

Chapter 327

Molecular Cloning, Sequence Analysis of Thioesterases from Wintersweet (*Chimonanthus Praecox*)

Li-Hong Zhang, Qiong Wu, Xian-Feng Zou, Li-Na Chen,
Shu-Yan Yu, Chang-Cheng Gao and Xing Chen

Abstract A coding region of wintersweet cDNA was cloned via screening a wintersweet cDNA library. The *CpFATB* cDNA is 1110 bp in length with an open reading frame (ORF) of 1107 bp encoding a protein of 369 amino acids. Molecular weight and isoelectric point of the protein is 41.72 kD and 7.72, respectively. Bioinformatics analysis shows that a signal peptide with 49 amino acid residues and Protein Localization Sites exists in the chloroplast stroma. This classifies the protein as stable.

Keywords Thioesterases · Bioinformatics · Wintersweet

327.1 Introduction

Fatty acyl–acyl carrier protein thioesterases (FAT) was initially identified in the *Escherichia coli* and has been found to be widespread in eukaryotes, bacteria, and archaea and to be involved in a range of cellular processes including fatty acid biosynthesis [1]. Two different classes of FATs have been described in plants, based upon their amino acid sequence and substrate specificity, namely FATA and FATB [2]. The FATA thioesterases have the highest *in vitro* activity for 18:1-ACP substrates and exhibit a much lower activity for saturated acyl-ACP substrates [3]. Conversely, FATBs prefer saturated acyl group substrates, but also have activity for unsaturated acyl-ACPs [4]. In the current study, a putative acyl–acyl carrier protein thioesterase cDNA (*CpFATB*) was cloned by screening a cDNA library of wintersweet (*Chimonanthus praecox*). The Bioinformatics Toolkit is a platform that integrates a great variety of tools for *CpFATB* sequence analysis to predict the possible functions, inquiry this protein properties.

L.-H. Zhang (✉) · Q. Wu · X.-F. Zou · L.-N. Chen · S.-Y. Yu · C.-C. Gao · X. Chen
Key Laboratory of Agricultural Products Processing, Changchun University,
Changchun, People's Republic of China
e-mail: zhlh2005@sohu.com

327.2 Methods

327.2.1 *CpFATB Cloning and Analysis*

A wintersweet cDNA library was constructed using a SMARTTM cDNA Library Construction Kit from wintersweet corolla material (Sambrook et al. 1989). The coding region of CpFATB was amplified by PCR from the wintersweet cDNA library using the following primers

sense primer: 5'-GCTCTAGAACCATGGCCGCTACT,
anti-sense primer: 5'-CCGCTCGAGTCTTTCATTCATTC ATCACAA.

Thirty thermal cycles were carried out, each consisting of 45 s at 95 °C, 1 min at 54 °C, and 90 s at 72 °C in an automatic thermal cycler. PCR products were separated on a 1 % agarose gel and visualized under UV light.

327.2.2 *Bioinformatics Analysis*

Physicochemical properties and molecular features of CpFATB were predicted by bioinformatic approaches including physical and chemical properties analysis, hydrophobicity analysis, domain analysis, phylogenetic tree analysis and subcellular localization analysis. Homology searches using the program BLAST (Basic Local Alignment Search Tool) from NCBI (National Center for Biotechnology Information, Washington, D.C.). BLAST can be used to infer functional and evolutionary relationships between sequences as well as help identify members of gene families. Protein Secondary Structure Prediction using SOPMA from <http://npsa-pbil.ibcp.fr/>. The Hydrophobicity profile predicted of FAT amino acid sequence by DNAMAN and MEGA4x1 software. SignalP server predicts the presence and location of signal peptide cleavage sites in amino acid sequences. PSORT software predicts protein subcellular location.

327.3 Results

327.3.1 *Cloning and Characterization of CpFATB cDNA Sequence*

A coding region of wintersweet cDNA (CpFATB) was cloned via screening a wintersweet cDNA library. The CpFAT cDNA is 1107 bp from a single open reading frame, predicting a 369 amino acid polypeptide (Fig. 327.1). Sequence homology was investigated using NCBI-BLAST provided by the National Center

Fig. 327.1 The ORF nucleotide sequence and putative amino acid sequence of CpFATB

```

1   M S M I A S S V G A A F F P A Q G I I K
21  S K P A G L H V K A N G R A S P S I D G
41  P K V T V G L E G T N A S S T R K F M N
61  L L P D W S M L L A A F T T I F E K Q K
81  V V V D Q F R F G H D R L V Y S E N F T
101 I R S Y E I G A D Q T A S I E T V M N L
121 L Q E T G I N C F R S L G L L L D G F D
141 S T V E M C K R D L I W V V T R M Q V I
161 V D H Y P S R G D T V E V E T H C G A Y
181 G K H G H R R E W L I R N S K T G Q I L
201 T R A T S V L V V M N K R T R R L S I L
221 P D E V R R E L E P Y F M E N L S V M K
241 D Q G R K L P K V D H S I A D Y V R Q G
261 L T C Q W S D L D I N Q H V N H I K Y V
281 K W I F E S V P V S I L E S H E I S S M
301 T L E F K R E C G K D S M L Q S L T A V
321 V S G R R V D G S V E E T D V E F Q H L
341 L Q L E D G P E V M R G T T K W R P K S
361 T L F P N S I S H *

```

for Biotechnology Information. Sequence analysis showed that the gene had typical characteristics encoding FAT protein family, the gene was named as CpFATB (Fig. 327.2).

327.3.2 *CpFAT* Conservative Regional and Phylogenetic Tree

Use of NCBI (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) this gene encoding amino acids are conservative regional analysis, CpFAT belongs the plant acyl–acyl carrier protein (ACP) thioesterases (TEs). Acyl–ACP thioesterase, This family consists of various acyl–acyl carrier protein (ACP) thioesterases (TE) these terminate fatty acyl group extension via hydrolysing an acyl group on a fatty acid (Fig. 327.3).

A simple phylogenetic tree to show the relatedness of CpFATB to other FATs also indicates the protein from wintersweet belongs to FATB class (Fig. 327.4). These results suggest that CpFATB may have similar function with B class members in plant.

Q95G13.1	BacName: Full-Myristoyl-acyl carrier protein thioesterase, chloroplast	303	409	97%	3e-112
AB195249.1	chloroplast acyl-ACP thioesterase [<i>Jatropha curcas</i>] >gb ACCT0568	302	407	97%	7e-112
AD29529.1	chloroplast acyl-ACP thioesterase [<i>Macadamia tetraphylla</i>]	305	406	95%	2e-111
MF_002248850.1	PREDICTED: hypothetical protein [<i>Vitis vinifera</i>]	304	404	95%	6e-111
CB12185.1	unnamed protein product [<i>Vitis vinifera</i>]	303	404	95%	1e-110
AA01982.1	palmitoyl-acyl carrier protein thioesterase [<i>Gossypium hirsutum</i>]	303	403	96%	2e-110
AB21729.1	acyl-ACP thioesterase [<i>Myristica fragrans</i>]	303	400	97%	1e-109
AC057489.1	acyl acyl-carrier-protein thioesterase type B [<i>Camellia oleifera</i>]	303	399	97%	3e-109
AC057489.1	acyl acyl-carrier-protein thioesterase type B [<i>Camellia oleifera</i>]	303	398	97%	7e-109
MF_002513564.1	palmitoyl-acyl carrier protein thioesterase [<i>Ricinus communis</i>] >gb	302	397	91%	1e-108
AD42220.2	palmitoyl-acyl carrier protein thioesterase [<i>Elaeis guineensis</i>]	302	397	99%	1e-108
AC03259.1	acyl acyl-carrier-protein thioesterase type B [<i>Camellia oleifera</i>]	305	396	97%	2e-108
AC037490.1	acyl acyl-carrier-protein thioesterase type B [<i>Camellia oleifera</i>]	303	393	90%	2e-107
CAN1819.1	hypothetical protein [<i>Vitis vinifera</i>]	303	393	90%	2e-107
ABC32311.1	FATB [<i>Populus tomentosa</i>]	303	393	92%	2e-107
AB13939.1	palmitoyl-ACP thioesterase [<i>Elaeis guineensis</i>]	303	391	99%	6e-107
AB19656.1	unknown [<i>Populus trichocarpa</i> x <i>Populus deltoides</i>]	300	390	97%	9e-107
MF_002324962.1	predicted protein [<i>Populus trichocarpa</i>] >gb EF03527.1 predicted	303	389	97%	2e-106
AB038556.1	acyl ACP-thioesterase [<i>Arachis hypogaea</i>] >gb ABC28557.1 acyl A	289	389	92%	4e-106
AB21581.1	acyl-acyl carrier protein thioesterase [<i>Cuphea catophylla</i> subsp. m]	288	388	97%	5e-106
AB038553.1	acyl ACP-thioesterase [<i>Arachis hypogaea</i>]	288	388	92%	5e-106
AB038554.1	acyl ACP-thioesterase [<i>Arachis hypogaea</i>]	288	388	92%	6e-106
AB038555.1	acyl ACP-thioesterase [<i>Arachis hypogaea</i>]	286	386	92%	2e-105
MF_002309244.1	predicted protein [<i>Populus trichocarpa</i>] >gb EE02767.1 predicted	284	384	93%	7e-105
XP_002309243.1	predicted protein [<i>Populus trichocarpa</i>] >gb EE02766.1 predicted	283	383	93%	2e-104

Fig. 327.2 Alignment of the *CpFATB* gene amino acid sequence in GenBank

Fig. 327.3 Analysis of conserved domain

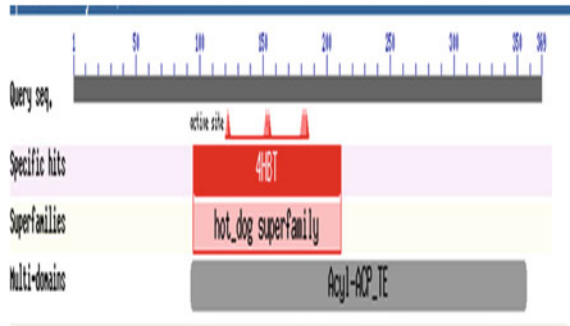
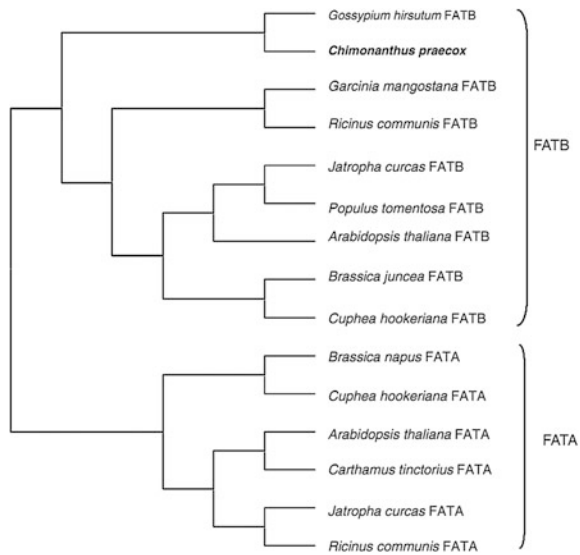


Fig. 327.4 Phylogenetic tree analysis of *Fat*



327.3.3 Fundamental Properties of Protein

Use <http://www.expasy.ch/tools/> Primary structure analysis, A coding region of wintersweet cDNA was cloned via screening a wintersweet cDNA library. The CpFATB cDNA is 1110 bp in length with an open reading frame (ORF) of 1107 bp encoding a protein of 369 amino acids. Molecular weight and isoelectric point of the protein is 41.72 kD and 7.72, respectively. Formula is C1843H2934N5200O548S18. Extinction coefficients are in units of M⁻¹ cm⁻¹, at 280 nm measured in water. The N-terminal of the sequence considered is M (Met). The estimated half-life is 30 h (mammalian reticulocytes, in vitro), >20 h (yeast, in vivo), >10 h (Escherichia coli, in vivo). The instability index (II) is computed to be 38.27. This classifies the protein as stable. Aliphatic index is 86.29.

327.3.4 Protein Secondary Structure Prediction and Analysis of Hydrophobicity

Protein secondary structure prediction using SOPMA. The gene containing 40.33 % Alpha helix (Hh), 4.63 % Beta bridge (Bb) 14.99 % Extended strand (Ee), 40.05 % Random coil (Ee) (Fig. 327.5).

The software of DNAMAN are applied to analyze to these FATs in clustering (Fig. 327.6). The hydrophobic protein of the gene expression of the maximum value of 2.91, is located in 205 bp, the minimum value is -3.60, located in 211, 18 bp as the hydrophobic region.

327.3.5 Chloroplast Transit Peptides and Protein Localization Sites

Use <http://www.cbs.dtu.dk/services/>, The ChloroP server predicts the presence of chloroplast transit peptides (cTP) in protein sequences and the location of potential cTP cleavage sites. The predicted length of the presequence is 49 bp.

Fig. 327.5 Secondary structure prediction of CpFATB

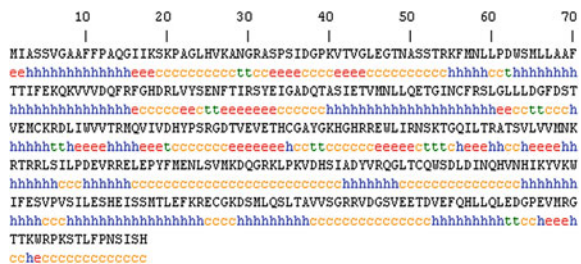
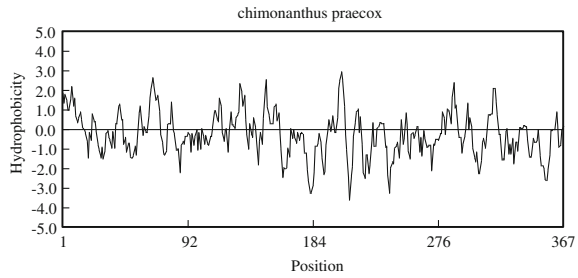


Fig. 327.6 Analysis of Hydrophobicity



Name	Length	Score	cTP	CS-score	cTP-length
Sequence	369	0.543	Y	2.303	49

PSORT—Prediction of Protein Localization Sites. The possibility exists in the chloroplast stroma reached 50.3 %

Final Results:

chloroplast stroma—Certainty = 0.503 (Affirmative) < succ >

mitochondrial matrix space—Certainty = 0.432 (Affirmative) < succ > chloroplast thylakoid membrane—Certainty = 0.255 (Affirmative) < succ > microbody (peroxisome)—Certainty = 0.245 (Affirmative).

327.4 Conclusion

A full-length complementary deoxy-ribonucleic acid (cDNA) of a fatty acyl–acyl carrier protein thioesterase (CpFATB) was isolated from a *Chimonanthus praecox* (wintersweet) cDNA library. Wintersweet is a hardy shrub native to Chinese montane regions and is known to be tolerant to many biotic and abiotic stresses including cold, drought, and a variety of plant pathogens [5]. Bioinformatic analysis also demonstrated that CpFATB contains the hotdog fold/domain [6]. The cDNA is 1,110 nucleotides in length, of which 1,107 bp form a single open-reading frame, predicting a 369 amino acid polypeptide. Molecular weight and isoelectric point of the protein is 41.72 kD and 7.72, respectively. Bioinformatics analysis shows that a signal peptide with 49 amino acid residues and Protein Localization Sites exists in the chloroplast stroma. This classifies the protein as stable. This gene may be of use in the production of drought-tolerant crop species and, as such, is worthy of further characterization and investigation.

References

1. Ohlrogge JB, Kuhn DN (1979) Subcellular localization of acyl carrier protein in leaf protoplasts of *Spinacia oleracea*. *Proc Natl Acad Sci USA* 76:1194–1198
2. Voelker TA, Jones A, Cranmer AM, et al (1997) Broad-range and binary-range acyl–acyl–carrier protein thioesterase suggest an alternative mechanism for medium-chain production in seeds. *Plant Physiol* 114:669–677
3. Dörmann P, Voelker TA, Ohlrogge JB (2000) Accumulation of palmitate in *Arabidopsis* mediated by the acyl–acyl carrier protein thioesterase FATB1. *Plant Physiol* 123:637–643
4. Salas JJ, Ohlrogge JB (2002) Characterisation of substrate specificity of plant FATA and FATB acyl-ACP thioesterases. *Arch Biochem Biophys* 403:25–34
5. Zhang SH, Wei Y, Liu JL, et al (2011) CpCHT1 an apoplastic chitinase isolated from the corolla of wintersweet (*Chimonanthus praecox* communis L.) exhibits both antifungal and antifreeze activities. *Biol Plantarum* 55:141–148
6. Dillon SC, Bateman A (2004) The Hotdog fold: wrapping up a superfamily of thioesterases and dehydratases. *BMC Bioinforma* 5:109