PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 4 : C12N 15/00, C07H 21/04 C12N 1/20, C12P 21/00 A23L 1/236	A1	(11) International Publication Number: WO 88/ 10303 (43) International Publication Date: 29 December 1988 (29.12.88)

(21) International Application Number: PCT/KR88/00010

(22) International Filing Date: 2 June 1988 (02.06.88)

(31) Priority Application Number: 064,341

(32) Priority Date: 19 June 1987 (19.06.87)

(33) Priority Country: US

(71) Applicants: LUCKY, LTD. [KR/KR]; 20, Yoido-dong, Yongdungpo-gu, Seoul 150-010 (KR). LUCKY BIOTECH CORPORATION [US/US]; 4560 Horton Street, Emeryville, CA 94608 (US).

(72) Inventor: CHO, Joong, Myung ; 4238 Treat Blvd., Concord, CA 94521 (US).

(74) Agent: KIM, Yoon, Bae; Y.B. Kim, Lees & Associates International Patent & Law Offices, 7th Floor, Garden Tower Bldg., 98-78, Unni-dong, Jongro-gu, Seoul 110-350 (KR).

(81) Designated States: AT (European patent), AU, BE (European patent), BR, CH (European patent), DE (European patent), FR (European patent), GB (European patent), IT (European patent), JP, KR, LU (European patent), NL (European patent), SE (European patent).

Published
With international search report.

(54) Title: PREPARATION OF NOVEL PROTEIN SWEETENERS

(57) Abstract

Novel proteinacious sweeteners are provided comprising an amino acid sequence based on the sequences of monellin subunits. The single protein may be prepared by recombinant techniques, to provide for a stable strong sweetening agent, which may be utilized in a wide variety of contexts.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GA**	Gabon	**MR**	Mauritania
AU	Australia	**GB**	United Kingdom	**MW**	Malawi
BB	Barbados	**HU**	Hungary	**NL**	Netherlands
BE	Belgium	**IT**	Italy	**NO**	Norway
BG	Bulgaria	**JP**	Japan	**RO**	Romania
BR	Brazil	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SN**	Senegal
CH	Switzerland	**LI**	Liechtenstein	**SU**	Soviet Union
CM	Cameroon	**LK**	Sri Lanka	**TD**	Chad
DE	Germany, Federal Republic of	**LU**	Luxembourg	**TG**	Togo
DK	Denmark	**MC**	Monaco	**US**	United States of America
FI	Finland	**MG**	Madagascar		
FR	France	**ML**	Mali		

PREPARATION OF NOVEL PROTEIN SWEETENERS

TECHNICAL FIELD

Novel proteinacious sweeteners are provided produced by recombinant techniques.

BACKGROUND ART

Monellin is an intensely sweet material present in the sap of "Serendipity Berries", the fruit of the West African plant, Dioscoreophyllum comminisii. The material has been purified to homogeneity and shown to be a basic protein with a molecular weight of about 1.1×10^4 and is completely free of carbohydrate. Monellin is the first well characterized material among several sweet or taste modifying substances found in tropical plants. It has been characterized and shown to have two subunits of about the same size held together by non-covalent bonds. The two subunits are not identical and the flavor modifying ability of monellin is dependent upon the presence of both subunits and a single mercaptan

group, which if blocked abolishes the sweetness.

Because of the uncertainties and cost of extracting natural products from plant sources, an alternative route to the production of protein sweetners is of substantial interest. Recombinant techniques offer an opportunity to synthesize proteins of varying types. However, in employing recombinant techniques, one is required to develop a strategy for producing the gene, demonstrate successful expression of the protein in a cellular host, and isolate a product which is shown to have physiological activity. In many instances, it is necessary or desirable to modify the naturally occurring sequence, which substantially increases the uncertainties of success of the production of a useful product.

The relevant literatures are as follows :

Morris et al., J. Biol. Chem. (1973) 248 : 534-539 describe the characterization of monellin. See also Cagan, Science (1973) 181 : 32-35 ; Wlodawer and Hodgson, Proc. Natl. Acad. Sci. USA (1975) 72: 398-399 ; Bohak and Li, Biochimica et Biophysica Acta (1976) 427 : 153-170 ;

Hudson and Bieman, Biochem. Biophys. Res. Comm. (1976) 71: 212-220 ; Jirgenson, Biochem. Biophys. Acta (1976) 446 : 255-261 ; and Van der Wel and Loeve, FEBS Lett. (1973) 29: 181-183 for further characterization. U.S. Patent No. 3,998,798 describes the preparation of natural monellin.

SUMMARY OF THE INVENTION

Novel DNA open reading frames, constructs employing the open reading frames and expression systems are provided for expressing novel proteins having sweetening capability, where the proteins employ a substantial proportion of the amion acid sequence of monellin. The proteins are a single molecule as distinct from the two subunits of monellin, so as to define a single sequence.

DISCLOSURE OF INVENTION

Novel proteinacious sweeteners, methods for their production, and intermediates used in the methods, particularly nucleic acid intermediates are provided. The sweeteners are modelled after the naturally occurring sweetener monellin, where the two independent subunits of

monellin are joined together in a continuous sequence. The two subunits may be joined end to end, by modifying the amino acids adjacent the juncture between the two subunits, or by introducing a short bridge extending the sequence.

The amino acid sequence will be in substantial part the amino acid sequence of the subunits of monellin, usually having at least about 80% homology with the monellin sequence, more usually at least about 90% homology with the monellin sequence. The sequence may be varied by insertions, deletions, or substitutions, where insertions and deletions will usually not exceed about 9 amino acids, more usually not exceed about 6 amino acids, and substitutions may be conservative or non-conservative, where the following table indicates as conservative substitutions those amino acid on the same line. For the most part, polar amino acids will not be substituted for non-polar amino acids and aliphatic amino acids will not be substituted for aromatic amino acids.

Amino Acids

Aliphatic

non-polar

G, A, P

V, I, L

polar

neutral

S, T, C, M, N, Q

charged

basic

K, R, H

acidic

D, E

Aromatic

F, W, Y

For the most part, conservative substitutions will be preferred and the cysteine and methionine at positions 41 and 42 of subunit II will be retained. (In the numbering of amino acids relating to monellin as distinct from the subject constructions, the numbering of the natural monellin subunits will be employed.)

The protein may have either subunit II or subunit I as the N-terminus, particularly subunit II. Depending on the construction, the product may or may not have an N-terminal methionine. The two subunits may be joined by a short bridge, usually of not more than 10, usually not more than 8 amion acids, or may be joined directly, or preferably the amino acids at the juncture will be modified. The amino acids at the juncture forming the bridge will provide for a polar juncture, that is, at least 50 number %, usually at least about 75 number % of the amino acids will be polar and conveniently, at least about 25 number %, generally about 50 number % will be amino acids naturally present at the subunit terminal. The amino acids may come from a loop of subunit I.

In referring to the juncture, the juncture will include as a bridge not more than about 10, usually not more than about 6 amino acids of the naturally occurring sequence of the subunits. For the joining of the C-terminus of subunit II with the N-terminus of subunit I, the juncture will be at Ile(46) of subunit II and Gly(6) of subunit I with the intervening amino acids, if any, as the bridge.

Where subunit II is the N-terminus, one or more of the wild-type amino acids at the juncture may be removed or substituted, usually not more than about 10 amino acids will be removed or substituted, more usually not more than about 6 amino acids. Generally not more than 75% of the removed or substituted amino acids will be associated with one of the subunits.

Bridges of interest will include :

$$aa_x^1 - aa_x^2 - aa_x^3 - aa_x^4 - aa_x^5 - aa_x^6 - aa_x^7 - aa_x^8$$

where only one amino acid need be present, and the individual amino acids as defined are as follows :

aa^1 is A, D, E, K, R or Y ;

aa^2 is Y, A, D, E, N, Q, R, T or S ;

aa^3 is N, Q, S, T, D, E, R or Y ;

aa^4 is F, W, Y, S, T, D, E, K or R ;

aa^5 is D, E, K, R, L or T ;

aa^6 is D, E, V, I, L, K or R ;

aa^7 is G, A, V, I, L, K or R ;

aa^8 is K or R ;

where x is 0 or 1, at least one x being 1.

Compositions of interest include sequences where :

aa^1 is Y or E ;

aa^2 is D, E, Y or K ;

aa^3 is N, T, A or Y ;

aa^4 is R, S, K or E ;

aa^5 is E, D or T ;

aa^6 is K, D or R ;

aa^7 is G, I or L ;

aa^8 is K or R ;

For sequences having the first two amino acids Y and E, there may be from 0 to 4 x's plus y that are 0, while for chains having different amino acids as the first two amino acids there may be form 0 to 5 x's plus y that are 0. That is, the above chains will usually be from 3 to 8, more usually 4 to 8 amino acids.

Of particular interest is removal of the phenylalanine where the juncture will be Y-E-N-E-R-E-I-K. Other bridges include Y-E-N-R-E-D-I-K ; Y-K-T-R-E-D-I-K ; Y-E-R-E-I-K ; Y-E-N-I-K ; Y-E-I-K ; Y-Y-A-S-D-K-L-K ; Y-A-S-D-K-L ; Y-A-S-D-K ; Y-S-D-K ; E-D-Y-K-T-R-G-R ; and E-D-Y-T-R. Usually there will be at least one Y, E, D, K or R present in the chain, more usually at least one of E, D, K or R. Preferred amino

acids for the bridge are Y, I, S, T, D, E, K, R, N or Q, where greater than 50% of the amino acids of the bridge will be selected from this group.

The total number of changes, insertions, deletions, and substitutions will generally not exceed a total of 12, more usually 10 amino acids, where substitutions will be counted first, followed by deletions or insertions to arrive at the total.

The subject compositions can be prepared by recombinant technology. In order to provide for expression, a gene must be provided. Sequences for subunits I and II may be obtained from the natural source as genomic DNA or cDNA. Alternatively, a strategy may be developed for preparing single stranded sequences which may be ligated together to provide the desired double-strand. The sequences are designed to minimize heteroduplexing, so as to substantially insure that the resulting ligated double-strand DNA has the proper open reading frame. The strategy employed in the Experimental section is particularly preferred.

Once the double-stranded sequence has been designed, the various single-stranded fragments may be synthesized and

ligated together in accordance with conventional techniques. The coding region may then be used to prepare an expression cassette. The expression cassette will comprise a transcriptional and translational initiation regulatory region at the 5' terminus in the direction of transcription of the open reading frame and a translational and transcriptional termination region at the 3' terminus of the open reading frame in the direction of transcription.

Today, there are many vectors which include transcriptional and translational regulatory regions of a wide variety of genes, where the initiation and termination regions are separated by a polylinker, so that an open reading frame may be inserted between the initiation and termination regions to be under their transcriptional and translational regulation. Depending upon the particular expression host, vectors are commercially available or have been described in the literature and may be prepared from available segments having the necessary functions. For the most part, the vectors will include a replication system, which may be low or high copy number, usually having copy numbers of fewer than about 1000, although in certain situations, runaway vectors may be

employed. Alternatively, instead of having extrachromosomal maintenance, one may provide for homology between the vector and the host genome, to enhance the opportunity for integration. Where integration is involoved, one may provide for an amplifying gene in tandem with the expression cassette. Amplifying genes include dihydrofolate reductase, the metallothioneins, thymidine kinase, or the like. These genes will be accompanied with an appropriate transcriptional and translational regulatory region to provide for expression in the expression host. With prokaryotes, a polycistronic message may be employed, where the amplifying gene and the sweetener gene may be under the regulation of the same transcriptional and translational regulatory regions.

Usually, the vector will include a marker which allows for selection of those host cells containing the expression cassette for expressing the subject protein. Markers may include biocide resistance, particularly from antibiotics, heavy metals, or the like; complementarity to an auxotrophic host to provide prototrophy; resistance to viral infection; etc. One or more markers may be present, particularly where one marker is used for insertion of the construct, so that

loss of the particular capability will indicate the presence of the expression cassette.

Transcription intiation regions which may be employed include those associated with such genes as trp, lac, gal, his ; or viral promoters such as λP_L, λP_R and P_4 promoters, yeast promoters such as those associated with the genes adh-1, adh-2, mat, gal, pgk, pyk, pho5, mA, gapdh, amy or dbfr, etc. Joint promoter regions may be employed, such as the tac, adh-2/gapdh, gal/gapdh, cye/gal transcriptional initiation regions. See, for example, U. S. Patent Nos. 4,418,149 ; 4,304,863 ; 4,350,764 ; 4,363,877 and 4,366, 246.

Specialty sequences may also be used, such as enhancers, to enhance the level of transcription. A wide variety of enhancers have been reported in the literature associated with a wide variety of genes in a range of hosts.

Another specialty sequence is a signal leader, which provides for secretion and processing of the protein. Again, a large number of signal leaders have been described in the literature and have been shown to be effective with a broad spectrum of proteins. Thus, if one signal leader is not efficient, other available signal leaders may be tried. As

exemplary of signal sequences are U.S. Patent Nos. 4,336,336; 4,338,397 ; and 4,546,082.

The signal sequence, if any, will be joined to the open reading frame coding for the sweetener at its 5' terminus and will provide the methionine codon, where the open reading frame will be in proper reading phase with the methionine. Thus, the precursor protein will include the signal sequence, the processing signal, and the protein sweetener in going from the N- to the C-terminus, where the signal sequence and processing signal will be enzymatically removed as the precursor protein is secreted. A number of processing signals are known, based on the host and the enzymatic system employed for secretion and processing whereby the signal sequence is removed.

A wide variety of hosts may be employed, both prokaryotic and eukaryotic. Common hosts which are exemplary include E. coli, B. subtilis, B. licheniformis, S. cerevisie, K. lactis, N. crassa, Streptomyces, Aspergillis niger, and the like. Other members of each of the genera may also be employed. For the most part, microbial expression hosts will be employed, particularly procaryotic.

Depending upon the nature of the host, various techniques may be employed for transforming the expression host with the expression cassette, either by itself, or as part of a vector or other construct. The introduction of the expression cassette may be as a result of conjugation, transformation, transfection, transduction, fusion, etc. Intact host cells, protoplasts, partially regenerated protoplasts, or the like may be employed for the introduction of the exogenous DNA.

Once the host has been transformed, it may then be grown in a selective medium, so as to select for those hosts having the marker or associated expression cassette. Where antibiotic resistance is involved, the nutrient may contain a level of the antibiotic cytotoxic in the absence of the antibiotic resistance gene. In the case of auxotrophy complementation, the nutrient medium lacks the necessary metabolite. Where the product is produced and retained in the cytoplasm, after sufficient time for the cells to grow, the cells may be lysed and the desired protein obtained by conventional purification procedures. These procedures included liquid-liquid extraction, HPLC,

chromatography, electrophoresis, etc. The product may then be subjected to further purification, such as gel exclusion, chromatography, etc.

The resulting product may be used in a variety of ways as a sweetener. It may be used in canned products, in conjunction with various carbonated drinks, as a powder or liquid for addition to various beverages, such as coffee, tea, or the like, in cooking, chewing gum, toothpaste, mouthwash, dental hygiene products, pharmaceuticals, meat products, e.g. ham, sausage, etc., instant soups, yogurt, desserts, cereals, animal food, etc.

The subject proteanase sweeteners may be formulated as a liquid or powder. As a liquid, other additives may be combined, such as stabilizers, buffers, bactericides, protease inhibitors, or the like. An aqueous medium will mormally be used where the sweetener will be from about 0.1 to 90 weight % of the composition. For powders, various excipients may be added which are conventional food extenders.

Rather than providing the sweetener as an independent product, the expression cassette can be prepared for use

in plants. Particularly, expression cassettes can be prepared where a constitutive or regulated transcriptional initiation region functional in a plant may be employed, so that products, such as fruit, vegetables, melons, or the like may have enhanced sweetening. A wide variety of constructs are described in the literature, demonstrating expression of a wide variety of genes in plants, using either constitutive or regulated transcriptional initiation regions. Transcriptional initiation regions include the various opine initiation regions, such as octopine, mannopine, nopaline, etc. Alternatively, plant viral transcription initiation regions may be employed, such as the cauliflower mosaic virus 35S promoter. Other transcription initiation regions, particularly inducible regions, more particularly regions associated with cell differentiation, include the small subunit or large subunit transcriptional initiation regions of ribulose-1,3-biphosphate carboxylase, fruit specific promoters, heat shock promoters, etc.

The following examples are offered by way of illustration and not by way of limitation.

Example

1. Oligonucleotide Synthesis and Purification

The following oligomers were synthesized using Applied Biosystems 380B DNA Synthesizer.

5' ——> 3'

U1 : TATGGGAGAATGGGAAATTATCGATATTGGACCATTCACTCAAAAC (46mer)

U2 : TTGGGTAAGTTCGCTGTTGACGAAGAAAACAAGATTGGTCAATAT (45mer)

U3 : GGTAGATTGACTTTCAACAAGGTTATTAGACCATGTATGAAGAAG (45mer)

U4 : ACTATTTACGAAAACGAAAGAGAAATTAAGGGGTACGAATACCAA (45mer)

U5 : TTGTATGTTTACGCTTCTGACAAGCTTTTCAGAGCTGACATTTCT (45mer)

U6 : GAAGACTACAAGACCCGCGGTAGAAAGTTGTTGAGATTCAACGGT (45mer)

U7 : CCAGTTCCACCACCATAATAG (21mer)

L1 : CGATAATTTCCCATTCTCCCA (21mer)

L2 : CGTCAACAGCGAACTTACCCAAGTTTTGAGTGAATGGTCCAATAT (45mer)

L3 : CCTTGTTGAAAGTCAATCTACCATATTGACCAATCTTGTTTTCTT (45mer)

L4 : CTCTTTCGTTTTCGTAAATAGTCTTCTTCATACATGGTCTAATAA (45mer)

L5 : TGTCAGAAGCGTAAACATACAATTGGTATTCGTACCCCTTAATTT (45mer)

L6 : TACCGCGGGTCTTGTAGTCTTCAGAAATGTCAGCTCTGAAAAGCT (45mer)

L7 : TCGACTATTATGGTGGTGGAACTGGACCGTTGAATCTCAACAACTTTC (48mer)

The oligomers were isolated by urea-polyacrylamide gel electrophoresis and purified by passing through a

Seppak C18 column (Whatman).

The following is the amino acid sequence encoded by gene indicating the monellin subunits, the bridge, and the ligation strategy.

(Ligation strategy)

Fused Monellin

Met Gly Glu Trp Glu Ile Ile Asp Ile Gly

Pro Phe Thr Gln Asn Leu Gly Lys Phe Ala

Val Asp Glu Glu Asn Lys Ile Gly Gln Tyr

Gly Arg Leu Thr Phe Asn Lys Val Ile Arg

Subunit II

Pro Cys Met Lys Lys Thr Ile Tyr Glu Asn

Glu Arg Glu Ile Lys Gly Tyr Glu Tyr Gln

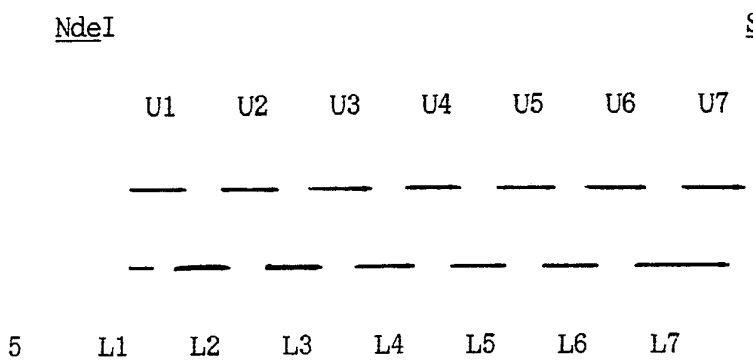
Subunit I

Leu Tyr Val Tyr Ala Ser Asp Lys Leu Phe

Arg Ala Asp Ile Ser Glu Asp Tyr Lys Thr

Arg Gly Arg Lys Leu Leu Arg Phe Asn Gly

Pro Val Pro Pro Pro



2. Annealing, Ligation of oligomers, and Isolation of Fused Monellin Gene

Each oligomer was phosphorylated at 37 °C for 45 min. in a reaction mixture of 30 μl containing 50 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP, and 5 units of T4 polynucleotide kinase. Each reaction mixture was pooled, extracted by phenol/chloroform, precipitated with ethanol, and dried under Speed-Vac. The dried pellet was dissolved in 50 μl distilled water and 7 μl ligation buffer (0.2 M Tris-HCl, pH 7.5, 0.1 M $MgCl_2$, 0.1 M DTT) added. The solution was placed in a 95 °C water-bath and cooled slowly to room-temperature overnight. To the mixture was added 7 μl of 10 mM ATP, 40 units of T4 DNA ligase (New England Biolab Inc.) and 2 μl of water.

The reaction mixture was kept at room temperature for

10 min., extracted by phenol/chloroform, precipitated, dried and redissolved in 85 μl water. The ligated oligomer mixture was treated with restriction endonuclease NdeI and SalI (New England Biolabs, Inc.). The 290 base pair fragment was isolated by electrophoresis with a 7% polyacrylamide gel, the band electroeluted and purified using the Elutip-D column (S & S Co.).

3. Molecular Cloning

M13mp19RF was used for cloning the fused synthetic monellin gene. First, M13mp19RF was cut with XbaI/SalI (New England Biolabs, Inc.). The large fragment was isolated and purified as described previously. A synthetic XbaI/NdeI adaptor was synthesized.

XbaI NdeI

5' - CTAGAAACTGCAATGTTGAATAAACGCTGATTTTCGATCA - 3' (40mer)

3' - TTTGACGTTACAACTTATTTGCGACTAAAAGCTAGTAT - 5' (38mer)

The adaptor was purified as described above. The NdeI/SalI digested, annealed fused synthetic monellin DNA fragment was combined with XbaI/SalI-treated M13mp19RF

and XbaI/NdeI adaptor in 10 μl of 20 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 10 mM DTT, 200 units T4 DNA ligase (N. E. Biolabs, Inc.) and incubated at 4°C overnight to provide M13mp19 MON-1RF. The transformation was done by adding 5 μl of the ligation mixture to 200 μl of E. coli JM101 competent cells (Messing, J. Methods in Enzymology (1983) 101 : 20-78). The dideoxy DNA sequencing and M13mp19 NON-1RF preparation were done as described in Messing, J. (1983) Methods in Enzymology ; and Sanger, T. et al. Proc. Natl. Acad. Sci. USA (1985) 74 : 5463-5467.

4. Construction of Expression Vector

Synthetic fused monellin DNA (293 bp) was isolated from M13mp19 MON-1RF and purified. The vector pDR720 containing trp O, P (Pharmacia, Inc.;Cat. # 27-4930-01) was digested with SmaI/PvuII and blunt-end ligated to produce ptrp322. The ptrp322 was digested with HpaI/SalI and a 2.5 kbp large fragment isolated. A synthetic HpaI/NdeI adaptor,

5' - AACTAGTACGCAAGTTCACGTAAAAAGGGTAATACA - 3' (36mer)

3' - TTGATCATGCGTTCAAGTGCATTTTTCCCATTATGTAT - 5' (38mer)

HpaI NdeI

was synthesized using Applied Biosystems DNA Synthesizer Model 380B. The ligation reaction of the 293 bp synthetic fused monellin, HpaI/SalI-treated ptrp322 vector and the HpaI/NdeI synthetic adaptor was carried out in the presence of 10 μl of 20 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 10 mM DTT, and 200 units of T4 DNA ligase (N.E. Biolabs, Inc.) at 14 °C overnight to give ptrp322H MON-1. The transformation of E.coli W3110 (ATCC 27325) with this plasmid and screening of recombinant clones was done as described above.

5. Identification of Expression of Synthetic Fused Monellin Gene

For a gene expression study, an overnight culture of 50 μl of ptrp322H MON-1 in W3110 with Luria Broth was inoculated into 5 ml of M9 media containing 0.4% casamino acid, 10μg/ml vitamin B_1, 40μg/ml ampicillin and cultured at 37 °C in a temperature controlled-

shaking incubator until OD_{650nm} reached about 0.5. Then 0.1 mg of indoleacrylic acid was added to the reaction mixture to a concentration of 50 μg/ml and the mixture incubated further for about 8 hrs. The cultured cells were pelleted at 2500 rpm for 5 min. in a Beckman J6 centrifuge. Laemmli protein sample buffer was added to the cell pellet, followed by heating at 95 °C for 5min. and the DNA loaded onto a 15% Laemmli SDS-polyacrylamide gel (Laemmli, Nature (1970) 227 : 680-685). The electrophoresis was run at 300 for 2.5 hours. The gel was stained with Coomassie blue brilliant dye demonstrating a product having the correct molecular weight. The expressed product was isolated and shown to have a sweet taste.

Following the above procedures, modified DNA sequences were prepared, where the amino acids at the juncture were varied. The following sequences indicate the sequence joining the isoleucine of subunit II(amino acid 46), (Bohak and Lee, supra, numbering) to the glycine (amino acid 6) of subunit I.

2. Y-E-N-R-E-D-I-K 8. Y-A-S-D-K-L

3. Y-K-T-R-E-D-I-K
4. Y-E-R-E-I-K
5. Y-E-N-I-K
6. Y-E-I-K
7. Y-Y-A-S-D-K-L-K
9. Y-A-S-D-K
10. Y-S-D-K
11. E-D-Y-K-T-R-G-R
12. E-D-Y-K-T-R

The codons employed were the same as the codons indicated for the MON-1 construct, with the exception of the sequence indicated as 3, where codons preferred by S. cerevisiae glycolytic enzymes were employed. That sequence is AAG, ACT, AGA.

It is evident from the above results, that novel proteinacious sweetners based on the monellin sequence may be produced as a stable single-chain protein for use in a wide variety of ways. The product can be produced efficiently and economically by employing microbial hosts, so that a stable uniform supply of the sweetener can be obtained, as distinct from isolation from natural sources. In addition, various changes may be made in the structure of the amino acid, without affecting its sweetening characteristic, while providing for other advantages, such as chemical and physical stabilty, storage life, ease

of formulation and purification, enhancement of sweetness, etc.

All publications and patent applications mentioned in this specification are indicative of the level of skill of those skilled in the art to which this invention pertains. All publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious that certain changes and modifications may be practiced within the scope of the appended claims.

WHAT IS CLAIMED IS ;

1. A DNA sequence encoding a polypeptide comprising an amino acid sequence having at least 80% of the amino acid sequence of the subunits of monellin covalently bonded together through a peptide linkage.

2. A DNA sequence according to Claim 1, wherein said polypeptide has a sequence of subunit II as the N-terminus and subunit I as the C-terminus.

3. A DNA sequence according to Claim 2, wherein in said polypeptide sequence the interventing sequence from tyr(41) to leu(3) of the monellin wild-type sequence is modified by substitution, deletion or insertion consisting of a total of not more than ten amino acids.

4. A DNA sequence according to Claim 3, wherein said intervening sequence is of the formula :

$$aa_x^1 - aa_x^2 - aa_x^3 - aa_x^4 - aa_x^5 - aa_x^6 - aa_x^7 - aa_x^8$$

wherein :

aa^1 is D or E ;

aa^2 is K, R, N or Q ;

aa^3 is D, E, S or T ;

aa^4 is F, W or Y ;

aa^5 is K or W ;

aa^6 is D or E ;

wherein x is 0 or 1, at least one x being 1.

5. A DNA sequence according to Claim 4, wherein said interventing sequence is Y-E-N-E-R-E-I-K ; Y-E-N-R-E-D-I-K ; Y-K-T-R-E-D-I-K ; Y-E-R-E-I-K ; Y-E-N-I-K ; Y-E-I-K ; Y-Y-A-S-D-K-L-K ; Y-A-S-D-K-L ; Y-A-S-D-K ; Y-S-D-K ; E-D-Y-K-T-R-G-R ; or E-D-Y-T-R.

6. An expression cassette comprising in the direction of transcription, a transcriptional and translational initiation regulatory region, a DNA sequence according to any of Claims 1 to 5, wherein said DNA sequence has methionine at its 5' terminus, and a translational and transcriptional termination region, wherein said regul-

atory regions are functional in a microbial host.

7. A mincrobial host comprising an expression cassette comprising in the direction of transcription, a transcriptional and translational initiation regulatory region, a DNA sequence according to any of Claims 1 to 5, wherein said DNA sequence has a methionine at its 5' terminus, and a translational and transcriptional termination region, wherein said regulatory regions are functional in a microbial host.

8. A method for producing a proteinaceous sweetener having at least about 80% homology with the two subunits of monellin, said sweetener consisting of a single chain, said method comprising

growing in a suitable nutrient a microbial host comprising an expression cassette comprising in the direction of transcription, a transcriptional and translational initiation regulatory region, a DNA sequence according to any of Claims 1 to 5, wherein said DNA sequence has a methionine at its 5' terminus,

and a translational and transcriptional termination region, wherein said regulatory regions are functional in a microbial host, whereby said DNA sequence is expressed to produce said proteinaceous sweetener ; and

isolating said proteinaceous sweetener.

INTERNATIONAL SEARCH REPORT

International Application No PCT/KR 88/00010

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC[4]: C 12 N 15/00, C 07 H 21/04, C 12 N 1/20, C 12 P 21/00, A 23 L 1/236.

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
Int.Cl.[4]:	C 12 N, C 07 H, C 12 P, A 23 L.

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category [10]	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
D,Y	FEBS LETTERS, Volume 29, no. 2, issued January 1973 (Amsterdam), H. VAN DER WEL et al. "Characterization of the Sweet--Tasting Protein from Dioscoreophyllum cumminisii (Stapf) Diels", see pages 181-184; especially page 181, introduction; page 183, tables 1,3.	(1)
D,Y	SCIENCE, Volume 181, no. 4094, issued July 6, 1973 (Washington DC.), R.H. CAGAN "Chemostimulatory Protein: A New Type of Taste Stimulus", see pages 32-35.	(1)
Y	EP, A2, 0 054 330 (UNILEVER NV), 23 June 1982 (23.06.82), see abstract; fig. 1.	(1)
Y	WO, A1, 87/03 007 (BEATRICE COMPANIES, ING.), 21 May 1987 (21.05.87), see claims 1-12.	(1)
D,A	BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, Volume 71, no. 1, issued 1976 (New York, London), G. HUDSON et al. "Mass Spectrometric Sequencing of Proteins: The Structure of Subunit I of Monellin", see pages 212-220.	(1,2)

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
03 August 1988 (03.08.88)	10 August 1988 (10.08.88)
International Searching Authority	**Signature of Authorized Officer**
AUSTRIAN PATENT OFFICE	

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No
D,A	US, A, 3 998 798 (CAGAN et al.), 21 December 1976 (21.12.76), see abstract; claim.	(1,2)

Anhang zum internationalen Recherchenbericht über die internationale Patentanmeldung Nr.

Annex to the International Search Report on International Patent Application No. PCT/KR 88/00010

Annexe au rapport de recherche internationale relatif à la demande de brevet international n°.

In diesem Anhang sind die Mitglieder der Patentfamilien der im obengenannten internationalen Recherchenbericht angeführten Patentdokumente angegeben. Diese Angaben dienen nur zur Unterrichtung und erfolgen ohne Gewähr.

This Annex lists the patent family members relating to the patent documents cited in the above-mentioned International search report. The Austrian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

La présente annexe indique les membres de la famille de brevets relatifs aux documents de brevets cités dans le rapport de recherche internationale visé ci-dessus. Les renseignements fournis sont donnés à titre indicatif et n'engagent pas la responsabilité de l'Office autrichien des brevets.

Im Recherchenbericht angeführtes Patentdokument Patent document cited in search report Document de brevet cité dans le rapport de recherche	Datum der Veröffentlichung Publication date Date de publication	Mitglied(er) der Patentfamilie Patent family member(s) Membre(s) de la famille de brevets	Datum der Veröffentlichung Publication date Date de publication
EP-A2-0 054 330	23/06/82	AT-E - 22 327	15/10/86
		EP-A3-0 054 330	22/09/82
		EP-B1-0 054 330	17/09/86
		BR-A -8 108 074	21/09/82
		CA-A1-1 192 151	20/08/85
		IE-B - 52 417	28/10/87
		JP-A2-58-004799	11/01/83
		JP-B4-61-005392	18/02/86
		JP-A2-61-043996	03/03/86
WO-A1-87/03 007	21/05/87	AU-A1-65 433/86	02/06/87
		JP-T2-63-501264	19/05/88
		EP-A1-0 255 823	17/02/88
US-A -3 998 798	21/12/76	None	