) and 474 (Thr) as reported in [50] are most probably occurring after the proteolytic processing of the 66-kDa precursor [9,22].

	5	10	15	20	25	30	35	40	45	50	55	60	65
1: Vincillin_Monomer5vF5_a03													
	72	77	82	87	92	97	102	107	112	117	122	127	132
1: Vincillin_Monomer5vF5_a03													NR <mark>S</mark> EE
	139	144	149	154	159	164	169	174	179	184	189	194	199
1: Vincillin_Monomer5vF5_a03	EEGQQ	RNNPYY	FPKRR	SFQTR	RDEEG	NFKI	LQRFAE	NSPPL	KGINE	YRLAM	FEANP	NTFIL	PHHCD
	206	211	216	221	226	231	236	241	246	251	256	261	266
1: Vincillin_Monomer5vF5_a03	AEAIYI	VTNGK	GTITF	VTHEN	KESYNV	/QRGT	V V <mark>S</mark> V P A	GSTVY	VVSQ	DNQEKL	TIAVL	ALPVN	N <mark>S</mark> PGKY
	273	278	283	288	293	298	303	308	313	318	323	328	333
1: Vincillin_Monomer5vF5_a03	ELFFP	AGNNKP	ESYYG	AF <mark>S</mark> YE	LETVE	NTQR	EKLEEI	LEEQF	GQKR	QGQQG	MFRKA	KPEQI	RAISQ
	340	345	350	355	360	365	370	375	380	385	390	395	400
1: Vincillin_Monomer5vF5_a03	QAISPI	RHRGGE	RLAIN	L L SQ <mark>S</mark> I	^{SVY} SNC	NGRF	FEACPE	DF <mark>S</mark> QF	QNMD	AVSAF	KLNQG	AIFVE	PHYNSK
4. Minerillin, Managerer Fuff, 200	407	412	417	422	427	432	437	442	447	452	457	462	467
1: VInciliin_MonomerSVF5_a03	ATFVVI	VIDGY	GYAQM	ACPHL	skusud	5056	RUDKKE	QEEES	DEEEIR	GEFQU	VKAPL	BPGD	FVAPA
	474	470	494	490	404	400	504	500	514	510	524	520	534
1: Vincillin Monomer5vE5, a03		FLASKD										329	334
1. VInchini_Wohomer3VF5_a03		FRAND	GELNA	VAPOLI	MAGNING	RATEL	AGKEEE	LINHKG		INFIVE			