

①9 RÉPUBLIQUE FRANÇAISE
INSTITUT NATIONAL
DE LA PROPRIÉTÉ INDUSTRIELLE
PARIS

①1 N° de publication :
(à n'utiliser que pour les
commandes de reproduction)

2 575 162

②1 N° d'enregistrement national :

85 18965

⑤1 Int Cl⁴ : C 07 H 21/04.

⑫

DEMANDE DE BREVET D'INVENTION

A1

②2 Date de dépôt : 20 décembre 1985.

③0 Priorité : JP, 26 décembre 1984, n° 59-281645.

④3 Date de la mise à disposition du public de la
demande : BOPI « Brevets » n° 26 du 27 juin 1986.

⑥0 Références à d'autres documents nationaux appa-
rentés :

⑦1 Demandeur(s) : *NIPPON SHINYAKU CO., LTD.* — JP.

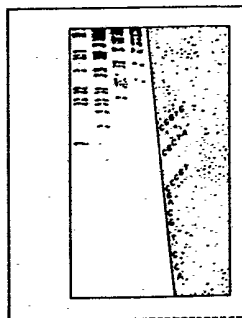
⑦2 Inventeur(s) : Junichi Yano et Tadaaki Ogi.

⑦3 Titulaire(s) :

⑦4 Mandataire(s) : Novapat, Cabinet Chereau.

⑤4 Procédé de synthèse d'un DNA à longue chaîne.

⑤7 Procédé de synthèse d'un DNA à longue chaîne, caracté-
risé en ce que des blocs ayant 4 à 8 séquences de bases sont
liés par voie purement chimique par un procédé « en phase
solide » (procédé au triester) en utilisant un CPG aminé comme
support.



FR 2 575 162 - A1

D

1.

La présente invention concerne un nouveau procédé de synthèse d'un DNA à longue chaîne portant de l'information pour la synthèse de protéines déterminées, et plus particulièrement un procédé de synthèse d'un DNA à longue chaîne
5 par voie purement chimique, c'est-à-dire sans utiliser d'enzymes, par un procédé dit "en phase solide", en utilisant un bloc de 4 à 8 séquences de base et un CPG aminé comme support.

Il est connu que la synthèse de polypeptides par
10 la technologie des gènes en utilisant un gène synthétique est possible par les stades de (1) synthèse d'un gène structural; (2) recombinaison du gène dans un plasmide approprié; (3) transformation d'un hôte approprié par le plasmide chimère formé; et (4) préparation du polypeptide désiré par
15 culture de la substance transformée.

Depuis une époque récente, la mise au point d'une sonde à DNA attire l'attention du public en tant que moyen nouveau de la technologie des gènes. Il s'agit d'un procédé d'identification d'un DNA et d'un RNA inconnus qui
20 est un produit transcrit par une hybridation d'un DNA et

2.

d'un RNA à toron unique, lesquels sont connus dans la technique, en utilisant les propriétés du DNA et du RNA de former un duplex en choisissant une substance complémentaire, relation qui est exactement celle du gabarit de moulage et de l'objet moulé. Comme une identification très sensible et très rapide est possible en utilisant le procédé d'hybridation, ce procédé peut être appliqué au diagnostic précis d'une maladie en trouvant un DNA et un RNA particuliers au niveau des gènes dans le sang et les cellules des malades et dans les bactéries pathogènes. En conséquence, le DNA est important comme agent de diagnostic étant donné son utilisation comme sonde à DNA.

En ce qui concerne le DNA indiqué ci-dessus comme gène structural et son utilisation comme sonde à DNA, il est connu, pour ce qui est de sa nature, que plus la séquence de bases est longue, plus la source d'information est importante et plus le domaine d'utilisation comme sonde à DNA est vaste. Cependant, il est également connu que plus la séquence des bases est longue, et plus sa synthèse est difficile.

En conséquence, il était souhaitable de disposer d'une technique pour préparer par synthèse d'une manière aisée un DNA à longue chaîne.

Le procédé classique de synthèse du DNA est le suivant. On prépare d'abord par synthèse chimique des fragments de DNA relativement courts comprenant 10 à 20 radicaux basiques, puis on les combine pour préparer des fragments ayant une structure totale de DNA à double toron présentant l'information sur la synthèse peptidique désirée, et enfin on les combine en utilisant une enzyme appelée DNA ligase.

Par ce procédé, cependant, on ne fabrique que des fragments relativement courts comportant 1 base (monomère), 2 bases (dimère) ou 3 bases (trimère) avant la condensation en blocs, et il n'est pas possible de préparer par synthèse un DNA long avec par exemple 80 radicaux.

3.

En outre, dans ce procédé, il est essentiel d'utiliser une enzyme appelée DNA ligase. Par conséquent, lorsqu'on prépare par synthèse un DNA à double toron comme gène, il est nécessaire que toutes les séquences de bases constituant le DNA à double toron soient préparées par synthèse en même temps. En conséquence, le procédé ci-dessus n'est pas aussi efficace dans un procédé de synthèse d'un DNA à double toron.

La demanderesse a poursuivi des études pour surmonter la difficulté technique ci-dessus, et elle a réussi à préparer par synthèse un DNA comportant par exemple 46 bases, en utilisant un procédé dit "procédé au triester" (parmi les procédés dits "solides"), dans lequel on utilise comme support du polystyrène à 1 %, et les composés de 4 bases (tétramère) ou 5 bases (pentamère) sont soumis à une condensation répétée.

Même par ce procédé, cependant, le nombre de bases dans le DNA obtenu est de 50 au maximum, et il reste difficile de préparer par synthèse un DNA ayant des chaînes comportant jusqu'à 80 à 150 radicaux.

Compte-tenu de ce qui précède, la demanderesse a poursuivi des études supplémentaires en portant son attention (1) sur la synthèse d'un DNA à longue chaîne portant autant d'information de gène que possible, et (2) sur la synthèse dans des conditions encore plus avantageuses, et elle a finalement réalisé la présente invention.

Les caractéristiques de la présente invention sont les suivantes :

(1) les nombres d'unité avant la condensation en blocs sont de 4 à 8 bases (octamère);

(2) parmi les procédés dits "en phase solide", on utilise le "procédé au triester"; et

(3) on utilise du CPG aminé comme support.

La présente invention sera encore illustrée par ce qui suit :

Avant la condensation, chaque bloc peut être

4.

obtenu de la manière classique, dans laquelle chaque base est soumise à une synthèse en phase liquide.

Le CPG aminé (verre à pores de taille réglée) (voir Tetrahedron, 24, 747-750, 1983) utilisé dans la présente invention est utilisé comme support dans le procédé en phase solide. Au groupe amino de cette substance, on combine de la désoxythymidine qui est transformée en 3'-succinate par le procédé habituel. Celui-ci est utilisé comme support pour le nucléoside. Chacun des blocs désirés est prolongé successivement, sur cette résine, en direction de l'élément 5'-terminal. Comme agent de condensation, on peut utiliser par exemple du mésitylène sulfonyl-3-nitrotriazolide (MSNT). Le DNA ainsi obtenu est à un seul toron, et le DNA à toron complémentaire qui est nécessaire pour la préparation du DNA duplex peut être aisément obtenu d'une manière similaire. On peut aussi obtenir très aisément ce DNA duplex au moyen de DNA polymérase, en utilisant un fragment court (10 paires de bases, par exemple), qui est complémentaire de la région 3'-terminale du DNA à toron unique obtenu. Le fait que la DNA polymérase puisse être utilisée est un des mérites les plus importants de la présente invention, puisque la réaction de condensation peut être accomplie sans l'aide de DNA ligase qui a été largement utilisée dans les procédés classiques.

Le DNA duplex obtenu est combiné pour donner un plasmide vecteur par le procédé connu, puis transformé dans des bactéries telles qu'Escherichia coli, et la souche est cultivée pour donner le polypeptide désiré. Dans les stades ci-dessus, on peut appliquer divers moyens de la technologie des gènes qui ont déjà fait leurs preuves.

Il est possible, conformément à la présente invention, de préparer par synthèse un DNA ayant jusqu'à 80 à 150 radicaux, et on peut par conséquent préparer par synthèse des polypeptides ayant 15 à 30 aminoacides par les moyens connus de la technologie des gènes. Par exemple, on peut

5.

préparer par synthèse les polypeptides suivants : le facteur d'inhibition de la libération des hormones (la somatostatine contenant 14 aminoacides), le stimulant de la sécrétion gastrique (gastrine, contenant 17 aminoacides), le remède
5 à l'ulcère duodénal (secrétine, contenant 27 aminoacides), le stimulant de la sécrétion de l'hormone de croissance, l'agent pour augmenter le taux d'insuline et de sucre dans le sang (glucagon, contenant 29 aminoacides), un agent analogue à la morphine (béta-endorphine, contenant 31 amino-
10 acides) et le remède contre l'hypercalcémie (calcitonine, contenant 32 aminoacides), etc.

En outre, le DNA à longue chaîne de la présente invention s'applique non seulement à des séquences de bases de DNA de parties structurales de gènes, mais aussi à la
15 préparation d'un DNA général comprenant des sites régulateurs et des séquences déterminées, ainsi qu'à une sonde de DNA à longue chaîne reconnaissant leurs structures. En conséquence, la présente invention peut être utilisée avec succès pour la mise au point d'agents de diagnostic.

20 Conformément à la présente invention, un DNA à longue chaîne peut être préparé par synthèse d'une manière simple et en grandes quantités. Le DNA à longue chaîne de la présente invention peut être utilisé efficacement (1) comme source d'information de gène concernant la synthèse
25 de polypeptides et (2) comme source pour l'application d'une sonde à DNA dans la technologie des gènes.

La production de DNA a été de 0,1 OD (1 OD équivalent à environ 50 microgrammes) pour un lot, au maximum. Cependant, conformément à la présente invention, il est
30 à présent possible de fabriquer dans des quantités allant jusqu'à 30 à 50 OD par lot. En conséquence, on peut s'attendre à une expansion du domaine d'utilisation du DNA à longue chaîne comme gène et comme sonde à DNA.

EXEMPLES

35 La présente invention sera encore illustrée par des exemples concernant la synthèse de l'endorphine dont les

6.

activités, physiologiques telles que l'action analgésique sur le système nerveux central et l'action d'hormone endocrine, sont connues.

- 5 (1) Synthèse de chacun des blocs constituant des séquences de base contenant le gène endorphine.

La séquence d'acides aminés des endorphines est connue, et la séquence de bases du DNA lui correspondant peut être librement choisie en se référant à un tableau de l'utilisation des codons. On les trouvera ci-dessous en même temps
10 que leur relation avec chacun des blocs constituant les séquences de bases de DNA utilisées dans la présente invention. Les colonnes supérieure, moyenne et inférieure correspondent respectivement à chaque bloc (les chiffres qui s'y trouvent sont les numéros des blocs), à la séquence des bases et à la
15 séquence d'acides aminés correspondante. Incidemment, les sites d'enzymes de restriction sont donnés aux deux extrémités des séquences de bases du DNA. Ces sites sont utilisés pour insérer le plasmide.

7.

① α - Endorphine

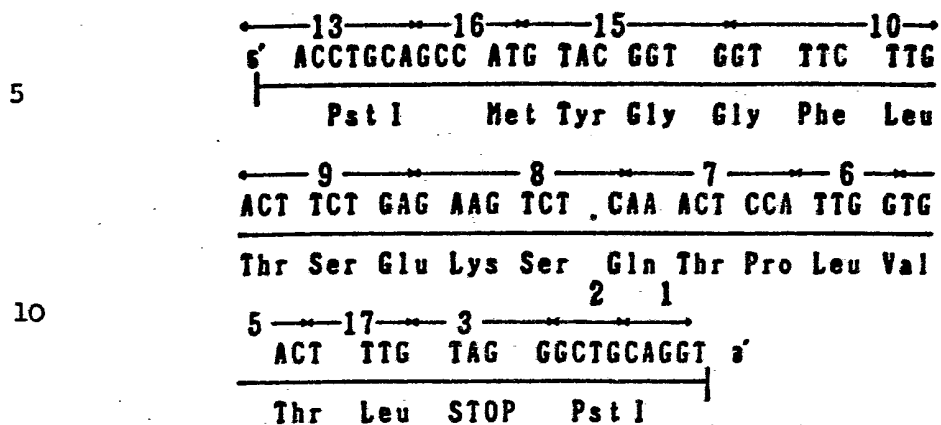
5
 13 12 11 10
 5' ACCTGCAGCC CGT CGC TAC GGT GGT TTC ATG
 Pst I Arg Arg Tyr Gly Gly Phe Met
 9 8 7 6
 ACT TCT GAG AAG TCT CAA ACT CCA TTG GTG
 10
 Thr Ser Glu Lys Ser Gln Thr Pro Leu Val
 5 4 3 2 1
 ACT TAA TAG GGCTGCAGGT 3'
 Thr STOP STOP Pst I

15

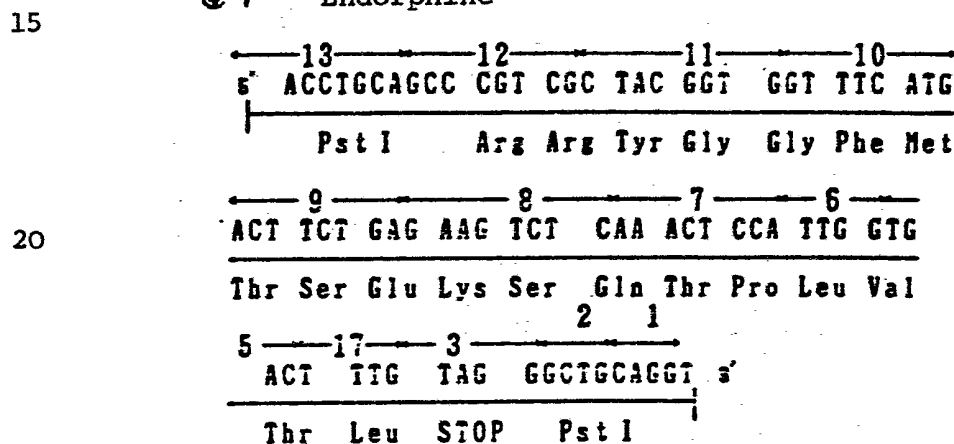
② α - (Leu⁵) - Endorphine

13 16 15 10
 5' ACCTGCAGCC ATG TAC GGT GGT TTC TTG
 Pst I Met Tyr Gly Gly Phe Leu
 20
 9 8 7 6
 ACT TCT GAG AAG TCT CAA ACT CCA TTG GTG
 Thr Ser Glu Lys Ser Gln Thr Pro Leu Val
 5 4 3 2 1
 ACT TAA TAG GGCTGCAGGT 3'
 25
 Thr STOP STOP Pst I

8.

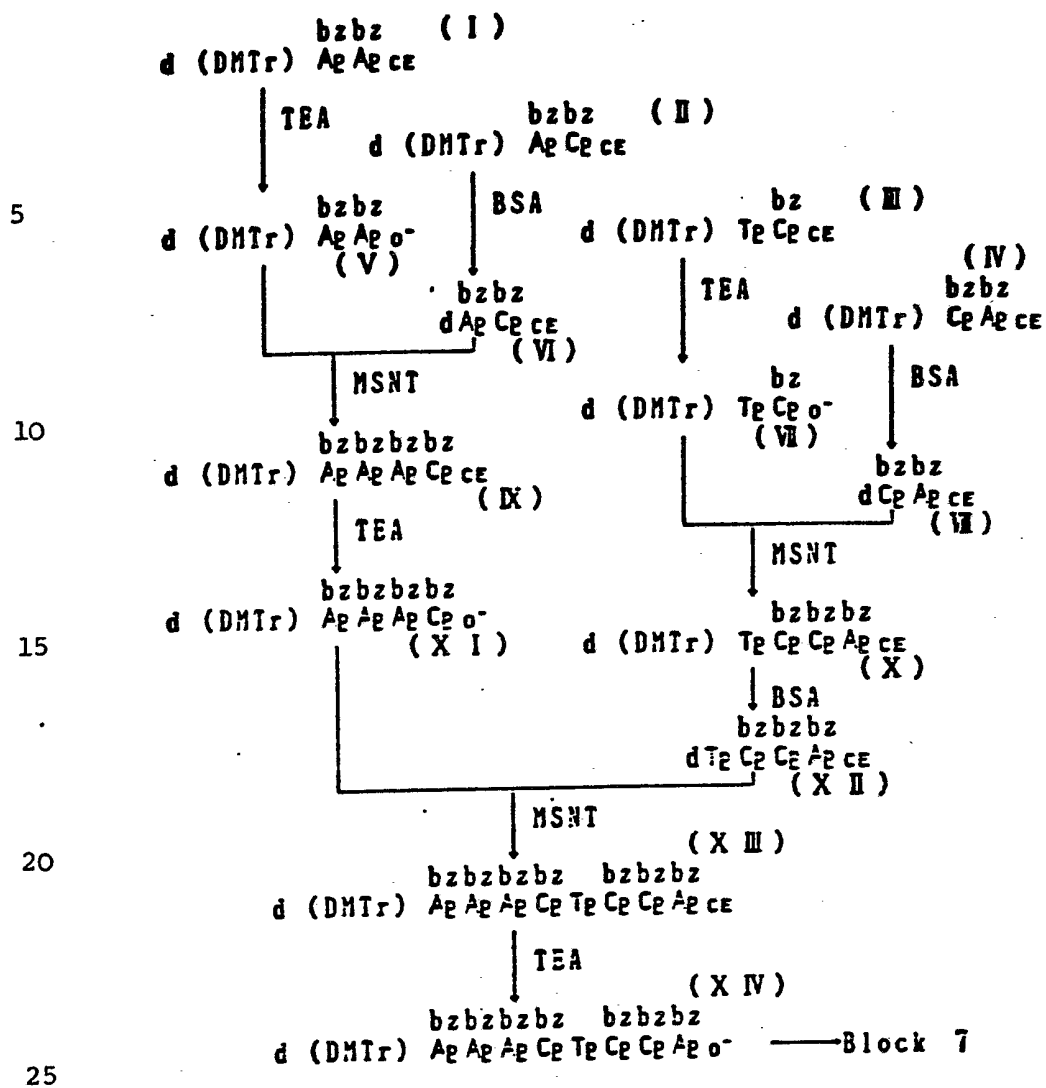
③ 7 - (Leu⁵) - Endorphine

④ 7 - Endorphine



Parmi les blocs constituant les gènes endorphine ci-dessus, le bloc 7 a été préparé par synthèse en passant par les stades indiqués ci-dessous.

9.



DMTr: 4,4'-diméthoxytolityle
 bz: N-benzoyladénosyle
 A: N-benzoylcytidyllyle
 C: Tymidyllyle
 T: Phosphate d'o-chlorophényle
 BSA: Acide benzènesulfonique
 TEA: Triéthylamine

10.

D'autres blocs (1-6, et 8-17) constituant des gènes du type endorphine peuvent être préparés par synthèse d'une manière similaire. Les rendements sont donnés ci-dessous.

5

	Blocs	Séquences de bases	Rendement
10	1	CACG	77
	2	GCTG	94
	3	TAGG	70
	4	TTAA	104
	5	TGAC	93
15	6	TTGG	67
	7	AAACTCCA	70
	8	GAAGTCTC	62
	9	ACTTCTGA	65
	10	GTTTCATG	65
20	11	CTACGGTG	70
	12	GCCCCGCA	65
	13	ACCTGCA	65
	14	GTTTCTTG	70
	15	GTACGGTG	67
25	16	GCCAT	75
	17	TTTG	95

(2) Synthèse des gènes endorphines :

Le gène alpha-endorphine (désoxy 80 mère) contenant des sites d'enzyme de restriction a été préparé par synthèse par un procédé en phase solide de la manière suivante :

1. Une résine désoxytimidine CPG est lavée avec $\text{CH}_2\text{Cl}_2/\text{MeOH}$.

2. La détritylation est effectuée avec de la BSA à 2 %/ CH_2Cl_2 (ceci est effectué à plusieurs reprises et

35

11.

rapidement jusqu'à ce que la coloration disparaisse).

3. On effectue un séchage azéotropique après substitution par de la pyridine.

5 On ajoute une solution de chacun des blocs, on la soumet à un séchage azéotropique et on ajoute du MSNT et de la pyridine pour la réaction. On laisse reposer à la température ambiante et on lave avec de la pyridine.

10 4. On ajoute une solution de diméthylaminopyridine 0,1M/pyridine et de l'anhydride acétique, on laisse reposer à la température ambiante et on lave à la pyridine.

On effectue plusieurs fois l'opération ci-dessus, 13 fois au total. Le rendement moyen de la réaction est de 84 %. Puis, on déprotège la résine, à la température ambiante, avec une solution de tétraméthylguanidine 0,1M-pyridine
15 aldoxime (voir C. B. Reese, et coll.: Tetrahedron Lett., 2727, 1978) dans du dioxane-eau, puis on la lave avec de la pyridine-eau, on concentre le liquide de lavage sous vide, on y ajoute de l'ammoniaque concentrée et on chauffe le mélange. On en fait évaporer l'ammoniac et une partie du résidu est prise en utilisant le groupe diméthoxytrityle comme
20 cible pour calculer le rendement du stade final.

On soumet la solution réactionnelle résiduelle à des chromatographies ouvertes en phase inversée (gel de silice C₁₈ pour Prep 500 fabriqué par Waters), par échange d'ions
25 (DEAE-toyopal) et en phase inversée (gel de silice C₁₈, Gel TSK 10-20 microns), ce qui donne le gène alpha-endorphine pur (contenant des sites d'enzyme de restriction) (= désoxy 80 mère).

La pureté est confirmée par HPLC (Nucléosil
30 300-7 C₁₈) et par électrophorèse, et ses séquences de base sont confirmées par la méthode de Maxam-Gilbert. Le résultat est donné dans les figures 1 à 3.

On prépare de la même manière le gène alpha-
(Leu⁵)-endorphine contenant un site d'enzyme de restriction)
35 (désoxy 77 mère), le gène gamma-(Leu⁵)-endorphine (contenant

12.

un site d'enzyme de restriction) (désoxy 77 mère) et le gène gamma-endorphine (contenant un site d'enzyme de restriction) (désoxy 80 mère).

(3) Synthèse du DNA duplet et sa combinaison avec
5 le plasmide vecteur.

On mélange chaque molécule de désoxy 80 mère et d'amorce de nucléotide synthétique qui est complémentaire de l'élément 3'-terminal du premier, on les chauffe à 65°C et on les refroidit à la température ambiante pour cuire
10 le désoxy 80 mère et l'amorce. Puis on ajoute de la E. coli polymérase I (fragment de Klenow) par un moyen classique et on la fait réagir à 37°C pendant 30 minutes, de façon à transformer le DNA en sa forme à double toron.

On récupère le DNA sous forme de précipité dans
15 l'éthanol, on fait réagir à 37°C pendant 30 minutes en utilisant de la polynucléotidokinase T_4 et on phosphoryle les deux éléments 5'-terminaux du DNA à double toron.

On coupe ensuite le plasmide vecteur pUC 8 DNA avec une enzyme de restriction Pst 1, on l'ajoute à la solution de DNA à double toron ci-dessus, on le fait réagir à 16°C
20 pendant une nuit avec de la T_4 DNA ligase, et on combine le DNA désoxy-80 mère à double toron avec le plasmide vecteur.

(4) Clonage du gène endorphine contenant le plasmide.
de.

On transforme le plasmide préparé comme ci-dessus
25 dans une souche de E. coli JM 103 par des moyens classiques, puis on le choisit en utilisant une déficience d'activité de bêta-galactosidase présente dans le pUC 8 comme cible, et on recueille les molécules de plasmide par clonage à partir de la souche.
30

Il est confirmé que le plasmide dans lequel le gène endorphine a été inséré dans l'orientation et la position correctes est tel que désiré conformément à la méthode de Maxam-Gilbert.

(5) Préparation d'endorphines
35 Une souche d'E. coli JM 103 transformée est

13.

pré-cultivée une nuit dans un milieu LB, elle est transplantée dans un milieu 2YT et soumise à une culture agitée à 37°C.

De l'IPTG est ajouté aux stades de la phase productive logarithmique (stades initial, moyen et final) pour le rendre 0,5 mM, et on amorce la synthèse de l'endorphine. Après avoir été induite par l'IPTG, la protéine condensée est extraite et analysée par HPLC, et l'on trouve qu'il s'est formé une quantité adéquate de protéine ($1-5,0 \times 10^5$ molécules par cellule) lorsque l'induction est appliquée au stade initial de la phase productive logarithmique.

En ce qui concerne l'endorphine alpha du type naturel et l'endorphine gamma ayant un radical méthionine dans une molécule, on les traite par la trypsine de la manière classique. En ce qui concerne l'alpha-(Leu⁵)-endorphine et la gamma-(Leu⁵)-endorphine ayant un radical leucine à la place de la méthionine, on la traite par BrCN. En tout cas, chacune des protéines endorphines désirée est soumise à une chromatographie sur colonne conformément au procédé général de purification des protéines, après quoi chacune d'elles est séparée et purifiée.

Le fait que chacune des molécules d'endorphine obtenues présente la séquence d'acides aminés désirée est confirmé par le fait qu'elles sont identiques à des échantillons déjà obtenus par synthèse peptidique dans l'essai par HPLC en utilisant un support de phase inversée.

La figure 1 est un autoradiogramme aux rayons X montrant le résultat d'une électrophorèse sur polyacrylamide à 20 % de désoxy 80 mère contenant le gène alpha-endorphine préparé par synthèse après détermination par la méthode Maxam-Gilbert.

La figure 2 est un autoradiogramme aux rayons X montrant le résultat d'une électrophorèse sur polyacrylamide à 8 % du désoxy 80 mère contenant le gène alpha-endorphine préparé par synthèse après détermination par la méthode Maxam-Gilbert.

14.

La figure 3 montre le résultat d'une chromatographie liquide à hautes performances (Nucléosil 300-7 C₁₈) du désoxy 80 mère contenant le gène alpha-endorphine préparé par synthèse après détermination par la méthode Maxam-Gilbert. En ordonnées et en abscisses, on a respectivement l'absorbance et le temps. Le système de solvants utilisé est l'acétate de triéthylamine-acétonitrile et le débit est de 1,0 ml/minute.

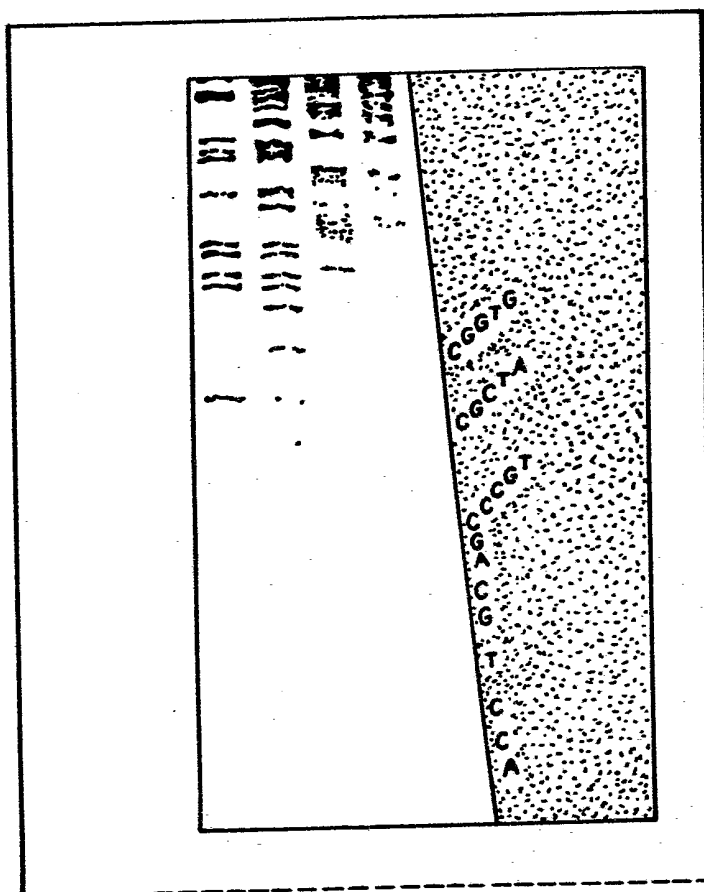
La présente invention n'est pas limitée aux exemples de réalisation qui viennent d'être décrits, elle est au contraire susceptible de modifications et de variantes qui apparaîtront à l'homme de l'art.

15.

REVENDEICATIONS

- 1 - Procédé de synthèse d'un DNA à longue chaîne, caractérisé en ce que des blocs ayant 4 à 8 séquences de base sont liés par voie purement chimique par un procédé
- 5 "en phase solide" (procédé au triester) en utilisant un CPG aminé comme support.

1/3

Fig.1

3/3

Fig. 3