



(72) QUERELLOU, Joël, FR

(72) CAMBON, Marie-Anne, FR

(71) APPLIGENE-ONCOR S.A., FR

(51) Int.Cl.<sup>6</sup> C12N 15/54, C12P 19/34, C12N 9/12

(30) 1996/07/10 (96/08631) FR

(54) **HEAT-STABLE DNA POLYMERASE OF ARCHAEOBACTERIA  
OF GENUS PYROCOCCUS SP.**

(54) **ADN POLYMERASE THERMOSTABLE  
D'ARCHAEBACTERIES DU GENRE PYROCOCCUS SP.**

(57) La présente invention concerne une ADN polymérase purifiée thermostable d'archaebactéries du genre Pyrococcus sp. ayant un poids moléculaire compris entre environ 89 000 et 90 000 daltons.

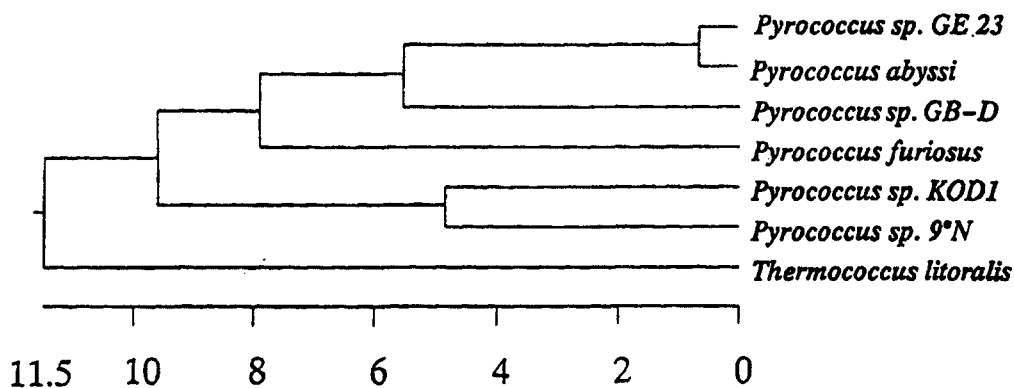
(57) A purified heat-stable DNA polymerase of archaeobacteria of genus Pyrococcus sp. having a molecular weight of around 89,000-90,000 daltons is disclosed.

PCT

ORGANISATION MONDIALE DE LA PROPRIÉTÉ INTELLECTUELLE  
Bureau international

DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITE DE COOPERATION EN MATIÈRE DE BREVETS (PCT)

|  |  |  |
|--|--|--|
| (51) Classification internationale des brevets <sup>6</sup> :<br>C12N 15/54, 15/70, 1/21, 9/12, C12Q 1/68  | A3   | (11) Numéro de publication internationale: <b>WO 98/01567</b><br>(43) Date de publication internationale: 15 janvier 1998 (15.01.98) |
| (21) Numéro de la demande internationale: PCT/FR97/01259<br>(22) Date de dépôt international: 10 juillet 1997 (10.07.97)<br>(30) Données relatives à la priorité:<br>96/08631 10 juillet 1996 (10.07.96) FR<br>(71) Déposant (pour tous les Etats désignés sauf US): APPLIGENE-ONCOR S.A. [FR/FR]; Parc d'Innovation, Rue Geiler de Kaysersberg, Boîte postale 72, F-67402 Illkirch (FR).<br>(72) Inventeurs; et<br>(75) Inventeurs/Déposants (US seulement): QUERELLOU, Joël [FR/FR]; L'Arc'Hantel, F-29280 Brest (FR). CAMBON, Marie-Anne [FR/FR]; 8, allée de Stèle, F-29280 Plouzane (FR).<br>(74) Mandataire: BREESE-MAJERWICZ; 3, avenue de l'Opéra, F-75001 Paris (FR). | (81) Etats désignés: CA, JP, US, brevet européen (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).<br><br>Publiée<br>Avec rapport de recherche internationale.<br>Avant l'expiration du délai prévu pour la modification des revendications, sera republiée si de telles modifications sont reçues.<br>(88) Date de publication du rapport de recherche internationale:<br>9 avril 1998 (09.04.98) |  |

(54) Title: HEAT-STABLE DNA POLYMERASE OF ARCHAEOBACTERIA OF GENUS *PYROCOCCUS SP.*(54) Titre: ADN POLYMERASE THERMOSTABLE D'ARCHAEBACTERIES DU GENRE *PYROCOCCUS SP.*

## (57) Abstract

A purified heat-stable DNA polymerase of archaeobacteria of genus *Pyrococcus sp.* having a molecular weight of around 89,000-90,000 daltons is disclosed.

## (57) Abrégé

La présente invention concerne une ADN polymérase purifiée thermostable d'archaebactéries du genre *Pyrococcus sp.* ayant un poids moléculaire compris entre environ 89 000 et 90 000 daltons.

ADN POLYMERASE THERMOSTABLE  
D'ARCHAEBACTÉRIES DU GENRE *PYROCOCCUS SP.*

5 La présente invention concerne de nouvelles  
ADN polymérasés thermostables provenant  
d'archaebactéries du genre *Pyrococcus sp.*

10 Les ADN polymérasés sont des enzymes  
impliquées dans la réplication et la réparation de  
l'ADN. On connaît aujourd'hui de nombreuses ADN  
polymérasés isolés de micro-organismes tel que *E. coli*;  
par exemple l'ADN polymérase I de *E. coli*, le fragment  
de klenow de l'ADN polymérase I de *E. coli*, l'ADN  
polymérase T4. Des ADN polymérasés thermostables ont  
15 également identifiées et purifiées à partir d'organismes  
thermophiles, tels que *Thermus aquaticus* (Chien, A. et  
al., *J. bacteriol.* 1976, 127:1550-1557 ; Kaladin et al.,  
*Biokhymiyay* 1980, 45:644-651), *Thermus thermophilus*, ou  
encore des espèces *Bacillus* (Demande de Brevet Européen  
No. 699 760), *Thermococcus* (Demande de Brevet Européen  
20 No. 455 430), *Sulfobus* et *Pyrococcus* (Demande de Brevet  
Européen No. 547 359). Parmi celles-ci, on peut citer  
plus particulièrement la Pfu de *Pyrococcus Furiosus*  
(Réf. Biblio.), la Vent DNA polymérase de *Thermococcus*  
*litoralis* (Kong, H. M., R. B. Kucera, and W. E. Jack,  
25 1993. *J. Biol. Chem.* 268(3):1965-1975), la 9°-NDNA  
polymérase de *Pyrococcus sp.* 9°-N et la Deep-Vent DNA  
polymérase de *Pyrococcus sp.* GB-D.

30 Le processus de réplication s'effectue selon  
un mécanisme bien connu consistant à fabriquer, à partir  
d'une matrice, de l'enzyme ADN polymérase et des quatre  
nucléotides triphosphates, un brin d'acide nucléique  
complémentaire de ladite matrice. Les enzymes ayant une  
activité ADN polymérase sont à présent largement  
utilisées *in vitro* dans de nombreux procédés de biologie  
35 moléculaire, tels que le clonage, la détection, le

marquage et l'amplification de séquences d'acide nucléique.

L'amplification de séquences d'acide nucléique par la méthode dite réaction de polymérisation en chaîne (PCR), décrite dans les Brevets Européens No. 200 362 et 201 184, est basée sur la réalisation de cycles successifs d'extensions d'amorces, mettant en oeuvre une ADN polymérase et les quatre nucléotides triphosphates, suivies d'une dénaturation des acides nucléiques double brin ainsi obtenus et servant de matrices pour le cycle suivant. Les températures utilisées à l'étape de dénaturation n'étant pas compatibles avec la conservation des activités de nombreuses ADN polymérases, des travaux de recherche important sont consacrés aux enzymes thermostables décrites précédemment. Il est notamment essentiel ne pas limiter la préparation de ces enzymes aux seuls procédés de purification à partir du microorganisme, mais de chercher à augmenter les rendements de production en utilisant les méthodes du génie génétique. Selon ces méthodes bien connues de l'homme du métier (Maniatis, et al., *Molecular Cloning: A laboratory Manual*, 1982), le gène codant pour l'ADN polymérase est cloné dans un vecteur d'expression, lequel vecteur est inséré dans un hôte cellulaire capable d'exprimer l'enzyme, l'hôte cellulaire est cultivée dans des conditions adéquates et l'ADN polymérase est extraite et récupérée. Cette méthode a par exemple été décrite dans la demande de brevet PCT WO89/06691 pour produire l'ADN polymérase de *Thermus aquaticus*.

Le développement des technologies d'ADN recombinant tant dans le domaine de la recherche que dans celui de la production industrielle nécessite de disposer de divers types d'ADN polymérase susceptibles d'améliorer quantitativement ou qualitativement des

techniques aussi diverses que le clonage, la détection, le marquage ou l'amplification de séquences d'acide nucléique.

5 La présente invention vise précisément à offrir de nouvelles enzymes hyperthermostables obtenues des espèces *Pyrococcus* sp. qui catalysent la polymérisation de l'ADN. Ces enzymes proviennent d'isolats d'échantillons d'archaebactéries hyperthermophiles (Woese, C. R., O. Kandler, and M. Wheelis, 1990. *Proc. Natl. Acad. Sci. USA.* 87:4576-4579) 10 prélevés dans des sources hydrothermales profondes du bassin Nord-Fiji dans le Pacifique sud (Desbruyères, D., A.-M. Alayse-Danet, and S. Ohta. 1994. *Geology.* 116:227-242 ; Marteinsson, V. T., L. Watrin, D. Prieur, J. C. Caprais, G. Raguènes, and G. Erauso. 1995. *International Journal of Systematic Bacteriology.* 15 45(4):623-632). Tous ces isolats sont hyperthermostables avec des températures d'isolement de l'ordre de 80 à 100°C. Une ADN polymérase de l'invention a un poids moléculaire d'environ 89000 daltons et présente une hyperthermostabilité permettant sa mise en oeuvre dans des réactions conduites à des températures de 70 à 90°C. 20

Les travaux réalisés dans le cadre de l'invention ont permis d'identifier deux nouvelles ADN 25 polymérases thermostables d'archaebactéries du genre *Pyrococcus* sp. dont la relation phylogénétique a été étudiée. Les gènes codant pour ces deux ADN polymérases ont été clonés et séquencés et leur comparaison a mis en évidence de fortes analogies de séquence et d'organisation. 30

L'invention a plus particulièrement pour objet une ADN polymérase purifiée thermostable d'archaebactéries du genre *Pyrococcus* sp. ayant un poids 35 moléculaire compris entre environ 89 000 et 90 000 daltons.

Une première ADN polymérase purifiée thermostable selon l'invention provient de la souche d'archaebactéries du genre *Pyrococcus sp.* déposée à la Collection National de Culture de Micro-organismes (CNCM) à l'Institut Pasteur, le 3 Juillet 1996 sous le No I-1764. Cette ADN polymérase sera dénommée dans ce qui suit *Pyrococcus sp.* GE 23. Sa séquence de 771 acides aminés est représentée dans la liste de séquence en annexe sous le numéro SEQ ID NO:1. Un poids moléculaire de 89 409 daltons et un pI de 8,37 ont été déduits de cette séquence. L'invention concerne donc l'ADN polymérase de *Pyrococcus sp.* GE 23 dont la séquence en acides aminés est représentée dans la liste de séquence en annexe sous le numéro SEQ ID NO:1 ou tout autre séquence constituant un dérivé enzymatiquement équivalent de celle-ci. On entend par dérivés enzymatiquement équivalents, les polypeptides et protéines constitués par ou comprenant la séquence en acides aminés représentée dans la liste de séquences en annexe sous le numéro SEQ ID NO:1 dès lors qu'ils présentent les propriétés de l'ADN polymérase de *Pyrococcus sp.* GE 23. A ce titre l'invention envisage plus particulièrement une ADN polymérase dont la séquence en acides aminés est un fragment de celle représentée dans la liste de séquences en annexe sous le numéro SEQ ID NO:1 ou encore un assemblage de tels fragments..

On entend aussi par dérivés enzymatiquement équivalents, les séquences en acides aminés ci-dessus modifiées par insertion et/ou délétion et/ou substitution d'un ou plusieurs aminoacides, pour autant que les propriétés de l'ADN polymérase de *Pyrococcus sp.* GE 23 qui en résultent ne soient pas significativement modifiées.

35

Une seconde ADN polymérase purifiée thermostable selon l'invention provient de la souche d'archaebactéries du genre *Pyrococcus sp.* déposée à la CNCM le 20 Avril 1993 sous le No I-1302. Cette ADN polymérase sera dénommée dans ce qui suit *Pyrococcus sp.* GE 5 (alias *Pyrococcus abyssi*). Sa séquence de 771 acides aminés est représentée dans la liste de séquence en annexe sous le numéro SEQ ID NO:2. Un poids moléculaire de 89 443 daltons et un pI de 8,13 ont été déduits de cette séquence.

L'invention concerne donc l'ADN polymérase de *Pyrococcus sp.* GE 5 dont la séquence en acides aminés est représentée dans la liste de séquence en annexe sous le numéro SEQ ID NO:2 ou tout autre séquence constituant un dérivé enzymatiquement équivalent de celle-ci. On entend par dérivés enzymatiquement équivalents, les polypeptides et protéines constitués par ou comprenant la séquence en acides aminés représentée dans la liste de séquences en annexe sous le numéro SEQ ID NO:2 dès lors qu'ils présentent les propriétés de l'ADN polymérase de *Pyrococcus sp.* GE 5. A ce titre l'invention envisage plus particulièrement une ADN polymérase dont la séquence en acides aminés est un fragment de celle représentée dans la liste de séquences en annexe sous le numéro SEQ ID NO:2 ou encore un assemblage de tels fragments.

On entend aussi par dérivés enzymatiquement équivalents, les séquences en acides aminés ci-dessus modifiées par insertion et/ou délétion et/ou substitution d'un ou plusieurs aminoacides, pour autant que les propriétés de l'ADN polymérase de *Pyrococcus sp.* GE 5 qui en résultent ne soient pas significativement modifiées.

Les ADN polymérases *Pyrococcus sp.* GE 23 et de *Pyrococcus sp.* GE 5 se distinguent l'une de l'autre

par 10 résidus d'acides aminés dont les positions sont indiqués dans le tableau I ci-dessous.

Tableau I

| Positions | <i>Pyrococcus sp.</i> GE 23 | <i>Pyrococcus sp.</i> GE 5 |
|-----------|-----------------------------|----------------------------|
| 263       | Val                         | Ala                        |
| 277       | Ala                         | Thr                        |
| 281       | Ala                         | Val                        |
| 320       | Phe                         | Ser                        |
| 339       | Gln                         | His                        |
| 359       | Arg                         | Thr                        |
| 391       | Lys                         | Asn                        |
| 532       | Ser                         | Arg                        |
| 553       | Pro                         | His                        |
| 554       | Asn                         | Glu                        |

5 L'invention concerne également une séquence d'ADN constituée par ou comprenant la séquence codant pour une ADN polymérase purifiée thermostable de l'invention.

10 Une première séquence d'ADN selon l'invention comprend ou est constituée par les nucléotides 1547 à 3862 de la SEQ ID NO:1 codant pour les 771 acides aminés de l'ADN polymérase de *Pyrococcus sp.* GE 23.

15 Une seconde séquence d'ADN selon l'invention comprend ou est constituée par les nucléotides 678 à 2994 de la SEQ ID NO:2 codant pour les 771 acides aminés de l'ADN polymérase de *Pyrococcus sp.* GE 5.

20 L'invention concerne autant l'ADN polymérase thermostable définie précédemment isolée et purifiée d'une souche de *pyrococcus sp.* que l'ADN polymérase thermostable préparée par les méthodes du génie génétique. En conséquence, l'invention a aussi pour objet un vecteur comprenant une séquence d'ADN définie  
25 précédemment, ainsi qu'un procédé de production ou d'expression dans un hôte cellulaire des ADN polymérases thermostables de l'invention.

Un procédé de production d'une ADN polymérase thermostable conforme à l'invention consiste:

- à transférer une molécule d'acide nucléique codant pour une ADN polymérase thermostable ou un vecteur contenant ladite molécule dans un hôte cellulaire,

- à cultiver l'hôte cellulaire obtenu à l'étape précédente dans des conditions permettant la production de l'ADN polymérase,

- à isoler, par tous moyens appropriés ladite ADN polymérase.

L'hôte cellulaire mis en oeuvre dans les procédés précédents peut être choisi parmi les procaryotes ou les eucaryotes et notamment parmi les bactéries, les levures, les cellules de mammifères, de plantes ou d'insectes.

Le vecteur utilisé est choisi en fonction de l'hôte dans lequel il sera transféré, il peut s'agir de tout vecteur comme un plasmide.

Les ADN polymérases thermostables de l'invention sont utiles notamment dans des procédés d'amplification enzymatique de séquences d'acides nucléiques. En conséquence, l'invention a pour objet, de tels procédés mettant en oeuvre une ADN polymérase thermostable décrite précédemment, ainsi que les kits d'amplification comprenant, outre les réactifs généralement utilisés, une quantité adéquate de cette ADN polymérase.

D'autres avantages et caractéristiques de l'invention apparaîtront à la lecture des exemples qui suivent donnés à titre non limitatif et concernant le clonage, l'expression, la caractérisation et l'activité des ADN polymérases thermostables de l'invention.

I - Matériels et méthodes.1) Conditions de culture, plasmides et souches utilisées.

5 Les souches *Pyrococcus furiosus* (DSM 5262) et *Thermococcus litoralis* (DSM 5474) ont été obtenues auprès de la collection du Deutsche Sammlung von Mikroorganismen (DSM) Braunschweig - Stoeckheim, Allemagne. Les souches *Pyrococcus* sp. G 23 et G 5 ont  
10 été isolées de cheminées de sources hydrothermales profondes découvertes lors de la campagne franco-japonaise STARMER effectuée en 1989 à 2000 mètres de profondeur dans le bassin Nord-Fiji.

15 *Pyrococcus* sp. G 5 a été cultivée anaerobiquement dans un fermenteur de 8 litres dans un milieu BHI (DIFCO) supplémenté avec du soufre à un pH de 6,5 et à 90°C, comme décrit dans la littérature (Erauso, G., A. L. Reysenbach, A. Godfroy, J.-R. Meunier, B. Crump, F. Partensky, J. A. Baross, V. Marteinsson, G.  
20 Barbier, N. R. Pace, and D. Prieur. 1993. *Archives of Microbiology*.160:338-349).

*Pyrococcus* sp. G 23 a été cultivée à 85°C dans un fermenteur identique dans un milieu 2216S (DIFCO) à un pH de 6,5.

25 La souche *E. coli* SURE, XL1-Blue (Stratagene, La Jolla, Calif. USA) a été utilisée comme hôte pour les plasmides recombinants dérivés de pUC18 et pBluescript. Les souches *E. coli* NOVABLUE, BL21(DE3) et BL21(DE3)pLysS (Novagen, Madison, Wi. USA) ont été  
30 utilisées comme hôte pour les plasmides recombinants dérivés. *E. coli* SURE, XL1-Blue, NOVABLUE, BL21(DE3) et BL21(DE3)pLysS ont été cultivées dans un milieu LB (DIFCO) ou LB supplémenté avec des antibiotiques appropriés (DIFCO).

35

2) Isolement de l'ADN, hybridation et ADN recombinants.

L'ADN de haut poids moléculaire de *Pyrococcus* sp. G 23 et G 5 a été isolé par la méthode de Charbonnier modifiée (Charbonnier, F., G. Erauso, T. Barbeyron, D. Prieur, and P. Forterre. 1992. *Journal of Bacteriology*. 174, 19 :6103-6109). Les cellules centrifugées ont été resuspendues dans un tampon TE-Na-1x, puis lysées à 40°C pendant 3 heures avec un mélange de N-lauryl sarcosine 1 %, sulfate de sodium de dodecyl 1% et 0,4 mg/ml de protéinase K. Après centrifugation à 5000 g pendant 10 minutes, l'ADN est extrait par PCI (25-24-1), puis traité par la RNAase à 5 µg/ml à 60°C pendant une heure. Ces étapes sont suivies par une extraction additionnelle par PCI et une extraction au chloroforme. L'ADN est précipité avec de l'éthanol 100 % et les pastilles sont séchées et suspendues dans du TE-1x. La concentration et la pureté de l'ADN ont été estimées par spectrophotométrie à 230, 260 et 280 nm avec un appareil GeneQuantII (Pharmacia, Upsalla, SWeden). Pour la construction de la mini-librairie pUC18 (Sutherland, K. J., C. M. Henneke, P. Towner, and D. W. Hough. 1990. *European Journal of Biochemistry*. 194:839-844) de *Pyrococcus* sp. G 23, l'ADN génomique a été digéré par une série d'enzymes de restriction (BamHI, BgII, EcoRI, EcoRV, HindIII, PvuII, SalI, SacI, XbaI et XhoI) par simple ou double digestion. Puis les fragments d'ADN ont subi une migration sur gel d'agarose 0,8 % dans du TBE-1x et ont été transférés sur une membrane de nylon Hybond-N+ (Amersham, UK) et hybridés avec des sondes d'ADN préparées par PCR avec des amorces spécifiques sélectionnées à partir des gènes d'ADN polymérase de *P. furiosus* et *T. litoralis*, marqués au <sup>32</sup>P par random-priming conformément aux recommandations du fabricant (Magaprime, Amersham, UK). Deux sondes de *P. Furiosus* ont été utilisées, pFuΣ et PfuF, couvrant

respectivement les régions délimitées par les paires de base 8 à 2316 et 819 à 1915 de la section codante du gène de la polymérase *Pfu*, comme défini par Uemori et al. (Uemori, T., Y. Ishino, H. Toh, K. Asada, and I. Kato. 1993. *Nucleic Acids Research*. 21(2):259-265). Deux sondes de *T. litoralis*, *TliI* et *TliT*, couvrant respectivement les régions délimitées par les paires de base 297 à 1768 et 4631 à 5378, comme défini par Hodges et al. (Hodges, R. A., F. B. Perler, C. J. Noren, and W. E. Jack. 1992. *Nucleic Acids Res.* 20(23):6153-6157. Les fragments positifs, identifiés par hybridation ADN-ADN (Southern, E. M. 1975. *Journal of molecular biology*. 98:503), ont été ensuite préparés par des digestions appropriées de 100 µg d'ADN génomique, purifiés dans des sacs de dialyse à partir des gels d'agarose, et précipités à l'éthanol 100 %. Les fragments ont été ligaturés dans pUC18, lequel avait été coupé par les enzymes de restriction appropriés pour une digestion simple et déphosphorylé. La transformation des souches hôtes a été réalisée par électroporation (Gene Pulser, Biorad). Le criblage des clones recombinants a été effectué par hybridation de colonie selon les techniques décrites dans la littérature (Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor). La température des Southern blots et d'hybration de colonie était de 55°C dans un tampon standard sans formamide. L'ADN de plasmide a été isolé par la méthode de Birnboim et Doly (Birnboim, H. C., and J. Doly. 1979. *Nucleic Acids Research*. 7:1503), puis purifié par chromatographie échangeuse d'anions en phase solide (Quiagen, Chatsworth, Calif. USA). Les fragments de restriction des plasmides dérivés de pUC ont été purifiés sur gels d'agarose par la méthode de GeneClean (Bio 101, La Jolla, Calif.) pour un clonage ultérieur. Les réaction de polymérisation en chaîne

(PCR) de longs fragments ont été effectuées conformément à la procédure définie par Barnes (Barnes, W. M. 1994. *Proc. Natl. Acad. Sci. USA.* 91:2216-2220) avec un mélange réactionnel Taq Extender (Stratagene, La Jolla, Calif.).

- 5 Les ADNr de 16S et 23S de *Pyrococcus* sp. GE 23 ont été amplifiés par PCR avec un thermocycleur Stratagene 96-gradient en utilisant les conditions suivantes :
- amorce directe Aa (5'-TCCGGTTGATCCTGCCGGA-3'),
  - amorce indirecte 23Sa (5'-CTTTCGGTCGCCCCTACT-3'),
  - 10 - étape initiale 3 minutes à 94°C suivi de 30 cycles (94°C, 1 mn/49°C, 1mn/72°C, 2mn) et,
  - élongation finale de 5 mn à 72°C.

Les produits de PCR ont été clonés dans le vecteur pBluescript pour séquençage.

15 Le gène de l'ADN polymérase de *Pyrococcus* sp. GE 5 a été isolé par PCR en utilisant les amorces mises au point pour exprimer le gène de l'ADN polymérase de *Pyrococcus* sp. GE 23 :

- GE23DIR (5'-TGGGGCATATGATAATCGATGCTGATTAC-3')
- 20 - GE23REV (5'-GACATCGTCGACTCTAGAACTTAAGCCATGGTCCG-3').

### 3) Séquençage des ADN.

Les séquences d'ADN ont été obtenues par la méthode de terminaison de chaîne didéoxy (Sanger, F., S. Nicklen, and A. R. Coulson. 1977. *Proc. Natl. Acad. Sci. USA.* 74:5467-5473) en utilisant un système d'analyse automatique d'ADN Applied Biosystems. Les deux brins des gènes codant l'ADN polymérase de *Pyrococcus* sp. G23 et G5 ont été séquencés alors que l'ADNr 16S de *Pyrococcus* sp. G23 a été séquencé sur un seul brin, en utilisant des amorces universelles localisées sur des vecteurs et des amorces internes.

L'analyse de séquence a été réalisée avec un logiciel DNASTAR (Madison, Wis., USA) et le programme de Genetics Computer Group (University of Wisconsin Biotechnology Center, Madison, Wis., USA) accessible en

ligne sur INFOBIOGEN. Les recherches informatisées de similitudes ont été réalisées avec le programme BLAST, les alignements multiples avec CLUSTAL V, et les arbres phylogénétiques ont été établis selon la méthode dite  
5 "Neighbour-joining (Saitou, N., and M. Nei. 1987. *Mol. Biol. Evol.* 4(4):406-425).

4) Construction des recombinants exprimant l'ADN polymérase de *Pyrococcus sp.* G 23 et G5.

10 Les ADN polymérases de *Pyrococcus sp.* G 23 et G 5 ont été exprimées dans *E. coli* en utilisant le système d'expression pET12, appartenant au système d'expression T7 (Studier, F. W., A. H. Rosenberg, F. J. Dunn, and J. W. Dubendorff. 1990. *Methods Enzymol.*  
15 185:60-89) acquis auprès de Novagen (Madison, Wi; USA). La PCR a été utilisée pour préparer les fragments de *Pyrococcus sp.* G 23 et G 5 contenant les sites de restriction NdeI et SalI compatibles avec l'extrémité des amorces GE23DIR et GE23REV. Le mélange de PCR  
20 contenait l'ADN polymérase GOLDSTAR (Eurogentec, B), l'enzyme Taq extender Taq (*Pfu* de Stratagene), le tampon d'extension avec les quatre dNTP (chacun à 0,2 mM) et les amorces GE23DIR et GE23REV à 50 pmoles dans un volume final de 50 µl. L'amplification a été effectuée  
25 sur 20 cycles : 1mn à 90°C, 1mn à 50°C et 3 mn à 72°C, en utilisant un thermocycleur Stratagene 96-gradient. Les fragments de PCR digérés par NdeI et SalI ont été ligaturés aux sites NdeI et SalI de pET12a digérés avec les mêmes enzymes, réétablissant ainsi le codon  
30 d'initiation. Les constructions obtenues ont été nommées respectivement pETPAB1 et pETGE1. Ces constructions ont été séquencées toutes les deux aux sites de jonction et entièrement pour pETPAB1. Les tests d'expression ont été réalisés conformément aux recommandations du fabricant  
35 du système d'expression pET : sélection des clones dans

*E. coli* Novablue, culture of *E. coli* BL21(DE3), induction (IPTG 0,2-1mM).

750 µl de cellules de cultures induites à DO<sub>600</sub>=1 et non induites ont été centrifuguées, resuspendues dans un tampon de lyse B (Tris-HCl 10 mM pH 7,5 ; NaCl 10 mM ; MgCl<sub>2</sub> 2 mM ; Triton X-100 1% v/v). L'ADN polymérase a été observée par SDS-PAGE après une dénaturation thermique partielle des protéine hôtes (71°C, 20 mn) et concentrée sur filtres Millipore (Ultrafree MC). L'activité a été testée par mesure de l'incorporation de <sup>3</sup>H-dTTP (Amersham, UK) en utilisant de l'ADN de thymus de veau activé comme substrat (Sigma-Aldrich, F) dans le milieu de réaction C (Tris-HCl 50 mM pH 8,8 ; DTT 1mM ; MgCl<sub>2</sub> 10mM ; KCl 10 mM ; BSA 0,4 mg/ml, chaque dNTP à 0,4 mM).

## II - Résultats.

### 1) Isolement du gène de l'ADN polymérase de *Pyrococcus* sp. GE 23.

Les ADN génomiques des souches *Pyrococcus* sp. GE 23 et GE 5, digérés par une série d'enzymes de restriction, ont été hybridés à l'ADN de *P. furiosus* et à des sondes de *T. litoralis* préparées par PCR. Comme montré aux figures 1 et 2, l'analyse en Southern blots a révélé un fragment positif HindIII-HindIII de 2,9 kb avec une sonde *TliI* marquée au <sup>32</sup>P (figure 2), alors que la sonde *PfuF* a révélé un fragment positif XbaI-XbaI de 8 kb (figure 1) (exposition de 24 heures). Le fragment de 2,9 kb a été récupéré et purifié comme décrit précédemment dans Matériels et Méthodes et cloné dans le vecteur pUC18 digéré par HindIII. Environ 600 recombinants (*E. coli* SURE) ont été criblés avec une sonde *TliI* radiomarquée et 6 d'entre-eux ont donné un signal positif d'hybridation. Des profils de restriction identiques ont été obtenus pour ces clones indiquant

qu'ils contiennent probablement le même insert. Le séquençage ultérieur de l'un de ces clones (désigné pGE23a) et la comparaison de séquence (FASTA) ont démontré qu'il correspond aux 1358 premières paires de base du gène de l'ADN polymérase appartenant à la famille B (Braithwaite, D. K., and J. Ito. 1993. *Nucleic Acids Research*.21(4):787-802) et qu'il se termine au site HindIII. L'analyse de la séquence nucléotidique montre que le site XbaI est localisé en amont du codon d'initiation et qu'en conséquence le fragment positif XbaI-XbaI de 8 kb devrait contenir la partie 3' du gène. Selon la même méthode, ce fragment de 8 kb a été cloné dans le vecteur pUC18 et 3 clones recombinants positifs parmi 800 ont été identifiés par hybridation de colonie comme précédemment en utilisant une sonde homologe préparée à partir de la partie 5' du gène de l'ADN polymérase de *Pyrococcus sp.* GE 23. Les cultures dans un milieu liquide (LB-ampicillin) de ces clones ont été systématiquement lysées bien avant  $DO_{600}=0,5$ . Une culture a été arrêtée à  $DO_{600}=0,1$  et l'ADN plasmidique (désigné pGE23b) a été extrait. La PCR longue distance (avec le Taq Extender) a été effectuée sur cet ADN plasmidique en utilisant une amorce directe localisée sur la partie 5' de la séquence obtenue précédemment et une amorce réverse localisée en aval sur le vecteur pUC18, pour obtenir un produit d'amplification de 8 kb. La partie 3' a été séquencée ultérieurement. Puis, une nouvelle PCR longue distance a été effectuée en utilisant la même amorce directe localisée sur la partie 5' du gène et une amorce réverse localisée sur la partie 3' disponible. Les produits d'amplification de 10 réactions de PCR séparées ont été rassemblés, purifiés par électrophorèse sur gel d'agarose et séquencés sur les deux brins en utilisant des amorces internes. Ces produits contiennent la partie 3' du gène de l'ADN polymérase de *Pyrococcus sp.* G23.

2) Position phylogénétique de *Pyrococcus* sp. G23 et GE5.

*Pyrococcus* sp. G23 et G5 appartiennent au même groupe de PCR-RFLP d'ADNr 16S (Meunier, J. R. 1994. Biodiversité et systématique des populations de microorganismes hermophiles isolés d'écosystèmes hydrothermaux océaniques abyssaux. Thèse, Paris 7) indiquant qu'il pourrait s'agir de deux différentes souches de la même espèce. Toutefois, comme montré sur la figure 3, les profils RFLP différents des gènes de l'ADN polymérase de *Pyrococcus* sp. G23 et G5, obtenus avec les différentes sondes (*Pfu*Σ, *Pfu*P, *Tli*I) indiquent qu'au moins les gènes de l'ADN polymérase présentent des différences significatives. Des hybridations ADN-ADN en slot blots (Marteinsson, V. T., L. Watrin, D. Prieur, J. C. Caprais, G. Raguènes, and G. Erauso. 1995. *International Journal of Systematic Bacteriology*. 45(4):623-632) ont conduits à des niveaux contradictoires de relation : 79 % entre GE 23 non marqué et GE 5 marqué, et 98 % entre GE 5 non marqué et GE 23 marqué. L'amplification des gènes des ADNr 16S-23S a été effectuée comme décrit précédemment dans Matériels et Méthodes. La bande amplifiée de 1,9 kb a été purifiée et séquencée sur un brin. La séquence complète de l'ADNr 16S a été comparée avec sa contrepartie de *Pyrococcus* sp. GE 5 et d'autres espèces de *Pyrococcus* sp. et *Thermococcus*. Le niveau de similitude entre *Pyrococcus* sp. GE 23 et *Pyrococcus* sp. GE 5 est de 97,8 % et *Pyrococcus* sp. GE 23 a été regroupé avec *Pyrococcus* sp. GE 5 (*Pyrococcus abyssi*) dans l'arbre phylogénétique des ADN polymérases de *Thermococcales* de la figure 4. De même la figure 5 représente la situation phylogénétique de *Pyrococcus* sp. GE23 et *Pyrococcus* sp. GE 5 (*Pyrococcus abyssi*) sur la base des séquences d'ADNr 16S. Dans ces deux figures, le dendogramme a été établi

par la méthode "Neighbour-joining" et l'échelle représente la distance relative entre les séquences.

5                   3) Identification du gène de l'ADN polymérase de *Pyrococcus sp.* GE 5.

Déduisant de la relation entre ces deux souches que l'isolement direct du gène de l'ADN polymérase de *Pyrococcus sp.* GE 5 était possible par PCR, les deux amorces GE23DIR et GE23REV, préparées pour  
10 cloner le gène de l'ADN polymérase de *Pyrococcus sp.* GE 23 dans le vecteur pET12, ont été utilisées dans une réaction de PCR contenant l'ADN génomique de *Pyrococcus sp.* GE 5. Les produits d'amplification de 5 réactions de PCR ont été rassemblés, purifiés et séquencés sur les  
15 deux brins. La séquence nucléotidique présente une forte homologie (97 %) avec le gène de l'ADN polymérase de *Pyrococcus sp.* GE 23. Afin d'obtenir la séquence du gène l'ADN polymérase de *Pyrococcus sp.* GE 5, différents essais de clonage ont été réalisés. Tout d'abord, un  
20 fragment HindIII-HindIII de 2,9 kb, identifié par hybridation ADN-ADN, comme montré à la figure 3, a été cloné dans le vecteur pUC18, transformé dans *E. coli* SURE et séquencé. Il se présente comme homologue au fragment HindIII-HindIII de 2,9 kb de *Pyrococcus sp.* GE  
25 23 cloné précédemment et contenant la partie 5' du gène cible. Deuxièmement, un fragment positif XbaI-XbaI de 2,6 kb, contenant potentiellement la totalité de la portion codante, a été identifié par hybridation ADN-ADN en utilisant le gène de l'ADN polymérase de *Pyrococcus sp.* GE 5 produit par PCR comme sonde radio-marquée. Ce  
30 fragment a été cloné dans les vecteurs pUC18, pBluescript et pET12 et transformé dans *E. coli* Novablue (souche recA<sup>-</sup>). Trois clones recombinants positifs parmi 600 ont été obtenus avec pET12 et aucun avec pUC18 et  
35 pBluescript. Les profils de restriction des clones recombinants pET12 démontrent que l'intégrité de la

construction n'a pas été préservée. La figure 3 montre que la majeure partie de l'insert a été déléetée dans le procédé et seul subsiste le fragment 3' de 342 pb HindIII-XbaI. La séquence définitive du gène de l'ADN polymérase de *Pyrococcus* sp. GE 5 est composée d'une partie 5' génomique HindIII-HindIII, d'une portion interne HindIII-HindIII de 611 pb issue de produits de PCR et de la partie 3' de 342 pb génomique.

4) Séquences nucléotidique et polypeptdique des ADN polymérase de *Pyrococcus* sp. GE 23 et GE 5.

a) ADN polymérase de *Pyrococcus* sp. GE 23.

Le fragment de *Pyrococcus* sp. GE 23 HindIII-HindIII de 2,9 Kb qui a été cloné ainsi qu'un fragment contigu dans le sens 3' de 1,6 Kb obtenu par PCR longue distance ont été séquencés sur les deux brins et assemblés. La séquence de 4447 pb obtenue a été étudiée pour déterminer les régions susceptibles d'être traduites. Deux cadres de lecture ouverts ont été mis en évidence. Le premier, dans le cadre 2, s'étendant de la paire de bases 1547 à la paire de bases 3862 correspond au gène de l'ADN polymérase codant pour une protéine de 771 acides aminés, dont le poids moléculaire déduit de la séquence est de 89 409 Daltons et le point isoélectrique théorique de 8,37. Ce poids moléculaire correspond au poids moléculaire apparent estimé par SDS-PAGE de l'ADN polymérase recombinante. Le second ORF a été localisé sur le cadre 4 entre les paires de bases 1439 et 627.

La séquence déduite a une longueur de 270 aminés et les recherches de similitude effectuées avec les programmes BLAST et FASTA n'ont révélé aucune homologie significative avec les séquences disponibles dans les bases de données Swiss-Prot et PIR.

b) ADN polymérase de *Pyrococcus* sp. GE 5.

La séquence codante complète du gène de l'ADN polymérase de *Pyrococcus* sp. GE 5 produit par PCR en utilisant les amorces préparées à partir du gène de *Pyrococcus* sp. GE 23 a été obtenue par séquençage des deux brins. Afin de confirmer la séquence de ce produit de PCR, le fragment de *Pyrococcus* sp. GE 5 HindIII-HindIII de 2,9 Kb qui a été cloné a également été séquencé sur les deux brins, et les mêmes résultats ont été obtenus. La région 3' du gène a également été séquencée à partir de l'insert HindIII-XbaI restant de 342 pb résultant de la délétion d'une partie du fragment XbaI-XbaI de 2,6 kb. Seule la région interne HindIII-HindIII de 611 pb de la séquence du gène obtenue à partir des produits de PCR n'a pas été confirmée au niveau génomique. Un cadre de lecture ouvert entre les paires de bases 679 et 2994 a été identifié dont la traduction produit une séquence polypeptidique de 771 acides aminés, longueur identique à celle de l'ADN polymérase de *Pyrococcus* sp. GE 23.

c) Comparaison des séquences des deux polymérases.

L'alignement des parties codantes des deux gènes montre leur forte parenté, puisqu'ils ne diffèrent que de 64 nucléotides dispersés le long des séquences à deux exceptions près. On observe tout d'abord 14 substitutions entre les paires de bases 1541 et 1603, faisant de cette région, la plus variable de la toute séquence, puis on constate un nombre limité de substitutions dans la région 3' entre les paires de bases 1890 et 2316. La plupart de ces substitutions sont sans conséquence au niveau des séquences polypeptidiques, notamment les 23 premières substitutions sont sans effet sur la composition des protéines. La comparaison avec d'autres gènes d'ADN polymérases de Thermococcales

disponibles dans les bases de données révèle que la région hautement variable entre les paires de bases 1541 et 1603 de *Pyrococcus* sp. GE 23 et GE 5 semble être une caractéristique de ces deux souches, la diversité étant bien distribuée sur les autres séquences de gènes codant pour des ADN polymérases. La comparaison au niveau des séquences polypeptidiques montre la forte similitude des deux ADN polymérases, 98 % d'homologie avec la méthode CLUSTAL V, et seulement 10 résidus différents. Les substitutions d'acides aminés sont donnés dans le tableau I et les alignements avec les ADN polymérases d'autres Thermococcales sont représentés aux figures 6 et 7.

La figure 6 représente l'alignement de la séquence polypeptidique de l'ADN polymérase de *Pyrococcus* sp. GE 23 avec celle des ADN polymérases de : *Deep Vent* (intein-), *Pyrococcus* sp. KOD1, *Pyrococcus furiosus*, *Vent* (Tli intein-), *Pyrococcus* sp. 9°N.

La figure 7 représente l'alignement de la séquence polypeptidique de l'ADN polymérase de *Pyrococcus* sp. GE 5 avec celle des ADN polymérases de : *Pyrococcus* sp. GB-D, *Pyrococcus* sp. KOD1, *Pyrococcus furiosus*, *Thermococcus litoralis*, *Pyrococcus* sp. 9°N.

Le tableau II ci-dessous indique les pourcentage de similarités entre les séquences polypeptidiques de : (1) *Pyrococcus* sp. GE 23, (2) *Pyrococcus* sp. GE 5, (3) *Pyrococcus* sp. GB-D, (4) *Pyrococcus* sp. KOD1, (5) *Pyrococcus furiosus*, (6) *Thermococcus litoralis*, (7) *Pyrococcus* sp. 9°N.

Tableau II

| (1) | (2)  | (3)  | (4)  | (5)  | (6)  | (7)  |     |
|-----|------|------|------|------|------|------|-----|
| *** | 98,6 | 89,2 | 81,3 | 83,5 | 76,3 | 81,5 | (1) |
|     | ***  | 88,4 | 80,6 | 82,8 | 75,5 | 80,9 | (2) |
|     |      | ***  | 81,6 | 85,1 | 77,1 | 83,0 | (3) |
|     |      |      | ***  | 79,2 | 78,1 | 90,4 | (4) |
|     |      |      |      | ***  | 74,0 | 80,1 | (5) |
|     |      |      |      |      | ***  | 77,0 | (6) |
|     |      |      |      |      |      | ***  | (7) |

5 La plupart des substitutions d'acides aminés  
entre les ADN polymérase de *Pyrococcus sp.* GE 23 et GE  
5 sont non conservatives, mais aucune n'est localisée  
dans les motifs connus pour être impliqués dans l'action  
catalytique de l'enzyme, ni dans les domaines 3'-5'  
exonucléase ou polymérase. Les deux gènes ne possèdent  
10 aucune séquence intronique (IVS ou intéines)  
contrairement aux gènes d'ADN polymérase de *T.*  
*litoralis*, *Thermococcus sp.* KOD1, *Pyrococcus sp.* GB-D,  
et présentent la même organisation que les gènes d'ADN  
polymérase de *P. furiosus* et *Pyrococcus sp.* 9°N.

15

5) Expression, caractérisation et activité  
de l'ADN polymérase de *Pyrococcus sp.* GE 5.

a) Clonage et expression.

20

Un insert de 2320 pb couvrant la séquence  
d'ADN de 2316 pb de SEQ ID NO:I codant pour l'ADN  
polymérase de *Pyrococcus sp.* GE 5 a été cloné aux sites  
NdeI et BamHI d'un vecteur pour transformer la souche  
*E. coli* BL21(DE3). Des mini-préparations d'ADN  
25 plasmidien ont été réalisées sur une vingtaine de clones  
transformants et un seul a été sélectionné sur la base  
de la taille des fragments d'ADN libérés après digestion  
par les enzymes de restriction NdeI et BamHI.

Des essais d'expression ont ensuite été réalisés à 37°C, en erlenmeyer, avec le clone sélectionné. L'expression est induite en phase exponentielle de croissance (DO 600nm = 0,6 - 0,7) avec  
5 les concerntations en IPTG de 0 ; 0,5 ; 1 et 1, 5 mM. Des échantillons sont prélevés à différents moments au cours de la culture et les protéines analysés par électrophorèse sur gel de polyacrylamide et coloration au bleu de Coomassie. La figure 8 représente les  
10 résultats des essais d'expression, dans lesquels :

- Le puits No. 1 correspond à la protéine Ladder de 10 kDA (Gibco BRL).
- Le puits No. 2 correspond à l'échantillon induit sans IPTG, temps 0 (phase exponentielle de  
15 croissance).
- Le puits No. 3 correspond à l'échantillon induit sans IPTG, 1 heure après l'induction.
- Le puits No. 4 correspond à l'échantillon induit sans IPTG, 4 heure après l'induction.
- Le puits No. 5 correspond à l'échantillon induit sans IPTG, 18 heure après l'induction.  
20
- Le puits No. 6 correspond à la protéine Ladder de 10 kDA (Gibco PRL).
- Le puits No. 7 correspond à l'échantillon induit avec 1 mM d'IPTG, temps 0 (phase exponentielle de  
25 croissance).
- Le puits No. 8 correspond à l'échantillon induit avec 1 mM d'IPTG, 1 heure après l'induction.
- Le puits No. 9 correspond à l'échantillon induit avec 1 mM d'IPTG, 4 heure après l'induction.  
30
- Le puits No. 10 correspond à l'échantillon induit avec 1 mM d'IPTG, 18 heure après l'induction.

On observe que seules les cellules cultivées en l'absence d'IPTG expriment l'ADN polymérase et le  
35 taux est maximum après une nuit de culture. Le poids moléculaire de la protéine exprimée est estimé à 89 Kda.

b) Fermentation et extraction des cellules.

La culture de la souche *Pyrococcus sp.* GE 5 a été réalisée selon un protocole standard. On a ajouté le facteur de sélection du plasmide au milieu R12 choisi pour la préculture et la culture. Le transfert de la préculture vers le fermenteur a été effectué lorsque la densité optique à 600 nm de la préculture était d'environ 0,8. Le fermenteur NBS MICROS (25 litres) a été employé pour cette production, avec un volume de culture de 24 litres. Les conditions de fermentation ont été les suivantes : température = 37°C ; agitation = 300 rpm ; aération = 30 l/m ; oxygène dissous = 15 %. Le pH est régulé à 6,8 avec NaOH pendant la phase d'acidification. Lors de la phase de basification, le pH est demeuré en libre évolution. Les bactéries ont été récoltées près 16 à 17 heures de culture. La densité optique finale était d'environ 10 unités et le pH final de 8. La culture a ensuite été concentrée par filtration sur fibres creuses (AMICON), jusqu'à obtenir un volume final de 2 litres. Ce concentré de bactéries a alors été utilisé pour la purification des protéines. Après concentration cellulaire, les cellules sont broyées en continu au moyen d'un broyeur à billes et le broyat cellulaire est additionné de PMSF 1 mM.

c) Purification de l'ADN polymérase de *Pyrococcus sp.* GE 5.

Après centrifugation du broyat cellulaire (15 minutes à 10000 g), le surnageant (fraction soluble) est récupéré et chauffé 15 minutes à 75°C. Le précipité formé après dénaturation thermique est éliminé par centrifugation (15 minutes à 10000 g). Le surnageant est chargé sur une colonne d'échange d'anions :

- Matrice Q Sépharose Fast Flow dans une colonne XK 16/30 (Pharmacia).

- Système chromatographique : Bio Pilot (Pharmacia).
- Tampon A : Tris-HCl 50 mM, pH 8 et tampon B : Tris-HCl 50 mM, pH 8 + NaCl 0,5 M.
- 5 - Débit : 10 ml/mn ; gradient 0-50% B en 60 minutes.
- Fractions : 10 ml/tube.

10 La figure 9 représente le profil d'éluion de l'échantillon constitué de 450 ml de protéines solubles après dénaturation thermique à 75°C.

La figure 10 représente l'analyse par SDS-PAGE (Phast System, Pharmacia), en présence de  $\beta$ -mercaptoéthanol, de différentes fractions issues de la purification :

- 15 - Gels : Phast gel 8-15% (Phast System, Pharmacia).
- Lignes 1 et 8 : marqueur de poids moléculaire (10 Kda, Bio-Rad).
- Ligne 2 : Protéines totales.
- 20 - Ligne 3 : protéines solubles après broyage.
- Ligne 4 : protéines solubles après dénaturation thermique.
- Lignes 5, 6 et 7 : fractions 1, 2 et 3
- 25 après chromatographie sur QFF de la figure 2.

Comme illustré par la figure 10, l'ADN polymérase de *Pyrococcus* sp. GE 5 est soluble et thermostable. Elle est éluée vers 0,15 M NaCl lors de la chromatographie d'échange d'ions. Les fractions

30 contenant l'ADN polymérase de *Pyrococcus* sp. GE 5 ont été rassemblées (fraction 2; 50 ml) et chromatographiées sur hydroxyapatite (type II, Biorad) :

- Matrice : hydroxyapatite type II dans une colonne 10/5 (Biorad).
- 35 - Système chromatographique : FPLC (Pharmacia).

- Tampon A: Phosphate 50 mM pH 8; tampon B : Phosphate 500 mM pH 8.

- Débit : 10 ml/min; gradient 0-100% B en 30 minutes.

5                    La figure 11 représente le profil d'élu-  
de l'échantillon constitué de 50 ml de la fraction 2  
après chromatographie d'échange d'ions. Les différentes  
fractions ont été analysées par spectrométrie UV et par  
SDS-PAGE. La fraction non absorbée contient  
10                   principalement des acides nucléiques tandis que les  
fractions 1 à 3 contiennent l'ADN polymérase de  
*Pyrococcus sp.* GE 5. Après chromatographie, 1 mM de  
mercaptoéthanol, 0,1 % Tween 20 et 50 % glycérol ont été  
ajoutés avant conservation à -20°C.

15

d) Caractérisation des fractions purifiées.

Les 3 fractions obtenues après chromatographie sur hydroxyapatite ont été analysées par SDS-PAGE en présence de 2-mercaptoéthanol (figure 12) :

20

- Gels : gel 15 % (Laemli, Babygel).  
- Lignes 1 et 5: Marqueur de poids moléculaire (Bio-Rad, Broad Range).

25

- Ligne 2 : fraction 1 (5 µl).
- Ligne 3 : fraction 2 (5 µl).
- Ligne 4 : fraction 3 (5 µl).
- Ligne 6 : fraction 1 (10 µl).
- Ligne 7 : fraction 2 (10 µl).
- Ligne 8 : fraction 3 (10 µl).

30

Comme illustré par la figure 12 et à la figure 13, la fraction 1 contient de l'ADN polymérase de *Pyrococcus sp.* GE 5 homogène en SDS-PAGE, alors que les fractions 2 et 3 sont contaminées par 2 protéines de 35 Kda et 65 Kda.

35

e) Activité des fractions purifiées.

L'activité d'ADN polymérase des différentes fractions a été testée dans une réaction d'amplification de type PCR. Un fragment de 1300 pb environ a été amplifié à partir d'un ADN cible et d'amorces spécifiques. Le tampon utilisé est composé de :

- Tris-HCl : 75 mM pH 9,0.
- $(\text{NH}_4)_2\text{SO}_4$  : 20 mM.
- 0,01 % (w/v) Tween 20.
- $\text{MgCl}_2$  : 1,5 mM.

Trente cinq cycles ont été réalisés, chacun comprenant une étape de dénaturation de 1 minute à 94°C, une étape d'appariement des oligonucléotides de 1 minute à la température appropriée et une étape d'élongation de 3 minutes à 72°C. La figure 13 rapporte les résultats obtenus avec un volume réactionnel de 100  $\mu\text{l}$  pour des quantités d'ADN polymérase de *Pyrococcus sp.* GE 5 de 2,5  $\mu\text{g}$ , 1  $\mu\text{g}$ , 0,1  $\mu\text{g}$  et 0,01  $\mu\text{g}$  :

- Puits No. 1 : 1 kb DNA ladder (Gibco BRL).
- Puits No. 2 : fraction 1 (2,5  $\mu\text{g}$ ).
- Puits No. 3 : fraction 1 (1  $\mu\text{g}$ ).
- Puits No. 4 : fraction 1 (0,1  $\mu\text{g}$ ).
- Puits No. 5 : fraction 1 (0,01  $\mu\text{g}$ ).
- Puits No. 6 : fraction 2 (2,5  $\mu\text{g}$ ).
- Puits No. 7 : fraction 2 (1  $\mu\text{g}$ ).
- Puits No. 8 : fraction 2 (0,1  $\mu\text{g}$ ).
- Puits No. 9 : fraction 2 (0,01  $\mu\text{g}$ ).
- Puits No. 10 : fraction 3 (2,5  $\mu\text{g}$ ).
- Puits No. 11 : fraction 3 (1  $\mu\text{g}$ ).
- Puits No. 12 : fraction 3 (0,1  $\mu\text{g}$ ).
- Puits No. 13 : fraction 3 (0,01  $\mu\text{g}$ ).
- Puits No. 14 : 1 kb DNA ladder (Gibco BRL).
- Puits No. 15 : Contrôle positif avec la Taq polymérase.

6) Conditions d'optimisation de la PCR pour les polymérases de l'invention.

5 Afin de simplifier l'exposé des résultats ci-après et des figures qui s'y rapportent, les ADN polymérases thermostables de *Pyrococcus abyssi* (*Pyrococcus* sp. GE 5) et *Pyrococcus* sp. GE 23 seront désignées respectivement ci-après Pab et Ppr.

10 Des essais de PCR ont été effectués dans les conditions suivantes :

- 10 mM Tris HCl, pH 9,0
- 50 mM KCl
- 3 mM MgSO<sub>4</sub>
- 0,1 % Tween
- 15 - 0,312 mM de chaque dNTP
- 50 picomoles de chaque oligonucléotide de formule :

5' TCACCTTAGGGTTGCCCATAA3'

5' TGGGCATAAAAGTCAGGGCAG3'

- 20 - 1 U. de Pab (figure 14) ou de Ppr (figure 15)

- Quantité croissante d'ADN génomique humain de  $\beta$ -globine (de 0,5 ng à 3  $\mu$ g) ; volume réactionnel : 50  $\mu$ l.

25 Le programme de PCR suivant a été utilisé :

- 5 minutes à 93°C
- 37 fois (1 minute à 62°C ; 2 minutes à 72°C ; 1 minute à 91°C)
- 1 minute à 62°C
- 30 - 10 minutes à 72°C.

Les figures 14 et 15 représentent les produits de PCR déposés sur gel d'agarose 2% en présence du marqueur pBR digéré par HaeIII.

35 Les amorces choisies permettent d'amplifier un fragment de 420 pb. Pab et Ppr permettent d'obtenir un produit de PCR de 420 pb quelque soit la quantité de

matrice comprise entre 0,5 ng et 1 µg. Un produit de PCR a également été obtenu dans ces mêmes conditions en présence de 10 picogrammes de matrice d'ADN.

5                                   7) Aptitude à l'amplification de la Pab et de la Ppr à hautes températures.

10                                   Les tests d'activité des polymérases recombinantes de *Pyrococcus abyssi* (Pab) et de *Pyrococcus* sp. GE23 (Ppr) réalisés avec la méthode d'incorporation de dNTP marqués ont démontré, dans les conditions de tampons retenues, une activité résiduelle jusqu'à 85°C. Au delà de cette température, une possible dégradation du substrat pourrait masquer l'activité des polymérases de l'invention. Il est cependant intéressant d'apprécier si cette aptitude peut être mise à profit pour de l'amplification de gène *in vitro*, sachant que certaines applications spécifiques pourraient en découler :

- 20                                   - PCR sur des matrices possédant des structures secondaires bloquantes aux températures usuelles d'élongation de 72 à 74°C,  
                                      - PCR inverse directe sur matrice circulaire double brin (chromosome).

25                                   Les principes de la PCR inverse (iPCR) et de la méthode nouvelle élaborée avec la Pab, dénommée High-Temperature inverse PCR (HT-iPCR) sont représentés dans les schémas de la figure 16. Un double brin d'ADN, de séquence partiellement connue représentée "en gras" dans la figure 16, est coupé par une enzyme de restriction RI, dilué et ligaturé pour former une molécule d'ADN circulaire double brin.

30                                   Dans la PCR inverse conventionnelle (procédure indiquée 1, 2 et 3 dans la figure 16), les molécules circulaires sont linéarisées par RII et l'amplification s'effectue sur l'ADN linéarisé. En HT-

iPCR, l'amplification est effectuée sur les molécules circulaires (procédure indiquée 2' dans la figure 16).

a) Matériels et méthodes.

5 La matrice d'ADN retenue pour ces travaux est un fragment circulaire d'ADN de 2,2 kb. Ce fragment est obtenu par digestion de l'ADN génomique de l'isolat *Thermococcus* sp. GE8 par l'enzyme Xho, purification sur gel d'agarose à 0,8% et élution de la bande  
10 2 - 2,5 kb par la méthode de GeneClean, puis ligature des extrémités des fragments linéaires par la ligase T4. Les amorces oligonucléotidiques sont choisies sur la partie de séquence connue de ce fragment circularisé de manière à copier depuis la région connue vers la région  
15 de séquence inconnue (figure 16).

Les cycles d'amplification ont été réalisés avec un thermocycleur à gradient de la société Stratagene dans les conditions suivantes :

HT-iPCR avec un seul cycle à 85°C :

20 94°C, 15s/1 cycle  
94°C, 15s/56°C, 45s/85°C, 4mn/1cycle  
94°C, 15s/56°C, 45s/82-78°C, 4mn/1 cycle  
94°C, 15s/56°C, 45s/76-72°C, 4mn/1cycle  
94°C, 15s/56°C, 45s/72°C, 4 mn/27 cycles

25 HT-iPCR continue à 85°C :

94°C, 15s/1 cycle  
94°C, 15s/56°C, 45s/85°C, 4mn/30 cycles  
stockage à 4°C

30 Les résultats sont visualisés sur gels d'agarose à 0,8%.

La figure 17 représente la comparaison des résultats de i-PCR et de HT-iPCR en utilisant la Taq et la Pab :

35 - ligne 1 : marqueur "Raoul" (société Appligene-Oncor)

- ligne 2 : HT-iPCR avec la Pab

- ligne 3 : HT-iPCR avec la Taq
- ligne 4 : i-PCR avec la Taq.

La figure 18 représente la PCR à haute température (élongation à 85°C sur 30 cycles) :

- 5
- ligne 1 : marqueur "Raoul" (société Appligene-Oncor)
  - ligne 2 : Pab
  - ligne 3 : Ppr
  - ligne 4 : témoin sans ADN
  - 10 - ligne 5 : Vent.

b) Résultats.

Les résultats de la PCR inverse avec un seul cycle initial d'élongation à haute température (85°C) mettent en évidence un excellent comportement de la Pab, et une inaptitude de la Taq (Perkin-Elmer) à effectuer ce type de réaction. Le témoin de PCR inverse après linéarisation de la matrice par la Taq fournit un signal de faible amplitude dans les conditions retenues (figure 17).

15

20

Les résultats obtenus avec 30 cycles d'amplification avec des températures d'élongation de 85°C mettent en évidence une excellente aptitude de la Pab à effectuer ce type de réaction (figure 18). La Ppr est également en mesure d'amplifier l'ADN cible dans ces conditions, mais la Vent polymérase en est incapable de même que la Pfu.

25

Les résultats rapportés ci-dessus permettent des applications remarquables des polymérases de l'invention.

30

D'une part, l'amplification de fragments d'ADN présentant des structures secondaires rendant difficile l'élongation à des températures de 72-74°C. Le fait de monter en température jusqu'à 85°C permet de résoudre ces structures secondaires qui peuvent se

35

former sur du simple brin lors de l'étape d'hybridation des amorces.

D'autre part, sur le chromosome lorsqu'une faible partie de séquence d'un gène est connue. La méthode exposée ci-dessus s'applique d'une manière générale comme suit :

- digestion de l'ADN génomique par plusieurs enzymes de restriction,

- ligature par un ligase, comme la ligase T4, des produits de digestion,

- HT-iPCR à partir d'amorces de sens opposé sur la partie de séquence connue en utilisant la Pab ou la Ppr,

- éventuellement séquençage des produits d'amplification.

#### 8) Étude de la thermostabilité des enzymes recombinantes de l'invention.

##### a) Matériels et méthodes.

Pour réaliser les tests nécessaires à cette étude, les ADN polymérases Pab et Ppr sont diluées dans un tampon de conditionnement (20 mM Tris-HCl pH=8,0 ; 1 mM EDTA ; 1 mM dithiothréitol ; 50 % glycérol ; 0,5 % Tween 20 ; 0,5 % Nonidet 40 ; 0,2 mg/ml BSA), de façon à ajouter 0,01 U. de polymérase par test et à se placer dans les conditions non saturantes. Ainsi, 0,01 U. de Pab ou de Ppr sont incubés en 20 µl de tampon d'incubation (10 mM Tris HCl pP=9,0 ; 50 mM KCl ; 3 mM MgSO<sub>4</sub> ; 0,1 % Tween) à différents temps (de 0 à 36 heures), à 92°C, à 95°C ou à 100°C. Autant de tubes que de temps différents d'incubation ont été préparés à partir d'un même mélange initial. Les différents tubes sont incubés dans un bain à sec (bloc PCR). A la fin de l'incubation, les tubes sont mis dans la glace et l'activité résiduelle de l'enzyme sera dosée à 72°C comme décrit ci-après.

Le test d'activité est effectué dans un volume final de 100 µl de tampon d'incubation (10 mM Tris HCl pP=9,0 ; 50 mM KCl ; 3 mM MgSO<sub>4</sub> ; 0,1 % Tween) dans les conditions suivantes :

- 5                   - 13 µg d'ADN activé de thymus de veau  
                   - 500 µM de chacun de 4 dNTP  
                   - 10 µCi de l'un des quatre dNTP marqué au <sup>32</sup>P utilisé comm etraceur (du dATP ou dNTP ont été utilisés)
- 10                  - les 20 µl de la solution correspondant au test d'inactivation par la température.

Les essais sont incubés 30 minutes à 70°C. 10 µl de chaque test sont prélevés après 10 minutes, 20 minutes et 30 minutes d'incubation, puis sont déposés sur papier DE81 (Whatmann). Un compactage avant lavage est alors réalisé à sec par la méthode de Cerenkov. Les papiers DE81 sont ensuite lavés trois fois 5 minutes dans une solution 0,5 mM de Na<sub>2</sub>HPO<sub>4</sub> afin d'éliminer les dNTP non incorporés. Un passage à l'alcool permet un séchage rapide des papiers qui sont remis à compter. Un témoin négatif (To) est réalisé sans ADN polymérase et un témoin positif est réalisé avec de l'ADN polymérase non traité à la température, correspondant au temps 0

15

20                  heure. Une unité polymérase est la quantité d'enzyme nécessaire à l'incorporation de 10 nmoles de nucléosides en produits acido-solubles en 30 minutes à 72°C dans les conditions d'essais. L'activité polymérase est calculé en faisant la moyenne des coups totaux avant lavage en 10 µl de dépôt. Ceux-ci correspondent à 4 x 5 nanomoles

25

30                  de chaque dNTP, soit 20 nanomoles de dNTP totaux. A chaque valeur après lavage (Ci), il faut retrancher la valeur du témoin négatif To (sans enzyme).

Soit, U/µl =

35                  (Ci-To) x (30min/temps d'incubation) x (100µl/10µl dépôt) x dilution polymérase

---

(Ct pour 10 nanomoles) x µl de polymérase ajouté

b) Résultats.

Les deux enzymes ont été étudiés à trois températures 92°C, 95°C et 100°C. Tous les travaux ont été réalisés en parallèle avec les mêmes substrats. Les résultats sont représentés sur les courbes des figures 19 à 22. Pour les figures 21 (tableau V ci-après) et 22 (tableau VI ci-après), les différentes polymérases étudiées sont ramenées à une même base au niveau activité initiale, pour en faciliter l'étude comparative. Une quatrième température de 105°C a également été étudiée pour la Pab mais n'est pas illustrée dans les figures.

Figure 19 : activité Pab.

Le tableau III ci-dessous rapporte les résultats de l'inactivation de la Pab en fonction du temps à 100°C, 95°C et 92°C.

Tableau III

| Temps (heure) | Inactivation à 100°C | Inactivation à 95°C | Inactivation à 92°C |
|---------------|----------------------|---------------------|---------------------|
| 0             | 23,8                 | 22,3                | 25                  |
| 1             | 23,6                 |                     |                     |
| 2             | 21,3                 | 25,8                | 27,1                |
| 3             | 19,7                 |                     |                     |
| 4             | 17,2                 | 19,9                | 23,1                |
| 5             | 14,1                 |                     |                     |
| 6             | 15,9                 | 17,1                | 22,8                |
| 7             | 12,6                 |                     |                     |
| 8             | 11,2                 | 14,9                | 21,8                |
| 9             |                      |                     |                     |
| 10            | 10,5                 | 13,6                | 19,5                |
| 12            |                      | 12,7                | 18                  |
| 14            |                      | 11,6                | 14,05               |
| 16            | 8,6                  | 11,6                | 11,2                |
| 18            | 7,5                  | 10,6                | 10,3                |
| 20            |                      | 9,4                 | 9,9                 |
| 22            |                      | 9,7                 | 7,5                 |
| 24            | 6,2                  | 8                   | 7                   |
| 29            | 7,1                  |                     |                     |
| 33            | 7                    |                     |                     |
| 36            | 4,3                  | 5,3                 | 6,5                 |

Figure 20 : activité Ppr.

Le tableau IV ci-dessous rapporte les résultats de l'inactivation de la Ppr en fonction du temps à 100°C, 95°C et 92°C.

5

Tableau IV

| Temps (heure) | Inactivation à 100°C | Inactivation à 95°C | Inactivation à 92°C |
|---------------|----------------------|---------------------|---------------------|
| 0             | 6,5                  | 7,66                | 7,88                |
| 1             | 4,9                  | 5,5                 | 8,98                |
| 2             | 3,75                 | 5,63                | 7,09                |
| 3             | 1,9                  | 5,23                | 6,36                |
| 4             | 1                    | 4,8                 | 4,94                |
| 4,5           |                      | 5,3                 |                     |
| 5             |                      |                     | 5,25                |
| 6             | 0                    |                     | 5,27                |
| 7             |                      |                     | 4,8                 |
| 8             |                      |                     | 4,44                |
| 9             |                      | 3,75                |                     |
| 10            |                      |                     | 4,34                |
| 11            |                      | 3,28                |                     |
| 12            |                      |                     | 3,92                |
| 13            |                      | 3,35                |                     |
| 14            |                      |                     | 4,11                |
| 15            |                      | 2,14                |                     |
| 16            |                      |                     | 2,63                |
| 17            |                      | 1,94                |                     |
| 18            |                      |                     | 2,8                 |
| 24            |                      |                     | 1,65                |
| 29            |                      |                     | 1,84                |
| 33            |                      | 1,06                | 2,46                |
| 36            |                      |                     | 1,23                |

Figure 21 : activité Pab et Ppr après inactivation à 100°C.

Le tableau V ci-dessous rapporte l'activité de Pab et Ppr après inactivation à 100°C.

5

Tableau V

| Temps (heure) | Pab  | Ppr  |
|---------------|------|------|
| 0             | 23,8 | 23,8 |
| 1             | 23,6 | 17,9 |
| 2             | 21,3 | 13,7 |
| 3             | 19,7 | 6,9  |
| 4             | 17,2 | 3,7  |
| 5             | 14,1 |      |
| 6             | 15,9 | 0    |
| 7             | 12,6 |      |
| 8             | 11,2 |      |
| 10            | 10,5 |      |
| 16            | 8,6  |      |
| 18            | 7,5  |      |
| 24            | 6,2  |      |
| 29            | 7,1  |      |
| 33            | 7    |      |
| 36            | 4,3  |      |

Figure 22 : activité Pab, Ppr et polymérase de thermus aquaticus (Taq) après inactivation à 92°C.

Le tableau VI ci-dessous rapporte l'activité de Pab, Ppr et de la Taq après inactivation à 92°C.

5

Tableau VI

| Temps (heure) | Pab   | Taq  | Ppr  |
|---------------|-------|------|------|
| 0             | 25    | 27,7 | 25   |
| 1             |       | 26,5 | 28,5 |
| 2             | 27,1  | 22,5 | 22,5 |
| 3             |       | 20,2 | 20,2 |
| 4             | 23,1  | 19   | 15,6 |
| 5             |       | 15,7 | 16,6 |
| 6             | 22,8  | 14,5 | 16,7 |
| 7             |       |      | 15,2 |
| 8             | 21,8  | 12,3 | 14,1 |
| 10            | 19,5  |      | 13,7 |
| 12            | 18    |      | 12,5 |
| 14            | 14,05 |      | 13   |
| 16            | 11,2  |      | 8,3  |
| 18            | 10,3  |      | 8,9  |
| 20            | 9,9   |      |      |
| 22            | 7,5   |      |      |
| 24            |       | 2,3  | 5,2  |
| 29            |       |      | 5,8  |
| 36            | 6,5   | 0,5  | 3,9  |
| 48            |       | 0,8  |      |
| 56            |       | 2,2  |      |
| 72            |       | 1    |      |
| 84            |       | 0,7  |      |

10 L'analyse des courbes des figures 19 à 22 montrent que :

- Pab conserve 90% de son activité après un traitement de 5 heures à 92°C, 70% après 5 heures à 95°C et 60% après 5 heures à 100°C. La Pab conserve 50% de son activité après un traitement de 90 minutes à 105°C.

15 - Ppr conserve 68% de son activité après 5 heures à 92°C, 65% après 5 heures à 95°C et 10% après 5 heures à 100°C.

En comparaison Taq pol conserve 55% de son activité après 5 heures à 92°C, mais elle n'a qu'une demi-vie de 90 minutes à 95°C et de 5 minutes à 100°C.

5 D'après les courbes des figures 19 à 22 et les commentaires ci-dessus, l'enzyme Pab est plus thermostable que la Ppr et ce, malgré une différence seulement dix résidus entre leur séquence. L'enzyme Pab est la plus thermostable actuellement décrite avec une  
10 demi-vie à 95°C de 18 heures et de plus de 8 heures à 100°C. L'enzyme Ppr présente une thermostabilité comparable à celle publiée par la Vent voire même meilleure avec 2 heures de demi-vie à 100°C.

## LISTE DE SÉQUENCES

- (1) INFORMATION GÉNÉRALES :
- 5 (i) DEPOSANT : APPLIGENE - ONCOR
- (ii) TITRE DE L'INVENTION: ADN POLYMERASE  
THERMOSTABLE D'ARCHAEBACTÉRIES DU GENRE  
PYROCOCCUS SP.
- (iii) NOMBRE DE SEQUENCES: 2
- 10 (2) INFORMATION POUR LA SEQ ID NO:1 :
- (i) CARACTÉRISTIQUES DE LA SEQUENCE:
- (A) LONGUEUR: 4446 paires de bases
- (B) TYPE: nucléotide
- 15 (C) NOMBRE DE BRIN: double
- (D) CONFIGURATION: linéaire
- (ii) TYPE DE MOLECULE: ADN
- (ix) CARACTÉRISTIQUES
- (A) NOM/CLE: séquence codante de l'ADN  
polymérase de *Pyrococcus* sp. GE.23
- 20 (B) EMPLACEMENT: de 1547 à 3862
- (ix) CARACTÉRISTIQUES
- (A) NOM/CLE: codon stop
- (B) EMPLACEMENT: de 3860 à 3862
- (xi) DESCRIPTION DE LA SEQUENCES: SEQ ID NO:1 :
- 25 CGAAGATGAG AGATTTGGTG GAATGCCGAC TTACGGGCAA GAAATTTGAG AGAGATAAAA 60
- TCAACGTTAA GATAGCGGTG GCCTATTCTG GAGGAAGCGA TAGCTCAGCC ACAGTAAAGA 120
- 30 TACTGAGATG GGCTGGCTTT GATGTGGTCC CAATAACGGC GAGGCTTCCC CACATAAGCA 180
- AAGAGGAGTT ACGGGAAGAA ACTCTATTCTG TGGAAAGTTC TGGGTACCTT GAGGAGATGG 240
- AGAGGTTAAT AGAAAAGAGG GCCCCTATCT GTGGAAGGTG CCACTCTATG GTTATGAGAG 300
- 35 CTGTTGCGAG AAAAGGTTAG GGAGCTTAAA ATAAGAATAC TCGCTACTGG AGACATGCTC 360
- AGCATAGGAA GCGGGTCAAT CTACGAGAAA GAAAATCTTG TGATTTTGAA CTTACCAGCT 420
- 40 TTCCTATCAC TAAACAAGGT TGACCTTCTG AGCATACTAG GCTGGGAGGA TTATGAGTTT 480
- AAGTATGGAT GCCCCCTTTG GAGGGAGGCC GTGAAAAGGG CTCCAATAAT GAAGAGGTTT 540
- GCAATCCAGA GGGTTCTGAG GGAATTGAGG GCAGGGGCAA TAAACGAGAA TATTGCTAAG 600
- 45 AAAC TTATTT TTGATATATT AAGGGCCTAA ACGAACCTCG CCGGTCTGAG GGTTTTCACT 660
- TTAATTTTCT TGCTATTCGT AACCCCTGAAC CCTTCCTTGG CCACATAGGT TTTCAGCCG 720
- 50 GTAACGTTTT GGATATACTG GGCCTCCTTA TAGGGACCAG CGAAGTGCAT CTTTCATGCCG 780
- AGATGCGTCA TCACGAGAAC TTCAGGCCTT TGCTTCATTG CCTTTAGCAT GTAAACTATG 840
- TCGTCGGTTG ATAAGTGGTA GGGAAATCTTC ATGTCCCTGG GCCTCGTTAC TGAGGCTATC 900
- 55 AAAACTCTCG ACCCATCATG CCAGCTTACC AGCTCTGGAA AATACTCAGT ATCCGCTATG 960

|    |   |            |            |            |            |             |      |
|----|---|------------|------------|------------|------------|-------------|------|
|    | TACGAGATAT  | CCCCAAGGCT | GGTTTTTAAT | CTAAAGCCTA | TCGTTGTTGA | GTCGCTGTGC  | 1020 |
|    | TGGGATGGAG  | TTATTATCAT | CTCCTCATTT | CCTAACCTAA | ACCTGTCCCC | AGGATTGGGG  | 1080 |
| 5  | GCATGGACTT  | CCTCTAATGC | CTCCAAGTGG | TACTTGCTCA | GGGCTGGGGT | ATGATCCTCG  | 1140 |
|    | CCCCCATAAA  | ACACGCTTCT | AGAACCTATT | AGGGTTCCCC | TCTTCTTAGT | AACACCGTAT  | 1200 |
| 10 | GTCATTCCCA  | CAACGAACAC | CTCGGCGGCG | GTGGAGTGAT | CGGTGTGTCT | ATGCGAAATG  | 1260 |
|    | AAGAGAACAT  | CTATCTTCCT | GGGGTCTAAC | TTGTATCTAA | TCATCCTAAC | TAGCGCTCCA  | 1320 |
|    | GGCCCAGGAT  | CCACAAAGAT | ATTTTTGCTT | GCCTTGATGA | AGAATCCCCC | TGTGGATCTT  | 1380 |
| 15 | ACTTGAGTTA  | TCGTCACGAA | CCTGCCCCCA | CCGGCACCCA | AGAACGTAAT | CTCTATCATT  | 1440 |
|    | TTTAGTCCCG  | AAATTAAGT  | GCGAGGCTTA | TGCTTTTAAG | GATGTATGGC | GAAAGGTGAA  | 1500 |
| 20 | GTTTATTAGA  | AGTTAGAATC | TAAAGATTC  | AGATTGGGTG | GGGGTA     | ATG ATA ATC | 1555 |
|    |   |            |            |            |            | Met Ile Ile |      |
|    |   |            |            |            |            | 1 3         |      |
| 25 | GAT GCT GAT TAC ATA ACG GAA GAT GGC AAG CCG ATA ATA AGA ATA TTC | 1603       |            |            |            |             |      |
|    | Asp Ala Asp Tyr Ile Thr Glu Asp Gly Lys Pro Ile Ile Arg Ile Phe |            |            |            |            |             |      |
|    | 5 10 15   |            |            |            |            |             |      |
| 30 | AAA AAG GAA AAG GGA GAG TTT AAG GTA GAA TAC GAT AGG ACG TTT AGA | 1651       |            |            |            |             |      |
|    | Lys Lys Glu Lys Gly Glu Phe Lys Val Glu Tyr Asp Arg Thr Phe Arg |            |            |            |            |             |      |
|    | 20 25 30 35   |            |            |            |            |             |      |
| 35 | CCC TAC ATT TAC GCT CTT TTA AAG GAT GAT TCG GCC ATA GAT GAG GTT | 1699       |            |            |            |             |      |
|    | Pro Tyr Ile Tyr Ala Leu Leu Lys Asp Asp Ser Ala Ile Asp Glu Val |            |            |            |            |             |      |
|    | 40 45 50  |            |            |            |            |             |      |
| 40 | AAG AAG ATA ACC GCC GAG AGG CAC GGA AAG ATA GTC AGG ATA ACC GAA | 1747       |            |            |            |             |      |
|    | Lys Lys Ile Thr Ala Glu Arg His Gly Lys Ile Val Arg Ile Thr Glu |            |            |            |            |             |      |
|    | 55 60 65  |            |            |            |            |             |      |
| 45 | GTT GAG AAA GTC CAG AAG AAA TTC CTA GGA AGG CCA ATA GAA GTC TGG | 1795       |            |            |            |             |      |
|    | Val Glu Lys Val Gln Lys Lys Phe Leu Gly Arg Pro Ile Glu Val Trp |            |            |            |            |             |      |
|    | 70 75 80  |            |            |            |            |             |      |
| 50 | AAG CTG TAC CTT GAG CAT CCA CAA GAC GTT CCA GCT ATC AGA GAG AAG | 1843       |            |            |            |             |      |
|    | Lys Leu Tyr Leu Glu His Pro Gln Asp Val Pro Ala Ile Arg Glu Lys |            |            |            |            |             |      |
|    | 85 90 95  |            |            |            |            |             |      |
| 55 | ATA AGG GAA CAT CCA GCT GTA GTT GAT ATA TTC GAA TAC GAC ATA CCC | 1891       |            |            |            |             |      |
|    | Ile Arg Glu His Pro Ala Val Val Asp Ile Phe Glu Tyr Asp Ile Pro |            |            |            |            |             |      |
|    | 100 105 110 115   |            |            |            |            |             |      |
| 55 | TTT GCG AAA CGC TAC CTA ATA GAT AAG GGA TTG ACT CCA ATG GAG GGG | 1939       |            |            |            |             |      |
|    | Phe Ala Lys Arg Tyr Leu Ile Asp Lys Gly Leu Thr Pro Met Glu Gly |            |            |            |            |             |      |
|    | 120 125 130   |            |            |            |            |             |      |
| 55 | AAC GAG GAG CTA ACG TTT CTA GCA GTT GAC ATA GAA ACA TTG TAC CAT | 1987       |            |            |            |             |      |
|    | Asn Glu Glu Leu Thr Phe Leu Ala Val Asp Ile Glu Thr Leu Tyr His |            |            |            |            |             |      |
|    | 135 140 145   |            |            |            |            |             |      |

|    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
|    | GAA | GGA | GAG | GAG | TTC | GGG | AAA | GGC | CCT | ATA | ATC | ATG | ATC | AGC | TAC | GCC | 2035 |
|    | Glu | Gly | Glu | Glu | Phe | Gly | Lys | Gly | Pro | Ile | Ile | Met | Ile | Ser | Tyr | Ala |      |
|    |     |     | 150 |     |     |     |     | 155 |     |     |     |     | 160 |     |     |     |      |
| 5  | GAC | GAG | GAA | GGG | GCC | AAG | GTG | ATA | ACT | TGG | AAG | AGC | ATA | GAC | TTA | CCT | 2083 |
|    | Asp | Glu | Glu | Gly | Ala | Lys | Val | Ile | Thr | Trp | Lys | Ser | Ile | Asp | Leu | Pro |      |
|    |     | 165 |     |     |     |     | 170 |     |     |     |     | 175 |     |     |     |     |      |
| 10 | TAC | GTT | GAA | GTG | GTT | TCA | AGC | GAG | AGG | GAG | ATG | ATA | AAG | AGG | CTC | GTG | 2131 |
|    | Tyr | Val | Glu | Val | Val | Ser | Ser | Glu | Arg | Glu | Met | Ile | Lys | Arg | Leu | Val |      |
|    | 180 |     |     |     |     | 185 |     |     |     |     | 190 |     |     |     |     | 195 |      |
| 15 | AAG | GTA | ATT | AGA | GAG | AAG | GAT | CCC | GAC | GTG | ATA | ATA | ACG | TAC | AAT | GGT | 2179 |
|    | Lys | Val | Ile | Arg | Glu | Lys | Asp | Pro | Asp | Val | Ile | Ile | Thr | Tyr | Asn | Gly |      |
|    |     |     |     |     | 200 |     |     |     |     | 205 |     |     |     |     | 210 |     |      |
| 20 | GAT | AAT | TTC | GAC | TTT | CCG | TAC | CTC | TTA | AAG | AGG | GCT | GAA | AAG | CTC | GGA | 2227 |
|    | Asp | Asn | Phe | Asp | Phe | Pro | Tyr | Leu | Leu | Lys | Arg | Ala | Glu | Lys | Leu | Gly |      |
|    |     |     |     | 215 |     |     |     |     | 220 |     |     |     |     | 225 |     |     |      |
| 25 | ATA | AAG | CTC | CCC | CTT | GGA | AGG | GAC | AAT | AGC | GAG | CCG | AAG | ATG | CAG | AGG | 2275 |
|    | Ile | Lys | Leu | Pro | Leu | Gly | Arg | Asp | Asn | Ser | Glu | Pro | Lys | Met | Gln | Arg |      |
|    |     |     | 230 |     |     |     |     | 235 |     |     |     |     | 240 |     |     |     |      |
| 30 | ATG | GGG | GAT | TCA | TTA | GCT | GTA | GAG | ATA | AAG | GGC | AGA | ATA | CAC | TTC | GAT | 2323 |
|    | Met | Gly | Asp | Ser | Leu | Ala | Val | Glu | Ile | Lys | Gly | Arg | Ile | His | Phe | Asp |      |
|    |     | 245 |     |     |     | 250 |     |     |     |     |     | 255 |     |     |     |     |      |
| 35 | TTA | TTC | CCC | GTC | ATA | AGA | AGA | ACG | ATC | AAC | CTT | CCA | ACA | TAC | ACC | CTC | 2371 |
|    | Leu | Phe | Pro | Val | Ile | Arg | Arg | Thr | Ile | Asn | Leu | Pro | Thr | Tyr | Thr | Leu |      |
|    | 260 |     |     |     | 265 |     |     |     |     |     | 270 |     |     |     |     | 275 |      |
| 40 | GAA | GCG | GTT | TAT | GAG | GCT | ATA | TTT | GGA | AAG | TCT | AAG | GAG | AAA | GTC | TAT | 2419 |
|    | Glu | Ala | Val | Tyr | Glu | Ala | Ile | Phe | Gly | Lys | Ser | Lys | Glu | Lys | Val | Tyr |      |
|    |     |     |     |     | 280 |     |     |     |     | 285 |     |     |     |     | 290 |     |      |
| 45 | GCC | CAT | GAG | ATA | GCT | GAG | GCC | TGG | GAA | ACC | GGG | AAA | GGG | CTA | GAG | AGG | 2467 |
|    | Ala | His | Glu | Ile | Ala | Glu | Ala | Trp | Glu | Thr | Gly | Lys | Gly | Leu | Glu | Arg |      |
|    |     |     |     | 295 |     |     |     | 300 |     |     |     |     | 305 |     |     |     |      |
| 50 | GTA | GCT | AAG | TAT | TCA | ATG | GAA | GAT | GCG | AAG | GTA | ACC | TTT | GAG | CTC | GGA | 2515 |
|    | Val | Ala | Lys | Tyr | Ser | Met | Glu | Asp | Ala | Lys | Val | Thr | Phe | Glu | Leu | Gly |      |
|    |     |     | 310 |     |     |     |     | 315 |     |     |     |     | 320 |     |     |     |      |
| 55 | AAG | GAG | TTC | TTC | CCA | ATG | GAA | GCC | CAG | CTA | GCT | AGG | CTC | GTT | GGC | CAG | 2563 |
|    | Lys | Glu | Phe | Phe | Pro | Met | Glu | Ala | Gln | Leu | Ala | Arg | Leu | Val | Gly | Gln |      |
|    |     | 325 |     |     |     | 330 |     |     |     |     |     | 335 |     |     |     |     |      |
| 60 | CCA | GTT | TGG | GAC | GTT | TCA | AGG | TCG | AGC | ACC | GGA | AAC | CTC | GTT | GAG | TGG | 2611 |
|    | Pro | Val | Trp | Asp | Val | Ser | Arg | Ser | Ser | Thr | Gly | Asn | Leu | Val | Glu | Trp |      |
|    | 340 |     |     |     |     | 345 |     |     |     |     |     | 350 |     |     |     | 355 |      |
| 65 | TTT | CTC | CTT | AGG | AAG | GCC | TAC | GAG | AGA | AAT | GAG | CTC | GCG | CCC | AAT | AAA | 2659 |
|    | Phe | Leu | Leu | Arg | Lys | Ala | Tyr | Glu | Arg | Asn | Glu | Leu | Ala | Pro | Asn | Lys |      |
|    |     |     |     |     | 360 |     |     |     |     | 365 |     |     |     |     | 370 |     |      |
| 70 | CCG | GAC | GAG | AGG | GAA | TAC | GAG | AGA | AGG | CTA | AGA | GAG | AGC | TAT | GAA | GGG | 2707 |
|    | Pro | Asp | Glu | Arg | Glu | Tyr | Glu | Arg | Arg | Leu | Arg | Glu | Ser | Tyr | Glu | Gly |      |
|    |     |     |     | 375 |     |     |     | 380 |     |     |     |     |     | 385 |     |     |      |

|    |   |      |
|----|---|------|
|    | GGT TAC GTT AAG GAG CCA GAG AAG GGA TTG TGG GAA GGG ATA GTC AGC | 2755 |
|    | Gly Tyr Val Lys Glu Pro Glu Lys Gly Leu Trp Glu Gly Ile Val Ser |      |
|    | 390 395 400   |      |
| 5  | TTA GAC TTT AGG TCC CTA TAT CCG TCT ATA ATT ATA ACT CAC AAC GTC | 2803 |
|    | Leu Asp Phe Arg Ser Leu Tyr Pro Ser Ile Ile Ile Thr His Asn Val |      |
|    | 405 410 415   |      |
| 10 | TCA CCA GAC ACT TTG AAT AGA GAA AAT TGC AAG GAA TAC GAC GTT GCC | 2851 |
|    | Ser Pro Asp Thr Leu Asn Arg Glu Asn Cys Lys Glu Tyr Asp Val Ala |      |
|    | 420 425 430 435   |      |
| 15 | CCC CAA GTG GGG CAC AGA TTC TGC AAG GAT TTC CCA GGA TTC ATA CCA | 2899 |
|    | Pro Gln Val Gly His Arg Phe Cys Lys Asp Phe Pro Gly Phe Ile Pro |      |
|    | 440 445 450   |      |
| 20 | AGC TTA CTG GGT AAC TTA CTG GAG GAG AGA CAA AAG ATA AAA AAG AGA | 2947 |
|    | Ser Leu Leu Gly Asn Leu Leu Glu Glu Arg Gln Lys Ile Lys Lys Arg |      |
|    | 455 460 465   |      |
| 25 | ATG AAA GAA AGT AAA GAT CCC GTC GAG AAG AAA CTC CTT GAT TAC AGA | 2995 |
|    | Met Lys Glu Ser Lys Asp Pro Val Glu Lys Lys Leu Leu Asp Tyr Arg |      |
|    | 470 475 480   |      |
| 30 | CAG AGA GCT ATA AAA ATA CTT GCA AAC AGC TAT TAT GGC TAT TAT GGA | 3043 |
|    | Gln Arg Ala Ile Lys Ile Leu Ala Asn Ser Tyr Tyr Gly Tyr Tyr Gly |      |
|    | 485 490 495   |      |
| 35 | TAT GCA AAG GCC AGA TGG TAC TGT AAG GAG TGT GCA GAG AGC GTA ACT | 3091 |
|    | Tyr Ala Lys Ala Arg Trp Tyr Cys Lys Glu Cys Ala Glu Ser Val Thr |      |
|    | 500 505 510 515   |      |
| 40 | GCA TGG GGG AGG CAA TAC ATA GAT CTA GTT AGA AGA GAG CTT GAA AGC | 3139 |
|    | Ala Trp Gly Arg Gln Tyr Ile Asp Leu Val Arg Arg Glu Leu Glu Ser |      |
|    | 520 525 530   |      |
| 45 | AGC GGA TTC AAG GTT CTG TAC ATA GAC ACT GAT GGC CTC TAC GCG ACC | 3187 |
|    | Ser Gly Phe Lys Val Leu Tyr Ile Asp Thr Asp Gly Leu Tyr Ala Thr |      |
|    | 535 540 545   |      |
| 50 | ATT CCT GGG GCC AAG CCA AAT GAG ATA AAA GAA AAG GCC CTT AAG TTC | 3235 |
|    | Ile Pro Gly Ala Lys Pro Asn Glu Ile Lys Glu Lys Ala Leu Lys Phe |      |
|    | 550 555 560   |      |
| 55 | GTC GAG TAC ATA AAC TCC AAG TTA CCT GGG CTT CTT GAA TTG GAA TAC | 3283 |
|    | Val Glu Tyr Ile Asn Ser Lys Leu Pro Gly Leu Leu Glu Leu Glu Tyr |      |
|    | 565 570 575   |      |
| 55 | GAA GGT TTC TAC GCG AGA GGG TTC TTC GTG ACG AAG AAA AAG TAC GCA | 3331 |
|    | Glu Gly Phe Tyr Ala Arg Gly Phe Phe Val Thr Lys Lys Lys Tyr Ala |      |
|    | 580 585 590 595   |      |
| 55 | CTA ATC GAC GAG GAA GGA AAG ATA GTT ACG AGG GGG CTC GAA ATA GTA | 3379 |
|    | Leu Ile Asp Glu Glu Gly Lys Ile Val Thr Arg Gly Leu Glu Ile Val |      |
|    | 600 605 610   |      |
| 55 | AGG AGA GAT TGG AGT GAA ATA GCA AAG GAG ACC CAA GCT AAG GTT CTC | 3427 |
|    | Arg Arg Asp Trp Ser Glu Ile Ala Lys Glu Thr Gln Ala Lys Val Leu |      |
|    | 615 620 625   |      |

|    |  |      |
|----|--|------|
|    | GAG GCA ATA CTC AAG CAC GGT AAC GTT GAT GAG GCC GTA AAA ATA GTA    | 3475 |
|    | Glu Ala Ile Leu Lys His Gly Asn Val Asp Glu Ala Val Lys Ile Val    |      |
|    | 630 635 640  |      |
| 5  | AAG GAG GTT ACA GAA AAA CTC AGT AAA TAT GAA ATA CCA CCC GAA AAG    | 3523 |
|    | Lys Glu Val Thr Glu Lys Leu Ser Lys Tyr Glu Ile Pro Pro Glu Lys    |      |
|    | 645 650 655  |      |
| 10 | CTT GTA ATT TAT GAG CAG ATA ACG AGG CCT CTG AGC GAG TAT AAA GCG    | 3571 |
|    | Leu Val Ile Tyr Glu Gln Ile Thr Arg Pro Leu Ser Glu Tyr Lys Ala    |      |
|    | 660 665 670 675  |      |
| 15 | ATA GGC CCT CAC GTT GCA GTA GCT AAA AGG CTC GCA GCG AAG GGA GTA    | 3619 |
|    | Ile Gly Pro His Val Ala Val Ala Lys Arg Leu Ala Ala Lys Gly Val    |      |
|    | 680 685 690  |      |
| 20 | AAA GTT AAG CCA GGG ATG GTT ATC GGT TAC ATA GTT TTG AGG GGA GAC    | 3667 |
|    | Lys Val Lys Pro Gly Met Val Ile Gly Tyr Ile Val Leu Arg Gly Asp    |      |
|    | 695 700 705  |      |
|    | GGG CCA ATA AGC AAG AGG GCC ATA GCT ATA GAG GAG TTC GAT CCC AAA    | 3715 |
|    | Gly Pro Ile Ser Lys Arg Ala Ile Ala Ile Glu Glu Phe Asp Pro Lys    |      |
|    | 710 715 720  |      |
| 25 | AAG CAT AAG TAC GAT GCC GAA TAC TAC ATA GAG AAC CAA GTT CTG CCA    | 3763 |
|    | Lys His Lys Tyr Asp Ala Glu Tyr Tyr Ile Glu Asn Gln Val Leu Pro    |      |
|    | 725 730 735  |      |
| 30 | GCG GTG GAG AGG ATA TTG AGA GCA TTT GGT TAT CGC AAA GAA GAT TTG    | 3811 |
|    | Ala Val Glu Arg Ile Leu Arg Ala Phe Gly Tyr Arg Lys Glu Asp Leu    |      |
|    | 740 745 750 755  |      |
| 35 | AGG TAT CAA AAA ACT AAA CAA GTG GGC CTC GGA GCA TGG CTT AAG TTC    | 3859 |
|    | Arg Tyr Gln Lys Thr Lys Gln Val Gly Leu Gly Ala Trp Leu Lys Phe    |      |
|    | 760 765 770 771  |      |
|    | TAG CTACCCAGAT GTCACCGTAT CTCAACAGGT ATTCTTGGAG ATCTCTTAAA         | 3912 |
| 40 | TCAACTACAA GCTCTTCCTC AAGTTCCATA AAGTTTATTG ACTTTATCGG TTTAATTATG  | 3972 |
|    | AGCTTATAGG AGCCTAGAAC CCCAGAAATC TTAACTCTAA AGACTCTTGA AGCTAGCTCT  | 4032 |
| 45 | ATCAGTTCAA GAACTATGTC CTTCTTAAGG AACGAGGAAT TAATGAAAAC TATTCCTTTA  | 4092 |
|    | CCGTTCCGGAT CCTGGAGAGC CATTTTCCCA ACTAATGTGA AGAAGAGTTC GCTTTCAATG | 4152 |
|    | TAGTTACTCT CCCTTGTTTT TAGAAGTCTC TCTAAGCCGA CGTTTATAGT GAACCGTCTT  | 4212 |
| 50 | TTCCCCGTGC TTCTCAAGGG TAGTGAAAAG TTGTTTTCTC CAAACTCCGA TATCCGAGCC  | 4272 |
|    | TATTTCTATT CTCTTCACTA CGCTGCCCGT TTTTATAAAT CCACCAAGTT TAACGACCCT  | 4332 |
| 55 | GGCACTATCT ATAACATCGG TCTTCAGACC CAATAGCCTT AAGGTGGTTT CTCAAAACTG  | 4392 |
|    | AGACCTTCCC TTTAGAGCAT TCAAGACCAT TAGAAGATCA GGTCTATTGG CTCG        | 4446 |

## (2) INFORMATION POUR LA SEQ ID NO:2 :

(i) CARACTERISTIQUES DE LA SEQUENCE:  
 (A) LONGUEUR: 2995 paires de bases  
 (B) TYPE: nucléotide  
 (C) NOMBRE DE BRIN: double  
 (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(ix) CARACTERISTIQUES  
 (A) NOM/CLE: séquence codante de l'ADN polymérase de *Pyrococcus* sp. GE 5  
 (B) EMPLACEMENT: de 678 à 2994

(ix) CARACTERISTIQUES  
 (A) NOM/CLE: codon stop  
 (B) EMPLACEMENT: de 2992 à 2994

(xi) DESCRIPTION DE LA SEQUENCES: SEQ ID NO:2 :

```

TCATGTCCTG GGGCTTGGTT ACGGAGGCTA TCAAATATT GGACCXTTCG TGCCAGCTTA      60
CCAGCTCTGG AAAATACTCA GTATCCGCTA TGTACGAGAT ATCCCCAAGG CTAGTTTTTA      120
ATCTAAAGCC TATAGTTGTT GGGTCGCTGT GCTGGGATGG AGTTATTATC ATCTCCTCAT      180
TTCCTAGCCT AAACCTGTCC CCAGGATTAG GAGCATGGAC TTCCTCTAAT GCCTCCAAGT      240
GGTACTTGCT CAGGGCTGGG GTATGATCCT CGTCCCCATA AACCACGCTT CTAGAACCTA      300
TTAGGGTTCC CCTCTTCTTA GTWACCCCAT AGGTCATCCC YTCAACGATC ACYTCGGCAT      360
CGTTGCAGTG ATCGGTGTGT CTATGCGAGA TGAAGAGAAC ATCTATCTTC CTGGGGTCTA      420
GCTTATATCT AATCATCCTA ACTAGCGCTC CAGGCCCAGG GTCCACAAAG ATATTTTTTGC      480
TTGCCTTGAT GAAGAATCCA CCCGTAGATC TTA CTGAGT TATCGTCACG AACCTGCCCC      540
CACCGGCACC CAGGAACGTA ATCTCTATCA TTTT TAGTCC CGAAATTAAA GTGCGAGGCT      600
TATGCTTTTA AGGATGTATG GCGAAAGGTG AAGTTTATTA GAAGTTAGAA TCTAAAGATT      660
TCAGATTGGG TGGGGGTA ATG ATA ATC GAT GCT GAT TAC ATA ACG GAA GAT      711
                Met Ile Ile Asp Ala Asp Tyr Ile Thr Glu Asp
                1                5                10
GGC AAG CCG ATA ATA AGG ATA TTC AAA AAG GAA AAG GGA GAG TTT AAG      759
Gly Lys Pro Ile Ile Arg Ile Phe Lys Lys Glu Lys Gly Glu Phe Lys
                15                20                25
GTA GAA TAC GAT AGG ACG TTT AGA CCC TAC ATT TAT GCT CTT TTA AAG      807
Val Glu Tyr Asp Arg Thr Phe Arg Pro Tyr Ile Tyr Ala Leu Leu Lys
                30                35                40
GAT GAT TCG GCC ATA GAT GAG GTT AAG AAG ATA ACC GCC GAG AGG CAC      855
Asp Asp Ser Ala Ile Asp Glu Val Lys Lys Ile Thr Ala Glu Arg His
                45                50                55
GGA AAG ATA GTC AGG ATA ACC GAG GTT GAG AAA GTC CAG AAG AAA TTC      903
Gly Lys Ile Val Arg Ile Thr Glu Val Glu Lys Val Gln Lys Lys Phe
60                65                70                75

```

|    |   |      |
|----|---|------|
|    | CTA GGA AGG CCA ATA GAA GTC TGG AAG CTC TAT CTT GAG CAT CCC CAG | 951  |
|    | Leu Gly Arg Pro Ile Glu Val Trp Lys Leu Tyr Leu Glu His Pro Gln |      |
|    | 80 85 90  |      |
| 5  | GAT GTT CCA GCC ATA AGA GAG AAG ATA AGG GAA CAT CCA GCT GTA GTT | 999  |
|    | Asp Val Pro Ala Ile Arg Glu Lys Ile Arg Glu His Pro Ala Val Val |      |
|    | 95 100 105  |      |
| 10 | GAT ATA TTT GAA TAC GAC ATA CCC TTT GCG AAG CGC TAC CTC ATA GAC | 1047 |
|    | Asp Ile Phe Glu Tyr Asp Ile Pro Phe Ala Lys Arg Tyr Leu Ile Asp |      |
|    | 110 115 120   |      |
| 15 | AAG GGA TTG ACT CCA ATG GAG GGG AAC GAG GAG CTA ACG TTT CTA GCC | 1095 |
|    | Lys Gly Leu Thr Pro Met Glu Gly Asn Glu Glu Leu Thr Phe Leu Ala |      |
|    | 125 130 135   |      |
| 20 | GTT GAT ATA GAA ACA TTG TAC CAT GAA GGA GAG GAG TTC GGG AAA GGG | 1143 |
|    | Val Asp Ile Glu Thr Leu Tyr His Glu Gly Glu Glu Phe Gly Lys Gly |      |
|    | 140 145 150 155   |      |
|    | CCA ATA ATA ATG ATC AGC TAC GCC GAC GAG GAA GGG GCC AAG GTG ATA | 1191 |
|    | Pro Ile Ile Met Ile Ser Tyr Ala Asp Glu Glu Gly Ala Lys Val Ile |      |
|    | 160 165 170   |      |
| 25 | ACT TGG AAG AGC ATA GAC TTA CCT TAC GTT GAA GTG GTT TCG AGC GAG | 1239 |
|    | Thr Trp Lys Ser Ile Asp Leu Pro Tyr Val Glu Val Val Ser Ser Glu |      |
|    | 175 180 185   |      |
| 30 | AGG GAG ATG ATA AAG AGG CTC GTG AAG GTA ATT AGA GAG AAA GAT CCC | 1287 |
|    | Arg Glu Met Ile Lys Arg Leu Val Lys Val Ile Arg Glu Lys Asp Pro |      |
|    | 190 195 200   |      |
| 35 | GAC GTG ATA ATA ACG TAC AAT GGT GAT AAT TTC GAC TTT CCG TAC CTC | 1335 |
|    | Asp Val Ile Ile Thr Tyr Asn Gly Asp Asn Phe Asp Phe Pro Tyr Leu |      |
|    | 205 210 215   |      |
| 40 | TTA AAG AGG GCT GAA AAG CTC GGA ATA AAG CTC CCC CTT GGA AGG GAC | 1383 |
|    | Leu Lys Arg Ala Glu Lys Leu Gly Ile Lys Leu Pro Leu Gly Arg Asp |      |
|    | 220 225 230 235   |      |
|    | AAT AGC GAG CCG AAA ATG CAG AGG ATG GGG GAT TCA TTA GCC GTA GAG | 1431 |
|    | Asn Ser Glu Pro Lys Met Gln Arg Met Gly Asp Ser Leu Ala Val Glu |      |
|    | 240 245 250   |      |
| 45 | ATA AAG GGC AGA ATA CAC TTC GAT TTA TTC CCC GCC ATA AGA AGA ACG | 1479 |
|    | Ile Lys Gly Arg Ile His Phe Asp Leu Phe Pro Ala Ile Arg Arg Thr |      |
|    | 255 260 265   |      |
| 50 | ATC AAC CTT CCA ACA TAC ACC CTC GAA ACG GTT TAT GAG GTT ATA TTT | 1527 |
|    | Ile Asn Leu Pro Thr Tyr Thr Leu Glu Thr Val Tyr Glu Val Ile Phe |      |
|    | 270 275 280   |      |
| 55 | GGA AAG TCT AAG GAG AAA GTC TAT GCC CAT GAG ATA GCT GAG GCC TGG | 1575 |
|    | Gly Lys Ser Lys Glu Lys Val Tyr Ala His Glu Ile Ala Glu Ala Trp |      |
|    | 285 290 295   |      |

|    |   |      |
|----|---|------|
|    | GAA ACC GGG AAA GGG CTA GAG AGG GTA GCT AAG TAT TCA ATG GAA GAT | 1623 |
|    | Glu Thr Gly Lys Gly Leu Glu Arg Val Ala Lys Tyr Ser Met Glu Asp |      |
|    | 300 305 310 315   |      |
| 5  | GCG AAG GTA ACC TCT GAG CTC GGA AAG GAG TTC TTC CCG ATG GAA GCC | 1671 |
|    | Ala Lys Val Thr Ser Glu Leu Gly Lys Glu Phe Phe Pro Met Glu Ala |      |
|    | 320 325 330   |      |
| 10 | CAG CTA GCT AGG CTC GTT GGC CAT CCA GTT TGG GAC GTT TCA AGG TCG | 1719 |
|    | Gln Leu Ala Arg Leu Val Gly His Pro Val Trp Asp Val Ser Arg Ser |      |
|    | 335 340 345   |      |
| 15 | AGC ACC GGA AAC CTC GTT GAG TGG TTT CTC CTT ACG AAG GCC TAC GAG | 1767 |
|    | Ser Thr Gly Asn Leu Val Glu Trp Phe Leu Leu Thr Lys Ala Tyr Glu |      |
|    | 350 355 360   |      |
| 20 | AGA AAT GAG CTC GCG CCC AAT AAA CCG GAC GAG AGG GAA TAC GAG AGA | 1815 |
|    | Arg Asn Glu Leu Ala Pro Asn Lys Pro Asp Glu Arg Glu Tyr Glu Arg |      |
|    | 365 370 375   |      |
|    | AGG CTA AGA GAG AGC TAT GAA GGG GGT TAC GTT AAC GAG CCA GAG AAG | 1863 |
|    | Arg Leu Arg Glu Ser Tyr Glu Gly Gly Tyr Val Asn Glu Pro Glu Lys |      |
|    | 380 385 390 395   |      |
| 25 | GGA TTG TGG GAA GGG ATA GTC AGC TTA GAC TTT AGG TCC CTA TAT CCC | 1911 |
|    | Gly Leu Trp Glu Gly Ile Val Ser Leu Asp Phe Arg Ser Leu Tyr Pro |      |
|    | 400 405 410   |      |
| 30 | TCT ATA ATT ATA ACT CAC AAC GTC TCA CCA GAC ACT TTG AAT AGA GAA | 1959 |
|    | Ser Ile Ile Ile Thr His Asn Val Ser Pro Asp Thr Leu Asn Arg Glu |      |
|    | 415 420 425   |      |
| 35 | AAT TGC AAG GAA TAT GAC GTT GCC CCC CAA GTG GGG CAC AGA TTC TGC | 2007 |
|    | Asn Cys Lys Glu Tyr Asp Val Ala Pro Gln Val Gly His Arg Phe Cys |      |
|    | 430 435 440   |      |
| 40 | AAG GAT TTC CCA GGA TTC ATA CCA AGC TTA CTG GGT AAC CTA CTG GAG | 2055 |
|    | Lys Asp Phe Pro Gly Phe Ile Pro Ser Leu Leu Gly Asn Leu Leu Glu |      |
|    | 445 450 455   |      |
|    | GAG AGA CAA AAG ATA AAA AAG AGG ATG AAA GAA AGT AAA GAT CCC GTC | 2103 |
|    | Glu Arg Gln Lys Ile Lys Lys Arg Met Lys Glu Ser Lys Asp Pro Val |      |
|    | 460 465 470 475   |      |
| 45 | GAG AAG AAA CTC CTT GAT TAC AGA CAG AGA GCT ATA AAA ATA CTT GCA | 2151 |
|    | Glu Lys Lys Leu Leu Asp Tyr Arg Gln Arg Ala Ile Lys Ile Leu Ala |      |
|    | 480 485 490   |      |
| 50 | AAC AGC TAT TAT GGC TAT TAT GGA TAT GCA AAG GCC AGA TGG TAC TGT | 2199 |
|    | Asn Ser Tyr Tyr Gly Tyr Tyr Gly Tyr Ala Lys Ala Arg Trp Tyr Cys |      |
|    | 495 500 505   |      |
| 55 | AAA GAG TGT GCA GAG AGC GTA ACC GCA TGG GGA AGG CAG TAC ATA GAC | 2247 |
|    | Lys Glu Cys Ala Glu Ser Val Thr Ala Trp Gly Arg Gln Tyr Ile Asp |      |
|    | 510 515 520   |      |
|    | CTG GTT AGG AGG GAA CTT GAG AGC AGA GGA TTT AAA GTT CTC TAC ATA | 2295 |
|    | Leu Val Arg Arg Glu Leu Glu Ser Arg Gly Phe Lys Val Leu Tyr Ile |      |
|    | 525 530 535   |      |



## REVENDICATIONS

- 5 1) ADN polymérase purifiée thermostable  
d'archaebactéries du genre *Pyrococcus sp.* ayant un poids  
moléculaire compris entre environ 89 000 et 90 000  
daltons.
- 10 2) ADN polymérase selon la revendication 1  
de la souche d'archaebactéries du genre *Pyrococcus sp.*  
GE 5 déposée à la CNCM sous le No. I-1302
- 15 3) ADN polymérase selon la revendication 2  
ayant un poids moléculaire d'environ 89 443 daltons et  
un pI d'environ 8,13.
- 20 4) ADN polymérase selon la revendication 1  
purifiée de la souche d'archaebactéries du genre  
*Pyrococcus sp.* GE 23 déposée à la CNCM sous le No. I-  
1764.
- 25 5) ADN polymérase selon la revendication 4  
ayant un poids moléculaire d'environ 89 409 daltons et  
un pI d'environ 8,37.
- 30 6) Une ADN polymérase purifiée thermostable  
dont la séquence en acides aminés est représentée dans  
la liste de séquence en annexe sous le numéro SEQ ID  
NO:1.
- 35 7) Une ADN polymérase purifiée thermostable  
dont la séquence en acides aminés est représentée dans  
la liste de séquence en annexe sous le numéro SEQ ID  
NO:2.

8) Une séquence d'ADN constituée par ou comprenant la séquence codant pour l'ADN polymérase purifiée thermostable de l'une quelconque des revendications 1 à 7.

5

9) Une séquence d'ADN selon la revendication 8 comprenant ou constituée par les nucléotides 1547 à 3862 de la SEQ ID NO:1.

10

10) Une séquence d'ADN selon la revendication 8 comprenant ou constituée par les nucléotides 678 à 2994 de la SEQ ID NO:2.

15

11) Un vecteur contenant la séquence d'ADN de l'une quelconque des revendications 8 à 10.

12) Un hôte transformé par un vecteur selon la revendication 11.

20

13) Procédé de préparation d'une ADN polymérase d'archaebactéries du genre *Pyrococcus sp.* caractérisé en ce que l'on cultive l'hôte selon la revendication 12 dans des conditions permettant l'expression de ladite ADN polymérase et en ce que l'on extrait et récupère celle-ci par tout moyen approprié.

25

14) Procédé d'amplification enzymatique d'une séquence d'acide nucléique caractérisé en ce que l'on met en oeuvre une ADN polymérase thermostable selon l'une quelconque des revendications 1 à 7.

30

ADN POLYMERASE THERMOSTABLE  
D'ARCHAEBACTÉRIES DU GENRE *PYROCOCCUS SP.*

5 La présente invention concerne une ADN  
polymérase purifiée thermostable d'archaebactéries du  
genre *Pyrococcus sp.* ayant un poids moléculaire compris  
entre environ 89 000 et 90 000 daltons.

1/30

Fig.1

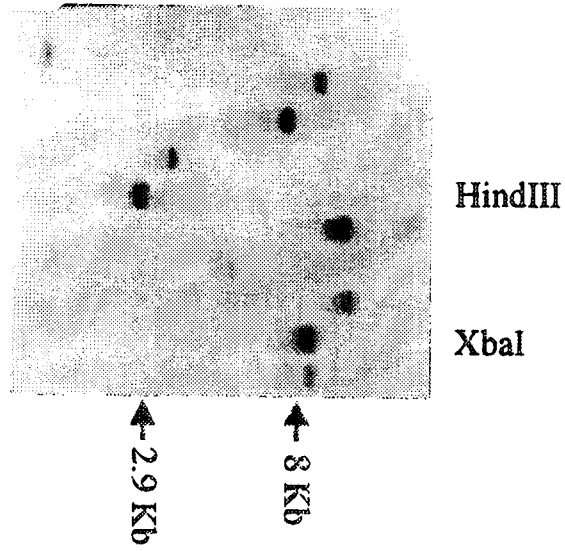


Fig.2

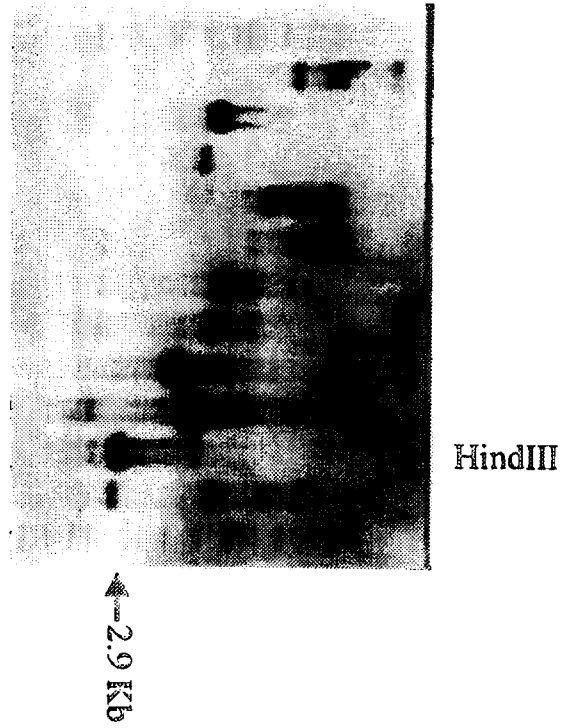


Fig.3

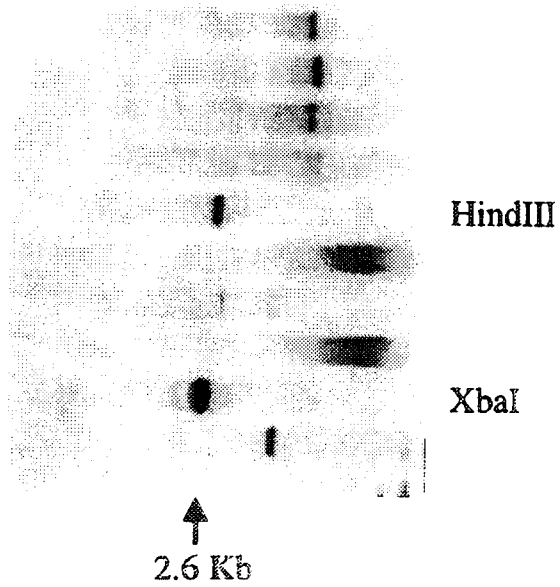
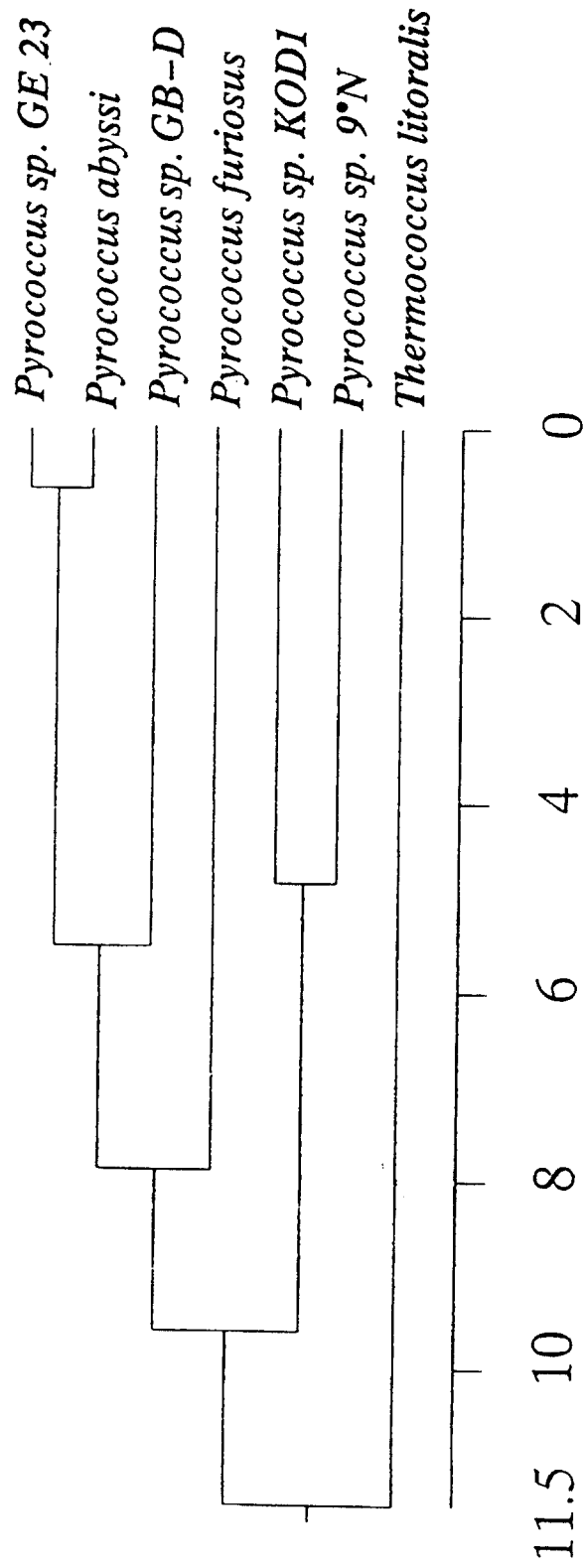


Fig.4



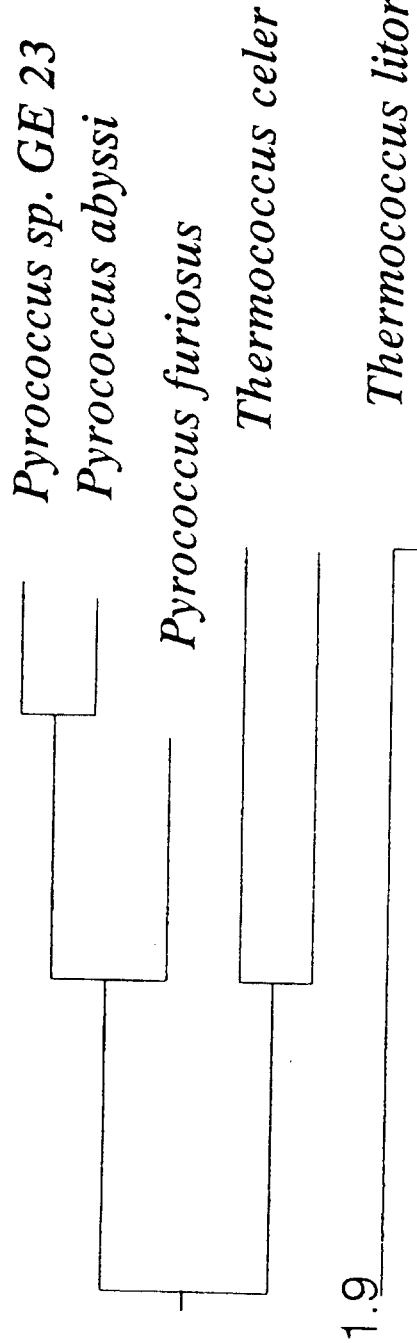


Fig.5

Fig.6a

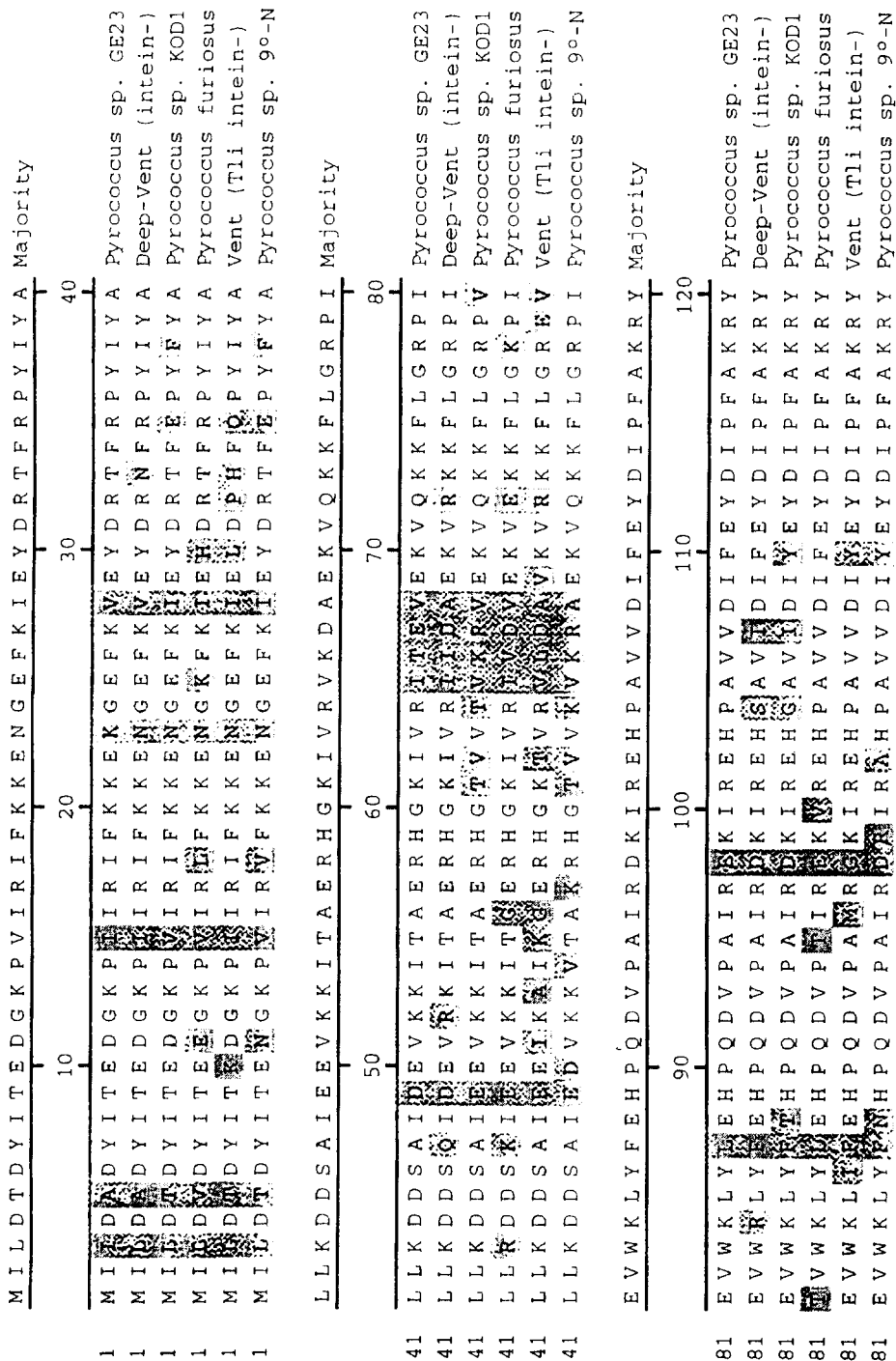
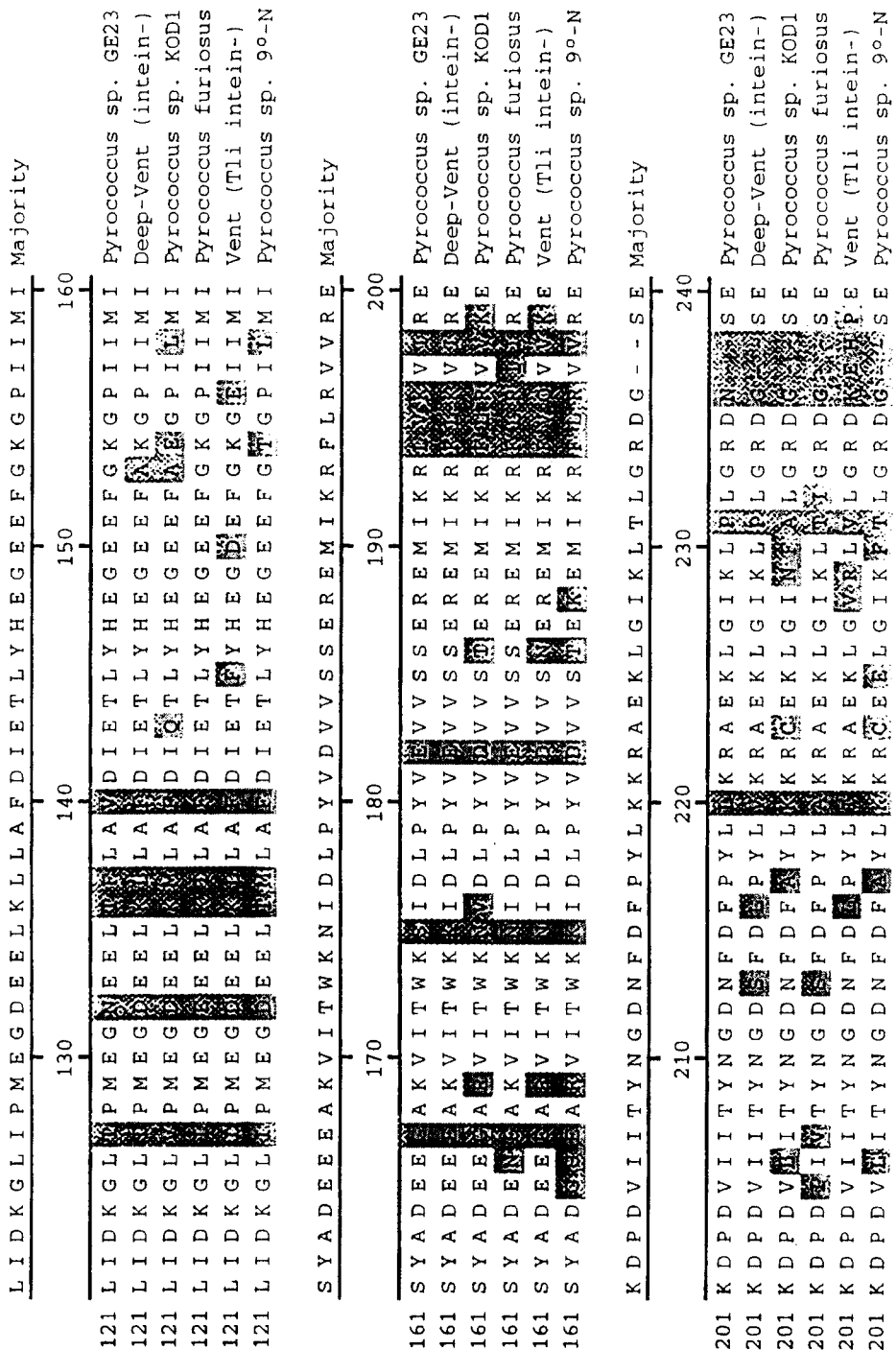


Fig. 6b



6/30

Fig.6c

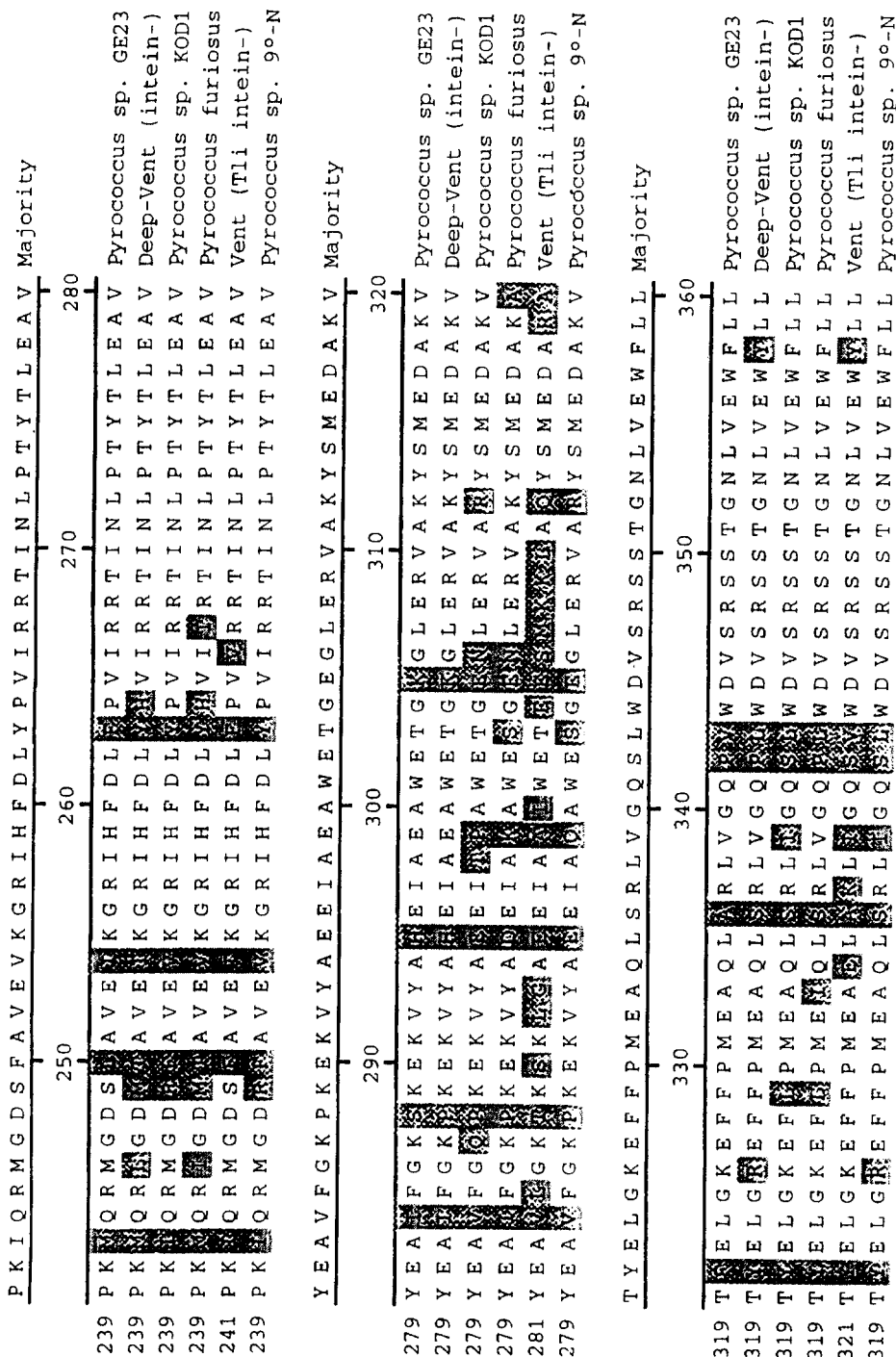


Fig.6d

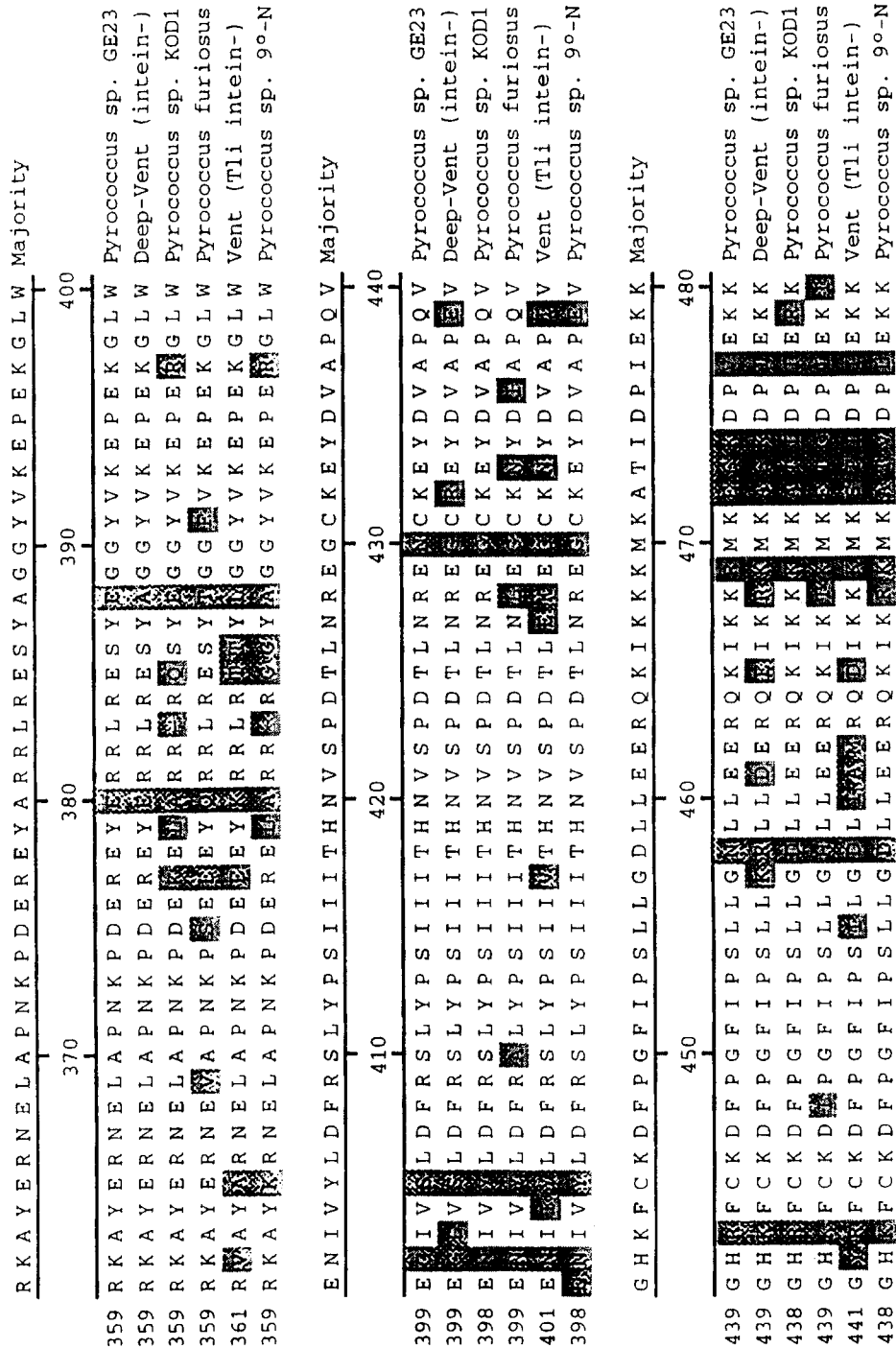
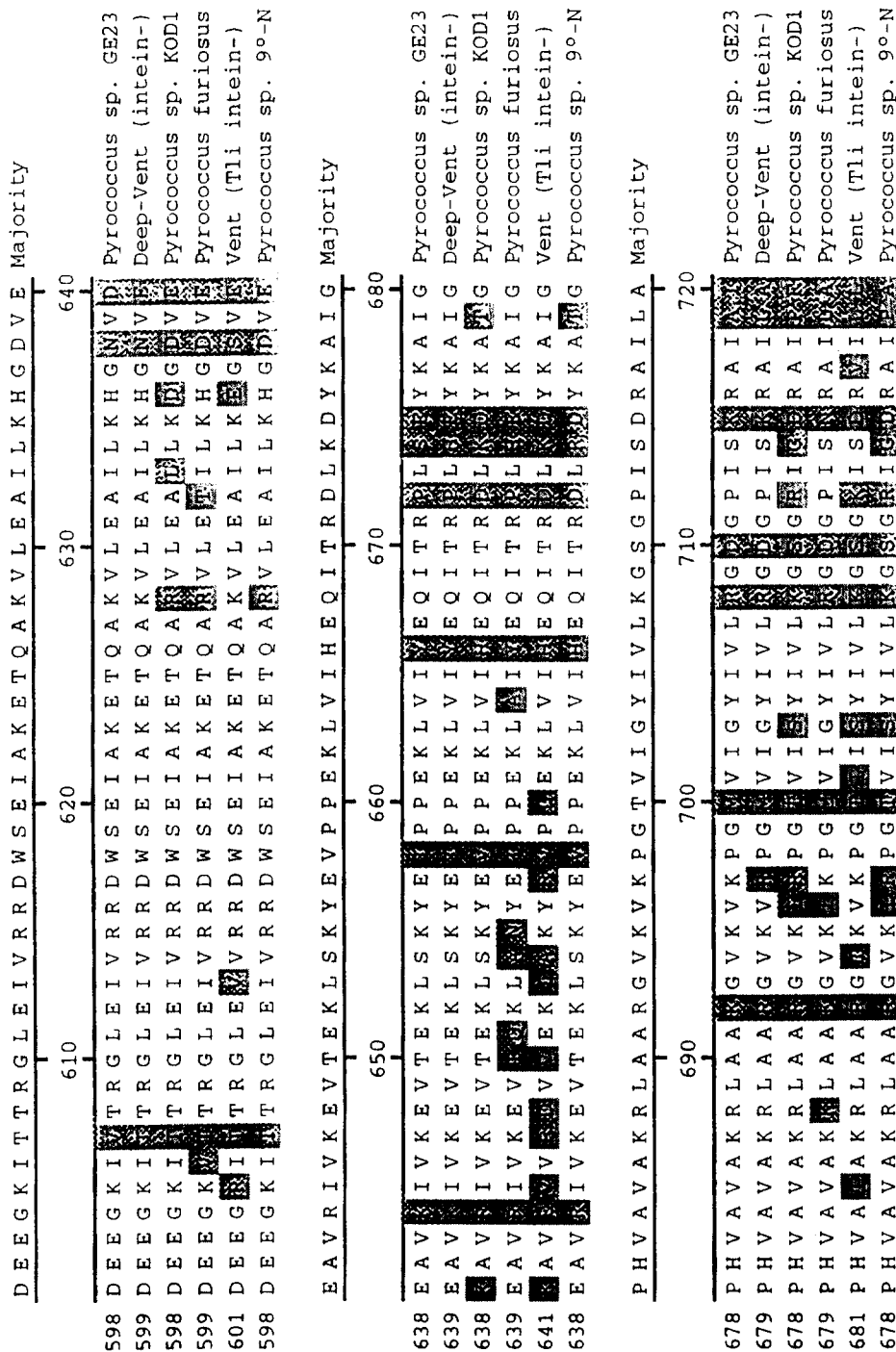


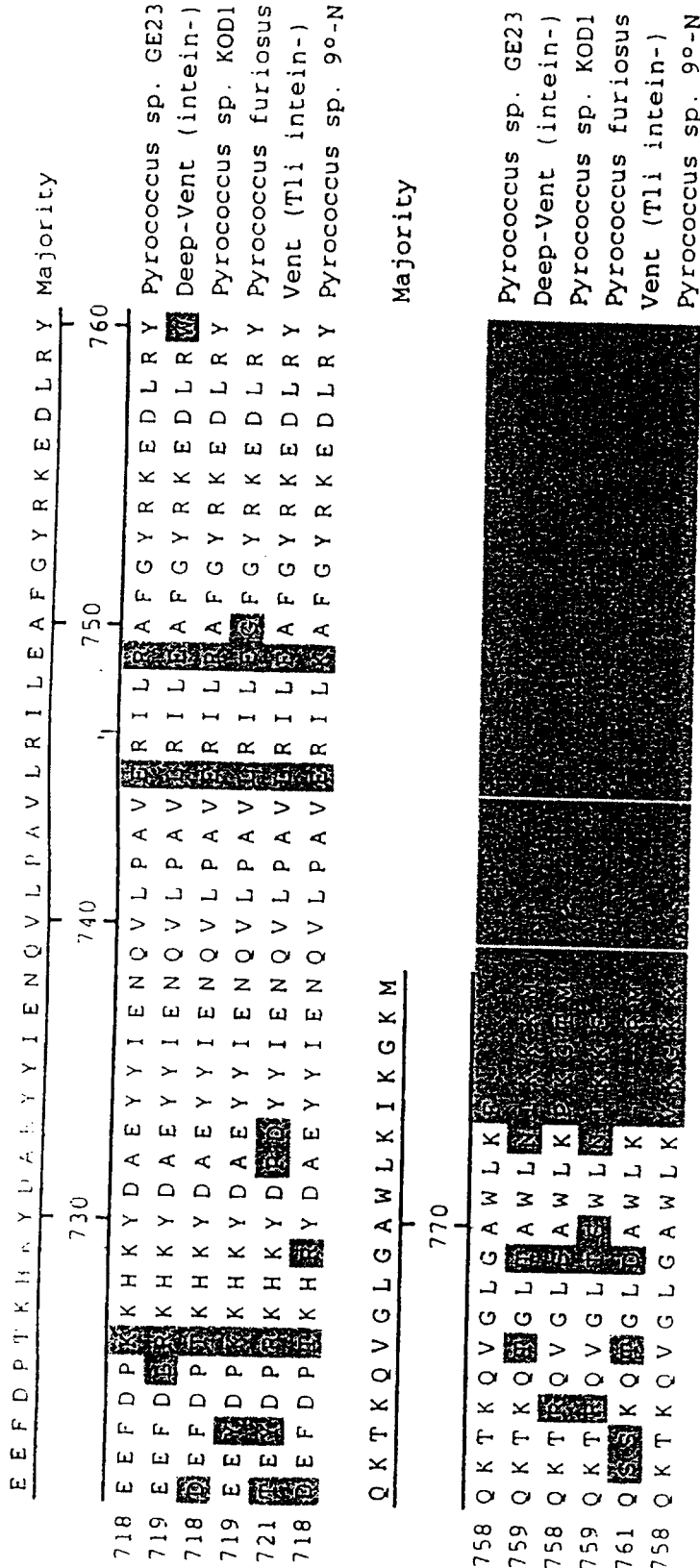


Fig.6f



10/30

Fig. 6g



Decoration 'Decoration #1': Shade (with black at 25% fill) residues that differ from the Consensus.

Decoration 'Decoration #2': Shade (with black at 25% fill) residues that differ from Pyrococcus sp. GE23.

Fig.7a

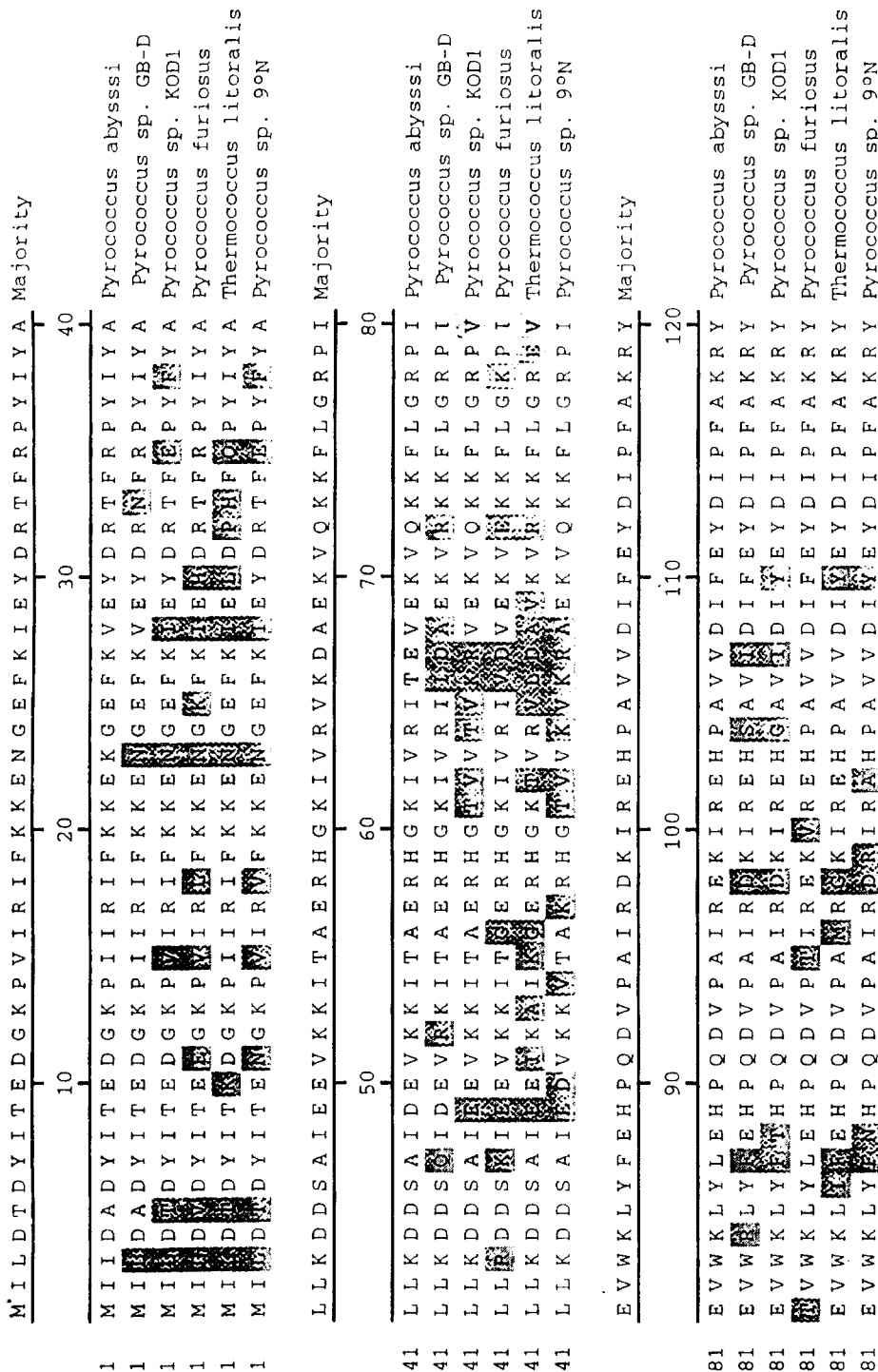




Fig.7c

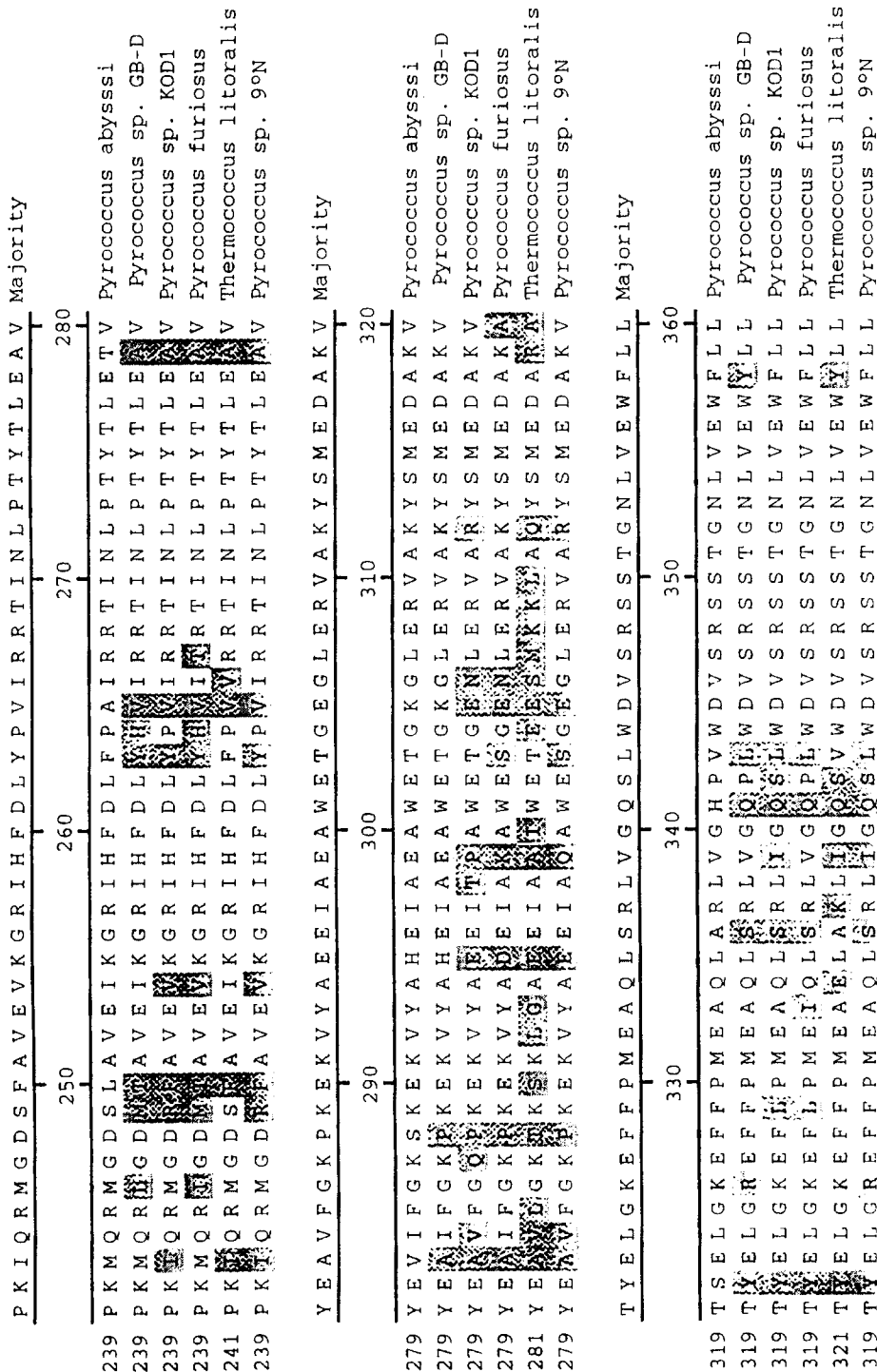


Fig.7d

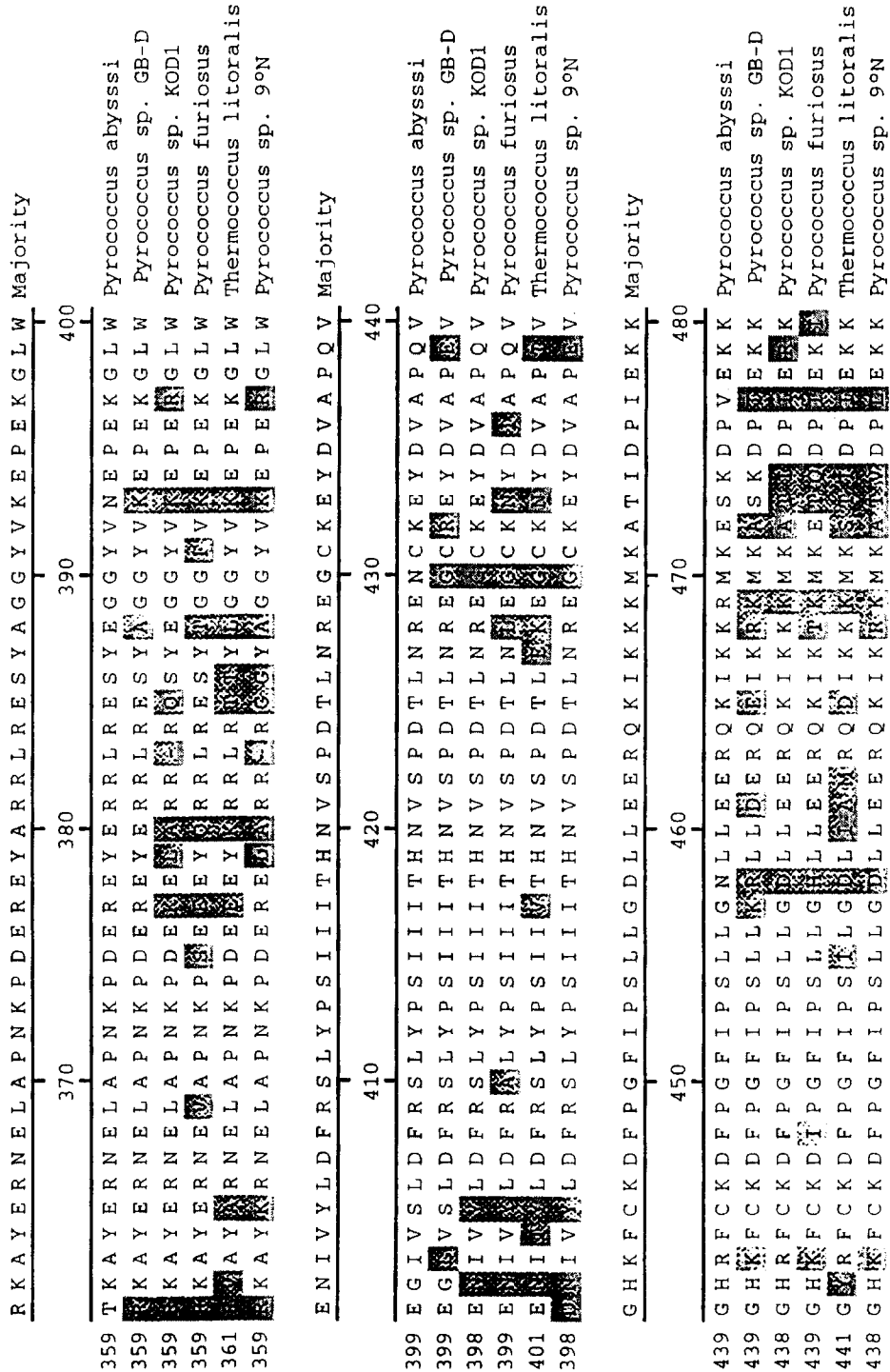


Fig.7e

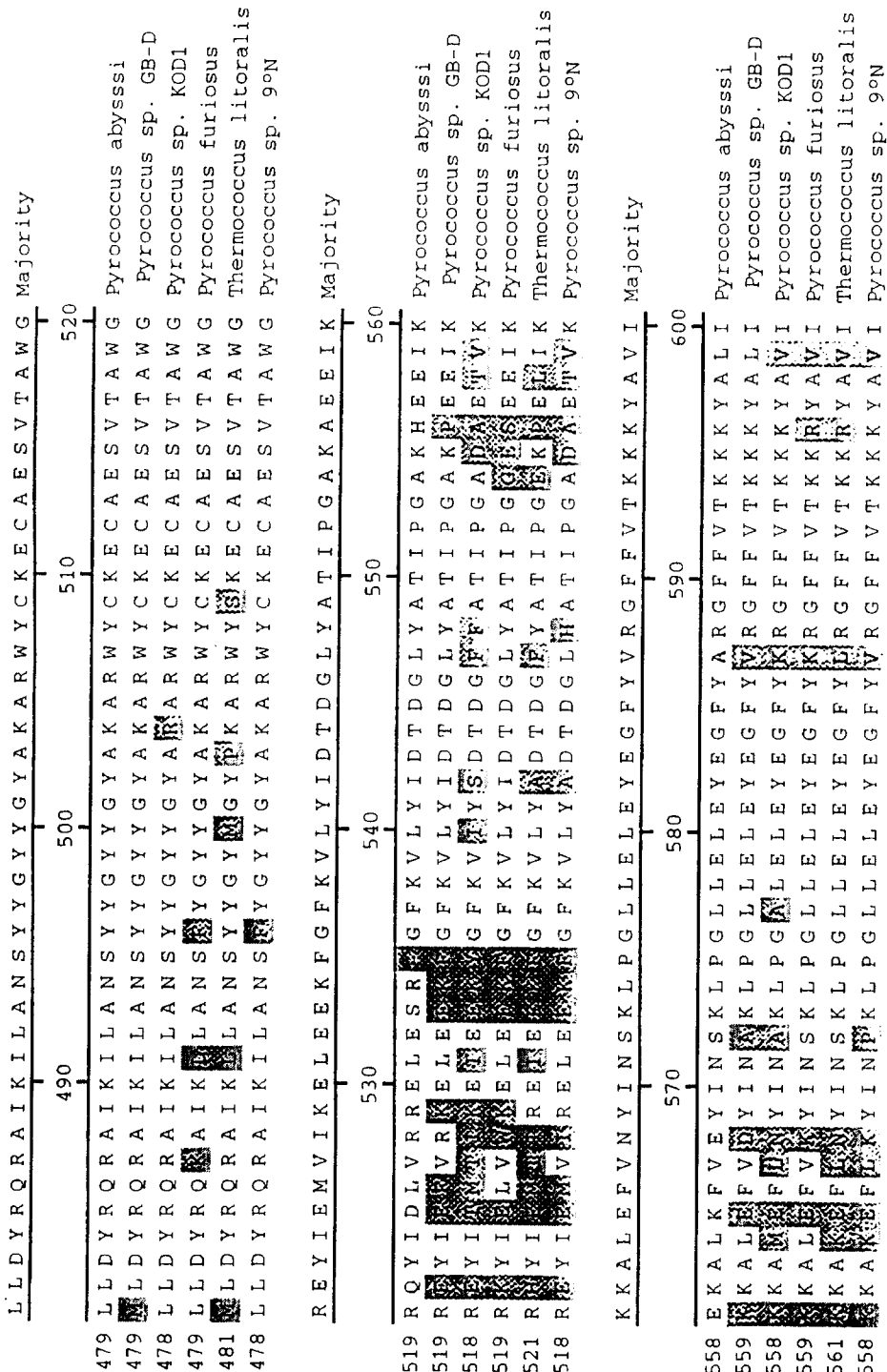


Fig.7f

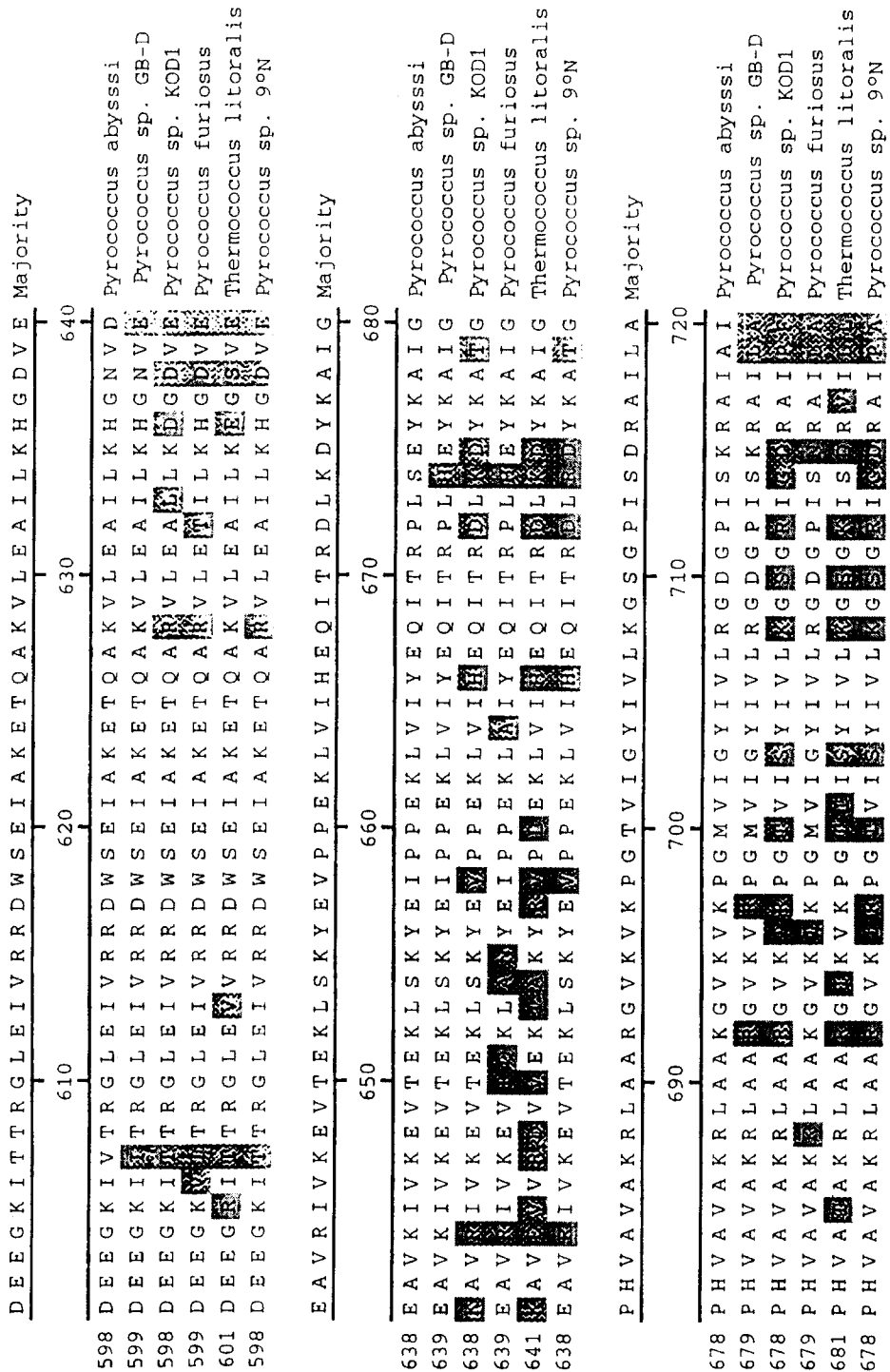
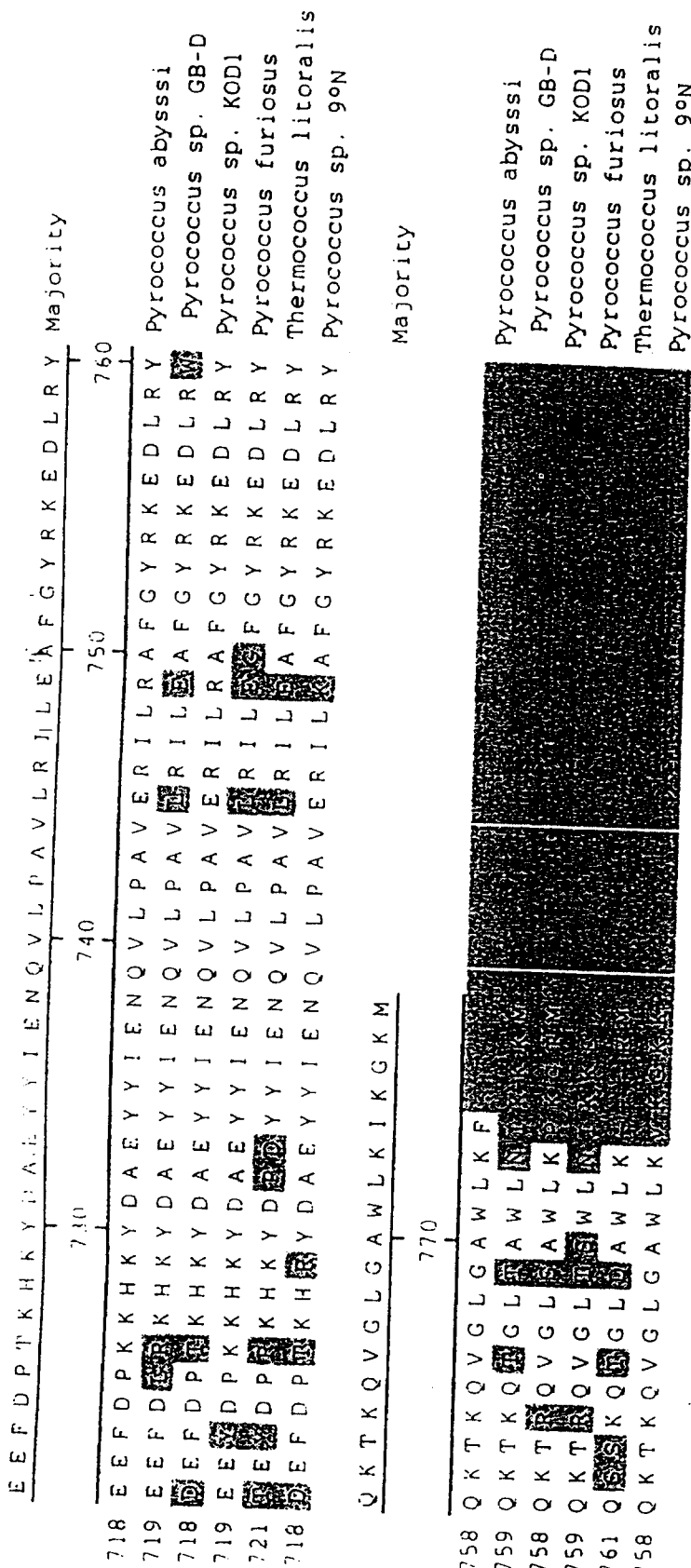


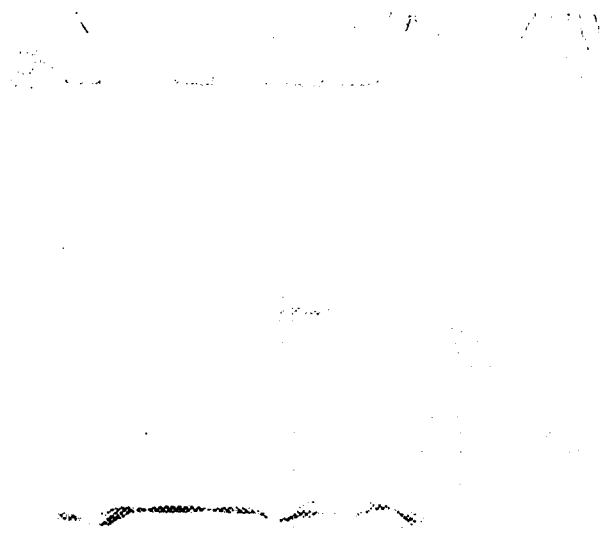
Fig. 7g



Decoration #1: Shade (with black at 25% fill) residues that differ from Pyrococcus abyssssi

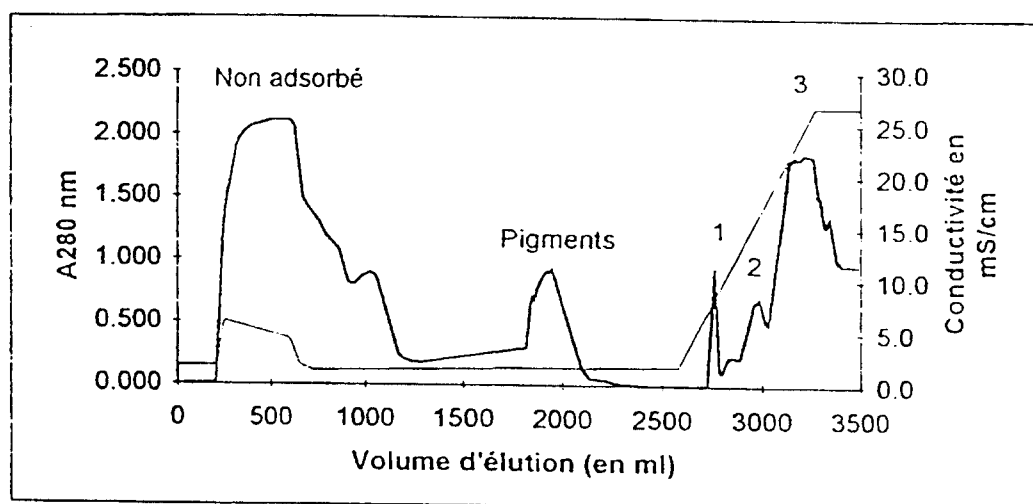
18/30

Fig. 8



19/22

Fig. 9



20/30

Fig. 10

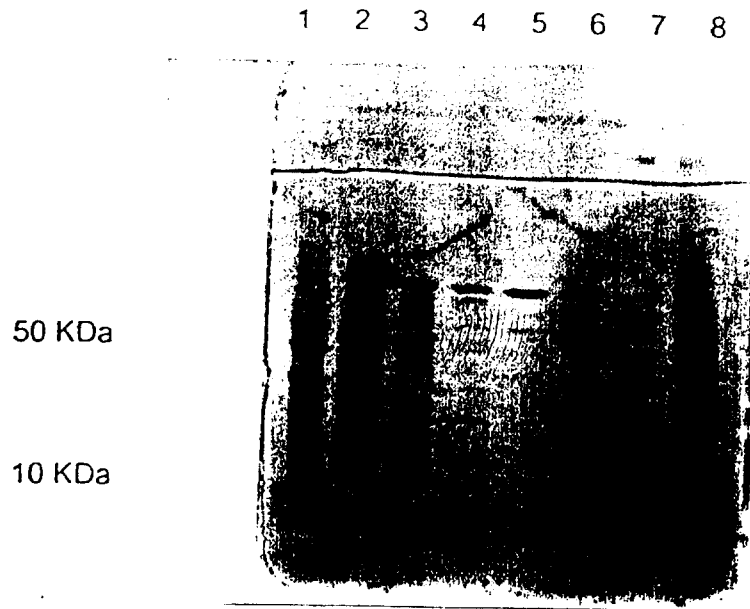
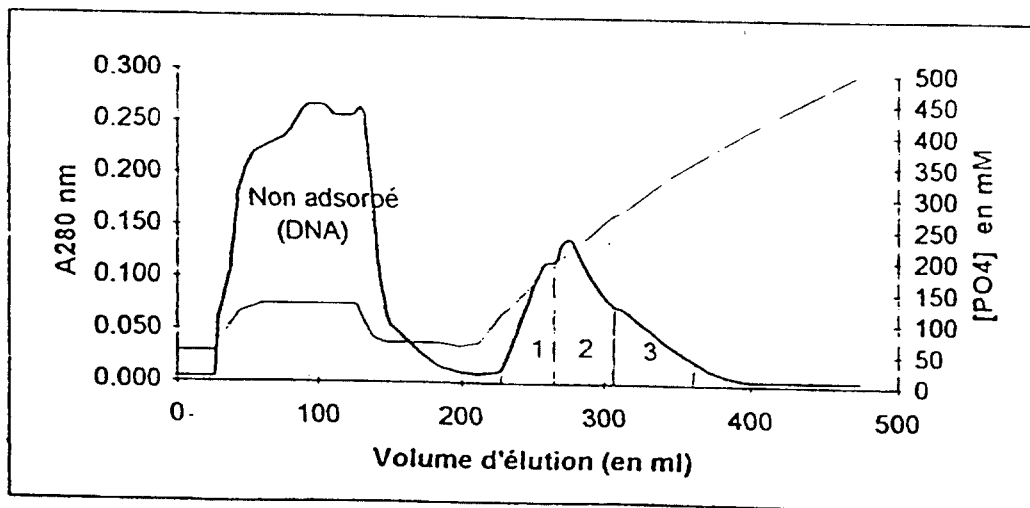
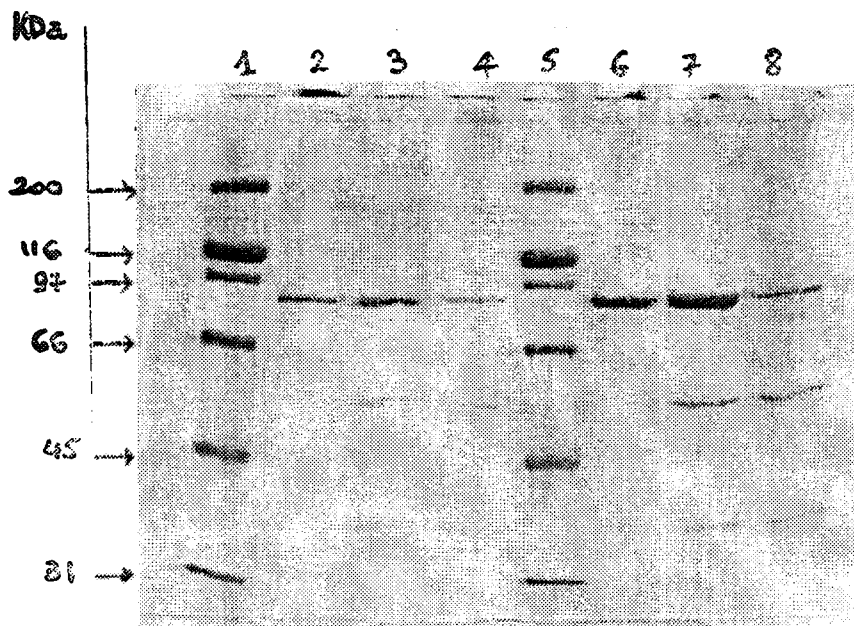


Fig. 11



21/22

Fig. 12



22/30

Fig.13

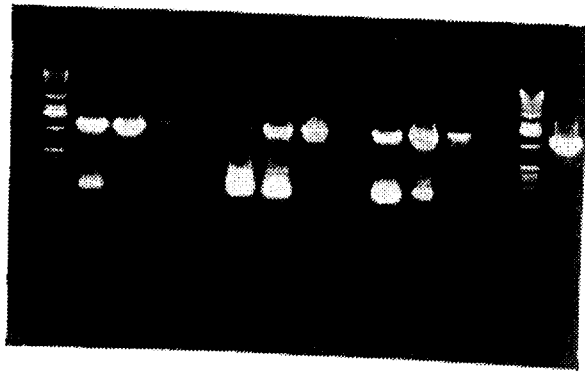


Fig. 14

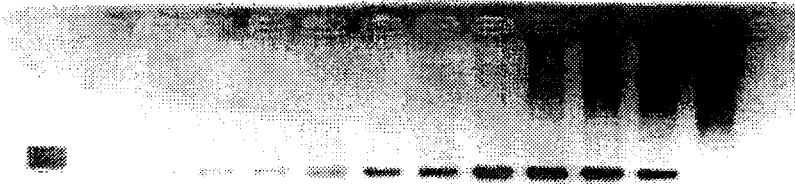


Fig. 15

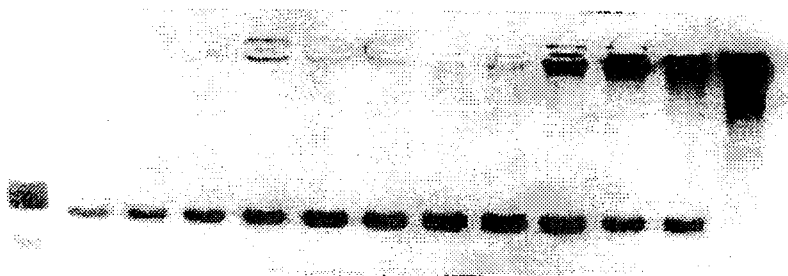


Fig.16

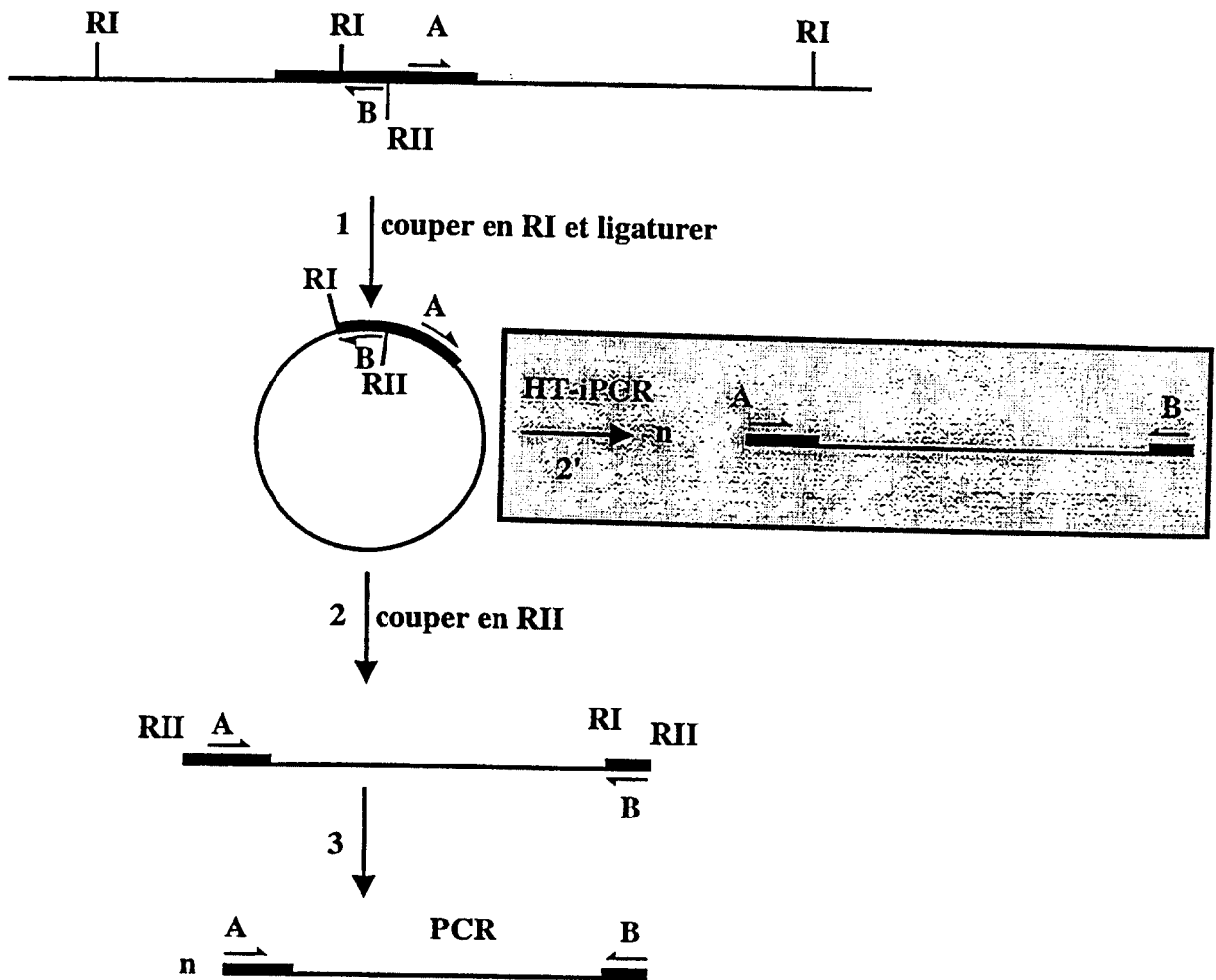
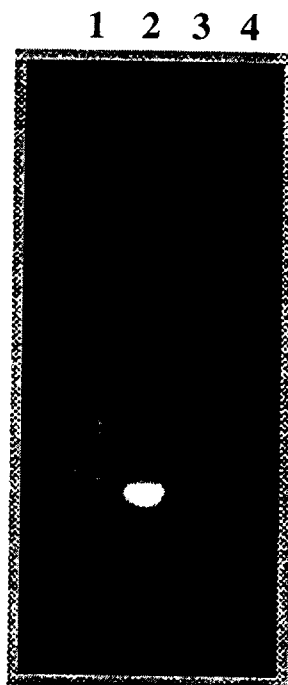


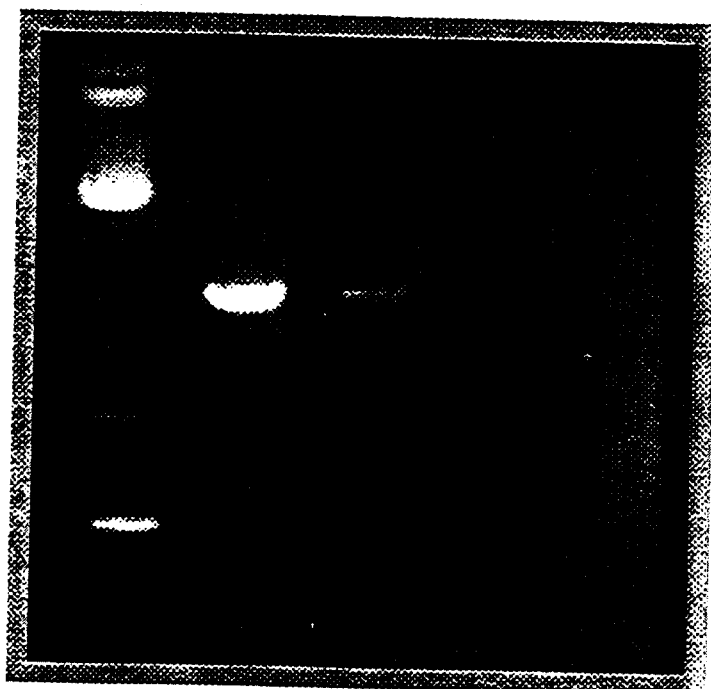
Fig.17



26/30

Fig. 18

1 2 3 4 5

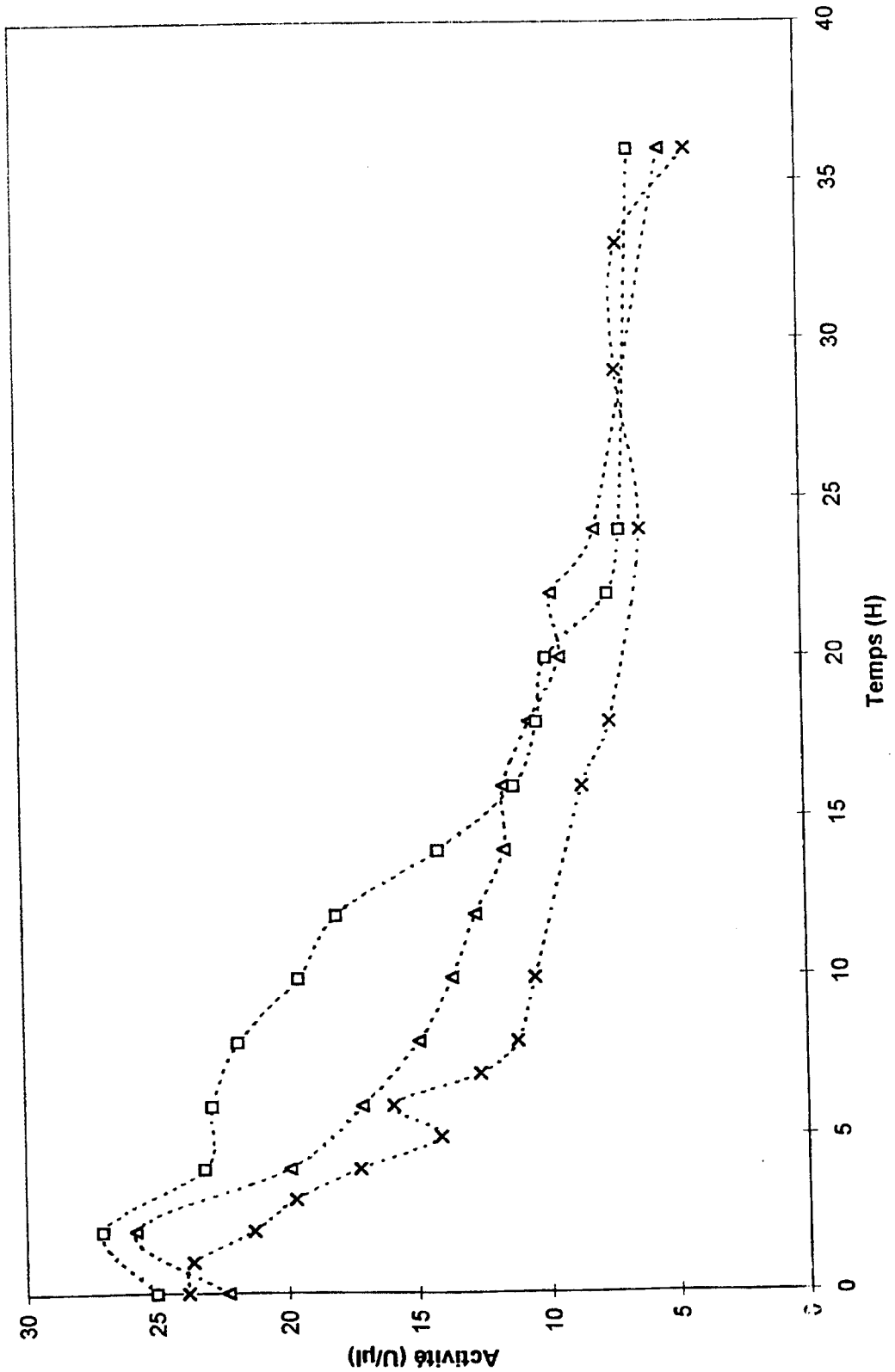


Activité Pab Pol

Fig.19

27/30

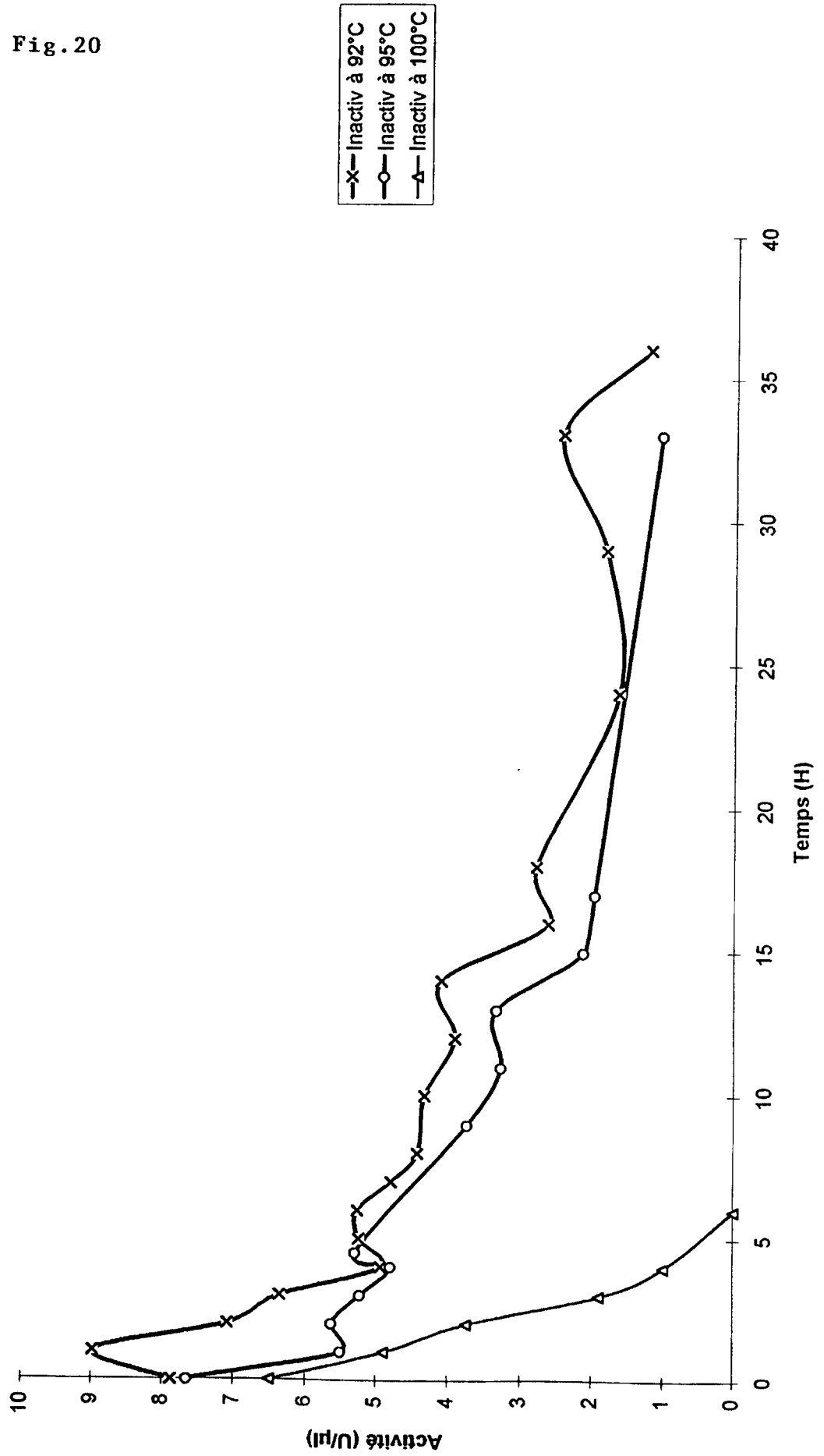
- ...△... Inactiv à 95°C
- ...□... Inactiv à 92°C
- ...x... Inactiv à 100°C



28/30

Fig. 20

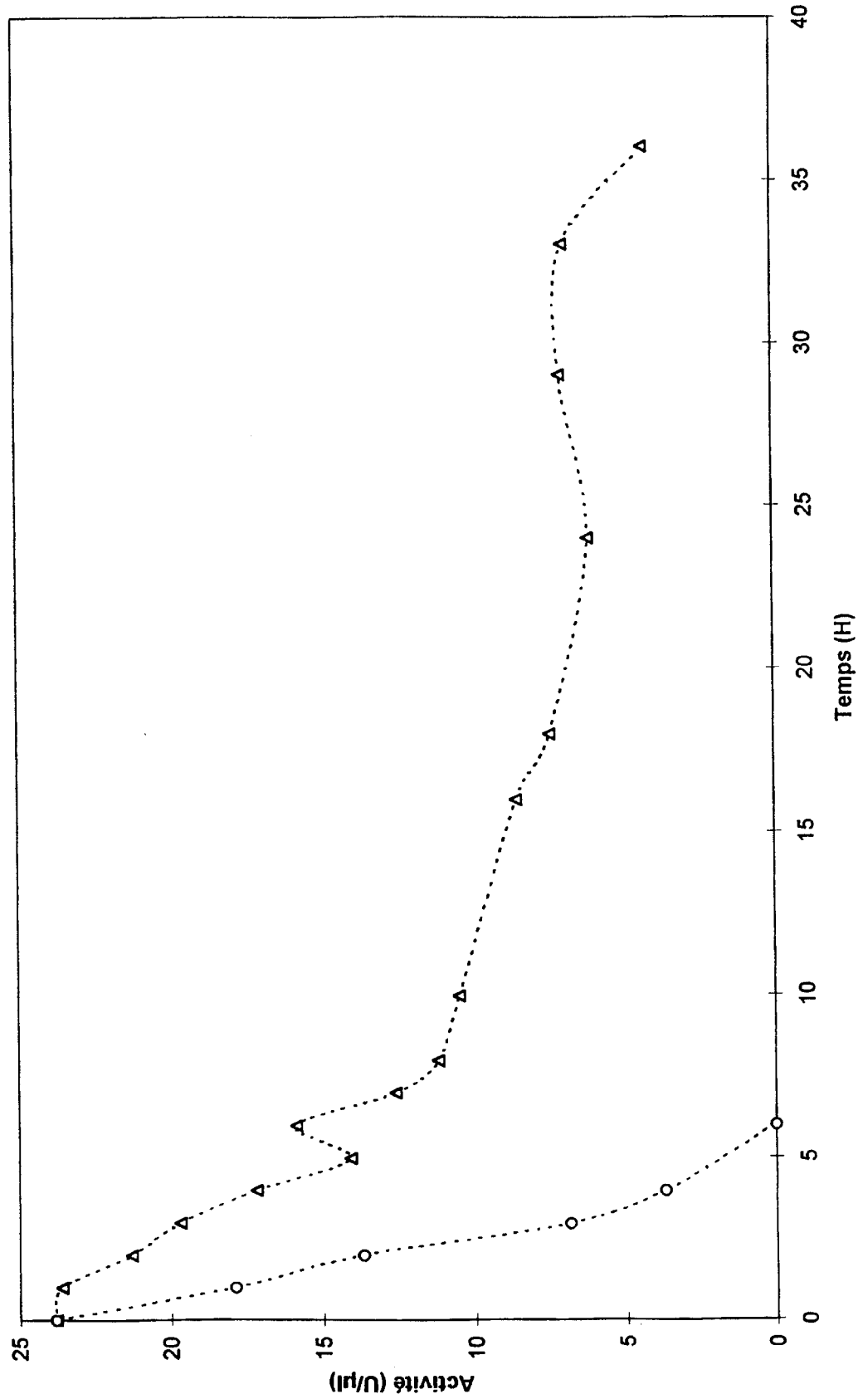
Activité Ppr Pol



29/30

Fif.21

Activité Pab - Ppr après inactivation à 100°C



Activité Pab Taq Ppr après inactivation à 92°C

Fig. 22

30/30

