

①⑨ RÉPUBLIQUE FRANÇAISE
INSTITUT NATIONAL
DE LA PROPRIÉTÉ INDUSTRIELLE
PARIS

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

2 632 321

②① N° d'enregistrement national :

89 03952

⑤① Int Cl⁴ : C 12 N 15/00; C 07 K 15/00; C 07 H 21/00.

①②

DEMANDE DE BREVET D'INVENTION

A1

②② Date de dépôt : 24 mars 1989.

③⑦ Priorité : US, 24 mars 1988, n° 172653.

④③ Date de la mise à disposition du public de la
demande : BOPI « Brevets » n° 49 du 8 décembre 1989.

⑥⑦ Références à d'autres documents nationaux appa-
rentés :

⑦① Demandeur(s) : Société dite : ONCOGEN. — US.

⑦② Inventeur(s) : Joseph P. Brown ; Daniel R. Twardzik ;
Hans Marquardt ; George J. Todaro ; A. Gregory Bruce ;
Timothy M. Rose ; Anthony F. Purchio ; Shiu-Lok Hu ;
Wen Chang ; Suo Win Liu ; Thomas J. Franceschini ;
Mohammed Shoyab ; Greg Plowman.

⑦③ Titulaire(s) :

⑦④ Mandataire(s) : Cabinet Weinstein.

⑤④ Séquence d'ADN codant pour un polypeptide, cassette d'expression comprenant une telle séquence, cellule hôte
comprenant la cassette d'expression et procédé pour produire un tel polypeptide.

⑤⑦ La présente invention se rapporte à une séquence d'ADN
codant pour un polypeptide, à une cassette d'expression com-
prenant une telle séquence, à une cellule hôte comprenant la
cassette d'expression et à un procédé pour produire un tel
polypeptide.

Selon l'invention, la séquence d'ADN code pour un polypep-
tide qui est capable de se lier à un récepteur de EGF, qui a
sensiblement la même structure en boucle qu'un facteur de
croissance produit naturellement et qui est capable de se lier à
un récepteur de EGF et qui a au moins une homologie de
30 % avec un ligand naturel, ledit polypeptide comprenant
trois domaines, chacun de ceux-ci ayant une séquence d'acides
aminés avec au moins une homologie de 30 % avec les
fragments des ligands naturels identiques ou différents et
ayant la même position relative que lesdits fragments desdits
ligands naturels, lesdits domaines étant définis spécifiquement.

Les composés de la présente invention trouvent une utilisation
comme mitogènes, additifs dans des milieux nutritifs, dans
des essais pour des récepteurs de EGF, et comme des agents
de compétition pour d'autres facteurs de croissance. Les com-
posés trouvent également une utilisation dans la cicatrisation
des plaies.

FR 2 632 321 - A1

D

La présente invention se rapporte à de nouvelles compositions de polypeptides ayant une activité de facteur de croissance, comprenant des polypeptides chimériques et à des procédés pour leur préparation
5 utilisant les techniques de l'ADN recombinant.

Un nombre significatif de polypeptides secrétés par des cellules de mammifères sont trouvés comme ayant une activité de facteur de croissance. La séquence de ces polypeptides est sensiblement conservée sur un éventail
10 important de mammifères. L'intérêt important de ces composés est aiguë par leur association avec l'oncogénicité. Il est également intéressant de savoir comment la production des facteurs de croissance est régulée et comment ils régulent à leur tour l'activité
15 cellulaire.

Il faut également noter que l'infection des cellules de mammifères par des virus entraîne une prolifération de la croissance des cellules infectées.

Dans le but de cette invention, il y a des
20 membres particulièrement intéressants de la famille des virus de la variole, tels que le virus de la variole, le virus de la vaccine et les virus associés à des maladies particulières tels que le virus du fibrome de Shope, le virus de la tumeur Yaba, et le Molluscum contagiosum
25 virus (MCV).

Il serait intéressant de déterminer si des virus provoquant une prolifération cellulaire par infection, produisent des polypeptides impliqués avec des facteurs de croissance, soit en agissant comme facteur de
30 croissance ou comme récepteur du facteur de croissance. Ces composés pourraient ensuite être utilisés dans les études des actions virales, dans des essais pour la présence du virus, dans des milieux nutritifs comme mitogènes, et dans le développement d'agents
35 thérapeutiques pour traiter l'infection virale. Il est également intéressant de développer des composés qui

puissent être agonistes ou antagonistes des facteurs de croissance pour une utilisation in vitro dans les cultures de cellules en croissance, dans la recherche de processus mitotiques, et en thérapie.

- 5 Venkatesan et al., J. Virol. (1982) 44:637-646 décrit la séquence d'ADN du gène structuré codant des protéines du virus de la vaccine mais il n'était pas admis qu'une quelconque de ces protéines avait une activité de facteur de croissance. Cooper et al., ibid
10 (1981) 37:284-294, rapporte la traduction d'ARNm codé dans la répétition terminale inversée du génome du virus de la vaccine. Des maladies prolifératives pour des membres de la famille du virus de la variole ont été rapportées pour le virus du fibrome de Shope (Shope, J. Exp. Med. (1932) 56:793-822 ; le virus de la tumeur de Yaba (Niven et al., J. Path. Bacteriol. (1961) 81:1-14) et le Molluscum contagiosum virus (MCV) (Postlethwaite, Arch. Environ. Health (1970) 21:432-452). Les
15 descriptions de facteur de croissance de l'épiderme (EGF) peuvent être trouvées dans Scott et al., Science (1983) 221:236-240 et Gray et al., Nature (1983) 303:722-725. La présence de trois ponts disulfures dans EGF et dans le facteur de croissance transformant (TGF) est rapporté par Savage et al., J. Biol. Chem. (1973) 248: 7669-7692. Voir
20 également Doolittle et al., Nature (1984) 307:558-560 et la demande de brevet australien No. 85006288. La région de liaison au récepteur de la molécule EGF a été suggérée comme s'étendant dans la boucle entre les troisième et quatrième résidus de cystéine (Komoriya et al., Proc. Natl. Acad. Sci. USA (1983) 81:1351-1355).

- De nouvelles descriptions du facteur de croissance du virus de la vaccine (VGF) peuvent être trouvées dans Brown et al., Nature (1985) 313:491-492 ; Reisner, Nature (1985) 313:801-803 et Blomquist et al.,
25 Proc. Natl. Acad. Sci. USA (1984) 81:7263-7367. Le facteur de croissance du virus de Shope Naturel (SFGE)

est décrit par Chang, W.C., et al., dans Molecular and Cellular Biol. (1987) 7 : 535-540. Shope, J. Exp. Med. (1932) 56:793-802 rapporte qu'une infection de lapins adultes par le virus du fibrome de Shope provoque une prolifération rapide des fibroblastes conduisant à un fibrome bénin. Une infection de lapins adultes immunosuppressés ou de nouveaux nés provoque une prolifération rapide des fibroblastes conduisant à un fibrosarcome atypique envahissant. Allison, A.C., J. Natl. Cancer Inst. (1966) 36 : 869-876 ; Smith, et al., J. Natl. Cancer Inst. (1973) 50:1529-1539 ; Allison, A.C. J. Natl. Cancer Inst. (1966) 35:859-868.

Le facteur de croissance du virus du Myxome Naturel (MGF) est décrit par Upton et al., J. Virol. (1987) 61 : 1271-1275. Une infection de lapins adultes par le virus du Myxome provoque une tumeur de myxosarcome envahissant proliférant rapidement. Strayer, D.S. et Sell, S., J. Natl. Cancer Inst. (1983) 71:105-116. Une séquence d'ADN capable de coder un facteur de croissance conservé (facteur de croissance de Molluscum contagiosum virus Naturel (MCGF)) a été développée. Porter, C.D. et Archard, L.C., J. Gen. Virol. (1987) 68 : 673-682. Les infections des humains par Molluscum contagiosum virus induit des tumeurs de la peau bénignes caractérisées par des cellules proliférant rapidement. Brown, et al Sex. Transm. Dis. (1981) 8 : 227-234. Le virus de la tumeur de Yaba provoque des histocytomes subcutanés chez les singes et les humains, Bearcroft, et al., Nature (London) (1958) 182 : 195-196, et à cause de sa relation avec le virus du fibrome du Shope, le virus du Myxome et le Molluscum contagiosum virus sont supposés comme ayant un facteur de croissance apparenté.

Un EGF de souris (mEGF) (Scott et al., 1983, supra Gray et al., 1983. supra), et un EGF d'humain (hEGF) (Gregory et al., Int. J. Pept. Protein Res. (1977) 9:107-118), sont connus pour être homologues aux TGF

d'humains, de souris et de rats) (Marquardt et al., Proc. Natl. Acad. Sci. USA (1983) 80:4684-4688) et aux résidus 45 à 85 d'un polypeptide de 140 résidus, le facteur de croissance de la vaccine (VGF) codé par le génome du virus de la vaccine (Venkatesan et al., 1982, supra).

L'expression de peptides étrangers employant un vecteur de baculovirus dans des cellules d'insectes est décrit par Maeda et al., Nature (1985) 315:592-594, et Carbonell et al., J. of Virology (1985) 56 : 153-160.

Les descriptions de ces références sont incorporées ici à titre de référence.

De nouvelles compositions de polypeptides sont fournies trouvant une analogie dans des fragments de protéines virales, lesquelles compositions agissent comme mitogènes et peuvent être utilisées dans des milieux nutritifs, comme réactifs pour la détection des récepteurs du facteur de croissance ou pour la présence de facteurs de croissance, et comme agents de compétitions pour un facteur de croissance transformant et un facteur de croissance de l'épiderme. Les compositions trouvent une utilisation thérapeutique, par exemple pour favoriser une épithélisation et une cicatrisation de brûlures et blessures. Les nouvelles compositions peuvent être synthétisées. Les peptides du sujet peuvent être formés comme des oligopeptides ou des protéines fusionnées en employant des techniques de recombinaison dans une grande variété d'hôtes.

D'autres buts, détails et avantages de la présente invention apparaîtront plus clairement à la lecture de la description explicative qui va suivre faite aux références aux dessins dans lesquels :

- la figure 1 est une comparaison de la séquence d'acides aminés de la protéine du virus de la vaccine avec des facteurs de croissance connus où acide

aminé initial du VGF est l'acide aspartique (D) à la position 20 à partir de la méthionine, laquelle position correspond à aa⁻²⁴ les peptides révélés ; et

5 - la figure 2 montre la structure proposée de VGF nature, avec des remplacements (N) d'asparagine. Les sites de glycosylation potentiels dans MGF et SFGF sont indiqués par des astérisques.

Les nouvelles compositions qui peuvent agir comme agonistes ou antagonistes du facteur de croissance de l'épiderme (EGF) ou du facteur de croissance transformant (TGF) sont fournies et permettent une grande variété d'applications dans la culture de cellules, les diagnostics, en thérapie in vivo, et en combinaison avec d'autres peptides dans la formation de polypeptides
15 hybrides comme immunogènes pour la production d'anticorps. Des séquences de polynucléotides peuvent être isolées ou préparées, qui peuvent être introduites dans un vecteur d'expression pour l'expression des compositions du sujet.

20 Ces compositions trouvent une analogie dans des fragments de protéines virales, particulièrement des protéines du virus de la variole. Comme montré à la figure 1, VGF a des résidus hydrophobes et non chargés vers le N-terminal entre les résidus 5 et 15 et vers le
25 C-terminal entre les résidus 100 et 124. Ces résidus peuvent être considérés par analogie aux glycoprotéines de membrane intégrales comme ayant une séquence signal hydrophobe N-terminal, qui est enlevée de façon protéolytique pendant ou immédiatement après la
30 traduction, et une séquence transmembranaire C-terminal, qui sert à accrocher la protéine mature dans la membrane. Le clivage du polypeptide viral à arg-43 et arg-90 du peptide mature conduirait à libérer un polypeptide soluble. Le polypeptide VV de 140 résidus peut être
35 supposé comme donnant naissance premièrement à une protéine associée à la membrane d'approximativement 121

résidus après enlèvement de la séquence d'acides aminés 1g. Suivant un processus intracellulaire, un peptide de facteur de croissance soluble d'environ 77 résidus est obtenu. Stroobant, P et al, Cell (1985) 42:383-393.

5 Une comparaison du facteur de croissance du Myxome (MGF) et du facteur de croissance du fibrome de Shope (SFGF) avec d'autres membres de la famille des facteurs de croissance (voir figure 1) révèle plusieurs différences visibles. MGF et SFGF, synthétisés comme
10 molécules précurseurs, ont une séquence signal au N-terminal mais au contraire des autres membres de facteurs de croissance n'ont pas de domaine transmembranaire au C-terminal, indiquant qu'il y a des molécules secrétées. Une analyse de séquence montre une utilisation
15 significative de l'acide aminé asparagine à la fois dans MGF et SFGF. Dans six positions, MGF et SFGF ont à la fois conservé des asparagines où les autres membres n'ont pas d'asparagine. A ces positions, les asparagines remplacent deux résidus de glycine hautement conservés
20 aux positions 8 et 31, où la première cystéine est à la position 2. Ces deux résidus glycine se trouvent en combinaison avec des résidus de tyrosine grandement conservés aux positions 9 et 32. A cause de la nature hautement conservée des acides aminés dans cette face de
25 la molécule, on suppose que les boucles de cystéine présentes dans cette face sont impliquées avec la liaison au récepteur et avec la fonction des facteurs de croissance dans cette famille.

 Les acides aminés glycine et asparagine sont
30 supposés comme ayant le potentiel de rotation Béta le plus élevé (Chou et Fassman, Ann. Rev. Biochem. (1978) 47 : 251-276). Donc il apparaît que la conservation de ce potentiel de rotation est extrêmement importante pour l'activité du facteur de croissance. Les facteurs de
35 croissance autres que Myxome et Shope utilisent presque exclusivement la glycine dans des régions de rotation

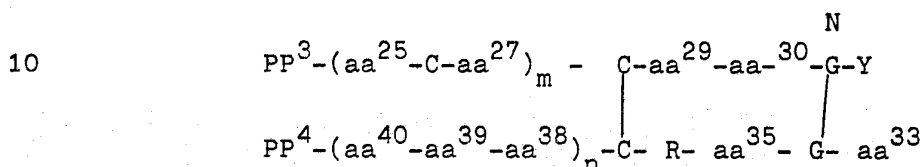
importante de la molécule. Dans MGF et SFGF, ces glycines ont presque toutes été converties en asparagines (voir figure 2). Cela indique une fonction biologique importante dû au remplacement de la glycine grandement conservée avec le résidu d'asparagine. Il est possible que le remplacement de la glycine dans les régions de rotation conservées dans la structure du facteur de croissance par l'asparagine puisse conduire à une fonctionnalité augmentée du facteur de croissance soit en augmentant la liaison au récepteur d'EGF soit à un récepteur apparenté, en augmentant l'effet direct du facteur de croissance dans les cellules affectées ou en augmentant la stabilité de la molécule.

Deux des asparagines nouvellement présentes dans MGF et SFGF fournissent des sites de glycosylation potentiels liés à N. De plus, ces sites sont présents dans les deux boucles de cystéine qui tiennent compte des nombres variables d'acides aminés. Donc, les asparagines nouvellement présentes peuvent offrir de multiples avantages au facteur de croissance au-delà d'une fonctionnalité accrue. Un de ceux-ci pourrait être la stabilisation de la protéine in vivo par protection contre la dégradation où des sites immunogènes par glycosylation.

Les polypeptides du sujet sont caractérisés en ce qu'ils sont capables d'avoir au moins une boucle ou cercle, habituellement trois boucles ou cercles, par suite des cystéines qui forment des ponts, où la partie active physiologiquement de la molécule fournissant une activité en rapport avec le facteur de croissance sera d'environ 12 à 65 acides aminés, habituellement 15 à 60 acides aminés. De ces trois boucles, deux des boucles sont de 12 à 15 acides aminés (14-17 acides aminés annulaires) à l'exclusion du pont de cystéine, plus particulièrement l'une est de 12 ou 13 acides aminés (14 ou 15 acides aminés annulaires) et les autres de 15

acides aminés (17 acides aminés annulaires), où la boucle proche du N-terminal sera habituellement de 12 à 13 acides aminés, plus habituellement de 12 acides aminés, et le milieu de la boucle sera de 15 acides aminés.

5 La troisième boucle ou boucle C-proximale sera de 8 acides aminés (10 acides aminés annulaires) et comprenant des acides aminés de chaque côté qui auront la formule suivante :



où :

Les symboles à une lettre pour les acides aminés ont leur signification habituelle, où C est cystéine, G est la glycine, N est l'asparagine, Y est la tyrosine et R est l'arginine ;

15 aa^{25} est un acide aminé neutre qui peut être aliphatique, particulièrement de 3 à 4 atomes de carbone, ou aromatique, particulièrement de 9 atomes de carbone, et ayant de 0 à 1 groupement hydroxyle, tel que l'alanine, la sérine, la thréonine, la tyrosine ou la phénylalanine;

20 aa^{27} est un acide aminé neutre ou basique, ayant particulièrement de 3 à 6 atomes de carbone, et quand il est neutre ayant de 0 à 1 groupement carboxamide, tel que l'arginine, l'alanine, la valine, la leucine, l'isoleucine, l'asparagine ou la glutamine ;

25 aa^{29} est un acide aminé neutre ou basique, qui est aliphatique ou aromatique, où l'aromatique est exemplifié par l'histidine et l'aliphatique est de 2 à 6, plus habituellement de 3 à 6 atomes de carbone, ayant de

0 à 1 groupement hydroxyle, tel que la sérine, la thréonine, la leucine, la valine, l'isoleucine, ou l'arginine ;

5 aa³⁰ est un acide aminé neutre, qui est aliphatique ou aromatique, ou l'aromatique est exemplifié par l'histidine et l'aliphatique est de 3 à 6 atomes de carbone et ayant de 0 à 1 groupement hydroxyle, particulièrement la sérine, l'isoleucine et la valine .

10 aa³³ signifie un acide aminé aliphatique neutre de 2 à 6, plus habituellement de 3 à 6 atomes de carbone ayant de 0 à 1 groupement hydroxyle, plus particulièrement la sérine, la thréonine, la valine, la leucine ou l'isoleucine, ou un acide aminé acide, tel que représenté par l'acide glutamique et l'acide aspartique ;

15 aa³⁵ est un acide aminé neutre ou acide, de 3 à 6 atomes de carbone, où le neutre est exemplifié par l'alanine, la valine, la leucine, l'isoleucine et la sérine, et où l'acide est exemplifié par les acides aspartique et glutamique ;

20 aa³⁸ est un acide aminé aliphatique neutre substitué où le substituant est un carboxamide et a de 4 à 5 atomes de carbone, où un acide aminé acide, tel que l'asparagine, la glutamine, l'acide aspartique ou l'acide glutamique ;

25 aa³⁹ est un acide aminé aromatique ou un acide aminé neutre aliphatique de 3 à 4 atomes de carbone ayant habituellement un substituant hydroxyle, de préférence aromatique, tel que la phénylalanine, l'histidine, la tyrosine, la sérine ou la thréonine ;

30 aa⁴⁰ est un acide aminé aliphatique neutre substitué ou non substitué de 3 à 6 atomes de carbone, ou un acide aminé basique, tel que l'alanine, la valine, l'isoleucine, la glutamine ou l'arginine ;

m et n sont 0 ou 1 ;

PP³ et PP⁴ peuvent être identiques ou différents et sont soit hydrogènes, indiquant la terminaison du polypeptide, où peuvent être des chaînes de polypeptides pas plus grandes qu'un total de 1000 acides aminés, habituellement pas plus grandes qu'un total d'environ 500 acides aminés, de préférence dans lesquelles au moins 90% des acides aminés, de façon plus préférentielle au moins environ 95% des acides aminés présents sont dans une des deux chaînes de polypeptides ; dans quelques cas, la chaîne peut être d'un acide aminé seulement et pas plus que de 100 acides aminés, fréquemment pas plus qu'environ 50 acides aminés, dépendant de l'utilisation du polypeptide et du rôle de la chaîne étendue, des chaînes de polypeptides peuvent être apparentées aux chaînes des polypeptides se produisant naturellement associées au facteur de croissance se produisant naturellement et aux protéines des virus de la variole ou peuvent être autres que les chaînes se produisant naturellement ou leurs fragments associés avec la chaîne des polypeptides spécifiquement décrite dans la formule, et habituellement non apparentées.

les définitions des acides aminés sont décrites ci-dessous.

25	<u>Neutre</u> (Ne)	
	aliphatique (Al)	
	non substitué	G, A, V, L, I, P
	substitué	
	oxy	S, T
30	thio	C, M
	amido	N, Q
	aromatique (Ar)	
	non substitué	F
	substitué	Y
35	hétérocyclique	H, W

Chargé (à pH 6,0)

basique (Ba)

K, R

Acide (Ac)

D, E

- Les abréviations entre parenthèses se rapportent aux groupes d'acides aminés particuliers. Par non substitué, il est entendu aucun autre hétérosubstituant que le groupement carboxy et amino présent dans la glycine. Les acides aminés sont les L-acides aminés se produisant naturellement.
- Les acides aminés neutres peuvent également être décrits comme ayant des groupements R non polaires ou polaires (mais non chargés). Les acides aminés qui entrent dans ces définitions sont les suivants :

- | | |
|----------------------------|------------------------|
| Acides aminés non polaires | A, V, L, I, P, M, F, W |
| Acide amines polaires | G, S, T, C, Y, N, Q |

- La boucle entre les cystéines aux positions 28 et 37 dans la formule ci-dessus aura de 0 à 1 acide aminé acide (S). Les acides aminés immédiatement de chaque côté (c'est-à-dire les acides aminés 27 et 38) des cystéines en dehors de la boucle peuvent comprendre au moins un acide aminé chargé et au moins un acide aminé aliphatique substitué par un amido. Des acides aminés dans la boucle (28-37) de 4 à 6, habituellement 5 acides aminés seront des acides aminés aliphatiques neutres et pas plus que deux, habituellement un, seront des acides aminés aromatiques.

- Les composés particulièrement intéressants sont les composés où le polypeptide est de moins que 130 acides aminés, plus particulièrement de moins que 55 acides aminés, et d'au moins 44 acides aminés, de préférence au moins 53 acides aminés. Les polypeptides particulièrement intéressants sont les polypeptides comprenant les acides aminés 44 à 88 et leurs fragments, de VGF (aa¹ à aa⁴⁷, voir figure 1).

- De préférence aa²⁹ est un acide aminé aliphatique substitué ou non substitué de 3 à 6 atomes de carbone ayant de 0 à 1 groupement hydroxyle, particulièrement la sérine ou valine ; ou un acide aminé aromatique, particulièrement l'histidine.
- 5 aa³⁰ est de préférence l'histidine, la sérine ou un acide aminé aliphatique non substitué de 5 à 6 atomes de carbone, particulièrement l'isoleucine ;
- 10 aa³³ est de préférence un acide aminé aliphatique substitué ou non substitué ayant de 0 à 1 groupement hydroxyle et de 3 à 6 atomes de carbone ;
- aa³⁵ est de préférence un acide aminé aliphatique de 3 à 6 atomes de carbone ;
- 15 aa³⁸ est de préférence la glutamine ou l'acide glutamique ;
- aa³⁹ est de préférence un aromatique, de façon plus préférentielle l'histidine ou la phénylalanine ;
- aa⁴⁰ est de préférence un acide aminé neutre ou basique.
- 20 La présence de deux boucles est particulièrement intéressante, avec la boucle C-proximale unie à la seconde boucle, où les acides aminés de la seconde boucle peuvent être largement variés. La seconde
- 25 boucle est de préférence d'environ 14 à 16 acides aminés à l'exclusion du pont de cystéines, de façon plus préférentielle d'environ 15 acides aminés. Les acides aminés, de 6 à 9, de préférence de 7 à 9, de façon plus préférentielle environ 8 sont des acides aminés
- 30 aliphatiques soit substitués ou non substitués, de préférence pas plus qu'environ 3 des acides aminés étant substitués, de façon plus préférentielle de 1 à 2 acides aminés étant substitués ; il y aura de 2 à 4 acides
- 35 aminés aromatiques, plus particulièrement 3 acides aminés aromatiques, avantageusement l'histidine et la tyrosine ; il y aura de 2 à 4 acides aminés acides, de préférence 3 acides aminés acides, plus particulièrement l'acide

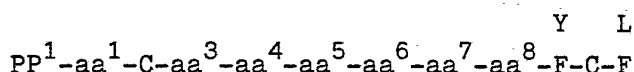
aspartique, et il aura de 0 à 2, plus habituellement un acide aminé basique, particulièrement l'arginine.

Avantageusement, la cystéine formant la boucle du sujet la plus proche de la cystéine de l'autre boucle sera

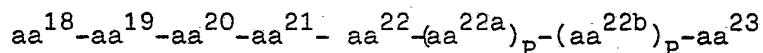
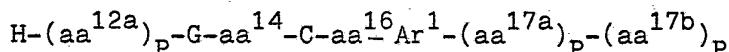
5 séparée par 0 à 2, plus habituellement de 0 à 1 acide aminé, particulièrement l'arginine.

Les composés du sujet particulièrement intéressants pour des applications sont les composés de la formule suivante :

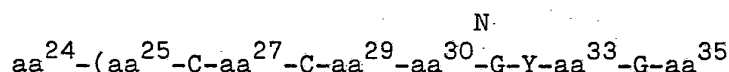
10



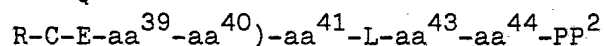
N



15



Q



où :

20 Les symboles aa^{25} à aa^{40} dans la formule ont été décrits ;

PP^1 et PP^2 sont identiques ou différents et peuvent être hydrogènes, indiquant la partie terminale du polypeptide indiqué ou peuvent être des polypeptides
25 ayant un total de jusqu'à environ 1000 acides aminés, plus habituellement jusqu'à 500 acides aminés et qui peuvent avoir un total aussi petit qu'un acide aminé, ou peuvent individuellement ou séparément être des polypeptides de 1 à 100 acides aminés, plus

habituellement d'environ 1 à 75 acides aminés, plus particulièrement d'environ 5 à 50 acides aminés ; ces polypeptides auront des applications spécifiques en modifiant la séquence décrite spécifiquement pour un but prédéterminé ; PP¹ et PP² peuvent être identiques ou
5 différents du polypeptide du virus de la variole naturel et sont habituellement différents ;
aa¹ peut être tout acide aminé, plus particulièrement un acide aminé aliphatique, un acide aminé basique ou un acide aminé acide, de préférence un
10 acide aminé aliphatique non substitué de 2 à 6 atomes de carbone, comprenant la glycine, la leucine, la valine, la lysine, l'acide glutamique et l'acide aspartique ;
aa³ est un acide aminé neutre, particulièrement de 2 à 4 atomes de carbone, plus particulièrement
15 l'asparagine, la glycine et la proline ;
aa⁴ peut être tout acide aminé, aliphatique ou aromatique, particulièrement de 2 à 6 atomes de carbone, qui peut être neutre ou acide, plus particulièrement de 3
20 à 6 atomes de carbone, comprenant la proline, l'acide aspartique, l'histidine, la sérine et la leucine ;
aa⁵ est un acide aminé neutre aliphatique substitué ou acide, particulièrement l'acide glutamique ou l'acide aspartique ou un acide aminé substitué par un
25 hydroxy de 3 à 4 atomes de carbone, plus particulièrement la sérine ;
aa⁶ est un acide aminé aliphatique non substitué de 2 à 6, habituellement 2 à 3 atomes de carbone, ou un acide aminé aromatique, particulièrement
30 la glycine, l'histidine ou la tyrosine ;
aa⁷ est un acide aminé acide, basique ou neutre, particulièrement de 3 à 6 atomes de carbone, tel que l'acide glutamique, l'acide aspartique, la lysine et la thréonine ;

- aa⁸ est un acide aminé neutre aliphatique non substitué de 2 à 3 atomes de carbone ou un acide aminé substitué avec un carboxamide de 4 à 5 atomes de carbone, tel que l'asparagine, la glutamine ou la glycine ;
- 5 aa¹¹ est un acide aminé neutre, soit aliphatique ou aromatique, plus particulièrement aliphatique, particulièrement de 5 à 6 atomes de carbone, plus particulièrement la leucine ou la phénylalanine ;
- 10 aa¹² est un acide aminé aromatique ou un acide aminé aliphatique substitué par un carboxamide de 4 à 5 atomes de carbone, particulièrement l'histidine ou l'asparagine ;
- 15 aa^{12a} est un acide aminé neutre aliphatique ou un acide aminé acide, plus particulièrement la glycine, l'asparagine ou l'acide aspartique ;
- 20 aa¹⁴ est un acide aminé acide ou un acide aminé neutre aliphatique, substitué ou non substitué, de 4 à 5 atomes de carbone, ayant de 0 à 1 groupement hydroxyle, particulièrement l'acide aspartique, la thréonine ou la valine ;
- 25 aa¹⁶ est une liaison ou un acide aminé aliphatique neutre ou basique, soit substitué ou non substitué, de 4 à 6 atomes de carbone, particulièrement l'isoleucine, l'arginine ou la méthionine ;
- 30 Ar¹ signifie un acide aminé aromatique, qui peut avoir un cycle carboxylique ou hétérocyclique et comprend l'histidine, la phénylalanine ou la tyrosine ;
- 35 aa^{17a} est un acide aminé aliphatique neutre, soit substitué ou non substitué, de 3 à 6, habituellement 3 à 5 atomes de carbone, ayant 0 à 1 groupement hydroxyle et comprend la valine, la leucine, l'isoleucine, et la thréonine ;
- aa^{17b} est un acide aminé neutre non substitué de 5 à 6 atomes de carbone, ou un acide aminé acide, plus particulièrement la valine, l'isoleucine ou l'acide glutamique ;

- 5 aa¹⁸ est un acide aminé aliphatique neutre,
substitué ou non substitué, de 2 à 6, de préférence de 3
à 6 atomes de carbone ayant de 0 à 1 groupement
hydroxyle, particulièrement l'alanine, la sérine ou la
glutamine ;
- 10 aa¹⁹ peut être tout acide aminé,
particulièrement les acides aminés aliphatiques acides ou
basiques, plus particulièrement de 4 à 6 atomes de
carbone, de façon plus préférentielle de 5 à 6 atomes de
carbone, particulièrement l'arginine, la leucine, la
valine et l'acide glutamique ;
- 15 aa²⁰ est un acide aminé acide ou un acide aminé
aliphatique substitué par un carboxamide et comprend
l'acide aspartique, l'asparagine, l'acide glutamique ou
la glutamine ;
- 20 aa²¹ est un acide aminé aliphatique non
substitué neutre ou un acide aminé basique de 3 à 6
atomes de carbone ayant de 0 à 1 groupement hydroxyle ou
carboxamide, et comprend l'isoleucine, la leucine,
l'asparagine, la sérine, la thréonine, la lysine et
l'arginine ;
- 25 aa²² est une liaison ou un acide aminé acide
qui est l'acide aspartique ou l'acide glutamique ou
l'acide aminé aliphatique neutre, particulièrement la
valine ;
- 30 aa^{22a} est un acide aminé aliphatique neutre,
ayant 0 ou 1 groupement hydroxyle, particulièrement un
groupement hydroxyle, plus particulièrement la sérine ;
aa^{22b} est un acide aminé aliphatique neutre
ayant 4 à 6 atomes de carbone, particulièrement la
leucine ou l'isoleucine ;
- 35 aa²³ est une liaison ou un acide aminé
aliphatique neutre, substitué ou non substitué, de 2 à 6
atomes de carbone ayant de 0 à 1 substituant hydroxyle,
comprenant la glycine, la sérine, la thréonine et
l'asparagine ;

aa²⁴ est un acide aminé aliphatique, particulièrement un acide aminé aliphatique thiosubstitué, plus particulièrement la méthionine, la proline ou un acide aminé aromatique, particulièrement la tyrosine ;

aa⁴¹ est un acide aminé acide ou aliphatique non substitué ou substitué neutre de 4 à 6 atomes de carbone, particulièrement la valine, l'asparagine, l'acide aspartique ;

aa⁴³ est un acide aminé aliphatique neutre, acide ou basique, qui peut être substitué ou non substitué de 4 à 6 atomes de carbone, et comprend la valine, la leucine, l'isoleucine, l'arginine, la lysine et l'acide aspartique ;

aa⁴⁴ peut être tout acide aminé, particulièrement autre qu'un acide aminé basique, et peut être acide, neutre ou aromatique, quand il est autre qu'aromatique, ayant de 3 à 6 atomes de carbone, particulièrement l'acide aspartique, l'alanine, la leucine, le tryptophane ou la thréonine ; et

p est 0 ou 1.

Il y a un intérêt particulier quand ^{PP1} a la séquence suivante :

PP1, -aa⁻²⁴ -aa⁻²³ -aa⁻²² -aa⁻²¹ -aa⁻²⁰ -aa⁻¹⁹ -aa⁻¹⁸

-aa⁻¹⁷ -aa⁻¹⁶ -aa⁻¹⁵ -aa⁻¹⁴ -aa⁻¹³ -aa⁻¹² -aa⁻¹¹

aa⁻¹⁰ -aa⁻⁹ -aa⁻⁸ -aa⁻⁷ -aa⁻⁶ -aa⁻⁵ -aa⁻⁴ -aa⁻³ -aa⁻² -aa¹

où :

PP¹ en combinaison avec les symboles d'acides aminés suivants dans la séquence ci-dessus est l'équivalent de PP¹. La séquence ci-dessus entre dans la définition de PP¹ ; PP¹ peut être hydrogène ou une séquence d'acides aminés. Un ou plusieurs des acides

aminés, symbolisés par aa^{-x} (x est tout nombre), peut être une liaison, afin de servir à réduire le nombre d'acides aminés dans la chaîne N-terminal. De plus, toute ou une partie de la séquence d'acides aminés indiquée peut être présente. Quand une partie de la séquence est présente, de préférence la séquence impliquera des acides aminés contigus, c'est-à-dire des acides aminés dans leur ordre numérique sans délétion. Habituellement, il y aura moins un acide aminé, plus habituellement au moins trois, plus habituellement au moins six, où les acides aminés restants en amont peuvent être absents, afin que $PP^{1'}$ puisse être uni à aa^{-1} , aa^{-6} , ou analogues ;

aa^{-1} peut être tout acide aminé, particulièrement aliphatique, d'environ 3 à 6 atomes de carbone, de préférence neutre ou basique, plus particulièrement comprenant la lysine, l'arginine, la proline, l'asparagine ou la sérine ;

aa^{-2} peut être tout acide aminé, soit aliphatique ou aromatique, particulièrement aliphatique plus particulièrement d'environ 3 à 6 atomes de carbone ;

aa^{-3} peut être aliphatique ou aromatique, particulièrement aliphatique, soit polaire ou non polaire, de 2 à 6, plus habituellement de 2 à 5 atomes de carbone, généralement ayant de 0 à 1 hydroxyle ;

aa^{-4} peut être tout acide aminé aliphatique, neutre ou basique, généralement de 3 à 6 atomes de carbone, habituellement de 5 à 6 atomes de carbone, particulièrement l'asparagine, la proline ou la lysine ;

aa^{-5} peut être tout acide aminé aliphatique, particulièrement neutre, généralement de 3 à 6 atomes de carbone, particulièrement la valine ou l'isoleucine ;

aa^{-6} est un acide aminé neutre ou acide, qui lorsqu'il est aliphatique a de d'environ 3 à 6, plus habituellement environ 4 à 6 atomes de carbone, plus particulièrement, l'isoleucine, la valine ou l'acide aspartique ;

- aa⁻⁷, aa⁻⁸, aa⁻⁹, aa⁻¹², aa⁻¹⁴, aa⁻¹⁵, aa⁻¹⁶,
 aa⁻¹⁷, aa⁻¹⁹, aa⁻²⁰, aa⁻²¹ et aa⁻²³ sont des acides aminés
 aliphatiques de 3 à 6 atomes de carbone, soit polaires ou
 non polaires, particulièrement polaires ayant de 0 à 1
 5 substituant hydroxyle ou aromatiques ;
 aa⁻¹⁰ et aa⁻²⁵ sont des acides aminés
 aliphatiques non polaires ou des acides aminés acides de
 2 à 6, plus habituellement de 2 à 3 atomes de carbone ;
 aa⁻¹³, aa⁻²², et aa⁻²⁴ sont des acides aminés
 10 aliphatiques polaires ou des acides aminés acides de 4 à
 5 atomes de carbone, particulièrement ayant un
 substituant carboxamido ;
 aa⁻¹¹ et aa⁻¹⁸ sont des acides aminés
 aliphatiques, soit polaires ou non polaires, ou des
 15 acides aminés acides.

- Un fragment N-terminal particulièrement
 intéressant est un fragment N-terminal depuis aa⁻¹ à
 aa⁻²⁴ ou aa⁻¹ à aa⁻⁷, plus particulièrement aa⁻¹ à aa⁻⁶.
 PP1' peut de manière convenable être une séquence d'acide
 20 aminé non apparentée qui peut servir comme une protéine
 de fusion, particulièrement pour fournir un peptide
 immunogène pour la production d'anticorps au composé.

- Les aspects primaires des compositions du sujet
 sont la séquence aa²⁵ à aa⁴⁰, particulièrement la
 25 séquence à partir de la cystéine à la position 28 à la
 cystéine à la position 37 dans les boucles produites par
 les cystéines aux positions 2 et 10 et les cystéines aux
 positions 15 et 26. Avantageusement, les compositions du
 sujet ont la boucle créée par les cystéines aux positions
 30 28 et 37 conjointement avec la boucle créée par les
 cystéines aux positions 15 et 26.

- Les séquences de polypeptides chimériques
 peuvent être préparées en combinant les fragments de
 divers polypeptides ayant une séquence sensiblement
 35 similaire aux chaînes des polypeptides se produisant
 naturellement associées avec des facteurs de croissance

se produisant naturellement et des protéines des virus de la variole. Les polypeptides chimériques résultants auront d'environ 40 à 65 acides aminés, habituellement environ 45 à 60 acides aminés, particulièrement 49 à 53 acides aminés. Dans chaque cas, la structure du squelette des cystéines est maintenue avec les ponts de cystéines définissant des boucles des tailles décrites antérieurement. Donc, un fragment peut être employé à partir de tout facteur de croissance ayant une homologie substantielle avec VGF (voir par exemple figure 1) où un autre facteur de croissance ayant la même structure de squelette que les compositions du sujet et une activité physiologique similaire. Ceci proviendra de manière indifférente de la source des mammifères, tels que des primates, tels que humains, des rongeurs, tels que rats et souris, des bovins, des aviens, des porcins, etc... .

Il peut être requis jusqu'à trois jonctions, selon la source des fragments et de la longueur du polypeptide désiré. Avantageusement, les jonctions seront faites à un point entre aa⁻⁶ et aa¹ ; aa¹¹ et aa¹⁵ ; aa²⁵ et aa²⁹ ; et aa⁴² et aa⁵³. La séquence entre environ aa⁻⁶ et environ aa¹⁵ est référée comme la séquence N-terminale ; celle entre environ aa¹⁴ et environ aa²⁹ est référée comme le fragment central ; et celle entre environ aa²⁵ et l'extrémité du peptide (généralement aa⁴⁴ à aa⁴⁷) est référée comme le domaine C-terminal. Le point exact de jonction variera en fonction de la localisation des sites de restriction, etc... . De plus, une séquence extrême N-terminale comprenant aa⁻²⁴ à aa⁻⁷, plus habituellement aa⁻²³ à aa⁻⁷, peut également être jointe aux polypeptides chimériques.

Les fragments seront généralement d'environ 15 à 50, habituellement 15 à 45 acides aminés, puisque aa²⁰ ne représente pas le vingtième acide aminé du composé, mais seulement de la séquence spécifique qui a été spécifiquement définie (voir figure 1).

Plusieurs sites de restriction convenables sont désignés dans les gènes synthétiques utilisés pour construire les polypeptides chimériques. Quand cela est possible, les sites de restriction laissent la séquence d'acides aminés du gène du facteur de croissance non altérée. Cependant, dans certains cas, l'incorporation du nouveau site de restriction produit une séquence d'acides aminés altérée. La séquence codante particulièrement intéressante est la séquence codante où la séquence codante pour VGF est modifiée pour introduire un site de restriction KpnI en changeant la séquence d'acides aminés à aa¹³ à aa¹⁵ à partir de GDC à GTC et pour introduire un site SphI à aa²³ à aa²⁸ en changeant la séquence GMYCRC en GYACVC. Ces changements fournissent des sites appropriés pour des fragments de liaison du gène de VGF aux fragments d'autres facteurs de croissance. Par exemple, le fragment de gène aa²⁶ à aa⁴⁷ peut être joint de manière appropriée à un fragment de gène de α -TGF codant les acides aminés correspondants aux résidus aa⁻⁶ à aa²⁵ pour produire un polypeptide hybride TGF/VGF. Donc, la séquence du peptide chimérique comprendra aa⁻⁶ à aa¹³ (N-terminal) et aa¹⁴ à aa²⁵ (fragment central) dérivé de α TGF, et aa²⁶ à aa⁴⁷ (C-terminal) dérivé de VGF, modifié comme ci-dessus.

La préparation de plasmides capables d'exprimer des protéines chimériques ayant des séquences d'acides aminés dérivées des fragments de plus d'un facteur de croissance ou d'un polypeptide du virus de la variole avec des changements de séquence où il est nécessaire d'introduire un site de restriction approprié comme décrit en détail dans la partie expérimentale.

Des protéines de fusion peuvent également être préparées où un facteur de croissance ou un polypeptide chimérique est exprimé joint aux acides aminés de N-terminal d'une séquence signal ou aux acides aminés de N-terminal d'un gène grandement exprimé de bactérie,

bactériophage ou eucaryote. De plus, un site de clivage enzymatique ou chimique peut être fourni suivant la séquence leader pour permettre le clivage du polypeptide mature à partir de la séquence leader. La préparation des
5 plasmides capables d'exprimer ces polypeptides de fusion et des procédés pour le clivage et l'isolement des polypeptides désirés sont décrits en détail dans la partie expérimentale.

Préparation de facteurs de croissance.

10 Les compositions du sujet peuvent être préparées par une variété de voies dépendant de la taille de la composition. Particulièrement en dessous d'environ 80, plus particulièrement en dessous d'environ 60 acides aminés, la composition peut être préparée par synthèse
15 suivant des voies classiques. Voir, par exemple, Merrifield, Solid-Phase Peptide Synthesis, "The Peptides Analysis, Synthesis Biology, "Special Methods in Peptide Synthesis, Part A, Vol. 2, Gross and Meinhofer eds., Academic Press, NY, 1980, pp. 1-284. Voir aussi, Brevet
20 U.S. No. 4 127 526.

Autrement, l'utilisation d'une technologie d'ADN hybride peut être employée, où les séquences d'ADN qui codent pour le polypeptide désiré ou son précurseur peuvent être utilisées. Des séquences d'ADN codant pour
25 le facteur de croissance intéressant peuvent être synthétisées en employant des techniques classiques telles que des brins simples se chevauchant qui peuvent être liés ou ligaturés ensemble pour définir la séquence codante souhaitée. Les terminaux peuvent être indiqués
30 pour fournir des sites de restriction où un ou deux terminaux peuvent être coupés en bouts francs pour la ligation aux extrémités complémentaires d'un vecteur d'expression. Pour l'expression de la séquence, on fournit une méthionine initiale. Les vecteurs
35 d'expression sont généralement disponibles et sont amplement décrits dans la littérature.

Au lieu de synthétiser le gène intéressant, le gène du facteur de croissance peut être isolé par diverses techniques. Par exemple, quand le facteur de croissance est un facteur de croissance viral, ARNm peut être isolé à partir du virus qui code pour le polypeptide contenant le facteur de croissance, l'ARNm peut être transcrit de façon inverse, l'ADN simple brin (ss) résultant peut être utilisé comme matrice pour préparer l'ADN double brin (ds) et le gène de l'ADN ds peut être isolé. Une autre technique consiste à isoler une partie de l'ADN viral et à utiliser une sonde dégénérée de façon appropriée, comprenant une région des séquences les plus conservées dans un gène du facteur de croissance, à identifier les séquences codant pour un facteur dans le génome viral. La sonde peut être considérablement plus courte que la séquence entière, mais sera au moins de 10, de préférence au moins 14, de façon plus préférentielle d'au moins 20 nucléotides en longueur. Des oligonucléotides plus longs sont également utilisables, jusqu'à la longueur complète du gène du facteur de croissance. A la fois les sondes d'ADN et d'ARN peuvent être utilisées.

Dans l'utilisation, les sondes sont typiquement marquées d'une manière détectable, (par exemple avec ^{32}P ou des nucléotides biotinylés) et sont incubés avec l'ADN simple brin ou l'ARN du micro-organisme dans lequel un gène doit être cherché. L'hybridation est détectée au moyen du marqueur après que l'ADN simple brin et double brin (hybridé) (ou ADN/ARN) ait été séparé, typiquement en utilisant une feuille de nitrocellulose. Les techniques d'hybridation appropriées pour une utilisation avec des oligonucléotides sont bien connues de ceux qui sont familiers de la technique.

Bien que les sondes soient normalement utilisées avec un marquage détectable qui permet une identification aisée, des oligonucléotides non marqués

sont également utilisables, à la fois comme précurseurs de sondes marquées et pour une utilisation dans des procédés qui permettent une détection directe de l'ADN double brin (ou ADN/ARN). En conséquence, le terme

5 "oligonucléotide" fait référence à la fois aux formes marquées et non marquées.

Quand la séquence d'ADN désirée a été obtenue, elle peut ensuite être manipulée selon une variété de voies pour permettre une expression ; par exemple des

10 séquences de polypeptides chimériques peuvent être préparées en combinant des fragments de gène d'au moins deux polypeptides ayant des séquences qui sont au moins homologues à plus de 30% aux facteurs de croissance se produisant naturellement qui se lient au récepteur d'EGF

15 ("ligands naturels"). Il est très souhaitable que la structure tridimensionnelle, particulièrement la structure de la boucle, du polypeptide soit maintenue, particulièrement celle de la partie ou des parties de la structure qui peuvent être chargées de la liaison au

20 récepteur d'EGF et de l'activité biologique des facteurs de croissance se produisant naturellement qui se lient au récepteur d'EGF.

En fonction de la source des fragments et de la longueur du polypeptide désiré, des sites de restriction

25 appropriés peuvent être désignés dans les gènes de synthèse utilisés pour construire des polypeptides chimériques, tel que décrit ci-dessus. Quand cela est possible, le ou les sites de restriction laissent la séquence d'acides aminés du polypeptide inaltérée.

30 Cependant, dans certains cas, l'incorporation de nouveau(x) site(s) de restriction peut fournir une séquence d'acides aminés altérée sans changer l'activité de la protéine.

Quand le gène doit être exprimé dans un hôte

35 qui reconnaît les régions de régulation de type sauvage de la transcription et de la traduction du facteur de

croissance, le gène entier avec ses régions de régulation 5'- et 3'- de type sauvage peut être introduit dans un vecteur d'expression approprié. Il existe divers vecteurs d'expression employant des systèmes de replication à
5 partir des virus de mammifères, tels que Simian Virus 40, l'adénovirus, le virus du papillome des bovins, le virus de la vaccine, le baculovirus des insectes, etc... . Ces systèmes de replication ont été développés pour fournir des marqueurs qui permettent une sélection des
10 transfectants, aussi bien que pour fournir des sites de restriction appropriés dans lesquels le gène peut être inséré.

Quand le gène doit être exprimé dans un hôte qui ne reconnaît pas les régions de régulation de type sauvage de la transcription et de la traduction se
15 produisant naturellement, une manipulation supplémentaire sera requise. D'une manière convenable, une variété de régions régulatrices de la transcription à partir de 3' sont connues et peuvent être insérées en aval des codons d'arrêt. La région -5' non codante en
20 amont du gène intéressant peut être enlevée par restriction par une endonucléase, par résection par Bal31 ou analogues. Alternativement, quand un site de restriction approprié est présent à proximité du
25 5'-terminal du gène intéressant, le gène intéressant peut être restreint et un adaptateur peut être employé pour lier le gène intéressant à la région du promoteur, où l'adaptateur pourvoit aux nucléotides perdus du gène intéressant. Diverses stratégies peuvent être employées
30 pour fournir une cassette d'expression, qui dans la direction 5'-3' de transcription a une région régulatrice de la transcription et une région d'initiation de la traduction, qui peut également comprendre des séquences de régulation permettant
35 l'induction d'une régulation; le gène intéressant sous le

contrôle de la transcription et de la traduction de la région d'initiation; et une région de terminaison de la transcription et de la traduction.

Le choix des séquences de régulation appropriées tiendra compte des facteurs suivants qui affectent l'expression. En termes de régulation de la transcription, la quantité et la stabilité du ARN messenger sont des facteurs importants qui influencent l'expression des produits des gènes. La quantité ARNm est déterminée par le nombre de copies du gène particulier, par l'efficacité relative de son promoteur et par les facteurs qui régulent le promoteur, tels que des stimulateurs ou activateurs ou des répresseurs. On pense que l'initiation survient dans la région juste en amont du commencement de la séquence codante.

Le promoteur dans des cellules procaryotiques comprend des séquences de nucléotides qui peuvent affecter l'efficacité de la transcription. Les séquences comprennent les régions de régulation à environ -35 et -10 nucléotides à partir du départ de la chaîne d'ADN. Des promoteurs efficaces comprennent ceux dans lesquels la séquence des nucléotides des régions de régulation -35 et -10 est sensiblement la même que celle des séquences consensus pour ces régions dans des promoteurs bactériens à partir de gène grandement efficaces. Généralement, ces régions sont respectivement d'environ 5 nucléotides et 6 nucléotides en longueur, et chaque séquence peut varier d'environ un nucléotide en longueur et/ou en séquence. Une séquence préférée pour la séquence de régulation consensus -35 est à partir du promoteur trp, c'est-à-dire TGACA, et dans la séquence de régulation consensus -10 est à partir du promoteur lac, c'est-à-dire TATAAT.

Non seulement les séquences de nucléotides mais également l'espacement des séquences consensus des régions de régulation -35 et -10, l'une par rapport à l'autre, est important pour obtenir une transcription

optimale du gène intéressant. Généralement, les séquences consensus des régions de régulation -35 et -10 sont séparées par environ 16 à 18 nucléotides, de préférence environ 17 nucléotides. Chaque séquence peut
5 varier d'environ un nucléotide.

Des exemples de régions régulatrices de la transcription ou des promoteurs comprennent, pour les bactéries, le promoteur β -gal, des promoteurs gauche et droit lambda, les promoteurs trp et lac, les
10 promoteurs de fusion trp-lac, etc..., les promoteurs de synthèse ayant des séquences sensiblement similaires à ces séquences peuvent également être utilisés. Un promoteur préféré pour bactérie est un promoteur de fusion comprenant la région de régulation -35 à partir du
15 promoteur trp et la région de régulation -10 à partir du promoteur lac. De façon la plus préférentielle, le promoteur est un dans lequel la séquence consensus trp -35 est localisée à environ 17 nucléotides en amont de la séquence consensus lac -10; pour la levure, les
20 promoteurs d'enzymes glycolytiques, tels que les promoteurs ADH-I et -II, le promoteur GPK et le promoteur PGI, le promoteur trp, etc... ; pour des cellules de mammifères, les promoteurs retard et précoce SV40, les promoteurs retard principaux de l'adénovirus, le
25 promoteur I5 ou les séquences de stimulateur et analogues.

La région régulatrice de la transcription peut de plus comprendre des séquences de régulation qui permettent, au temps d'expression du gène intéressant,
30 d'être modulé, tel que par la présence ou l'absence des produits nutritifs ou de produits d'expression dans le milieu de croissance, de la température, etc... . Par exemple, l'expression du gène intéressant peut être
35 régulée par la température du milieu de croissance de la cellule hôte en incluant l'utilisation d'une séquence de régulation comprenant le promoteur P_L du bactériophage

lambda, l'opérateur O_L du bactériophage lambda et le gène CI857 qui code pour le répresseur C_I sensible à la température dans le vecteur d'expression. Cela permettra une régulation du promoteur par interaction entre le
5 répresseur et l'opérateur aux températures faibles, par exemple à environ 30°C. Une augmentation de la température à environ 42°C inactiverait le répresseur et permettrait l'expression du gène intéressant.

Comme exemple d'une modulation utilisant les
10 éléments nutritifs du milieu de croissance, la régulation du promoteur lac ou du promoteur hybride trp-lac peut être accomplie par une utilisation du gène pour le répresseur lacI, qui se lie à la région du promoteur lac en aval de la région de régulation -10. Le gène du
15 répresseur lacI peut être présent sur un épisode, de préférence le mutant amélioré lacI, ou peut être inclus dans la cassette d'expression elle-même. La présence d'une concentration significative de la molécule du répresseur dans le milieu de croissance inhibe la
20 fonction du promoteur en l'absence d'inducteurs. Donc l'addition de IPTG ou de lactose au milieu de la croissance de la cellule hôte accroît la fonction du promoteur. Quand la souche bactérienne est Lac⁺, le lactose peut être utilisé comme un inducteur au lieu de
25 IPTG. Dans les systèmes à la fois eucaryotique et procaryotique, les régions de terminaison peuvent également contenir des séquences ou des structures qui augmentent la stabilité des espèces de ARNm et permettent une expression plus grande. Donc, la région de régulation
30 de la transcription peut de plus comprendre des séquences de régulation qui terminent la transcription et qui fournissent des séquences ou des structures qui inhibent la dégradation du ARNm et donc augmentent la stabilité des espèces d'ARNm et permettent une plus grande
35 expression. Plusieurs exemples de séquences procaryotiques sont connues, par exemple le terminateur

Trp, le terminateur du gène 32 (T4) du gène ou des terminateurs de synthèse qui sont similaires en séquence au gène 32.

Dans des systèmes eucaryotiques, la terminaison
5 de la transcription, des régions de clivage de l'ARN et
de polyadénylation fournissent une maturation appropriée
pour les transcrits ARNm et sont nécessaires pour une
expression efficace. La région native non traduite 3'
10 peut suffire, mais le signal de polyadénylation à partir
de, par exemple SV40, particulièrement comprenant un site
d'épissure qui fournit une expression plus efficace, peut
également être utilisé. Alternativement, la région 3'
-non traduite dérivée d'un gène grandement exprimé dans
15 un type de cellule particulier (par exemple, Ig dans les
cellules de Myelome) pourrait être fusionnée avec le gène
intéressant. De telles séquences 3' pourraient également
être appariées avec des séquences de régulation 5' du
même gène grandement exprimé et pourraient même
20 comprendre des régions codant dans le cadre de lecture
pour produire des protéines de fusion comme décrit
ci-dessous.

En termes de régulation de la traduction, étant
donné la présence d'ARNm, l'expression peut être régulée
en influençant la vitesse d'initiation (liaison du
25 ribosome au RNAm), la vitesse d'élongation (translocation
du ribosome le long de l'ARNm), la vitesse des
modifications de post-traduction et la stabilité du
produit des gènes. La vitesse d'élongation est
probablement affectée par le traitement du codon ;
30 l'utilisation de codons pour des tARNs rares peut réduire
la vitesse de traduction. Il est donc préférable
d'utiliser des codons qui apparaissent fréquemment dans
les gènes normalement exprimés par la cellule hôte pour
augmenter la vitesse de traduction.

Chez les procaryotes, en aval des régions de régulation -35 et -10, il y a une séquence de nucléotides consensus, généralement AGGA, pour désigner la séquence de Shine-Dalgarno, qui est supposée être impliquée dans la liaison au ribosome. Une liaison au ribosome optimale et une initiation de la traduction peut être exécutée en utilisant un site de liaison au ribosome fonctionnel dans la cellule hôte d'un gène grandement exprimé. Il existe également une preuve de la présence de séquences de nucléotides entourant la séquence de Shine-Dalgarno et des séquences dans la région codante qui peuvent affecter les liaisons au ribosome, éventuellement par la formation de motifs structurels à travers lesquels le ribosome reconnaît le site d'initiation. L'altération de séquences de nucléotides de la région codante peut donc être utilisée pour exécuter une liaison optimale et une initiation de la traduction. La séquence des premiers environ 7 à 30 codons après le codon d'initiation ATG peut également affecter la liaison et l'expression. De préférence, la séquence leader et la séquence de Shine-Dalgarno sont obtenues à partir du même gène, où quand ils sont obtenus à partir de gènes différents, les codons de la séquence leader peuvent être modifiés en utilisant leur dégénération des codons pour rapprocher la séquence nucléotidique du gène naturel qui suit la séquence leader, comme décrit dans la demande de brevet PCT/US 88/03972, dont la description est incorporée ici à titre de référence.

La position de la séquence AGGA par rapport au codon d'initiation ATG peut influencer l'expression. Généralement la séquence de Shine-Dalgarno est localisée à partir d'environ 5 à 9 nucléotides à partir du codon d'initiation, bien que, de manière inattendue, des niveaux élevés d'expression puissent être obtenus en utilisant des cassettes d'expression où la séquence de

Shine-Dalgarno est localisée à partir d'environ 10 à 13 nucléotides, de préférence 11 à 12 nucléotides à partir du codon d'initiation.

La stabilité du ARNm est gouvernée par la sensibilité de l'ARNm aux enzymes ribonucléases. En général, la digestion par exonucléase est inhibée par la présence de motifs structuraux aux extrémités du ARNm ; structures de palindrome, nucléotides altérés ou nucléotides spécifiques. La digestion par endonucléase est supposée survenir aux sites de reconnaissance spécifiques dans l'ARNm et l'ARNm stable serait dépourvu de ces sites. Il est également évident que des ARNm subissant des niveaux élevés de traduction soient également protégés de la dégradation par la présence des ribosomes sur l'ARNm.

La stabilité du produit d'expression peut être obtenue en fournissant la synthèse d'une protéine de fusion dans laquelle le polypeptide désiré est exprimé conjointement avec un second polypeptide ou son fragment. De préférence, la stabilité du produit d'expression est obtenue en fournissant des synthèses d'une protéine de fusion dans laquelle le polypeptide intéressant est exprimé joint à une séquence leader. Une séquence d'ADN codant pour une séquence d'acides aminés N-terminale à partir par exemple d'un gène bactériophage ou de bactérie hautement exprimé tel que le gène de protéine -N du bactériophage lambda, le gène cro ou le gène de la β -galactosidase ou un gène eucaryotique qui est joint en amont de et dans leur cadre de lecture avec le gène d'intérêt. La séquence leader comprend habituellement environ 8 à environ 50, de préférence d'environ 15 à environ 45, de façon plus préférentielle 18 à 25 acides aminés de N-terminal.

L'expression du polypeptide d'intérêt comme une protéine fusionnée avec une séquence leader d'un autre gène à plusieurs avantages supplémentaires pour fournir

la stabilité. Par exemple, la présence des acides aminés du N-terminal fournit un moyen pour utiliser des techniques de purification générales pour la purification de toute une variété de polypeptides. Par exemple, les
5 acides aminés de N-terminal de la protéine -N sont particulièrement antigéniques, et donc des anticorps spécifiques élevés contre les acides aminés de N-terminal de la protéine- N peuvent être utilisés comme moyen pour purifier des protéines de fusion contenant le N-terminal
10 de la protéine-N. De plus, le N-terminal de la protéine-N a une charge positive élevée, qui facilite la purification de la protéine désirée par chromatographie par échange d'ions et analogues.

La séquence leader peut également être une
15 séquence d'acides aminés hydrophobes qui peut agir de plus comme une séquence signal pour la sécrétion. Une séquence d'ADN codant pour la séquence signal est jointe en amont de et dans le cadre de lecture avec le gène d'intérêt. Typiquement, la séquence signal comprend un
20 site de clivage qui est reconnu par une peptidase de la séquence signal. Donc, le position du polypeptide d'intérêt directement après le site de clivage de la séquence signal permettra au polypeptide d'intérêt d'être clivé de façon spécifique de la séquence signal et d'être
25 sécrété comme un polypeptide mature. Des exemples de séquence d'acides aminés hydrophobes comprennent la séquence signal phosphatase alcaline bactérienne ; des séquences signal OMP-A-B-C-D-E ou F ; la séquence signal LPP ; la séquence signal B-lactamase ; et les séquences
30 signal de toxine et leurs mutants. Pour des cellules eucaryotiques, les séquences signal peuvent comprendre la séquence signal du TGF- β , du gène P97 simien (antigène de mélanome humain), d'une séquence de toxine tueuse de K. lactis, du facteur de type α -appariement et d'autres

séquences signal de toxines tueuses ou la séquence signal normale associée avec le gène d'intérêt ensemble avec des mutants de la séquence signal.

D'autres séquences leader qui peuvent être
5 utilisées comprennent des séquences hydrophiles, par exemple des résidus de 41 acides aminés de N-terminal de l'amphiréguline qui peut fournir une modification de la fonction du polypeptide d'intérêt. En plus, un agent cytotoxique tel qu'un fragment de la chaîne-A d'une
10 toxine, de la chaîne- A de ricine, d'un peptide arrêtant la croissance du venin de serpent, ou une molécule ciblant telle qu'un hormone ou un anticorps peut être couplé de façon covalente avec la séquence leader avec, dans la plupart des cas, un effet minimal sur l'activité
15 biologique du produit du gène d'intérêt. Comme décrit pour les autres séquences leader, une séquence ADN codant la séquence leader et jointe en amont de et dans le cadre de lecture avec le gène d'intérêt.

Quand la séquence leader n'est pas une séquence
20 signal ou ne contient pas un site de clivage naturel convenable, des acides aminés supplémentaires peuvent être insérés pour fournir un site de clivage enzymatique ou chimique pour le clivage du peptide leader, suit une purification de la protéine de fusion, pour permettre une
25 purification subséquente du polypeptide mature. Par exemple, l'introduction de liaisons aspartyl-proline instables à l'acide entre deux segments de la protéine de fusion facilite leur séparation à pH faible. Cette méthode n'est pas appropriée si le polypeptide désiré est
30 instable aux acides. Comme autre exemple, la protéine de fusion peut être clivée avec, par exemple du bromure de cyanogène qui est spécifique pour le côté carboxy des résidus de méthionine. Le positionnement méthionine entre la séquence leader et le polypeptide désiré permet la

libération du polypeptide désiré. Ce procédé n'est pas approprié quand le polypeptide souhaité contient des résidus de méthionine.

Quand la séquence leader comprend une séquence signal, les gènes d'intérêt avec des séquences leader 5 secrétrices peuvent être exprimées avec ou sans la séquence leader dans des conditions où la séquence peut être maintenue ou clivée. En plus, pour obtenir une proportion élevée du polypeptide désiré comme un peptide 10 mature, clivé et replié ou ré-enroulé, secrété dans le milieu, il est préférable d'utiliser un promoteur, tel que le promoteur tac, qui peut agir à une température inférieure, par exemple d'environ 30°C. D'une façon inattendue, des niveaux élevés de secrétions peuvent être 15 obtenus à des températures inférieures. Des niveaux d'expression extrêmement élevés peuvent prévenir l'apparition des modifications de traduction entière de la protéine, aboutissant à une aggrégation et à une accumulation de précurseur non clivée (c'est-à-dire la 20 protéine structurelle et le leader sécréteur). De façon similaire, la croissance aux températures élevées, par exemple 42°C, tend à aboutir à une aggrégation et accumulation du précurseur non clivé. La préparation de protéines fusionnées comprenant des procédés de clivage 25 et de repliage ou ré-enroulement du polypeptide d'intérêt sont mieux décrits dans la demande de brevet US copendante No. 264,098. supra.

Après chaque manipulation de l'ADN dans le développement de la cassette, le plasmide sera cloné et 30 isolé et comme requis, le constituant particulier de la cassette sera analysé jusqu'à sa séquence pour s'assurer que la séquence appropriée a été obtenue. En fonction de la nature de la manipulation, la séquence désirée peut être exciser du plasmide et introduite dans un vecteur 35 différent ou le plasmide peut être restreint ou limité et le constituant de la cassette d'expression manipulé,

comme approprié. Dans certains cas, un vecteur navette sera employé lorsque le vecteur est capable de se repliquer dans des hôtes différents requérant des systèmes de répliquation différents. Cela peut ou peut ne pas exiger des marqueurs supplémentaires qui sont fonctionnels dans les deux hôtes. Quand de tels marqueurs sont exigés, ils peuvent être inclus dans le vecteur. Le plasmide contenant la cassette, deux systèmes de répliquation et le marqueur (s) peuvent être transférés d'un hôte à un autre comme requis. Pour une sélection, tout marqueur approprié peut être utilisé. De façon souhaitable, les résistances à la néomycine ou à la tétracycline sont intéressantes. Cependant, bien qu'un marqueur pour la sélection soit grandement désirable par commodité, d'autres processus pour cribler des cellules transformées ont été décrits. Voir par exemple G. Reipin et al., Current Genetics, 1982, 189-193. Les cellules transformées peuvent également être criblées par les produits spécifiques qu'elles font, par exemple la synthèse du produit désiré peut être déterminée par des méthodes immunologiques ou enzymatiques.

La cassette d'expression peut être incluse dans un système de répliquation pour le maintien épisomal dans un hôte cellulaire approprié ou peut être fournie sans système de répliquation, lorsqu'elle peut devenir intégrée dans le génome de l'hôte. L'ADN peut être introduit dans l'hôte selon des techniques connues, telles que la transformation, en utilisant de l'ADN précipité au phosphate de calcium, la transfection en mettant en contact les cellules avec le virus, la microinjection de l'ADN dans des cellules ou analogues. Un gène d'intérêt a été introduit dans l'hôte approprié, l'hôte peut être cultivé pour exprimer le gène d'intérêt.

Une variété de cellules hôtes peut être employée. Des exemples de cellules hôtes procaryotiques comprennent les organismes gram négatif tels que E. Coli,

par exemple JM109, JM101, et JM107 ; HB101, DH1 ou DH5. Des organismes gram positif sont particulièrement appropriés tels que B. subtilis qui n'ont pas d'espace périplasmique et qui peuvent sécréter directement des polypeptides dans le milieu de croissance. Les cellules hôtes eucaryotiques peuvent comprendre des cellules de levure, des cellules d'insectes et des cellules de mammifères, par exemple des cellules de COS, les cellules de CHO, les cellules de rein de singe et les cellules de vers à soie (Sf9).

La construction contenant le gène d'intérêt et des régions de chaque côté fournissant une régulation de l'expression peut être introduite dans l'hôte d'expression par tout moyen approprié, tel que transformation, avec par exemple de l'ADN précipité au phosphate de calcium, transfection, transduction, conjugaison, microinjection etc. L'hôte peut ensuite être cultivé à une grande densité dans un milieu restrictif approprié. Quand le promoteur est inductible, des conditions tolérées seront ensuite employées, telles que un changement de température, un excès ou une élimination d'un produit métabolique ou d'un nutriment, ou analogues.

Par exemple, quand la séquence régulatrice comprend un promoteur λP_L de bactériophage, l'opérateur λO_L de bactériophage et le répresseur CI857 sensible à la température, les cellules hôtes peuvent être cultivées à la température tolérée, généralement environ 30°C, à laquelle température la transcription du promoteur P_L est réprimée et les cellules hôtes peuvent croître sans être entravées par les demandes de synthèse du produit de gène étranger, qui en outre peut être toxique à l'organisme de l'hôte. Lorsque les cellules hôtes ont atteint une densité optimale, la température peut être augmentée jusqu'à une température non tolérée, par exemple environ 42°C, auquel moment le répresseur CI est rendu inactif, permettant la transcription du

promoteur P_L . Une sécrétion maximale peut être obtenue en utilisant le promoteur lac ou un promoteur trp-lac et une induction avec un inducteur métabolique tel que lactose pour une souche d'hôte lac⁺ et en fournissant lac I^q sur un vecteur. Les exemples de cellules hôtes qui pourraient être utilisées avec un tel système comprennent DH1, DH5 ou HB101.

Quand le produit est maintenu dans la cellule hôte, les cellules sont récoltées et lysées et le produit est isolé est purifié par extraction, précipitation, chromatographie, électrophorèse etc... Quand le produit est sécrété, le milieu nutritif peut être collecté et le produit isolé par des moyens conventionnels, tels que par chromatographie d'affinité. Pour produire une protéine active, il peut être nécessaire de permettre à la protéine de se replier ou ré-enrouler. Si la protéine est exprimée comme une protéine de fusion avec la séquence leader, la séquence leader peut être enlevée par traitement avec par exemple l'acide formique et du bromure de cyanogène. La séquence leader est de préférence enlevée après ré-enroulement de la protéine.

Les produits recombinants peuvent être glycosylés ou non glycosylés, ayant une glycosylation de type sauvage ou une autre glycosylation. En général, la glycosylation diffèrera de pas plus qu'environ 50% habituellement de pas plus qu'environ 20% de la glycosylation de type sauvage. La quantité de glycosylation dépendra en partie de la séquence du peptide particulier, aussi bien que de l'organisme dans lequel il est produit. Donc l'expression du produit en des cellules de E. coli aboutira à un produit non glycosylé et l'expression du produit dans des cellules d'insectes aboutira généralement à moins de glycosylation que l'expression du produit dans des cellules de mammifères.

Utilisation des facteurs de croissance

Les compositions du sujet trouvent une grande variété d'applications in vitro et in vivo comme agonistes ou antagonistes pour des facteurs de croissance, tels que le facteur de croissance épidermique (EGF), et le facteur de croissance transformant, particulièrement α -TGF. Une discussion sur les facteurs de croissance, utilisés par eux-mêmes ou en combinaison avec d'autres compositions, particulièrement des compositions de polypeptides, pour réguler la croissance des cellules, ou d'autres activités, peut être trouvée par exemple dans Handbook of Experimental Pharmacology, Tissue Growth Factors, ed. Baseraga, Vol. 57, Sprinter-Verlag, Berlin, 1981, Chapitre 3, particulièrement pages 98-109 ; et Carpenter, Ann. Rev. Biochem. (1979) 48:193-216.

L'EGF d'humain apparaît comme étant identique à l'uurogastrone humaine et exerce une variété d'effets sur la croissance prénatale et néonatale des tissus. Parmi les effets, il y a l'ouverture précoce de l'oeil, la cicatrisation des plaies, la poussée des incisives et la maturation accélérée du poumon. Des récepteurs d'EGF sont trouvés dans une large variété de tissus adultes. EGF est trouvé pour simuler la phosphorylation de son propre récepteur. EGF est également trouvé comme étant en relation avec une résorption des os augmentée.

Un facteur de croissance transformant, particulièrement α -TGF, a de nombreuses activités analogues à celles de EGF. TGF se lie au récepteur d'EGF conduisant à une phosphorylation du récepteur, à un accroissement et son activité tyrosine-kinase spécifique et à une stimulation de la croissance des cellules Cohen, dans : Biological Response Mediators and Modulators (ed. August, J.T.). Academic, New York, 1983, pp. 7-12 ; Tam et al., Nature (1984) 309:376-378; Ibbotson et al., Science (1983) 221-1292-1294.

Les virus du fibrome de Shope et du myxome induisent une augmentation significative dans le potentiel de prolifération des cellules dans des zones affectées. Tandis que la vaccine induit un effet de prolifération mineur du site infection, le virus du fibrome de Shope induit un effet de prolifération majeur conduisant à la formation de tumeur, qui chez les lapins adultes est bénigne, mais chez des adultes immunosuppressés ou chez les nouveaux nés est proliférative et invasive. Le virus du myxome induit une tumeur de myxosarcome se disséminant rapidement. Bien que ces effets de prolifération puissent être dus à d'autres facteurs viraux, les facteurs de la croissance virale apparaissent être grandement impliqués dans la prolifération cellulaire induite causée par infection virale.

Les composés du sujet ont une application particulière en tant que médicaments, comme agonistes pour EGF, et pour la cicatrisation de plaies, telles que épithélisation de plaies, telles que des brûlures, des plaies des yeux, des incisions chirurgicales, et analogues. L'ingrédient actif peut être employé dans un véhicule approprié, tel que Silvadene, dans une quantité suffisante pour favoriser l'épithélisation et/ou la cicatrisation de plaies, généralement dans des quantités s'étendant d'environ 0,01 à 10,0 $\mu\text{g/ml}$, habituellement d'environ 0,075 à 05,0 $\mu\text{g/ml}$ équivalents d'EGF, de préférence 0,5 à 5 $\mu\text{g/ml}$. La formulation est étendue sur la plaie afin de fournir un revêtement complet de la plaie avec la formulation. Les traitements peuvent être aussi fréquents que quatre fois par jour ou aussi peu fréquents que chaque jour ou moins, en fonction de la nature de la blessure, de sa réponse au traitement, et de la concentration de l'ingrédient actif et analogues.

Les compositions du sujet peuvent être utilisées comme des réactifs dans les essais de diagnostic ou pour la préparation de réactifs, tels que des anticorps monoclonaux ou polyclonaux. Comme réactifs, elles peuvent être utilisés pour la détection de facteurs de croissance analogues ou pour la détection d'anticorps aux facteurs de croissance dans des fluides physiologiques tels que le sang.

En fonction du protocole particulier et du but de réactif, le polypeptide peut être marqué ou non marqué. Une grande variété de marqueurs ont été utilisés qui fournissent, directement ou indirectement, un signal détectable. Ces marqueurs comprennent les radionuclides, les enzymes, les produits fluorescents, les particules, les produits chemiluminescents, les substrats ou cofacteurs d'enzymes, les inhibiteurs d'enzymes, les particules magnétiques etc.... Voir par exemple Brevets US Nos. 3 654 090, 3 817 837, 3 935 074, 3 996 345, 4 277,437, 4 374 925 et 4 366 241.

Une grande variété de procédés existe pour lier les marqueurs aux polypeptides, qui peuvent impliquer une utilisation du groupement amino N-terminal pour une fonctionnalisation pour former un pyrazolone, tandis que d'autres groupements amino libres sont protégés, où le pyrazolone peut être ensuite mis en contact avec divers réactifs, tels que les groupements amino, ou se lier à la partie générant un signal détectable. En protégeant les acides aminés d'arginine associés avec la troisième boucle ou proximale, d'autres arginines peuvent être mises en action pour une conjugaison aux groupements amino ou aux groupements thio selon des voies connues. Alternativement, le polypeptide peut être mis en contact avec un agent actif, tel qu'un acide carboxylique activé et substitué au hasard, où le matériau actif biologiquement peut être séparé du matériau biologiquement inactif comme résultat de la substitution au hasard. Finalement,

selon le procédé de synthèse, le polypeptide peut être modifié pour fournir la fonctionnalité désirée comme partie du processus de synthèse.

Les compositions du sujet peuvent également être utilisées pour contrôler les récepteurs d'EGF. Les compositions du sujet peuvent également être utilisées pour contrôler une réponse cellulaire à EGF et/ou TGF en fournissant une compétition entre ces matériaux se produisant naturellement et une composition selon l'invention du sujet. Dans cette voie, des changements dans la conformation du récepteur peuvent être contrôlés.

Les compositions du sujet peuvent également être utilisées pour divers buts thérapeutiques impliquant une stimulation ou un contrôle de la croissance de la formation des os. Ces composés peuvent être administrés dans des véhicules physiologiques appropriés de façon intrapéritonéale, subcutanée, intraveineuse, intra-artérielle, ou par application au site d'intérêt. De plus, les compositions du sujet peuvent être introduites dans des liposomes, qui peuvent ou non impliquer l'utilisation d'anticorps pour l'administration en site. Les divers véhicules comprennent une solution saline tamponnée en phosphate, une solution saline, l'eau et analogues. La concentration de l'additif variera largement en fonction de son utilisation finale et de son activité.

D'autres additifs peuvent également être inclus dans les formulations, tels que EGF, TGF, d'autres facteurs de croissance, des bactériocides, tels que des antibiotiques, des bactériostatiques, des tampons etc...

Pour préparer des anticorps, les polypeptides du sujet où PP¹⁻⁴ sont hydrogènes, ou des chaînes d'oligopeptides courtes (moins de cinq acides aminés), peuvent être joints aux polypeptides ou protéines antigéniques, pour une injection dans les hôtes mammifères. La protéine antigénique aura au moins environ

60 acides aminés et sera habituellement pas supérieure à 100 kilodaltons (kDal). Des techniques nombreuses existent pour union aux polypeptides, soient à un site spécifique ou au hasard, en utilisant des réactifs

5 bifonctionnels, tels que l'acide p-maléimidobenzoïque, le glutaraldéhyde, la p,p'-benzidine, etc... . Les protéines antigéniques communes comprennent l'albumine de sérum de bovins, la keyhole limpet hémocyanine, le toxoïde du tétanos, etc... . Les polypeptides du sujet sont joints

10 aux protéines antigéniques en nombre suffisant pour fournir la réponse immunogénique désirée. Habituellement il y aura deux ou plusieurs injections de relance après l'injection initiale. Pour des antisérums, le sang est enlevé de l'hôte immunisé et l'infraction d'immunoglobulines est isolée. Pour des anticorps monoclonaux, la

15 rate est isolée et les splénocytes sont fusionnés avec un partenaire de fusion approprié selon des voies classiques. Les hybridomes résultants sont ensuite criblés pour des anticorps se lient aux sites épitopiques

20 du polypeptide du sujet. Ces anticorps peuvent être utilisés dans une grande variété de buts, tels que des réactifs de diagnostic, en thérapie, etc... . Les anticorps quand ils sont utilisés comme réactifs peuvent être marqués ou non marqués comme cela fut décrit pour

25 les polypeptides.

Les exemples qui suivent sont donnés à titre d'illustration de la présente invention et ne constituent nullement une limitation.

EXPERIMENTATIONTableau des Contenus

<u>Exemple I :</u>	<u>Préparation de Gènes de synthèse d'intérêt.</u>
5	A. Oligonucléotides de synthèse de TGF B. Oligonucléotides de synthèse de VGF C. Oligonucléotides de synthèse d'EGF
<u>Exemple II:</u>	<u>Description de Clonage et Plasmides d'Expression</u>
10	A. Plasmide pLEBam B. Plasmide pBM11 C. Plasmide pBM11M4 D. Plasmide pBM11M5 E. Plasmide pBM11/NDP
15	F. Plasmide pBM11/PAD G. Plasmide pBM11/PAK H. Plasmides pTCPt
	I. Plasmide plac/cro- β gal J. Plasmide ptac/cro- β gal
20	K. Plasmide pRSV L. Plasmide pAc610 M. Plasmides pTox1 et pTox2 N. TakPak
<u>Exemple III:</u>	<u>Assemblage de gènes du facteur de Croissance pour une Expression dans des cellules Procaryotiques</u>
25	A. Préparation du plasmide pLEBam/TVV B. Préparation de VGF par plasmide pLEBam/TVV et ses fragments

Exemple IV: Expression des Polypeptides d'Intérêt
 comme une Protéine de fusion avec la
 Protéine-N

- 5 A. TGV de Synthèse Modifié
 B. Hybride TGF-VGF de Synthèse Modifié

Exemple V: Préparation du Polypeptide d'Intérêt
 comme une Protéine de Fusion avec la
 Protéine-N et un Site de Clivage

- 10 A. VGF de Synthèse Modifié
 B. Hybrides TGF-VGF de Synthèse Modifié
 C. EGF de synthèse

Exemple VI: Préparation du Polypeptide d'Intérêt
 comme une Protéine de Fusion avec la
 Séquence Phosphatase Alcaline modifiée

- 5 A. Préparation de pBM11/PAD/EGF
 B. Préparation de pBM11/PAD/nVGfα

Exemple VII: Expression du Polypeptide d'Intérêt
 comme une Protéine de Fusion avec un
 Signal Phosphate Alcaline

- 10 A. Préparation de pBM11/PAK/nVGfα
 (séquence signal phosphatase
 alcaline/nVGfα avec N-terminal de VGF
 naturel et séquences GTC et
 GYACRC)
- 15 B. Préparation de pBM11/PAK/EGF
 C. Préparation de TacPak/EGF (séquence
 signal phosphatase alcaline/
 EGF d'humain)

20 Exemple VIII: Expression d'un Polypeptide d'Intérêt
 comme une Protéine de Fusion avec la
 Séquence Signal Phosphatase Alcaline
 en Utilisant une Cassette d'Expression
 Comprenant une Région de Terminaison de
 transcription

- 25 A. Préparation de pTCPt/EGF ([trp-35]
 16bp [lac-10] [lacSD] 11bp [ATG]
 signal phosphatase alcaline/
 EGF d'humain/term. trans-NEO

5 B. Préparation de pTCPt/nVGfA ([trp-35] 16bp [lac-10][lacSD][croSD] 11bp [ATG]/signal phosphatase alcaline/VGfA N-terminal avec séquence GTC et GYACRC)/term.trans.-NEO)

C. Préparation de pTNPt/EGF ([trp-35] 17bp [lac-10][nSD] 8bp [ATG]/signal phosphatase alcaline/EGF d'humain/term. trans. -NEO)

10 Exemple IX : Assemblage des Gènes de Facteur de Croissance pour Expression dans des Cellules Eucaryotiques

15 A. Préparation du plasmide PRSV/VGF
B. Préparation du plasmide pAc/VGF
C. Préparation de pAc/SFGF

Exemple X: Préparation de Polypeptides

20 A. Synthèse en phase solide de VGF et TGF
B. Isolement de Polypeptides Recombinants Préparés dans des Cellules Procaryotiques
C. Isolement de VGF et SFGF Produit par Expression Eucaryotique des Gènes Naturels
D. EGF Naturel

25 Exemple XI : Essais d'Activité

- 5
- A. Essai Mitogène
 - B. Essai de Stimulation de Croissance de Colonies sur Agar mou
 - C. Essai d'Inhibition de Liaison au Récepteur d'EGF
 - D. Radioimmuno essai
 - E. Cicatrisation de Plaies

10

Exemple XII : Activité Biologique de Facteurs de Croissance Recombinant Préparés dans des Cellules Procaryotiques

- A. Liaison au Récepteur d'EGF
- B. Activité Mitogène
- C. Cicatrisation de Plaies

15

Exemple XIII : Activité Biologique de VGF Préparé dans des Cellules Eucaryotiques

- 20
- A. VGF Préparé dans des Cellules de Reins de Singe (BSC-1)
 - B. VGF préparé dans des Cellules de CHO
 - C. VGF Préparé dans le Ver à Soie
 - D. Comparaison Immunologique de VGF et de TGF.

Dépôts biologiques.

25

Les plasmides d'expression suivants, tous transformés dans E. Coli HB101, ont été déposés à la date indiquée à l'American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD 20852 et ont leur identification, leurs numéros de dépôt à l'ATCC donnés ci-dessous :

	<u>Identification</u>	<u>Numéros de dépôt à l'ATCC</u>	<u>Date de dépôt</u>
	PBM11	67366	
	PBM14	67367	
5	PBM11/PA/VGF	67417	3 Juin 1987
	PBM11/DP/VGFa	67418	3 Juin 1987
	PBM11/PA/EGF	67419	3 Juin 1987
	PMM11/M5	67436	
	PMM11/C2	67437	
10	PBM11/NDP/EGF	67547	23 Octobre 1987

Procédés

Des techniques de clonage générales ont été utilisées comme décrites dans Maniatis et al, 1982, "Molecular Cloning : A Laboratory Manual", Cold Spring Harbor Laboratory, CSH, New York. Toutes les enzymes modifiant l'ADN ont été obtenues chez les fournisseurs du commerce. Elles ont été utilisées selon les instructions du fabricant. Les matières et l'appareil pour la purification et la séparation de l'ADN ont été utilisés conformément aux instructions du fournisseur.

Exemple 1.Préparation de Gènes de Synthèse d'Intérêt.

Des gènes de synthèse de facteurs de croissance ont été représentés qui utilisent des codons de cellules hôtes optimisés pour des niveaux élevés d'expression. De plus, plusieurs sites de restriction convenables ont été représentés dans les gènes de synthèse. Quand cela est possible, les nouveaux sites de restriction laissent la séquence d'acides aminés du gène de facteur de croissance non altérée, cependant, dans certains cas l'incorporation de nouveaux sites de restriction fournissent une séquence d'acides aminés altérée. Ces

sites divisent grossièrement les gènes de synthèse en fournissant trois domaines N-terminal, milieu et C-terminal.

Le produit du gène de VGF naturel contient un domaine N-terminal extrême qui n'a pas de contre partie dans TGF mature. Les fragments de VGF où ce domaine manque, sont référencés comme tronqués. Les sites de restriction ont été utilisés pour la construction initiale des gènes définitifs à partir de fragments partiels d'oligonucléotides de synthèse s'étendant d'un site de restriction à l'autre. Les oligonucléotides ont été synthétisés sur un synthétiseur d'oligonucléotides d'Applied Biosystems et ont été purifiés sur un gel d'acrylamide. Les oligonucléotides ont été phosphorylés à l'extrémité 5' en utilisant une polynucléotide kinase de T4 et chaque oligonucléotide a ensuite été cyclisé à son complément.

A. Oligonucléotides de synthèse de TGF

1. Domaine N-terminal de TGF d'humain

20 TGF ->

BssHIINcoI

M V V S H F N D C P D S H T Q F C
5' CGCGCCATGGTTGTTTCTCACTTTAACGACTGCCCGGACTCTCATACTCAGTTTTC
3' GGTACCAACAAAGAGTGAAATTGCTGACGGGCCTGAGAGTATGAGTCAAAACG

25

KpnI

F H G T

TTTCATGGTAC 3' TGF104
AAAGTAC 5' TGF103

2. Domaine du milieu de TGF d'humain modifié avec la séquence humaine QEDK qui est altérée en QEEK, la séquence trouvée dans TGF de rat

30

KpnI

SphI

C R F L V Q E E K P A C
5' CTGCCGTTTTCTGGTTCAGGAAGAAAAACCGGCATG 3' TGF101
35 3' CATGGACGGCAAAAGACCAAGTCCTTCTTTTGGCC 5' TGF102

3. Domaine C-terminal de TGF d'humain :

SphI
 V C H S G Y V G A R C E H A D L L
 5' CGTTTGCCATTCTGGCTACGTTGGCGCACGTTGCGAACACGCTGACCTGCTG
 5 3' GTACGCAAACGGTAAGACCGATGCAACCGCGTGCAACGCTTGTGCGACTGGACGAC

BamHI
 A Ter
 GCTTAAG 3' TGF205
 CGAATTCCTAG 5' TGF206

10 B. Oligonucléotides de synthèse de VGF1. Domaine N-terminal extrême de VGF

HindIII
 E D S G N A I E T T S P E I T N A T T
 5' AGCTGACTCTGGTAACGCTATCGAACTACTTCTCCGGAAATCACTAACGCTACTACT 3' \
 15 3' CTGAGACCATTGCGATAGCTTTGATGAAGAGGCCTTTAGTGATTGCGATGATGA 5' \

2. Domaine N-terminal de VGF modifié comprenant
un site de clivage Asp-Pro, avec la séquence
HGT remplaçant la séquence naturelle HGD :

BamHI
 I D P M D I P A I R L C G P E G D G Y
 5' GATCGATCCCATGGACATCCCGGCTATCCGTCTGTGCGGCCGGAAGGCGACGGCTAC
 20 3' CTAGGGTACCTGTAGGGCCGATAGGCAGACACGCCGGGCCTTCCGCTGCCGATG

KpnI
 C L H G T
 25 TGCCTGCATGGTAC 3' VGF104a
 ACGGACGTAC 5' VGF103a

5' KpnI T C I H A R D I D G Y A SphI C 3' .VGf101a
3' CATGGACGTAGGTACGTGCACTGTAGCTGCCGATGC. 5' .VGf102a

SphI EcoRI

C R C S H G Y T G

10 5' CCGTTGCTCTCATGGCTACACTGG 3' VGF1A
 3' GTACGGCAACGAGAGTACCGATGTGACCTAA 5' VGF2A

15 SphI EcoRI

 V C S H G Y T G

5' CGTTTGCTCTCATGGCTACACTGG 3' VGF1

3' GTACGCAAACGAGAGTACCGATGTGACCTTA 5' VGF2

EcoRI
BamHI

	I	R	C	Q	H	V	V	L	V	D	Y	Q	R	Ter	
5'	AATTCGTTGCCAGCATGTTGTTCTGGTCGACTACCAGCGTTAAGGATC													3'	VGF3
3'	GCAACGGTCGTACAACAAGACCAGCTGATGGTCGCAATTC													5'	VGF4

C. Oligonucléotides de synthèse d'EGF

Trois ensembles d'oligonucléotides de synthèse se chevauchant 1(A,B), 2(A,B) et 3(A,B) codant pour EGF humain ont été synthétisés sur un synthétiseur

5 d'oligonucléotides Applied Biosystems et purifiés sur un gel d'acrylamide. Les oligonucléotides ont été phosphorylés à l'extrémité 5' en utilisant une polynucléotide kinase de T4. Chaque oligonucléotide a été cyclisé à son complément.

10 1.

NcoIEcoRI
M N S D S E C P L S H D G Y
5' CATGAATTCTGACTCTGAATGCCCGCTGTCTCATGACGGCTAC 3' EGF1A
3' TTAAGACTGAGACTTACGGGCGACAGAGTACTGCCGATGACGGAC 5' EGF2A

15 2.

C L H D G V NsiI C M Y I E A L D K Y A SphI C

5' TGCCTGCATGACGGCGTATGCATGTACATCGAAGCTCTGGACAAGTACGCATG 3' EGF1B
3' GTACTGCCGCATACGTACATGTAGCTTCGAGACCTGTTTCATGC 5' EGF2B

20 3.

SphI
N C V V G Y I G E R C Q Y R D L K
5' CAACTGCGTTGTTGGCTACATCGGCGAACGTTGCCAGTACCGTGACCTGAAA
3' GTACGTTGACGCAACAACCGATGTAGCCGCTTGCAACGGTCATGGCACTGGACTTT

25

BamHI
W W E L R *
TGGTGGGAACTGCGTTAAG 3' EGF3
ACCACCCTTGACGCAATTCCTAG 5' EGF4

Exemple IIDescription de clonage et plasmides
d'expression

5 A. Un plasmide pLEBam a été utilisé pour cloner
des fragments d'oligonucléotides de synthèse à cause de
ses sites de restriction appropriés BssHII et BamHI. Un
plasmide avec des sites de restriction NcoI et BamHI tel
que pBM11 ou pBM11/NDDP (décrit ci-dessous) peut être
utilisé pour cloner les fragments de nucléotides de
10 synthèse et autres séquences d'ADN d'intérêt.

B. Le plasmide pBM11, décrit dans la demande de
brevet US copendante No. 264.098, déposée le 28 Octobre
1988 permet le clonage d'un gène d'intérêt en aval des
séquences d'ADN codant pour les 33 acides aminées de
15 N-terminal du gène-N du bactériophage λ à un site de
restriction BamHI. Sur induction du promoteur λ PL par
inactivation du répresseur C1857 sensible à la
température à 42°C, le produit de gène étranger est
exprimé comme la partie C-terminale d'une protéine de
20 fusion dont la séquence N-terminal est celle du gène-N.

C. Le plasmide pBM11M4, décrit dans la demande
de brevet US copendante No. 264.098, déposée le 28
Octobre 1988, est dérivé de pBM11 et permet à un gène
d'intérêt d'être cloné à un site de restriction BamHI
25 directement après la méthionine d'initiation du gène-N.
Le plasmide pBM11/M4 contient également un site NcoI dans
le gène de la néomycine.

D. Le plasmide pBM11M5, décrit dans la demande
de brevet US copendante No. 264.098, déposée le 28
30 Octobre 1988, est dérivé de pBM11 dans lequel un site
NcoI présent dans le gène de résistance à la néomycine a

été enlevé par une mutagénèse dirigée vers le site. Le clonage d'un gène d'intérêt dans pBM11/M5 n'exige donc pas de digestion partielle du vecteur avec NcoI.

5 E. Le plasmide pBM11/NDP, décrit dans la demande de brevet US copendante No. 264.098, déposée le 28 Octobre 1988, est dérivé de pBM11 et à des séquences d'ADN codant pour un dipeptide acide aspartique-proline instable aux acides inséré entre les séquences codant pour le gène-N et un gène d'intérêt. Le plasmide contient
10 un site NcoI et un site ClaI pour cloner un gène d'intérêt en aval promoteur P_L.

F. Le plasmide pBM11/PAD, décrit dans la demande de brevet US copendante No. 264.098, déposée le 28 Octobre 1988, est dérivé du plasmide pBM11M4 et permet
15 à un gène d'intérêt d'être cloné à HindIII, SmaI ou BamHI en aval d'une séquence signal phosphatase alcaline modifiée.

G. Le plasmide pBM11/PAK, décrit dans la demande de brevet US copendante No. 264.098, déposée le 28 Octobre 1988, est dérivé du plasmide pBM11M4 et permet
20 à un gène d'intérêt d'être cloné à HindIII, SmaI ou BamHI en aval de la séquence signal phosphatase alcaline.

H. Le plasmide pTCPt, est désigné comme ayant les éléments du promoteur tac et utilise le cro SD pour
25 exprimer le gène d'intérêt derrière la séquence signal phosphatase alcaline. Un exemple de la construction de ce plasmide est donné ci-dessous dans la construction de pTCPt/EGF.

30 Construction de pTNPt ([trp-35]17bp[lac-10]
[nSD] 8bp[ATG]/signal phosphatase alcaline/linker/term.
trans.-NEO)

Ce plasmide est désigné comme ayant des éléments du promoteur tac et utilise le SD du gène-N pour exprimer un gène donné derrière la séquence signal

phosphatase alcaline. Il a un fond de pBR22 avec le gène de résistance à la néomycine. Un plasmide pTNPt a été construit comme suit :

5 (a) Préparation du fragment de 2,8kb par
EcoRI-BamHI de pBM16t/VGFa sans le site HindIII.

Un plasmide pBM16/VGFa a été digéré avec EcoRI et BamHI et le fragment de 2,8kb a été isolé. Le fragment de 2,8kb a été lié à un liant ou linker EcoRI-BamHI et une construction correcte a été isolée par analyse de
10 restriction et est référencée comme Intermédiaire I.

Le site unique HindIII proche du gène de résistance de la néomycine a été enlevé du plasmide intermédiaire I par digestion avec HindIII, en créant des bouts francs en utilisant un fragment de Klenow, et en
15 reliant. Cela conduit à un plasmide Intermédiaire II qui est dépourvu du site HindIII.

Le fragment de 2,8kb obtenu par EcoRI-BamHI de pBM16t/VGFa sans le site HindIII a été isolé en digérant p.Intermédiaire II avec EcoRI et BamHI. Le fragment de 2,8kb
20 résultant a été isolé par une électrophorèse sur gel d'agarose.

(b) Préparation du fragment de 150bp par
BamHI-BsmI de pBM11/PAK

Le plasmide pBM11/PAK est identique à
25 pBM11/PAK/EGF sauf qu'il contient une région de liaison ou linker avec les sites HindIII, SmaI et BamHI en aval de la séquence signal phosphatase alcaline au lieu du gène EGF. pBM11/PAK a été digéré avec BsmI et BamHI et le fragment de 150bp contenant le SD du gène-N, la séquence
30 signal phosphatase alcaline et la région linker a été isolée.

(c) Préparation des olifonucléotides TacA+ et
TacA-

Les oligonucléotides TacA+ et TacA- ont été
35 synthétisés sur un synthétiseur d'oligonucléotides Applied Biosystems et ont été désignés comme ayant la

partie en saillie EcoRI à l'extrémité 5' avec la séquence consensus trp-35 séparée de la séquence consensus lac-10 par 17 nucléotides dans lesquels était positionné un site SstI. La séquence contenait également l'extrémité 5' du ARNm, le site de liaison ou répresseur lac et une partie en saillie BsmI.

TacA+ 5' AATTACTCCCCATCCCCCTGTTGACAATTAATCATCGAGCTC
GTATAATGTGTGGAATTGTGAGCGGATAACAATTTACACAG 3'

10 TacA- 5' GTGTGAAATTGTTATCCGCTCACAATTCCACACATTATA
CGAGCTCGATGATTAATTGTCAACAGGGGGATGGGGAGT 3'

(d) Ligation et isolement de pTNPt

Le fragment de 2,8kb obtenu par EcoRI-BamHI, le fragment de 150bp obtenu par BsmI-BamHI et les oligonucléotides TacA+ et TacA- ont été liés ensemble en utilisant une ADN ligase et l'ADN a été utilisé pour transformer JM109 (lacIq) compétente. Une construction correcte a été isolée par analyse de restriction et séquençage d'ADN.

(Site EcoRI de pBM11)
|
GAATTACTCCCCATCC
20 trp-35 (17bp) SstI lac-10
CCCTG [TTGACA] ATTAATCATCGAGCTCG (TATAATG)
5' lac ARNm -> BsmI n ARNm ->
25 TGTGG/AATTGTGTGAGCGGATAACAATTTACACAGCATTCAAAGCAGAAGGCT
TTGGGGTGTGTGATACGAAACGAAGCATTGGCCGTAAGTCCGATTCCGGATTAGC
TGCCAATGTGCCAATCGCGGGGGTTCGTTTCAGGACTACAACCTGCCACACACC
30 ACCAAAGCTAACTGAC {AGGA} MSD (8bp) Signal Sequence -> PvuI
M K Q S T I A L
TCGCACTTCTCCCACTGCTGTTCACTCCAGTGACAAAAGCTTCCCGGGATCCGCTG
35 A L L P L L F T P V T K HindIII SmaI BamHI

Term. Trans. site (BamHI of de pBMII)
ACTAATTGGGGACCCCTAGAGGTCCCCTTTTTTATTTTAAAACGATCC

I. Le plasmide plac/cro- β gal comprend la région opérateur-promoteur de l'opéron lactose (lac) de E. Coli, aussi bien que les sites de liaison aux ribosomes de lac et cro. Le plasmide a été construit
5 comme décrit dans la demande US copendante No. 264,098 déposée le 28 Octobre 1988. Les protéines de fusion exprimées par ce vecteur comprennent le N-terminal de la protéine Cro du bactériophage λ , la séquence d'acides aminés codée par l'ADN inséré, et le C-terminal de
10 β -galactosidase. Les éléments de contrôle de ce vecteur comprennent la région opérateur-promoteur de l'opéron (lac) lactose de E. Coli, aussi bien que les sites de liaison aux ribosomes de lac et cro.

J. Le plasmide ptac/cro- β gal permet à un gène
15 d'intérêt d'être cloné en aval des 21 acides aminés de N-terminal de la protéine de Cro bactérienne. Le plasmide a été construit comme décrit dans la demande US copendante No. 264 098 déposée le 28 Octobre 1988. Le vecteur d'expression ptac/cro- β gal est similaire à
20 ptac/cro- β gal, sauf que le promoteur de ptac/cro- β gal comprend la région -35 du promoteur de l'opéron tryptophane et la "Pribnow box" de l'opéron lac. Le promoteur hybride permet un taux plus élevé d'expression que plac/cro- β gal.

25 K. Le plasmide pRSV permet l'expression d'un gène d'intérêt dans des cellules eucaryotiques avec le promoteur RSV LTR, Gorman et al., Proc. Natl. Acad. Sci. USA (1982) 79:6777-6781.

30 L. Le plasmide pAc610, Smith et al., Molec. Cell. Biol. (1983) 12: 2156-2165, permet l'expression d'un gène étranger dans les cellules d'insectes.

M. Les plasmides pTox1 et pTox2 permettent l'expression d'un gène d'intérêt dans des cellules de levure en utilisant la séquence signal toxine tueuse de
35 X.lactis (EMBO 6, 229-234 (1987) ; Biochem, Biophys. Res. Comm. 144, 613-619 (1987)). Cette construction a le gène

URA3 pour une sélection dans la levure et le gène AMP pour une sélection dans les bactéries. La transcription s'initie avec le promoteur CYC1 et "la séquence d'activation en aval" GAL (Methods Enzymol., 101, 181-191 (1983) et se termine dans l'extrémité 3' du gène FLP (Mol. Cell. Biol., 5, 2770-2780 (1985). La construction a à la fois l'origine de 2 microns et l'origine de pBR322 pour la replication chez la levure et chez E. Coli, respectivement. La région N-terminale de toxine tueuse avec la séquence signal et le site de clivage Kex2 a été introduite sur les oligonucléotides et contenait la région riche en A-T précédant immédiatement le codon initiateur pour maintenir les signaux optimum d'initiation de la traduction et de l'élongation. Plusieurs sites de restriction ont été obtenus par génie génétique pour permettre l'insertion des gènes d'intérêt en aval à la fois du site de clivage signal du site de clivage Kex2.

1. Préparation du fragment de 3kb par EcoRI-EspI de pBR322 contenant l'origine et le gène de résistance à l'AMP.

Le plasmide pLGSD5(-ATG), un vecteur d'expression de levure (Methods Enzymol. 101:181-191 (1983) a été digéré avec EcoRI et EspI et le fragment de 3,0kb a été isolé.

2. Préparation du fragment de 2,1 kb par EcoRI-EspI du plasmide 2 microns contenant l'extrémité 3' du gène FLP et l'origine.

Le plasmide pLGSD5(-ATG) a été digéré avec EcoRI et EspI et le fragment de 2,1 kb a été isolé.

3. Ligation et isolement du plasmide intermédiaire I.

Les fragments de 2,1 kb et 3 kb obtenus par EcoRI-EspI ont été liés ensemble en utilisant une ADN ligase et l'ADN a été utilisé pour transformer HB101 compétente. Une construction correcte a été isolée par analyse de restriction.

4. Préparation du fragment de 1,8 kb par EcoRI-BamHI de pLGSD5(-ATG) contenant le gène URA3, le GAL UAS et le promoteur CYC1.

Le plasmide pLGSD5(-ATG) a été digéré avec EcoRI et BamHI et le fragment de 1,8 kb a été isolé. Le site BamHI est présent dans un "linker" ou liant inséré en aval de 4 bp de l'initiateur CYC1.

5. Préparation des oligonucléotides pTox1A+, 2A-, 1B+, 2B-.

Les oligonucléotides ont été désignés pour lier les codons pour la séquence signal de la toxine tueuse K. lactis au site BamHI en aval de l'initiateur CYC1. Pour conserver une initiation optimale de la traduction, 13bp de la séquence riche en A-T en amont du codon initiateur de la toxine tueuse ont été inclus. Cette séquence se conforme à la séquence consensus déterminée pour les codons d'initiation de la levure. Un site BglII a été placé directement en amont de la région A-T pour permettre une mutagénèse de la région d'initiation et de la séquence signal. Un site HindIII a été placé 10 nucléotides en amont du site de clivage signal pour permettre une mutagénèse des séquences en aval et une ligation du gène d'intérêt. Dans pTox1, un site AvaI, XhoI a été localisé directement en aval du site de clivage signal, tandis que dans pTox2, un site NaeI a été placé au site de clivage signal pour cloner des bouts francs directement sur le signal de clivage qui a été changé de Gly-Leu à Ala-Gly. La détermination de l'hétéroduplex au site de clivage pourrait aboutir à une

génération de clones contenant la séquence de pTox1 et celle contenant la séquence de pTox2. Dans les deux constructions, les séquences ultérieures codent pour le reste de la séquence précurseur de la toxine tueuse

5 jusqu'au site de clivage Lys-Arg de Kex2, auquel point plusieurs sites de restriction (StuI, SalI, AccI, HincII, et BamHI) ont été placés pour faciliter le clonage d'un gène d'intérêt. Le site de StuI permet d'obtenir un

10 clonage en bouts francs directement après le codon Arg du site de clivage de Kex2. Les oligonucléotides ont une partie qui dépasse ou en saillie BamHI à l'extrémité 5' et une partie qui dépasse ou en saillie HindIII à l'extrémité 3' pour permettre de cloner dans le vecteur intermédiaire. La séquence résultante pas plus longue a

15 l'un des deux sites.

(BamHI) BglII

HindIII

M N I F Y I F L F L L S

1A+ 5'GATCAGATCTAATAATTATAAAATGAATATATTTTACATATTTTGTGTTTGTGCTAAGC

2A- 3' TCTAGATTATTAAATATTTTACTTATATAAAATGTATAAAAACAAAACGATTCGAAG

Site de clivage

Site de clivage Kex2

(pTox1) / Aval, XhoI

/ StuI SalI, AccI BamH

F V Q G L E E T E R R G S L V K R P L S T D P

1B+ 5'TTCGTTTCAGGGCCTCGAGCATACTCATAGAAGAGGCTCCTTAGTCAAAGGCCTTTGTGACGGATCC

2B- 3' AAGTCCGGCCGCTCGTATGAGTATCTTCTCCGAGGAATCAGTTTCCGGAACAGCTGCCTAGGTCC

A G

(pTox2) NaeI

Ces oligonucléotides ont été synthétisés par un synthétiseur d'Applied Biosystems et ont été purifiés par électrophorèse sur gel. Postérieurement, ils ont été phosphorylés en utilisant une kinase de T4 et les paires complémentaires ont été appariées ensemble, liées et purifiées sur gel.

6. Préparation du vecteur Intermédiaire I par EcoRI-HindIII.

Le vecteur Intermédiaire I a été digéré avec EcoRI et HindIII et a ensuite été traité avec la phosphatase alcaline de veau. Ce traitement enlevait un petit fragment obtenu par EcoRI-HindIII et laissait un site HindIII en amont de la séquence 3' de FLP.

7. Ligation et isolement de pTox1 et pTox2.

Le fragment de 1,8kb obtenu par EcoRI-BamHI a été lié au fragment oligonucléotides en utilisant une ADN ligase et le fragment résultant a été purifié sur gel et a été postérieurement lié au vecteur Intermédiaire I obtenu par EcoRI-HindIII et l'ADN a été utilisé pour transformer HB101 compétente. Les constructions correctes ont été identifiées en utilisant une analyse de restriction et en séquençant l'ADN et un clone ayant la séquence de pTox1 et un ayant la séquence de pTox2 ont été isolés. Les deux clones étaient dépourvus du site BamHI à la jonction des oligonucléotides du vecteur Intermédiaire I.

Exemple III.

Assemblage des gènes de facteur de croissance pour une expression dans des cellules procaryotiques.

A. Préparation du plasmide pLEBam/TTV

Le facteur de croissance chimérique de synthèse, noté TTV ou (TGF/TGF/VGF) a été assemblé dans le vecteur de clonage pLEBam. Le facteur de croissance hybride contenait la séquence d'acides aminés du TGF humain dans les deux tiers amino terminal du gène à l'exception de la séquence QEEK qui était altérée à

15 TGF +
 NcoI
 CCATGGTTGTTTCTCACTTTAACGACTGCCCCGACTCTCATACTCAGTTTTGCTT
 M V V S H F N D C P D S H T Q F C F

20 KpnI VGf +
 SphI
- TCATGGGTACCTGCCGTTTTCTGGTTCAGGAAGAAAAACCGGCATGCGTTTGCTCT
H G T C R F L V Q E E K P A C V C S

25 CATGGCTACACTGGAATTCGTTGCCAGCATGTTGTTCTGGTCGACTACCAGCGT
H G Y T G I R C Q H V V L V D Y Q R

BamHI
TAAGGATCC
Ter

B. Préparation du plasmide pLEBam/TVV.

Le facteur de croissance chimérique de synthèse, noté TVV (TGF/VGF/VGF) a été assemblé dans le vecteur de clonage pLEBam. Le facteur de croissance hybride contenait la séquence d'acides aminés de TGF humain dans le domaine N-terminal du gène. Les domaines du milieu et C-terminal sont dérivés de la séquence non tronquée de VGF et de l'extrémité avec la séquence YQR en amont de la séquence naturelle PNT. De plus, le gène de synthèse présente la modification, GYACVC pour GMYCRC.

(a) Préparation d'un fragment de 4,3 kb par KpnI-SphI de pLEBam/TTV.

Le plasmide pLEBam/TTV a été digéré avec KpnI et SphI et le fragment de 4,3 kb obtenu par KpnI-SphI a été purifié sur gel. La digestion enlève le domaine du milieu de TGF du gène de synthèse TTV dans le plasmide de clonage pLEBam.

(b) Ligation et isolement de pLEBam/TVV.

Des oligonucléotides de VGF101a et 102a sont liés ou ligaturés au fragment de 4,3 kb obtenu par KpnI-SphI de pLEBam/TTV en utilisant une ADN ligase et le mélange résultant a été utilisé pour transformer HB101 compétente. Les transformants ont été sélectionnés sur l'ampicilline et ont été criblés par séquençage des nucléotides en utilisant la méthode didésoxy de Sanger. Une construction correcte a été isolée et notée pLEBam/TVV.

TGF +

NcoI

CCATGGTTGTTTCTCACTTTAACGACTGCCCGGACTCTCATACTCAGTTTGGCTT
M V V S H F N D C P D S H T Q F C F

VGF +

KpnI

SphI

TCATGGTACCTGCATCCATGCACGTGACATCGACGGCTACGCATGCGTTTGCTCT
H G T C I H A R D I D G Y A C V C S

EcoRI

CATGGCTACACTGGAATTCGTTGCCAGCATGTTGTTCTGCTCGACTACCAGCGT
H G Y T G I R C Q H V V L V D Y Q R

BamHI

TAAGGATCC

Ter

Exemple IV.

Expression du polypeptide d'intérêt comme une protéine de fusion avec la protéine -N.

A. TGF de synthèse modifié5 1. Préparation de pBM11/N/TGF

Le TGF humain modifié a été exprimé dans ce système comme partie d'une fusion avec les 33 acides aminés N-terminaux du gène-N et la séquence QEEK remplaçant la séquence humaine QEDK.

10 (a) Préparation d'un fragment de 780 bp par SphI-PvuI de pBM11/N/TTV

Le plasmide pBM11/N/TTV a été digéré avec SphI et PvuI et le fragment de 780 bp obtenu par SphI-PvuI a été purifié sur gel. Ce fragment contient une partie du plasmide pBM11 à l'extrémité du PvuI et à l'extrémité SphI, le gène-N et les deux tiers N-terminaux du gène de TGF humain.

15 (b) Préparation du fragment de 5kb par BamHI-PvuI de pBM11/N/TTV.

20 Le plasmide pBM11/N/TTV a été digéré avec BamHI et PvuI et le fragment de 5kb obtenu par BamHI-PvuI a été purifié sur gel.

25 (c) Ligation et isolement de pBM11/N/TGF:

Les oligonucléotides de TGF 205 et 206, le fragment de 780 bp obtenu par SphI-PvuI et le fragment de 5 kb obtenu par BamHI-PvuI de pBM11/N/TTV ont été liés ou ligaturés ensemble et utilisés pour transformer HB101 compétente. Les transformants ont été sélectionnés sur la néomycine et ont été criblés par l'analyse de restriction en utilisant EcoRI et en séquençant les nucléotides 30 suivant la méthode didésoxy de Sanger. Une construction correcte a été isolée et notée pBM11/N/TGF.

Gène-N →

ATGGATGCACAAACACGCCGCCGCGAACGTCGCGCAGAGAAACAGGCTCAATGGA
M D A Q T R R R E R R A E K Q A Q W K

AAGCAGCAAATCCCCTGTTGGTTGGGGTAAGCGCAAACAGTTCCGGATCCGCAT
A A N P L L V G V S A K P V R I R M

TGF -
GGTTGTTTCTCACTTTAACGACTGCCCCGACTCTCATACTCAGTTTGGCTTTCAT
V V S H F N D C P D S H T Q F C F H

KpnI SphI
GGTACCTGCCGTTTTCTGGTTCAGGAAGAAAAACCGGCATGCGTTTGCCATTCTG
G T C R F L V Q E E K P A C V C H S G

BamHI
GCTAGGTTGGCGCACGTTGCCAACACGCTGACCTGCTGGCTTAAGGATCC
Y V G A R C E H A D L L A Ter

B. Hybride EGF-VGF de synthèse modifié.

A. Préparation de pBM11/N/TTV

Dans cette construction, un gène chimérique TTV modifié de synthèse a été exprimé comme une partie
5 C-terminale d'une protéine de fusion ayant les premiers
33 acides aminés du gène-N au N-terminal. Ce facteur de
croissance hybride contenait la séquence d'acides aminés
de TGF humain dans les deux tiers amino terminaux du gène
à l'exception de la séquence QEEK qui a été altérée à
10 partir de la séquence naturelle humaine QEDK. Le carboxy
terminal a été dérivé de la séquence d'acides aminés de
VGF et a été terminée avec la séquence YQR en amont de la
séquence naturelle PNT.

(a) Préparation de gène de synthèse TTV à
15 partir du NcoI(franc)-BamHI

Le plasmide pLEBam/TTV a été digéré avec NcoI
et les bouts ont été coupés en bouts francs en
remplissant les saillies en utilisant un fragment de
Klenow d'ADN polymérase. L'ADN a ensuite été digéré avec
20 BamHI et le fragment TTV obtenu par NcoI (franc)-BamHI a
été purifié sur gel.

Le plasmide pBM11 a été digéré avec BamHI.

pBM11 digéré avec BamHI, le fragment TTV obtenu

Gène-N →

AAGCAGCAAATCCCCTGTTGGTTGGGGTAAGCGCAAACCGATTCCGGATCCGCAT
A A N P L L V G V S A K P V R I R M

20 TGF +
GGTGTGTTTCTCACTTTAACGACTGCCCCGACTCTCATACTCAGTTTTGCTTTTCAT
V V S H F N D C P D S H T Q F C F H

GGTACCTGCCGTTTTCTGGTTCAGGAAGAAAAACCGGCATGCGTTTGCTCTCATG
G T C R F L V Q E E K P A C V C S H G

25 EcoRI
GCTACACTGGAATTCGTTGCCAGCATGTTGTTCTGGTCGACTACCAGCGTTAAG
Y T G I R C Q H V V L V D Y Q R Ter

BamHI
GATCC

Exemple V.

Préparation du polypeptide d'intérêt comme une protéine de fusion avec la protéine-N et un site de clivage.

5 A. VGf de synthèse modifié.1. Préparation de pBM11/NDP/VGFA :

La séquence N-terminale du gène VGFA de
synthèse est une version tronquée de la séquence de VGf
naturelle et commence avec la séquence DIPAIR. Dans ce
10 plasmide le fragment de VGFA est localisé en aval des 32
acides aminés de la protéine-N de lambda et du dipeptide
acide aspartique-proline. Pour préserver le site de
clonage KpnI, la séquence de synthèse a été altérée pour
coder pour CLHCGTC au lieu de la séquence naturelle de
15 VGf CLHGDC et se termine avec la séquence YQR en aval de
la séquence naturelle PNT. De plus, le gène de VGFA code
pour la séquence GYACVC qui remplace la séquence
naturelle GMYCRC.

20 a. Préparation d'un fragment C-terminal de 80bp
par KpnI-BamHI du gène VGf de synthèse.

Le plasmide pLEBam/TVV a été digéré avec KpnI
et BamHI et le fragment de 80 bp obtenu KpnI-BamHI a été
purifié sur gel. Ce fragment contient les deux tiers
C-terminaux du gène de synthèse VGf avec le site de KpnI
25 l'extrémité 5'.

b. Préparation de pBM11 déphosphorylé et digéré
par BamHI

Le plasmide pBM11/N/TTV a été digéré avec BamHI
et les phosphates 5' ont été enlevés par traitement avec
30 de la phosphatase alcaline d'intestin de veau. Le
fragment de plasmide obtenu par BamHI de 5,6 kb a été
purifié sur gel.

c. Ligation et isolement de pBM11/NDP/VGFA.

Les oligonucléotides VGf 103a, 104a, le
35 fragment obtenu par BamHI de 5,6 kb de pBM11 et le
fragment obtenu par KpnI-BamHI de 80 bp de pLEBam/TVV ont

été liés ou ligaturés ensemble en utilisant une ADN
ligase et ont ensuite été utilisés pour transformer HB101
compétente. Les transformants ont été sélectionnés sur la
néomycine et ont été criblés par analyse de restriction
5 en utilisant ClaI et en séquençant les nucléotides
suivant la technique didéoxy de Sanger. Une construction
correcte a été isolée et notée pBM11/NDP/VGFA. Cette
construction a les séquences GTC et GYACVC au lieu des
séquences authentiques de VGF GDC et GMYCRC.

Gène-N →

ATGGATGCACAAACACGCCCGCCGCGAACGTCGCGCAGAGAAACAGGCTCAATGGA
M D A Q T R R R E R R A E K Q A Q W K

ClaI

AAGCAGCAAATCCCCTGTTGGTTGGGGTAAGCGCAAACAGTTCGGATCGATC
A A N P L L V G V S A K P V R I D P

NcoI VGF →

CCATGGACATCCCGGCTATCCGTCTGTGCGGCCCGGAAGGCGACGGCTACTGCCT
M D I P A I R L C G P E G D G Y C L

KpnI

SphI

GCATGGTACCTGCATCCATGCACGTGACATCGACGGCTACGCATGCGTTTGCTCT
H G T C I H A R D I D G Y A C V C S

EcoRI

CATGGCTACACTGGAATTCGTTGCCAGCATGTTGTTCTGGTCGACTACCAGCGT
H G Y T G I R C Q H V V L V D Y Q R

BamHI

TAAGGATCC

Ter

2. Préparation de pBM11/NDP/VGFa

La séquence N-terminale de VGFA est une version tronquée de la séquence naturelle de VGF et démarre avec les séquences DIPAIR. De plus, la séquence de VGFA
5 contient les séquences altérées GTC et GYACRC au lieu des séquences naturelles de VGF, GDC et GMYCRC. Dans ce plasmide, le gène de VGFA est localisé en aval des 32^e acides aminés de la protéine- N de lambda et du dipeptide
10 acide aspartique-proline. Ce traitement de la protéine de fusion purifiée avec l'acide formique aboutit à un clivage à la liaison du peptide acide aspartique-proline, instable aux acides permettant une séparation de la
15 protéine de VGFA à partir du terminal amino de la protéine- N de lambda. Le clivage est tel que la protéine de VGFA est laissée avec le résidu proline au terminal amino.

a. Préparation de pBM11/DP/VGFA déphosphorylé et digéré par SphI.

Le plasmide pBM11/DP/VGFA (10 µg) a été digéré
20 avec 30 unités de SphI et les phosphates 5' ont été enlevés par traitement avec de la phosphatase alcaline d'intestin de veau. Le fragment de plasmide de 5kb a été récupéré après électrophorèse sur un gel d'agarose.

b. Préparation d'un fragment de 70 bp par EcoRI-SphI de PBM11/DP/VGA.

Le plasmide pBM11/DP/VGFA (10 µg) a été digéré
25 avec 30 unités de EcoRI et ensuite 30 unités de SphI. Le fragment de 70bp a été récupéré après électrophorèse sur un gel d'agarose.

30 3. Ligation et isolement de pBM11/NDP/VGFa

Le fragment de 24bp contenant des
oligonucléotides de VGF 1A et 2A, le fragment de 5kb
obtenu par SphI et le fragment de 70 bp obtenu par
EcoRI-SphI de pBM11/DP/VGFA ont été liés ou ligaturés
35 ensemble et le mélange a été utilisé pour transformer les cellules de E. coli HB101 compétentes. Les transformants

ont été criblés par séquençage des nucléotides en utilisant la méthode pour nucléotides de didésoxy de Sanger. Un clone correct a été isolé et noté pBM11/NDP/VGFa.

5 Gène-N →
 ATGGATGCACAAACACGCCGCCGCGAACGTCGCGCAGAGAAACAGGCTCAATGGA
 M D A Q T R R R E R R A E K Q A Q W K

 ClaI
 10 AAGCAGCAAATCCCCTGTTGGTTGGGGTAAGCGCAAAACCAGTTCGGATCGATCC
 A A N P L L V G V S A K P V R I D P

NcoI VGF -
 CCATGGACATCCCGGCTATCCGTCTGTGCGGCCCGGAAGGCGACGGCTACTGCCT
 M D I P A I R L C G P E G D G Y C L

KpnI SphI
 15 GCATGGTACCTGCATCCATGCACGTGACATCGACGGCTACGCATGCCGTTGCTCT
 H G T C I H A R D I D G Y A C R C S

EcoRI
 CATGGCTACACTGGAATTCGTTGCCAGCATGTTGTTCTGGTCTCGACTACCAGCGT
 H G Y T G I R C Q H V V L V D Y Q R

20 BamHI
 TAAGGATCC
 Ter

B. Hybrides TGF-VGF de synthèse modifiés

1. Préparation de pBM11/NDP/TTV

Dans cette construction, le gène chimérique de TTV modifié de synthèse est exprimé comme la partie
5 C-terminale d'une protéine de fusion ayant les 32 premiers acides aminés du gène-N au N-terminal. Un dipeptide acide aspartique-proline instable à l'acide sépare les deux parties de la fusion. Le facteur de croissance hybride contient la séquence d'acides aminés
10 de TGF humain dans les deux tiers amino-terminaux du gène à l'exception de la séquence QEEK qui a été altérée à partir de la séquence de TMR humaine naturelle QEDK. Le terminal carboxy était dérivé de la séquence d'acides aminés de VGF et se terminait avec la séquence YQR en
15 amont de la séquence naturelle PNT.

a. Préparation d'un fragment de plasmide pBM11 de 5 kb par NcoI.

Le plasmide pBM11/NDP/VGFA a été digéré avec NcoI et le fragment du plasmide de 5kb obtenu par NcoI a
20 été purifié sur gel. Ce fragment a une saillie NcoI au site de clivage acide aspartique-proline en aval des séquences codant pour les premiers 32 acides aminés du gène-N. L'autre site NcoI est dans le gène de résistance à la néomycine.

25 b. Préparation d'un fragment de pBM11 de 0,6 kb par NcoI-BamHI.

Le plasmide pBM11/N/TTV a été digéré avec NcoI et BamHI et le fragment du plasmide de 0,6 kb obtenu par NcoI-BamHI a été purifié sur gel. Ce fragment a la
30 saillie NcoI dans le gène de résistance à la néomycine.

c. Préparation du fragment de TGF/TGF/VGF de synthèse de 170bp.

Le plasmide pLEBam/TTV a été digéré avec NcoI et BamHI et le fragment de 170bp obtenu par NcoI-BamHI
35 contenant le gène de synthèse de TGF/TGF/VGF a été

d. Ligation et isolement de pBM11/NDP/TTV

Gène-N →

NcoI TGF →

30

BamHI
TAAGGATCC
Ter

2. Préparation de pBM11/NDP/VTV

Dans cette construction, le gène chimérique de VTV modifié de synthèse a été exprimé comme une partie C-terminale d'une protéine de fusion ayant les 32 premiers acides aminés du gène-N au N-terminal. Un dipeptide acide aspartique-proline instable aux acides sépare les deux parties de la fusion. Le facteur de croissance hybride contenait la séquence d'acides aminés de TGF humain dans le domaine du milieu avec la séquence d'acides aminés QEEK remplaçant la séquence naturelle QEDK. Les domaines N-terminal et C-terminal ont été dérivés de la séquence de VGF non tronquée et commence avec la séquence DIPAIR et finit avec la séquence YQR qui est en aval de la séquence naturelle PNT.

15 a. Préparation d'un fragment de pBM11 de 5 kb par BamHI-NcoI.

Le plasmide pBM11/N/TTV a été digéré avec BamHI et NcoI et le fragment de 5kb obtenu par BamHI-NcoI a été purifié sur gel. Ce fragment contient une saillie BamHI à l'extrémité 3' des séquences codant pour les premiers 32 acides aminés du gène -N et leur site NcoI dans le gène de résistance à la néomycine.

20 b. Préparation d'un fragment de pBM11/N/TTV de 700 bp par KpnI-NcoI

Le plasmide pBM11/N/TTV a été digéré avec KpnI et NcoI et le fragment de 700 bp obtenu par KpnI-NcoI a été purifié sur gel. Ce fragment est composé d'une partie du plasmide pBM11 contenant une partie du gène de résistance à la néomycine à la partie qui dépasse NcoI, et le domaine C-terminal de VGF du gène de synthèse de TTV à la partie qui dépasse KpnI.

30 c. Ligation et isolement de pBM11/NDP/VTV

Les oligonucléotides VGF 103a et 104a, le fragment de 5kb obtenu par BamHI-NcoI de pBM11 et le fragment de 700 bp obtenu par KpnI-NcoI de pBM11/N/TTV ont été liés ou ligaturés ensemble en utilisant une ADN

ligase et ont ensuite été utilisés pour transformer HB101 compétente. Les transformants ont été sélectionnés sur de la néomycine et ont été criblés par analyse de restriction en utilisant ClaI et un séquençage de
 5 nucléotides utilisant la technique didésoxy de Sanger.

Gène-N →

ATGGATGCACAAACACGCCGCCGCGAACGTCGCGCAGAGAAACAGGCTCAATGGA
 M D A Q T R R R E R R A E K Q A Q W K

ClaI

10 AAGCAGCAAATCCCCTGTTGGTTGGGGTAAGCGCAAAACCAGTTCGGATCGATC
 A A N P L L V G V S A K P V R I D P

NcoI VGF -

CCATGGACATCCCGCTATCCGTCTGTGCGGCCCGGAAGGCGACGGCTACTGCCT
 M D I P A I R L C G P E G D G Y C L

KpnI TGF -

15 GCATGGGTACCTGCCGTTTTCTGGTTCAGGAAGAAAAACCGGCATGCGTTTGCTCT
 H G T C R F L V Q E E K P A C V C S

SphI

VGF -

EcoRI

CATGGCTACACTGGAATTCGTTTCGAGCATGTTGTTCTGGTCGACTACCAGCGT
 H G Y T G I R C Q H V V L V D Y Q R

BamHI

20 TAAGGATCC
 Ter

3. Préparation de pBM16/NDP/TVV

Dans cette construction, le gène chimérique de TVV modifié de synthèse a été exprimé comme la partie C-terminale d'une protéine de fusion ayant les 32 premiers acides aminés du gène-N au N-terminal. Un dipeptide acide aspartique-proline instable à l'acide sépare les deux parties de la fusion. Le facteur de croissance hybride contenait une séquence d'acides aminés de TGF humain dans le domaine N-terminal. Les domaines du milieu et de C-terminal ont été dérivés de la séquence de VGF non tronquée et de l'extrémité de la séquence YQR. En plus, le gène de synthèse a la modification : GYACVC pour GMYCRC.

a. Préparation de fragment de 4,3 kb par NcoI-BglII de pBM11/NDP/VGFa.

Le plasmide pBM11/NDP/VGFa a été digéré avec NcoI et BglII et le fragment de 4,3 kb a été purifié sur gel. La partie en saillie NcoI est positionnée au site de clivage acide aspartique-proline juste en aval des premiers 32 acides aminés du gène N.

b. Préparation du fragment de 1,2 kb par BamHI-BglII de pBM11M5.

Le plasmide pBM11M5 a été digéré avec BamHI et BglII et le fragment de 1,2kb a été purifié sur gel. Ce fragment diffère du fragment de pBM11 normal en ce que le site NcoI dans le gène de résistance à la néomycine a été enlevé, et tous les vecteurs par la suite perdant ce site NcoI sont référés comme pBM16.

c. Préparation du gène de synthèse de TVV de 170 bp par NcoI-BamHI.

Le plasmide pLEBam/TVV a été digéré avec NcoI et BamHI et le fragment de 170 bp obtenu par NcoI-BamHI a été purifié sur gel. Le fragment du gène de synthèse a le site NcoI à l'extrémité 5', le site de BamHI à l'extrémité 3'.

Le fragment de 4,3 kb obtenu par NcoI-BglIII de pMB11/NDP/VGFA et le fragment de 1,2 kb obtenu par BamHI-BglIII de pBM11M5, et le fragment du gène de synthèse de TVV de 170 bp obtenu par NcoI-BamHI ont été liés ou ligaturés ensemble en utilisant une ADN ligase et le mélange résultant a été utilisé pour transformer HB101 compétente. Les transformants ont été sélectionnés sur la néomycine et criblés par analyse de restriction et par séquençage des nucléotides en utilisant la technique du didésoxy de Sanger. Le plasmide est marqué pBM16 pour indiquer la perte du site de restriction NcoI dans le gène de résistance à la néomycine.

15 Gène-N →
ATGGATGCACAAACACGCCGCCGCGAACGTCGCGCAGAGAAACAGGCTCAATGGA
M D A Q T R R R E R R A E K Q A Q W K

 ClaI
 AAGCAGCAAATCCCTGTTGGTTGGGGTAAGCGCAAAACCAAGTTCGGATCGATC
 A A N P L L V G V S A K P V R I D P

20 NcoI TGF +
 CCATGTTGTTTCTCACTTTAACGACTGCCGGACTCTCATACTCAGTTTTGCTT
 M V V S H F N D C P D S H T Q F C F

TCATGGTACCTGCATCCATGCACGTGACATCGACGGCTACGCATGCGTTTGCTCT
H G T C I H A R D I D G Y A C V C S

EcoRI
 CATGGCTACACTGGAATTCGTTGCCAGCATGTTGTTCTGGTCGACTACCAGCGT
 H G Y T G I R C Q H V V L V D Y Q R

30 BamHI
TAAGGATCC
Ter

C. EGF de synthèse

1. Préparation de PBM11/NDP/EGF

Dans cette construction le gène d'EGF humain est exprimé comme partie d'une fusion avec 32 acides aminés N-terminaux du gène N qui est en aval d'un site de clivage Asp-Pro.

a. Préparation d'un fragment de 5kb par NcoI de pBM11.

Le plasmide pBM11/DP/VGFA a été digéré avec NcoI et les phosphates 5' ont été enlevés par traitement avec de la phosphatase alcaline d'intestin de veau. Le fragment de plasmide de 5kb a été purifié sur gel. Ce fragment a une saillie NcoI du site de clivage Asp-Pro en aval des séquences codant pour les 32 premiers acides aminés du gène N. L'autre site NcoI est dans le gène de résistance à la néomycine.

b. Préparation d'un fragment de 0,6 kb par NcoI-BamHI de pBM11.

Le plasmide pBM11/N/TTV a été digéré avec NcoI et BamHI et le fragment de plasmide de 0,6 kb obtenu par NcoI-BamHI a été purifié sur gel. Ce fragment a une saillie NcoI dans le gène de résistance à la néomycine.

c. Ligation et isolement de pBM11/DP/EGF

Les trois ensembles d'oligonucléotides d'EGF cyclisés avec une saillie NcoI à l'extrémité 5' et une saillie BamHI à l'extrémité 3', le fragment de 5kb obtenu par NcoI de pBM11 et le fragment de 0,6 kb obtenu par NcoI-BamHI de pBM11 étaient liés ou ligaturés ensemble en utilisant l'ADN ligase T4 et le mélange résultant a été utilisé pour transformer E. coli HB101 compétente. Les transformants ont été sélectionnés sur la néomycine de façon à ce que seulement des colonies avec un gène de résistance à la néomycine correctement reconstruit puisse survivre. Les transformants ont été criblés par analyse de restriction en utilisant EcoRI et BamHI et par le séquençage d'ADN comme décrit ci-dessus.

Gène-N →

ATGGATGCACAAACACGCCGCCGCGAACGTCGCGCAGAGAAACAGCGTCAATGGA
M D A Q T R R R E R R A E K Q A Q W K

AAGCAGCAAATCCCCTGTTGGTTGGGGTAAGCGCAAAACCAGTTCGGATCGATCC
A A N P L L V G V S A K P V R I D P

EGF1 → EcoRI

EGF2 →

CATGAATTCTGACTCTGAATGCCCGCTGTCTCATGACGGCTACTGCCTGCATGAC
M N S D S E C P L S H D G Y C L H D

EGF3 →

NsiI

SphI

GGCGTATGCATGTACATCGAAGCTCTGGACAAGTACGCATGCAACTGCGTTGTTG
G V C M Y I E A L D K Y A C N C V V G

GCTACATCGGCGAACGTTGCCAGTACCGTGACCTGAAATGGTGGGAACTGCGTTA
Y I G E R C Q Y R D L K W W E L R *

BamHI

AGGATCC

Exemple VI.

Préparation du polypeptide d'intérêt comme une protéine de fusion avec une séquence signal de phosphatase alcaline modifiée.

5 A. Préparation de pBM11/PAD/EGF

Des oligonucléotides de synthèse ont été désignés pour permettre une insertion d'ADN codant pour un peptide signal de phosphatase alcaline modifié et une région "linker" avec trois sites de clonage (HindIII,
10 SmaI et BamHI) dans le vecteur d'expression pBM11 en aval du promoteur P_L et du site de liaison au ribosome du gène N. La séquence de nucléotides a été optimisée pour être aussi similaire que possible à la séquence de nucléotides de l' amino terminal du gène N lambda comme la séquence du
15 gène N lambda a été déroulée avec celle de son site de liaison au ribosome pour une initiation et une traduction du ribosome efficace. De plus, le second acide aminé de la séquence signal phosphatase alcaline, l'acide aminé lysine basique a été changé en un acide aminé acide,
20 l'acide aspartique.

1. Préparation du fragment de 0,17kb par EcoRI-BamHI de EGF.

Le plasmide pBM11/NDP/EGF (30 μ g) a été digéré avec 30 unités de EcoRI et a ensuite été traité avec 4
25 unités de fragment de Klenow d'ADN polymérase pour créer des bouts francs. L'ADN a finalement été digéré avec 30 unités de BamHI et le fragment de 0,17 kb du gène EGF a été récupéré après électrophorèse sur gel d'agarose. L'ADN ainsi purifié a un site EcoRI à bord franc à
30 l'extrémité 5' et une partie en saillie BamHI à l'extrémité 3'.

2. Préparation du fragment de 0,5kb par PvuI -HindIII de pBM11/PAD.

Le plasmide pBM11/PAD (18 μ g) a été digéré avec
35 30 unités de HindIII et a ensuite été traité avec le fragment de Klenow pour couper les extrémités. L'ADN a

ensuite été digéré avec PvuI et le fragment de 0,5kb obtenu par PvuI-HindIII (bout franc) a été récupéré après électrophorèse sur un gel d'agarose.

5 3. Préparation de fragment de 5,2 kb par PvuI -BamHI de pBM11/PAD.

Le plasmide pBM11/PAD (18µg) a été digéré avec 30 unités de PvuI suivies par 30 unités de BamHI. Le fragment de 5,2kb a été récupéré après électrophorèse sur un gel d'agarose.

10 4. Ligation et isolement de pBM11/PAD/EGF.

Le fragment de 0,17kb obtenu par EcoRI (bout franc) -BamHI, le fragment de 0,5kb obtenu par PvuI-HindIII (bout franc) et le fragment de 5,2 kb obtenu par PvuI-BamHI ont été ligaturés ou liés ensemble et le mélange résultant a été utilisé pour transformer E. coli HB101 compétente. Les transformants ont été criblés en utilisant le séquençage d'ADN comme décrit ci-dessus. La région souhaitée séquence signal/EGF avait la séquence suivante :

20 Séquence Signal

ATGGATCAATCTACAATCGCCCTCGCACTTCTCCCACTGCTGTTCACT
M D Q S T I A L A L L P L L F T

EGF

25 CCAGTGACAAAAGCTAATTCTGACTCTGAATGCCCGCTGTCTCATGAC
P V T K A N S D S E C P L S H D

NsiI

GGCTACTGCCTGCATGACGGCGTATGCATGTACATCGAAGCTCTG
G Y C L H D G V C M Y I E A L

SphI

30 GACAAGTACGCATGCAACTGCGTTGTTGGCTACATCGGCGAACGT
D K Y A C N C V V G Y I G E R

BamHI

TGCCAGTACCGTGACCTGAAATGGTGGGAAGTGGCTTAAGGATCC
C Q Y R D L K W W E L R *

L'efficacité de la production de protéine étrangère dans le système d'expression pBM11/PAD et la capacité pour purifier des protéines étrangères fonctionnellement actives à partir du produit de fusion a été montrée en utilisant pBM11/PAD/EGF comme exemple. 5 Après chromatographie d'exclusion (TSK-250), 10,3 mg équivalents de polypeptide de fusion d'EGF actif a été recueilli à partir de 23 g (8 litres) de E. coli 10 déréprimée pour exprimer le gène de EGF. 40% de l'activité de EGF étaient dérivés de EDF clivés à partir de la séquence signal.

B. Préparation de pBM11/PAD/nVGFa

Des oligonucléotides de synthèse ont été désignés pour se lier au gène de synthèse de VGFa avec 15 une séquence signal phosphatase alcaline modifiée pour fournir un site de clivage optimal d'une séquence signal en codant pour des résidus supplémentaires N-terminaux survenant immédiatement en aval du site de clivage de la 20 séquence signal dans le VGF naturel, désigné ci-dessus N-terminal extrême. La séquence nVGFa contient les séquences altérées GTC et GYACRC au lieu des séquences naturelles GDC et GMYCRC de VGF et se terminent avec la séquence YQR en amont de la séquence naturelle PNT. Dans 25 ce système d'expression, dans la majorité des molécules, la séquence signal reste attachée au nVGFa formant une protéine de fusion avec nVGFa au C-terminal.

1. Préparation d'un fragment de 0,5 kb de pBM1/PAD digéré par HindIII-PvuI.

Le plasmide pBM11/PAD a été digéré avec HindIII 30 et PvuI et le fragment de 0,5kb a été purifié sur gel. Le site de HindIII est localisé au C-terminal de la séquence signal phosphatase alcaline modifiée..

2. Préparation du fragment de plasmide pBM11 de 5,2 kb par PvuI-BamHI.

Le plasmide pBM11/NDP/VGFa a été digéré avec PvuI et BamHI et le fragment de plasmide de 5,2kb a été purifié sur gel.

3. Préparation du gène de VGFa de synthèse de 170 bp par NcoI(bout franc)-BamHI.

Le plasmide pMB11/NDP/VGFa a été digéré avec NcoI et les parties qui dépassent ou bouts cohésifs 5' ont été enlevées par traitement avec une S1-nucléase. Cela créait un bout franc au premier codon du gène de synthèse tronqué de VGFa. L'ADN a ensuite été digéré avec BamHI et le fragment de 170 bp obtenu par NcoI (bout franc)-BamHI a été purifié sur gel.

4. Ligation et isolement de pBM11/PAD/nVGFa.

Des oligonucléotides VGF105 et 106, le fragment de 0,5kb obtenu par HindIII-PvuI de pBM11/PAD, le fragment de pBM11 de 5,2 kb obtenu par PvuI-BamHI et le gène de synthèse de VGFa de 170 bp obtenu par NcoI (bout franc)-BamHI ont été liés ou ligaturés ensemble en utilisant une ADN ligase et le mélange résultant a été utilisé pour transformer HB101 compétente. Les transformants ont été sélectionnés sur la néomycine et criblés par analyse de restriction et séquençage de nucléotide en utilisant la technique de didésoxy de Sanger. Une construction correcte a été isolée contenant la séquence signal phosphatase alcaline modifiée dans un cadre avec le gène de nVGFa.

Séquence Signal -
30 ATGGATCAATCTACAATCGCCCTCGCACTTCFCCCACTGCTGTTCCTCCAGTGACAAAA
M D Q S T I A L A L L P L L F T P V T K

nVGf -
GCTGACTCTGGTAACGCTATCGAACTACTTCTCCGGAAATCACTAACGCTACTACT
A D S G N A I E T T S P E I T N A T T

35 GACATCCCGGCTATCCGTCTGTGCGGCGCGGAAGGCGACGGCTACTGCCTGCATGGT^{KpnI}
D I P A I R L C G P E G D G Y C L H G

EcoRI
CTCATGGCTACACTGGAATTCGTTGCCAGCATGTTGTTCTGGTCCGACTACCAGCG
H G Y T G I R C Q H V V L V D Y Q R

BamHI
TTAAGGATCC
Ter

Expression du polypeptide d'intérêt comme une protéine de fusion avec une séquence signal phosphatase alcaline.

Le plasmide pBM11/PAD/nVG a subi une mutagénèse in vitro (Morinaga et al., Biotechnology (1984) 2:636--643) pour altérer les codons codants pour le second acide aminé dans la séquence signal, c'est-à-dire pour changer le codon Asp (D) pour celui pour Lys (K), le résidu trouvé dans la séquence naturelle. Cette mutagénèse introduisait également un site PvuI dans la séquence signal.

Séquence Signal →
ATGGATCAATCTACAATCGCCCTCGCACTTCTCCCACTGCTGTTCACTCCAGTGA
M D O S T I A L A L L P L L F T P V T

20 nVGF -
CAAAAGCTGACTCTGGTAACGCTATCGAACTACTTCTCCGGAAATCACTAACGC
 K A D S G N A I E T T S P E I T N A

TACTACTGACATCCCGGCTATCCGTCTGTGCGGCCCGGAAGGCGACGGCTACTGC
T T D I P A I R L C G P E G D G Y C

CTGCATGGTACCTGCATCCATGCACGTGACATCGACGGCTACGCATGCCGTTGCT
25 L H G T C I H A R D I D G Y A C R C S

ACCTGCATCCATGCACGTGACATCGACGGCTACG^{SphI}CATGCCGTTGCTCTCATGGCTACACT
T C I H A R D I D G Y A C R C S H G Y T

EcoRI BamHI

30 GGAATTCGTTGCCAGCATGTTGTCTGGTCTCGACTACCAGCGTTAAGGATCC
G I R C Q H V V L V D Y Q R Ter

Dans cette cassette d'expression, le gène d'EGF est partie d'une fusion avec la séquence signal phosphatase alcaline.

Séquence Signal PvuI
 ATGAAACAATCTACGATCGCCCTCGCACTTCTCCCACTGCTGTTCACTCCAGTGA
 M K Q S T I A L A L L P L L F T P V T

TGACGGCGTATGCATGTACATCGAAGCTCTGGACAAGTACGCATGCCAACTGCGTT
D G V C M Y I E A L D K Y A C N C V

BamHI
GTTAAGGATCC

25 *

C. Préparation de TacPak/EGF (séquence signal phosphatase alcaline/EGF d'humain).

1. Préparation de fragments de plasmide.

Le plasmide p135-1 a été dérivé du plasmide pDR540 (Pharmacia) et contenait le SD du gène Cro et un site BglII en aval du SD de lac. pDR540 est un vecteur d'expression contenant le promoteur hybride trp-lac. p135-1 a été digéré avec BglII et BamHI et traité avec une phosphatase alcaline bactérienne.

Le plasmide pBM11/PAK/EGF a été digéré avec PvuII et BamHI et le fragment d'environ 230 bp codant pour une partie de la séquence signal de phosphatase alcaline et EGF humain a été isolé.

2. Préparation des oligonucléotides TacPak1 et TacPak2.

Des oligonucléotides de synthèse TacPak1 et TacPak2 ont été désignés avec une partie qui dépasse ou en saillie, compatible avec le site de BglII de p135-1 et une partie en saillie de PvuII, compatible avec le site PvuII dans le fragment de phosphatase alcaline/EGF PvuII/BamHI. Les oligonucléotides ont été synthétisés sur un synthétiseur d'oligonucléotides d'Applied Biosystems.

		<u>BglII</u>	<u>PvuI</u>
TacPak1	5'	GATCTATGAAACAATCTACGAT	3'
TacPak2	3'	ATACTTTGTTAGATGC	5'

3. Ligation et isolement du clone TacPak/EGF.

Le p135-1 digéré par BglII-BamHI, le fragment PAK/EGF 230 bp et les oligonucléotides TacPak1 et TacPak2 ont été liés en utilisant une ADN ligase, transformés dans HB101 compétente et une construction correcte a été isolée par séquençage d'ADN.

(Site HindIII de pDR540)

AAGCTTACTCCC

trp-35 (16bp) lac-10
CATCCCCCTG [TTGACA] ATTAATCATCGGCTCG (TATAATG)

5 ARNm site de liaison lacI lacSD
TGTTGG/AATTGTG AGCGGATAACAATTTACAC {AGGA} AACAGGATCACTA

PvuI
croSD (11bp) BglII
{AGGA} GGTTCAGATCT

10 Séquence Signal →
ATGAAACAATCTACGATCGCCCTCGCACTTCTCCCACTGCTGTTCCTCCAGTGA
M K Q S T I A L A L L P L L F T P V T

EGF →
15 CAAAAGCTAATTCTGACTCTGAATGCCCGCTGTCTCATGACGGCTACTGCCTGCA
K A N S D S E C P L S H D G Y C L H

NsiI SphI
TGACGGCGTATGCATGTACATCGAAGCTCTGGACAAGTACGCATGCAACTGCGTT
D G V C M Y I E A L D K Y A C N C V

20 GTTGGCTACATCGGCGAACGTTGCCAGTACCGTGACCTGAAATGGTGGGAAGTGC
V G Y I G E R C Q Y R D L K W W E L R

BamHI
GTTAAGGATCC
*

Exemple VIII.

Expression d'un polypeptide d'intérêt comme une protéine de fusion avec une séquence signal de phosphatase alcaline en utilisant une cassette d'expression comprenant une région de terminaison de la traduction.

A. Préparation de pTCPt/EGF ([trp-35] 16bp [lac-10] [lacSD] 11bp [ATG] /signal phosphatase alcaline/EGF humain/term.trad. -NEO)

Ce plasmide est désigné comme ayant les éléments du promoteur tac et utilise le SD de cro pour exprimer EGF humain derrière la séquence signal phosphatase alcaline. Il a un fond de pBR322 avec le gène de résistance à la néomycine.

1. Préparation du fragment de 420 bp par HindIII (bout franc)-BamHI de TacPak/EGF.

TacPak/EGF a été digéré avec HindIII et ensuite traité avec le fragment de Klenow de l'ADN polymérisase pour créer des bouts francs. L'ADN a ensuite été digéré avec BamHI et le fragment de 420bp contenant les éléments du promoteur tac et la région codante pour la séquence signal phosphatase alcaline et EGF humain a été isolé par électrophorèse sur gel d'agarose.

2. Préparation du fragment de 2,8 kb par EcoRI (bout franc)-BamHI de pBM16t/NDP/VGFa.

pBM16t/NDP/VGFa a été digéré avec EcoRI et ensuite traité avec le fragment de Klenow pour créer des extrémités ou bouts francs. L'ADN a ensuite été digéré avec BamHI et le fragment de 2,8kb a été isolé. Ce fragment d'ADN contient l'origine de pBR322, le gène de résistance à la néomycine avec son site NcoI enlevé, et le terminateur de transcription semblable au gène 32 en aval du site BamHI.

3. Ligation et isolement de pSTCPT/EGF.

Le fragment de 2,8kb obtenu par EcoRI(bout franc)-BamHI de pBM16t/NDP/VGFa a été lié ou ligaturé au fragment de 420bp obtenu par HindIII (bout franc)-BamHI

de TacPak/EGF et l'ADN résultant a été utilisé pour transformer JM109 (ladIq) compétente. Une construction correcte a été isolée par sa résistance à la néomycine et par séquençage d'ADN.

5

(Site HindIII de pDR540)

AAGCTTACTCCC

trp-35 (16bp) lac-10
CATCCCCCTG [TTGACA] ATTAATCATCGGCTCG (TATAATG)

10

ARNm 5' site de liaison lacI lacSD
TGTGG/AATTGTG AGCGGATAACAATTTTCACAC {AGGA} AACAGGATCACTA

15

PvuI
croSD (11bp) BglIII Signal Sequence ->
{AGGA} GGTTCAGATCT ATGAAACAATCTACGATCGCCCTCGCACTTCTCC
M K Q S T I A L A L L P

EGF ->

CACTGCTGTTCACCTCCAGTGACAAAAGCTAATTCTGACTCTGAATGCCCGCTGTC
L L F T P V T K A N S D S E C P L S

20

NsiI
TCATGACGGCTACTGCCTGCATGACGGCGTATGCATGTACATCGAAGCTCTGGAC
H D G Y C L H D G V C M Y I E A L D

SphI
AAGTACGCATGCAACTGCGTTGTTGGCTACATCGGC
K Y A C N C V V G Y I G

25

BamHI
GAACGTTGCCAGTACCGTGACCTGAAATGGTGGGAAGTGCCTTAAGGATCCGTGA
E R C Q Y R D L K W W E L R *

Term. Trad.

CTAATTGGGGACCCTAGAGGTCCCCTTTTTTATTTTAAAACGATC

B. Préparation de pTCPt/nVGFa ([trp-35] 16bp [lac-10] [lacSD] [croSD] 11bp [ATG] /signal phosphatase alcaline/VGFa N-terminal avec séquence GTC et GYACRC)/term.trad. -NEO).

5 Ce plasmide a les éléments promoteurs tac et utilisent le SD de cro pour exprimer le gène modifié de VGF avec la prolongation N-terminale en aval de la séquence signal phosphatase alcaline. Le plasmide a un fond de pBR322 avec le gène de résistance à la néomycine.

10 1. Préparation du fragment 350 bp par PvuI-BamHI de pBM11/PAK/nVGFa.

15 Le plasmide pBM11/PAK/nVGFa a été digéré avec PvuI et BamHI et le fragment de 350 bp a été isolé par électrophorèse sur gel. Ce fragment contient la plupart de la séquence signal phosphatase alcaline et le gène nVGFa.

2. Préparation du fragment de 2,8 kb de pTCPt/EGF obtenu par PvuI-BamHI.

20 Le plasmide pTCPt/EGF a été digéré avec PvuI et BamHI et le fragment de 2,8 kb a été isolé par électrophorèse sur gel.

Ligation et isolement de pTCPt/nVGFa

25 Le fragment de 2,8kb et le fragment de 350bp ont été liés en utilisant une ADN ligase et l'ADN a été utilisé pour transformer JP109(lacIq) compétente. Une construction correcte a été isolée en utilisant une analyse par restriction.

3. Ligation et isolement de pTCPt/nVGFa

30 Le fragment de 2,8 kb et le fragment de 350 bp ont été liés en utilisant une ADN ligase et l'ADN a été utilisé pour transformer JM109 (lacIq) compétente. Une construction correcte a été isolée en utilisant une analyse par restriction.

(Site HindIII de pDR540)

AAGCTTACTCCC

trp-35 (16bp) lac-10
 CATCCCCCTG [TTGACA] ATTAATCATCGGCTCG (TATAATG)

5 ARNm 5' site de liaison lacI lacSD
 TGTGG/AATTGTG AGCGGATAACAATTTACAC {AGGA} AACAGGATCACTA

croSD (11bp) BglIII Séquence Signal -> PvuI
 {AGGA} GGTTTCAGATCT ATGAAACAATCTACGATCGCCCTCGCACTTCTCCC
 10 M K Q S T I A L A L L P

nVGF ->
 ACTGCTGTTCACTCCAGTGACAAAAGCTGACTCTGGTAACGCTATCGAAACTACT
 L L F T P V T K A D S G N A I E T T

15 TCTCCGGAAATCACTAACGCTACTACTGACATCCCGGCTATCCGTCTGTGCGGCC
 S P E I T N A T T D I P A I R L C G P

KpnI
 CGGAAGGCGACGGCTACTGCCTGCATGGTACCTGCATCCATGCACGTGACATCGA
 E G D G Y C L H G T C I H A R D I D

SphI EcoRI
 20 CGGCTACGCATGCCGTTGCTCTCATGGCTACACTGGAATTCGTTGCCAGCATGTT
 G Y A C R C S H G Y T G I R C Q H V

BamHI Term.
 GTTCTGGTCGACTACCAGCGTTAAGGATCCGTGACTAATTGGGGACCCTAGAGGT
 V L V D Y Q R *

25 Trad.
 CCCCTTTTTTATTTTAAAACGATC

C. Préparation de pTNPt/EGF ([trp-35] 17bp
[lac-10] [nSD] 8bp [ATG] /signal phosphatase alcaline/EGF
humain/term. trad. -NEO).

5 Ce plasmide est désigné comme ayant les
éléments promoteurs tac et utilisent le SD du gène-N pour
exprimer EGF humain derrière la séquence signal de
phosphatase alcaline. Il a un fond de pBR322 avec un gène
de résistance à la néomycine.

10 1. Préparation d'un fragment de pTNPt de 2,8 kb
par PvuI-BamHI.

Le plasmide pTNPt a été digéré avec PvuI et
BamHI et le fragment de 2,8 kb a été isolé par
électrophorèse sur gel.

15 2. Préparation du fragment de 300 bp de
pBM11/PAK/EGF par PvuI-BamHI.

Le plasmide pBM11/PAK/EGF a été digéré avec
PvuI et BamHI et le fragment de 300 bp a été isolé.

3. Ligation et isolement de pTNPt/EGF.

20 Le fragment de 2,8 kb et le fragment de 300bp
ont été liés en utilisant une ADN ligase et l'ADN a été
transformé dans JM109 (lacIq) compétente. Une
construction correcte a été isolée par analyse par
restriction et par séquençage d'ADN.

GAATTACTCCCCATCC

ARNm -> BsmI
n ARNm ->
TGTGAGCGGATAACAATTTACACAGCATTCAAAGCAGAAGGCTTTGGGGTGTGT

10 CAATCGCGGGGGGTTTCGTTCAGGACTACAAGTCCACACACCACCAAAGCTAA

CTGAC {AGGA} GAATCCAG ATGAAACAATCTACGATCGCCCTCGCACTTCTC
M K Q S T I A L A L L

1.5 EGF ->
CCACTGCTGTTCACTCCAGTGACAAAAGCTAATTCTGACTCTGAATGCCCGCTGT
P L L F T P V T K A N S D S E C P L S

CTCATGACGGCTACTGCCTGCATGACGGCGTATGCATGTACATCGAAGCTCTGGA
20 H D G Y C L H D G V C M Y I E A L D

CAAGTACGCATGCAACTGCGTTGTTGGCTACATCGGCGAACGTTGCCAGTACCGT
K Y A C N C V V G Y I G E R C Q Y R

25 BamHI BamHI
GACCTGAAATGGTGGGAAGTTCGTTAAGGATCCGTGACTAATTGGGGA
D L K W W E L R *

Term. Trad. (Site BamHI de pBM11)
CCCTAGAGGTCCCCTTTTTTATTTTAAAACGATCC

Exemple IX.Assemblage de gènes de facteur de croissance
pour l'expression dans des cellules eucaryotiques.A. Préparation du plasmide pRSV/VGF :

5 La transfection de ce plasmide des cellules
d'ovaires de hamsters chinois aboutissait à la production
de VGF naturel.

(a) Préparation de pRSV par HindIII(bout
franc)-BglIII (bout franc).

10 Le plasmide pRSV a été digéré avec HindIII et
BglIII et le fragment de 4kb a été purifié sur gel. Les
extrémités qui dépassent ont été faites en bouts francs
avec un fragment de Klenow d'ADN polymérase et ensuite
les phosphates 5' ont été enlevés avec de la phosphatase
15 alcaline intestinale de veau.

(b) Préparation d'un fragment DdeI contenant le
gène VGF.

Un fragment sous-cloné du génome d'ADN de la
vaccine a été digéré avec DdeI, les extrémités qui
20 dépassent ont été transformées en bouts francs utilisant
le fragment de Klenow et un fragment de 550 bp obtenu par
DdeI a été purifié sur gel.

(c) Ligation et isolement de pRSV/VGF.

Le fragment de pRSV de 4kb obtenu par HindIII
25 (bout franc)-BglIII (bout franc) et le fragment d'ADN de
la vaccine de 550bp obtenu par DdeI (bout franc) ont été
liés ensemble en utilisant une ADN ligase et les
transformants ont été sélectionnés avec l'ampicilline et
criblés par analyse par restriction. Une construction
30 correcte a été isolée et notée pRSV/VGF.

(d) Transfection de cellules CHO par pRSV/VGF

Le plasmide pRSV/VGF a été cotransfecté avec un
plasmide pRSV par précipitation au phosphate de calcium
dans des cellules d'ovaires de hamsters chinois (CHO).

Les transfectants ont été criblés par hybridation par Southern dot blot et un clone positif désigné pRSV/VGF52 a été isolé.

B. Préparation du plasmide pAc/VGF.

5 Un système d'expression supplémentaire qui a
été utilisé pour exprimer la protéine recombinante de VGF
est un système insecte. Voir Maeda et al. et Carbonell et
al., supra. Dans un tel système, le virus de la
polyhidrose nucléaire d'Autographa Californica (AcNPV)
10 est utilisé comme vecteur pour exprimer des gènes
étrangers. Le virus croît dans des cellules de Spodoptera
frugiperda. Le gène d'enveloppe peut être cloné dans des
régions non essentielles (par exemple, le gène de la
polyhédrine) du virus et est placé sous contrôle d'un
15 promoteur d'AcNPV (par exemple, le promoteur de
polyhédrine). Une insertion avec succès de la construc-
tion du gène de VGF aboutit à une inactivation du gène de
polyhédrine et à une production de virus recombinant non
occlué (c'est-à-dire, virus dépourvu de revêtement
20 protéinique codé pour par le gène de polyhédrine). Ces
virus recombinants sont ensuite utilisés pour infecter
des cellules de Spodoptera frugiperda dans lesquelles le
gène inséré est exprimé.

Le gène de VGF d'intérêt a été placé sous le
25 contrôle d'un promoteur approprié pour un système de
cellules d'insectes. Le plasmide pAc610 contient le gène
polyhédrine cloné dans un vecteur plasmide possédant un
marqueur de résistance à l'ampicilline. Des polylinkers
sont insérés dans ce gène qui sont 50 bases en aval du
30 site de départ de la transcription du gène de polyhédrine et
7 bases avant le premier ATG. Le gène de VGF est cloné
dans un site polylinker approprié afin qu'il soit sous le
contrôle du promoteur de polyhédrine. Le codon
d'initiation ATG pour la méthionine et les codons de

terminaison de la traduction sont ceux du gène de VGF; les sites de départ de la transcription et les signaux de polyadénylation sont ceux du gène de la polyhedrine.

(a) Préparation de pAc610 par SmaI-BamHI.

5 Le plasmide pAc610 a été digéré avec SmaI et BamHI.

(b) Préparation du fragment du gène de VGF par HindIII (bout franc)-BglII.

10 Le plasmide pRSV/VGF a été digéré avec HinIII et les extrémités qui dépassent en saillie ont été faites en bouts francs en utilisant le fragment de Klenow. L'ADN a ensuite été différé avec BglII et le fragment de 560 bp a été purifié sur gel.

(c) Ligation et isolement du pAc/VGF.

15 Le fragment de pAc610 obtenu par SmaI-BamHI et le fragment de 560 bp de pRSV/VGF contenant le gène de VGF ont été liés ensemble en utilisant l'ADN ligase et les transformants ont été criblés par analyse par restriction. Une construction correcte a été isolée et
20 déterminée pAc/VGF.

(d) Transfection des cellules Sf9 avec pAc/VGF.

Le plasmide pAc/VGF a été cotransfecté avec l'ADN baculovirus du type sauvage de AcNPV dans des cellules d'insectes Sf9 (Spodoptera frugiperda). Les
25 transfectants ont été criblés pour le phénotype négatif d'occlusion dans des essais en plaque. Une plaque négative d'occlusion a été isolée et après cinq opérations de purification successives de la plaque, un ensemble de virus recombinant à titre élevé a été
30 préparé.

C. Préparation de pAc/SFGF

L'expression de baculovirus du facteur de croissance du virus de fibrome de Shope.

(a) Préparation de pAc610 digéré par BamHI

Le plasmide pAc610 a été digéré avec BamHI et SmaI. Le site SmaI dans ce plasmide est localisé en aval du promoteur du gène de polyhédrine de baculovirus.

(b) Préparation d'un fragment de 430 bp par HincII-Sau3a de l'ADN du génome du virus de fibrome de Shope.

Le fragment de 3,7 kb par BglIII du fragment-C de 19 kb par BamHI de l'ADN du génome de virus du fibrome de Shope a été cloné dans un plasmide SP64 digéré par BamHI et noté SP64/3,7. Le plasmide SP64/3,7 a été digéré avec HincII et le fragment de 1kb obtenu par HincII a été purifié sur gel. Le fragment de 1kb a ensuite été digéré avec Sau3a et le fragment de 430 bp obtenu par HincII-Sau3a a été purifié sur gel.

(c) Ligation et isolement de pAc/SFGF. Le pAc610 digéré par BamHI-SmaI a été lié au fragment de 430 bp obtenu par HincII-Sau3a contenant le gène de SFGF et le mélange résultant a été transformé dans E. coli HB101 compétente. Les transformants ont été criblés par analyse par restriction et une construction correcte a été isolée et notée pAc/SFGF.

Exemple X.

Préparation de polypeptides.

A. Synthèse en phase solide de VGF et TGF.

1. Fragment de VGF de synthèse. La séquence correspondant au VGF tronqué, commençant avec la séquence DIPAIR et se terminant avec la séquence LVDY a été assemblée sur Synthétiseur Peptide d'Applied Biosystems Modèle 430A en utilisant le protocole standard. Le traitement de la résine de peptide final avec HF liquide sous des conditions standards "basse-haute" donnait le peptide déprotégé brut. Le peptide a été soumis à une chromatofocalisation sur PBE94 (Pharmacia). Le peptide partiellement purifié a été élué à pH 6,5. Un bref traitement avec de l'hydroxyde de sodium à 0,2 N enlevait les groupements protégeant la

cystéine (éthylcarbamoyl) et le peptide était oxydé en présence de glutathione oxydée et réduite. La purification par filtration sur gel et HPLC donnait un VGF grandement purifié avec la séquence suivante :

5 DIPAIRLCGPEGDGYCLHGDCIHARDIDGMYCRCSHGTYGIRCQHVVLVDY

2. α -TGF de synthèse humain et de rat.

α -TGF d'humain et de rat ont été synthétisés chimiquement, essentiellement comme décrit ci-dessus pour VGF, et sont fournis par Peninsula Labs.

10 B. Isolement de polypeptides recombinants préparés dans des cellules procaryotiques.

1. Facteurs de croissance produits dans des vecteurs sur la base de pBM en utilisant le promoteur PL le répresseur CI ts.

15 E. coli B (HB101) contenant les plasmides pBM11/NDP/facteur de croissance sont mis en culture sur Bouillon Luria à 30°C. La densité de la culture a été mesurée à 550 nm et quand la densité atteignait une absorbance de 0,7 à 0,9, la synthèse de la protéine de fusion du facteur de croissance a été induite en
20 augmentant la température à 42°C. La culture a été incubée à cette température pendant 5-20 heures, ensuite les bactéries ont été isolées par centrifugation et gelées à -70°C jusqu'à utilisation.

25 Pour l'isolement de la protéine recombinante, les cellules ont été dégelées dans un tampon contenant NaH_2PO_4 de 0,05 M, pH 7,2, NaCl 0,5 M, EDTA 0,01 M. 150 ml de tampon a été utilisé pour une préparation de 50g de bactéries. Les cellules ont été brisées par traitement
30 par ultrasons sur glace pendant 15 minutes en utilisant une sonde de 0,635 cm (1/4-pouce), 50% de pulsation à 60 watts de puissance. Après rupture des cellules, la protéine insoluble a été collectée par centrifugation à 12000 rpm dans un rotor GSA pendant 90 minutes. La

pastille contenant la protéine insoluble a été ensuite resuspendue dans 50 ml d'hydrochlorure de guanidine 6 M. Le matériau insoluble a été collecté par centrifugation pendant 2 heures à 25 000 rpm dans un rotor
5 ultracentrifugeuse Beckman type 30. Le surnageant a été collecté et stocké à -20°C jusqu'à utilisation postérieure.

La purification de la protéine de fusion a été exécutée soit sur une colonne de Séphacryl S300 ou de
10 Fractogel HW-55 équilibré avec de l'hydrochlorure de guanidine 1 M. Les fractions contenant la protéine de fusion ont été identifiées comme les fractions contenant un polypeptide ayant un poids moléculaire compatible avec le poids moléculaire du polypeptide codé par le gène de
15 synthèse comme déterminé sur un gel polyacrylamide-urée 15%.

Pour obtenir une forme active du facteur de croissance recombinant, la protéine de fusion a été soumise à réenroulement en l'incubant dans un tampon
20 Tris-HCl 50 mM, pH 8,7, contenant de l'hydrochlorure de guanidine 1 M, et de la glutathione réduite 1,25 mM et de la glutathione oxydée 0,25 mM à 4°C pendant 3-10 jours. L'activité biologique du facteur de croissance a été contrôlée par un essai de liaison à un récepteur
25 compétitif comme décrit ci-dessus (voir exemple I.C.). Quand un niveau maximum d'activité a été obtenu, la protéine a été dialysée contre de l'eau distillée et lyophilisée par séchage.

S'il était désiré d'enlever la séquence leader,
30 la protéine était clivée soit par resuspension dans l'acide formique 70% et incubation à 40°C pendant trois jours soit par incubation pendant une nuit à température ambiante dans un excès molaire de 100 fois de bromure de cyanogène. Le produit clivé était dialysé contre de l'eau
35 distillée et lyophilisée par séchage.

Pour encore purifier le facteur de croissance recombinant, le facteur de croissance était resuspendu dans de l'acétonitrile 40%, TFA 0,1% et purifié par HPLC en utilisant une colonne BioRad TSK-250. Les fractions
5 contenant le facteur de croissance étaient mises en commun et encore purifiées en utilisant HPLC en phase inverse, soit Waters μ Bondapak C-18, soit Rainin Dynamax C-8. L'éluant est un gradient linéaire de 20-40% d'acétonitrile contenant 0,1% TFA. Les fractions
10 contenant l'activité de liaison au récepteur ont été mises en commun, lyophilisées et stockées à -20°C jusqu'à utilisation.

(a) TGF et TGF modifié

(i) N/TGF

15 TGF d'humain modifié recombinant a été produit à partir du plasmide pBM11/N/TGF et contenait 33 acides aminés du gène N au N-terminal et la modification de séquence QEEK au lieu de la séquence humaine naturelle QUEDK.

20 (b) VGF modifié et tronqué.

(i) PAD/nVGFa

VGF recombinant modifié a été produit à partir du plasmide pBM11/PAD/nVGFa contenant la séquence extrême N-terminal de VGF, et les séquences modifiées GTC et
25 GYACRC au lieu de la séquence GDC et GMYGCRC de VGF naturel. Le fragment de nVGFA a été exprimé comme une protéine de fusion avec une séquence signal phosphatase alcaline modifiée au N-terminal et était tronquée à la séquence YQR C-terminal.

30 (ii) NDP/VGFA

VGF modifié recombinant a été produit à partir du plasmide pBM11/NDP/VGFA commençant à la séquence DIPAIR et se terminant à la séquence YKQR dans VGF. Il
35 avait les séquences modifiées GTC et GYACRC au lieu de la séquence GDC et GMYCRC de VGF naturel. Le fragment de

VGFa était exprimé comme une protéine de fusion avec 32 acides aminés du gène N au N-terminal, et le dipeptide acide aspartique-proline instable à l'acide.

(iii) VGFa

5 Le fragment de VGF a été préparé comme décrit dans 2.b ci-dessus, et, après clivage de la protéine de fusion par traitement acide, a été postérieurement purifié par HPLC.

(iv) NDP/VGFA

10 VGF modifié recombinant a été produit à partir du plasmide pBM11/NDP/VGFA commençant à la séquence DIPAIR et se terminant à la séquence YKQR dans VGF et ayant des séquences modifiées GTC et GYACVC au lieu de la séquence naturelle GDC et GMYCRC de VGF. Le fragment de
15 VGFA était exprimé comme une protéine de fusion avec 32 acides aminés du gène -N du N-terminal et le dipeptide acide aspartique-proline instable à l'acide.

(c) Hybrides TGF/VGF chimériques.

(i) N/TTV (TGF/TGF/VGF)

20 TTV modifié recombinant a été produit à partir de pBM11/N/TTV et contenait la séquence d'acides aminés TGF humain dans les deux tiers amino terminal du gène à l'exception de la séquence QEEK qui était altérée à partir de la séquence humaine naturelle QEDK. Le terminal
25 carboxy était dérivé de la séquence d'acides aminés de VGF et se terminait avec la séquence YQR en amont de la séquence naturelle PNT. Le fragment de TTV était exprimé pour une protéine de fusion avec 33 acides aminés du gène -N au N-terminal.

(ii) NDP/TTV

30 TTV recombinant a été produit à partir du plasmide pBM11/N/TTV et modifié comme décrit dans (a) sauf que le fragment de TTV était exprimé comme une protéine de fusion avec 32 acides aminés du gène -N au
35 N-terminal et le dipeptide acide aspartique-proline instable à l'acide.

(iii) NDP/VTV

VTV modifié recombinant a été produit à partir du plasmide pBM11/NDP/VTV et contenait la séquence d'acides aminés de TGF humain dans le domaine du milieu avec la séquence d'acides aminés QEEK remplaçant la séquence naturelle QEDK. Les domaines N-terminal et C-terminal étaient dérivés de la séquence de VGF tronquée et commençaient avec la séquence DIPAIR et terminaient avec la séquence YQR. Le fragment de VTV était exprimé comme une protéine de fusion avec 32 acides aminés du gène -N au N-terminal et au dipeptide acide aspartique-proline instable à l'acide.

(iv) NDP/TVV

VTV modifié recombinant a été produit à partir du plasmide pBM11/NDP/TVV et contenait la séquence d'acides aminés de TGF humain dans le domaine N-terminal du gène. Les domaines du milieu et C-terminal étaient dérivés de la séquence de VGF tronquée et se terminaient avec la séquence YQR. De plus, le gène de synthèse a la modification GYACVC pour GMYCRC. Le fragment de TVV a été exprimé comme une protéine de fusion avec 32 acides aminés du gène -N au N-terminal et au dipeptide acide aspartique-proline instable à l'acide.

2. Facteurs de croissance produits dans des vecteurs comprenant les promoteurs tac ou lac.

Les hôtes bactériens contenant des cassettes d'expression qui comprennent les promoteurs tac ou lac ont été cultivés de 30 à 37°C à une densité optique $A_{600}=0,2$ à $0,8$ dans soit du Bouillon LB ou un milieu défini chimiquement tel qu'un milieu M9 supplémenté avec de la thiamine et du glucose. Un antibiotique approprié a été inclus dans le milieu de croissance pour sélectionner les hôtes contenant la cassette d'expression. Les cultures bactériennes ont été induites avec une concentration d'IPTG de 100 à 1000 mM et cultivées à 30°C pendant 16 à 24 heures. Pour les cassettes d'expression

auxquelles manquaient un gène lacI, les hôtes bactériens portaient un facteur-F avec le gène lac et lacqI, tel que JM109, XL1, JM103, etc... . Pour les cassettes d'expression qui portent le gène lacI, des exemples d'hôtes bactériens sont HB101, DH1, DH5, etc... Dans le cas où l'hôte bactérien a un opéron fonctionnel lac (lac+), la cassette d'expression peut être enduite avec du lactose à 1%. Après la période d'induction, les facteurs de croissance peuvent être isolés soit du milieu, soit de la pastille de cellules.

(a) PAK/EGF

EGF d'humain produit à partir des cassettes d'expression TacPak/EGF et pTcCpt/EGF a été isolé à partir du milieu sous une forme active avec la séquence signal phosphatase alcaline enlevée. Approximativement 85% d'EGF actif a été trouvé dans le milieu, avec le reste associé dans la pastille de cellules. Ces cassettes d'expression fournissaient 4 mg/l d'EGF actif. Les cellules sont enlevées du milieu par centrifugation et le milieu est placé à travers un filtre en spirale de coupure d'Amicon SY30, 30 000 M_r et ensuite passées sur une colonne Q-Sepharose et l'EGF humain grandement purifié a été élué dans 20 mM de NaPO₄ pH 7 avec un gradient de NaCl de 0 à 0,5 M. Alternativement, les facteurs de croissance peuvent être isolés à partir de la pastille de cellules par choc osmotique ou traitement par ultrasons et purifiés par pratiquement le même processus que celui décrit ci-dessus.

(b) PAK/nVGFa

nVGFa recombinant a été produit à partir de la cassette d'expression pTCpt/nVGFa. Le nVGFa a été isolé à partir de la pastille de cellule par traitement par ultrasons et a été montré comme constituant approximativement 40% de la protéine bactérienne totale.

35 C. Isolement de VGF de SFGF produit par l'expression eucaryotique des gènes naturels.

1. Préparation de VGF dans des cellules de reins de singe.

VGF naturel a été produit par infection des cellules de reins de singe avec le virus de la vaccine. Des couches monomoléculaires de cellules de reins de singe *Cercopithecus* (BSC-1) ont été infectées avec 15 unités formant plaque (pfu) par cellule de VV purifié (souche WR cultivée dans les cellules Hela et purifiée par sédimentation par gradient de densité de sucrose (Moss, B., (1981) dans Gene Amplification and Analysis, eds. Chirickjian, J.G. et Papis, T.S. (Elsevier/North-Holland, NY), Vol. 2, pp.253-266)). Les cellules infectées ont été incubées à 37°C avec approximativement 1 ml du milieu de base de Eagle supplémenté avec du sérum foetal de veau 2% à 2×10^6 cellules. Les cellules infectées par imitation ont été traitées de manière identique. Les surnageants de culture cellulaire ont été clarifiés par une centrifugation à faible vitesse et lyophilisés. Le résidu a ensuite été resuspendu dans de l'acide acétique 1M et dialysé de façon extensive contre de l'acide acétique 0,2M. Le matériau insoluble a été enlevé par centrifugation et le surnageant a été lyophilisé et resuspendu dans 1/100 ème du volume d'origine d'acide acétique 1M et stocké à 4°C.

2. Préparation de VGF dans des cellules CHO.

VGF naturel a également été produit par transfection du plasmide pSRSV/VGF contenant le fragment du gène de VGF des cellules d'ovaires d'hamster chinois. Les cellules transfectées ont été étalées et le milieu et la pastille de cellules ont été soumis à des essais pour la présence de VGF par immunoprécipitation en utilisant un antisérum contre le peptide de VGF N-terminal.

3. Préparation de VGF dans le ver à soie en utilisant un promoteur de Baculovirus.

(a) Transformation. La cellule hôte pour AcNPV est Spodoptera frugiperda (sf9). Pour produire un ensemble de virus recombinant, le plasmide contenant VGF est mélangé avec de l'ADN d'AcNPV et transfecté dans des cellules de sf9 en utilisant la technique du phosphate de calcium. Les virus recombinant sont isolés à partir du milieu et purifiés en plaques sur des cellules sf9. Les plaques recombinantes sont identifiées par hybridation en utilisant un ADN de VGF biomarqué comme sonde.

(b) Expression Un virus recombinant a été identifié, il est ensuite étendu sur des cellules de sf9. L'ensemble de virus recombinants est ensuite utilisé pour infecter les cellules sf, les cellules sont lysées et les surnageants sont criblés pour la production de protéine de VGF, qui est purifié, comme décrit ci-dessus, pour des cellules de mammifères infectées par le virus de la vaccine.

La séquence attendue de VGF produite dans les cellules d'insecte est :

DSGNAIETTSPEITNATTDIPAIRLCGPEGDGYCLHGDCIHARDIDGMYCRCSHGYTGIR
CQHVVL MYQRSENPN TTTSYIPSPGIMLVLVGIIIIITCCLLSVYRFRTRTKLP IQDMVVP

Un site de glycosylation potentiel survient à l'asparagine dans la position 15 à partir de l'extrémité N-terminal du polypeptide.

Cette séquence de 121 résidus démarre avec la séquence N-terminal qui a été déterminée directement pour VGF purifié à partir de cellules de singe infectées par VV (c'est-à-dire, au résidu 20 le cadre de lecture ouvert de VGF) et se continue au dernier résidu du cadre de lecture ouvert. Il perd donc le peptide signal (résidu 1-19 du cadre de lecture ouvert) mais comprend la séquence transmembranaire (YIPSPGIMLVLVGIIIIITCCLLSVY).

Le poids moléculaire attendu du peptide résidu 121 est 13 304, en ne comptant pas de carbohydre. Le poids moléculaire apparent du VGF produit par baculovirus est 17 000, de manière significative plus petit que celui obtenu à partir des cellules infectées par VV, indiquant que les deux formes ont probablement été produites différemment. Il peut manquer au VGF produit par baculovirus une partie supplémentaire de la séquence N-terminale ou une partie de la séquence C-terminale, ou les deux et/ou pourrait différer dans le type et l'étendue de la glycosylation.

4. Préparation de SFGF dans le ver à soie en utilisant un promoteur de baculovirus.

(a) transfection des cellules de Sf9 avec pAc/SFGF.

Un plasmide pAc/SFGF a été cotransfecté avec l'ADN de baculovirus de type sauvage d'AcNPV dans des cellules d'insectes de Sf9 (Spodoptera frugiperda). Les transfectants ont été criblés pour un phénotype négatif d'occlusion dans des essais en plaques. Une plaque négative d'occlusion a été isolée et après 5 opérations de purification successives des plaques, un ensemble de virus recombinant à haut titre a été préparé. Le virus recombinant a été montré par analyse Southern comme contenant le gène de SFGF.

D. EGF naturel.

1. EGF naturel a été isolé à partir des glandes salivaires de souris et a été acheté à Collaborative Research.

30 Exemple XI.

Essais d'activité.

A. Essai mitogène.

Les essais de mitogénèse ont été réalisés comme suit : les fibroblastes diploïdes humains obtenus à partir d'explants de prépuces de nouveau nés ont été ensemencés à une densité de 3×10^4 cellules/puits (plaque

de 96 puits, Nunclon, Roskilde, Denmark) et ont été cultivés ensemble dans un milieu Eagle modifié de Dulbecco (GIBCO)/sérum de nouveau nés de veau 10%. Les cultures ont ensuite été placées dans un milieu contenant du sérum de veau nouveau né à 0,2%, et deux jours plus tard EGF (10 ng/ml) ou le facteur de croissance devant être testé (10 ng équivalents de EGF/ml) est ajouté. Après 8 heures, les cultures ont été marquées avec 5- $[^{125}\text{I}]$ iodo-2'-désoxyuridine (Amersham, 10 $\mu\text{Ci/ml}$, 5 Ci/mg ; 1 Ci : 37 GBq), et la quantité d'isotope incorporée dans le matériau insoluble TCA est déterminée comme décrit (Twardzik et al., Proc. Natl. Acad. Sci. USA (1985) 182 : 5300-5304).

15 B. Essai de stimulation de croissance d'une colonie Agar doux.

Une couche de base de 0,5ml de 0,5% (Agar Noble; Difco Laboratories, Détroit, Michigan) dans un milieu de croissance a été ajoutée aux plaques de cultures de tissu Costar à 24 puits. Une moitié de ml d'agar 0,3% dans le milieu de croissance contenant 1 à $1,5 \times 10^4$ de cellules/ml de cellules NRK ou une autre lignée de cellules d'intérêt de diverses concentrations de facteur devant être testée a été étendue sur la couche de base d'agar. Les plaques ont été incubées à 37°C dans une atmosphère humidifiée de CO_2 à 5% dans l'air et référencé après 7 jours par addition de 0,5 ml d'agar de 0,3% dans le milieu de croissance et contenant la même concentration du facteur devant être testé. Les colonies non fixées et non teintées ont été dénombrées. Le nombre de colonies avec plus de 6 cellules a été compté.

30 C. Essai d'inhibition de liaison au récepteur d'EGF.

Les essais de radiorécepteur ont été analysés comme suit : la liaison de EGF marqué au ^{125}I (^{125}I -EGF) à son récepteur sur des couches monomoléculaires de cellules A431 a été modifiée par le processus décrit par

Cohen et Carpenter, Proc. Natl. Acad. Sci. USA (1975) 72:1317-1321. Des cellules (1×10^3 par puits) ont été fixées sur des plaques à 24 puits (Linbro, Flow Laboratories) avec la formaline à 10% dans une solution saline tamponnée au phosphate avant essai. Les cellules fixées à la formaline ne se retirent pas des plaques aussi facilement que le font des cellules non fixées et les valeurs répliquées sont donc les plus conséquentes. Dans ces conditions d'essais, ^{125}I -EGF (1×10^{10} cpm/nmol) sature l'essai de liaison à 3nM ; les essais ont été réalisés à 10% à la valeur de saturation. Les concentrations de facteur de croissance sont exprimées comme ng équivalents de EGF/ml, c'est-à-dire la quantité requise pour produire une inhibition de la liaison de ^{125}I -EGF équivalent à celle que produisait une quantité connue d'EGF.

D. Radioimmunoessai.

Chaque milieu de réaction de 50 μl contenait ce qui suit : 20 mM de phosphate de sodium, pH 7,4, NaCl 200 mM dithio-thréitol, 40 mM, albumine de sodium de bovin 0,1%, NaN_3 0,1%, peptide marqué ^{125}I (2×10^4 cpm) correspondant aux 17 résidus carboxyl-terminal d' α -TGF (Linsley et al., Proc. Natl. Acad. Sci. USA (1985) 82:356-369), l'antisérum de dilution finale de 1:5000, et autres additifs comme spécifié. La réaction était commencée par l'addition d'antisérum et a été continué à 23°C pendant 90 minutes. Un volume égal de S. aureus fixée par la formaline à 10% (Pansorbine, Calbiochem) a été ensuite ajouté, l'incubation a été continuée pendant 30 minutes supplémentaires à 23°C. L'immunoabsorbant a été enlevé par sédimentation, et la quantité de peptide marquée au ^{125}I lié a été mesurée. La quantité de peptide lié comme corrigée pour les liaisons non spécifiques mesurés en l'absence d'anticorps (moins de 5% du total) est exprimée comme un pourcentage de liaison maximale.

E. Cicatrisation des plaies

1. Lésions du mésoderme par la chaleur

Des liaisons du mésoderme par la chaleur ont été faites sur le dos du torax de cochons femelles Yorkshire anesthésiés (30 livres) dont les dos ont été rasés et épilés avec une crème dépilatoire du commerce. Une forme en laiton (3x3 cm, 147 gm) a été équilibrée dans un bain d'eau à 70°C et a été placée en contact ferme avec la peau pendant exactement 10 secondes. La cloque résultante a ensuite été enlevée. Cinq brûlures du mésoderme ont été faites sur chaque côté de la colonne vertébrale et ont été séparées l'une de l'autre par approximativement 2,54 cm (un pouce). Les brûlures ont été traitées deux fois par jour avec approximativement 3 ml d'une crème véhicule (Silvadene^R) seule ou contenant un facteur de croissance ou ont été non traitées. Après 9 ou 10 jours de traitement avec VGF naturel, les porcs ont été anesthésiés et un escarre a été enlevé des brûlures. Des biopsies ont été prises de chaque brûlures à partir des surfaces de ré-épithélisation.

2. Lésions du mésoderme au cours d'une greffe sur un donneur.

Un petit porc de 20,5 kg âgé de 5 mois a été anesthésié avec la kétamine 20 mg/kg et du Rompum 2 mg/kg. Le dos du torax a été rasé, préparé avec de la bétadine et rincé minutieusement avec une solution saline. Une série de six emplacements sur le donneur de 5x5 cm ont été faits sur chaque côté du dos du torax avec un dermatome Padgett à 0,15 cm (60/1000 pouces) en donnant deux coups à 0,05 cm (30/1000 pouces). La thérapie topique comprenait 1 ml de solution saline dans 20 g de Silvadene distribué de façon égale entre les six plaies sur le côté gauche. Le côté droit a été traité avec 1 ml du facteur de croissance pour être testé, dans 20 g de Silvadene divisé de façon égale

entre les six plaies. Toutes les plaies ont été couvertes avec un grand pansement contre les brûlures, une bande Velpeau (marque enregistrée) ou analogues. L'animal a été anesthésié comme décrit ci-dessus, aux jours

5 post-opératoires 1, 2, 3, 4, 7, 8, 9, 10 et 11. Les pansements ont été enlevés. Les plaies ont été doucement essuyées avec de la bétadine et soigneusement rincées avec une solution saline. L'agent approprié a été appliqué et les blessures ont été recouvertes comme

10 décrit ci-dessus.

Exemple XII

Activité biologique des facteurs de croissance recombinants dans des cellules procaryotiques

A. Liaison au récepteur d'EGF

15 Cet essai détermine la capacité d'une molécule à se lier au récepteur d'EGF telle que mesurée par sa capacité à inhiber la liaison EGF à son récepteur. Tous les facteurs de croissance et facteurs de croissance chimériques, si modifiés ou tronqués, isolés à ce jour

20 étaient actifs dans l'essai d'inhibition de liaison au récepteur d'EGF. Un résumé de ces résultats est montré ci-dessous :

TABLEAU

Liaison au récepteur d'EGF et facteurs de croissance
recombinants

	peptide	cassette d'expression	Liaison au récepteur d'EGF purifié	
5				
	N/TGF - de fusion tronqué modifié	pBM11/N/TGF	95%	oui
10	PAD/nVGFa - de fusion tronqué modifié	pBM11/PAD/nVGFa	95%	oui
	NDP/VGFa - de fusion tronqué modifié	pBM11/NDP/VGFa	95%	oui
15	NDP/VGFa - de fusion tronqué modifié	pBM11/NDP/VGFa	95%	oui
	N/TTV - de fusion chimérique tronqué modifié	pBM11/N/TTV	95%	oui
20	NDP/TTV - fusion chimérique tronqué modifié	pBM11/NDP/TTV	95%	oui

	NDP/VTV - fusion chimérique tronqué modifié	pBM11/NDP/VTV	95%	oui
5	NDP/TVV - fusion chimérique tronqué modifié	pBM11/NDP/TVV	95%	oui.
	PAK/EGF	pBM11/PAK/EGF	95%	oui
	EGF	pBM11/PAK/EGF	95%	oui
	PAD/EGF	pBM11/PAD/EGF	95%	oui
10	EGF	pBM11/PAD/EGF	95%	oui
	PAK/EGF	TacPak/EGF	95%	oui
	EGF	TacPak/EGF	95%	oui
	PAK/EGF	pTCPt/EGF	95%	oui
	EGF	pTCPt/EGF	95%	oui

- 15 Une comparaison des courbes d'inhibition de la
liaison pour EGF de souris naturel et du
facteur de croissance chimérique se combinant N/TTV exprimé chez
les bactéries (un polypeptide de fusion de 32 acides
aminés N-terminaux du gène N de lambda et de l'hybride TGF/VGF
20 tronqué et modifié) suggèrait qu'il n'y avait pas de
différence dans l'activité de liaison.

B. Activité mitogène

L'activité de plusieurs facteurs de croissance purifiés a été testée et l'activité déterminée dans tous les cas était comparable à l'effet causé par EGF. Les
5 composés testés sont indiqués ci-dessous :

TABLEAU

Activité mitogène des facteurs de croissance

<u>Peptide</u>	<u>Activité mitogène*</u>
TTV - chimérique tronqué modifié	oui
10 N/TTV- fusion chimérique tronquée modifié	oui

* comme mesuré par l'incorporation de ^3H -thymidine ou ^{125}I -IdU.

C. Cicatrisation des plaies

15 1. Lésions du mésoderme

L'effet de TGF, EGF et VGF naturels ou de
synthèse, aussi bien que des facteurs de croissance
recombinants sur les lésions du mésoderme a été estimé
comme décrit dans l'exemple VE1. Le pourcentage de la
20 surface de brûlure d'origine qui avaient cicatrisées a été
mesuré par télémétrie assistée d'un ordinateur, et le
pourcentage de ré-épithélisation des plaies été
déterminé. Les plaies non traitées étaient
approximativement ré-épithélisées à 15%. Le traitement
25 avec silvadene seul ou Silvadene avec EGF provoquait
approximativement une ré-épithélisation de 50%, tandis

que le traitement avec TGF synthèse ou VGF naturel aboutissait à une ré-épithélisation approximativement à 90%. La concentration optimale pour stimuler une ré-épithélisation d'EGF était de 1-10 $\mu\text{g/ml}$, tandis que
5 TGF de synthèse et VGF naturel produisaient une réponse maximale à 0,1 $\mu\text{g/ml}$.

Des expériences similaires à celles décrites ci-dessus étaient faites pour tester l'effet dans la cicatrisation de plaies de soit TGF, une forme tronquée et
10 modifiée de VGF (VGFa), ou d'une fusion chimérique tronquée et modifiée de TGF et VGF (TTV), tous étant produits par la technologie de recombinaison dans les bactéries. Ces facteurs de croissance recombinants et les facteurs de croissances hybrides accélèrent la cicatrisation des plaies à une même
15 ampleur que TGF de synthèse ou VGF naturel, avec une concentration optimale à 0,1 $\mu\text{g/ml}$.

2. Lésions du mésoderme en cours d'une greffe sur un donneur

VGFa tronqué modifié a été testé pour sa
capacité à accélérer la cicatrisation de plaie dans un
20 modèle de greffe sur donneur. Le régime du traitement était comme décrit ci-dessus (voir Exemple I) en utilisant 1 ml VGFa, 5 $\mu\text{g/ml}$, dans 20g de Silvadene. Des photographies ont été prises sur une base journalière. le résumé des résultats est fourni ci-dessous :

TABLEAUEffet de VGFa recombinant sur la cicatrisation des plaies

		<u>condition des plaies</u>			
<u>traitement*</u>	<u>POD**7</u>	<u>POD 8</u>	<u>POD 9</u>	<u>POD 10</u>	
5	Solution saline	très ouvertes	ouvertes avec des cicatrisations	princi- palement cicatrisées	cicatrisées
10	VGFa - modifié, tronqué	ouvertes épithéli- sation apparente	princi- palement cicatrisées	cicatrisées	cicatrisées

* Silvadene comme véhicule

** POD = jour post-opératoire

- 15 Le VGFa modifié et tronqué accélérerait la cicatrisation de la plaie comparée au véhicule contrôle.
- Les photographies (non fournies) à POD 8 montraient des différences sensibles entre la solution saline et VGFa.

EXEMPLE XIIIActivité biologique de VGF préparé dans des cellules eucaryotiquesA. VGF préparé dans cellules de reins de singes (BSC-1)

5 1. Le milieu dérivé de cellules de BSC-1
24 heures après infection avec VV a été testé pour la
présence de matériau qui pourrait entrer en compétition avec
125I-EGF pour une liaison aux cellules de carcinome de kiste
épidermique humain riches en récepteurs d'EGF (A431). Les
10 cellules infectées par VV libéraient une activité puissante
qui entraînait en compétition avec EGF, cette activité est
désignée VGF. Le volume du milieu de contrôle de cultures de
contrôle BSC-1 par imitation ou simulation infectées
contenait une activité minimale en compétition avec EGF.

15 En contrôlant la production de VGF, examinée au
temps le plus tôt, deux heures après infection, des niveaux
améliorés de VGF étaient observés dans le milieu de culture.
En douze heures, des quantités maximales de cette activité
étaient trouvées dans le surnageant de culture avec
20 seulement une légère augmentation notée à 24 heures.

Le taux de production de VGF a été trouvé comme
étant une fonction de la multiplicité d'infection par le
virus comme démontré par le tableau suivant.

TABLEAU

Effet de la multiplicité d'infection sur la libération de VGF

	<u>Multiplicité du Virus</u>	<u>VGF libéré</u>
	<u>pfu par cellule</u>	<u>ng eq. d'EGF par ml</u>
5	0	---
	10	2,7
	20	4,5
	40	6,1
10	80	10,1

Les cultures de BSC-1 ont été infectées au rapport pfu-à-cellules et les surnageants (approximativement 1ml./2x10⁶ cellules) ont été récoltés 24 heures après infection. Des échantillons de chacun ont été acidifiés, lyophilisés, et testés dans des essais à radiorécepteurs doubles pour EGF comme décrit (Delarco et Torado, Proc. Natl. Acad. Sci. USA (1978) 75:401-405). ng eq., équivalents nanogrammes.

2. Purification partielle de VGF

EGF avec activité de compétition trouvé dans des cellules de BSC-1 infectées par VV était partiellement purifié à partir de surnageants de culture extraits à l'acide 24 heures après infection comme décrit ci-dessus. Les polypeptides solubilisés à l'acide (10,5 mg) à partir des surnageants étaient appliqués à une colonne Bio-Gel P10 équilibrée avec de l'acide acétique 1M et des échantillons de chaque fraction étaient testés pour une activité en compétition

avec EGF. Le pic actif principal (fraction 42) s'élueait légèrement après le marqueur d'anhydrase carbonique M_r de 29.000 avec un M_r apparent de 25.000. Le poids moléculaire a été confirmé en utilisant, liées en tandem, des colonnes de calibration Bio-Sil TSK 250 HPLC, avec toute l'activité éluante comme un pic principal dans la région du marqueur de protéine M_r 25.000. Un microgramme de VGF partiellement purifié était équivalent à 90ng d'EGF dans l'essai de compétition au récepteur.

10 3. Purification supplémentaire de VGF

Les fractions mises en commun à partir de la colonne de filtration sur gel (25-35) ont été concentrées par centrifugation sous vide, resuspendues dans l'acide trifluoroacétique à 0,05% (TFA), clarifiées et injectées dans une colonne μ Bondapak C_{18} de 3,9 x 30cm (Waters, Milford, MA). Les peptides ont été élués gradient linéaire de 20-60% d'acétonitrile dans TFA 0,05% à une vitesse d'écoulement de 1,0ml/mn à 22°C. Les aliquotes de chaque fraction ont été testées dans un essai à récepteur pour une activité de compétition avec EGF. Le peptide correspondant à l'activité du pic a été collecté et dilué avec TFA 0,05% et réinjecté dans une colonne μ Bondapak et élué en utilisant des conditions isocratiques. La concentration d'acétonitrile était environ 22 à 25%.

25 La comparaison des niveaux de VGF produits par des cellules BSC-1 infectées avec 15pfu de VV par cellule avec celui de TGF produit des cellules transformées par un rétrovirus est montré dans le tableau suivant.

TABLEAU

Comparaison de l'activité biologique de VGF avec α TGF et EGF

5	Liaison au récepteur d'EGF ²			induction de synthèse d'ADN ³	Stimulation de la croissance des cellules ancrage-indépendant ⁴
	facteur de croissance ¹	ng eq. EGF/ml du milieu	[¹²⁵ I] IdU incorporé (cpm/plaque)	agar mou	colonies/plaque
10	aucun	—	1,779	20	
	VGF	2,3	3,760	108	
	TGF—	0,2	NT	294	
	EGF	—	8,482	346	

NT = non testé.

- 15 1 VGF (90 ng équivalents (eq) de EGF/ μ g de protéine) a été purifié par filtration sur gel suivi par l'élution à partir d'une colonne μ Bondapak C₁₈ (Waters Associates). α -TGF a été purifié à partir de cellules embryonnaires de rats de Fisher transformées par le virus du sarcome des
- 20 félins de Snyder-Theilen, comme décrit dans Marquardt et al., Proc. Natl. Acad. Sci.. USA (1983) 80:4684-4688; EGF a été purifié à partir de la glande submaxillaire de souris (Cohen et al., J. Biol. Chem. (1980) 255:41834 -41842).

2 La quantification des équivalents d'EGF a été basée sur une courbe de compétitions ou la liaison ^{125}I -EGF standard comme décrit.

5 3 Les essais des mitogénèses ont été comme décrits ; les cultures au repos de fibroblastes d'humains diploïdes ont été traitées avec 10ng de EGF/ml purifié ou le même nombre équivalent d'EGF de VGF/ml. Les valeurs pour ^{125}I iodo-désoxyuridine (^{125}I IdU) incorporée
10 représentant la moyenne des déterminations en trois exemplaires.

4 Le nombre de colonies sur agar mou représente le nombre moyen de colonies contenant un minimum de 20 cellules NRK/ six champs à puissance faible au hasard
10 jours après ensemencement ($1,5 \times 10^4$ cellules/ml) avec EGF purifié (5ng/ml) ou le même nombre d'équivalents
15 EGF de VGF/ml et 2,0 ng de TGF- β /ml purifié à partir de plaquettes humaines comme décrit dans Assoian et al., J.Biol.Chem. (1983) 258:7155-7160. Les plaques de cellules NRK traitées avec TGF- β seul ci-dessus ne formaient pas de colonies.

20 B. VGF préparé dans des cellules de CHO

L'activité biologique de VGF préparé dans des cellules de CHO a été évaluée en utilisant un essai de liaison au récepteur d'EGF. Le milieu de croissance de la culture a été utilisé sans autre purification. Le facteur
25 de croissance se liait au facteur d'EGF.

C. EGF préparé dans le ver à soie

L'activité biologique de VGF partiellement purifié (approximativement 50%) préparé dans le ver à soie a été évalué en utilisant un essai de liaison au
30 récepteur d'EGF et a été montré comme se liant au récepteur d'EGF.

TABLEAU

pourcentage d'épithélisation de plaies suivant un traitement avec
facteur de croissance

<u>porc 1</u>					
5	(9 jours après brûlure)			Naturel	
				<u>VGf ug/ml</u>	
	Côté droit	<u>Silvadene</u>	<u>non traité</u>	<u>0,1</u>	<u>0,1</u>
		15%	0%	70%	65%
				<u>h-EGF ug/ml</u>	
10	Côté gauche	<u>Silvadene</u>	<u>non traité</u>	<u>0,1</u>	<u>0,1</u>
		75%	55%	60%	70%
<u>Porc 2</u>					
(10 jours après brûlure)					
15				VGf* ug/ml	
	Côté droit	<u>Silvadene</u>	<u>non traité</u>	<u>0,1</u>	<u>0,5</u>
		50%	0%	95%	95%
	Côté gauche	<u>Silvadene</u>	<u>non traité</u>	<u>0,1</u>	<u>0,5</u>
		50%	0%	60%	75%
20	<u>Porc 3</u>			<u>TGF de rat</u>	
	(9 jours après brûlure)			<u>ug/ml</u>	
				<u>0,1</u>	<u>1,0</u>
25	Côté droit	<u>Silvadene</u>	<u>non traité</u>	<u>0,1</u>	<u>1,0</u>
		20%	15%	90%	65%

Côté gauche			TGF d'humain $\mu\text{g/ml}$		
	<u>Silvadene</u>	<u>non traité</u>	<u>0,1</u>	<u>1,0</u>	<u>10</u>
	25%	5%	90%	85%	65%

* VGF : VGF naturel préparé à partir des cellules affectées par le virus de la vaccine.

5 D. Comparaison immunologique VGF et TGF

Dans le radioimmunoessai décrit ci-dessus, un déplacement de 50% d'antigène à partir de l'anticorps au 17 acides aminés du carboxy N-terminal de la molécule de TGF- α de rat a été observé à une concentration d'antigène d'approximativement 0,2 à 0,3 équivalents ng d'EGF, où α -EGF ^{125}I -marqué rivalise avec α -TGF. Quand VGF a été testé à des concentrations équivalentes, aucune compétition n'a été observée. Dans un radioimmunoessai de compétition pour des préparations d'EGF et de VGF natifs, même quand testés à 50ng d'équivalents par ml d'EGF, il apparaissait un déplacement minimal (<10%) de ^{125}I -EGF à partir d'un anticorps polyclonal à EGF natif.

A partir des résultats ci-dessus, il est évident que VGF est un agent d'épithélisation puissant en le comparant de manière satisfaisante avec d'autres facteurs de croissance qui ont été antérieurement testés et montrés comme ayant une activité mitogène. Les polypeptides du sujet peuvent être utilisés avec divers hôtes pour induire une réponse immunogène, améliorer la prolifération cellulaire, cicatriser les plaies ou analogues. Pour des plaies, une épithélisation et vascularisation est observée aussi bien qu'un rétablissement rapide de la résistance de la plaie, c'est-à-dire une résistance de la séparation ou déchirement. Les hôtes comprennent les mammifères, tels que rongeurs, animaux domestiques, primates et humains.

- Selon l'invention du sujet, de nouvelles compositions sont fournies ayant une aptitude s'étendant largement dans une variété de domaines, tels que des diagnostics, effets in vivo et in vitro sur des cellules, des actions comme mitogènes, des additifs dans des milieux nutritifs, des utilisations comme agonistes et antagonistes à TGF et EGF, des actions comme immunogènes, des actions comme agents thérapeutiques, des améliorations dans la cicatrisation de plaies, et analogues. De plus, en fournissant un petit polypeptide ayant une activité de liaison,
- 10 l'oligopolypeptide peut être utilisé par lui-même ou en combinaison avec d'autres polypeptides, fusionné à d'autres polypeptides pour varier les propriétés des autres polypeptides, aboutissant à la liaison du
- 15 polypeptide aux cellules ayant des récepteurs de facteurs croissance. Donc on peut lier de manière réversible divers polypeptides aux cellules ayant des récepteurs de facteurs de croissance, affectant les propriétés des cellules dans des voies prédéterminées.
- 20 Toutes les publications et demandes de brevet mentionnées dans cette description sont indicatives et sont à la portée de l'homme du métier. Toutes les publications et demandes de brevet sont incorporées ici à titre de référence.
- 25 Bien entendu il sera évident pour un homme du métier que certains changements et modifications peuvent être faits sans modifier l'esprit ou la portée de la présente invention.

RE V E N D I C A T I O N S

1. Séquence d'ADN codant pour un polypeptide, caractérisée en ce qu'il est capable de se lier à un récepteur d'EGF, en ce qu'il a sensiblement la même structure en boucle qu'un facteur de croissance se produisant naturellement capable de se lier à un récepteur d'EGF, ou ligand naturel, et en ce qu'il a au moins environ une homologie de 30% avec un ligand naturel où ledit polypeptide comprend trois domaines, chacun de ceux-ci est une séquence d'acides aminés ayant au moins une homologie de 30% avec les fragments des ligands naturels identiques ou différents, et ayant la même position relative que lesdits fragments desdits ligands naturels, lesdits domaines étant définis comme :
- 15 un premier domaine d'au moins environ onze acides aminés N-terminaux commençant environ à un acide aminé aa⁻⁶ à aa¹ et se terminant environ à un acide aminé aa¹¹ à aa¹⁵ ;
- 20 un second domaine d'au moins environ onze acides aminés d'un fragment central commençant environ à un acide aminé aa¹¹ à aa¹⁵ et se terminant à environ un acide aminé aa²⁵ à aa²⁹ ;
- 25 un troisième domaine d'au moins environ quatorze acides aminés C-terminaux commençant à environ un acide aminé aa²⁵ à aa²⁹ et se terminant à environ aa⁴⁰ à aa⁵³ ; et
- 30 facultativement ladite séquence d'ADN code de plus pour un quatrième domaine d'au moins environ six acides aminés de N-terminal extrême commençant à environ aa⁻⁷ à aa⁻¹ et se terminant à environ un acide aminé aa⁻⁴⁵ à aa⁻⁷, avec facultativement une séquence d'acides aminés capable d'être sélectivement clivée et insérée au carboxy terminal dudit quatrième domaine, unie au terminal amino dudit premier domaine à la condition que lorsque ledit

polypeptide a une séquence d'acides aminés identique à un ligand naturel, ladite séquence d'ADN codant pour ledit quatrième domaine soit présente et code pour d'autres que les acides aminés de N-terminal extrême du ligand naturel.

5

2. Séquence d'ADN selon la revendication 1, caractérisée en ce que ledit ligand naturel est VGF, alpha-TGF, MGF, EGF, SFGF ou amphiréguline.

10 3. Séquence d'ADN selon la revendication 1, caractérisée en ce que la séquence d'ADN code en plus pour une séquence leader, avec facultativement une séquence d'acides aminés capable d'être sélectivement clivée et insérée au carboxy terminal de ladite séquence leader, unie à l' amino terminal dudit premier domaine, ou
15 quand ledit quatrième domaine est présent, à l' amino terminal dudit quatrième domaine.

4. Séquence d'ADN selon la revendication 3, caractérisée en ce que ladite séquence leader est un fragment N-terminal d'une protéine-N de lambda, d'une
20 protéine cro de lambda ou de l'amphiréguline

5. Séquence d'ADN selon la revendication 3, caractérisée en ce que ladite séquence leader est une séquence signal ou son mutant.

25 6. Séquence d'ADN selon la revendication 5, caractérisée en ce que ladite séquence signal est une séquence signal phosphatase alcaline, une séquence signal beta-TGF de simien ou une séquence signal d'une toxine tueuse de K. Lactis.

30 7. Séquence d'ADN selon l'une quelconque des revendications 1 à 6, caractérisée en ce que au moins un desdits domaines est au moins homologue à environ 30% à un fragment de VGF, alpha-TGF, EGF ou SFGF.

35 8. Séquence d'ADN selon l'une des revendications 1, 4 et 6, caractérisée en ce que ladite séquence d'ADN code pour VGF.

9. Séquence d'ADN selon l'une quelconque des revendications 1, 4 et 6, caractérisé en ce que ladite séquence d'ADN code pour TGF.

10. Séquence d'ADN selon l'une des
5 revendications 1, 4 et 6, caractérisé en ce que ladite séquence d'ADN code pour EGF.

11. Séquence d'ADN selon la revendication 1, caractérisé en ce que la séquence d'ADN code pour SFGE.

12. Séquence d'ADN selon la revendication 1,
10 caractérisée en ce que ladite séquence comprend au moins une altération de la séquence codant pour un ligand naturel aboutissant aux changements suivants dans la séquence d'acides aminés : HGD en HGT ; GMYC en GYAC ; RCS en MCS ; CLHGDC en CLHGTC ; et GMYCRC en GYACVC.

13. Séquence d'ADN codant pour un polypeptide
15 d'intérêt et pour une séquence d'environ 8 à 45 acides aminés N-terminaux, comprenant des modifications de celle-ci à partir d'une protéine-N d'un bactériophage lambda ou d'une protéine de cro, de l'amphiréguline, du gène 32 de T4 ou une séquence signal phosphatase
20 alcaline, jointe au terminal amino dudit polypeptide d'intérêt, facultativement via une séquence d'acides aminés capable d'être sélectivement clivée, ledit polypeptide d'intérêt comprenant trois domaines, la séquence de
25 chaque domaine correspondant au moins sensiblement à la séquence d'un fragment d'un facteur de croissance se produisant naturellement capable de se lier à un récepteur d'EGF, ou ligand naturel, et ayant la même position relative que lesdits fragments dudit ligand naturel, lesdits domaines étant
30 à au moins environ 11 acides aminés N-terminaux commençant environ à un acide aminé aa⁻⁶ à aa¹ et se terminant à environ un acide aminé aa¹¹ à aa¹⁵ ;

un second domaine d'au moins environ 11 acides
aminés d'un fragment central commençant à environ un
35 aminé aa¹¹ à aa¹⁵ et se terminant à environ un acide aminé aa²⁵ à aa²⁹ ;

un troisième domaine d'au moins environ 14 acides aminés de l'acide C-terminal commençant à environ un acide aminé aa²⁵ à aa²⁹ et se terminant à environ aa⁴⁰ à aa⁴⁷.

5 14. Séquence d'ADN selon la revendication 13, caractérisée en ce que ladite séquence leader est environ 32 ou 33 acides aminés N-terminaux du gène-N.

10 15. Séquence d'ADN selon la revendication 14, caractérisée en ce que ledit polypeptide d'intérêt est alpha-TGF, VGF, VGFA, VGFa ou EGF.

16. Séquence d'ADN selon la revendication 13, caractérisée en ce que ladite séquence leader est une séquence signal de phosphatase alcaline ou une séquence signal phosphatase alcaline modifiée.

15 17. Séquence d'ADN selon la revendication 16, caractérisée en ce que ledit polypeptide d'intérêt est VGFa, nVGFa, ou EGF.

20 18. Séquence d'ADN selon la revendication 14, caractérisée en ce que ledit polypeptide d'intérêt a la séquence d'acides aminés suivante :

M V V S H F N D C P D S H T Q F C F H G
T C R F L V Q E E K P A C V C S H G Y T
G I R C Q H V V L V D Y Q R.

25 19. Séquence d'ADN selon la revendication 14, caractérisée en ce que ledit polypeptide d'intérêt a la séquence d'acides aminés suivante :

M D I P A I R L C G P E G D G Y C L H G
T C R F L V Q E E K P A C V C S H G Y T
G I R C Q H V V L V D Y Q R.

30 20. Séquence d'ADN selon la revendication 14, caractérisée en ce que ledit polypeptide a la séquence d'acides aminés suivante :

M V V S H F N D C P D S H T Q F C F H G T
 C I H A R D I D G Y A C V C S H G Y T G I
 R C Q H V V L V D Y Q R.

21. Séquence d'ADN selon la revendication 13,
 5 caractérisée en ce que ladite séquence leader est
 environ 21 acides aminés N-terminaux de ladite protéine de cro.

22. Séquence d'ADN selon la revendication 19,
 caractérisée en ce que ledit polypeptide d'intrêt a la
 séquence d'acides aminés suivante :

10 M V V S H F N D C P D S H T Q F C F H G
 T C R F L V Q E E K P A C V C S H G Y T
 G I R C Q H V V L V D Y Q R.

23. Séquence d'ADN selon la revendication 13,
 caractérisée en ce que ladite séquence leader est d'environ 8
 15 acides aminés N-terminaux du gène 32 de T4

24. Séquence d'ADN selon la revendication 23,
 caractérisée en ce que ledit polypeptide d'intérêt a la
 séquence d'acides aminés suivante :

20 V V S H F N D C P D S H T Q F C F H G T
 C R F L V Q E E K P A C V C S H G Y T G
 I R C Q H V V L V D Y Q R.

25. Cassette d'expression qui comprend dans la
 direction de la transcription, une région de régulation
 d'initiation de la transcription et de la traduction
 25 fonctionnelle dans une cellule hôte ; une séquence d'ADN
 codant pour un polypeptide, caractérisée en ce qu'il est capable
 de se lier à un récepteur d'EGF, en ce qu'il a sensiblement la
 même structure en boucle qu'un facteur de croissance se
 produisant naturellement capable de se lier à un
 30 récepteur d'EGF, ou ligand naturel, et en ce qu'il y a au moins
 environ une homologie de 30% avec un ligand naturel, où

ledit polypeptide comprend trois domaines, chacun de ceux-ci est une séquence d'acides aminés ayant au moins environ une homologie de 30% aux fragments des ligands naturels identiques ou différents, et ayant la même position relative que lesdits fragments desdits ligands naturels, lesdits domaines étant définis comme :

5 un premier domaine d'au moins environ 11 acides aminés N-terminaux commençant à environ un acide aminé aa⁻⁶ à aa¹ et se terminant à environ un acide aminé aa¹¹ à aa¹⁵ ;

10 un second domaine d'au moins environ 11 acides aminés du fragment central commençant à environ un acide aminé aa¹¹ à aa¹⁵ et se terminant à environ un acide aminé aa²⁵ à aa²⁹ ;

15 un troisième domaine d'au moins environ 14 acides aminés de C-terminal amino commençant à environ un acide aminé aa²⁵ à aa²⁹ et se terminant à environ aa⁴⁰ à aa⁵³ ; et

20 facultativement ladite séquence d'ADN code de plus pour un quatrième domaine d'au moins environ 6 acides aminés de N-terminal extrême commençant à environ à aa⁻⁷ à aa⁻¹ et se terminant à environ un acide aminé aa⁻⁴⁵ à aa⁻⁷, avec facultativement une séquence d'acides aminés capable d'être sélectivement clivée et insérée au carboxy terminal dudit quatrième domaine, jointe au terminal amino dudit premier domaine à la condition que lorsque ledit polypeptide à une séquence acides aminés identique à un ligand naturel, une séquence d'ADN codant pour ledit quatrième domaine soit présente et code pour d'autres que les acides aminés de N-terminal extrême du ligand naturel et une région de régulation de la terminaison de la transcription et de la traduction fonctionnelle dans ladite cellule hôte.

26. Cellule hôte comprenant une cassette
35 d'expression qui comprend dans la direction de transcription, une région de régulation de l'initiation de la

transcription et de la traduction fonctionnelle dans une cellule hôte ; une séquence d'ADN codant pour un polypeptide caractérisée en ce qu'il est capable de se lier à un récepteur d'EGF, en ce qu'il a sensiblement la même structure en boucle qu'un facteur de croissance se produisant naturellement capable de se lier à un récepteur d'EGF, ou ligand naturel, en ce qu'il a au moins environ une homologie d'environ 30% avec un ligand naturel, où ledit polypeptide comprend trois domaines, chacun de ceux-ci est une séquence d'acides aminés ayant au moins environ une homologie de 30% aux fragments des ligands naturels identiques ou différents, et ayant la même position relative que lesdits fragments desdits ligands naturels, lesdits domaines étant définis comme :

un premier domaine d'au moins environ 11 acides aminés N-terminaux commençant à environ un acide aminé aa⁻⁶ à aa¹ et se terminant environ à un acide aminé aa¹¹ à aa¹⁵ ;

un second domaine d'au moins environ 11 acides aminés du fragment central commençant à environ un acide aminé aa¹¹ à aa¹⁵ et se terminant à environ un acide aminé aa²⁵ à aa²⁹ ;

un troisième domaine d'au moins environ 14 acides aminés du C-terminal amino commençant à environ un acide aminé aa²⁵ à aa²⁹ et se terminant à environ aa⁴⁰ à aa⁵³ ; et

facultativement ladite séquence d'ADN code en plus pour un quatrième domaine d'au moins environ 6 acides aminés de N-terminal extrême commençant à environ aa⁻⁷ à aa⁻¹ et se terminant environ à un acide aminé aa⁻⁴⁵ à aa⁻⁷, avec facultativement une séquence d'acides aminés capable d'être sélectivement clivée et insérée au terminal carboxy du quatrième domaine, jointe au terminal amino dudit premier domaine à la condition que lorsque ledit polypeptide a une séquence d'acides aminés identique à un ligand naturel, une séquence d'ADN codant pour ledit

quatrième domaine soit présente et code pour d'autres que les acides aminés de N-terminal extrême du ligand naturel et une région de régulation de la terminaison de la transcription et de la traduction dans ladite cellule hôte.

27. Procédé pour produire un polypeptide caractérisé en ce qu'il est capable de se lier à un récepteur d'EGF, en ce qu'il a sensiblement la même structure en boucle et qu'un facteur de croissance se produisant naturellement capable de se lier à un récepteur d'EGF, ou ligand naturel, et en ce qu'il a au moins environ une homologie de 30% avec un ligand naturel au récepteur d'EGF, où ledit polypeptide comprend trois domaines, chacun de ceux-ci est une séquence d'acides aminés ayant au moins environ une homologie de 30% aux fragments des ligands naturels identiques ou différents et ayant la même position relative que lesdits fragments desdits ligands naturels, lesdits domaines étant définis comme :

un premier domaine d'au moins environ 11 acides aminés N-terminaux commençant environ à un acide aminé aa⁻⁶ à aa¹ et se terminant environ à un acide aminé aa¹¹ à aa¹⁵ ;

un second domaine d'au moins environ 11 acides aminés du fragment central commençant à environ un acide aminé aa¹¹ à aa¹⁵ et se terminant à environ un acide aminé aa²⁵ à aa²⁹ ;

un troisième domaine d'au moins environ 14 acides aminés du C-terminal amino commençant à environ un acide aminé aa²⁵ à aa²⁹ et se terminant à environ aa⁴⁰ à aa⁵³ ; et

facultativement un quatrième domaine d'au moins environ 6 acides aminés de N-terminal extrême commençant à environ aa⁻⁷ à aa⁻¹ et se terminant environ à un acide aminé aa⁻⁴⁵ à aa⁻⁷, avec facultativement une séquence d'acides aminés capable d'être sélectivement clivée et insérée au carboxy terminal dudit quatrième

domaine, jointe au terminal amino dudit premier domaine à condition que lorsque le polypeptide a une séquence d'acides aminés identique à un ligand naturel, ledit quatrième domaine soit présent et code pour d'autres que
 5 les acides aminés de N-terminal extrême du ligand naturel , et une région de régulation de la terminaison de la transcription et de la traduction fonctionnelle dans ladite cellule hôte, ledit procédé consistant à :
 cultiver une cellule hôte contenant une
 10 cassette d'expression qui comprend, dans la direction de la transcription, une région de régulation de l'initiation de la transcription et de la traduction fonctionnelle dans ladite cellule hôte, une séquence d'ADN codant pour ledit polypeptide et une région de
 15 régulation de la terminaison de la transcription et de la traduction, dans un milieu de culture approprié de façon à ce que ledit polypeptide soit exprimé ; et à isoler ledit polypeptide.

28. Plasme comprenant les domaines
 20 fonctionnels de pBM11/NDP/EGF ; pBM11/PAD/EGF ; pBM11/PAK/EGF ; pBM11/N/TGF ; pBM11/NDP/VGFa ; pBM11/NDP/VGFa ; pBM11/PAD/nVGfA ; pBM11/PAK/nVGfA ; pRSV/VGF ; pAc/SFGF ; pAc/VGF ; ptac/TTV ; TacPak/EGF ; pTCPt/EGF ; pTCPt/nVGfA ; pBM11/N/TTV ; pBM11/NDP/TTV ;
 25 pBM11/NDP/VTV ; ou pM16/NDP/TVV.

29. Polypeptide caractérisé en ce qu'il est capable de se lier à un récepteur d'EGF, en ce qu'il a sensiblement la même structure en boucle qu'un facteur de croissance se produisant naturellement capable de se
 30 lier à un récepteur d'EGF, ou ligand naturel, en ce qu'il a au moins une homologie d'environ 30% avec un ligand naturel au récepteur d'EGF, où ledit polypeptide comprend trois domaines, chacun de ceux-ci est une séquence d'acides aminés ayant au moins environ une homologie de 30% aux
 35 fragments des ligands naturels identiques ou différents et

ayant la même position relative que lesdits fragments
desdits ligands naturels, lesdits domaines étant définis
comme :

- 5 un premier domaine d'au moins environ 11 acides
aminés N-terminaux commençant environ à un acide aminé
aa⁻⁶ à aa¹ et se terminant environ à un acide aminé aa¹¹
à aa¹⁵ ;
- 10 un second domaine d'au moins environ 11 acides
aminés du fragment central commençant à environ un acide
aminé aa¹¹ à aa¹⁵ et se terminant à environ un acide
aminé aa²⁵ à aa²⁹ ;
- 15 un troisième domaine d'au moins environ 14
acides aminés de l' amino C-terminal commençant à environ un
acide aminé aa²⁵ à aa²⁹ et se terminant à environ aa⁴⁰ à
aa⁵³ ; et
- 20 facultativement un quatrième domaine d'au moins
environ 6 acides aminés de N-terminal extrême commençant à
environ aa⁻⁷ à aa⁻¹ et se terminant environ à un acide
aminé aa⁻⁴⁵ à aa⁻⁷, avec facultativement une séquence
d'acides aminés capable d'être sélectivement clivée et au
insérée au terminal carboxy du quatrième domaine, unie
terminal amino dudit premier domaine à la condition que
lorsque ledit polypeptide a une séquence d'acides aminés
identique aux ligands naturels, ledit quatrième domaine
25 soit présent et a une séquence autre que celle des acides
aminés de N-terminal extrême du ligand naturel.
30. Polypeptide comprenant la séquence d'acides
aminés suivante ou son fragment d'au moins 37 acides
aminés comprenant au moins environ 1 à 37 acides aminés :

aa -24 -20 -15 -10 -5 -1
 D S G N A I E T T S P E I T N A T T D I P A I R

1 5 10 15 20
 L C G P E G D G Y C L H X¹ G D C I H X² A R D

5 25 30 35 40
 I D G X³ M Y C R C S H G Y T G I R C Q H V V

45 50 53
 L V D Y Q R S E N P N T

où X¹ est une liaison ou un acide aminé ; X² est une
 10 liaison ou de 1 à 2 acides aminés ; X³ est une liaison
 ou de 1 à 4 acides aminés ; et X¹, X² et X³ peuvent être
 identiques ou différents.

31. Polypeptide selon la revendication 30,
 caractérisé en ce qu'au moins un dudit aa¹⁴ est la
 15 thréonine, ledit aa²⁴ est tyrosine, ledit aa²⁵ est
 alanine ; et ledit aa²⁷ est valine.

32. Polypeptide selon la revendication 30,
 caractérisé en ce que ladite séquence d'acides aminés
 comprend aa⁻⁶ à aa⁴⁷.

20 33. Polypeptide selon la revendication 32,
 caractérisé en ce que aa¹⁴ est thréonine, ledit aa²⁴ est
 tyrosine et ledit aa²⁵ est alanine.

34. Polypeptide selon la revendication 33,
 caractérisé en ce que ledit aa⁻⁶ à aa¹³ de ladite
 25 séquence d'acides aminés est remplacé par ce qui suit :

aa -6 -1 1 5 10 13
 V V S H F N D C P D S H T Q F C F H G

35. Polypeptide selon la revendication 33,
 caractérisé en ce que ledit aa¹⁴ à aa²⁷ de ladite séquence
 30 d'acides aminés sont remplacés par ce qui suit :

aa 14 20 25 27
T C R F L V Q E D K P A C V

5 36. Polypeptide selon la revendication 35,
caractérisé en ce que ledit aa²² est remplacé avec de
l'acide glutamique.

37. Séquence d'ADN codant pour un polypeptide
ayant une séquence selon l'une quelconque des
revendications 30 à 36.

1/2

FAMILLE DES FACTEURS DE CROISSANCE

MGF Facteur de croissance du virus du myxome
 SFGF Facteur de croissance du virus du fibrome de Shope
 VGF Facteur de croissance de la vaccine
 hTGF TGF d'humain
 rTGF TGF de rat
 hEGF EGF d'humain
 mEGF EGF de souris
 rEGF EGF de rat
 gpEGF EGF du porc de Guinée
 amphi amphiréguline

1*

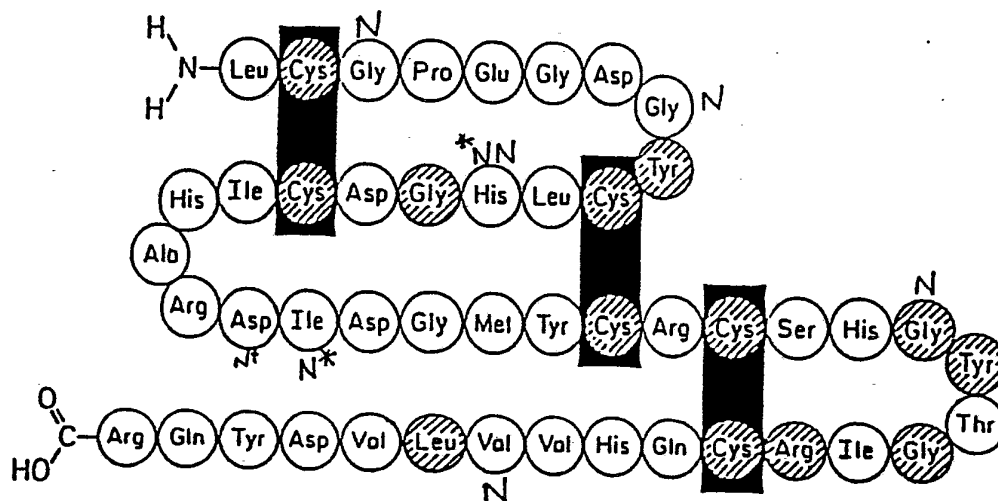
MGF	MVPRDLVATLLCANCIVQA-----TMPSLDNYLYIIKRIKL
SFGF	NATRNLVASLLCINAVHA-----MNDYLYIVKHVKV
VGF	MSMKYLMLLFAAMIIISFA-DSGNAIETTSPEITNATTDIPAIRL
hTGF	MVPSAGQLALPALGIVLAACQAEENSTSPLSADPPVAAAVVSHFND
rTGF	MVPAAGQLALLALGILVAVCAEENSTPPLSADSPVAAAVVSHFNK
hEGF	NSDSE
mEGF	NSYPS
rEGF	NSYPS
gpEGF	QDAPG
amphi	SVRVEQVVKKPPQDKTESENTSOKPKRKKKGGKNGKNRRNRKKKNP

	2	5	10	15	20	25	30	35	40	45																																																	
MGF	C	N	D	D	Y	K	N	Y	C	L	N	N	G	T	C	-	F	T	V	A	L	N	N	V	S	L	N	P	F	C	A	C	H	I	N	Y	V	G	S	R	C	O	F	I	N	L	I	T	I	K	*								
SFGF	C	N	H	D	Y	E	N	Y	C	L	N	N	G	T	C	-	F	T	I	A	L	D	N	V	S	I	T	P	F	C	V	C	R	I	N	Y	E	G	S	R	C	O	F	I	N	L	V	T	Y	*									
VGF	C	G	P	E	G	D	G	Y	C	L	H	-	G	D	C	I	H	-	-	A	R	D	I	D	-	-	G	M	Y	C	R	C	S	H	G	Y	T	G	I	R	C	H	N	V	L	V	D	Y	R	S	E	N	P	N	T	T	T	S	
hTGF	C	P	D	S	H	T	D	F	C	F	H	-	G	T	C	R	F	L	V	Q	E	D	K	-	-	-	-	P	A	C	V	C	H	S	G	Y	V	G	A	R	C	E	H	A	D	L	L	A	V	V	A	S	Q	K	K	Q	A	I	T
rTGF	C	P	D	S	H	T	D	F	C	F	H	-	G	T	C	R	F	L	V	Q	E	E	K	-	-	-	-	P	A	C	V	C	H	S	G	Y	V	G	A	R	C	E	H	A	D	L	L	A	V	V	A	S	Q	K	K	Q	A	I	T
hEGF	C	P	L	S	H	D	G	Y	C	L	H	D	G	V	C	M	Y	I	E	A	L	D	K	-	-	-	-	Y	A	C	N	C	V	V	G	Y	I	G	E	N	C	Q	Y	R	D	L	K	N	W	E	L	R							
mEGF	C	P	S	S	Y	D	G	Y	C	L	N	G	G	V	C	M	H	I	E	S	L	D	S	-	-	-	-	Y	T	C	N	C	V	I	G	Y	S	G	D	R	C	Q	T	R	D	L	R	N	W	E	L	R							
rEGF	C	P	P	S	Y	D	G	Y	C	L	N	G	G	V	C	M	Y	V	E	S	V	D	R	-	-	-	-	Y	V	C	N	C	V	I	G	Y	I	G	E	R	C	Q	H	R	D	L	R												
gpEGF	C	P	P	S	H	D	G	Y	C	L	N	G	G	V	C	M	H	I	E	S	L	N	T	-	-	-	-	Y	A	C	N	C	V	I	G	Y	V	G	E	R	C	E	H	Q	D	L	D	L	N	E									
amphi	C	N	A	E	F	Q	N	F	C	I	H	-	G	E	C	K	Y	I	E	H	L	E	A	-	-	-	-	V	T	C	K	C	Q	Q	E	Y	F	G	E	R	C	G	E	K	*														

VGF	Y	I	P	S	P	G	I	N	L	V	L	V	G	I	I	I	T	C	C	L	S	V	Y	R	F	T	R	A	T	K	L	T	I	Q	D	M	V	P	*																					
hTGF	A	L	V	V	V	S	I	V	A	L	A	V	L	I	I	T	C	V	L	I	H	C	C	Q	V	R	K	H	C	E	N	C	R	A	L	I	C	R	H	E	K	P	S	A	L	L	K	G	R	T	A	C	C	H	S	E	T	V	V	*
rTGF	A	L	V	V	V	S	I	V	A	L	A	V	L	I	I	T	C	V	L	I	H	C	C	Q	V	R	K	H	C	E	N	C	R	A	L	V	C	R	H	E	K	P	S	A	L	L	K	G	R	T	A	C	C	H	S	E	T	V	V	*

* Les nombres correspondent aux résidus d'acides aminés dans VGF, comme indiqué dans la description.

FIG. 1

$$\frac{2}{2}$$


+ Dans SFGF c'est l'acide aspartique.

* Site de glycosylation potentiel dans SFGF et/ou MGF.

N L'acide aminé indiqué est remplacé avec l'asparagine.

Fig. 2