8
Sweet-tasting Proteins

Dr. Ignacio Faus[1] and Ms. Heidi Sisniega[2]
[1] Uriach Group, Avda. Camí Reial 51-57, 08184 Palau-solità i Plegamans (Barcelona), Spain; Tel.: + 34-93-8630351; Fax: + 34-93-8630561; E-mail: bd-ifaus@uriach.com
[2] Uriach Group, Avda. Camí Reial 51-57, 08184 Palau-solità i Plegamans (Barcelona), Spain; Tel.: + 34-902-471511; Fax: + 34-93-8646606;
E-mail: pharm-sisniega@uriach.com

α-amy	α-amylase
β-gal	β-galactosidase
CaMV	cauliflower mosaic virus
Da	dalton
E8	tomato fruit-ripening promoter
FDA	U.S. Food and Drug Administration
Gapdh	glyceraldehyde-3-dehydrogenase promoter
Gla	glucoamylase
Lac	lactose
PgK	3-phosphoglicerate promoter
PR-5	pathogenesis-related group 5 proteins
Trp	tryptophan

1 Introduction

Traditional methods to sweeten foodstuffs, feed, and other consumer products have relied on the use of low-molecular-weight sweetening agents, especially sucrose. In recent years, however, there has been an increasing demand for "low-calorie" sweeteners. Together with this trend, there is also an increase in the demand for "healthy" and "natural" eating products. Therefore, in order to address this need, there is an intense and ongoing search for alternative sweeteners. The alternative sweeteners that have been developed show a very intense sweetening profile. They are several orders of magnitude sweeter than sucrose on both a molar and weight basis. Therefore, in order to supply the same "sweetening" effect as sucrose, these compounds can be used in minute amounts. This results in almost negligible addition to the calorie count, as well as a product that does not contribute to tooth decay. Moreover, the alternative and intense sweetening additives also can be used in diabetic foods, since they do not trigger a demand for insulin in these patients. Finally, some intense sweeteners also have been found to be strong flavor enhancers, thus expanding their range of applications.

Currently, there are six alternative, high-intensity sweeteners that have been approved by European Union regulatory bodies (Grijspaardt-Vink, 1996): aspartame, saccharin, cyclamate, neohesperidine DC, acesulfame-K, and thaumatin. The first five compounds in this list are low-molecular-weight entities, obtained by traditional organic synthesis technology, although it should be pointed out that aspartame is a peptide (albeit an unnatural one). The last one, thaumatin, is a protein, i.e., a "natural" product.

Besides thaumatin, several other sweet-tasting proteins exist in nature. Some of them have been isolated, purified, and characterized. Their genes have been cloned, and in some cases recombinant versions of the natural protein have been obtained. In this chapter we will review our current knowledge of these sweet-tasting proteins.

2 Historical Perspectives

The existence in nature of sweet-tasting proteins has been known for many years. The first such protein was initially identified in 1968 (Kurihara and Breidler, 1968). It was

termed miraculin and was isolated from a West African plant. Since then, other sweet-tasting proteins have been isolated: thaumatin and monellin in 1972, mabinlin in 1983, pentadin in 1989, curculin in 1990, and brazzein in 1994.

All of these proteins have been found in the fruits of tropical plants, and natives of the areas where the plants producing these proteins grow naturally have frequently used them to sweeten their foodstuffs. Yet, it has been only in the last 30 years that efforts have been made to commercially exploit the use of sweet-tasting proteins. With the commercialization of thaumatin as both a sweetener and a flavor enhancer, there has been increased interest in these compounds. In recent years, an extraordinary number of sweet-tasting proteins have been discovered, studied, purified, and characterized. Their genes have been cloned and sequenced and, in many cases, have been expressed in foreign hosts.

As mentioned before, there are six known sweet-tasting proteins: thaumatin, monellin, mabinlin, pentadin, brazzein, and curculin. A seventh protein, miraculin, is by itself not a sweet-tasting protein, but rather a taste-modifier. Nonetheless, it will be included in this review because of its importance. Of these seven proteins, only one, thaumatin, has become a marketed product (Witty and Higginbotham, 1994). Several of the others are undergoing thorough testing, and it is likely that at least some of them will reach the market soon. Table 1 shows a comparison of their properties.

Tab. 1 Comparison of thaumatin, monellin, mabinlin, pentadin, brazzein, curculin and miraculin

	Thaumatin	***Monellin***	***Mabinlin***	***Pentadin***	***Brazzein***	***Curculin***	***Miraculin***
Source	*Thaumatococcus danielli* Benth	*Dioscoreophyllum cumminsii* Diels	*Capparis masakai* Levl	*Pentadiplandra brazzeana* Baillon	*Pentadiplandra brazzeana* Baillon	*Curculingo latifolia*	*Richadella dulcifica*
Geographic distribution	West Africa	West Africa	China	West Africa	West Africa	Malaysia	West Africa
Variants	I, II, a, b, c[a]	–	I, II-a, III, IV[a]	–	–	–	–
Sweetness factor (weight basis)	3,000	3,000	100	500	2000	550	–
Molecular mass (active form, kDa)	22.2	10.7	12.4	12.0[b]	6.5	24.9	98.4
Amino acids	207	45 (A chain) 50 (B chain)	33 (A chain) 72 (B chain)	?	54	114	191
Active form	Monomer	Dimer (A + B)	Dimer (A + B)	?	Monomer	Dimer (A+A)	Tetramer (A+A+A+A)

Source: Adapted from Kurihara (1994). [a] At least five different forms of thaumatin (Lee et al., 1988) and four different forms of mabinlin (Nirasawa et al., 1994) have been identified. [b] A chromatographic fraction containing a 12-kDa protein was sweet. This same fraction, when subjected to electrophoresis under non-reducing conditions showed bands in the region between 22 and 41 kDa, suggesting the presence of subunits.

3
Thaumatin

The thaumatins are a family of very sweet proteins present in the fruits of the tropical plant *Thaumatococcus daniellii* Benth (for a

Fig. 1 The *Thaumatococcus danielli* Benth plant. The leaves of a *T. Danielli* Benth shrub are shown. The inset shows an opened fruit. The creamy-white arils taste sweet and contain the thaumatin proteins

thorough review on all aspects of thaumatin, see Witty and Higginbotham, 1994). The plant was originally described almost 150 years ago (Daniell, 1855). It grows in the rain forest areas of West Africa, predominantly Sierra Leone to Zaire. Figure 1 shows a picture of the plant and the fruit. The sweet-tasting component of *T. daniellii* Benth was identified as a protein by van der Wel and Loeve (1972), who were the first to isolate it. Since then, because of their intense sweetness (3000 times sweeter than sucrose on a weight basis), the thaumatins have been identified as a potential replacement for sucrose. Currently, thaumatin is extracted from the fruits of the plant and commercialized as both a sweetener and a flavor enhancer, being present in foodstuffs and feed products. It has undergone thorough toxicological and safety tests and has received approval by several regulatory bodies.

Due to the difficulties in obtaining thaumatin from its natural source, attempts have been made to cultivate *T. daniellii* Benth in habitats different from the natural one. In these foreign habitats (including greenhouses), the plant grows but bears no fruit. Therefore, as an alternative process to the production of thaumatin from its natural source, active efforts have been made to produce thaumatin in recombinant hosts (Witty and Higginbotham, 1994).

All the forms of thaumatin are intensely sweet, and have 207 amino acids. The two predominant forms, thaumatin I and II differ by five amino acids (Figure 2).

The thaumatin II gene has been cloned and sequenced (Edens et al., 1982). Attempts have been made to produce recombinant thaumatin in several microorganisms and in transgenic plants, using either the cloned natural gene for thaumatin II or, alternatively, a synthetic gene that contains

```
ThaumatinI     1 ATFEIVNRCSYTVWAAASKGDAALDAGGRQLNSGESWTINVEPGTNGGKIWARTDCYFDD
ThaumatinII    1 ATFEIVNRCSYTVWAAASKGDAALDAGGRQLNSGESWTINVEPGTKGGKIWARTDCYFDD
                 ******************************************** **************

ThaumatinI    61 SGSGICKTGDCGGLLRCKRFGRPPTTLAEFSLNQYGKDYIDISNIKGFNVPMNFSPTTRG
ThaumatinII   61 SGRGICRTGDCGGLLQCKRFGRPPTTLAEFSLNQYGKDYIDISNIKGFNVPMDFSPTTRG
                 ** *** ******** ************************************ *******

ThaumatinI   121 CRGVRCAADIVGQCPAKLKAPGGGCNDACTVFQTSEYCCTTGKCGPTEYSRFFKRLCPDA
ThaumatinII  121 CRGVRCAADIVGQCPAKLKAPGGGCNDACTVFQTSEYCCTTGKCGPTEYSRFFKRLCPDA
                 ************************************************************

ThaumatinI   181 FSYVLDKPTTVTCPGSSNYRVTFCPTA
ThaumatinII  181 FSYVLDKPTTVTCPGSSNYRVTFCPTA
                 ***************************
```

Fig. 2 Sequence alignment of the most predominant forms of thaumatin, forms I and II. As described in the text, they differ by 5 amino acids, which are highlighted in the figure.

codons optimized for expression in the particular host being used. The results of these attempts are summarized in Table 2. Overbeeke (1989) has suggested that given the current price of thaumatin obtained by extraction from its natural source, microbial production of this protein would be economically feasible only if the recombinant microorganism could produce 1 g L^{-1} of product. So far, none of the attempts at expressing recombinant thaumatin have been capable of reaching these expression levels.

Efforts have been made to express thaumatin in transgenic plants. Witty (1990) obtained transgenic *Solanum tuberosum* that produced recombinant thaumatin with a sweet phenotype. The production of thaumatin in plants that are more amenable to cultivation under non-tropical conditions might be an alternative to the extraction of this product from *T. daniellii* Benth fruits.

Thaumatin has a lingering aftertaste that might not be considered acceptable to some palates. In order to circumvent this problem, Blair et al. (1990) performed site-directed mutagenesis on a thaumatin I synthetic gene, resulting in a collection of thaumatin I analogues containing selected amino acid substitutions. Some of these variants were shown to have a reduced life in their sweet aftertaste, resulting in a product with more marketable qualities.

4 Monellin

Monellin is present in the red berries of the West African plant *Dioscoreophyllum cumminsii* Diels. This protein, originally purified by Morris and Cagan (1972), is about 3000 times sweeter than sucrose on a weight basis. Unlike the single-chain thaumatin, monellin consists of two polypeptides of 45 and 50 amino acid residues, respectively, that are associated through non-covalent interactions. The solution structure of monellin has been elucidated (Lee et al., 1999; Sung et al., 2001). Structure-function studies also have revealed that certain amino acids in the protein are critical to the sweet-tasting properties of monellin (Mizukoshi et al., 1997).

Attempts have been made to express monellin in both microorganisms and transgenic plants (see Table 3 for a summary). Monellin has been shown to lose its sweet-

Tab. 2 Most relevant published results on the expression of recombinant thaumatin

Host	*Promoter*[a]	*Secretion*	*Yield*	*Sweet phenotype*	*Reference*
E. coli	Trp/lac	No	Very low	No	Edens et al. (1982)
S. cerevisiae	PgK	No	Low	No	Lee et al. (1988)
K. lactis	Gapdh	Yes	Low	No	Edens and van der Wel (1985)
B. subtilis	α-amy	Yes	1 mg L^{-1}	Yes	Illingworth et al. (1988)
S. lividans	β-gal	Yes	0.2 mg L^{-1}	?	Illingworth et al. (1989)
P. roquefortii	Gla	Yes	1–2 mg L^{-1}	Yes	Faus et al. (1997)
A. awamori	Gla	Yes	5–7 mg L^{-1}	Yes	Faus et al. (1998)
S. tuberosum	CaMV	No	Low	Yes	Witty (1990)

[a] Trp/lac: *E. coli* tryptophan and lactose promoters; PgK: *S. cerevisiae* 3-phosphoglycerate promoter; Gapdh: *K. lactis* glyceraldehyde-3-phosphate-dehydrogenase promoter; α-amy: *B. subtillis* α-amylase promoter; β-gal: *S. lividans* β-galactosidase promoter; Gla: *A. niger* glucoamylase promoter; CaMV: cauliflower mosaic virus promoter for the 35S RNA.

Tab. 3 Monellin gene expression in biotechnology

Host	*Promoter*[a]	*Yield*	*Reference*
E. coli	Trp	Low	Kim et al. (1989)
Tomato	E8	23.9 μg monellin per g weight wet	Peñarrubia et al. (1992)
Lettuce	CaMV	Low	Peñarrubia et al. (1992)
C. utilis	gapdh	10 mg monellin per g weight wet	Kondo et al. (1997)
S. cerevisiae	gapdh	Low	Kondo et al. (1997)

[a] Trp: *E. coli* tryptophan promoter; E8: tomato fruit-ripening specific promoter; CaMV: cauliflower mosaic virus promoter for the 35S RNA; gapdh: *C. utilis* or *S. cerevisiae* glyceraldehyde-3-phosphate dehydrogenase promoter. All the recombinant monellin proteins have been expressed as single-chain analogues.

ness when heated above 50°C under acidic pH. To circumvent this lack of stability, Kim et al. (1989) prepared single-chain analogues of monellin, where the two chains were joined together by several linkers. One of these single-chain analogues was expressed in *Escherichia coli* and proved to be as potent a sweetener as the natural product, as well as being more stable under extreme heat and pH conditions.

Monellin also has been expressed in transgenic plants, although at very low levels (Peñarrubia et al., 1992). Recently, attempts have been made to express monellin in the yeasts *Saccharomyces cerevisiae* and *Candida utilis* (Kondo et al., 1997).

Given the complexity of recombinant production of a two-chain protein, an artificial protein, termed MNEI, has been designed. It is comprised of 96 amino acids by linking with glycine-phenylalanine dipeptide chains A and B of monellin (Spadaccini et al., 2001). This protein also has been shown to elicit sweetness.

5 Mabinlin

The Chinese plant *Capparis masaikai* Levl bears fruits that also contain a sweet-tasting protein, first characterized by Liu et al. (1993). It is suspected that at least four sweet-tasting polypeptides exist in the *C. masaikai* Levl plant. The most studied protein has been named mabinlin II. This protein also is comprised of two polypeptide chains, of 33 and 72 amino acids, respectively, which are tightly associated through non-covalent interactions. It is about 100 times sweeter than sucrose on a weight basis. The possible sweetness of the mabinlin precursor has not been investigated.

Other variants of mabinlin have been isolated (Nirasawa et al., 1994). They have been named mabinlin I-1, III, and IV. The heat stability (and hence the sweet profile) of the mabinlin proteins is quite different and has been ascribed to single amino acid replacements. For example, it has been suggested that the difference in the heat stability of the different mabinlin homologues is due to the presence of an arginine residue (heat-stable homologue) or a glutamine (heat-unstable homologue) at position 47 in the B-chain (Nirasawa et al., 1994).

The cDNAs representing the four known mabinlin isoforms (I-1, II, III, and IV) have been cloned and sequenced (Sun et al., 1997). Mabinlin IV could be derived from mabinlin III, since the C-terminus of the smaller chain is shorter by 4 amino acids than that of mabinlin III. Sequence analysis also has shown that mabinlin II is originally synthesized as a single-chain protein with a 14-amino-acid linker sequence that connects

both chains. This linker is eventually cleaved in the process of protein maturation (Sun et al., 1996).

Recently, mabinlin has been crystallized (Guan et al., 2000). Attempts have been made to express mabinlin II in transgenic potato tubers (Xiong and Sun, 1996), but no explicit results have been reported yet.

6 Pentadin

Fruits of the plant *Pentadiplandra brazzeana* Baillon, a climbing shrub found in some countries of tropical Africa (such as Gabon), contain a 12-kDa sweet-tasting protein, first isolated by van der Wel et al. (1989). Electrophoretic studies in the presence and absence of 2-mercaptoethanol suggested that the mature protein consists of subunits coupled by disulfide bonds. The sweetness intensity was estimated to be around 500 times that of sucrose on a weight basis, the sweet profile resembling monellin more than thaumatin. No further work has been reported towards the characterization of this sweet-tasting protein.

7 Brazzein

Brazzein is also contained in the fruit of *P. brazzeana* Baillon. It was first isolated by Ming and Hellekant (1994). Since pentadin has not been sequenced, we cannot rule out the possibility that brazzein might be homologous to pentadin and that the two proteins might, in fact, be related or belong to the same family. The molecular mass of brazzein is 6473, and its three-dimensional structure has been solved (Caldwell et al., 1998). Like thaumatin, brazzein is a single-chain protein (54 amino acids). Its sweetness profile remains even after incubation at 353 K for 4 hours, probably because of the compact structure afforded by its four disulfide bridges (Ishikawa et al., 1996).

The brazzein gene has been sequenced (Hellekant and Ming, 1994). Attempts have been made to express this protein in recombinant hosts (Hellekant and Ming, 1994; Tomes, 1997), and its expression in *E. coli* recently has been reported (Assadi-Porter et al., 2000).

8 Curculin

Curculin has been isolated from the plant *Curculigo latifolia*, which grows in some parts of Malaysia (Yamashita et al., 1990). The protein contains 114 amino acid residues, with a calculated molecular mass of 12,491. Curculin is a dimer of two identical polypeptides, which are assembled together through two disulfide bridges. It is 550 times sweeter than sucrose on a weight basis. Curculin is also a taste modifier, having the remarkable property of turning a sour taste (e.g., lemon) into sweet (e.g., orange).

cDNA clones for curculin have been isolated and sequenced (Abe et al., 1992), but to this date there are no reports of expression of this protein in recombinant hosts.

9 Miraculin

The red berries of *Richadella dulcifera*, a shrub native to West Africa, contain a taste-modifier protein called miraculin. Figure 3 is a picture of the plant and its fruit. It was first isolated by Theerasilp and Kurihara (1988). The complete amino acid sequence of miraculin has been determined (Theer-

asilp et al., 1989). It contains 191 amino acid residues, with a calculated molecular mass of 24,600. Native miraculin appears to be a tetramer, held together by several disulfide bridges. Miraculin by itself does not elicit a sweet response. Like curculin, though, it can modify a sour taste into a sweet taste.

The cDNA corresponding to miraculin has been cloned and sequenced (Masuda et al., 1995). A synthetic gene encoding miraculin was assembled and inserted into an *E. coli* expression vector (Kurihara, 1994). Production of recombinant miraculin was detected, but a complete characterization has not been reported.

10 Comparison of Different Sweet-tasting Proteins: The Thaumatin-like Proteins

The properties of the seven sweet-tasting and sweet-modifying proteins are compared in Table 1, and Figure 4 shows the amino acid sequences of the most relevant sweet proteins described. Sequence comparisons between the known genes encoding sweet-tasting proteins have shown no obvious homology sequences. Moreover, the three-dimensional structures of thaumatin, monellin, and brazzein have been resolved (Caldwell et al., 1998 and references therein), and they show little mutual resemblance. The lack of both sequence and structural similarities raises the question of whether these proteins all elicit the sweet-taste response through similar or completely different mechanisms.

Fig. 3 The *Richadella dulcifera* plant, a small evergreen shrub native to tropical West Africa. The picture shows the red berries of this plant called miracle fruits. They contain the taste-modifier protein miraculin.

On the other hand, all of the sweet-tasting proteins that have been isolated so far are found in different tropical rain forests. It has been suggested by van der Wel et al. (1989) that this is no coincidence and that the biosynthesis of these proteins might help their host plants to mimic the sweetness of sucrose. This could make it more likely that wild animals ingest these plants, thus leading to more efficient seed dispersal. Primates, especially chimpanzees, are very efficient seed dispersers. Some sweet-tasting proteins (e.g., thaumatin, monellin, pentadin) have been shown to elicit a sweet response in Old World primates, as measured by recording the summated activity in the chorda tympani proper nerve of a rhesus monkey during stimulation of the tongue by solutions containing different compounds (van der Wel et al., 1989). It is therefore possible that the sweet-taste receptors of primates were the original targets in the evolution of the sweet-tasting proteins. From there, the ability to evoke a sweet taste would have been passed on to humans.

One intriguing aspect of thaumatin is that while it does not share any homology to the other sweet-tasting proteins, it is highly homologous to the so-called PR-5 family of proteins ("pathogenesis-related group 5