



(86) Date de dépôt PCT/PCT Filing Date: 2001/05/11  
(87) Date publication PCT/PCT Publication Date: 2001/11/15  
(85) Entrée phase nationale/National Entry: 2002/11/12  
(86) N° demande PCT/PCT Application No.: FR 2001/001441  
(87) N° publication PCT/PCT Publication No.: 2001/085955  
(30) Priorité/Priority: 2000/05/11 (00/06030) FR

(51) Cl.Int.<sup>7</sup>/Int.Cl.<sup>7</sup> C12N 15/55, A61K 48/00, A61K 38/46,  
A61K 38/14, C12N 9/14, C07K 16/40, C12Q 1/34,  
C12Q 1/68, C12N 15/63, C12N 9/90, C12N 15/85,  
G01N 33/53

(71) Demandeurs/Applicants:  
ISTAC, FR;  
INSTITUT PASTEUR DE LILLE, FR

(72) Inventeurs/Inventors:  
BAHR, GEORGES, FR;  
COCUDE, CECILE, FR;  
CAPRON, ANDRE, FR

(74) Agent: OGILVY RENAULT

(54) Titre : POLYPEPTIDE RH116 ET SES FRAGMENTS ET POLYNUCLEOTIDES CODANT LESDITS POLYPEPTIDES  
ET APPLICATIONS THERAPEUTIQUES  
(54) Title: RH116 POLYPEPTIDES AND ITS FRAGMENTS AND POLYNUCLEOTIDES ENCODING SAID  
POLYPEPTIDES AND THERAPEUTIC USES

(57) **Abrégé/Abstract:**

La présente invention concerne un nouveau polypeptide de 116 kDa présentant des homologues de séquences avec des hélicases à ARN (DEXH box) dénommé RH116 et ses fragments, le clonage de l'ADNc et les polynucléotides codant lesdits polypeptides, des vecteurs de clonage et/ou d'expression incluant lesdits polynucléotides, des cellules transformées par lesdits vecteurs et des anticorps spécifiques dirigés contre lesdits polypeptides. L'invention concerne également des procédés de détection et/ou de dosage desdits polypeptides et polynucléotides, les trousse de diagnostic correspondantes, et un procédé de criblage de ligands, ainsi que des composés utilisables à titre de médicament pour la prévention et/ou le traitement thérapeutique.

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION  
EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété  
Intellectuelle  
Bureau international



(43) Date de la publication internationale  
15 novembre 2001 (15.11.2001)

PCT

(10) Numéro de publication internationale  
**WO 01/85955 A1**

(51) Classification internationale des brevets<sup>7</sup> :

**C12N 15/55**, 9/14,  
15/63, 15/85, C07K 16/40, C12Q 1/68, 1/34, G01N 33/53,  
A61K 48/00, 38/46, 38/14, C12N 9/90

(21) Numéro de la demande internationale :

PCT/FR01/01441

(22) Date de dépôt international : 11 mai 2001 (11.05.2001)

(25) Langue de dépôt :

français

(26) Langue de publication :

français

(30) Données relatives à la priorité :

00/06030 11 mai 2000 (11.05.2000) FR

(71) Déposants (pour tous les États désignés sauf US) : **IS-TAC** [FR/FR]; 1, rue du Professeur Calmette, F-59000 Lille (FR). **INSTITUT PASTEUR DE LILLE** [FR/FR]; 1, rue du Professeur Calmette, F-59000 Lille (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants (pour US seulement) : **BAHR, Georges** [FR/FR]; 266, rue Nationale, F-59800 Lille (FR). **COCUDE, Cécile** [FR/FR]; 2, rue Charles Giounod, F-59112 Annoeullin (FR). **CAPRON, André** [FR/FR]; 8, rue du Capitaine Jasmin, F-59133 Phalempin (FR).

(74) Mandataires : **MARTIN, Jean-Jacques** etc.; Cabinet Regimbeau, 20, rue de Chazelles, F-75847 Paris Cedex 17 (FR).

(81) États désignés (national) : AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) États désignés (régional) : brevet ARIPO (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), brevet eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Publiée :

— avec rapport de recherche internationale  
— avant l'expiration du délai prévu pour la modification des revendications, sera republiée si des modifications sont reçues

En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.

(54) Title: RH116 POLYPEPTIDES AND ITS FRAGMENTS AND POLYNUCLEOTIDES ENCODING SAID POLYPEPTIDES AND THERAPEUTIC USES

(54) Titre : POLYPEPTIDE RH116 ET SES FRAGMENTS ET POLYNUCLEOTIDES CODANT LESDITS POLYPEPTIDES ET APPLICATIONS THÉRAPEUTIQUES

(57) Abstract: The invention concerns a novel 116 kDa polypeptide exhibiting homologies of sequences with RNA helicases (DEXH box) called RH116 and its fragments, the cloning of cDNA and the polynucleotides encoding said polypeptides, cloning and/or expression vectors including said polynucleotides, cells transformed by said vectors and specific antibodies directed against said polypeptides. The invention also concerns methods for detecting and/or assaying said polypeptides and polynucleotides, corresponding diagnosis kits, and a method for screening ligands, as well as compounds for use as medicine for preventing and/or therapeutic treatment.

(57) Abrégé : La présente invention concerne un nouveau polypeptide de 116 kDa présentant des homologies de séquences avec des hélicases à ARN (DEXH box) dénommé RH116 et ses fragments, le clonage de l'ADNc et les polynucléotides codant lesdits polypeptides, des vecteurs de clonage et/ou d'expression incluant lesdits polynucléotides, des cellules transformées par lesdits vecteurs et des anticorps spécifiques dirigés contre lesdits polypeptides. L'invention concerne également des procédés de détection et/ou de dosage desdits polypeptides et polynucléotides, les trousse de diagnostic correspondantes, et un procédé de criblage de ligands, ainsi que des composés utilisables à titre de médicament pour la prévention et/ou le traitement thérapeutique.



WO 01/85955 A1

« POLYPEPTIDE RH116 ET SES FRAGMENTS ET POLYNUCLEOTIDES  
CODANT LESDITS POLYPEPTIDES ET APPLICATIONS  
THERAPEUTIQUES ».

La présente invention concerne un nouveau polypeptide de 116 kDa  
5 présentant des homologies de séquences avec des hélicases à ARN (DEXH  
box) dénommé RH116 et ses fragments, le clonage de l'ADNc et les  
polynucléotides codant lesdits polypeptides, des vecteurs de clonage et/ou  
d'expression incluant lesdits polynucléotides, des cellules transformées par  
lesdits vecteurs et des anticorps spécifiques dirigés contre lesdits  
10 polypeptides. L'invention concerne également des procédés de détection  
et/ou de dosage desdits polypeptides et polynucléotides, les trousse de  
diagnostic correspondantes, et un procédé de criblage de ligands, ainsi que  
des composés utilisables à titre de médicament pour la prévention et/ou le  
traitement thérapeutique.

15 Les muramylpeptides sont, parmi les immunomodulateurs de  
synthèse, ceux qui ont montré un grand nombre d'effets  
immunopharmacologiques sur les cellules de la lignée monocyttaire  
/macrophagique potentialisant leur résistance non spécifique à l'infection,  
augmentant l'activité tumoricidale des macrophages, et aussi agissant  
20 comme adjuvants de vaccins. Le Murabutide (MB), analogue du  
muramyldipeptide (MDP) a été sélectionné pour son profil biologique  
particulièrement prometteur et sa bonne tolérance chez l'animal et chez  
l'homme. En effet, contrairement au MDP et à de nombreux autres  
analogues, il est démontré que le MB n'est pas pyrogène, n'induit pas de  
25 réactions inflammatoires et n'a pas montré de toxicité sévère lors des  
études cliniques chez des volontaires sains et de patients atteints de cancer.

De par ses capacités biologiques, le MB est un agent antiviral  
prometteur dans le domaine du SIDA (Syndrome d'Immunodéficience  
acquise). En effet, le MB inhibe la réplication du virus de  
30 l'immunodéficience humaine (VIH) dans les macrophages et les cellules  
dendritiques mais également dans les cellules mononuclées du sang  
périphérique (Peripheral Blood Mononuclear Cells, PBMC) de patients  
infectés. Ainsi, compte tenu de ces caractéristiques biologiques

l'immunomodulateur MB a fait l'objet d'un brevet français délivré N° FR 2 724 845 et intitulé « Compositions de Muramyl peptides capables d'inhiber jusqu'à 100% la réplication d'un virus de l'immunodéficience acquise tel que le VIH ». De plus, les essais cliniques de phase I et de phase IIa menés à  
5 terme sur des patients VIH+ ont démontré une bonne tolérance clinique du MB.

Les inventeurs ont démontré que le MB exerce une forte inhibition de la réplication virale dans les PBMCs de patients déplétées en lymphocytes CD8, activées par la phytohéماغglutinine (PHA) et mises en culture avec de  
10 l'interleukine 2 (IL-2). En effet, le MB inhibe de 70 à 100% le taux de protéine virale p24 du VIH dans les surnageants de culture. Cet effet est corrélé au taux d'expression des ARN messagers viraux (non épissés et simple épissés). De plus, l'analyse du profil de cytokines et chimiokines sécrétées a démontré que le MB induisait la production de chimiokines  
15 connues comme inhibitrices de la réplication du VIH. Toutefois, cette induction ne semble pas être corrélée en totalité avec l'effet inhibiteur du MB. L'inhibition de la réplication du VIH par le MB ne passe pas seulement par l'induction de la production de  $\beta$ -chimiokines puisque le MB intervient aussi au niveau de l'ADN proviral et de la transcription virale. L'absence de  
20 toxicité du MB dans ces mêmes cultures cellulaires a été vérifiée par les inventeurs qui ont constaté que non seulement le nombre de cellules vivantes reste inchangé en début de culture mais semble, de plus, augmenter en fin de culture.

Les résultats obtenus par les inventeurs suggèrent donc que le MB  
25 induit la production de cytokines ou d'autres facteurs non identifiés à ce jour, et qui possèdent une activité suppressive de la réplication du VIH.

Afin d'identifier ces nouveaux facteurs impliqués dans la régulation de la réplication virale, les inventeurs ont utilisé la méthodologie du « Differential Display-RT-PCR » (DD-RT-PCR) qui repose sur 2 étapes  
30 essentielles. Une première étape de reverse transcription (RT) de l'ARN total cellulaire afin d'obtenir des ADN complémentaires de tous les ARN qui possède une queue poly A. Ensuite, une deuxième étape d'amplification par

Polymerase Chain Reaction (PCR) à partir des ADNc servant de matrice et de différents couples d'amorces en présence d'un nucléotide radio-marqué. Les produits de PCR sont ensuite séparés sur gel par électrophorèse. Les fragments différentiellement amplifiés sont coupés du gel, ré-amplifiés puis  
5 clonés et séquencés.

La DD-RT-PCR, réalisée à partir de PBMCs d'un patient VIH+, a permis aux inventeurs de sélectionner plus de 130 fragments d'ADNc différentiellement exprimés après traitement au MB. Ces fragments ont été sous-clonés dans le vecteur pCR2.1 (Invitrogen), puis séquencés par  
10 séquençage automatique (ABI Prism 377, Perkin-Elmer). Les séquences ont été analysées pour les recherches d'homologies en utilisant les banques de données et le serveur Basic Local Alignment Search Tool (Blast 2) du NCBI.

Les inventeurs ont identifié un nouveau polypeptide présentant des homologies de séquences avec des hélicases à ARN.

15 Les hélicases à ARN (pour revue voir Critical Rev. In Biochemistry and Molecular Biology (1998) 33 (4) : 259-296) représentent une large famille de protéines présentes dans tous les types de systèmes biologiques où l'ARN joue un rôle central. Elles sont distribuées de manière ubiquitaire dans une large gamme d'organismes et sont impliquées dans le processus d'épissage  
20 mitochondrial et nucléaire, l'édition de l'ARN (RNA editing), la maturation de rRNA (rRNA processing), l'initiation de la traduction, l'export des ARNm nucléaires et la dégradation des ARNm.

Les hélicases à ARN constituent des facteurs indispensables à la différenciation et au développement cellulaire, et certaines d'entre elles  
25 jouent un rôle dans la transcription et la réplication de génome viral à ARN simple brin.

Les hélicases à ARN appartiennent au large groupe des enzymes susceptibles d'hydroliser les nucléotides 5'-trisphosphates.

Une étude de comparaison de séquence des ATPases-ADN dépendantes  
30 a conduit à une nouvelle classification des NTPases selon le motif ATPase A. Ainsi, les hélicases à ADN se caractérisant par un motif A, encore appelé

motif "Walker" (G-X-X-X-X-G-K-T), et appartiennent à la superfamille I, alors que les hélicases à ARN présentent des variations dans ce domaine (A-X-X-G-X-G-K-T) et forment la superfamille II voisine (Gorbalenya *et al.*, 1988).

- 5 Des comparaisons des séquences conservées ont montré des liens étroits entre les différentes hélicases à ARN, suggérant que ces protéines dérivent d'un ancêtre commun.

En effet, l'alignement de la séquence en acides aminés de différents membres de la superfamille II des hélicases à ARN a permis de mettre en  
10 évidence une région centrale commune qui se caractérise par la présence de huit domaines hautement conservés. A l'heure actuelle, la fonction biochimique de quatre des huit domaines conservés (domaines I, II, VI, et VIII) ont été élucidés. Compte tenu de la forte homologie de séquence dans le cinquième des huit éléments structuraux, appelé motif "boîte DEAD" (ou  
15 "DEAD Box") (D-E-A-D : Asp-Glu-Ala-Asp), les hélicases à ARN appartenant à la superfamille II sont également appelés protéines "Boîte DEAD" (Linder *et al.*, 1989). L'existence de motifs DEAD divergents ont permis de subdiviser en sous-groupes la superfamille II des hélicases à ARN. A ce jour, trois sous-groupes ont été identifiés. Le premier sous-groupe est formé  
20 par les protéines à boîte DEAD classique, les deux autres sous-groupes sont appelés DEAH et DEXH du fait de leur motif ATPase B divergent.

De part et d'autre de la séquence centrale conservée, les extrémités amino- et carboxy-terminales des hélicases à ARN se caractérisent par des séquences de longueur et de contenu variables. Il est suggéré que ces  
25 régions divergentes sont responsables des fonctions protéiques individuelles, alors que les domaines hautement conservés sont impliqués dans l'activité hélicase à ARN.

Le domaine I (A/G-X-X-G-X-G-K-T : Ala/Gly-X-X-Gly-X-Gly-lys-Thr) est décrit comme le motif A des ATPases (Walker *et al.*, 1982).

- 30 Le domaine V, ou boîte DEAD (L-D-E-A-D-X-X-Leu : Leu-Asp-Glu-Ala-Asp-X-X-leu représente une forme spécifique du motif ATPase B (Walker *et*

*al.*, 1982) qui semble impliqué dans l'hydrolyse de l'ATP (Pause and Sonenbery, 1992).

Le domaine VI ou motif SAT (Ser-Ala-Thr) est localisé à proximité de la boîte DEAD et semble spécifique des hélicases à ARN.

5 Le domaine VIII est caractérisé par la boîte YIHRIGRXXXR (Tyr-Ile-His-Arg-Ile-Gly-Arg-X-X-Arg) qui représente un motif qui est, comme SAT, spécifique aux hélicases à ARN. Des expériences *in vitro* avec le facteur d'initiation de la traduction eIF-4A indiquent que ce domaine est critique pour la liaison à l'ARN.

10 La présente invention a donc pour objet un polypeptide isolé dénommé « RH116 » (pour hélicase à ARN de 116 kDa) de séquence d'acides aminés SEQ ID N°2. Cette séquence comprend des domaines consensus conservés qui sont aisément identifiables par l'homme du métier et qui permettent de classer le polypeptide RH116 de l'invention dans le sous-groupe DEAH ou  
15 DEXH (pour revue voir Luking. *et al.*, (1998)) Parmi ces domaines consensus conservés, il convient de citer :

- la séquence GSGKT qui correspond aux acides aminés 332 à 336 de la séquence SEQ ID N°2, et qui constitue le domaine conservé I (G-GKT) de la superfamille des hélicases à ARN.

20 • la séquence DECH qui correspond aux acides aminés 443 à 446 de la séquence SEQ ID N°2 et qui constitue le domaine conservé V (DE-H) de la superfamille des hélicases à ARN du sous-groupe DEXH.

Ces deux domaines consensus conservé sont impliqués dans la fonction ATPase.

25 • la séquence TAS qui correspond aux acides aminés 488 à 490 de la séquence SEQ ID N°2 et qui constitue le domaine conservé VI (-A-) de la superfamille des hélicases à ARN du sous-groupe DEAH.

• la séquence RGRAR qui correspond aux acides aminés 820 à 824 de la séquence SEQ ID N°2 et qui constitue le domaine conservé VIII de la  
30 superfamille des hélicases à ARN du sous-groupe DEAH.

Ces deux domaines conservés (domaines VI et VIII) sont plus spécifiques des hélicases à ARN et sont responsables de la liaison et du déroulement de l'ARN cible.

Le polypeptide isolé se caractérise en ce qu'il comprend un polypeptide  
5 choisi parmi :

- a) un polypeptide de séquence SEQ ID N°2 ;
- b) un polypeptide variant de polypeptide de séquences d'acides aminés défini en a) ;
- c) un polypeptide homologue au polypeptide défini en a) ou  
10 b) et comportant au moins 80 % d'identité, de préférence 85 %, 87 %, 90 %, 95 %, 97 %, 98 %, 99 % d'identité avec ledit polypeptide de a) ;
- d) un fragment d'au moins 15 acides aminés consécutifs de préférence 17, 20, 23, 25, 30, 40, 50, 100, 250 amino-acides  
15 consécutifs d'un polypeptide défini en a), b) ou c) ;
- e) un fragment biologiquement actif d'un polypeptide défini en a), b) ou c).

Dans la présente description, on utilisera le terme polypeptide pour désigner également une protéine ou un peptide.

20 On entendra par polypeptide variant l'ensemble des polypeptides mutés pouvant exister naturellement, en particulier chez l'être humain, et qui correspondent notamment à des troncatures, substitutions, délétions et/ou additions de résidus d'acides aminés.

Par polypeptide homologue, on entendra désigner les polypeptides  
25 présentant, par rapport au polypeptide naturel RH116, certaines modifications comme en particulier une délétion, addition ou substitution d'au moins un acide aminé, une troncature, un allongement et/ou une fusion chimérique. Parmi les polypeptides homologues, on préfère ceux dont la séquence d'acides aminés présente au moins 80 % d'identité, de  
30 préférence d'au moins 85 %, 87 %, 90 %, 93 %, 95 %, 97 %, 98 %, 99 % d'identité avec les séquences d'acides aminés des polypeptides selon l'invention. Dans le cas d'une substitution, un ou plusieurs acides aminés

consécutifs ou non consécutifs, peuvent être remplacés par des acides aminés « équivalents ». L'expression acide aminé « équivalent » vise ici à désigner tout acide aminé susceptible d'être substitué à l'un des acides aminés de la structure de base sans cependant modifier les caractéristiques ou propriétés fonctionnelles essentielles, comme leurs activités biologiques, des polypeptides correspondants telles que l'induction in vivo d'anticorps capables de reconnaître le polypeptide dont la séquence d'acides aminés est comprise dans la séquence d'acides aminés SEQ ID N°2, ou l'un de ses fragments. Ces acides aminés équivalents peuvent être déterminés soit en s'appuyant sur leur homologie de structure avec les acides aminés auxquels ils se substituent, soit sur les résultats des essais d'activité biologique croisée auxquels les différents polypeptides sont susceptibles de donner lieu. A titre d'exemple, on mentionnera les possibilités de substitutions susceptibles d'être effectuées sans qu'il en résulte une modification approfondie des activités biologiques des polypeptides modifiés correspondants, les remplacements, par exemple, de la leucine par la valine ou l'isoleucine, de l'acide aspartique par l'acide glutamique, de la glutamine par l'asparagine, de l'arginine par la lysine etc., les substitutions inverses étant naturellement envisageables dans les mêmes conditions.

Par fragment biologiquement actif, on entendra désigner en particulier un fragment de séquence d'acides aminés de polypeptide selon l'invention présentant au moins une des caractéristiques ou propriétés fonctionnelles du polypeptide selon l'invention, notamment en ce qu'il comporte une activité hélicase à ARN. Le polypeptide variant, le polypeptide homologue ou le fragment de polypeptide selon l'invention possède au moins 10 %, de préférence 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 % de l'activité hélicase à ARN. Différents protocoles connus de l'homme de l'art ont été décrits pour mesurer l'activité hélicase à ARN des polypeptides selon l'invention ; il convient de citer les articles de Lain *et al.* (1990) et Lee et Hurwitz (1993). Les exemples ci-après proposent des fonctions biologiques pour la protéine RH116 en fonction des domaines

peptidiques de cette protéine et permettent ainsi à l'homme du métier d'identifier les fragments biologiquement actifs.

Par fragment de polypeptide, on entend désigner un polypeptide comportant au minimum 15 acides aminés consécutifs, de préférence 17,  
5 20, 23, 25, 30, 40, 50, 100, 250 amino-acides consécutifs. Les fragments de polypeptide selon l'invention obtenus par clivage dudit polypeptide par une enzyme protéolytique, par un réactif chimique, ou encore en plaçant ledit polypeptide dans un environnement très acide font également partie de l'invention.

10 De préférence un polypeptide selon l'invention est un polypeptide constitué de la séquence SEQ ID N° 2 ou d'une séquence possédant au moins 80 % d'identité, de préférence au moins 85 %, 90 %, 95 %, 98 % et 99 % d'identité avec la SEQ ID N° 2 après alignement optimal. Par polypeptide dont la séquence d'acides aminés présentant un pourcentage  
15 d'identité d'au moins 80 %, de préférence d'au moins 85 %, 90 %, 95 %, 98 % et 99 % après alignement optimal avec une séquence de référence, on entend désigner les polypeptides présentant certaines modifications par rapport au polypeptide de référence, comme en particulier une ou plusieurs délétions, troncations, un allongement, une fusion chimérique, et/ou une  
20 ou plusieurs substitutions.

Parmi les polypeptides dont la séquence d'acides aminés présente un pourcentage d'identité d'au moins 80 %, de préférence d'au moins 85 %, 90 %, 95 %, 98 % et 99 % après alignement optimal avec les séquences SEQ ID N° 2 ou avec l'un de leurs fragments selon l'invention, on préfère les  
25 polypeptides variants codés par les séquences peptidiques variantes telles que précédemment définies, en particulier les polypeptides dont la séquence d'acides aminés présente au moins une mutation correspondant notamment à une troncation, délétion, substitution et/ou addition d'au moins un résidu d'acide aminé par rapport aux séquences SEQ ID N°2 ou  
30 avec l'un de leurs fragments, de manière plus préférée les polypeptides variants présentant une mutation liée à une pathologie.

Le polypeptide selon l'invention se caractérise en ce qu'il est comporte au moins un domaine conservé appartenant à la superfamille des hélicases à ARN, ceux-ci étant choisi de préférence parmi les séquences G-GKT correspondant au domaine I, DEAD, DE-D, DEAH, correspondants au  
5 domaine V, SAT, -A-, correspondants au domaine VI, YIHRIGRR, HRIGR—R, -R-GR—R, ---GR, correspondants au domaine VIII.

L'invention concerne également un polynucléotide purifié ou isolé caractérisé en ce qu'il code pour un polypeptide de séquence SEQ ID N°2 tel que défini précédemment. De manière préférée, le polynucléotide selon  
10 l'invention possède la séquence SEQ ID N°1.

Le polynucléotide purifié ou isolé selon l'invention se caractérise en ce qu'il comprend un polynucléotide choisi parmi :

- a) SEQ ID N° 1 ;
- b) la séquence d'un fragment d'au moins 15 nucléotides consécutifs,  
15 de préférence d'au moins 18, 21, 24, 27, 30, 35, 40, 50, 75, 100 nucléotides consécutifs de la séquence SEQ ID N°1 à l'exception des séquences nucléiques identifiées sous le N° d'accèsion AC 007750 et N°AC0108176 dans la banque de données GenBank, ainsi qu'à l'exception de la séquence génomique de 95417 pb, ainsi qu'à  
20 l'exception des séquences nucléiques enregistrées identifiées sous numéro d'accès N°AW589567, N°AW152541 et AW189584 dans la banque de données EMBL ;
- c) une séquence nucléique présentant un pourcentage d'identité d'au moins 80 %, de préférence d'au moins 85 %, 90 %, 95 %, 98 % et 99  
25 % après alignement optimal avec une séquence définie en a) ou b) ;
- d) la séquence complémentaire ou la séquence d'ARN correspondant à une séquence telle que définie en a), b) ou c).

Par acide nucléique, séquence nucléique ou d'acide nucléique, polynucléotide, oligonucléotide, séquence de polynucléotide, séquence  
30 nucléotidique, termes qui seront employés indifféremment dans la présente description, on entend désigner un enchaînement précis de nucléotides, modifiés ou non, permettant de définir un fragment ou une région d'un

acide nucléique, comportant ou non des nucléotides non naturels, et pouvant correspondre aussi bien à un ADN double brin, un ADN simple brin que des produits de transcription desdits ADN, et/ou un fragment d'ARN.

5 Il doit être compris que la présente invention ne concerne pas les séquences nucléotidiques dans leur environnement chromosomique naturel, c'est-à-dire à l'état naturel. Il s'agit de séquences qui ont été isolées et/ou purifiées, c'est-à-dire qu'elles ont été prélevées directement ou indirectement, par exemple par copie, leur environnement ayant été au  
10 moins partiellement modifié. On entend ainsi également désigner les acides nucléiques obtenus par synthèse chimique.

Par polynucléotide de séquence complémentaire, on entend désigner tout ADN dont les nucléotides sont complémentaires de ceux de la SEQ ID N° 1 ou d'une partie de la SEQ ID N°1 et dont l'orientation est inversée.

15 Par « pourcentage d'identité » entre deux séquences d'acides nucléiques ou d'acides aminés au sens de la présente invention, on entend désigner un pourcentage de nucléotides ou de résidus d'acides aminés identiques entre les deux séquences à comparer, obtenu après le meilleur alignement, ce pourcentage étant purement statistique et les différences entre les deux  
20 séquences étant réparties au hasard et sur toute leur longueur. On entend désigner par "meilleur alignement" ou "alignement optimal", l'alignement pour lequel le pourcentage d'identité déterminé comme ci-après est le plus élevé. Les comparaisons de séquences entre deux séquences d'acides nucléiques ou d'acides aminés sont traditionnellement réalisées en  
25 comparant ces séquences après les avoir alignées de manière optimale, ladite comparaison étant réalisée par segment ou par « fenêtre de comparaison » pour identifier et comparer les régions locales de similarité de séquence. L'alignement optimal des séquences pour la comparaison peut être réalisé, outre manuellement, au moyen de l'algorithme d'homologie  
30 locale de Smith et Waterman (1981), au moyen de l'algorithme d'homologie locale de Needleman et Wunsch (1970), au moyen de la méthode de recherche de similarité de Pearson et Lipman (1988), au moyen de logiciels

informatiques utilisant ces algorithmes (GAP, BESTFIT, BLAST P, BLAST N, FASTA et TFASTA dans le Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI). Afin d'obtenir l'alignement optimal, on utilise de préférence le programme BLAST, avec la matrice  
5 BLOSUM 62. On peut également utiliser les matrices PAM ou PAM250.

Le pourcentage d'identité entre deux séquences d'acides nucléiques ou d'acides aminés est déterminé en comparant ces deux séquences alignées de manière optimale, la séquence d'acides nucléiques ou d'acides aminés à comparer pouvant comprendre des additions ou des délétions par rapport à  
10 la séquence de référence pour un alignement optimal entre ces deux séquences. Le pourcentage d'identité est calculé en déterminant le nombre de positions identiques pour lesquelles le nucléotide ou le résidu d'acide aminé est identique entre les deux séquences, en divisant ce nombre de positions identiques par le nombre total de positions comparées et en  
15 multipliant le résultat obtenu par 100 pour obtenir le pourcentage d'identité entre ces deux séquences.

Par séquences nucléiques présentant un pourcentage d'identité d'au moins 80 %, de préférence d'au moins 85 %, 90 %, 95 %, 98 % et 99 % après alignement optimal avec une séquence de référence, on entend  
20 désigner les séquences nucléiques présentant, par rapport à la séquence nucléique de référence, certaines modifications comme en particulier une délétion, une troncation, un allongement, une fusion chimérique, et/ou une substitution, notamment ponctuelle, et dont la séquence nucléique présente au moins 80 %, de préférence au moins 85 %, 90 %, 95 %, 98 % et 99 %  
25 d'identité après alignement optimal avec la séquence nucléique de référence. Il s'agit de préférence de séquences dont les séquences complémentaires sont susceptibles de s'hybrider spécifiquement avec les séquences SEQ ID N° 1 de l'invention. De préférence, les conditions d'hybridation spécifiques ou de forte stringence seront telles qu'elles  
30 assurent au moins 80 %, de préférence d'au moins 85 %, 90 %, 95 %, 98 % et 99 % d'identité après alignement optimal entre l'une des deux séquences et la séquence complémentaire de l'autre.

Une hybridation dans des conditions de forte stringence signifie que les conditions de température et de force ionique sont choisies de telle manière qu'elles permettent le maintien de l'hybridation entre deux fragments d'ADN complémentaires. A titre illustratif, des conditions de forte stringence de l'étape d'hybridation aux fins de définir les fragments polynucléotidiques décrits ci-dessus, sont avantageusement les suivantes.

L'hybridation ADN-ADN ou ADN-ARN est réalisée en deux étapes : (1) préhybridation à 42°C pendant 3 heures en tampon phosphate (20 mM, pH 7,5) contenant 5 x SSC (1 x SSC correspond à une solution 0,15 M NaCl + 0,015 M citrate de sodium), 50 % de formamide, 7 % de sodium dodécyl sulfate (SDS), 10 x Denhardt's, 5 % de dextran sulfate et 1 % d'ADN de sperme de saumon ; (2) hybridation proprement dite pendant 20 heures à une température dépendant de la taille de la sonde (i.e. : 42°C, pour une sonde de taille > 100 nucléotides) suivie de 2 lavages de 20 minutes à 20°C en 2 x SSC + 2 % SDS, 1 lavage de 20 minutes à 20°C en 0,1 x SSC + 0,1 % SDS. Le dernier lavage est pratiqué en 0,1 x SSC + 0,1 % SDS pendant 30 minutes à 60°C pour une sonde de taille > 100 nucléotides. Les conditions d'hybridation de forte stringence décrites ci-dessus pour un polynucléotide de taille définie, peuvent être adaptées par l'homme du métier pour des oligonucléotides de taille plus grande ou plus petite, selon l'enseignement de Sambrook *et al.*, 1989.

Parmi les séquences nucléiques présentant un pourcentage d'identité d'au moins 80 %, de préférence d'au moins 85 %, 90 %, 95 %, 98 % et 99 % après alignement optimal avec la séquence selon l'invention, on préfère également les séquences nucléiques variantes de SEQ ID N° 1, ou de leurs fragments, c'est-à-dire l'ensemble des séquences nucléiques correspondant à des variants alléliques, c'est-à-dire des variations individuelles des séquences SEQ ID N° 1. Ces séquences mutées naturelles correspondent à des polymorphismes présents chez les mammifères, en particulier chez l'être humain et, notamment, à des polymorphismes pouvant conduire à la survenue d'une pathologie.

On entend également désigner par séquence nucléique variante tout ARN ou ADNc résultant d'une mutation et/ou variation d'un site d'épissage de la séquence nucléique génomique dont l'ADNc a pour séquence SEQ ID N° 1.

5 Plus particulièrement, l'invention concerne un acide nucléique purifié ou isolé selon la présente invention, caractérisé en ce qu'il comprend ou est constitué de l'une des séquences SEQ ID N° 1, de leurs séquences complémentaires ou des séquences de l'ARN correspondant à SEQ ID N° 1. Les amorces ou sondes, caractérisées en ce qu'elles comprennent une  
10 séquence d'un acide nucléique selon l'invention, font également partie de l'invention. Ainsi, la présente invention pour la détection, l'identification, le dosage ou l'amplification de séquence d'acide nucléique concerne également les amorces ou les sondes selon l'invention qui peuvent permettre en particulier de mettre en évidence ou de discriminer les séquences  
15 nucléiques variantes, ou d'identifier la séquence génomique des gènes dont l'ADNc est représenté par SEQ ID N° 1, en utilisant notamment une méthode d'amplification telle que la méthode PCR, ou une méthode apparentée. Selon l'invention, les polynucléotides pouvant être utilisés comme sonde ou comme amorce dans des procédés de détection,  
20 d'identification, de dosage ou d'amplification de séquence nucléique, présentent une taille minimale de 15 bases, de préférence d'au moins 18, 20, 25, 30, 40, 50 bases.

Les polynucléotides selon l'invention peuvent ainsi être utilisés comme amorce et/ou sonde dans des procédés mettant en oeuvre notamment la  
25 technique de PCR (amplification en chaîne par polymérase) (Rolfs *et al.*, 1991). Cette technique nécessite le choix de paires d'amorces oligonucléotidiques encadrant le fragment qui doit être amplifié. On peut, par exemple, se référer à la technique décrite dans le brevet américain U.S. N° 4 683 202. Les fragments amplifiés peuvent être identifiés, par exemple  
30 après une électrophorèse en gel d'agarose ou de polyacrylamide, ou après une technique chromatographique comme la filtration sur gel ou la chromatographie échangeuse d'ions, puis séquencés. La spécificité de

l'amplification peut être contrôlée en utilisant comme amorces les séquences nucléotidiques de polynucléotides de l'invention et comme matrices, des plasmides contenant ces séquences ou encore les produits d'amplification dérivés. Les fragments nucléotidiques amplifiés peuvent être  
5 utilisés comme réactifs dans des réactions d'hybridation afin de mettre en évidence la présence, dans un échantillon biologique, d'un acide nucléique cible de séquence complémentaire à celle desdits fragments nucléotidiques amplifiés. L'invention vise également les acides nucléiques susceptibles d'être obtenus par amplification à l'aide d'amorces selon l'invention.

10 D'autres techniques d'amplification de l'acide nucléique cible peuvent être avantageusement employées comme alternative à la PCR (PCR-like) à l'aide de couple d'amorces de séquences nucléotidiques selon l'invention. Par PCR-like on entend désigner toutes les méthodes mettant en œuvre des reproductions directes ou indirectes des séquences d'acides nucléiques, ou  
15 bien dans lesquelles les systèmes de marquage ont été amplifiés, ces techniques sont bien entendu connues. En général il s'agit de l'amplification de l'ADN par une polymérase ; lorsque l'échantillon d'origine est un ARN il convient préalablement d'effectuer une transcription reverse. Il existe actuellement de très nombreux procédés permettant cette  
20 amplification, comme par exemple la technique SDA (Strand Displacement Amplification) ou technique d'amplification à déplacement de brin (Walker *et al.*, 1992), la technique TAS (Transcription-based Amplification System) décrite par Kwoh *et al.* (1989), la technique 3SR (Self-Sustained Sequence Replication) décrite par Guatelli *et al.* (1990), la technique NASBA (Nucleic  
25 Acid Sequence Based Amplification) décrite par Kievitis *et al.* (1991), la technique TMA (Transcription Mediated Amplification), la technique LCR (Ligase Chain Reaction) décrite par Landegren *et al.* (1988), la technique de RCR (Repair Chain Reaction) décrite par Segev (1992), la technique CPR (Cycling Probe Reaction) décrite par Duck *et al.* (1990), la technique  
30 d'amplification à la Q-béta-réplase décrite par Miele *et al.* (1983). Certaines de ces techniques ont depuis été perfectionnées.

Dans le cas où le polynucléotide cible à détecter est un ARNm, on utilise avantageusement, préalablement à la mise en oeuvre d'une réaction d'amplification à l'aide des amorces selon l'invention ou à la mise en oeuvre d'un procédé de détection à l'aide des sondes de l'invention, une enzyme de  
5 type transcriptase inverse afin d'obtenir un ADNc à partir de l'ARNm contenu dans l'échantillon biologique. L'ADNc obtenu servira alors de cible pour les amorces ou les sondes mises en oeuvre dans le procédé d'amplification ou de détection selon l'invention.

La technique d'hybridation de sondes peut être réalisée de manières  
10 diverses (Matthews *et al.*, 1988). La méthode la plus générale consiste à immobiliser l'acide nucléique extrait des cellules de différents tissus ou de cellules en culture sur un support (tels que la nitrocellulose, le nylon, le polystyrène) pour réaliser par exemple des puces à ADN, puis à incubé, dans des conditions bien définies, l'acide nucléique cible immobilisé avec la  
15 sonde. Après l'hybridation, l'excès de sonde est éliminé et les molécules hybrides formées sont détectées par la méthode appropriée (mesure de la radioactivité, de la fluorescence ou de l'activité enzymatique liée à la sonde).

Selon un autre mode de mise en oeuvre des sondes nucléiques selon l'invention, ces dernières peuvent être utilisées comme sondes de capture.  
20 Dans ce cas, une sonde, dite « sonde de capture », est immobilisée sur un support et sert à capturer par hybridation spécifique l'acide nucléique cible obtenu à partir de l'échantillon biologique à tester et l'acide nucléique cible est ensuite détecté grâce à une seconde sonde, dite « sonde de détection », marquée par un élément facilement détectable.

25 Parmi les fragments d'acides nucléiques intéressants, il convient par ailleurs de citer en particulier les oligonucléotides anti-sens, c'est-à-dire dont la structure assure, par hybridation avec la séquence cible, une inhibition de l'expression du produit correspondant. Il faut également citer les oligonucléotides sens qui, par interaction avec des protéines impliquées  
30 dans la régulation de l'expression du produit correspondant, induiront soit une inhibition, soit une activation de cette expression. Les oligonucléotides

selon l'invention présente une taille minimale de 9 bases, de préférence d'au moins 10, 12, 15, 17, 20, 25, 30, 40, 50 bases.

Les sondes, amorces et oligonucléotides selon l'invention peuvent être marqués directement ou indirectement par un composé radioactif ou non radioactif, par des méthodes bien connues de l'homme du métier, afin d'obtenir un signal détectable et/ou quantifiable. Les séquences de polynucléotides selon l'invention non marquées peuvent être utilisées directement comme sonde ou amorce.

Les séquences sont généralement marquées pour obtenir des séquences utilisables pour de nombreuses applications. Le marquage des amorces ou des sondes selon l'invention est réalisé par des éléments radioactifs ou par des molécules non radioactives. Parmi les isotopes radioactifs utilisés, on peut citer le  $^{32}\text{P}$ , le  $^{33}\text{P}$ , le  $^{35}\text{S}$ , le  $^3\text{H}$  ou le  $^{125}\text{I}$ . Les entités non radioactives sont sélectionnées parmi les ligands tels la biotine, l'avidine, la streptavidine, la dioxygénine, les haptènes, les colorants, les agents luminescents tels que les agents radioluminescents, chémoluminescents, bioluminescents, fluorescents, phosphorescents.

La présente invention concerne également les vecteurs de clonage et/ou d'expression comprenant un acide nucléique ou codant pour un polypeptide selon l'invention. Un tel vecteur peut également contenir les éléments nécessaires à l'expression et éventuellement à la sécrétion du polypeptide dans une cellule hôte. Une telle cellule hôte est également un objet de l'invention.

Selon un autre aspect, l'invention concerne un vecteur d'expression antisens. Un tel vecteur d'expression contient une séquence de polynucléotide selon l'invention, insérée en orientation inverse dans le vecteur d'expression. Ainsi, l'homme du métier reconnaît aisément qu'un ARNm correspondant à l'ADN dans le vecteur antisens hybride avec un ARNm correspondant à de l'ADN dans le vecteur sens. Un vecteur d'expression antisens est un vecteur qui exprime un ARN antisens d'intérêt dans une cellule hôte appropriée, soit de manière constitutive soit après induction. Le terme antisens se réfère à toute composition contenant une

séquence d'acide nucléique spécifique. Les molécules antisens peuvent être produites par des méthodes telles que la synthèse ou la transcription. Lorsque de telles molécules sont introduites dans la cellule, les nucléotides complémentaires se combinent avec les séquences naturelles produites par la cellule pour former des duplexes et ainsi bloquer soit la transcription ou la traduction du polypeptide selon l'invention. Il entre également dans l'étendue de l'invention, de réaliser des molécules antisens susceptibles de s'apparier avec la molécule d'ARN qui est le substrat de l'hélicase à ARN RH116, afin d'en bloquer l'activité biologique.

10 Les nouveaux composés identifiés susceptibles de diminuer ou d'abolir le taux d'expression et/ou l'activité hélicase à ARN du polypeptide selon l'invention constituent des antagonistes du polypeptide selon l'invention. Le terme "antagoniste" se réfère à une molécule qui, lorsqu'elle se lie au polypeptide selon l'invention, diminue la quantité ou la durée des effets de l'activité biologique ou immunologique du polypeptide RH116. Les antagonistes incluent les protéines, les acides nucléiques, les carbohydrates, et toutes les molécules susceptibles de diminuer les effets de RH116.

Lesdits vecteurs comportent de préférence un promoteur, des signaux d'initiation et de terminaison de la traduction, ainsi que des régions appropriées de régulation de la transcription. Ils doivent pouvoir être maintenus de façon stable dans la cellule et peuvent éventuellement posséder des signaux particuliers spécifiant la sécrétion de la protéine traduite. Selon un mode particulier de réalisation de l'invention, le promoteur peut être le promoteur naturellement présent en amont du gène codant pour le polypeptide RH116 humain de l'invention.

Les différents signaux de contrôle sont choisis en fonction de l'hôte cellulaire utilisé. A cet effet, les séquences d'acide nucléique selon l'invention peuvent être insérées dans des vecteurs à réplication autonome au sein de l'hôte choisi, ou des vecteurs intégratifs de l'hôte choisi.

Parmi les systèmes à réplication autonome, on utilise de préférence en fonction de la cellule hôte, des systèmes de type « plasmide », « cosmide » ou

« mini-chromosome » ou des systèmes de type viral, les vecteurs viraux pouvant notamment être des adénovirus (Perricaudet *et al.*, 1992), des rétrovirus, des lentivirus, des poxvirus ou des virus herpétiques (Epstein *et al.*, 1992). L'homme du métier connaît les technologies utilisables pour  
5 chacun de ces systèmes.

Lorsque l'on souhaite l'intégration de la séquence dans les chromosomes de la cellule hôte, on peut utiliser par exemple des systèmes de type plasmidique ou viral ; de tels virus sont, par exemple, les rétrovirus (Temin, 1986), ou les AAV (Carter, 1993).

10 Parmi les vecteurs non viraux, on préfère les polynucléotides nus tels que l'ADN nu ou l'ARN nu selon la technique développée par la société VICAL, les chromosomes artificiels de bactérie (BAC, bacterial artificial chromosome), les chromosomes artificiels de levure (YAC, yeast artificial chromosome) pour l'expression dans la levure, les chromosomes artificiels  
15 de souris (MAC, mouse artificial chromosome) pour l'expression dans les cellules murines et de manière préférée les chromosomes artificiels d'homme (HAC, human artificial chromosome) pour l'expression dans les cellules humaines.

De tels vecteurs sont préparés selon les méthodes couramment  
20 utilisées par l'homme du métier, et les clones en résultant peuvent être introduits dans un hôte approprié par des méthodes standard, telles que par exemple la lipofection, l'électroporation, le choc thermique, la transformation après perméabilisation chimique de la membrane, la fusion cellulaire.

25 L'invention comprend en outre les cellules hôtes, notamment les cellules eucaryotes et procaryotes, transformées par les vecteurs selon l'invention. Parmi les cellules utilisables aux sens de la présente invention, on peut citer les cellules bactériennes (Olins et Lee, 1993), mais aussi les cellules de levure (Buckholz, 1993), de même que les cellules animales, en  
30 particulier les cultures de cellules de mammifères (Edwards et Aruffo, 1993), et notamment les cellules d'ovaire de hamster chinois (CHO) et les cellules humaines. On peut citer également les cellules d'insectes dans

lesquelles on peut utiliser des procédés mettant par exemple en œuvre des baculovirus (Luckow, 1993). Un hôte cellulaire préféré pour l'expression des protéines de l'invention est constitué par les cellules COS et les cellules Hela.

5 L'invention comprend également les animaux transgéniques, de préférence les mammifères, excepté l'Homme, comprenant une desdites cellules transformées selon l'invention. Ces animaux peuvent être utilisés en temps que modèles, pour l'étude de l'étiologie de pathologie liée à une altération de l'homologue animal de la protéine RH116 naturelle humaine  
10 ou pour l'étude des effets d'une infection virale provoquée par un virus à ARN, tel le VIH, sur l'expression de la protéine RH116 en présence ou non d'un traitement anti-viral, tel qu'un antagoniste du RH116, comme par exemple le murabutide.

Parmi les mammifères selon l'invention, on préfère des animaux tels  
15 que les rongeurs, en particulier les souris, les rats ou les lapins, exprimant un polypeptide selon l'invention.

Les animaux transgéniques selon l'invention peuvent surexprimer le gène codant pour la protéine selon l'invention, ou leur gène homologue, ou exprimer ledit gène dans lequel est introduite une mutation. Ces animaux  
20 transgéniques, en particulier des souris, sont obtenus par exemple par transfection de copie de ce gène sous contrôle d'un promoteur fort de nature ubiquitaire, ou sélectif d'un type de tissu, ou après transcription virale.

Alternativement, les animaux transgéniques selon l'invention peuvent  
25 être rendus déficients pour le gène codant pour le polypeptide de séquence SEQ ID N° 2, ou leurs gènes homologues, par inactivation ciblée par recombinaison homologue en utilisant ou non le système LOX-P/CRE recombinase (Rohlmann *et al.*, 1996) ou par tout autre système d'inactivation de l'expression de ce gène. Ces animaux transgéniques sont  
30 obtenus par exemple par recombinaison homologue sur cellules souches embryonnaires, transfert de ces cellules souches à des embryons, sélection

des chimères affectées au niveau des lignées reproductrices, et croissance desdites chimères.

Les cellules ou mammifères transformés tels que décrits précédemment peuvent aussi être utilisés à titre de modèles afin d'étudier  
5 les interactions entre les polypeptides selon l'invention, et les composés chimiques ou protéiques, impliqués directement ou indirectement dans les activités des polypeptides selon l'invention, ceci afin d'étudier les différents mécanismes et interactions mis en jeu. Ils peuvent en particulier être utilisés pour la sélection de produits interagissant avec les polypeptides  
10 selon l'invention, notamment la protéine de séquence SEQ ID N°2 ou leurs variants selon l'invention, à titre de cofacteur, ou d'inhibiteur, notamment compétitif, ou encore ayant une activité agoniste ou antagoniste de l'activité des polypeptides selon l'invention. De préférence, on utilise lesdites cellules transformées ou animaux transgéniques à titre de modèle notamment pour  
15 la sélection de composés permettant de diminuer le taux d'expression ou l'activité hélicase à ARN du polypeptide RH116 de l'invention.

En plus de leur utilité à titre de modèle d'analyse, les cellules et mammifères selon l'invention sont utilisables dans une méthode de production d'un polypeptide selon l'invention, comme décrit ci-dessous. La  
20 méthode de production d'un polypeptide de l'invention sous forme recombinante, elle-même comprise dans la présente invention, se caractérise en ce que l'on cultive les cellules transformées, notamment les cellules de la présente invention, dans des conditions permettant l'expression et éventuellement la sécrétion d'un polypeptide recombinant  
25 codé par une séquence d'acide nucléique selon l'invention, et que l'on récupère ledit polypeptide recombinant. Les polypeptides recombinants susceptibles d'être obtenus par cette méthode de production font également partie de l'invention. Ils peuvent se présenter sous forme glycosylée ou non glycosylée et peuvent présenter ou non la structure tertiaire de la protéine  
30 naturelle. Les séquences des polypeptides recombinants peuvent être également modifiées afin d'améliorer leur solubilité, en particulier dans les solvants aqueux. De telles modifications sont connues de l'homme du

métier comme par exemple la délétion de domaines hydrophobes ou la substitution d'acides aminés hydrophobes par des acides aminés hydrophiles.

Ces polypeptides peuvent être produits à partir des séquences d'acide  
5 nucléique définies ci-dessus, selon les techniques de production de polypeptides recombinants connues de l'homme du métier. Dans ce cas, la séquence d'acide nucléique utilisée est placée sous le contrôle de signaux permettant son expression dans un hôte cellulaire.

Un système efficace de production d'un polypeptide recombinant  
10 nécessite de disposer d'un vecteur et d'une cellule hôte selon l'invention. Ces cellules peuvent être obtenues par l'introduction dans des cellules hôtes d'une séquence nucléotidique insérée dans un vecteur tel que défini ci-dessus, puis la mise en culture desdites cellules dans des conditions permettant la réplication et/ou l'expression de la séquence nucléotidique  
15 transfectée.

Les procédés utilisés pour la purification d'un polypeptide recombinant sont connus de l'homme du métier. Le polypeptide recombinant peut être purifié à partir de lysats et extraits cellulaires, du surnageant du milieu de culture, par des méthodes utilisées individuellement ou en combinaison,  
20 tels que le fractionnement, les méthodes de chromatographie, les techniques d'immunoaffinité à l'aide d'anticorps monoclonaux ou polyclonaux spécifiques, etc... . Une variante préférée consiste à produire un polypeptide recombinant fusionné à une protéine « porteuse » (protéine chimère). L'avantage de ce système est qu'il permet une stabilisation et une  
25 diminution de la protéolyse du produit recombinant, une augmentation de la solubilité au cours de la renaturation in vitro et/ou une simplification de la purification lorsque le partenaire de fusion possède une affinité pour un ligand spécifique.

Les polypeptides selon la présente invention peuvent aussi être  
30 obtenus par synthèse chimique en utilisant l'une des nombreuses synthèses peptidiques connues, par exemple les techniques mettant en œuvre des phases solides (voir notamment Stewart *et al.*, 1984) ou des

techniques utilisant des phases solides partielles, par condensation de fragments ou par une synthèse en solution classique. Les polypeptides obtenus par synthèse chimique et pouvant comporter des acides aminés non naturels correspondants sont également compris dans l'invention.

5 L'invention concerne également un anticorps monoclonal ou polyclonal et ses fragments, caractérisés en ce qu'ils lient sélectivement et/ou spécifiquement un polypeptide selon l'invention. Les anticorps chimériques, les anticorps humanisés et les anticorps simple chaîne font également partie de l'invention. Les fragments d'anticorps selon l'invention sont de  
10 préférence des fragments Fab, F(ab')<sub>2</sub> ou Fv.

Les polypeptides selon l'invention permettent de préparer des anticorps monoclonaux ou polyclonaux. Les anticorps monoclonaux pourront avantageusement être préparés à partir d'hybridomes selon la technique décrite par Kohler et Milstein en 1975.

15 Les anticorps polyclonaux pourront être préparés, par exemple par immunisation d'un animal, en particulier une souris, avec un polypeptide selon l'invention associé à un adjuvant de la réponse immunitaire, puis purification des anticorps spécifiques contenus dans le sérum des animaux immunisés sur une colonne d'affinité sur laquelle a préalablement été fixé le  
20 polypeptide ayant servi d'antigène. Les anticorps polyclonaux selon l'invention peuvent aussi être préparés par purification sur une colonne d'affinité, sur laquelle a préalablement été immobilisé un polypeptide selon l'invention.

Selon un mode particulier de réalisation de l'invention, l'anticorps est  
25 capable d'inhiber l'interaction entre le polypeptide RH116 et la séquence d'ARN sur laquelle celle-ci se lie afin d'altérer la fonction physiologique dudit polypeptide RH116.

L'invention concerne également des méthodes pour la détection et/ou la purification d'un polypeptide selon l'invention, caractérisées en ce  
30 qu'elles mettent en œuvre un anticorps selon l'invention. L'invention comprend en outre des polypeptides purifiés, caractérisés en ce qu'ils sont obtenus par une méthode selon l'invention.

Par ailleurs, outre leur utilisation pour la purification des polypeptides, les anticorps de l'invention, en particulier les anticorps monoclonaux, peuvent également être utilisés pour la détection de ces polypeptides dans un échantillon biologique.

5 Pour ces différentes utilisations, les anticorps de l'invention pourront également être marqués de la même manière que décrit précédemment pour les sondes nucléiques de l'invention et de manière préférée avec un marquage de type enzymatique, fluorescent ou radioactif.

Les anticorps de l'invention constituent également un moyen d'analyse  
10 de l'expression de polypeptide selon l'invention, par exemple par immunofluorescence, marquage à l'or, immunoconjugués enzymatiques. Plus généralement, les anticorps de l'invention peuvent être avantageusement mis en oeuvre dans toute situation où l'expression d'un polypeptide selon l'invention doit être observée, et plus particulièrement en  
15 immunocytochimie, en immunohistochimie ou dans des expériences de « western blotting ».

Ils peuvent permettre notamment de mettre en évidence une expression anormale de ces polypeptides dans les tissus ou prélèvements biologiques.

20 Plus généralement, les anticorps de l'invention peuvent être avantageusement mis en oeuvre dans toute situation où l'expression d'un polypeptide selon l'invention, normal ou muté, doit être observée. Ainsi, un procédé de détection d'un polypeptide selon l'invention dans un échantillon biologique, comprenant les étapes de mise en contact de l'échantillon  
25 biologique avec un anticorps selon l'invention et de mise en évidence du complexe antigène-anticorps formé est également un objet de l'invention.

Entre également dans le cadre de l'invention, une trousse de réactif pour la détection et/ou le dosage d'un polypeptide selon l'invention dans un échantillon biologique, caractérisé en ce qu'il comprend les éléments  
30 suivants : (i) un anticorps monoclonal ou polyclonal tel que décrit précédemment ; (ii) le cas échéant, les réactifs pour la constitution du milieu propice à la réaction immunologique ; (iii) les réactifs permettant la

détection des complexes antigène-anticorps produits par la réaction immunologique. Cette trousse est notamment utile à la réalisation d'expériences de Western Blotting ; celles-ci permettent d'étudier la régulation de l'expression du polypeptide selon l'invention à partir de tissus  
5 ou de cellules. Cette trousse est également utile aux expériences d'immunoprécipitation pour mettre en évidence notamment les protéines interagissant avec le polypeptide selon l'invention. Cette trousse est également utile pour réaliser la détection et/ou le dosage d'un polypeptide selon l'invention en utilisant une méthode qui met en jeu la technique  
10 ELISA, l'immunofluorescence, la radio-immunologie (technique RIA) ou une technique équivalente.

L'invention comprend également une méthode de détection et/ou de dosage d'un polynucléotide selon l'invention, dans un échantillon biologique, caractérisé en ce qu'il comporte les étapes suivantes : (i)  
15 d'isolement de l'ADN à partir de l'échantillon biologique à analyser, ou obtention d'un ADNc à partir de l'ARN de l'échantillon biologique ; (ii) d'amplification spécifique de l'ADN codant pour le polypeptide selon l'invention à l'aide d'amorces ; (iii) d'analyse des produits d'amplification. C'est également un objet de l'invention de fournir une trousse pour la  
20 détection et/ou le dosage d'un acide nucléique selon l'invention, dans un échantillon biologique, caractérisé en ce qu'il comprend les éléments suivants : (i) un couple d'amorces nucléiques selon l'invention, (ii) les réactifs nécessaires pour effectuer une réaction d'amplification d'ADN, et éventuellement (iii) un composant permettant de vérifier la séquence du  
25 fragment amplifié, plus particulièrement une sonde selon l'invention.

L'invention comprend aussi une méthode de détection et/ou de dosage d'acide nucléique selon l'invention, dans un échantillon biologique, caractérisé en ce qu'il comporte les étapes suivantes : (i) de mise en contact d'un polynucléotide selon l'invention avec un échantillon biologique ; (ii) de  
30 détection et/ou de dosage de l'hybride formé entre ledit polynucléotide et l'acide nucléique de l'échantillon biologique. C'est également un objet de l'invention de fournir une trousse pour la détection et/ou le dosage d'acide

nucléique selon l'invention, dans un échantillon biologique, caractérisé en ce qu'il comprend les éléments suivants : (i) une sonde selon l'invention, (ii) les réactifs nécessaires à la mise en oeuvre d'une réaction d'hybridation, et/ou le cas échéant, (iii) un couple d'amorces selon l'invention, ainsi que  
5 les réactifs nécessaires à une réaction d'amplification de l'ADN.

De préférence, l'échantillon biologique selon l'invention dans lequel sont réalisés la détection et le dosage est constitué par un fluide corporel, par exemple un sérum humain ou animal, du sang, ou par des biopsies.

Font également partie de l'invention, les méthodes de détermination  
10 d'une variabilité allélique, d'une mutation, d'une délétion, d'une perte d'hétérozygotie ou de toute anomalie génétique du gène codant le polypeptide selon l'invention, caractérisées en ce qu'elles mettent en oeuvre une séquence d'acide nucléique, un polypeptide ou un anticorps selon l'invention.

15 On peut détecter ces mutations directement par analyse de l'acide nucléique et des séquences selon l'invention (ARN, ou ADNc), mais également par l'intermédiaire des polypeptides selon l'invention. En particulier, l'utilisation d'un anticorps selon l'invention qui reconnaît un épitope portant une mutation permet de discriminer entre une protéine  
20 « saine » et une protéine « associée à une pathologie ».

Cette méthode de diagnostic et/ou d'évaluation pronostique peut être utilisée de façon préventive, ou afin de servir à l'établissement et/ou la confirmation d'un état clinique chez un patient. L'analyse peut être effectuée par séquençage de tout ou d'une partie du gène (i.e. les exons), ou  
25 par d'autres méthodes connues de l'homme du métier. On peut en particulier utiliser des méthodes basées sur la PCR, par exemple la PCR-SSCP qui permet de détecter des mutations ponctuelles. On peut également effectuer l'analyse par fixation d'une sonde selon l'invention sur une puce à ADN contenant au moins un polynucléotide selon l'invention et  
30 l'hybridation sur ces microplaques. Une puce à ADN contenant une séquence selon l'invention est également un des objets de l'invention.

De même, une puce à protéines contenant une séquence d'acides aminés selon l'invention est aussi un objet de l'invention. Une telle puce à protéines permet l'étude des interactions entre les polypeptides selon l'invention et d'autres protéines ou des composés chimiques, et peut ainsi  
5 être utile pour le criblage de composés interagissant avec les polypeptides selon l'invention. On peut également utiliser les puces à protéines selon l'invention pour détecter la présence d'anticorps dirigés contre les polypeptides selon l'invention dans le sérum de patients. On peut aussi mettre en œuvre une puce à protéines contenant un anticorps selon  
10 l'invention.

Des substances peuvent également être testées et identifiées pour leur aptitude à moduler les activités enzymatiques du polypeptide de l'invention. On entend désigner par "une substance qui module une activité enzymatique", une substance qui change l'activité enzymatique par rapport  
15 à l'activité enzymatique mesurée en absence de la substance à tester. Par exemple une telle substance peut inhiber partiellement ou totalement l'activité hélicase à ARN ; une telle activité peut être mesurée par des méthodes connues par l'homme de l'art. Par exemple des oligonucléotides synthétiques peuvent être immobilisés sur une matrice et hybridé avec un  
20 oligoribonucléotide complémentaire marqué. Les oligoribonucléotides hybridés sont ensuite mis en œuvre avec un polypeptide de l'invention, qui relargue une certaine quantité mesurable de l'oligoribonucléotide marqué, non attaché à la matrice, compte tenu de son activité hélicase à ARN. Les effets de la présence ou de l'absence de modulateurs potentiels de l'hélicase  
25 à ARN RH116 de l'invention ou l'un de ses fragments peuvent ainsi être testés. Une autre alternative consiste à utiliser le protocole décrit par Jaramillo *et al.* (1991) ; cette méthode consiste à mélanger un substrat ARN duplexe marqué au  $^{32}\text{P}$ , avec le polypeptide de l'invention dans une solution tamponnée ; la réaction est arrêtée par addition d'un mélange de  
30 glycérol/SDS/EDTA/bleu de bromophénol ; le mélange est ensuite préparé sur un gel SDS-PAGE (8%) selon les conditions standards. L'activité hélicase à ARN est estimée par le rapport entre la quantité d'ARN monomère

sur la quantité d'ARN duplexe. D'autres protocoles sont également disponibles (Rozen *et al.*(1990) ; Pause *et al.* (1992) ; Lain *et al.* (1993), Lee et Hurwitz (1993)). Ces essais peuvent être réalisés dans des microplaques afin de tester simultanément de larges quantités de modulateurs,  
5 d'inhibiteurs du polypeptide selon l'invention.

De telles substances peuvent être au préalable développées et sélectionnées par modelage moléculaire afin d'identifier des substances susceptibles de réagir avec un polypeptide de l'invention. Il est possible d'identifier une substance qui affecte l'aptitude de la protéine de l'invention  
10 à lier avec l'ATP ou d'autres substrats, tels que l'ARN, l'ADN ou les complexes ARN/protéines. Les substrats qui interfèrent avec un site de liaison au substrat de la protéine de l'invention ou avec un site qui affecte de tels épitopes fonctionnels peuvent être identifiés par des méthodes connues de l'homme de l'art (pour revue, voir Fruehleis *et al.*(1987) ; Perun  
15 *et al.* (1989) ; Van de Waterbeemd (1994), Blundell (1996).

L'invention concerne donc une procédé de criblage d'un composé susceptible d'affecter le taux d'expression cellulaire et/ou l'activité hélicase à ARN d'un polypeptide RH116 selon l'invention et qui comporte les étapes suivantes de (i) mise en contact d'une cellule choisie parmi la cellule hôte de  
20 l'invention et une cellule eucaryote, de préférence humaine, exprimant ou contenant le polypeptide selon l'invention, et d'un ou plusieurs composés ou ligands potentiels susceptibles de pénétrer ou d'être introduits dans ladite cellule, et de (ii) détection et/ou mesure du taux d'expression cellulaire et/ou de l'activité hélicase à ARN. Le gène codant pour le  
25 polypeptide selon l'invention qui est présent dans la cellule hôte ou dans une cellule eucaryote, de préférence humaine, correspond au moins à une séquence polynucléotidique codant pour le polypeptide de l'invention de préférence sous la forme d'ADN génomique ou d'ADNc, lié de manière opérationnelle à la séquence promotrice du gène du RH116 humaine ou du  
30 gène homologue d'une espèce animale telle la souris. Les composés criblés par un tel procédé et susceptibles d'affecter le taux d'expression de l'hélicase à ARN RH116 sont de préférence des composés susceptibles

d'interagir avec les séquences polynucléotidiques régulatrices (promoteur, séquence amont, « enhancer », « silencer », « insulator », etc...) du gène codant naturellement pour le polypeptide selon l'invention ou les composés susceptibles d'interagir avec des facteurs de transcription (facteurs de transcription généraux ou facteurs tissu-spécifiques) impliqués dans la régulation de la transcription du gène codant pour le polypeptide selon l'invention, pour former un complexe susceptible d'affecter la transcription du gène codant le polypeptide RH116 de l'invention, c'est-à-dire d'augmenter, de diminuer, de moduler ou d'annuler la transcription dudit gène. Les techniques de détection et/ou de mesure de l'activité transcriptionnelle sont connues de l'homme du métier. Il convient notamment de citer les technologies de Northern Blotting et de RT-PCR qui peuvent être mises en œuvre avec les polynucléotides de l'invention utilisés comme respectivement comme sonde ou comme amorce.

15 L'invention concerne également un procédé de criblage d'un composé susceptible d'affecter l'activité hélicase à ARN d'un polypeptide selon l'invention et qui comporte les étapes suivantes de (i) mise en contact dudit polypeptide et d'un ou plusieurs composé(s) ou ligands potentiel(s) en présence de réactifs nécessaires à la mise en œuvre de l'activité hélicase à  
20 ARN, et de (ii) détection et/ou mesure de l'activité hélicase à ARN.

Selon un mode préféré de réalisation, les procédés de criblage de composés précédemment décrits se caractérisent en ce que ledit composé criblé diminue ou annihile le taux d'expression et/ou l'activité hélicase à ARN du polypeptide RH116 de l'invention. Le composé susceptible d'être  
25 obtenu par les procédés précédemment décrits et qui diminue le taux d'expression et/ou l'activité hélicase à ARN du polypeptide RH116 de l'invention est un antagoniste de la RH116 et constitue également un des objets de l'invention ; ce composé se caractérise en ce qu'il est choisi parmi un polynucléotide selon l'invention utilisé en tant que séquence  
30 d'acide nucléique antisens, un vecteur d'expression antisens selon l'invention, un anticorps selon l'invention, les muramylpeptides, de

préférence le Murabutide, ou parmi tout autre antagoniste du polypeptide selon l'invention.

En effet, le blocage partiel ou total de l'activité biologique du polypeptide RH116 de l'invention et de ses fragments constitue un moyen  
5 d'inhiber ou d'annihiler la réplication virale. En effet, l'ARN génomique non épissé et l'ARNm incomplètement épissé des rétrovirus tels le VIH nécessitent d'être exportés dans le cytoplasme pour l'empaquetage et/ou la traduction. Un tel processus est médié soit par un élément de transport agissant en CIS pour les rétrovirus simples (CIS-acting constitutive  
10 transport element ou CTE), soit par la protéine virale Rev agissant en TRANS avec des éléments de réponse (Rev Responsive element pour RRE) pour les rétrovirus complexes tels le VIH. L'hélicase à ARN RH116 selon l'invention est susceptible de constituer un cofacteur pour le CTE et jouer également un rôle dans l'expression de gène médié par le RRE et ainsi de  
15 jouer un rôle dans la réplication du VIH.

La présente invention se propose donc de fournir des composés susceptibles de bloquer partiellement ou totalement l'activité hélicase à ARN du polypeptide de l'invention pour diminuer ou annihiler la réplication de virus. Parmi ceux-ci il convient de citer un polynucléotide selon l'invention  
20 utilisé en tant que séquence d'acide nucléique antisens, un vecteur d'expression antisens selon l'invention, un anticorps selon l'invention, les muramylpeptides, de préférence le Murabutide, ou tout autre antagoniste du polypeptide selon l'invention. De préférence, les muramylpeptides et le Murabutide constituent des antagonistes de la RH116 dans les cellules de  
25 patients VIH+.

La présente invention ne se limite pas au seul virus VIH mais à l'ensemble des virus faisant intervenir directement ou indirectement une hélicase à ARN dans sa réplication. Par virus, on entend désigner les virus à ADN ou à ARN, monocaténaires ou bicaténaires, enveloppés ou non-  
30 enveloppés. De préférence, les virus appartiennent à la famille des rétroviridae, des orthomyxoviridae, des rhabdoviridae, des bunyaviridae,

des adénoviridae, des hépadnaviridae, des herpesviridae, des poxviridae. De préférence, les hépadnavirus sont les virus de l'hépatite B et de l'hépatite C.

L'activité biologique du polypeptide de l'invention n'est pas restreinte au seul contrôle post-transcriptionnel des ARN viraux, mais participe également au contrôle transcriptionnel des ARN en général. Ainsi, le polypeptide de l'invention constitue-t-il une cible thérapeutique d'intérêt pour des composés susceptibles d'altérer l'activité hélicase à ARN intervenant dans d'autres processus biologiques normaux ou pathologiques.

L'hélicase à ARN constitue une cible de choix pour des composés destinés au traitement du cancer. En effet, différents articles de la littérature scientifique traitent de l'implication des hélicases à ARN dans la tumorigénèse. Ainsi une surexpression de la protéine DEAD-box1 (DDX1) est susceptible de jouer un rôle dans la progression des tumeurs telles que le neuroblastome (Nb) et le rétinoblastome (Godbout *et al.* 1998) en altérant la structure secondaire normale et le taux d'expression des ARN des cellules cancéreuses. D'autres hélicases à ARN ont été impliquées directement ou indirectement dans la tumorigénèse. Ainsi la protéine murine p68 est mutée dans les tumeurs induites par la lumière ultra-violette ; le gène de l'hélicase à ARN DDX6 est situé au niveau du point de cassure chromosomique associé avec le lymphome des cellules B. De la même manière, une protéine chimérique comprenant DDX10 et la nucléoporine NUP98 semble être impliquée dans la pathogénèse de certaines maladies myéloïdes. Les composés susceptibles de diminuer ou d'annuler l'activité hélicase à ARN de RH116 constituent donc des composés d'intérêt destinés au traitement préventif ou curatif du cancer. Parmi les cancers susceptibles d'être traités par les composés antagonistes de la RH116, il convient de citer de manière non les cancers tels que l'adénocarcinome, la leucémie, le lymphome, le mélanome, le myélome, le sarcome, le gliome, le tératocarcinome et plus particulièrement les cancers de la glande adrénaie, de la vessie, des os, de la moelle, du sein, du tractus gastro-intestinal, du foie, du poumon, du

pancréas, de l'ovaire, de l'utérus, des testicules, de la prostate et de la gorge.

De même les composés susceptibles de diminuer ou d'annuler l'activité hélicase à ARN de RH116 constituent donc des composés d'intérêt destinés  
5 au traitement préventif ou curatif d'autres pathologies telles que le rhumatisme, les maladies héréditaires, l'arthrite, l'athérosclérose, l'ostéoporose, les maladies infectieuses aiguës et chroniques, les maladies auto-immunes, les diabètes, ainsi que les problèmes liés à la rejet des organes lors de la transplantation. Plus particulièrement, l'invention  
10 concerne les maladies immunes et auto-immunes dont fait également partie le SIDA (Syndrome d'immunodéficience acquise).

Les inhibiteurs, antagonistes et autres composés susceptibles de diminuer ou d'annuler l'activité hélicase à ARN de RH116 constituent des composés d'intérêt pour le traitement préventif ou curatif de maladies auto-  
15 immunes. En effet, il a été récemment rapporté par Takeda *et al.* (1999) que l'hélicase à ARN A agit comme un auto-antigène chez les patients atteints de lupus systémique érythémateux. Ainsi, le polynucléotide selon l'invention utilisé en tant que séquence d'acide nucléique antisens, le vecteur d'expression antisens selon l'invention, l'anticorps selon l'invention,  
20 les muramylpeptides, et de préférence le Murabutide, ou tout autre antagoniste du polypeptide RH116 selon l'invention, peuvent être utilisés pour le traitement des maladies auto-immunes. Parmi les maladies auto-immunes, il convient de citer plus particulièrement, l'uvéite, la maladie de Bechet, la sarcoïdose, le syndrome de Sjögren, la polyarthrite rhumatoïde,  
25 la polyarthrite juvénile, le syndrome de Fiessinger-Leroy-Reiter, la goutte, l'ostéoarthrose, le lupus érythémateux systémique, le lupus érythémateux aigu disséminé, la polymyosite, la myocardite, la cirrhose biliaire primitive, la maladie de Crohn, la colite ulcéreuse, la sclérose en plaques et autres maladies démyélinisantes, l'anémie aplasique, le purpura  
30 thrombocytopénique, le myélome multiple et le lymphome à lymphocytes B, le panhypopituitarisme de Simmonds, la maladie de Basedow-Graves et l'ophtalmopathie de Graves, la thyroïdite subaiguë et la maladie de

Hashimoto, la maladie d'Addison, le diabète sucré insulino-dépendant (type 1), la maladie d'Addison, le syndrome de détresse respiratoire de l'adulte, les allergies, l'anémie, l'asthme, l'anémie hémolytique auto-immune, la bronchite, la dermatite atopique, l'emphysémie, la lymphopénie épisodique.

5        Selon un autre mode de réalisation de l'invention, le procédé de criblage de composé de l'invention se caractérise en ce que ledit composé criblé augmente le taux d'expression et/ou l'activité hélicase à ARN du polypeptide RH116 de l'invention. Le composé ainsi criblé, également objet de l'invention, constitue un agoniste du polypeptide selon l'invention. Parmi  
10    les composés agonistes susceptibles d'être obtenu par le procédé de l'invention, il convient de citer les muramylpeptides, et de préférence le Murabutide ; en effet, les inventeurs ont démontré expérimentalement que ce dernier composé, le Murabutide, augmente le taux d'expression de RH116 dans les cellules de donneurs sains, se comportant ainsi comme un  
15    agoniste du polypeptide de l'invention.

      Selon un autre mode de réalisation, l'invention porte sur un procédé de criblage de composés susceptibles d'affecter l'activité fonctionnelle d'un polypeptide selon l'invention. Un tel procédé comporte les  
20    étapes de (i) mise en contact dudit polypeptide et d'un ou plusieurs ligand(s) potentiel(s) en présence de réactifs nécessaires à la mise en œuvre d'une réaction choisie parmi la réaction d'épissage de l'ARN nucléaire et/ou mitochondrial, de réaction d'édition de l'ARN, de réaction de maturation de l'ARNr, de réaction d'initiation de la traduction, de la réaction d'export des  
25    ARNm nucléaire vers le cytoplasme, et la réaction de dégradation des ARNm et de (ii) détection et/ou mesure de ladite réaction. Le composé criblé susceptible d'être obtenu par le procédé entre également dans la portée de cette invention.

      L'invention concerne donc un composé caractérisé en ce qu'il est choisi  
30    parmi un anticorps selon l'invention, un polypeptide selon l'invention, un polynucléotide selon l'invention, un oligonucléotide selon l'invention, un vecteur selon l'invention, un vecteur d'expression antisens selon l'invention,

une cellule selon l'invention, un composé susceptible d'être obtenu par les différents procédés de criblage selon l'invention à titre de médicament et notamment en tant que principes actifs de médicament ; ces composés seront préférentiellement sous forme soluble, associés à un véhicule pharmaceutiquement acceptable. Par véhicule pharmaceutiquement acceptable, on entend désigner tout type de véhicule employé habituellement dans la préparation de compositions injectables, c'est-à-dire un diluant, un agent de suspension tel une solution saline isotonique ou tamponnée. De préférence, ces composés seront administrés par voie systémique, en particulier par voie intraveineuse, par voie intramusculaire, intradermique ou par voie orale. Leurs modes d'administration, posologies et formes galéniques optimaux peuvent être déterminés selon les critères généralement pris en compte dans l'établissement d'un traitement adapté à un patient comme par exemple l'âge ou le poids corporel du patient, la gravité de son état général, la tolérance au traitement et les effets secondaires constatés, etc.

De préférence, les composés de l'invention à titre de médicament sont destinés à la prévention et/ou au traitement de pathologies sélectionnées dans le groupe composé du cancer, des maladies infectieuses aiguës ou chroniques telles les infections par le VIH ou le virus de l'hépatite B ou C, des maladies génétiques héréditaires, des maladies immunes et auto-immunes, du rhumatisme, de l'arthrite, de l'athérosclérose, de l'ostéoporose, des diabètes et à la prévention des rejets de greffes d'organe.

Parmi les composés à titre de médicament de l'invention, les composés antagonistes du polypeptide RH116 sont particulièrement préférés pour la préparation d'un médicament destiné au traitement de pathologies virales tel que le syndrome d'immunodéficience acquise (SIDA) ou l'hépatite.

C'est également un des objets de la présente invention de fournir une composition pharmaceutique pour le traitement préventif et curatif de pathologies virales et notamment du SIDA ou de l'hépatite caractérisée en ce qu'elle contient une quantité thérapeutiquement efficace d'un composé antagoniste d'un polypeptide RH116 et d'un véhicule

pharmaceutiquement acceptable. Plus particulièrement l'invention vise également à fournir un produit comprenant au moins un composé antagoniste de la RH116 et au moins un autre agent antiviral comme produit de combinaison pour une utilisation simultanée, séparée ou étalée dans le temps en thérapie anti-virale, de préférence anti-VIH. Cet autre agent antiviral est de préférence choisi parmi (i) les inhibiteurs nucléotidiques ou non-nucléotidiques de la reverse transcriptase tel que par exemple le 3'-azido-3'déoxythymidine (AZT), le 2',3'-didéoxyinosine (ddI), le 2',3'-didéoxycytidine (ddC), le (-)2',3'-didéoxy-3'-théacytidine (3TC), le 2',3'-didéhydro-2',3'didéoxythymidine (d4T), le (-)2'-déoxy-5-fluoro-3'-théacytidine (FTC), le TIBO, le HEPT, le TSAO, l' $\alpha$ -APA, la névirapine, le BAHP, l'acide phosphonoformique (PFA) et (ii) les inhibiteurs de la protéase virale tels l'indinavir et le saquinavir.

Selon un autre mode de réalisation, il est également dans l'étendue de l'invention de fournir une méthode de traitement thérapeutique ou prophylactique de maladie associée à une augmentation de l'expression ou de l'activité du polypeptide RH116 selon l'invention. Cette méthode comprend d'administrer à un patient qui nécessite un tel traitement, une quantité thérapeutiquement efficace d'un antagoniste du polypeptide de l'invention.

L'invention porte également sur l'utilisation d'un polypeptide selon l'invention, d'un polynucléotide selon l'invention, d'un composé agoniste d'un polypeptide RH116 selon l'invention pour la préparation d'un médicament destiné au traitement prophylactique de pathologies. En effet, les vaccins en général induisent une immunité à une infection ou à une maladie en générant une réponse immune chez un patient contre un antigène spécifique associé avec l'infection ou avec la maladie. Cependant, dans beaucoup de cas, les antigènes purifiés sont de faibles immunogènes. Dans de tels cas, des agents immunomodulateurs tels un adjuvant ou un immunostimulant doivent être employés pour augmenter la réponse immune. Cependant, l'un des désavantages de nombreux adjuvants employés à l'heure actuelle est leur toxicité. Il existe donc un réel besoin

d'identifier de nouveaux composés qui permettent d'augmenter les réponses immunes spécifiques ; c'est ce que se propose de faire la présente invention. En effet, les polypeptides et les polynucléotides de l'invention et les composés agonistes d'un polypeptide RH116 selon l'invention peuvent être  
5 utilisés pour la préparation d'un médicament destiné à provoquer ou à augmenter la réponse immune à un vaccin chez un patient.

D'autres caractéristiques et avantages de l'invention apparaissent dans la suite de la description avec les exemples représentés ci-après. Dans ces  
10 exemples on se référera aux figures suivantes.

**FIGURE 1 : Stratégie utilisée pour l'obtention de l'extrémité 3' de l'ADNc codant pour le RH116.**

Le court fragment supérieur est le fragment initial de 164 pb obtenu  
15 par le DD-RT PCR et correspond à la région 3' de l'ADNc codant pour le polypeptide RH116 de l'invention. Ce fragment plus long de 1260 pb a été obtenu après un 3' RACE (trait inférieur).

**FIGURE 2 : Stratégie d'obtention de la région 5' de l'ADNc codant pour le RH116. (A)** La partie de l'extrémité 5' a été obtenue après deux  
20 réactions de PCR (5'RACE) pour Rapid Amplification of cDNA Ends, i.e. amplification rapide des extrémités des ADNc, les fragments obtenus sont symbolisés en (...) (R3) et en (----) (R8). Cette stratégie a donc permis aux inventeurs d'obtenir un cadre de lecture ouvert (ORF open reading frame)  
25 de 853 acides aminés. Les oligonucléotides HELI 1 et HELI 2 qui serviront d'amorces pour les RT-PCR, sont indiqués. **(B)** La partie 5' terminale de l'ADNc correspondant au polypeptide RH116 a été obtenue après une réaction de PCR à partir de l'amorce R16 (5'RACE). Le fragment obtenu de 964pb est symbolisé en pointillé (- - - - -).

30

**FIGURE 3 : Séquence en amino-acides de l'hélicase à ARN putative humaine RH 116 et comparaison avec l'hélicase ARN putative (RIG-1). (A)**

Représentation schématique des régions en amino-acides hautement conservées caractéristiques des ARN hélicases ATP dépendantes putatives de la famille des protéines « DEAD-BOX ». Le motif conservé est encadré (le « x » indique n'importe quel amino-acide), un motif SAT, -A- est l'acide aminé hautement conservé, S et T peuvent être remplacés par n'importe quel acide aminé dans n'importe quel cas individuel (Luking *et al.*, 1998). Dans le motif VI, (X)<sub>1-2</sub> indique n'importe lequel des 2 acides aminés. **(B)** Séquence en amino-acides de la protéine RH 116 déduite à partir de la séquence d'ADNc (N° d'accèsion EMBL AY017378). 6 motifs conservés caractéristiques de la famille des protéines « DEAD-BOX » sont indiqués par des lettres en gras. **(C)** Alignement de séquence du cadre de lecture ouvert de 1025 aa (clone 10.5) (séquence supérieure) avec une hélicase à ARN humaine (RIG-1) décrite par SUN Y.W. (Numéro d'accès AF 038963) (séquence inférieure). Les homologies de séquences sont présentées dans la ligne intermédiaire aux deux séquences. Les séquences encadrées correspondent aux domaines conservés des hélicases à ARN ; les domaines G-GKT et DEXH correspondent aux domaines ATPase et les domaines SAT (-A-) et RGR-R (GR--R), spécifiques des hélicases à ARN, sont responsables de la liaison et du déroulement de l'ARN cible.

**FIGURE 4 : Analyse en Northern-blot de l'expression de l'ARN messager RH 116 dans de multiples tissus humains (CLONTECH).**

L'ARN polyA<sup>+</sup> obtenu à partir de cerveau (1), de cœur (2), de muscle squelettique (3), de côlon (4), de thymus (5), de rate (6), de rein (7), de foie (8), d'intestin grêle (9), de placenta (10), de poumon (11), de leucocytes du sang périphérique (12) ont été hybridés avec une sonde oligonucléotidique dérivée de l'ARN messager de RH 116 et marquée à son extrémité au phosphore <sup>32</sup>P (partie supérieure de la figure). Des quantités équivalentes d'ARN polyA<sup>+</sup> sont également hybridées avec une sonde β-actine utilisée comme contrôle (partie inférieure de la figure). A la fois dans le cœur et dans le muscle squelettique, il existe deux formes d'ARN messager β-actine, une forme de 2 kb et une forme de 1,6-1,8 kb. Cette différence de taille n'est pas due à une dégradation de l'ARN messager mais à l'hybridation de la

sonde avec soit la forme  $\alpha$  soit la forme  $\gamma$  de l'actine (voir les instructions du fabricant CLONTECH).

**FIGURE 5 : Inhibition de l'expression de l'ARN messenger RH 116 induite par le murabutide dans des PBMC infectés de manière endogène et déplétés en CD8.** Les PBMC activés par la PHA et déplétés en CD8 sont maintenus en culture, en absence (medium) ou en présence de murabutide à 10  $\mu$ g/ml pendant 6 ou 24 heures. **(A)** Des échantillons d'ARN (20, 100 et 500 ng) sont sujets à une amplification par RT PCR avec une paire d'amorces spécifiques pour détecter l'ARN messenger RH 116. L'ARN messenger GAPDH (contrôle interne) exprimé de manière constitutive est également amplifié dans les mêmes échantillons. **(B)** Le pourcentage d'inhibition moyen de l'expression de l'ARN messenger de RH 116 est observé dans des échantillons de 5 patients infectés par le VIH-1.

15

**FIGURE 6 : Immuno-précipitation de la protéine RH 116 à partir d'extraits cellulaires marqués à l'iode  $^{125}$ I obtenues à partir de cellules U937.** **(A)** Extraits cellulaires totaux (1, 2) ou extraits nucléaires (3, 4) sont marqués et immuno-précipités en utilisant 20  $\mu$ g d'immunoglobulines purifiées à partir de sérum normal (1, 3) ou de sérum murin obtenu contre la protéine RH 116<sub>1-335</sub> (2,4). Les échantillons sont chargés sur un gel SDS-PAGE à 8%. **(B)** La qualité du fractionnement cellulaire est vérifiée par une analyse en Western-Blot en utilisant un anticorps dirigé contre la protéine de 58 kDa de l'appareil de Golgi (protéine cytoplasmique) et un anticorps anti-histone 1 (protéine nucléaire) dans des extraits totaux (1) et/ou des extraits nucléaires (2).

**FIGURE 7 : L'immuno-localisation de la protéine RH 116 dans les cellules HeLa-CD4+.** Les cellules Hela-CD4+ fixées au méthanol sont incubées avec du sérum murin normal **(A)**, des anticorps monoclonaux anti-histone H1 (Santa Cruz) **(B)** ou un anti-sérum murin dirigé contre la

30

protéine RH 116<sub>1-335</sub> **(C)**, et sont colorées avec des anticorps IgG anti-souris de chèvre conjugués au FITC et/ou avec l'iodure de propidium. L'immunofluorescence montre une coloration au FITC verte (1), une contre-coloration nucléaire rouge correspondant à l'iodure de propidium (2) et la double  
5 coloration (3).

**FIGURE 8 : Expression de l'antigène viral P24 par des cellules HeLa CD4+ transfectées et infectées par le virus VIH-1<sub>LAI</sub>.** Les cellules HeLa CD4+ sont transfectées de manière transitoire, d'une part avec le  
10 plasmide pEGFP (GFP) ou avec le plasmide pEGFP dans lequel sont respectivement clonés les ADNc de RH 116 ou de SSA (GFP-RH 116 et GFP-SSA respectivement) et, d'autre part, avec le plasmide pCR3 ou le plasmide pCR3 dans lequel est cloné Tat (pCR3-Tat) ; les cellules sont infectées 24 heures après par le virus VIH-1<sub>LAI</sub>. A partir du 2<sup>ème</sup> jour après l'infection  
15 des cellules, les milieux de culture sont collectés quotidiennement jusqu'au 5<sup>ème</sup> jour.

La réplication virale est évaluée par le relargage de l'antigène P24 par les cellules (concentration cellulaire normalisée à  $5 \times 10^5$  cellules/ml) dans les surnageants des cultures de cellules transfectées avec la GFP, GFP-RH  
20 116 ou GFP-SSA **(A)** et de cellules transfectées avec pCR3 ou pCR3-Tat **(B)**. **(C)** Augmentation de l'antigène P24 dans les cellules infectées et transfectées par GFP-RH 116, GFP-SSA ou Tat comparées avec les contrôles correspondants respectifs (GFP ou pCR3 respectivement). Les résultats correspondent aux valeurs moyennes plus ou moins l'écart type des valeurs  
25 obtenues pour 5 (GFP-RH 116 et GFP-SSA) et pour 3 expériences indépendantes (pCR3-Tat).

**FIGURE 9 : RT PCR semi-quantitative représentative de l'expression de l'ADN viral et de l'ARN messenger dans des cellules**  
30 **HeLa-CD4+ transfectées et infectées par le virus VIH-1<sub>LAI</sub>.** Les échantillons d'ARN (500, 100, 20 et 4 ng) sont sujets à des amplifications

RT PCR avec des amorces spécifiques pour détecter une surexpression des ARNm codant pour la protéine RH 116, SSA-56 ou TAT **(A)**. Les échantillons d'ARN sont soumis à des amplifications de RT PCR pour détecter l'ARNm viral **(B)** avec une paire d'amorces GAG04-GAG06 pour détecter les ARNm non épissés GAG ou POL et avec une paire d'amorces BSS/KPNA pour détecter des transcrits viraux de taille intermédiaire ou simplement épissés. Ces ARN messagers sont nommés sur la base des exons qu'ils contiennent et des protéines qu'ils produisent (Neuman, 1994). **(C)** L'ADN total est extrait 24 heures après l'infection et des concentrations variables (150, 30, 6 et 1,2 ng) sont utilisées comme matrice pour une amplification par PCR au moyen d'une paire d'amorces GAG04-GAG06 afin de détecter le gène GAG du VIH-1. L'équivalent cellulaire est déterminé par amplification des gènes  $\beta$ -actine ou GAPDH.

**FIGURE 10 : Détection d'auto-anticorps dirigés contre la protéine RH 116 dans des patients contrôle sains et des patients infectés par le VIH avant et après traitement anti-rétroviral.** Les barres horizontales reflètent les valeurs moyennes arithmétiques.

## **EXEMPLES**

### **MATERIELS ET METHODES**

#### **1. Cellules et conditions de culture :**

5 Les cellules HeLa-CD4<sup>+</sup> (fournies par le Docteur HUBER, LABORATOIRE DE VIROLOGIE, Lille, France) et les cellules U937 sont cultivées respectivement dans du milieu DMEM ou dans du milieu RPMI 1640 complémenté avec 10% de sérum de veau foetal inactivé par la chaleur (LIFE-TECHNOLOGIES-INVITROGEN, Cergy-Pontoise, France), 2 mM L-  
10 glutamine (LIFE-TECHNOLOGIES-INVITROGEN, Cergy-Pontoise, France) et 1% de gentamycine (SHERING-PLOUGH, Levallois-Perret, France).

#### **2. Préparation des cellules mononucléaires humaines du sang périphériques (PBMC) :**

15 Le sang veineux est collecté à partir de patients séropositifs au VIH-1 et des cellules mononucléaires du sang périphérique (PBMC) sont isolées par centrifugation sur Ficoll-Hypaque (AMERSHAM PHARMACIA BIOTECH, Uppsala, Suède). Les cellules sont lavées avec soin pour éliminer les plaquettes et sont déplétées en cellules CD8<sup>+</sup> en utilisant des billes  
20 électromagnétiques couvertes d'un anticorps anti-CD8 (DYNAL, Oslo, Norvège). Les PBMC déplétés obtenus par la mise en œuvre du protocole du fabricant contiennent moins de 3% de cellules CD8<sup>+</sup> après analyse par cytométrie de flux. Les cellules sont ensuite cultivées dans du RPMI 1640 supplémenté avec 10% de sérum AB inactivé par la chaleur  
25 (ETABLISSEMENT DE TRANSFUSION SANGUINE, Lille, France) puis stimulées avec de la phytohémaglutinine (PHA) (5 µg/ml pendant 3 jours). Les PBMC sont ensuite cultivés puis récupérés en présence ou en absence de murabutide (N-acétyl-muramyl-L-alanyl-D-glutamine-*n*-butyl ester) (fournis par ISTAC SA, Lille, France) (10 µg/ml) et d'interleukine 2 (IL-2) (10  
30 µg/ml) pendant 6 ou 24 heures.

#### **3. Expérience de « differential display-RT-PCR » (DD-RT-PCR) :**

La DD-RT-PCR a été réalisée sur des extraits d'ARN totaux de PBMC activés par la PHA et déplétés en CD8 obtenus de patients VIH-1, tels que précédemment décrits (Liang et Pardee, 1992 avec certaines modifications). Brièvement, 1 mg d'ARN total est incubé à 75°C pendant 10 minutes en présence de 2 mM d'amorces oligo (dT), T<sub>12</sub>BA (B étant un mélange de G, T, C), T<sub>12</sub>VT (V étant un mélange de A, G, C), T<sub>12</sub>HG (H étant un mélange de A, C, T), T<sub>12</sub>DC (D étant un mélange de A, G, T) et d'eau dépourvue de ribonucléase, dans un volume final de 11 µl. L'ADN complémentaire (ADNc) est ensuite synthétisé à l'aide de la reverse-transcriptase RNase H Superscript (LIFE-TECHNOLOGIES-INVITROGEN, Cergy-Pontoise, France) dans 25 µl contenant 10 mM de dithiotréitol, 20 µM de dNTP pendant une heure à 37°C avant une dénaturation de 10 min à 95°C. L'ADNc synthétisé est ensuite amplifié avec une unité de polymérase Ampli Taq Gold (APPLIED BIOSYSTEMS, Foster City, CA, Etats-Unis), 2,5 mM de MgCl<sub>2</sub>, 2 µM de dNTP, 2 µCi de <sup>33</sup>P-dATP (AMERSHAM PHARMACIA BIOTECH) en utilisant des amorces spécifiques en aval et des amorces aléatoires en amont : AP0: TAT CGA CTC CAA G (SEQ ID N° 27); AP1: TTA GCT AGC ATG G (SEQ ID N° 28); AP2: TGC TAA GAC TAG C (SEQ ID N° 29); AP3: TTG CAG TGT GTG A (SEQ ID N° 30); AP4: TGT GAC CAT TGC A (SEQ ID N° 31); AP5: TGT CTG CTA GGT A (SEQ ID N° 32); AP6 : TGC ATG GTA GTC T (SEQ ID N° 33); AP7: TGT GTT GCA CCA T (SEQ ID N° 34); AP8: TAG ACG CTA GTG T (SEQ ID N° 35); AP9: TTA GCT AGC AGA C (SEQ ID N° 36); AP10: TCA TGA TGC TAC C (SEQ ID N° 37); AP11: TAC TCC ATG ACT C (SEQ ID N° 38); AP12: TAT TAC AAC GAG G (SEQ ID N° 39); AP13: TAT TGG ATT GGT C (SEQ ID N° 40); AP14: TAT CTT TCT ACC C (SEQ ID N° 41); AP15: TAT TTT TGG CTC C (SEQ ID N° 42); AP16: TTA TCT ATA CAG G (SEQ ID N° 43); AP17: TTA TGG TAA AGG G (SEQ ID N° 44); AP18: TTA TCG GTC ATA G (SEQ ID N° 45); AP19: TTA GGT ACT AAG G (SEQ ID N° 46). Les paramètres de l'amplification sont 1 min à 94°C, 1 min à 40°C, 1 min à 72°C, suivies par une étape d'extension finale de 5 min à 72°C. Les produits PCR sont ensuite séparés par électrophorèse sur un gel à 6%

polyacrylamide-8 M urée. Après séchage, le gel est ensuite exposé sur un film Hyperfilm-HP. Des bandes ADNc exprimées de manière différente sont excisées du gel, éluées dans 100 µl d'eau stérile, précipitées puis resuspendues dans 10 µl d'eau stérile. Les fragments d'ADNc sont ensuite réamplifiés en utilisant le même couple d'amorces décrit précédemment et subséquemment clonés dans le vecteur de clonage TOPO TA PCR II (INVITROGEN, Groningen, Pays-Bas).

#### **4. Séquençage de l'ADN :**

Les ADNc clonés sont utilisés comme matrice pour le séquençage en utilisant le kit de séquençage « PRISM Ready Reaction Dye Deoxy Terminator » de APPLIED BIOSYSTEMS. Les échantillons sont soumis à électrophorèse sur un séquenceur d'ADN ABI 377 ; il sont lus automatiquement puis enregistrés en utilisant le logiciel ABI Prism, version 2.2.1, de APPLIED BIOSYSTEMS. Des recherches d'homologie dans des banques de données de nucléotides ont été réalisées en utilisant le programme Basic Local Alignment Search Tool (BLAST).

#### **5. RT PCR semi-quantitative pour la détection de l'expression de gènes :**

L'ARN cellulaire total extrait en utilisant le kit RNA plus (Q-BIOgene, Illkirch, France) est prétraité avec de la DNaseI. La synthèse du premier brin sur l'ARN polyA<sup>+</sup> a été amorcée avec une amorce p(dT)<sub>15</sub> (BOEHRINGER MANNHEIM, ROCHE DIAGNOSTICS, France) et réalisée par la reverse-transcriptase dérivée du virus de la leucémie murine de Moloney (M-MLV) (PROMEGA, Madison, CA, USA) (1 heure à 37°C, 3 min à 92°C). L'ADNc résultant est ensuite soumis à 25 à 35 cycles répétés d'amplification avec l'ADN polymérase AmpliTaq Gold (APPLIED BIOSYSTEMS). L'amplification PCR des séquences GAPDH est réalisée pour obtenir un contrôle interne. Les amorces oligonucléotidiques spécifiques sont les suivantes : GAPDH [Brin sens : 5'- GCC ATC AAT GAC CCC TTC ATT GAC-3' (SEQ ID N°19) ; brin antisens : 5'- TGA CGA ACA TGG GGG

CAT CAG CAG -3' (SEQ ID N° 20)], RH 116 (Brin sens : 5'- GGA AGT ACA  
 ATG AGG GCC TAC AAA-3' (SEQ ID N°13) ; brin antisens : 5'-TCC TCA GTC  
 CTA GTA TAT TGC TCC-3' (SEQ ID N° 14)], SSA-56 [Brin sens : 5'-GAA  
 AGA GAG GTC GCA GAG GCC TGT-3' (SEQ ID N°47); brin antisens : 5'-TGA  
 5 TAA GGC TGA GGA AGG GAA ATG-3' (SEQ ID N°48)], Tat (Brin sens : 5'-  
 CTA GAC CCC TGG AAG CAT CCA-3' (SEQ ID N°49); brin antisens : 5'-TCG  
 GGC CTG TCG GGT CCC CTC-3' (SEQ ID N°50)].

Tous les produits PCR sont ensuite séparés sur gels d'agarose à 2%  
 puis visualisés par une coloration au bromure d'éthidium. En utilisant des  
 10 systèmes d'imagerie (IMAGE MASTER 1D PRIME, AMERSHAM PHARMACIA  
 BIOTECH), le pourcentage de variation d'expression d'ARN messager a été  
 déduit après normalisation par rapport aux taux des standards internes  
 correspondants, tel que précédemment décrit (Amiel *et al.*, 1999).

## 15 **6. Clonage de l'ADNc pleine longueur :**

Le clonage de l'ADNc pleine longueur RH116 a été réalisé en utilisant  
 le kit d'amplification SMART<sup>TM</sup> RACE cDNA (CLONTECH, Palo Alto, CA,  
 USA) sur des échantillons d'ARN polyA+ à partir de PBMC de patient sain  
 déplétés en CD8. De plus, une banque humaine de rate ( $\lambda$  TriplEX<sup>TM</sup> Library  
 20 CLONTECH) a été criblée avec des sondes spécifiques marquées au  
 phosphore <sup>32</sup>P selon les instructions du fabricant.

## **7. Analyse par Northern Blotting :**

L'analyse par Northern Blotting a été réalisée en utilisant le kit  
 25 « Multiple Tissue Northern » (MTN<sup>TM</sup>) (CLONTECH) selon les instructions du  
 fabricant. L'ARN polyA+ a été hybridé avec les oligonucléotides 5'-CGT GCT  
 GAT TCC TCA GTC CTA GTA TAT TGC-3' (SEQ ID N° 26) and 5'-GCA TCT  
 GCA ATG GCA AAC TTC TTG CAT GGC-3' (SEQ ID N° 18) e dérivés de la  
 séquence ADNc de RH 116 et marqués aux extrémités par du phosphore  
 30 <sup>32</sup>P. Les membranes sont réhybridées avec une sonde ADNc de  $\beta$ -actine  
 humaine.

## **8. Expression de la protéine marquée avec un His-Tag et production d'anticorps polyclonaux de souris :**

Un fragment partiel correspondant au 335 premiers acides aminés de l'ARN messenger RH116 a été amplifié en utilisant des oligonucléotides spécifiques (SEQ ID N°21 et SEQ ID N°22) et a été sous-cloné dans le vecteur pQE-81 (QIAGEN, Courtabœuf, France), préalablement coupé avec les enzymes BamHI et Sall. Le plasmide résultant pQE-81-His6-RH116<sub>1-335</sub> est ensuite transformé dans les cellules *E. coli* TOP10F'. Les cellules transformées sont ensuite cultivées et induites par 1 mM d'isopropyl-1-thio-β-D-galactopyranoside pendant 5 heures. La purification en conditions dénaturantes est réalisée en utilisant des billes Ni-NTA selon les recommandations du fabricant (QIAexpressionist™ QIAGENE).

La protéine recombinante partielle RH 116<sub>1-335</sub> (50 µg), émulsifiée dans un adjuvant complet de Freund (SIGMA, Saint-Louis, MI, USA), est utilisée pour immuniser une souris par injection intrapéritonéale. Après 30 jours, les animaux sont stimulés avec 50 µg de la même protéine recombinante émulsifiée avec l'adjuvant incomplet de Freund (SIGMA) et 15 jours plus tard avec 50 µg de la même protéine sans adjuvant. Les sera sont collectés 10 jours après la dernière stimulation et sont testés pour le taux d'anticorps anti-RH116. Les immunoglobulines sont ensuite purifiées à partir des anti-séra selon la méthode utilisant l'acide caprylique (Steinbuch et Audran, 1969) et concentrées par précipitation au sulfate d'ammonium saturé. Avant l'immunisation, les sera sont collectés et utilisés comme contrôle négatif. Les activités d'anti-séra et d'immunoglobuline sont testées par ELISA, tel que décrit ci-dessous.

## **9. Analyses par immunofluorescence indirecte :**

Pour les analyses par immunofluorescence indirecte, les cellules HeLa-CD4<sup>+</sup> sont cultivées sur des lames à chambre (NALGE, Nunc, Rochester, NY, USA) et fixées et perméabilisées avec un mélange méthanol/acétone (2V/1V) pendant 10 min à -20°C. Après humidification dans du PBS contenant 1% de sérum albumine bovine (BSA) pendant 30

min, les échantillons sont incubés avec le premier anticorps à température ambiante pendant 45 min dans du PBS contenant 0,5% de BSA. Le sérum de souris normale et l'anti-sérum dirigé contre RH116<sub>1-335</sub> est utilisé à une dilution de 1/20ème ; l'anticorps monoclonal IgG2a anti-histone 1 (SANTA CRUZ BIOTECHNOLOGIES, USA) et un anticorps-contrôle qui correspond à ce type d'isotype sont utilisés à une concentration de 1/50ème. Après 3 lavages dans du PBS, les échantillons sont colorés avec un anticorps de chèvre IgG anti-souris conjugué à du FITC (fluorescéine isothiocyanate) (SIGMA). Après plusieurs lavages, les échantillons sont incubés 3 min dans du PBS contenant 0,5 µg/ml d'iodure de propidium puis examinés par microscopie de fluorescence.

#### 10. Fractionnement sub-cellulaire :

Le fractionnement sub-cellulaire est préparé tel que décrit dans Li-Ru *et al.* (1999). Brièvement, pour la préparation des lysats cellulaires totaux, les cellules sont lysées dans un tampon contenant 10 mM Tris-HCl pH 7,1, 1 mM EDTA, 1% triton X-100, 1 mM PMSF, 1 mM Na<sub>3</sub>VO<sub>4</sub>, 10 µM E-64 (Trans-Epoxy-Succinyl-L-Leucylamido-(4-guanidino)-butane), 1 µg/ml de pepstatine et 0,1% d'aprotinine. Pour la préparation des fractions nucléaires, les culots cellulaires sont incubés dans un tampon hypotonique (20 mM HEPES pH 7,4, 1 mM MgCl<sub>2</sub>, 10 mM KCl, 0,5% Nonidet P40, 0,5 mM dithiotréitol (DTT), 1 mM PMSF, 1 mM Na<sub>3</sub>VO<sub>4</sub>, 10 µM E64, 1 µg/ml pepstatine, 0,1 % aprotinine à 4°C pendant 30 min. Après centrifugation, les culots résultant contenant les noyaux sont resuspendus dans un tampon à haute force ionique (1 mM HEPES, pH7,4, 20% glycérol, 0,4 M NaCl, 1 mM MgCl<sub>2</sub>, 10 mM KCl, 0,5 mM DTT, 1 mM PMSF, 1 mM Na<sub>3</sub>VO<sub>4</sub>, 10 µM E64, 1 µg/ml pepstatine, 10 µg/ml leupeptine, 0,1 % aprotinine. La qualité du fractionnement est contrôlée par Western-blotting en utilisant des anticorps monoclonaux dirigés contre l'histone 1 (SANTA-CRUZ BIOTECHNOLOGY) (protéine nucléaire) et dirigée contre la protéine de Golgi 58K (SIGMA) (protéine cytoplasmique).

**11. Analyse par Western-blotting :**

Les lysats cellulaires sont mélangés avec un tampon de dépôt 2 X SDS incubés à 100°C pendant 5 min puis séparés par électrophorèse sur un gel polyacrylamide SDS-PAGE à 15% et enfin transférés sur une membrane de nitrocellulose (Hybond-C-Amersham). Les membranes sont saturées avec du PBS (phosphate Buffered Saline) contenant 5% de lait non gras pendant 1 heure à température ambiante, suivie par une incubation de l'anticorps primaire (dilution 1/100ème pour l'anticorps anti-histone 1 ; dilution 1/1000ème pour l'anticorps anti-Golgi 58 K) à 4°C pendant la nuit. Après des lavages avec du PBS 0,1% Tween-20, les membranes sont incubées avec un second anticorps conjugué à la peroxydase de raifort (1/1000) (SIGMA) à température ambiante pendant 1 heure. Les bandes transférées sur les membranes sont détectées par incubation avec un réactif de chemiluminescence augmentée (ECL) (Amersham) et exposées à un film KODAK XOMAT.

**12. Immunoprécipitation :**

500 µg de chaque extrait cellulaire sont radiomarqués avec 500 µCi d'iode <sup>125</sup>I, en utilisant la méthode à la chloramine T. 3 x 10<sup>6</sup> CPM d'extraits cellulaires sont pré-traités pendant 2 heures avec un sérum de souris normale et 50 µl de protéine G sépharose (PHARMACIA) dans un tampon TNSTEN (50 mM Tris-HCl pH8,2, 0,5 M NaCl, 0,1% SDS, 0,5% Triton X100, 5 mM EDTA, 0,02 % NaN<sub>3</sub>, avec 0,1% d'aprotinine). Les immunoprécipitations sont réalisées en ajoutant 20 µg d'immunoglobulines purifiées à partir du sérum anti-RH116<sub>1-335</sub> ou de sérum normal et 50 µl de protéine G sépharose. Après une incubation sur la nuit à 4°C, les billes de sépharose sont lavées 10 fois avec un tampon TNSTEN et ensuite resuspendues puis bouillies dans un tampon de dépôt SDS-PAGE 2X concentré. Les protéines sont séparées sur un gel SDS PAGE à 8% puis analysées par autoradiographie.

### 13. Transfection de l'ADN :

Tous les plasmides utilisés sont préparés en utilisant des matériaux dépourvus d'endotoxine (« EndoFree<sup>TM</sup> », Giga Kit, QIAGEN). Pour l'expression transitoire dans les cellules HeLa-CD4+, un fragment XhoI-BamHI comprenant l'ADNc RH116 est sous-cloné dans un vecteur d'expression de mammifère pEGFP-N1 (CLONTECH). La construction chimérique résultante consiste en une protéine RH116 pleine longueur fusionnée à l'extrémité carboxy-terminale de la protéine de fluorescence verte (GFP). La fusion dans un cadre correct de lecture de l'ADNc est contrôlée par séquençage. La séquence codante d'une protéine quelconque telle SSA (disponible dans le laboratoire des inventeurs) a été sous-clonée dans le vecteur pEGFP-N1 (GFP-SSA) et utilisées comme contrôle. Pour le contrôle positif, les inventeurs ont étudié les effets d'une expression transitoire de Tat cloné dans pCR3 tel que précédemment décrit (Billaut-Mulot *et al.*, 2001) ; le pCR3 natif constitue le contrôle correspondant. La transfection est réalisée en utilisant Effectene (QIAGEN) comme réactif de transfection et utilisé selon le protocole du fabricant. Brièvement, la veille de la transfection,  $2 \times 10^5$  cellules sontensemencées par puits dans des plaques de 12 puits et sont subséquemment transfectées en utilisant 1 µg d'ADN et 10 µl d'Effectene dans des expériences en triplicata. Les cellules sont infectées par le VIH-1<sub>LAI</sub> 24 heures après la transfection.

### 14. Infections *in vitro* :

La souche de cellules T présentant un tropisme pour le VIH-1<sub>LAI</sub> est obtenue à partir du Laboratoire de Virologie Centrale de Lille, France. Pour infecter des cellules transfectées,  $5 \times 10^4$  CPM d'activité reverse-transcriptase virale (RT) sont ajoutés dans chaque puits. Les cellules sont incubées sur la nuit à 37°C. Les virus sont ensuite retirés et les cellules sont lavées 2 fois et un milieu frais est ajouté. 1 jour après l'infection et pendant les 4 jours suivants, les surnageants sont collectés dans chacun des puits et les cellules sont récupérées et comptées. La réplication virale est évaluée par la détection de l'ADN et de l'ARN du VIH-1, 24 heures après

l'infection ou par la détection de la présence de la protéine p24 dans les surnageants entre 2 jours et 5 jours après l'infection.

### 15. Détection de l'ADN et de l'ARN du VIH-1

5 L'ADN cellulaire total est extrait des cellules infectées par le VIH-1 24 heures après l'infection, puis soumis à 35 cycles d'amplification avec l'ADN polymérase « AmpliTaq Gold ». L'amplification PCR des séquences de  $\beta$ -actine est réalisée avec les oligonucléotides 5'-GGG TCA GAA GGA TTC CTA TG-3' (SEQ ID N°51) et 5'-GGT CTC AAA CAT GAT CTG GG-3' (SEQ ID  
10 N°52) pour standardiser l'équivalence entre les cellules. L'ADN proviral VIH-1 de chaque échantillon est mesuré en utilisant l'amorce GAG06 (5'-GCI TTI AGC CCI GAA GTI ATA CCC ATG-3'; SEQ ID N° 53) et GAG04 (5'-CAT ICT ATT TGT TCI TGA AGG GTA CTA G-3'; SEQ ID N°54). Pour mesurer le taux d'ARN de VIH-1, l'ARN cellulaire total est extrait en utilisant le kit RNAlplus  
15 (Q-BIOgene) et ensuite amplifié en utilisant la rTth polymérase (Applied Biosystems) en présence de la paire d'amorces GAG06-GAG04 pour détecter l'ARNm GAG-POL non épissé du VIH-1 et en présence de la paire d'amorces BSS (5'-GGC TTG CTG AIG NGC ICA CIG CAA GAG G-3' ; SEQ ID N°55)-KPNA (5'-AGA GTI GTG GTT GNT TCN TTC CAC ACA G-3' ; SEQ ID N°56)  
20 pour détecter la taille intermédiaire de l'ARN messenger simplement épissé tel que précédemment décrit (Amiel *et al.*, 1999). Tous les produits PCR sont séparés sur gel d'acrylamide et visualisés par une coloration au bromure d'éthidium. En utilisant des systèmes d'imagerie (Image Master 1D Prime ; Amersham PHARMACIA BIOTECH), le changement d'expression de  
25 l'ADN VIH-1 et de l'ARN est déduit après normalisation par rapport aux contrôles standard internes correspondants (respectivement GAPDH ou  $\beta$ -actine). Brièvement, les taux d'ARN VIH sont normalisés au taux de  $\beta$ -actine en calculant le ratio du volume de la bande d'ARN du VIH sur celle de la  $\beta$ -actine. Ainsi, le taux d'augmentation de l'expression du VIH est déduit par  
30 comparaison des cellules transfectées par RH116, SSA ou Tat avec les cellules HeLa CD4+ transfectées par GFP ou pCR3.

**16. Test p24**

La réplication virale est évaluée en mesurant le taux d'antigène p24 dans les surnageants de culture en utilisant le kit d'essai d'antigène p24 de VIH-1 (Coulter, Miami, USA) selon les instructions du fabricant.

5

**17. Détection par ELISA des auto-anticorps dans les sera de patients VIH+ :**

Pour évaluer la présence d'anticorps dirigés contre RH 116 dans les sera de patients VIH-1, un groupe de 32 sujets infectés par le VIH-1 ont été testés avant et après traitement avec des agents antirétroviraux potentiels. Le groupe était constitué de 8 femmes et de 24 hommes ayant une moyenne d'âge de 37 ans (gamme 25-64). Avant le début de la thérapie antirétrovirale, la charge virale plasmatique moyenne était de 421326 copies/ml et le nombre moyen (+ l'écart type) de cellules CD4+ était de 150 (+ 14) cellules/ $\mu$ l. Après une durée moyenne de traitement de 14 mois avec 2 inhibiteurs de reverse-transcriptase et un inhibiteur de protéase, la charge virale est tombée à un taux moyen de 3272 copies/ml et le nombre de cellules CD4+ a augmenté jusqu'à 313 (+ 23) cellules/ $\mu$ l. Les taux d'anticorps dans le sérum des patients sont comparés avec ceux présents dans le sérum de 40 contrôles sains de sexes et d'âges correspondants. La présence d'autoanticorps contre RH116 dans les séra a été recherchée par ELISA. Brièvement, les plaques sont couvertes avec 1  $\mu$ g/ml de RH116<sub>1-335</sub> dans du tampon de coating (30 mM Na<sub>2</sub>CO<sub>3</sub>, 70 mM NaHCO<sub>3</sub>) pendant 3 heures à 4°C. Après 3 lavages dans du PBS contenant 0,05% de Tween-20 (PBS-Tween), les plaques sont incubées pendant 2 heures avec des sera dilués de 1/300 jusqu'à 1/2700. Les plaques sont ensuite lavées et incubées sur la nuit à 4°C avec un anticorps IgG humain (SIGMA) conjugué à la peroxydase. Après 3 lavages avec du PBS-Tween et un lavage avec du PBS, les plaques sont développées pendant 30 minutes en utilisant 0,5 mg/ml d'un substrat au O-phénylène-diamine-dihydrochloride (SIGMA) et 0,1% H<sub>2</sub>O<sub>2</sub> et la réaction est stoppée par l'addition de 50  $\mu$ l/puits de HCl 1N. Les valeurs d'absorbance sont lues à 492 nm en utilisant un lecteur de

10

15

20

25

30

microplaques automatiques (TITERTEK MULTISKAN, LABYSYSTEMS, Finlande). Des protocoles similaires sont utilisés pour évaluer la présence des anticorps antiRH116 dans les sera de souris immunisées, à l'exception que le conjugué utilisé est un anticorps antisouris à la place d'un anticorps antihumain IgG.

**EXEMPLE 1 : Identification du gène RH 116 modulé par le murabutide par DD-RT-PCR.**

Les échantillons d'ARN de PBMC non stimulé et de PBMC déplété en CD8 stimulé par le murabutide et obtenu à partir d'un patient VIH-1 sont analysés par DD-RT-PCR. Les inventeurs ont sélectionné plus de 130 fragments d'ADNc différentiellement exprimés après traitement des PBMC de patient VIH+ au murabutide. Ces fragments ont été sous-clonés dans le vecteur Pcr2.1 (Invitrogen), puis séquencés par séquençage automatique (ABI Prism 377, Perkin-Elmer). Les séquences ont été analysées pour les recherches d'homologie en utilisant les banques de données et le serveur Basic Local Alignment Search Tool (Blast 2) du NCBI.

L'expression d'un certain nombre de gènes apparaît être régulée de manière positive ou négative après un traitement avec le murabutide pendant 6 ou 24 heures. Parmi ceux-ci, 3 gènes correspondent à des répétitions alu, 20 séquences correspondent avec des EST ou des séquences des banques de données génomiques, et 49 séquences sont des gènes bien caractérisés. Parmi les gènes régulés par le murabutide, 28 présentent une expression accrue et 21 ont leur expression inhibée. Le profil d'expression de 14 gènes régulés de manière positive et de 7 gènes régulés de manière négative, par le murabutide a été confirmé par RT-PCR ou par reverse Northern-Blot sous forme de « dot-blot » (données non montrées) sur des extraits d'ARN de PBMC déplétés en CD8 obtenus à partir de 3 autres patients. La modulation de l'expression par le murabutide a également été vérifiée parmi les séquences identifiées par DD-RT-PCR qui correspondent à des EST ou les séquences génomiques des banques de données. Ainsi, 6 clones présentent une expression accrue : 4

clones (clones 1, 71, 87 et 99) suite à un traitement de 6 heures, 2 clones (clones 11 et 89) suite à un traitement de 24 heures. D'autre part, 4 séquences identifiées induisent une inhibition de l'expression : 2 clones (clones 7 et 10) suite à une exposition de 6 heures au murabutide et 2  
5 clones (clones 62 et 96) suite à une exposition de 24 heures au murabutide.

Le profil d'expression du nouveau gène régulé par le murabutide et correspondant au clone 10 a été étudié par RT PCR semi-quantitative et l'ADNc pleine longueur correspondant a été cloné et caractérisé.

## 10 **EXEMPLE 2 : Clonage pleine longueur du gène RH116 modulé par le murabutide.**

La séquence d'ADN pleine longueur correspondant au clone 10 a été obtenue stratégie 5' et 3' RACE. En se basant sur les fragments d'ADNc obtenus par DDRT PCR, un premier cycle de RACE 5' et de RACE 3' a  
15 permis aux inventeurs de caractériser l'extrémité polyA+ de ce nouveau gène et d'étendre la séquence d'ADNc dans la partie 5' du gène. L'extrémité 5' complète de l'ADNc a été obtenue suite à 3 étapes de 5' RACE. Le criblage d'une librairie d'ADNc a permis aux inventeurs de confirmer la séquence d'ADNc pleine longueur. La séquence nucléotidique complète a été obtenue  
20 en déterminant les séquences chevauchantes de 2 clones différents. La séquence d'ADNc pleine longueur a été soumise à GeneBank (N° d'accèsion AY 017378).

Plus précisément, à partir d'un fragment long de 164pb (SEQ ID N°3) obtenu par du DD-RT-PCR, les inventeurs ont synthétisé deux amorces  
25 spécifiques dont F1: 5' TGA TGA GGG TGG TGA TGA TGA GTA TTG TG 3' (SEQ ID N°4) afin de réaliser une première amplification par le 5' et 3' RACE. Les inventeurs ont ainsi pu obtenir un fragment de 1284pb grâce à l'amplification à partir de F1. Ce fragment a été séquencé en plusieurs étapes grâce aux amorces internes spécifiques F2 (5' -GCA GTG AGT TCA  
30 AAC CCA TGA CAC AGA ATG -3') (SEQ ID N°5) et R2 : (5'- CAG CAT TCT GAA TAG TCA AGA TTG GGA AAT G -3') (SEQ ID N°6) ; la séquence de ce fragment correspond à la séquence SEQ ID N°7. Celle-ci présente un cadre

de lecture ouvert de 380 acides aminés jusqu'à un codon Stop potentiel lequel est suivi après 119pb par une queue poly A+ (figure N°1).

Afin d'obtenir la partie 5' de l'ADNc, les inventeurs ont synthétisé une amorce R3 (SEQ ID N°8) qui correspond à la séquence complémentaire de F1 ; l'amorce R3 a permis après PCR d'obtenir une séquence de 194pb (symbolisé en pointillé .... dans la figure 2A).

A partir de cette séquence les inventeurs ont synthétisé une nouvelle amorce R8 (SEQ ID N°9): 5'-GTA GGG CCT CAT TGT ACT TCC TCA AAT-3') afin de déterminer la séquence en position 5' de l'ADNc : une réaction de PCR a permis d'obtenir un fragment d'environ 1200 pb (symbolisé en tiret -- -- dans la figure 2a) (SEQ ID N°10). Le séquençage complet du fragment a été réalisé en plusieurs étapes grâce à des oligonucléotides internes R8-seq 1 (5 'CTC CAA CAC CAG GTG AAG CTG 3') (SEQ ID N°11) et R8-seq 2 (5' CAG ATG AAG AGA ATG TGG CAG 3') (SEQ ID N°12). Le cadre de lecture reste ouvert et permet d'identifier un cadre de lecture de 853aa.

A partir de ces deux séquences les inventeurs ont synthétisé deux amorces HELI 1 (5'- GGA AGT ACA ATG AGG GCC TAC AAA-3') (SEQ ID N°13) et HELI 2 (5'-tcc tca gTc cta gta tat tgc tcc-3') (SEQ ID N°14), respectivement prises sur le fragment F1 et R3, afin de réaliser des RT-PCR pour l'analyse de l'expression différentielle de l'ARNm correspondant (Voir exemple N°4)

A partir de la séquence SEQ ID N°10, les inventeurs ont synthétisé une nouvelle amorce R16 (5'-CTA AGC AGC TGA CAC TTC CTT CTG CCA AAC TTG TGT CTG -3') (SEQ ID N°15) afin de remonter en 5' de l'ADNc ; une réaction de PCR a permis d'obtenir un fragment de 964pb (symbolisé en pointillé (- - - - -), figure 2b).

L'extrémité 5' de 964 pb obtenue après la réaction de PCR (5'RACE) contient un codon ATG potentiel qui est précédé d'un codon STOP dont la présence sera confirmé ultérieurement. Cette stratégie a donc permis aux inventeurs d'obtenir un cadre de lecture ouvert (ORF open reading frame) de 1025 acides aminés.

La séquence complète de l'ADNc correspondant à RH116 est maintenant de 3372pb et correspond à la SEQ ID N°1 et code pour une protéine de 1025 aa qui correspond à la SEQ ID N°2.

Une stratégie de PCR à l'aide de deux amorces F4: (5'-ggg ccc tgt gga  
5 caa cct cgt cat tgt-3') (SEQ ID N°16) et R14: (5'-CCA GAG TGG CTG TTT  
ACA TTG CCA AGG ATC ACT-3') (SEQ ID N°17) spécifiques de la séquence  
SEQ ID N°1 a été développée afin de confirmer la présence du codon  
d'initiation de la traduction ATG ainsi que la présence du codon STOP (TGA)  
en amont de celui-ci. Pour cela les inventeurs ont réalisé une PCR sur une  
10 matrice 5' RACE à l'aide des deux amorces citées précédemment, et ont  
obtenu un fragment de taille expectée, qui a été cloné et séquencé. La  
séquence de ce fragment a confirmé la présence du codon TGA avant le  
codon ATG initiateur confirmant ainsi que les inventeurs possèdent la copie  
complète de l'ADNc.

15 La séquence ADNc (3 372 paires de bases) contient un codon  
d'initiation en position 155 et un cadre ouvert de lecture (ORF) de 3 075  
paires de bases ainsi qu'une région 3' non transcrite de 141 paires de bases  
avec un signal de polyadénylation consensus AATAAAA situé 24 paires de  
bases en amont d'une queue polyA+ de 21 paires de bases. L'ORF code pour  
20 un polypeptide de 1 025 acides aminés (FIG. 2B) avec une masse  
moléculaire calculée de 116 kdaltons et un point isoélectrique de 5,2.

**EXEMPLE 3 : Etude de comparaison de la séquence pleine longueur de  
RH 116 avec les séquences des bases de données .**

25 Les inventeurs ont comparé la séquence en acides aminés déduite de  
1025 acides aminés avec les séquences présentes dans les banques de  
données.

Les inventeurs ont identifié dans cette séquence des domaines  
conservés appartenant à la superfamille des hélicases à ARN et plus  
30 précisément de la famille des protéines connues pour avoir des fonctions  
d'hélicase à ARN dépendantes de l'ATP. Parmi les classes d'hélicases  
basées sur cette homologie de séquence (Luking *et al.*, 1998) sont présentes

les protéines appelées « DEAD-BOX ». Une caractéristique importante de toutes les hélicases à ARN « DEAD-BOX » est la présence de motifs caractéristiques espacés les uns des autres par une distance conservée (Luking *et al.*, 1998 ; Linder et Daugeron, 2000). Malgré des différences mineures, ces motifs sont présents dans la nouvelle séquence clonée (FIG. 3A et 3B). De ce fait, les inventeurs considèrent que cette protéine, appelée RH 116 (pour RNA hélicase 116 kdaltons), constitue un membre de la sous-famille des protéines « DEAD-BOX » des ARN-hélicases dépendantes de l'ATP. Les inventeurs également ont noté la présence d'un motif KKKK à la position 349 amino-acides mais aucun motif bi-partite de signal de localisation nucléaire A n'a pu clairement être identifié. 8 sites de N-glycosylation ont été détectés et des sites potentiels de phosphorylation ont été trouvés pour la kinase cAMP dépendante ainsi que la protéine kinase C ; la signification biologique de ces sites n'a pas été étudiée.

Au cours des recherches effectuées dans les banques de données, les inventeurs ont noté que la protéine RH 116 présentait, en dehors de son homologie de séquence générale avec les membres de la famille des protéines « DEAD-BOX » un fort % d'homologie avec une hélicase à ARN humaine hypothétique et jusqu'à présent non caractérisée, appelée « RIG-1 » (FIG. 3b C) décrite par Sun, Y.W (Genbank, numéro d'accension : AF038963). L'alignement des séquences est présenté en figure 3b C. L'identité totale de séquence en amino-acides entre les deux protéines est de 31% et la similarité qui inclut des échanges conservatifs est de 44%.

L'interrogation de différentes banques de données a permis aux inventeurs de retrouver des parties de la séquence nucléotidique dans d'autres clones génomiques différents. Une partie de la séquence est retrouvée dans le clone NH0576116 enregistré sous le numéro d'accension gb/AC007750, la seconde partie est retrouvée dans le clone génomique RP11-214A4 enregistré sous le n°AC108176. Il convient également de citer les EST (expressed séquences tags) enregistrées sous les numéros d'accès AW 589567, AW 152541 et AW 189584 dans la banque de données EMBL qui présentent respectivement 100%, 99%, 97% d'homologie avec

respectivement les fragments de séquence 2879-3350, 2870-3350, 2818-3350 de la séquence SEQ ID N°1.

**EXEMPLE 4 : Etude par Northern blotting de l'expression de l'ARNm de**

**5 RH 116 dans différents tissus.**

Des ARN de polyA<sup>+</sup> de différents tissus ont été testés par Northern-Blot en utilisant des oligonucléotides marqués à leur extrémité et dérivés de la séquence en ADNc de RH 116 ; la séquence nucléotidique 5'- GCA TCT GCA ATG GCA AAC TTC TTG CAT GGC-3' (SEQ ID N°18) et la séquence  
10 nucléotidique 5'- CGT GCT GAT TCC TCA GTC CTA GTA TAT TGC-3' (SEQ ID N° 26) spécifique de la séquence codante a été synthétisée et marquée au <sup>32</sup>P (T4 Polynucléotide Kinase, Amersham) afin de servir de sonde pour réaliser un Northern blotting. L'hybridation d'une membrane contenant 2µg d'ARN poly A<sup>+</sup> (Clontech) a révélé le résultat présenté à la figure 4.

15 Un signal spécifique correspondant à un ARN messenger d'une taille estimée à environ 3,5 kb a été détecté indiquant que la séquence d'ADNc isolée était une séquence pleine longueur (figure N°4). Le taux d'expression de l'ARN messenger dans 12 tissus humains (cerveau, cœur, muscle squelettique, côlon, thymus, rate, rein, foie, intestin grêle, placenta,  
20 poumon, leucocytes du sang périphérique) a été étudié. L'intensité du signal radioactif est légèrement réduite dans le cerveau, le côlon, le thymus, l'intestin grêle et dans les leucocytes du sang périphérique et un signal radioactif très intense est présent dans le cœur et les reins.

25 **EXEMPLE 5 : Etude par RT PCR de l'expression différentielle de RH 116 dans les PBMCs de patients VIH<sup>+</sup> et de contrôles sains après traitement ou sans traitement par le murabutide.**

L'expression de l'ARN messenger de RH 116 est inhibée par le murabutide. Les inventeurs ont en outre étudié si l'ARN messenger RH 116  
30 était un nouveau gène modulé par le murabutide en utilisant la RT PCR semi-quantitative utilisant des oligonucléotides spécifiques dérivés de l'ADNc de RH 116.

Des PBMC de patients (P) infectés par le VIH ou de donneurs sains contrôles (C) sont isolées, déplétées en lymphocytes CD8<sup>+</sup> (Dynabeads, Dynal) et stimulées par la phytohémagglutinine PHA (5µg/ml) pendant 3 jours. Puis, les cellules sont traitées ou non par le Murabutide (10µg/ml) en présence d'interleukine 2 (IL2) (10U/ml) dans un milieu RPMI supplémenté par 10 % de sérum de veau fœtal (SVF) pendant 6 heures ou 24 heures à raison de 5.10<sup>6</sup> cellules minimum par condition. Après traitement, l'ARN des cellules est extrait (RNAplus, Quantum-bioprobe) puis traité à la DNase (Boehringer) et retrotranscrit (RT) à l'aide d'un oligo(dT) en présence de la reverse transcriptase Mu-MLV (Superscript II, Gibco). La qualité des RT est vérifiée par PCR (25 cycles) à l'aide d'amorces spécifiques de la GAPDH (5'GCC ATC AAT GAC CCC TTC ATT GAC 3') (SEQ ID N°19) et (5' TGA CGA ACA TGG GGG CAT CAG CAG 3') (SEQ ID N°20) à partir de 20, 100 et 500ng d'ARN total. Puis, les inventeurs ont réalisé les RT-PCR (35cycles) à l'aide d'amorces spécifiques Héli 1 et Héli 2 du nouveau polypeptide RH116 (5' GGA AGT ACA ATG AGG GCC TAC AAA 3') (SEQ ID N°13) et 5' TCC TCA GTC CTA GTA TAT TGC TCC 3' ) (SEQ ID N°14). Le nombre de cycles d'amplification a été mis au point au préalable (35 cycles). Les fragments amplifiés sont visualisés sur gel d'agarose (1%) en présence de bromure d'éthidium, puis quantifiés grâce au programme Imager master (Pharmacia). Pour chaque dilution, la valeur donnée pour le gène étudié est rapportée à celle de la GAPDH (Rapport=R). Pour chaque patient et chaque temps (6h ou 24h), le R des cellules traitées au Murabutide est rapporté à celui des cellules non traitées. Les résultats sont alors exprimés en % d'augmentation ou d'inhibition de l'expression du gène par rapport au cellules non traitées. Il est à noter que, pour chaque dilution testée, le R peut varier faiblement, la moyenne des R a été effectuée en prenant soin d'être toujours dans la phase linéaire d'amplification.

Cette étude a été menée sur 12 patients et 10 contrôles sains. L'étude sur les patients montrent une inhibition significative de l'expression du gène du RH116 après un traitement de 6 ou 24 h au Murabutide par rapport au cellules non traitées. D'autre part, l'étude réalisée sur des

PBMCs de donneurs sains révèlent une augmentation très significative de l'expression du gène de RH116 notamment après 6 heures de traitement au Murabutide.

La figure 5 représente les résultats de RT-PCR obtenus sur les PBMCs de patients après 6 et 24 heures de traitement.

Une RT PCR représentative réalisée sur des PBMC déplétés en CD8 obtenue à partir de patients infectés par le VIH-1 est présentée FIG. 5A et montre une inhibition significative de l'expression de l'ARN messenger suite à un traitement de 6 ou 24 heures avec le murabutide.

Les résultats présentés dans la figure 5B présentent un pourcentage moyen d'inhibition de l'expression de l'ARN messenger de RH 116 et démontrent que le traitement avec le murabutide des PBMC déplétés en CD8 de 5 patients induit une inhibition importante de l'expression de l'ARN messenger de RH 116 avec une inhibition moyenne maximale (90%) observée dans le patient N° 5 suite à 6 heures de traitement par le murabutide, et cet effet peut être maintenu après 24 heures, la moyenne maximum d'inhibition (57%) étant observée chez le patient N°3.

#### **EXEMPLE 6 : Expression de la protéine recombinante en système bactérien E. coli (pQE)**

Une PCR a été réalisée sur de l'ADNc de rate à partir d'amorces nucléotidiques correspondant à l'ATG: Heli ATG: (5'- TGA GAG GAT CCG ATG TCG AAT GGG TAT TCC 3') (SEQ ID N°21) et au STOP (5'-AAT GTC GAC CTA ATC CTC ATC ACT AAA TAA -3') (SEQ ID N°22).

Le fragment a été cloné dans le vecteur pQE80 (Qiagen) digéré par les enzymes de restriction BamHI/Sal I .

Après transformation de bactéries TOP 10F', l'expression de la protéine recombinante est induite par l'IPTG, le temps d'expression a été optimisé et estimé à deux heures, en effet après 5 heures d'induction les inventeurs ne détectent pas de protéine recombinante. La protéine est faiblement

exprimée (environ 500µg pour 500 ml de culture) et se présente sous forme soluble.

**EXEMPLE 7 : Expression de la protéine recombinante en système eucaryote.**

La protéine "RH116" étant exprimée dans le cytoplasme ou dans le noyau de cellules de mammifères les inventeurs ont développé une stratégie de surexpression de la protéine recombinante dans des cellules eucaryotes afin d'apprécier son rôle dans la régulation du VIH.

Pour cela, la copie complète de l'ADNc a été amplifiée à partir des amorces heli-GFP-ATG (Xho I) (5'-TGA GAG CTC GAG ATG TCG AAT GGG TAT TCC ACA GAC -3') (SEQ ID N°23) et Heli-GFP (Bam HI) (5'- TGT TTA TTT AGT GAT GAG GAT CGG GAT CCG ATT GAA-3') (SEQ ID N°24) ; le fragment amplifié est cloné dans le vecteur pEGFP digéré par Xho I/Bam HI puis séquencé. Le cDNA codant pour la "Green Fluorescent Protein" (GFP) est situé en 3' de l'insert cloné.

D'autre part, la copie complète de l'ADNc a été amplifiée à partir des amorces Heli ATG Bam (5'- TGA GAG GATCCG ATG TCG AAT GGG TAT TCC -3') (SEQ ID N°21) et Heli STOP XhoI (5'-ttc aat ctc gag atc ctc atc act aaa taa aga -3')(SEQ ID N°25) ; le fragment amplifié est cloné dans le vecteur pcDNA6 digéré par Bam HI /Xho I. Dans ce système la protéine est fusionnée à 6 Histidines et à une protéine V5 contre laquelle des anticorps monoclonaux sont disponibles.

**EXEMPLE 8 : Transfection en cellules eucaryotes.**

Des expériences de transfection ont été réalisées dans les cellules cos-7 et cellules Hela à l'aide des plasmides recombinants pEGFP et pcDNA6 contenant la séquence complète du polypeptide RH116 tels que décrits dans l'exemple 5.

Les transfections ont été réalisées (Effectene-Qiagen) avec 1µg de plasmide recombinant et 10µl d'Effectene. Une analyse en RT-PCR a permis

de confirmer la surexpression de l'ARNm codant pour le polypeptide RH116. L'expression des protéines recombinantes fusionnées a été vérifiée par cytofluorométrie (fusion avec la GFP) ou en western blotting (fusion avec V5-HIS6).

5

**EXEMPLE 9 : Caractérisation moléculaire de l'ADNc codant pour la protéine RH 116.**

La caractérisation moléculaire de la protéine RH 116 native a été réalisée en utilisant des immunoglobulines purifiées à partir de sérum de  
10 souris dirigé contre la protéine RH 116<sub>1-335</sub> recombinante partielle. Le marquage cellulaire et l'immuno-précipitation génèrent un produit protéique de 130 à 140 kdaltons en accord avec la masse moléculaire calculée de la protéine codée qui doit être hautement glycosylée. La protéine RH 116 est immuno-précipitée dans les lysats cellulaires totaux obtenus à  
15 partir de cellules U937 et dans des fractions enrichies en extraits nucléaires (FIG. 6A). Des résultats similaires sont obtenus sur des cellules HeLa (données non montrées). La figure 6B présente le contrôle de qualité de fractionnements sous-cellulaires et montre des extraits nucléaires enrichis qui présentent une expression accrue d'histone 1 (34 kdaltons) que dans  
20 l'extrait total et une expression moindre de protéines cytoplasmiques (60 kdaltons et 40 kdaltons environ).

**EXEMPLE 10 : Etude d'immunolocalisation de la protéine RH 116.**

Pour examiner la location cellulaire de la protéine RH 116 native, les  
25 inventeurs ont réalisé des expériences d'immunocytochimie sur des cultures en monocouches de cellules HeLa-CD4<sup>+</sup> (FIG. 7). Un anticorps souris dirigé contre la protéine RH 116<sub>1-335</sub> colore intensément le cytoplasme des cellules HeLa-CD4<sup>+</sup> (FIG. 7C). Un sérum de souris normale est utilisé comme contrôle négatif (FIG. 7A) et une coloration nucléaire est  
30 observée en utilisant un anticorps monoclonal anti-histone 1 (FIG. 7B). La localisation de la protéine RH 116 est également confirmée par des études d'immunolocalisation de cellules U937 (données non publiées). D'autre

part, la localisation intracellulaire de la protéine RH 116 dans les cellules HeLa-CD4<sup>+</sup> transfectées par une construction d'ADNc chimérique codant pour la protéine RH 116 pleine longueur fusionnée à une extrémité carboxyterminale de la GFP (GFP-RH 116) est différente. De fait, le marquage de la protéine GFP-RH 116 exprimée dans les cellules HeLa-CD4<sup>+</sup> transfectées montre une localisation cytoplasmique et nucléaire (données non montrées). Cette dernière observation est plus consistante que les analyses d'immunoprécipitation et l'activité putative de membres de la sous-famille des protéines « DEAD-BOX » des ARN hélicases.

**EXEMPLE 11 : RH 116 augmente l'expression de la protéine virale P24.**

Dans un but d'élucider le rôle putatif de la protéine RH 116, dont l'expression est inhibée par l'activité « inhibiteur du VIH » du murabutide, les inventeurs ont cherché à déterminer si RH 116 pouvait jouer un rôle dans la régulation de l'expression du virus. Ainsi, les inventeurs ont entamé une étude de l'expression de P24 par les cellules HeLa-CD4<sup>+</sup> transfectées par la protéine GFP-RH 116 et infectées 24 heures plus tard par le virus VIH-1<sub>LAI</sub>.

L'examen microscopique d'une culture à différents moments et l'analyse par la coloration au bleu trypan a révélé que l'infection de cellules transfectées par la GFP-RH 116 affecte la viabilité des cellules comparées à celle des cellules transfectées par la GFP ou des cellules GFP-SSA (une protéine sans rapport). Ainsi, les inventeurs ont normalisé les taux d'antigène P24 pour  $5 \times 10^5$  cellules. L'étude de l'expression de la protéine virale P24 dans les surnageants de cellules HeLa-CD4<sup>+</sup> exprimant la GFP, ou la GFP-RH 116, ou la GFP-SSA, a été réalisée au cours de 5 expériences indépendantes. L'étude de l'expression de la protéine virale P24 dans les surnageants de cellules HeLa-CD4<sup>+</sup> transformées par pCR3 et par pCR3-Tat a été réalisée dans 3 expériences indépendantes. Les figures 8A et 8B présentent des résultats obtenus. La figure 8A montre que l'expression de l'antigène P24 viral (en nanogrammes/ml  $\pm$  la déviation standard) dans les surnageants de cellules HeLa-CD4<sup>+</sup> transfectés par la GFP-RH 116 est

dramatiquement inhibée au jour 3 ( $12,1 \pm 1,48$  ng/ml) jusqu'au jour 5 ( $9\,770 \pm 648$  ng/ml) comparés aux cellules exprimant la GFP ( $2,65 \pm 0,5$  ng/ml et  $690 \pm 272$  ng/ml respectivement aux jours 3 et 5).

Le taux d'antigène P24 dans le surnageant des cellules transfectées par pCR3-Tat est présenté dans la figure 8B (en nanogrammes/ml + la déviation standard). Comme attendu, les résultats montrent une augmentation importante de la libération de P24 par les cellules transfectées par Tat à partir du jour 2 ( $2,2 \pm 0,18$  ng/ml) jusqu'au jour 5 ( $2,994 \pm 391$  ng/ml) comparé avec les cellules transfectées par pCR3 au jour 2 ( $1 \pm 0,08$  ng/ml) et au jour 5 ( $74 \pm 22$  ng/ml). La figure 8C présente le nombre de fois d'augmentation obtenu dans 3 ou 5 expériences indépendantes. Ces résultats indiquent clairement qu'une surexpression de la GFP-RH 116 induit une augmentation importante de P24 à partir du jour 2 ( $3,43 \pm 0,17$ ) jusqu'au jour 5 ( $18,2 \pm 10$ ) comparé à la surexpression de la GFP-SSA (jour 2 :  $1,18 \pm 0,25$  ; jour 5 :  $1,35 \pm 0,37$ ) et une augmentation équivalente est obtenue avec une surexpression de la TAT (jour 2 :  $3,0 \pm 8,42$  ; jour 5 :  $18,7 \pm 4,2$ ).

**EXEMPLE 12 : La protéine RH 116 régule l'expression de l'ARNm P24 viral.**

Pour déterminer le mécanisme de l'augmentation de l'expression de P24 en réponse à une surexpression de RH 116, les inventeurs ont isolé l'ARN total de cellules HeLa-CD4+ transfectées et infectées puis ont déterminé le taux d'ARN messager non épissé, simplement épissé ou de taille intermédiaire, par RT PCR semi-quantitative. Le nombre de fois d'augmentation d'expression d'ARN messager du VIH est obtenu dans 3 expériences séparées (moyenne + l'écart type) (données non présentées). Tout d'abord, les cellules transfectées par la GFP-RH 116 présentent un taux plus élevé d'ARN non épissé que les cellules transfectées par la GFP-SSA comparé aux cellules transfectées par la GFP seule (respectivement  $3,14 \pm 1,5$  et  $0,81 \pm 0,12$ ). Les ARNm de VIH-1 de taille intermédiaire ou simplement épissés sont également augmentés dans les cellules

transfectées par la GFP-RH 116 ( $2,81 \pm 0,73$ ) comparé aux cellules transfectées par la GFP-SSA ( $0,91 \pm 0,14$ ).

En ce qui concerne la surexpression de Tat, des résultats sont obtenus à partir de deux expériences indépendantes ; le taux d'ARN messager non  
5 épissé et de taille intermédiaire ou simplement épissé est respectivement de  $4,75 \pm 1,25$  et de  $4,18 \pm 0,31$ . Une analyse de RT PCR obtenue 3 jours après l'infection est présentée dans la figure 9. La figure 9A indique que la surexpression des ARNm de RH 116, SSA ou TAT est dans les cellules transfectées de manière transitoire. La figure 9B illustre l'augmentation des  
10 ARNm non épissés, ou de taille intermédiaire, ou simplement épissés des protéines surexprimées RH 116 (à droite) et TAT (à gauche) . L'étude de la formation de l'ADN proviral (figure 8C) révèle qu'il n'y a pas de différence significative dans aucune des cellules transfectées. Ces résultats suggèrent que la régulation de l'expression du VIH par RH 116 est réalisée au niveau  
15 transcriptionnel ou post-transcriptionnel.

**EXEMPLE 13 : La protéine RH 116 constitue un auto-antigène chez les patients infectés par le VIH-1.**

La question de savoir si RH 116 peut être un auto-antigène a été  
20 étudiée par analyse ELISA en utilisant des sera obtenus à partir de donneurs sains et de ceux obtenus à partir de patients VIH-1 testés avant et après un traitement anti-rétroviral. Les résultats présentés à la figure 10 démontrent de manière significative une augmentation des taux d'auto-anticorps dans les sera de patients comparés avec ceux de contrôle sain. De  
25 plus, les taux élevés d'auto-anticorps RH 116 observés dans les patients non traités VIH-1 sont toujours significatifs même après un traitement significatif avec des agents anti-rétroviraux efficaces.



**REFERENCES**

- Amiel et al. (1999) J. Infect Dis. 179:83-91.  
Billaut-Mulot et al. (2001) Vaccine In press.
- 5 Blundell (1996) Nature 384:23.  
Buckholz (1993) Curr. Op. Biotechnology 4:538.  
Carter (1993) Curr. Op. Biotechnology 3:533.  
Duck et al. (1990) Biotechniques 9:142.  
Edwards et Aruffo (1993) Curr. Op. Biotechnology 4:558.
- 10 Epstein (1992) Médecine/Sciences 8:902.  
Fruehleis et al. (1987) Int. Ed. Engl. 26:403.  
Godbout et al. (1998) J. Biol. Chem. 273:21161.  
Gorbalenya et al. (1988) Nature 333:22.  
Guatelli et al. (1990) Proc. Natl. Acad. Sci. USA 87:1874.
- 15 Jaramillo et al. (1991) Mol. Cell. Biol. 11:5992.  
Kievitis et al. (1991) J. Virol. Methods 35:273.  
Köhler et Milstein. (1975) Nature 256:495.  
Kwoh, et al. (1989) Proc. Natl. Acad. Sci. USA 86:1173.  
Lain et al. (1990) Nucleic Acid. Res. 18 :7003-7006.
- 20 Landegren et al. (1988) Science 241:1077.  
Lee et Hurwitz (1993) Drosophila J. Biol. Chem. 268:16822-16830.  
Liang et Pardee (1992) Science 257:967-971.  
Linder et al. (1989) Nature 337:121-122.  
Linder et Daugeron (2000) Nat. Struct. Biol. 7:97-99.
- 25 Li-Ru et al. (1999) J. Virol. 73:2841-2853.  
Luckow (1993) Curr. Op. Biotechnology 4:564.  
Luking. et al. (1998) Crit. Rev. Biochem Mol. Biol. 33:259-296.  
Matthews et al. (1988) Anal. Biochem. 169:1-25.  
Miele et al. (1983) J. Mol. Biol. 171: 281.
- 30 Neddleman et Wunsch (1970) J. Mol. Biol. 48:443.  
Olins et Lee (1993) Curr. Op. Biotechnology 4:520.

- Pause and Sonenbery (1992) EMBO J. 11:2643-2654.
- Pearson et Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444.
- Perricaudet et al. (1992) La Recherche 23:471.
- Perun et al. (1989) Eds Computer-aided drug design (Marcel Dekker  
5 Inc. New York.
- Rohlmann et al. (1996) Nature Biotech. 14:1562.
- Rolfs, A. et al. (1991) Berlin : Springer-Verlag.
- Rozen et al. (1990) Mol. Cell. Biol. 10 :1134.
- Segev (1992) Kessler C. Springer Verlag, Berlin, New-York, 197-205.
- 10 Smith et Waterman (1981) Ad. App. Math. 2:482.
- Steinbuch et Audran (1969) Rev. Fr. Etud. Clin. Biol. 14:1054-1058.
- Stewart et Yound (1984) Solid phase peptides synthesis, Pierce  
Chem. Company, Rockford.
- Takeda et al. (1999) J. Immunol. 163:6269-6274
- 15 Temin, (1986) Retrovirus vectors for gene transfer. In Kucherlapati  
R., Ed. Gene Transfer.
- Van de Waterbeemd, Advanced Computer-assisted techniques in  
drug discovery (Verlagsgesellschaft, Weinsheim, 1994)
- Walker et al. (1982), EMBO J. 1:945-951.

## LISTE DE SEQUENCES

<110> ISTAC & Institut Pasteur de Lille

<120> Polypeptide RH 116 et ses fragments et polynucléotides  
codant lesdits polypeptides et applications  
thérapeutiques

<130> D18868

<150> FR 00 06030

<151> 2000-05-11

<160> 56

<170> PatentIn Ver. 2.1

<210> 1

<211> 3372

<212> ADN

<213> Homo sapiens

<220>

<221> CDS

<222> (155)..(3229)

<400> 1

```

gggccctgtg gacaacctcg tcattgtcag gcacagagcg gtagaccctg cttctctaag 60
tgggcagcgg acagcggcac gcacatttca cctgtcccgc agacaacagc accatctgct 120
tgggagaacc ctctcccttc tctgagaaag aaag atg tcg aat ggg tat tcc aca 175
                               Met Ser Asn Gly Tyr Ser Thr
                               1                               5

gac gag aat ttc cgc tat ctc atc tcg tgc ttc agg gcc agg gtg aaa 223
Asp Glu Asn Phe Arg Tyr Leu Ile Ser Cys Phe Arg Ala Arg Val Lys
      10                               15                               20

atg tac atc cag gtg gag cct gtg ctg gac tac ctg acc ttt ctg cct 271
Met Tyr Ile Gln Val Glu Pro Val Leu Asp Tyr Leu Thr Phe Leu Pro
      25                               30                               35

gca gag gtg aag gag cag att cag agg aca gtc gcc acc tcc ggg aac 319
Ala Glu Val Lys Glu Gln Ile Gln Arg Thr Val Ala Thr Ser Gly Asn
      40                               45                               50                               55

atg cag gca gtt gaa ctg ctg ctg agc acc ttg gag aag gga gtc tgg 367
Met Gln Ala Val Glu Leu Leu Leu Ser Thr Leu Glu Lys Gly Val Trp
      60                               65                               70

cac ctt ggt tgg act cgg gaa ttc gtg gag gcc ctc cgg aga acc ggc 415
His Leu Gly Trp Thr Arg Glu Phe Val Glu Ala Leu Arg Arg Thr Gly
      75                               80                               85

agc cct ctg gcc gcc cgc tac atg aac cct gag ctc acg gac ttg ccc 463
Ser Pro Leu Ala Ala Arg Tyr Met Asn Pro Glu Leu Thr Asp Leu Pro
      90                               95                               100

tct cca tcg ttt gag aac gct cat gat gaa tat ctc caa ctg ctg aac 511
Ser Pro Ser Phe Glu Asn Ala His Asp Glu Tyr Leu Gln Leu Leu Asn

```

WO 01/85955

PCT/FR01/01441

2

| 105                             | 110                         | 115                     |      |
|---------------------------------|-----------------------------|-------------------------|------|
| ctc ctt cag ccc act             | ctg gtg gac aag ctt         | cta gtt aga gac gtc ttg | 559  |
| Leu Leu Gln Pro Thr             | Leu Val Asp Lys Leu         | Leu Val Arg Asp Val Leu |      |
| 120                             | 125                         | 130                     | 135  |
| gat aag tgc atg gag             | gag gaa ctg ttg aca att     | gaa gac aga aac cgg     | 607  |
| Asp Lys Cys Met                 | Glu Glu Glu Leu Leu Thr     | Ile Glu Asp Arg Asn Arg |      |
| 140                             | 145                         | 150                     |      |
| att gct gct gca gaa aac aat     | gga aat gaa tca ggt gta aga | gag cta                 | 655  |
| Ile Ala Ala Ala Glu Asn Asn     | Gly Asn Glu Ser Gly Val Arg | Glu Leu                 |      |
| 155                             | 160                         | 165                     |      |
| cta aaa agg att gtg cag aaa     | gaa aac tgg ttc tct gca ttt | ctg aat                 | 703  |
| Leu Lys Arg Ile Val Gln Lys     | Glu Asn Trp Phe Ser Ala Phe | Leu Asn                 |      |
| 170                             | 175                         | 180                     |      |
| gtt ctt cgt caa aca gga aac aat | gaa ctt gtc caa gag tta aca | ggc                     | 751  |
| Val Leu Arg Gln Thr Gly Asn Asn | Glu Leu Val Gln Glu Leu Thr | Gly                     |      |
| 185                             | 190                         | 195                     |      |
| tct gat tgc tca gaa agc aat     | gca gag att gag aat tta tca | caa gtt                 | 799  |
| Ser Asp Cys Ser Glu Ser Asn Ala | Glu Ile Glu Asn Leu Ser Gln | Val                     |      |
| 200                             | 205                         | 210                     | 215  |
| gat ggt cct caa gtg gaa gag     | caa ctt ctt tca acc aca gtt | cag cca                 | 847  |
| Asp Gly Pro Gln Val Glu Glu Gln | Leu Leu Ser Thr Thr Val Gln | Pro                     |      |
| 220                             | 225                         | 230                     |      |
| aat ctg gag aag gag gtc tgg     | ggc atg gag aat aac tca tca | gaa tca                 | 895  |
| Asn Leu Glu Lys Glu Val Trp Gly | Met Glu Asn Asn Ser Ser Glu | Ser                     |      |
| 235                             | 240                         | 245                     |      |
| tct ttt gca gat tct tct gta     | gtt tca gaa tca gac aca agt | ttg gca                 | 943  |
| Ser Phe Ala Asp Ser Ser Val     | Val Ser Glu Ser Asp Thr Ser | Leu Ala                 |      |
| 250                             | 255                         | 260                     |      |
| gaa gga agt gtc agc tgc tta     | gat gaa agt ctt gga cat aac | agc aac                 | 991  |
| Glu Gly Ser Val Ser Cys Leu     | Asp Glu Ser Leu Gly His Asn | Ser Asn                 |      |
| 265                             | 270                         | 275                     |      |
| atg ggc agt gat tca ggc acc     | atg gga agt gat tca gat gaa | gag aat                 | 1039 |
| Met Gly Ser Asp Ser Gly Thr     | Met Gly Ser Asp Ser Asp Glu | Glu Asn                 |      |
| 280                             | 285                         | 290                     | 295  |
| gtg gca gca aga gca tcc ccg     | gag cca gaa ctc cag ctc agg | cct tac                 | 1087 |
| Val Ala Ala Arg Ala Ser Pro     | Glu Pro Glu Leu Gln Leu Arg | Pro Tyr                 |      |
| 300                             | 305                         | 310                     |      |
| caa atg gaa gtt gcc cag cca     | gcc ttg gaa ggg aag aat atc | atc atc                 | 1135 |
| Gln Met Glu Val Ala Gln Pro     | Ala Leu Glu Gly Lys Asn Ile | Ile Ile                 |      |
| 315                             | 320                         | 325                     |      |
| tgc ctc cct aca ggg agt gga     | aaa acc aga gtg gct gtt tac | att gcc                 | 1183 |
| Cys Leu Pro Thr Gly Ser Gly     | Lys Thr Arg Val Ala Val Tyr | Ile Ala                 |      |
| 330                             | 335                         | 340                     |      |
| aag gat cac tta gac aag aag     | aaa aaa gca tct gag cct gga | aaa gtt                 | 1231 |
| Lys Asp His Leu Asp Lys Lys     | Lys Lys Ala Ser Glu Pro Gly | Lys Val                 |      |
| 345                             | 350                         | 355                     |      |

|                                                                 |      |
|-----------------------------------------------------------------|------|
| ata gtt ctt gtc aat aag gta ctg cta gtt gaa cag ctc ttc cgc aag | 1279 |
| Ile Val Leu Val Asn Lys Val Leu Leu Val Glu Gln Leu Phe Arg Lys |      |
| 360 365 370 375                                                 |      |
| gag ttc caa cca ttt ttg aag aaa tgg tat cgt gtt att gga tta agt | 1327 |
| Glu Phe Gln Pro Phe Leu Lys Lys Trp Tyr Arg Val Ile Gly Leu Ser |      |
| 380 385 390                                                     |      |
| ggt gat acc caa ctg aaa ata tca ttt cca gaa gtt gtc aag tcc tgt | 1375 |
| Gly Asp Thr Gln Leu Lys Ile Ser Phe Pro Glu Val Val Lys Ser Cys |      |
| 395 400 405                                                     |      |
| gat att att atc agt aca gct caa atc ctt gaa aac tcc ctc tta aac | 1423 |
| Asp Ile Ile Ile Ser Thr Ala Gln Ile Leu Glu Asn Ser Leu Leu Asn |      |
| 410 415 420                                                     |      |
| ttg gaa aat gga gaa gat gct ggt gtt caa ttg tca gac ttt tcc ttc | 1471 |
| Leu Glu Asn Gly Glu Asp Ala Gly Val Gln Leu Ser Asp Phe Ser Phe |      |
| 425 430 435                                                     |      |
| att atc att gat gaa tgt cat cac acc aac aaa gaa gca gtg tat aat | 1519 |
| Ile Ile Ile Asp Glu Cys His His Thr Asn Lys Glu Ala Val Tyr Asn |      |
| 440 445 450 455                                                 |      |
| aac atc atg agg cat tat ttg atg cag aag ttg aaa aac aat aga ctc | 1567 |
| Asn Ile Met Arg His Tyr Leu Met Gln Lys Leu Lys Asn Asn Arg Leu |      |
| 460 465 470                                                     |      |
| aag aaa gaa aac aaa cca gtg att ccc ctt cct cag ata ctg gga cta | 1615 |
| Lys Lys Glu Asn Lys Pro Val Ile Pro Leu Pro Gln Ile Leu Gly Leu |      |
| 475 480 485                                                     |      |
| aca gct tca cct ggt gtt gga ggg gcc acg aag caa gcc aaa gct gaa | 1663 |
| Thr Ala Ser Pro Gly Val Gly Gly Ala Thr Lys Gln Ala Lys Ala Glu |      |
| 490 495 500                                                     |      |
| gaa cac att tta aaa cta tgt gcc aat ctt gat gca ttt act att aaa | 1711 |
| Glu His Ile Leu Lys Leu Cys Ala Asn Leu Asp Ala Phe Thr Ile Lys |      |
| 505 510 515                                                     |      |
| act gtt aaa gaa aac ctt gat caa ctg aaa aac caa ata cag gag cca | 1759 |
| Thr Val Lys Glu Asn Leu Asp Gln Leu Lys Asn Gln Ile Gln Glu Pro |      |
| 520 525 530 535                                                 |      |
| tgc aag aag ttt gcc att gca gat gca acc aga gaa gat cca ttt aaa | 1807 |
| Cys Lys Lys Phe Ala Ile Ala Asp Ala Thr Arg Glu Asp Pro Phe Lys |      |
| 540 545 550                                                     |      |
| gag aaa ctt cta gaa ata atg aca agg att caa act tat tgt caa atg | 1855 |
| Glu Lys Leu Leu Glu Ile Met Thr Arg Ile Gln Thr Tyr Cys Gln Met |      |
| 555 560 565                                                     |      |
| agt cca atg tca gat ttt gga act caa ccc tat gaa caa tgg gcc att | 1903 |
| Ser Pro Met Ser Asp Phe Gly Thr Gln Pro Tyr Glu Gln Trp Ala Ile |      |
| 570 575 580                                                     |      |
| caa atg gaa aaa aaa gct gca aaa gaa gga aat cgc aaa gaa agt gtt | 1951 |
| Gln Met Glu Lys Lys Ala Ala Lys Glu Gly Asn Arg Lys Glu Ser Val |      |
| 585 590 595                                                     |      |

WO 01/85955

PCT/FR01/01441

4

|                                                                 |      |
|-----------------------------------------------------------------|------|
| tgt gca gaa cat ttg agg aag tac aat aag gcc cta caa att aat gac | 1999 |
| Cys Ala Glu His Leu Arg Lys Tyr Asn Lys Ala Leu Gln Ile Asn Asp |      |
| 600 605 610 615                                                 |      |
| aca att cga atg ata gat gcg tat act cat ctt gaa act ttc tat aat | 2047 |
| Thr Ile Arg Met Ile Asp Ala Tyr Thr His Leu Glu Thr Phe Tyr Asn |      |
| 620 625 630                                                     |      |
| gaa gag aaa gat aag aag ttt gca gtc ata gaa gat gat agt gat gag | 2095 |
| Glu Glu Lys Asp Lys Lys Phe Ala Val Ile Glu Asp Asp Ser Asp Glu |      |
| 635 640 645                                                     |      |
| ggg ggt gat gat gag tat tgt gat ggt gat gaa gat gag gat gat tta | 2143 |
| Gly Gly Asp Asp Glu Tyr Cys Asp Gly Asp Glu Asp Glu Asp Asp Leu |      |
| 650 655 660                                                     |      |
| aag aaa cct ttg aaa ctg gat gaa aca gat aga ttt ctc atg act tta | 2191 |
| Lys Lys Pro Leu Lys Leu Asp Glu Thr Asp Arg Phe Leu Met Thr Leu |      |
| 665 670 675                                                     |      |
| ttt ttt gaa aac aat aaa atg ttg aaa agg ctg gct gaa aac cca gaa | 2239 |
| Phe Phe Glu Asn Asn Lys Met Leu Lys Arg Leu Ala Glu Asn Pro Glu |      |
| 680 685 690 695                                                 |      |
| tat gaa aat gaa aag ctg acc aaa tta aga aat acc ata atg gag caa | 2287 |
| Tyr Glu Asn Glu Lys Leu Thr Lys Leu Arg Asn Thr Ile Met Glu Gln |      |
| 700 705 710                                                     |      |
| tat act agg act gag gaa tca gca cga gga ata atc ttt aca aaa aca | 2335 |
| Tyr Thr Arg Thr Glu Glu Ser Ala Arg Gly Ile Ile Phe Thr Lys Thr |      |
| 715 720 725                                                     |      |
| cga cag agt gca tat gcg ctt tcc cag tgg att act gaa aat gaa aaa | 2383 |
| Arg Gln Ser Ala Tyr Ala Leu Ser Gln Trp Ile Thr Glu Asn Glu Lys |      |
| 730 735 740                                                     |      |
| ttt gct gaa gta gga gtc aaa gcc cac cat ctg att gga gct gga cac | 2431 |
| Phe Ala Glu Val Gly Val Lys Ala His His Leu Ile Gly Ala Gly His |      |
| 745 750 755                                                     |      |
| agc agt gag ttc aaa ccc atg aca cag aat gaa caa aaa gaa gtc att | 2479 |
| Ser Ser Glu Phe Lys Pro Met Thr Gln Asn Glu Gln Lys Glu Val Ile |      |
| 760 765 770 775                                                 |      |
| agt aaa ttt cgc act gga aaa ata aat ctg ctt atc gct acc aca gtg | 2527 |
| Ser Lys Phe Arg Thr Gly Lys Ile Asn Leu Leu Ile Ala Thr Thr Val |      |
| 780 785 790                                                     |      |
| gca gaa gaa ggt ctg gat att aaa gaa tgt aac att gtt atc cgt tat | 2575 |
| Ala Glu Glu Gly Leu Asp Ile Lys Glu Cys Asn Ile Val Ile Arg Tyr |      |
| 795 800 805                                                     |      |
| ggg ctc gtc acc aat gaa ata gcc atg gtc cag gcc cgt ggt cga gcc | 2623 |
| Gly Leu Val Thr Asn Glu Ile Ala Met Val Gln Ala Arg Gly Arg Ala |      |
| 810 815 820                                                     |      |
| aga gct gat gag agc acc tac gtc ctg gtt gct cac agt ggt tca gga | 2671 |
| Arg Ala Asp Glu Ser Thr Tyr Val Leu Val Ala His Ser Gly Ser Gly |      |
| 825 830 835                                                     |      |
| gtt atc gaa cgt gag aca gtt aat gat ttc cga gag aag atg atg tat | 2719 |

Val Ile Glu Arg Glu Thr Val Asn Asp Phe Arg Glu Lys Met Met Tyr  
 840 845 850 855  
 aaa gct ata cat tgt gtt caa aat atg aaa cca gag gag tat gct cat 2767  
 Lys Ala Ile His Cys Val Gln Asn Met Lys Pro Glu Glu Tyr Ala His  
 860 865 870  
 aag att ttg gaa tta cag atg caa agt ata atg gaa aag aaa atg aaa 2815  
 Lys Ile Leu Glu Leu Gln Met Gln Ser Ile Met Glu Lys Lys Met Lys  
 875 880 885  
 acc aag aga aat att gcc aag cat tac aag aat aac cca tca cta ata 2863  
 Thr Lys Arg Asn Ile Ala Lys His Tyr Lys Asn Asn Pro Ser Leu Ile  
 890 895 900  
 act ttc ctt tgc aaa aac tgc agt gtg cta gcc tgt tct ggg gaa gat 2911  
 Thr Phe Leu Cys Lys Asn Cys Ser Val Leu Ala Cys Ser Gly Glu Asp  
 905 910 915  
 atc cat gta att gag aaa atg cat cac gtc aat atg acc cca gaa ttc 2959  
 Ile His Val Ile Glu Lys Met His His Val Asn Met Thr Pro Glu Phe  
 920 925 930 935  
 aag gaa ctt tac att gta aga gaa aac aaa gca ctg caa aag aag tgt 3007  
 Lys Glu Leu Tyr Ile Val Arg Glu Asn Lys Ala Leu Gln Lys Lys Cys  
 940 945 950  
 gcc gac tat caa ata aat ggt gaa atc atc tgc aaa tgt ggc cag gct 3055  
 Ala Asp Tyr Gln Ile Asn Gly Glu Ile Ile Cys Lys Cys Gly Gln Ala  
 955 960 965  
 tgg gga aca atg atg gtg cac aaa ggc tta gat ttg cct tgt ctc aaa 3103  
 Trp Gly Thr Met Met Val His Lys Gly Leu Asp Leu Pro Cys Leu Lys  
 970 975 980  
 ata agg aat ttt gta gtg gtt ttc aaa aat aat tca aca aag aaa caa 3151  
 Ile Arg Asn Phe Val Val Val Phe Lys Asn Asn Ser Thr Lys Lys Gln  
 985 990 995  
 tac aaa aag tgg gta gaa tta cct atc aca ttt ccc aat ctt gac tat 3199  
 Tyr Lys Lys Trp Val Glu Leu Pro Ile Thr Phe Pro Asn Leu Asp Tyr  
 1000 1005 1010 1015  
 tca gaa tgc tgt tta ttt agt gat gag gat tagcacttga ttgaagattc 3249  
 Ser Glu Cys Cys Leu Phe Ser Asp Glu Asp  
 1020 1025  
 ttttaaaata ctatcagtta aacatttaat atgattatga ttaatgtatt cattatgcta 3309  
 cagaactgac ataagaatca ataaaatgat tgttttactc tccaaaaaaa aaaaaaaaaa 3369  
 aaa 3372  
 <210> 2  
 <211> 1025  
 <212> PRT  
 <213> Homo sapiens  
 <400> 2  
 Met Ser Asn Gly Tyr Ser Thr Asp Glu Asn Phe Arg Tyr Leu Ile Ser

WO 01/85955

PCT/FR01/01441

6

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| 1   |     | 5   |     | 10  |     | 15  |     |     |     |     |     |     |     |     |     |  |
| Cys | Phe | Arg | Ala | Arg | Val | Lys | Met | Tyr | Ile | Gln | Val | Glu | Pro | Val | Leu |  |
|     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |     |     |  |
| Asp | Tyr | Leu | Thr | Phe | Leu | Pro | Ala | Glu | Val | Lys | Glu | Gln | Ile | Gln | Arg |  |
|     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |     |     |  |
| Thr | Val | Ala | Thr | Ser | Gly | Asn | Met | Gln | Ala | Val | Glu | Leu | Leu | Leu | Ser |  |
|     | 50  |     |     |     |     | 55  |     |     |     |     | 60  |     |     |     |     |  |
| Thr | Leu | Glu | Lys | Gly | Val | Trp | His | Leu | Gly | Trp | Thr | Arg | Glu | Phe | Val |  |
| 65  |     |     |     |     | 70  |     |     |     |     | 75  |     |     |     |     | 80  |  |
| Glu | Ala | Leu | Arg | Arg | Thr | Gly | Ser | Pro | Leu | Ala | Ala | Arg | Tyr | Met | Asn |  |
|     |     |     |     | 85  |     |     |     |     | 90  |     |     |     |     | 95  |     |  |
| Pro | Glu | Leu | Thr | Asp | Leu | Pro | Ser | Pro | Ser | Phe | Glu | Asn | Ala | His | Asp |  |
|     |     |     | 100 |     |     |     |     | 105 |     |     |     |     | 110 |     |     |  |
| Glu | Tyr | Leu | Gln | Leu | Leu | Asn | Leu | Leu | Gln | Pro | Thr | Leu | Val | Asp | Lys |  |
|     |     | 115 |     |     |     |     | 120 |     |     |     |     | 125 |     |     |     |  |
| Leu | Leu | Val | Arg | Asp | Val | Leu | Asp | Lys | Cys | Met | Glu | Glu | Glu | Leu | Leu |  |
|     | 130 |     |     |     |     | 135 |     |     |     |     | 140 |     |     |     |     |  |
| Thr | Ile | Glu | Asp | Arg | Asn | Arg | Ile | Ala | Ala | Ala | Glu | Asn | Asn | Gly | Asn |  |
| 145 |     |     |     |     | 150 |     |     |     |     | 155 |     |     |     |     | 160 |  |
| Glu | Ser | Gly | Val | Arg | Glu | Leu | Leu | Lys | Arg | Ile | Val | Gln | Lys | Glu | Asn |  |
|     |     |     |     | 165 |     |     |     |     | 170 |     |     |     |     | 175 |     |  |
| Trp | Phe | Ser | Ala | Phe | Leu | Asn | Val | Leu | Arg | Gln | Thr | Gly | Asn | Asn | Glu |  |
|     |     |     | 180 |     |     |     |     | 185 |     |     |     |     | 190 |     |     |  |
| Leu | Val | Gln | Glu | Leu | Thr | Gly | Ser | Asp | Cys | Ser | Glu | Ser | Asn | Ala | Glu |  |
|     |     | 195 |     |     |     |     | 200 |     |     |     |     | 205 |     |     |     |  |
| Ile | Glu | Asn | Leu | Ser | Gln | Val | Asp | Gly | Pro | Gln | Val | Glu | Glu | Gln | Leu |  |
|     | 210 |     |     |     |     | 215 |     |     |     |     | 220 |     |     |     |     |  |
| Leu | Ser | Thr | Thr | Val | Gln | Pro | Asn | Leu | Glu | Lys | Glu | Val | Trp | Gly | Met |  |
| 225 |     |     |     |     | 230 |     |     |     |     | 235 |     |     |     |     | 240 |  |
| Glu | Asn | Asn | Ser | Ser | Glu | Ser | Ser | Phe | Ala | Asp | Ser | Ser | Val | Val | Ser |  |
|     |     |     |     | 245 |     |     |     |     | 250 |     |     |     |     | 255 |     |  |
| Glu | Ser | Asp | Thr | Ser | Leu | Ala | Glu | Gly | Ser | Val | Ser | Cys | Leu | Asp | Glu |  |
|     |     |     | 260 |     |     |     |     | 265 |     |     |     |     | 270 |     |     |  |
| Ser | Leu | Gly | His | Asn | Ser | Asn | Met | Gly | Ser | Asp | Ser | Gly | Thr | Met | Gly |  |
|     |     | 275 |     |     |     |     | 280 |     |     |     |     | 285 |     |     |     |  |
| Ser | Asp | Ser | Asp | Glu | Glu | Asn | Val | Ala | Ala | Arg | Ala | Ser | Pro | Glu | Pro |  |
|     | 290 |     |     |     |     | 295 |     |     |     |     | 300 |     |     |     |     |  |
| Glu | Leu | Gln | Leu | Arg | Pro | Tyr | Gln | Met | Glu | Val | Ala | Gln | Pro | Ala | Leu |  |
| 305 |     |     |     |     | 310 |     |     |     |     | 315 |     |     |     |     | 320 |  |
| Glu | Gly | Lys | Asn | Ile | Ile | Ile | Cys | Leu | Pro | Thr | Gly | Ser | Gly | Lys | Thr |  |

|                                                                 |     |     |
|-----------------------------------------------------------------|-----|-----|
| 325                                                             | 330 | 335 |
| Arg Val Ala Val Tyr Ile Ala Lys Asp His Leu Asp Lys Lys Lys Lys |     |     |
| 340                                                             | 345 | 350 |
| Ala Ser Glu Pro Gly Lys Val Ile Val Leu Val Asn Lys Val Leu Leu |     |     |
| 355                                                             | 360 | 365 |
| Val Glu Gln Leu Phe Arg Lys Glu Phe Gln Pro Phe Leu Lys Lys Trp |     |     |
| 370                                                             | 375 | 380 |
| Tyr Arg Val Ile Gly Leu Ser Gly Asp Thr Gln Leu Lys Ile Ser Phe |     |     |
| 385                                                             | 390 | 400 |
| Pro Glu Val Val Lys Ser Cys Asp Ile Ile Ile Ser Thr Ala Gln Ile |     |     |
| 405                                                             | 410 | 415 |
| Leu Glu Asn Ser Leu Leu Asn Leu Glu Asn Gly Glu Asp Ala Gly Val |     |     |
| 420                                                             | 425 | 430 |
| Gln Leu Ser Asp Phe Ser Phe Ile Ile Ile Asp Glu Cys His His Thr |     |     |
| 435                                                             | 440 | 445 |
| Asn Lys Glu Ala Val Tyr Asn Asn Ile Met Arg His Tyr Leu Met Gln |     |     |
| 450                                                             | 455 | 460 |
| Lys Leu Lys Asn Asn Arg Leu Lys Lys Glu Asn Lys Pro Val Ile Pro |     |     |
| 465                                                             | 470 | 480 |
| Leu Pro Gln Ile Leu Gly Leu Thr Ala Ser Pro Gly Val Gly Gly Ala |     |     |
| 485                                                             | 490 | 495 |
| Thr Lys Gln Ala Lys Ala Glu Glu His Ile Leu Lys Leu Cys Ala Asn |     |     |
| 500                                                             | 505 | 510 |
| Leu Asp Ala Phe Thr Ile Lys Thr Val Lys Glu Asn Leu Asp Gln Leu |     |     |
| 515                                                             | 520 | 525 |
| Lys Asn Gln Ile Gln Glu Pro Cys Lys Lys Phe Ala Ile Ala Asp Ala |     |     |
| 530                                                             | 535 | 540 |
| Thr Arg Glu Asp Pro Phe Lys Glu Lys Leu Leu Glu Ile Met Thr Arg |     |     |
| 545                                                             | 550 | 555 |
| Ile Gln Thr Tyr Cys Gln Met Ser Pro Met Ser Asp Phe Gly Thr Gln |     |     |
| 565                                                             | 570 | 575 |
| Pro Tyr Glu Gln Trp Ala Ile Gln Met Glu Lys Lys Ala Ala Lys Glu |     |     |
| 580                                                             | 585 | 590 |
| Gly Asn Arg Lys Glu Ser Val Cys Ala Glu His Leu Arg Lys Tyr Asn |     |     |
| 595                                                             | 600 | 605 |
| Lys Ala Leu Gln Ile Asn Asp Thr Ile Arg Met Ile Asp Ala Tyr Thr |     |     |
| 610                                                             | 615 | 620 |
| His Leu Glu Thr Phe Tyr Asn Glu Glu Lys Asp Lys Lys Phe Ala Val |     |     |
| 625                                                             | 630 | 635 |
| Ile Glu Asp Asp Ser Asp Glu Gly Gly Asp Asp Glu Tyr Cys Asp Gly |     |     |
| 645                                                             | 650 | 655 |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|--|
| Asp | Glu | Asp | Glu | Asp | Asp | Leu | Lys | Lys | Pro | Leu | Lys | Leu | Asp | Glu | Thr |  |  |  |
|     |     |     | 660 |     |     |     |     | 665 |     |     |     |     | 670 |     |     |  |  |  |
| Asp | Arg | Phe | Leu | Met | Thr | Leu | Phe | Phe | Glu | Asn | Asn | Lys | Met | Leu | Lys |  |  |  |
|     |     | 675 |     |     |     |     | 680 |     |     |     |     | 685 |     |     |     |  |  |  |
| Arg | Leu | Ala | Glu | Asn | Pro | Glu | Tyr | Glu | Asn | Glu | Lys | Leu | Thr | Lys | Leu |  |  |  |
|     | 690 |     |     |     |     | 695 |     |     |     |     | 700 |     |     |     |     |  |  |  |
| Arg | Asn | Thr | Ile | Met | Glu | Gln | Tyr | Thr | Arg | Thr | Glu | Glu | Ser | Ala | Arg |  |  |  |
| 705 |     |     |     |     | 710 |     |     |     |     | 715 |     |     |     |     | 720 |  |  |  |
| Gly | Ile | Ile | Phe | Thr | Lys | Thr | Arg | Gln | Ser | Ala | Tyr | Ala | Leu | Ser | Gln |  |  |  |
|     |     |     |     | 725 |     |     |     |     | 730 |     |     |     |     | 735 |     |  |  |  |
| Trp | Ile | Thr | Glu | Asn | Glu | Lys | Phe | Ala | Glu | Val | Gly | Val | Lys | Ala | His |  |  |  |
|     |     |     | 740 |     |     |     |     | 745 |     |     |     |     | 750 |     |     |  |  |  |
| His | Leu | Ile | Gly | Ala | Gly | His | Ser | Ser | Glu | Phe | Lys | Pro | Met | Thr | Gln |  |  |  |
|     |     | 755 |     |     |     |     | 760 |     |     |     |     | 765 |     |     |     |  |  |  |
| Asn | Glu | Gln | Lys | Glu | Val | Ile | Ser | Lys | Phe | Arg | Thr | Gly | Lys | Ile | Asn |  |  |  |
|     | 770 |     |     |     |     | 775 |     |     |     |     | 780 |     |     |     |     |  |  |  |
| Leu | Leu | Ile | Ala | Thr | Thr | Val | Ala | Glu | Glu | Gly | Leu | Asp | Ile | Lys | Glu |  |  |  |
| 785 |     |     |     |     | 790 |     |     |     |     | 795 |     |     |     |     | 800 |  |  |  |
| Cys | Asn | Ile | Val | Ile | Arg | Tyr | Gly | Leu | Val | Thr | Asn | Glu | Ile | Ala | Met |  |  |  |
|     |     |     |     | 805 |     |     |     |     | 810 |     |     |     |     | 815 |     |  |  |  |
| Val | Gln | Ala | Arg | Gly | Arg | Ala | Arg | Ala | Asp | Glu | Ser | Thr | Tyr | Val | Leu |  |  |  |
|     |     |     | 820 |     |     |     |     | 825 |     |     |     |     | 830 |     |     |  |  |  |
| Val | Ala | His | Ser | Gly | Ser | Gly | Val | Ile | Glu | Arg | Glu | Thr | Val | Asn | Asp |  |  |  |
|     |     | 835 |     |     |     |     | 840 |     |     |     |     | 845 |     |     |     |  |  |  |
| Phe | Arg | Glu | Lys | Met | Met | Tyr | Lys | Ala | Ile | His | Cys | Val | Gln | Asn | Met |  |  |  |
|     | 850 |     |     |     |     | 855 |     |     |     |     | 860 |     |     |     |     |  |  |  |
| Lys | Pro | Glu | Glu | Tyr | Ala | His | Lys | Ile | Leu | Glu | Leu | Gln | Met | Gln | Ser |  |  |  |
| 865 |     |     |     |     | 870 |     |     |     |     | 875 |     |     |     | 880 |     |  |  |  |
| Ile | Met | Glu | Lys | Lys | Met | Lys | Thr | Lys | Arg | Asn | Ile | Ala | Lys | His | Tyr |  |  |  |
|     |     |     |     | 885 |     |     |     |     | 890 |     |     |     |     | 895 |     |  |  |  |
| Lys | Asn | Asn | Pro | Ser | Leu | Ile | Thr | Phe | Leu | Cys | Lys | Asn | Cys | Ser | Val |  |  |  |
|     |     |     | 900 |     |     |     |     | 905 |     |     |     |     | 910 |     |     |  |  |  |
| Leu | Ala | Cys | Ser | Gly | Glu | Asp | Ile | His | Val | Ile | Glu | Lys | Met | His | His |  |  |  |
|     |     | 915 |     |     |     |     | 920 |     |     |     |     | 925 |     |     |     |  |  |  |
| Val | Asn | Met | Thr | Pro | Glu | Phe | Lys | Glu | Leu | Tyr | Ile | Val | Arg | Glu | Asn |  |  |  |
|     | 930 |     |     |     |     | 935 |     |     |     |     | 940 |     |     |     |     |  |  |  |
| Lys | Ala | Leu | Gln | Lys | Lys | Cys | Ala | Asp | Tyr | Gln | Ile | Asn | Gly | Glu | Ile |  |  |  |
| 945 |     |     |     |     | 950 |     |     |     |     | 955 |     |     |     |     | 960 |  |  |  |
| Ile | Cys | Lys | Cys | Gly | Gln | Ala | Trp | Gly | Thr | Met | Met | Val | His | Lys | Gly |  |  |  |
|     |     |     |     | 965 |     |     |     |     | 970 |     |     |     |     | 975 |     |  |  |  |

Leu Asp Leu Pro Cys Leu Lys Ile Arg Asn Phe Val Val Val Phe Lys  
                   980                                  985                                  990

Asn Asn Ser Thr Lys Lys Gln Tyr Lys Lys Trp Val Glu Leu Pro Ile  
                   995                                  1000                                  1005

Thr Phe Pro Asn Leu Asp Tyr Ser Glu Cys Cys Leu Phe Ser Asp Glu  
           1010                                  1015                                  1020

Asp  
 025

<210> 3  
 <211> 164  
 <212> ADN  
 <213> Homo sapiens

<400> 3  
 aagatgatag tgatgagggt ggtgatgatg agtattgtga tggatgatgaa gatgaggatg 60  
 atttaaagaa accttgaaac tggatgaaac agatagattt ctcatgactt tatttttttga 120  
 aaacaataaa atgttgaaaa ggctggctga aaaccacagaa tatg 164

<210> 4  
 <211> 29  
 <212> ADN  
 <213> Homo sapiens

<400> 4  
 tgatgagggt ggtgatgatg agtattgtg 29

<210> 5  
 <211> 30  
 <212> ADN  
 <213> Homo sapiens

<400> 5  
 gcagtgagtt caaacccatg acacagaatg 30

<210> 6  
 <211> 31  
 <212> ADN  
 <213> Homo sapiens

<400> 6  
 cagcattctg aatagtcaag attgggaaat g 31

<210> 7  
 <211> 1284  
 <212> ADN  
 <213> Homo sapiens

<400> 7  
 tgatgagggt ggtgatgatg agtattgtga tggatgatgaa gatgaggatg atttaaagaa 60  
 acctttgaaa ctggatgaaa cagatagatt tctcatgact ttattttttg aaaacaataa 120  
 aatgttgaaa aggctggctg aaaccacaga atatgaaaat gaaaagctga ccaaattaag 180

```

aaataccata atggagcaat atactaggac tgaggaatca gcacgaggaa taatctttac 240
aaaaacacga cagagtgcac atgcgcctttc ccagtggatt actgaaaatg aaaaatttgc 300
tgaagtagga gtcaaagccc accatctgat tggagctgga cacagcagtg agttcaaacc 360
catgacacag aatgaacaaa aagaagtcac tagtaaattt cgcactggaa aaataaatct 420
gcttatcgct accacagtgg cagaagaagg tctggatatt aaagaatgta acattgttat 480
ccgttatggc ctcgtcacca atgaaatagc catgggtccag gcccggtggc gagccagagc 540
tgatgagagc acctacgtcc tggttgctca cagtgggttca ggagttatcg aacgtgagac 600
agttaatgat ttccgagaga agatgatgta taaagctata cattgtgttc aaaatatgaa 660
accagaggag tatgctcata agattttgga attacagatg caaagtataa tggaaaagaa 720
aatgaaaacc aagagaaata ttgccaagca ttacaagaat aacccatcac taataacttt 780
cctttgcaaa aactgcagtg tgctagcctg ttctggggaa gatatccatg taattgagaa 840
aatgcatcac gtcaatatga cccagaatt caaggaactt tacattgtaa gagaaaacaa 900
aacactgcaa aagaagtgtg ccgactatca aataaatggt gaaatcatct gcaaatgtgg 960
ccaggcttgg ggaacaatga tgggtgcacaa aggcttagat ttgccttgtc tcaaaataag 1020
gaattttgta gtggttttca aaaataattc aacaaagaaa caatacaaaa agtgggtaga 1080
attacctatc acatttccca atcttgacta ttcagaatgc tgtttattta gtgatgagga 1140
ttagcacttg attgaagatt cttttaaaat actatcagtt aaacatttaa tatgattatg 1200
attaatgtat tcattatgct acagaactga cataagaatc aataaaatga ttgttttacc 1260
ntcaaaaaaa aaaaaaaaaa aaaa 1284

```

<210> 8  
 <211> 29  
 <212> ADN  
 <213> Homo sapiens

<400> 8  
 cacaatactc atcatcacca ccctcatca 29

<210> 9  
 <211> 27  
 <212> ADN  
 <213> Homo sapiens

<400> 9  
 gtagggcctt attgtacttc ctcaaatt 27

<210> 10  
 <211> 1443  
 <212> ADN  
 <213> Homo sapiens

<400> 10

```

gaaagaaaac tggttctctg cttttctgaa tgttcttcgt caaacaggaa acaatgaact 60
tgtccaagag ttaacaggct ctgattgctc agaaagcaat gcagagattg agaatttatc 120
acaagttgat ggtcctcaag tggagaagca acttctttca accacagttc agccaaatct 180
ggagaaggag gtctggggca tggagaataa ctcacagaa tcatcttttg cagattcttc 240
tgtagtttca gaatcagaca caagtttggc agaaggaagt gtcagctgct tagatgaaag 300
tcttgacat aacagcaaca tgggcagtgat ttcaggcacc atgggaagtg attcagatga 360
agagaatgtg gcagcaagag catccccgga gccagaactc cagctcaggc cttaccaaat 420
ggaagttgcc cagccagcct tggaaaggaa gaatatcatc atctgcctcc ctacaggagg 480
tggaaaaacc agagtggctg tttacattgc caaggatcac ttagacaaga agaaaaaagc 540
atctgagcct ggaaaagtta tagttcttgt caataaggta ctgctagttg aacagctctt 600
ccgcaaggag ttccaaccat ttttgaagaa atggtatcgt gttattggat taagtgggtg 660
tacccaactg aaaatatcat ttccagaagt tgtcaagtcc tgtgatatta ttatcagtac 720
agctcaaact cttgaaaact ccctcttaaa cttggaaaat ggagaagatg ctggtgttca 780
attgtcagac ttttccttca ttatcattga tgaatgtcat cacaccaaca aagaagcagt 840
gtataataac atcatgaggc attatttgat gcagaagttg aaaaacaata gactcaagaa 900
agaaaacaaa ccagtgattc cccttcntca gatactggga ctaacagctt cacctggtgt 960

```

```

tggagggggcc acgaagcaag ccaaagctga agaacacatt ttaaaactat gtgccaatct 1020
tgatgcattt actattaata ctgttaaaga aaaccttgat caactgaaaa accaaataca 1080
ggagccatgc aagaagtttg ccattgcaga tgcaaccaga gaagatccat ttaaagagaa 1140
acttctagaa ataatagaca ggattcaaac ttattgtcaa atgagtccaa tgtcagattt 1200
tggaaactcaa ccctatgaac aatggggccat tcaaattggaa aaaaaagctg caaaagaagg 1260
aaatcgcaaa gaaagtgttt gtgcagaaca tttgaggaag tacaataagg ccctacaaat 1320
taatgacaca attcgaatga tagatgcgta tactcatctt gaaactttct ataataaga 1380
gaaagataag aagtttgcag tcatagaaga tgatagtgat gagggtggtg atgatgagta 1440
ttg 1443

```

```

<210> 11
<211> 21
<212> ADN
<213> Homo sapiens

```

```

<400> 11
ctccaacacc aggtgaagct g 21

```

```

<210> 12
<211> 21
<212> ADN
<213> Homo sapiens

```

```

<400> 12
cagatgaaga gaatgtggca g 21

```

```

<210> 13
<211> 24
<212> ADN
<213> Homo sapiens

```

```

<400> 13
ggaagtacaa tgagggccta caaa 24

```

```

<210> 14
<211> 23
<212> ADN
<213> Homo sapiens

```

```

<400> 14
tcctcagtcct agtatattgc tcc 23

```

```

<210> 15
<211> 39
<212> ADN
<213> Homo sapiens

```

```

<400> 15
ctaagcagct gacacttcct tctgccaac ttgtgtctg 39

```

```

<210> 16
<211> 27
<212> ADN
<213> Homo sapiens

```

<400> 16  
gggccctgtg gacaacctcg tcattgt 27

<210> 17  
<211> 33  
<212> ADN  
<213> Homo sapiens

<400> 17  
ccagagtggc tgtttacatt gccaaaggatc act 33

<210> 18  
<211> 30  
<212> ADN  
<213> Homo sapiens

<400> 18  
gcatctgcaa tggcaaactt cttgcatggc 30

<210> 19  
<211> 24  
<212> ADN  
<213> Homo sapiens

<400> 19  
gccatcaatg accccttcat tgac 24

<210> 20  
<211> 24  
<212> ADN  
<213> Homo sapiens

<400> 20  
tgacgaacat gggggcatca gcag 24

<210> 21  
<211> 30  
<212> ADN  
<213> Homo sapiens

<400> 21  
tgagaggatc cgatgtcgaa tgggtattcc 30

<210> 22  
<211> 30  
<212> ADN  
<213> Homo sapiens

<400> 22  
aatgtcgacc taatcctcat cactaaataa 30

<210> 23  
<211> 36  
<212> ADN

<213> Homo sapiens

<400> 23

tgagagctcg agatgtcgaa tgggtattcc acagac

36

<210> 24

<211> 36

<212> ADN

<213> Homo sapiens

<400> 24

tgtttattta gtgatgagga tcgggatccg attgaa

36

<210> 25

<211> 33

<212> ADN

<213> Homo sapiens

<400> 25

ttcaatctcg agatcctcat cactaaataa aga

33

<210> 26

<211> 30

<212> ADN

<213> Homo sapiens

<400> 26

cgtgctgatt cctcagtcct agtatattgc

30

<210> 27

<211> 13

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 27

tatcgactcc aag

13

<210> 28

<211> 13

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 28

ttagctagca tgg

13

<210> 29

<211> 13

<212> ADN  
<213> Séquence artificielle

<220>  
<223> Description de la séquence artificielle: Amorce

<400> 29  
tgctaagact agc

13

<210> 30  
<211> 13  
<212> ADN  
<213> Séquence artificielle

<220>  
<223> Description de la séquence artificielle: Amorce

<400> 30  
ttgcagtgtg tga

13

<210> 31  
<211> 13  
<212> ADN  
<213> Séquence artificielle

<220>  
<223> Description de la séquence artificielle: Amorce

<400> 31  
tgtgaccatt gca

13

<210> 32  
<211> 13  
<212> ADN  
<213> Séquence artificielle

<220>  
<223> Description de la séquence artificielle: Amorce

<400> 32  
tgtctgctag gta

13

<210> 33  
<211> 13  
<212> ADN  
<213> Séquence artificielle

<220>  
<223> Description de la séquence artificielle: Amorce

<400> 33  
tgcatggtag tct

13

<210> 34  
<211> 13  
<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 34

tgtgttgac cat

13

<210> 35

<211> 13

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 35

tagacgctag tgt

13

<210> 36

<211> 13

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 36

ttagctagca gac

13

<210> 37

<211> 13

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 37

tcatgatgct acc

13

<210> 38

<211> 13

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 38

tactccatga ctc

13

<210> 39

<211> 13

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 39

tattacaacg agg

13

<210> 40

<211> 13

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 40

tattggattg gtc

13

<210> 41

<211> 13

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 41

tatctttcta ccc

13

<210> 42

<211> 13

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 42

tatttttggc tcc

13

<210> 43

<211> 13

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 43

ttatctatac agg

13

<210> 44

<211> 13

<212> ADN

<213> Séquence artificielle

<400> 16  
gggccctgtg gacaacctcg tcattgt 27

<210> 17  
<211> 33  
<212> ADN  
<213> Homo sapiens

<400> 17  
ccagagtggc tgtttacatt gccaggatc act 33

<210> 18  
<211> 30  
<212> ADN  
<213> Homo sapiens

<400> 18  
gcatctgcaa tggcaaactt cttgcatggc 30

<210> 19  
<211> 24  
<212> ADN  
<213> Homo sapiens

<400> 19  
gccatcaatg accccttcat tgac 24

<210> 20  
<211> 24  
<212> ADN  
<213> Homo sapiens

<400> 20  
tgacgaacat gggggcatca gcag 24

<210> 21  
<211> 30  
<212> ADN  
<213> Homo sapiens

<400> 21  
tgagaggatc cgatgtcgaa tgggtattcc 30

<210> 22  
<211> 30  
<212> ADN  
<213> Homo sapiens

<400> 22  
aatgtcgacc taatcctcat cactaaataa 30

<210> 23  
<211> 36  
<212> ADN

<213> Homo sapiens

<400> 23

tgagagctcg agatgtcgaa tgggtattcc acagac 36

<210> 24

<211> 36

<212> ADN

<213> Homo sapiens

<400> 24

tgtttattta gtgatgagga tcgggatccg attgaa 36

<210> 25

<211> 33

<212> ADN

<213> Homo sapiens

<400> 25

ttcaatctcg agatcctcat cactaaataa aga 33

<210> 26

<211> 30

<212> ADN

<213> Homo sapiens

<400> 26

cgtgctgatt cctcagtcct agtatattgc 30

<210> 27

<211> 13

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 27

tatcgactcc aag 13

<210> 28

<211> 13

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 28

ttagctagca tgg 13

<210> 29

<211> 13

<212> ADN  
<213> Séquence artificielle

<220>  
<223> Description de la séquence artificielle: Amorce

<400> 29  
tgctaagact agc

13

<210> 30  
<211> 13  
<212> ADN  
<213> Séquence artificielle

<220>  
<223> Description de la séquence artificielle: Amorce

<400> 30  
ttgcagtggtg tga

13

<210> 31  
<211> 13  
<212> ADN  
<213> Séquence artificielle

<220>  
<223> Description de la séquence artificielle: Amorce

<400> 31  
tgtgaccatt gca

13

<210> 32  
<211> 13  
<212> ADN  
<213> Séquence artificielle

<220>  
<223> Description de la séquence artificielle: Amorce

<400> 32  
tgtctgctag gta

13

<210> 33  
<211> 13  
<212> ADN  
<213> Séquence artificielle

<220>  
<223> Description de la séquence artificielle: Amorce

<400> 33  
tgcatggtag tct

13

<210> 34  
<211> 13  
<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 34

tgtgttgac cat

13

<210> 35

<211> 13

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 35

tagacgctag tgt

13

<210> 36

<211> 13

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 36

ttagctagca gac

13

<210> 37

<211> 13

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 37

tcatgatgct acc

13

<210> 38

<211> 13

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 38

tactccatga ctc

13

<210> 39

<211> 13

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 39

tattacaacg agg

13

<210> 40

<211> 13

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 40

tattggattg gtc

13

<210> 41

<211> 13

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 41

tatctttcta ccc

13

<210> 42

<211> 13

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 42

tatttttggc tcc

13

<210> 43

<211> 13

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 43

ttatctatac agg

13

<210> 44

<211> 13

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 44

ttatggtaaa ggg

13

<210> 45

<211> 13

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 45

ttatcgggtca tag

13

<210> 46

<211> 13

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Amorce

<400> 46

ttaggtacta agg

13

<210> 47

<211> 24

<212> ADN

<213> Homo sapiens

<400> 47

gaaagagagg tcgcagaggc ctgt

24

<210> 48

<211> 24

<212> ADN

<213> Homo sapiens

<400> 48

tgataaggct gaggaaggga aatg

24

<210> 49

<211> 21

<212> ADN

<213> Homo sapiens

<400> 49

ctagaccctt ggaagcatcc a

21

<210> 50

<211> 21

<212> ADN

<213> Homo sapiens

<400> 50

tcgggcctgt cgggtcccct c

21

<210> 51

<211> 20

<212> ADN

<213> Homo sapiens

<400> 51

gggtcagaag gattcctatg

20

<210> 52

<211> 20

<212> ADN

<213> Homo sapiens

<400> 52

ggtctcaaac atgatctggg

20

<210> 53

<211> 27

<212> ADN

<213> Virus immunodeficient humain type 1

<220>

<223> GAG06

<220>

<223> n en position 3, 6, 12 et 18 correspond à l'inosine

<400> 53

gcnttnagcc cngaagtnat acccatg

27

<210> 54

<211> 28

<212> ADN

<213> Virus immunodeficient humain type 1

<220>

<223> GAG04

<220>

<223> n en position 4 et 15 correspond à l'inosine

<400> 54

catnctatatt gttcntgaag ggtactag

28

<210> 55

<211> 28

<212> ADN

<213> Virus immunodeficient humain type 1

<220>

<223> BSS

<220>

<223> n en position 11, 16 et 20 correspond à l'inosine

<220>

<223> n en position 13 correspond à a, g, c ou t/u

<400> 55

ggcttgctga ngngcncacn gcaagagg

28

<210> 56

<211> 28

<212> ADN

<213> Virus immunodeficient humain type 1

<220>

<223> KPNA

<220>

<223> n en position 6 correspond à l'inosine

<220>

<223> n en position 14 et 18 correspond à a, g, c ou t/u

<400> 56

agagtngtgg ttnttcntt ccacacag

28

**REVENDICATIONS**

1. Polypeptide isolé dénommé RH116 de séquence d'acides aminés  
SEQ ID N°2.  
5
2. Polypeptide isolé caractérisé en ce qu'il comprend un polypeptide  
choisi parmi :
  - a) un polypeptide de séquence SEQ ID N°2 ;
  - b) un polypeptide variant de polypeptide de séquences d'acides  
10 aminés défini en a) ;
  - c) un polypeptide homologue au polypeptide défini en a) ou b)  
et comportant au moins 80 % d'identité avec ledit  
polypeptide de a) ;
  - d) un fragment d'au moins 15 acides aminés consécutifs d'un  
15 polypeptide défini en a) ;
  - e) un fragment biologiquement actif d'un polypeptide défini en  
a).
3. Polypeptide selon l'une quelconque des revendications 1 à 2  
20 caractérisé en ce qu'il est comporte au moins un domaine conservé  
appartenant à la superfamille des hélicases à ARN.
4. Polypeptide selon la revendication 3 caractérisé en ce que ledit  
domaine conservé est choisi parmi :
  - 25 - les séquences G-GKT correspondant au domaine I de la  
superfamille des hélicases à ARN.
  - les séquences DEAD, DE-D, DEAH correspondant au domaine V  
de la superfamille des hélicases à ARN.
- 30 5. Polynucléotide purifié ou isolé caractérisé en ce qu'il code pour un  
polypeptide selon la revendication 1.
6. Polynucléotide selon la revendication 5 de séquence SEQ ID N°1.

7. Polynucléotide isolé caractérisé en ce qu'il comprend un polynucléotide choisi parmi :

- a) SEQ ID N° 1;
- 5 b) la séquence d'un fragment d'au moins 15 nucléotides consécutifs de la séquence SEQ ID N°1 à l'exception des séquences nucléiques enregistrées identifiées sous le numéro d'accès N°AC007750 et N°AC0108176 dans la banque de données GenBank et les séquences  
10 nucléiques enregistrées identifiées sous le numéro d'accès N°AW589567, N°AW152541 et AW189584 dans la banque de données EMBL ;
- c) une séquence nucléique présentant un pourcentage d'identité d'au moins 85 %, après alignement optimal  
15 avec une séquence définie en a) ou b) ;
- d) la séquence complémentaire ou la séquence d'ARN correspondant à une séquence telle que définie en a), b) ou c).

20 8. Utilisation d'un polynucléotide selon la revendication 7 en tant qu'amorce pour l'amplification ou la polymérisation de séquences nucléiques.

9. Utilisation in vitro d'un polynucléotide selon la revendication 7 en  
25 tant que sonde pour la détection de séquences nucléiques.

10. Utilisation in vitro d'un polynucléotide selon la revendication 7 en tant que séquence d'acide nucléique sens ou antisens pour contrôler l'expression du produit protéique correspondant.

30

11. Utilisation d'un polynucléotide selon l'une quelconque des revendications 8, 9, 10 caractérisé en ce que ledit polynucléotide est

marqué directement ou indirectement par un composé radioactif ou un composé non radioactif.

12. Vecteur recombinant de clonage et/ou d'expression comprenant  
5 un polynucléotide selon l'une des revendications 5 à 7 ou codant pour un polypeptide selon l'une quelconque des revendications 1 à 4.

13. Vecteur recombinant d'expression antisens comprenant un  
10 polynucléotide selon l'une des revendications 5 à 7 caractérisé en ce que ledit polynucléotide est inséré en orientation inverse dans ledit vecteur.

14. Cellule hôte, caractérisée en ce qu'elle est transformée par un  
15 vecteur selon l'une des revendications 12 et 13.

15. Animal, excepté l'homme, caractérisé en ce qu'il comprend une cellule selon la revendication 14.

20 16. Procédé de préparation d'un polypeptide recombinant caractérisé en ce que l'on cultive une cellule hôte selon la revendication 14 dans des conditions permettant l'expression et éventuellement la sécrétion dudit polypeptide recombinant et que l'on récupère ledit polypeptide recombinant.

25

17. Polypeptide recombinant obtenu par le procédé selon la revendication 16.

18. Anticorps monoclonal ou polyclonal et ses fragments caractérisé  
30 en ce qu'il lie sélectivement un polypeptide selon l'une des revendications 1 à 4 ou 17.

19. Procédé de détection et/ou de dosage d'un polypeptide selon l'une des revendications 1 à 4 ou 17, dans un échantillon biologique, caractérisé en ce qu'il comprend les étapes suivantes :

- 5 a) mise en contact de l'échantillon biologique avec un anticorps selon la revendication 18;
- b) mise en évidence du complexe antigène-anticorps formé.

20. Trousse de réactifs pour la mise en œuvre d'un procédé selon la revendication 19 dans un échantillon biologique par réaction immunologique, caractérisé en ce qu'elle comprend les éléments suivants :

- 15 a) un anticorps monoclonal ou polyclonal selon la revendication 18;
- b) le cas échéant, les réactifs pour la constitution du milieu propice à la réaction immunologique ;
- c) les réactifs permettant la détection du complexe antigène-anticorps produit lors la réaction immunologique.

21. Procédé de détection et/ou de dosage d'un polynucléotide selon l'une quelconque des revendications 5 à 7 dans un échantillon biologique, caractérisé en ce qu'il comporte les étapes suivantes :

- 20 a) isolement de l'ADN à partir de l'échantillon biologique à analyser, ou obtention d'un ADNc à partir de l'ARN de l'échantillon biologique ;
- 25 b) amplification spécifique de l'ADN à l'aide d'amorces selon la revendication 8 ;
- c) analyse des produits d'amplification.

22. Procédé de détection et/ou de dosage d'un polynucléotide selon l'une quelconque des revendications 5 à 7 dans un échantillon biologique, caractérisé en ce qu'il comporte les étapes suivantes :

- 5           a) Mise en contact d'un polynucléotide selon l'une des revendications 5 à 7 avec un échantillon biologique ;
- b) Détection et/ou dosage de l'hybride formé entre ledit polynucléotide et l'acide nucléique de l'échantillon biologique.

23. Puce à ADN caractérisée en ce qu'elle contient un polynucléotide selon l'une des revendications 5 à 7.

- 10   24. Puce à protéines caractérisée en ce qu'elle contient un polypeptide selon l'une des revendications 1 à 4 ou 17, ou un anticorps selon la revendication 18.

15   25. Procédé de criblage d'un composé qui affecte le taux d'expression cellulaire et/ou l'activité hélicase à ARN d'un polypeptide selon les revendications 1 à 4 ou 17 et qui comporte les étapes suivantes de:

- 20           a) mise en contact d'une cellule choisie parmi la cellule hôte de la revendication 14 et une cellule eucaryote, de préférence humaine, exprimant ou contenant le polypeptide selon la revendication 1 à 4 ou 17, et d'un ou plusieurs composés potentiels susceptibles de pénétrer ou d'être introduits dans ladite cellule;
- b) détection et/ou mesure du taux d'expression cellulaire et/ou de l'activité hélicase à ARN.

25

26. Procédé de criblage d'un composé qui affecte l'activité hélicase à ARN d'un polypeptide selon les revendications 1 à 4 ou 17 et qui comporte les étapes suivantes de:

- 30           a) mise en contact dudit polypeptide et d'un ou plusieurs composé(s) potentiel(s) en présence de réactifs nécessaires à la mise en œuvre de l'activité hélicase à ARN;
- b) détection et/ou mesure de l'activité hélicase à ARN.

27. Procédé selon les revendications 25 et 26 caractérisé en ce que ledit composé criblé diminue le taux d'expression et/ou l'activité hélicase à ARN du polypeptide selon les revendications 1 à 4 ou 17.

5

28. Composé obtenu par le procédé selon la revendication 27 caractérisé en ce qu'il est choisi parmi :

- a) un polynucléotide selon l'une quelconque des revendications 5 à 7 utilisé en tant que séquence d'acide nucléique antisens selon la revendication 10 ;
- b) un vecteur d'expression antisens selon la revendication 13 ;
- c) un anticorps selon la revendication 18 ;
- d) un antagoniste du polypeptide selon l'une quelconque des revendications 1 à 4 ou 17 ;
- e) les muramylpeptides.

20

29. Composé selon la revendication 28 caractérisé en ce que le muramylpeptide est le Murabutide.

30. Procédé selon les revendications 25 et 26 caractérisé en ce que ledit composé criblé augmente le taux d'expression et/ou l'activité hélicase à ARN du polypeptide selon les revendications 1 à 4 ou 17.

31. Composé agoniste du polypeptide selon l'une quelconque des revendications 1 à 4 ou 17 obtenu par le procédé selon la revendication 30.

32. Procédé de criblage de composés qui affectent l'activité fonctionnelle d'un polypeptide selon les revendications 1 à 4 ou 17 et qui comporte les étapes suivantes de:

- 5 a) mise en contact dudit polypeptide et d'un ou plusieurs composé(s) potentiel(s) en présence de réactifs nécessaires à la mise en œuvre d'une réaction choisie parmi la réaction d'épissage de l'ARN nucléaire et/ou mitochondrial, de réaction d'édition de l'ARN, de réaction de maturation de l'ARNr, de réaction d'initiation de la traduction, de la réaction d'export des ARNm nucléaires vers le cytoplasme, et la réaction de dégradation des ARNm;
- 10 b) détection et/ou mesure de ladite réaction.

10

33. Composé caractérisé en ce qu'il est choisi parmi :

- 15 a) un polypeptide selon l'une des revendications 1 à 4 ou 17;  
b) un polynucléotide selon l'une des revendications 5 à 7 ;  
c) un vecteur selon la revendication 12 ou 13 ;  
15 d) une cellule selon la revendication 14  
e) un anticorps selon la revendication 18 ;  
f) un composé selon l'une des revendications 28, 29 et 31 ;  
à titre de médicament.

- 20 34. Composé selon la revendication 33 à titre de médicament destiné à la prévention et/ou au traitement de pathologie sélectionnée dans le groupe composé du cancer, des maladies infectieuses aiguës ou chroniques, des maladies génétiques héréditaires, des maladies immunes et auto-immunes, du rhumatisme, de l'arthrite, de  
25 l'arthérosclérose, de l'ostéoporose, des diabètes et à la prévention des rejets de greffes d'organes.

35. Composé selon la revendication 34 caractérisé en ce que ladite maladie infectieuse est sélectionnée parmi le SIDA ou l'hépatite C.

30

36. Utilisation d'un composé selon la revendication 33 pour la préparation d'un médicament destiné au traitement de pathologies virales.

5 37. Utilisation selon la revendication 36 caractérisé en ce que la pathologie virale est le syndrome d'immunodéficience acquise (SIDA).

38. Composition pharmaceutique pour le traitement préventif et curatif du SIDA caractérisée en ce qu'elle contient une quantité  
10 thérapeutiquement efficace d'un composé selon la revendication 33 et un véhicule pharmaceutiquement acceptable.

39. Produit comprenant au moins un composé selon la revendication 33 et au moins un autre agent antiviral comme produit de  
15 combinaison pour une utilisation simultanée, séparée ou étalée dans le temps en thérapie anti-virale, de préférence anti-VIH.

40. Utilisation selon la revendication 36 caractérisé en ce que la pathologie virale est l'hépatite B ou C .  
20

41. Utilisation d'un polynucléotide selon l'une quelconque des revendications 5 à 7 et/ou d'un polypeptide selon l'une quelconque des revendications 1 à 4 et 17 et/ou d'un agoniste d'un polypeptide selon l'une quelconque des revendications 1 à 4 et 17 pour la  
25 préparation d'un médicament destiné à provoquer ou à augmenter la réponse immune à un vaccin chez un patient.

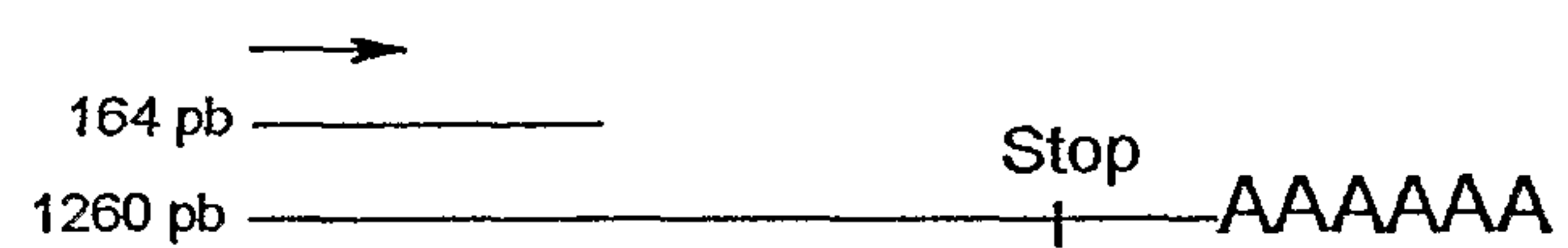


FIGURE 1

2 / 11

FIGURE 2a

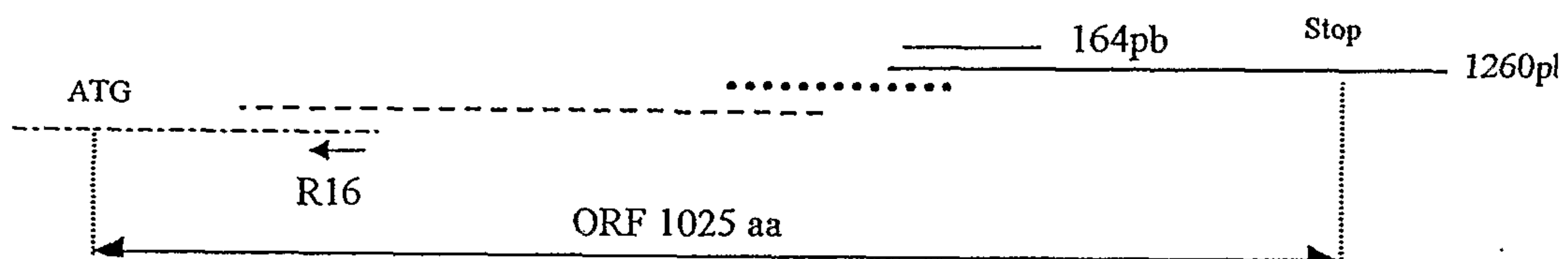
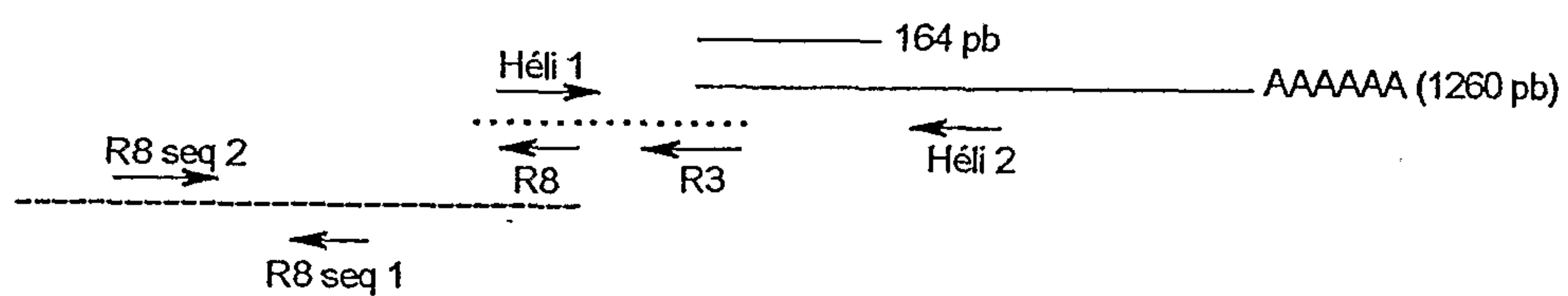


FIGURE 2b

3 / 11

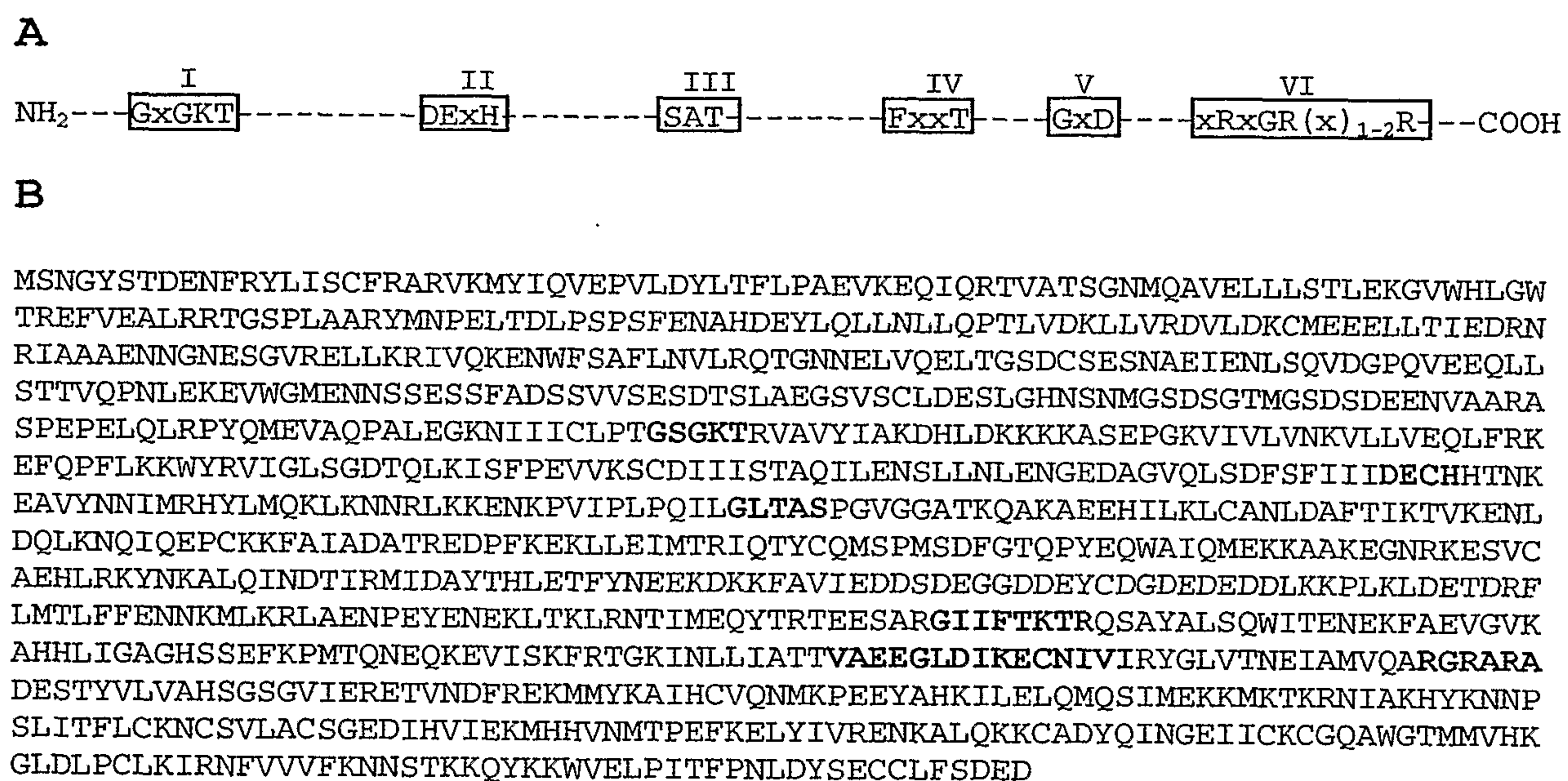


FIGURE 3a

4 / 11

C

```

MSGYSTDENFRYLISCFRARVKMYIQVEPVLDTLT-FLPAEVKEQIQRTVATSGNMQAVELLSTLEKG 69
M      T E R + F+ ++ +      +L Y+ + E + IQ      G M+A L L L +
M-----TTEQ-RRSLQAFQDYIRKTLDPITYILSYMAPWFFREEEVQYIQAENKNGPMEATLFLKFLLE- 63

VWHLGWTREFVEALRRTGSPLAARYMNPETDLPSPSFENAHDEYLQLLNLLQPTLVDKLLVRDV-LDKC 138
+   GW R F++AL   G   + Y   E D           +EY LL LQP   +++ D+ D
LQEEGWFRGFLDALDHAG--YSGLYEAIESWDFKKIE---KLEEYRLLLKRLQPEFKTRIIPDTIISDLS 128

MEEELLTIEDNRNRIAAAENN--GNESGVRELLKRIVQKENWFS AFLNVLROTGNNELVQELTGSDCSES 206
      E   I +      G E VE LL R.   KENW      L   N   EL
ECLINQECCEEILQICSTKGMMAGAEKLVECLL-RSD-KENWPKTLKLALKEKERNKFS--ELWIVEKGIKD 194

AEIENLSQVDGPQVEEQLLSTTVQPNLEKEVWGMENNSSESSFADSSVSESDTSLAEGSVSCLDESIGH 276
E E L      Q Q      Q NL      SE S      S VS      DT
VETEDLEDKMDIQIFYQEDPEC-Q-NL-----SENSCPP-SEVS---DTNLYSPFK----- 242

NSNMGSDSGTMGSDSDEENVAARASPEPELQLRPYQMEVAQPALEGKNIIICLPTGSGKTRVAVYIAKDH 346
      R YQ+E+A PA++GKN IIC PTG GKT V++ I+ H
-----RNYQLELALPAMKGKNTIICAPTGGCKTFVSLICEHH 281

LDKKKKASEPGKVIVLVNKVLLVEQ---LFRKEFQPFLLKQWYRVIGLSGDTQLKISFPEVVKSCDIIISTA 414
L KK      + GKV      N      EQ +F K F+      + YRV G+SG T      + ++V++ DIII T
L-KKFPQGQKGVVFFANQIPVYEQQKSVFSKYFE---RHGYRVTGISGATAENVPVEQIVENNDIILTP 348

QILENSLLNLENGEDAGVQLSDFSFIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKENKPVIPLPQI 484
QIL N L      LS F+ +I DECH-T+K+ YN IM +YL QKL      PLPQ+
QILVNNLKKGTIPS-----LSIFTLMIFDECHNTSKQHPYNMIMFNLYDQKLGGSSG-----PLPQV 405

LGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCKKFAIADATREDPFKEKLL 555
+GLTASGVG A      +A ++I KLCA+LDA I TVK NL++L+      + +P K F      ++ D FK      +
IGLTASGVGDAKNTDEALDYICKLCASLDASVIATVKHNLEELEQVVYKPKQFFRKVESRISDKFYIIA 475

EIMTRIQTICQ-----MSPMS-----DFGTQPYEQWAIQMEKKA-----AKEGNRKESVC-----AEHLRK 606
++M      ++      +      + +S      +FGTQ YEQW + ++K      + + + +C      HLRK
QLMRDTESLAKRICKDLENLSQIQNREFGTQKYEQWIVTVQKACMVFPMPDKDEESRICKALFLYTSHLRK 547

YNKALQINDTIRMIDAYTHLETIFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDEDDLKPLKLDDETDRFLM 677
YN AL I++ RM DA +L+ F++      + F IE D
YNDALIISEHARMKDALDYLDKDFFSNVRAAGFEEIEQDLTQ----- 588

TLFFENN-KMLKRLAENPEYENEKLTCLRNTIMEQYTRTEESARGIIFTKTRQSAYALSQWITENEKFAEV 747
FE      + L+ ++ +P EN KL L I E+Y      E+      I+F KTR      AL WI N K+
--RFEELQLELESVSRDPSNENPKLEDLCFILQEEYHLNPETITILFVKTRALVDALKNWIEGNPKLS--- 655

GVKAHHLIGAGHSSEFKPMTQNEQKEVISKFR-TGKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMV 817
+K      L G G +++ MT      QK ++ F+ +G N+LIAT+VA+EG+DI +CN+VI Y V N M+
FLKPGILTGRGKTNQNTGMTLPAQKCILDFAKASGDHNILIATSVADEGIDIAQCNIIVILYEVGNVIKMI 725

QARGRARADESTYVLVAHSGSGVIERETVNDFREKMMYKAHCVQNMKPEEYAHKILELQMQSIMEKKMK 887
Q RGR RA S L+      S +GVIE+E +N ++EKMM +I +Q      + KIL +Q      + +
QTRGRGRARGSKCFLLT-SNAGVIEKEQINMYKEKMMNDSILRLQTWDEAVFREKILHIQTHEKFIRDSQ 794

TKRNIKHYKNNPSLITFLCKNCSVLACSGEDIHVIEKMHVNMTPEFKELYIVRENKALQKKCADYQIN 957
K      +N      LC+ C LAC D+ VIE+ H+      + FKE ++ R      K      ++
EKPKPVPDKENK----KLICKCKALACYTADVRVIEECHYTVLGDAFKECFVSR-PHPKPKQFSSFEKR 859

GEIIC---KCGQAWGTMMVHKGLDLPCLKIRNFVVVFKNNSTKKQYKKWVELPITFPNLDYSECCLFSDE 1024
+I C      C WG + +K ++P +KI+ FVV      + Y KW +      D +E
AKIFCARQNCSDHWGIHVKYKTFEIPVIKIESFVVEDIATGVQTLYSKWKDFHFEKIPFDPAEMSK---- 925

DStop 1025
-Stop

```

FIGURE 3b

5 / 11

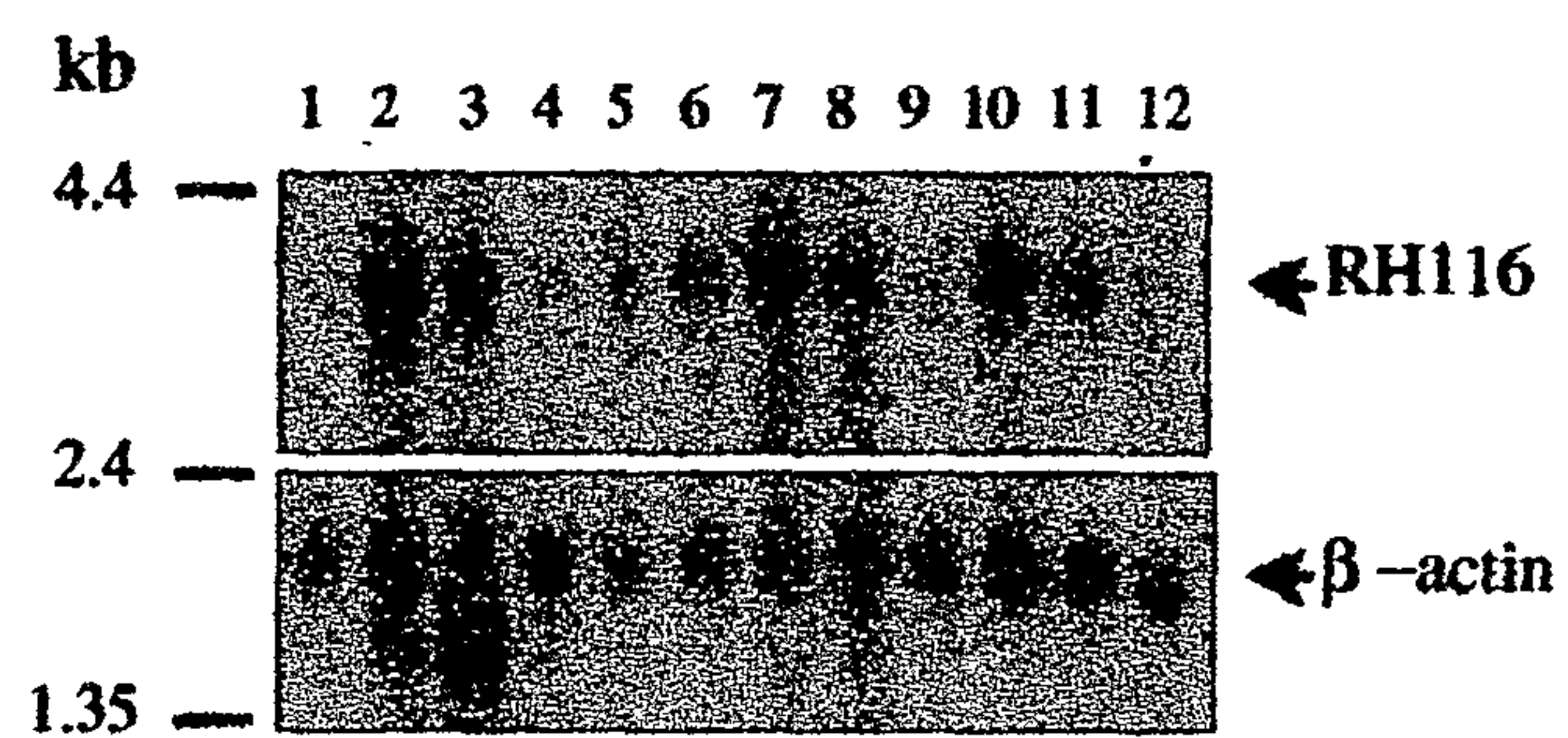


FIGURE 4

6 / 11

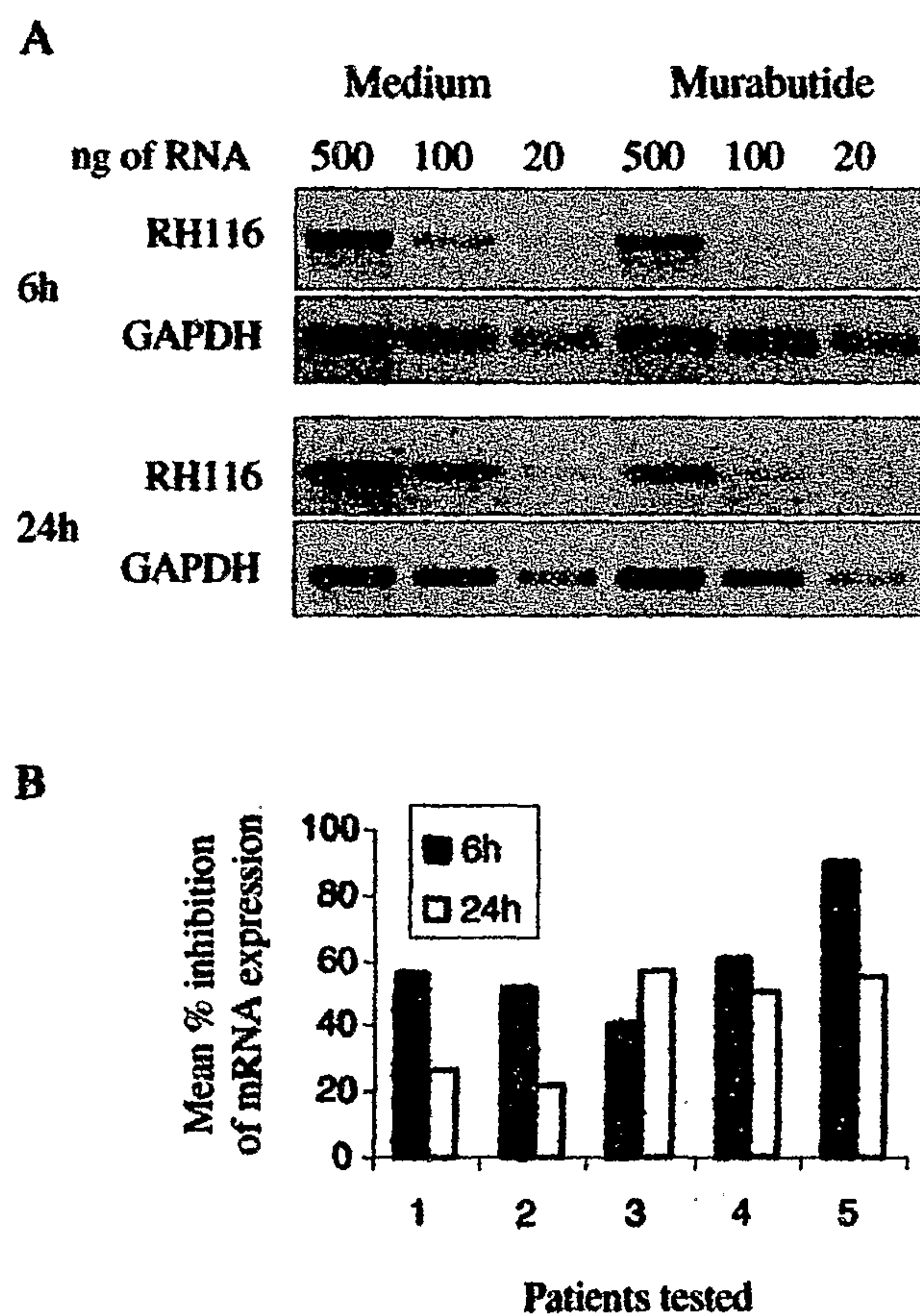


FIGURE 5

7 / 11

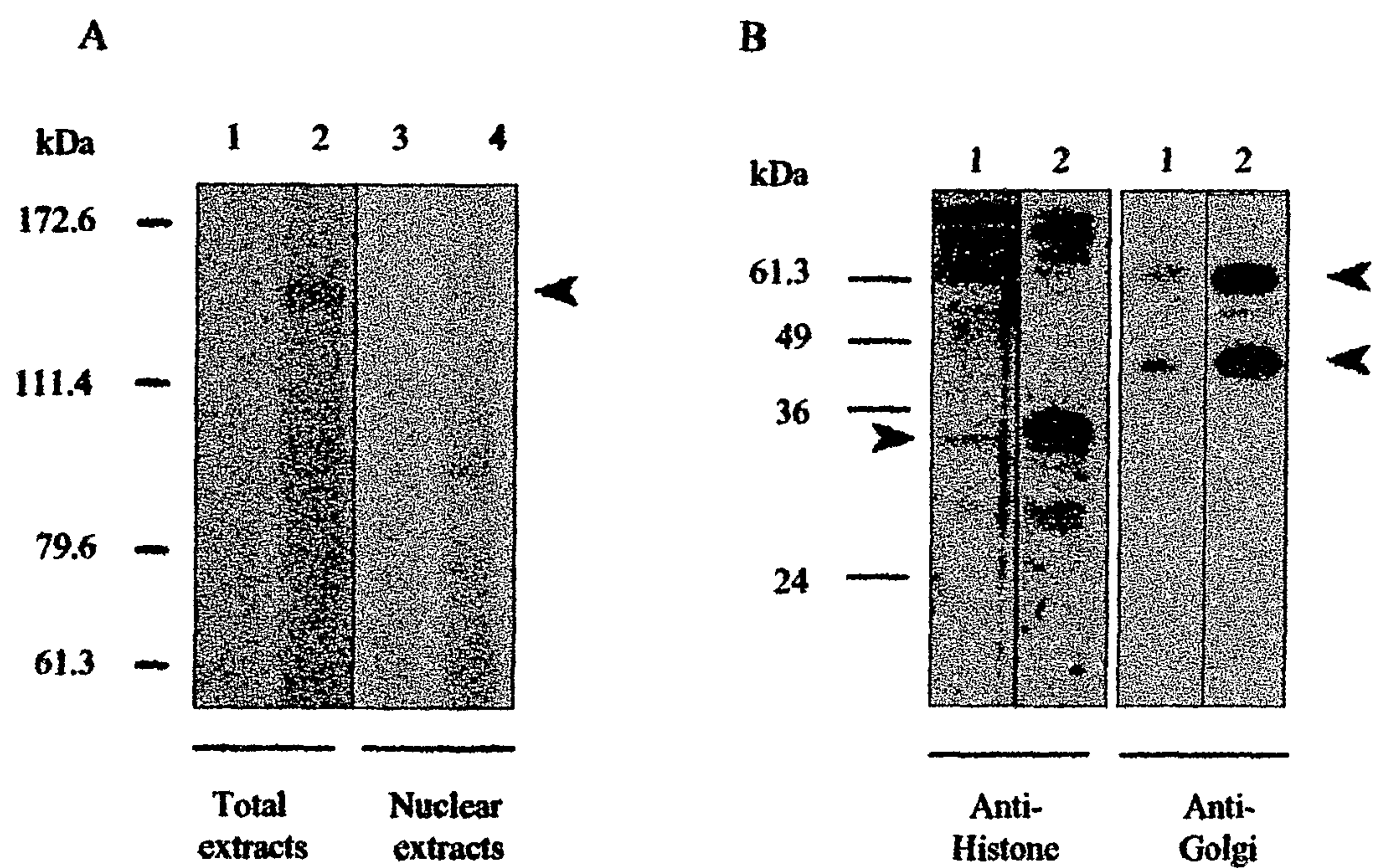


FIGURE 6

8 / 11

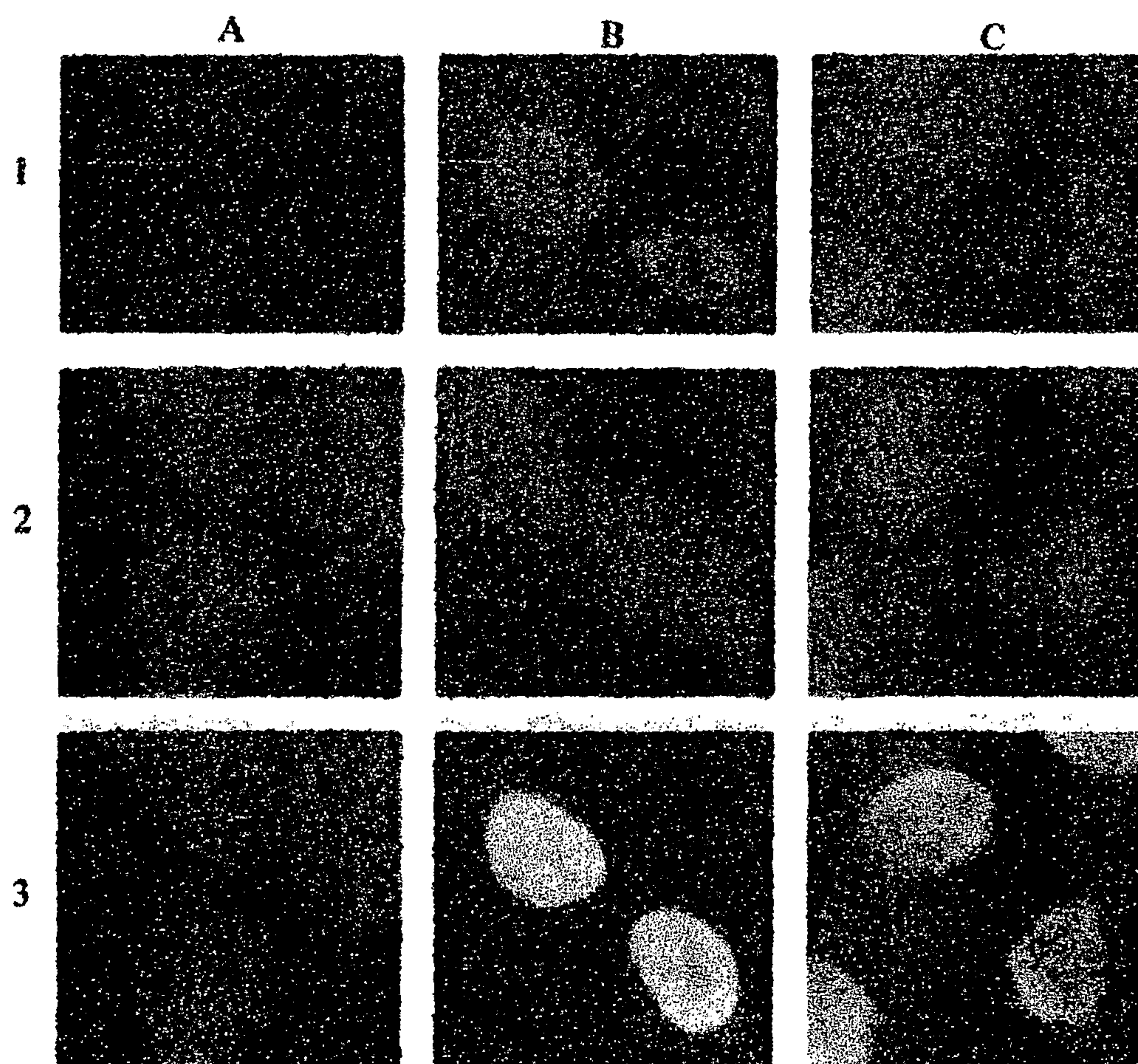


FIGURE 7.

9 / 11

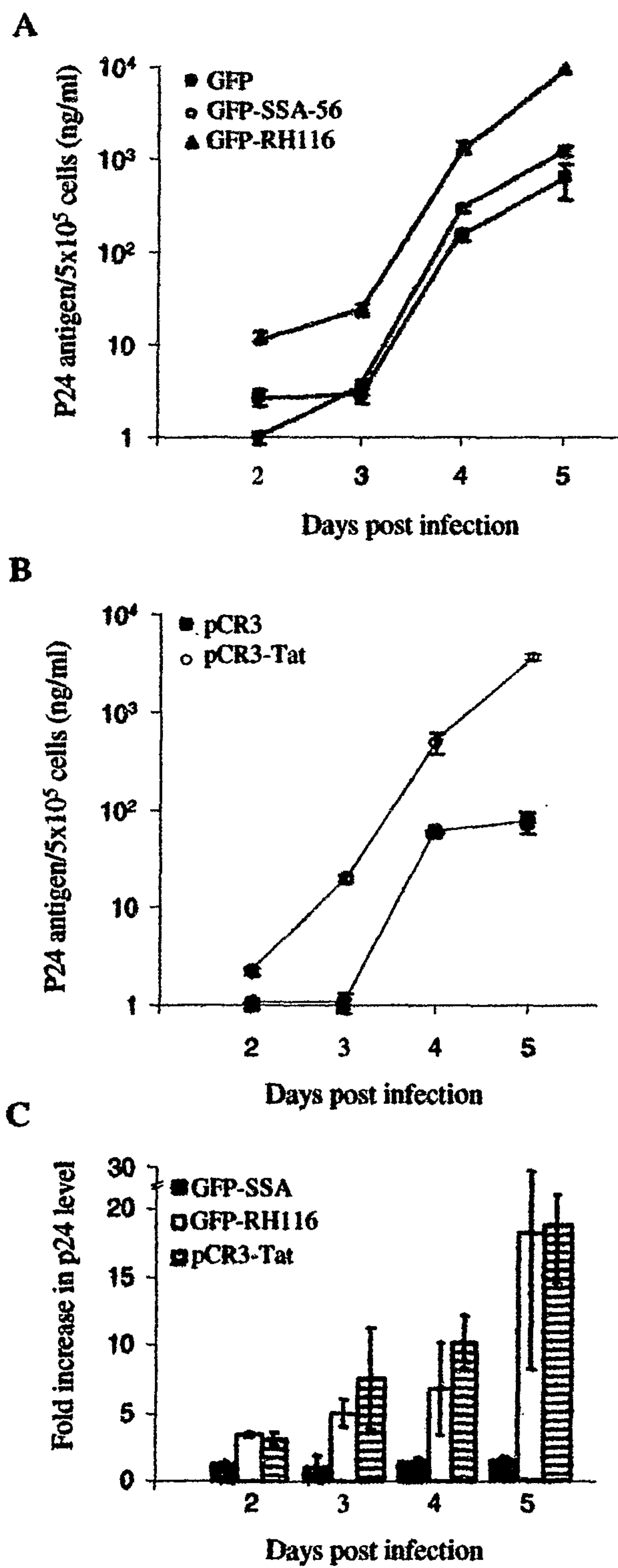


FIGURE 8

10 / 11

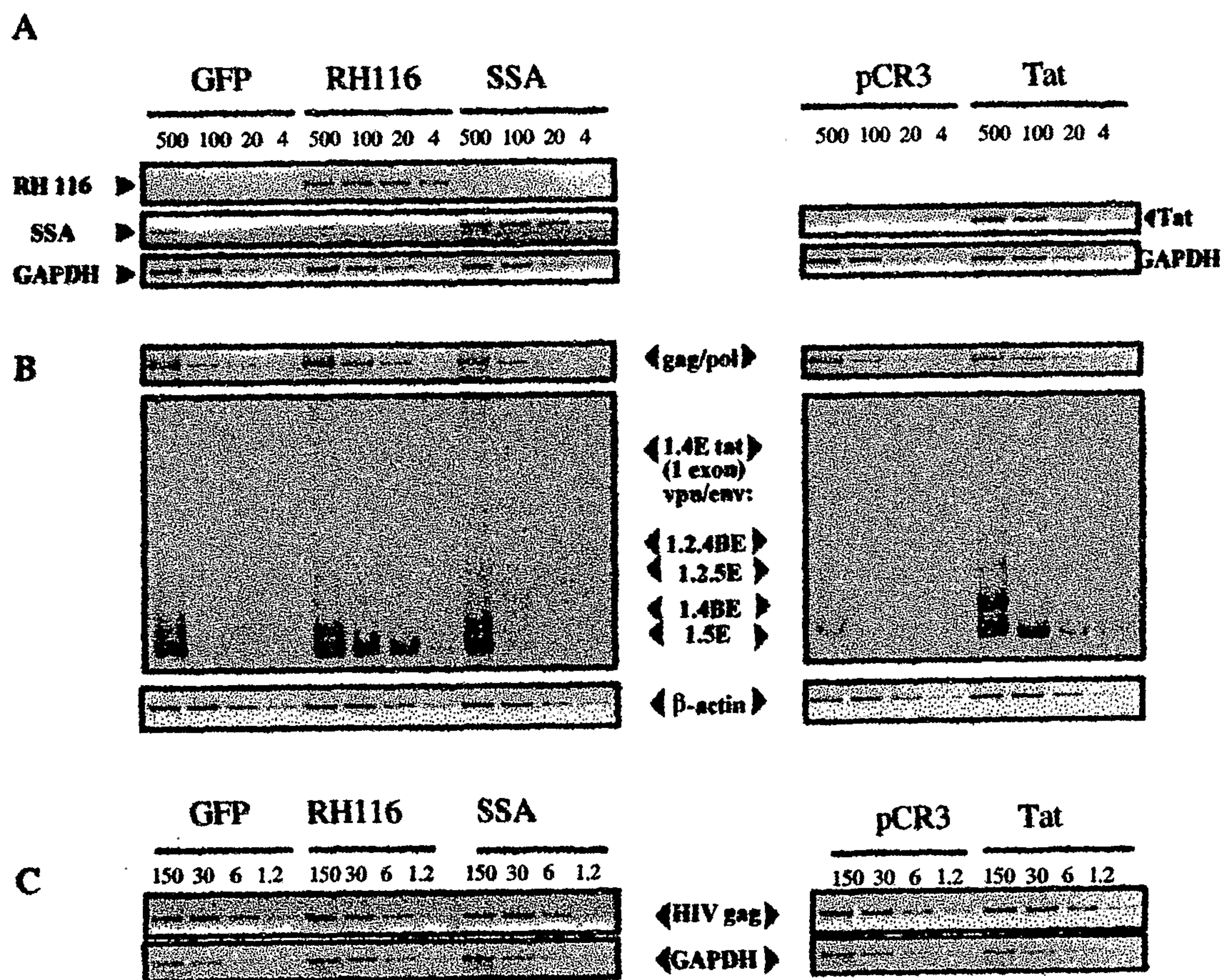


FIGURE 9

11 / 11

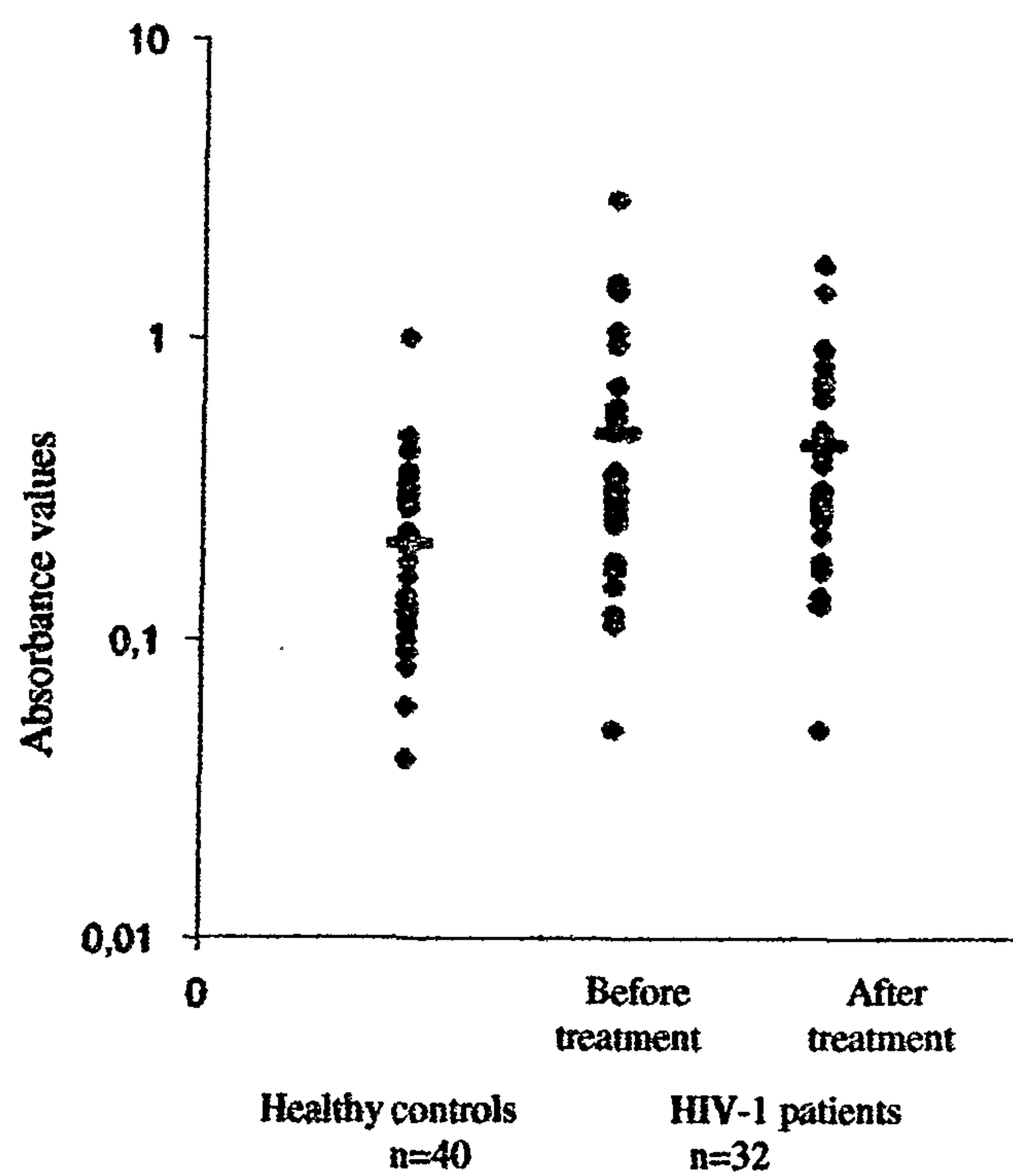


FIGURE 10