



(86) Date de dépôt PCT/PCT Filing Date: 2003/09/19
(87) Date publication PCT/PCT Publication Date: 2004/04/01
(85) Entrée phase nationale/National Entry: 2005/03/18
(86) N° demande PCT/PCT Application No.: DK 2003/000609
(87) N° publication PCT/PCT Publication No.: 2004/027067
(30) Priorité/Priority: 2002/09/20 (PA 2002 01391) DK

(51) Cl.Int.⁷/Int.Cl.⁷ C12N 15/57, C12N 1/19, C12P 21/02

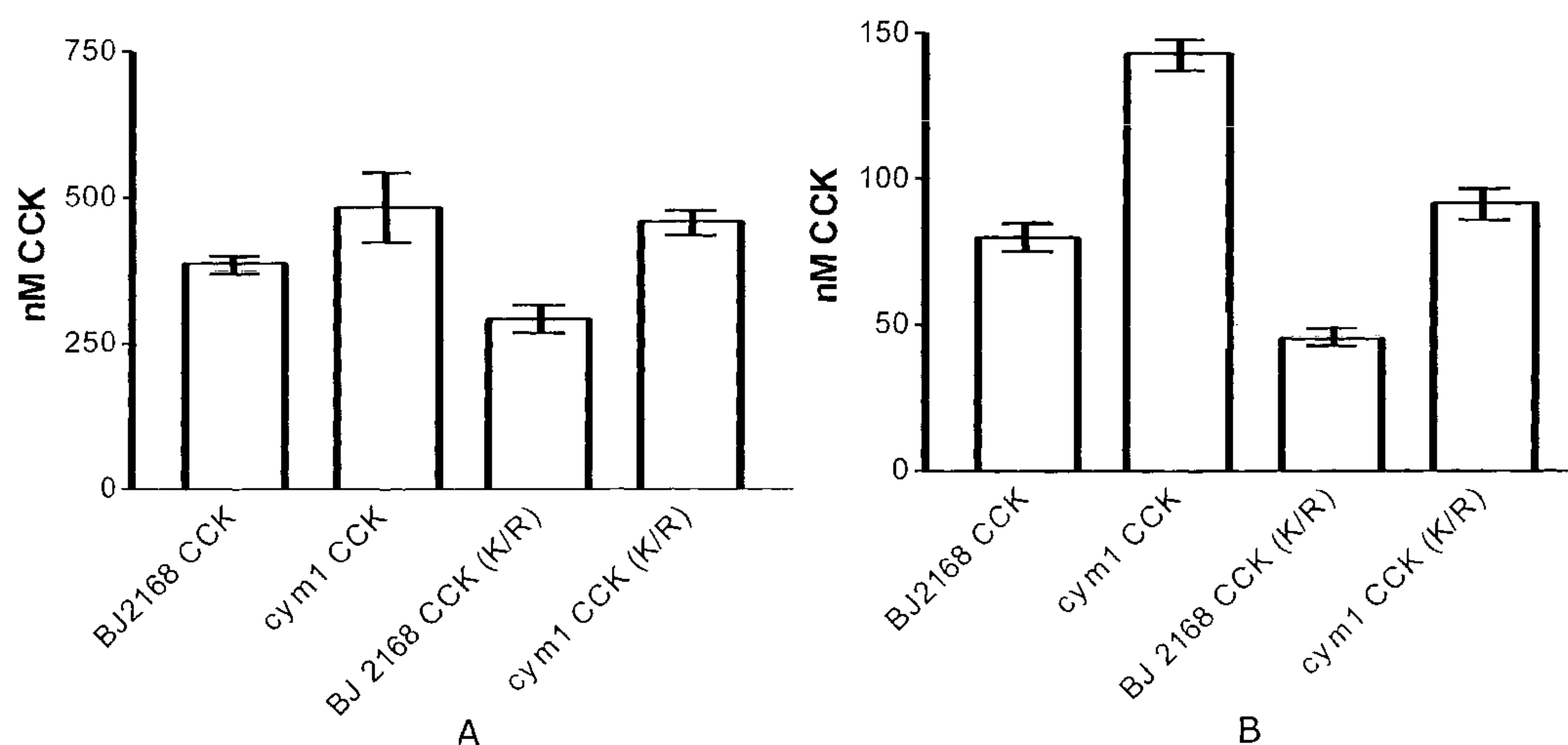
(71) Demandeur/Applicant:
CYM1P A/S, DK

(72) Inventeurs/Inventors:
JOHNSON, LARS, DK;
REHFELD, JENS F., DK;
JOHNSEN, ANDERS H., DK

(74) Agent: GOWLING LAFLEUR HENDERSON LLP

(54) Titre : PROCEDES POUR AUGMENTER LA PRODUCTION D'UN POLYPEPTIDE RECOMBINANT A PARTIR
D'UNE CELLULE HOTE

(54) Title: METHODS FOR INCREASING THE PRODUCTION OF A RECOMBINANT POLYPEPTIDE FROM A HOST
CELL



(57) **Abrégé/Abstract:**

The present invention provides a host cell comprising a nucleic acid sequence encoding a recombinant polypeptide in which the production of a naturally occurring metalloprotease comprising the sequenc HXXEH has been reduced or inhibited by genetic manipulation. The present invention also relates to methods for enhancing the production of a polypeptide from a cell by disrupting the synthesis or activity of the metalloprotease. In particular, the present invention relates to methods for enhancing the secretion of recombinant polypeptides from host cells such as, but not limited to, yeast and bacterial cells. The metalloprotease is a member of the pirtilysin subfamily of proteases, characterized by comprising the sequence HXXEH, where X is any amino acid.



(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
1 April 2004 (01.04.2004)

PCT

(10) International Publication Number
WO 2004/027067 A3

(51) International Patent Classification⁷: **C12N 15/57**,
C12P 21/02, C12N 1/19

(21) International Application Number:
PCT/DK2003/000609

(22) International Filing Date:
19 September 2003 (19.09.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PA 2002 01391 20 September 2002 (20.09.2002) DK

(71) Applicant (for all designated States except US): **CYM1P A/S** [DK/DK]; c/o Novi A/S, Bredgade 56, DK-1260 Copenhagen K (DK).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **JØNSEN, Lars** [DK/DK]; Vestergade 9, DK-4130 Viby Sjælland (DK). **REHFELD, Jens, F.** [DK/DK]; Esperance Allé 18, DK-2920 Charlottenlund (DK). **JOHNSEN, Anders, H.** [DK/DK]; Knudsvej 6, DK-4000 Roskilde (DK).

(74) Agent: **PLOUGMANN & VINGTOFT A/S**; Sundkrogs-
gade 9, P.O. Box 831, DK-2100 Copenhagen Ø (DK).

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

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

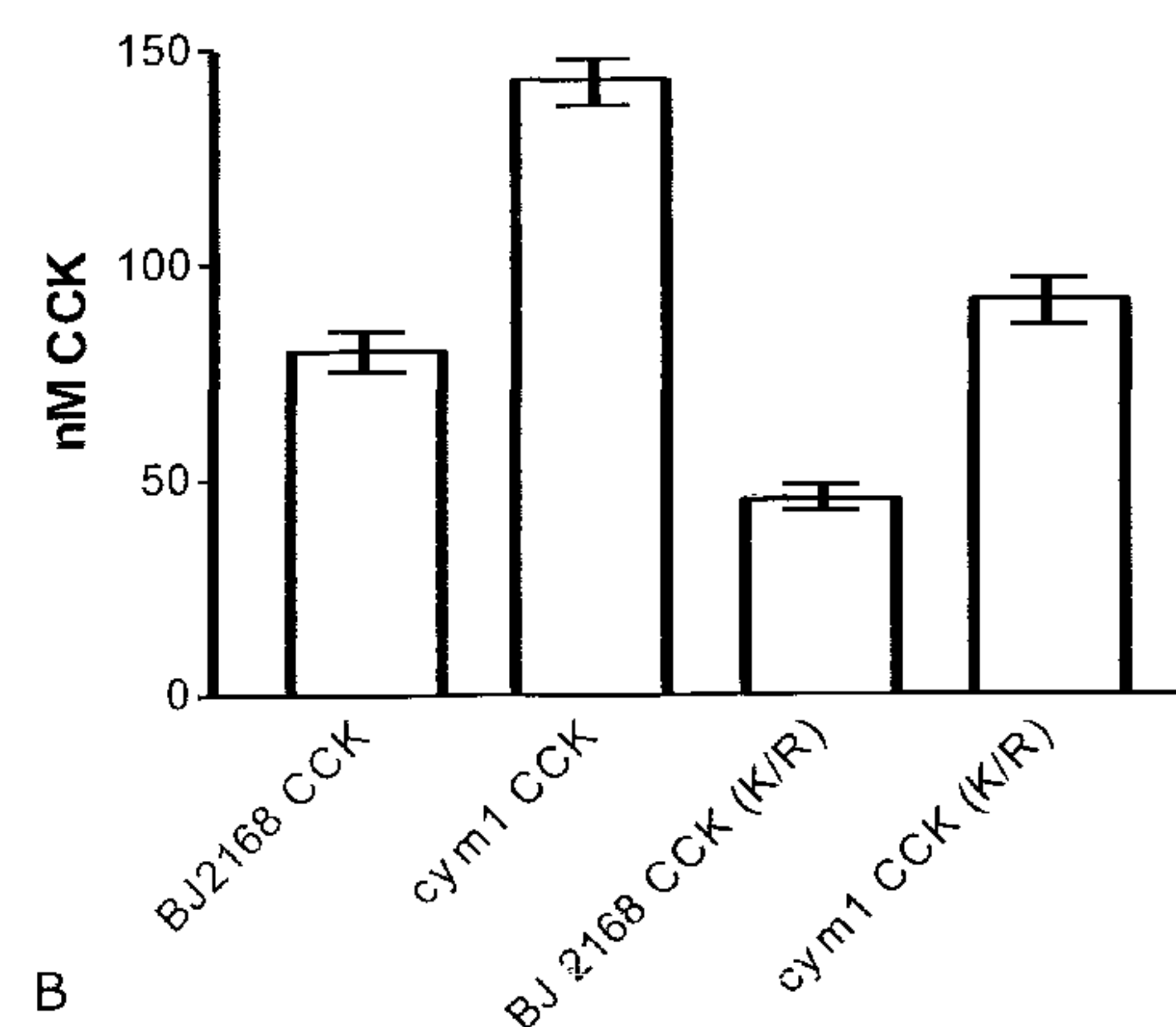
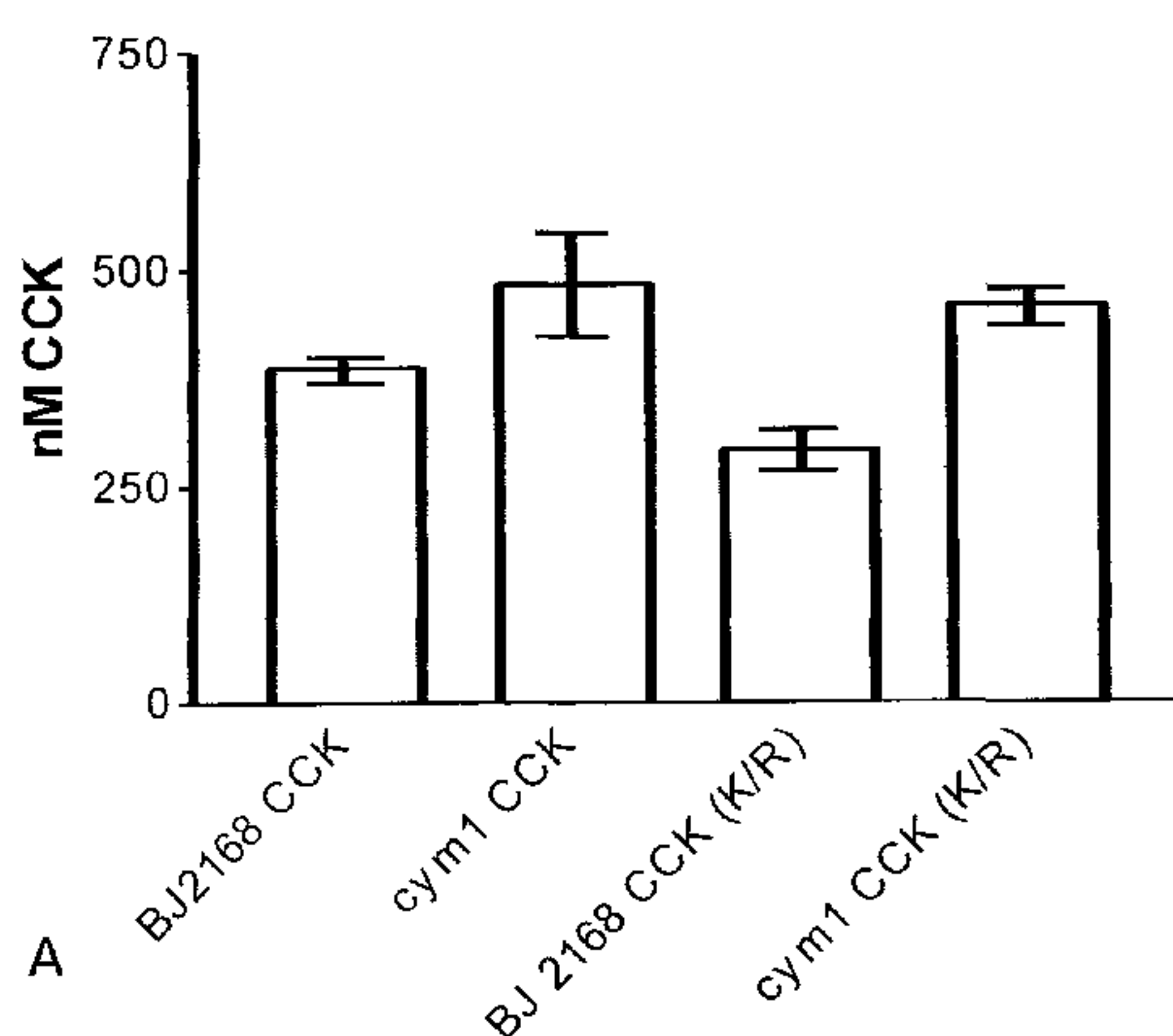
Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

(88) Date of publication of the international search report:
19 August 2004

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHODS FOR INCREASING THE PRODUCTION OF A RECOMBINANT POLYPEPTIDE FROM A HOST CELL



(57) **Abstract:** The present invention provides a host cell comprising a nucleic acid sequence encoding a recombinant polypeptide in which the production of a naturally occurring metalloprotease comprising the sequence HXXEH has been reduced or inhibited by genetic manipulation. The present invention also relates to methods for enhancing the production of a polypeptide from a cell by disrupting the synthesis or activity of the metalloprotease. In particular, the present invention relates to methods for enhancing the secretion of recombinant polypeptides from host cells such as, but not limited to, yeast and bacterial cells. The metalloprotease is a member of the ptilysin subfamily of proteases, characterized by comprising the sequence HXXEH, where X is any amino acid.

WO 2004/027067 A3

Methods for increasing the production of a recombinant polypeptide from a host cell

Field of the Invention

The present invention relates to methods for enhancing the production of a polypeptide
5 from a cell by disrupting the synthesis or activity of a metalloprotease from the clan ME (M16 family). In particular, the present invention relates to methods for enhancing the secretion of recombinant polypeptides from host cells such as, but not limited to, yeast and bacterial cells.

Background of the Invention

10 Cholecystokinin (CCK) is a vertebrate neuroendocrine peptide hormone that is expressed in both gut and brain tissues. The maturation of bioactive CCK peptides depends on post-translational tyrosine sulfation, endoproteolytic cleavages, exoproteolytic trimmings and carboxyterminal amidation. The endoproteolytic processing of the N-terminus varies with CCK-83, -58, -39, -33, -22, -8 and -5 being
15 identified. Most of the CCK peptides are synthesized after cleavage at a single Arg residue, however, CCK-22 requires processing after a single Lys residue.

Many recombinant polypeptides have been expressed in yeast as a fusion protein to the *Saccharomyces cerevisiae* α -factor prepro-peptide to direct secretion through the
20 secretory pathway. The best characterized yeast protease is the serine endoprotease, Kex2p (Fuller et al., 1989) which is involved in maturation of the α -mating pheromone and of killer toxin (Julius et al., 1984). Another yeast protease is Yps1p belonging to the yapsin family of glycosyl-phosphatidylinositol (GPI)-anchored aspartyl proteases, which is able to rescue mating deficiency when overexpressed in a *kex2* mutant (Egel-
25 Mitani et al., 1990). Expression of foreign proteins have shown that Yps1p and Yps2p contain endoprotease activity.

The use of host cells for the expression of recombinant polypeptides has greatly simplified the production of large quantities of commercially valuable polypeptides,
30 which otherwise are obtainable only by purification from their native sources. There is a varied selection of expression systems currently available from which to choose for the production of any given polypeptide, including eubacterial and eukaryotic hosts. One important factor in the selection of an appropriate expression system is the ability of the host cell to produce adequate yields of the polypeptide. However, a problem
35 frequently encountered is the high level of proteolytic enzymes produced by a given

host cell or in the culture medium. Accordingly, there is a need for further methods which enhance the production of a recombinant polypeptide from a host cell.

5 Metalloproteases are the most diverse of the four main types of protease, with more than 30 families identified to date. In these enzymes, a divalent cation, usually zinc, activates the water molecule. The zinc metalloproteases can be divided based on the zinc binding site into for example Zincins and Inverzincins (Hooper, N.M. 1994). The metal ion is held in place by amino acid ligands, usually three in number. The known metal ligands are His, Glu, Asp or Lys and at least one other residue which may play an
10 electrophilic role is required for catalysis,. Of the known metalloproteases, around half contain an HEXXH motif, which has been shown in crystallographic studies to form part of the metal-binding site.

A number of proteases dependent on divalent cations for their activity have been
15 shown to belong to a single family, peptidase M16. Included are insulinase, mitochondrial processing protease, pitrilysin, nardilysin and a number of bacterial proteins. These proteins do not share many regions of sequence similarity; the most noticeable is in the N-terminal section. This region includes a conserved histidine followed, two residues later, by a glutamate and another histidine. In pitrilysin, it has
20 been shown that this HXXEH motif is involved in enzymatic activity (Becker et al. 1992); the two histidines bind zinc and the glutamate is necessary for catalytic activity. The X can be any amino acid. Non active members of this family have lost from one to three of these active site residues.

25 It has previously been suggested that one could provide host cells and methods of producing proteins by expressing significantly reduced levels of a genetical modification in order to express significantly reduced levels of a metalloprotease containing an HEXXH motif in a filamentous fungal host cell, in e.g. US 5,861,280 (WO 98/12300).

30 Others have provided a protease deficient filamentous fungus which is characterised in that the filamentous fungus contains a site selected disruption of DNA that results in the filamentous fungus having reduced metalloprotease activity and isolated DNA sequences encoding a protein having metalloprotease activity, which is obtainable from a filamentous fungus (WO 97/46689). Again this metalloprotease contains an HEXXH
35 motif.

However, metalloproteases which can be reduced by a genetical modification in order to express significantly reduced levels of said metalloprotease in a non-filamentous

fungal host cell and other cells containing an motif other than HEXXH have never been described.

Summary of the Invention

Whilst investigating the role various proteases play in processing proCCK in
5 recombinant yeast, the present inventors surprisingly noted that deletion/disruption of *CYM1* enhanced recombinant polypeptide production and secretion. Furthermore, the present inventors have found that Cym1p belongs to a family of metalloproteases, the activity of which can be down-regulated to enhance the levels of recombinant polypeptide produced from a host cell.

10

Accordingly, in a first aspect the present invention provides a host cell comprising a nucleic acid sequence encoding a recombinant polypeptide in which the production of a naturally occurring metalloprotease comprising a sequence provided in SEQ ID NO:1 has been reduced or inhibited by genetic manipulation.

15

The host cell can be any cell which, in its native state, possesses the metalloprotease. Accordingly, the host cell can be a eukaryotic or prokaryotic cell. Examples of preferred eukaryotic cells include, but are not limited to, mammalian cells, plants cells and fungal cells. In a preferred embodiment, the host cell is a yeast cell. More preferably, the
20 yeast cell is selected from, but not limited to, the group consisting of: *Saccharomyces* sp. such as *Saccharomyces cerevisiae*, *Saccharomyces paradoxus*, *Saccharomyces mikatae*, *Saccharomyces bayanus*, *Saccharomyces castellii* and *Saccharomyces kluyveri*, *Schizosaccharomyces* sp. such as *Schizosaccharomyces pombe*, *Kluyveromyces lactis*, *Hansenula polymorpha*, *Pichia pastoris*, *Pichia methanolica*, *Pichia*
25 *kluyveri*, *Yarrowia lipolytica*, *Candida utilis*, *Candida cacaoi*, and *Geotrichum fermentans*.

Metalloproteases are among the hydrolases in which the nucleophilic attack on a peptide bond is mediated by a water molecule. This is a characteristic shared with
30 aspartic proteases, but in the metalloproteases a divalent metal cation, usually zinc, but sometimes cobalt or manganese, activates the water molecule. The metal ion is held in place by amino acid ligands usually 3 in number, the known metal ligands in metalloproteases are His, Glu, Asp or Lys residues.

35 Metalloproteases can be divided into two broad groups depending on the metal ions required for catalysis, and in the literature metalloproteases have been allocated into at least 8 different clans: MA, MB, MC, MD, ME, MF, MG and MH. Thus, illustrating the

complex diversity of this group of proteases. The allocation is based on different consensus sequences due to the ligand binding, and thus each family have different biological substrates and/or functions.

- 5 The metalloproteases which are to be down regulated according to the present invention is a member of the pepsin subfamily (ME) of proteases, characterized by comprising the sequence HXXEH (SEQ ID NO:1) where X is any amino acid. Presently more than 180 members are annotated in Swissprot to the ME clan. Thus, the most preferred embodiment, the metalloprotease comprises a consensus sequence provided
10 in SEQ ID NO:1. In another preferred embodiment, the metalloprotease comprises a consensus sequence provided in SEQ ID NO:2. Even more preferably, the metalloprotease comprises a consensus sequence provided in SEQ ID NO:3. In addition, it is preferred that the metalloprotease comprises SEQ ID NO:1 and a glutamic acid residue between 70 and 80 amino acids C-terminal of the second His
15 residue. Further, it is preferred that the metalloprotease comprises a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2 and SEQ ID NO:3 as well as a sequence selected from the group of:

- 20 i) any one of group consisting of SEQ ID NO's 4 to 15, and
ii) a sequence which is at least 80% identical to any one of SEQ ID NO's 4 to 15.

More preferably, the metalloprotease comprises SEQ ID NO:1 or SEQ ID NO:2 or SEQ ID NO:3 as well as a sequence selected from the group of:

- 25 i) any one of SEQ ID NO's 4 or 5, and
ii) a sequence which is at least 80% identical to any one of SEQ ID NO's 4 or 5.

Preferably, the metalloprotease comprises a sequence which is at least 85% identical,
30 such as at least 90% identical, such as at least 95% identical, and such as at least 99% identical to any one of SEQ ID NO's 4 to 15.

In a particularly preferred embodiment, the metalloprotease comprises a sequence as provided in SEQ ID NO:4, or a sequence at least 80% identical, such as at least 90%,
35 such as at least 95% and such as 99% identical, thereto.

The host cell can be genetically manipulated by any means known in the art as long as the production of the metalloprotease is reduced or inhibited when compared to a

parental host cell which has not been genetically manipulated. Such means of genetically manipulating the host cell include, but are not limited to; gene knockout, gene disruption, random or site directed mutagenesis, introduction of dominant-negative metalloproteases, RNA interference (RNAi) using dsRNA, catalytic nucleic acids (such as ribozymes and DNAzymes), and antisense nucleic acids. Preferably, the genetic manipulation acts directly upon the gene encoding the metalloprotease, the mRNA transcribed from the gene, or produces a protein that alters the activity of the metalloprotease such as a dominant negative mutant which competes with the metalloprotease for binding to a substrate but does not, for example, possess catalytic activity. However, the host cell may be genetically manipulated such that it indirectly affects the production or activity of the metalloprotease. For instance, the genetic manipulation can target a transcription factor involved in transcribing the mRNA encoded by the metalloprotease gene, thus at least reducing the levels of metalloprotease produced by the manipulated host cell.

Furthermore, the host cell may be further genetically manipulated such that it lacks at least one other naturally occurring protease of the host cell or has reduced activity for at least one other naturally occurring protease of the host cell. The protease can be any enzyme of which the inhibition increases the production of a recombinant polypeptide produced by the host cell. The protease can either be an endopeptidase, an aminopeptidase or a carboxypeptidase. Preferred proteases include serine proteases, aspartyl proteases, cysteine proteases and other metalloproteases.

In one embodiment, the host cell is a yeast cell and the other naturally occurring protease(s) is at least one protease encoded by any of the protease genes selected from the group consisting of; KEX2, YPS1 (previously known as YAP3), YPS2 (previously known as MKC7), YPS3, YPS6, YPS7, BAR1, STE13, KEX1, PRC1, PEP4 (also known as PRA1), APE1, APE2, APE3 and CPS1. Preferably, the host cell is a yeast cell and KEX2 production has been disrupted. Similar naturally occurring protease(s) within other host cells than yeast in addition to the metalloprotease specifically described here in could also be disrupted and/or genetically manipulated for an further additive enhancement.

The recombinant polypeptide can be any desired polypeptide which is capable of being produced in the host cell. The recombinant polypeptide can comprise a naturally occurring sequence or has been produced by the intervention of man (e.g. a mutant or truncation of a naturally occurring protein, or a fusion between at least two different

polypeptides). Typically, the recombinant polypeptide will be of commercial value, for example in the treatment of diseases.

The recombinant polypeptide can be any size. Typically, the recombinant polypeptide
5 will range in size from about 30 amino acids to about 4,500 amino acids. In one embodiment, the recombinant polypeptide is between about 30 to about 200 amino acids in length.

In at least some host cell expression systems for producing recombinant polypeptides,
10 it is desirable to direct the recombinant polypeptide to be secreted from the host cell. Thus, in a preferred embodiment, the nucleic acid comprises a sequence which encodes a signal for directing the recombinant polypeptide to be secreted from the host cell. Preferably, the signal is an N-terminal hydrophobic signal sequence. Such N-terminal hydrophobic signal sequences are known in the art, and include, for example but not
15 limited to, the leader sequence originating from the fungal amyloglucosidase (AG) gene such as galA - both 18 and 24 amino acid versions e.g. from *Aspergillus sp.*, the α -factor gene such as yeasts e.g. from *Saccharomyces sp.* and *Kluyveromyces sp.*, the P-factor of *Schizosaccharomyces sp.*, and the α -amylase gene from *Bacillus sp.* In one embodiment, the recombinant polypeptide is expressed as a fusion of an N-terminal
20 hydrophobic signal sequence and a second polypeptide sequence encoding the recombinant polypeptide which is from a different source than the signal sequence.

The nucleic acid encoding the recombinant polypeptide can be provided to the host cell using any technique known in the art. In one embodiment, the nucleic acid is inserted
25 into the genome of the host cell using, for example, homologous recombination based techniques. In another embodiment, the nucleic acid is transfected or transformed into the host cell in an expression vector which remains extrachromosomal. For example, the expression vector can be a plasmid or a virus. Further, it is preferred that the vector comprises a selectable marker which can be used to selectively propagate host
30 cells comprising the vector. Such selectable markers and the use thereof are also known in the art.

In a second aspect, the present invention provides a method of producing a recombinant polypeptide, the method comprising culturing a host cell according to the
35 second aspect under suitable conditions such that the recombinant polypeptide is produced, and recovering the recombinant polypeptide.

Since the proteolytic action arising from the metalloprotease has been reduced or inhibited, the method of the second aspect of the invention improves the stability of the recombinant polypeptide produced by the host cell.

- 5 In a preferred embodiment, the recombinant polypeptide is secreted from the host cell. Furthermore, it is preferred that the secreted protein is recovered during exponential growth of a culture comprising the host cell.

Preferably, the quantity of the recovered recombinant polypeptide is higher than if a
10 parental host cell was used. More preferably, the quantity of the recovered recombinant polypeptide is at 50% higher than if a parental host cell was used.

In a third aspect, the present invention provides a method of cleaving a polypeptide at a basic residue, the method comprising contacting the polypeptide, in the presence of a
15 divalent cation, with a metalloprotease comprising a sequence selected from the group of:

- i) any one of SEQ ID NO's 4 to 15, and
- ii) a sequence which is at least 80% identical to any one of SEQ ID NO's 4 to
15.

20

In a particularly preferred embodiment, the metalloprotease comprises a sequence as provided in SEQ ID NO:4, or a sequence at least 80% identical, such as at least 90%, such as at least 95% and such as 99% identical, thereto.

- 25 Preferably, the metalloprotease cleaves the polypeptide at the C-terminal side of an amino acid, or sequence of amino acids, selected from the group consisting of; Lys, Arg, ArgArg, LysLys, ArgLys and LysArg. Accordingly, it is preferred that the polypeptide comprises Lys, Arg, ArgArg, LysLys, ArgLys or LysArg. Other sequence requirements may also be necessary for cleavage, however, these can readily be
30 determined by routine experimentation.

Preferably, the divalent cation is selected from the group consisting of: Zn^{2+} , Co^{2+} and Mn^{2+} .

- 35 The method of the third aspect can be performed *in vivo*, within a recombinant host cell producing the metalloprotease, or *in vitro* in suitable reaction conditions. Considering the present disclosure, the skilled addressee could readily perform the method of the third aspect. An example of an *in vitro* system for cleaving a polypeptide with the

defined metalloprotease is provided herein. In this instance, the polypeptide is contacted with the metalloprotease provided in a crude yeast cell extract in 0.1 M NaH₂PO₄ (pH 4.5) and in the presence of 1 mM Mn²⁺ and 1 mM bestatin. In another example, the metalloprotease can be recombinantly produced as a fusion protein with a
5 suitable "tag", such as a His-tag, which enables easy purification of the fusion protein. Preferably, such a "tag" is removed (for example by enzymatic cleavage) before the metalloprotease is exposed to the substrate polypeptide.

In a fourth aspect, the present invention provides a method of identifying an agent that
10 inhibits the activity of a metalloprotease comprising a sequence provided in SEQ ID NO:1, the method comprising the steps of:

- a) incubating the metalloprotease with the agent, in the presence of a divalent cation and a suitable substrate;
- 15 b) determining the activity of the metalloprotease on the substrate;
- c) comparing the activity obtained in step b) with the activity of a control sample that has not been incubated with the agent; and
- 20 d) selecting an agent that inhibits the activity of the metalloprotease.

The substrate can be any polypeptide that can be cleaved by the metalloprotease and the cleavage event detected. One example disclosed herein is the use of CCK as a substrate, where the cleavage event is detected by the production of CCK-22. Similar
25 assays can readily be developed for other substrates.

In a preferred embodiment of the fourth aspect, the metalloprotease comprises a sequence as provided in SEQ ID NO:4, or a sequence at least 80% identical, such as at least 90%, such as at least 95% and such as 99% identical, thereto.

30

In a fifth aspect, the present invention provides a method of producing a recombinant polypeptide, the method comprising culturing a host cell comprising a nucleic acid sequence encoding a recombinant polypeptide under suitable conditions such that the recombinant polypeptide is produced, and recovering the recombinant polypeptide
35 wherein said culturing comprises the presence of an inhibitor of a metalloprotease comprising a sequence provided in SEQ ID NO:1.

Preferably, the inhibitor is identified according to a method of the fourth aspect.

As will be apparent, preferred features and characteristics of one aspect of the invention may be applicable to other aspects of the invention.

- 5 Throughout this specification the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element, integer or step, or group of elements, integers or steps, but not the exclusion of any other element, integer or step, or group of elements, integers or steps.
- 10 The invention will hereinafter be described by way of the following non-limiting Figures and Examples.

Brief Description of the Accompanying Drawings

Figure 1. Cholecystikinin expression construct. PreproMf α 1p-proCCK fusion protein
15 with the amino acid sequences around the fusion site and of the primary cleavage sites shown. The major forms of secreted CCK with their N- and C-terminal amino acid residues are shown below.

Figure 2. CCK-22 maturation in cells and media as a function of cell growth. BJ2168
20 expressing preproMf α 1p-proCCK fusion protein. The CCK-22 immuno-reactivity was measured by RIA using Ab 89009 and total CCK content measured with Ab 89009 after tryptic cleavage. Open circles represent the fraction of secreted CCK-22, whereas the intracellular fraction of CCK-22 is presented as filled triangles. The cell growth was measured by OD₆₀₀ (open squares). The data represent mean of two independent
25 experiments.

Figure 3. Chromatographic analyses of normal and K $\rightarrow$ A mutated CCK secreted from BJ2168. Media from yeast transformed with pRS426 preproMf α 1p-proCCK and pRS426 preproMf α 1p-proCCK (K $\rightarrow$ A) were subjected to G-50 gel chromatography and the CCK
30 immuno-reactivity was measured with Ab 7270 specific for Gly extended CCK (A and C) and Ab 89009, which is specific for the N-terminus of CCK-22 (B and D).

Figure 4. *In vitro* protease assay including inhibitors and activators. The fraction of CCK-22 was calculated from the immuno-reactivity using Ab 89009 divided by the total
35 amount of mature and N-terminal extended CCK-22 measured with Ab 89009 after trypsin treatment. A, Effect of different inhibitors. B, Protease reactivation by addition

of 1.2 mM divalent metal ions to extracts where the activity had been inhibited with 1 mM EDTA. The data represent mean $\pm$ SD of three independent experiments.

Figure 5. Protease reactivation by Zn^{2+} and Mn^{2+} . *In vitro* protease assays performed with cell extracts from LJY123, where the activity was inhibited with 1 mM EDTA (filled squares) and reactivated by addition of 1.2 mM Mn^{2+} (open circles) or 1.2 mM Zn^{2+} (filled circles). The activity was measured as the fraction of matured CCK-22 after 30, 60 and 120 min incubation. The fraction of CCK-22 was calculated from the immuno-reactivity using Ab 89009 divided by the total amount of mature and N-terminal extended CCK-22 measured with Ab 89009 after trypsin treatment. The data are represented by mean $\pm$ SD (n=3).

Figure 6. Extracellular CCK-22 maturation by members of the yapsin family. The ability of intact cells to process extracellular CCK was analysed as described under "Experimental Procedures" for BY4705 and the isogenic *yps1*, *yps1 yps3* and *yps1 yps2 yps3* strains. The fraction of CCK-22 was calculated from the immuno-reactivity using Ab 89009 divided by the total amount of total CCK measured with Ab 89009 after trypsin treatment. The data are represented by mean $\pm$ SD (n=4). Statistics were performed using unpaired t test as described in experimental procedures (*** = $P < 0.001$, ** = $P < 0.01$ and * = $P < 0.05$).

Figure 7. Increased proteolysis following *Cym1p* overexpression. *In vitro* protease assays performed with cell extracts from BJ2168 transformed with an empty pRS425 plasmid (A) and with pRS425 containing *CYM1* (B). The CCK-22 immuno-reactivity was measured over time using Ab 89009 (filled squares and circles) and the total amount of mature and N-terminal extended CCK-22 measured with Ab 89009 after trypsin treatment (open squares and circles). The data represent mean $\pm$ SD of three independent experiments.

Figure 8. Effects of *KEX2* and *CYM1* deletions on proCCK secretion and CCK-22 maturation. Yeast cells transformed with the proCCK expression construct were harvested during exponential phase and the media collected. The intracellular (A) and extracellular (B) amount of total CCK was measured with Ab 89009 after trypsin treatment. The fraction of intracellular (C) and secreted (D) CCK-22 was calculated as the immuno-reactivity measured with Ab 89009 before tryptic cleavage divided with the total amount of CCK measured in (A) and (B). The *kex2*, *cym1* and *kex2 cym1* strains are isogenic to BJ2168. The data are given as mean $\pm$ SD (n=4). Statistics were

performed using unpaired t test (** = $P < 0.01$ and * = $P < 0.05$). The stars enclosed in brackets are a comparison between the *kex2* and *kex2 cym1* strain.

Figure 9. Intracellular degradation of CCK depends on Cym1p cleavage to CCK-22.

5 Expression of wild type CCK, preproMf α 1p-proCCK, and the CCK mutant, preproMf α 1p-proCCK (K $\rightarrow$ A) in BJ2168 and a *CYM1* disrupted strain isogenic to BJ2168. The cells were sedimented during exponential growth and the total amount of CCK (hatched bars) was measured after trypsin and carboxypeptidase B treatment with Ab 7270 specific for Gly-extended CCK. The amount of mature Gly-extended (white bars), which
10 is dependent on translocation into the secretory pathway, Kex2p and carboxypeptidase activity is measured as the immuno-reactivity using Ab 7270 before tryptic cleavage and carboxypeptidase B treatment. The data are given as mean $\pm$ SD (n=3).

Figure 10. Aspartyl proteases involved in the maturation of CCK-22. Expression of wild

15 type proCCK in BY4705 and the isogenic yapsin deletion strains of *YPS1*, *YPS2* and *YPS3*. The intra- (A) and extracellular (B) fraction of synthesised CCK-22 was measured during exponential growth. The fraction of mature CCK-22 was calculated as the immuno-reactivity measured with Ab 89009 before tryptic cleavage divided with the total amount of CCK. The data are represented by mean $\pm$ SD (n=3). Statistics were
20 performed using unpaired t test as described in experimental procedures (** = $P < 0.001$, * = $P < 0.01$ and * = $P < 0.05$).

Figure 11. Cym1p processing C-terminally to both Lys and Arg residues. *CYM1*

deletion enhance the amount of secreted CCK more than two fold of both wild type CCK
25 and the Lys⁶¹ $\rightarrow$ Arg⁶¹ mutant. Expression of wild type CCK, preproMf α 1p-proCCK, and the CCK mutant, preproMf α 1p-proCCK (Lys⁶¹ $\rightarrow$ Arg⁶¹) in BJ2168 and a *CYM1* disruptant isogenic to BJ2168. Yeast cells were harvested during exponential phase and the media collected. The intracellular (A) and extracellular (B) amount of total CCK was measured with Ab 89009 after trypsin treatment. The data are given as mean $\pm$ SD (n=3).

30

Figure 12. Secreted proCCK fragments identified by mass spectrometry. The CCK-
numbers refer to C-terminal amidated CCK. The molecular masses are given as
monoisotopic values except for * which denote average value. Strain A, vacuolar
protease-deficient strain (BJ2168), and B, the isogenic strain with *KEX1 KEX2*
35 disruptions (LJY22).

Figure 13. Model for the production of the C-terminally extended CCK (A) and GLP2 (B). Expression of these fusion peptides should be performed in a *sec61* mutant, or the pre-sequence of the α -mating factor should be removed to avoid translocation into the ER. The amino acid sequences around the fusion sites are shown. Underlined are the N- and C-terminal amino acids of the Gly-extended CCK-22 and GLP1.

Figure 14

A. The preproMf α 1p-proBNP expression construct.

B. The preproMf α 1p-KREAEA-BNP-32 expression construct.

C. The preproMf α 1p-KR-BNP-32 expression construct.

Figure 15

A. The preproMf α 1p- proBNP expression construct transformed in BJ2168, LJY430 (*cym1* mutant), LJY431 (*yps1* mutant) and LJY432 (*cym1 yps1* double mutant). Media was analysed from cells that have reached stationary phase using Ab 98192 that is specific for the N-terminus of proBNP. The *cym1*, *yps1*, and *cym1yps1* strains are isogenic to BJ2168. The data are given as mean $\pm$ SD (n=3). Statistics were performed using unpaired t test as described in experimental procedures (*** = $P < 0.001$, ** = $P < 0.01$ and * = $P < 0.05$). The stars enclosed in brackets are comparisons are between the wild type strain, BJ2168 vs. *cym1*, BJ2168 vs. *yps1* and *yps1* vs. *yps1cym1*.

(ns, = not significant).

B. Analysis of proCCK fragments secreted from a *cym1* mutant. Media containing 10 pmole proBNP was applied to Superdex 200 column on a Akta purifier system. The proBNP content in the collected fractions were measured using Ab. 98192, that is specific for the N-terminus of proBNP.

Key to the Sequence Listing

SEQ ID NO:1 – Consensus sequence for pitrilysin proteases.

SEQ ID NO:2 - Consensus sequence for at least some pitrilysin proteases.

SEQ ID NO:3 - Consensus sequence for at least some pitrilysin proteases.

SEQ ID NO:4 – *Saccharomyces cerevisiae* Cym1p (Swissprot Accession No. P32898).

SEQ ID NO:5 - *Schizosaccharomyces pombe* C119.7 (Swissprot Accession No. O42908).

SEQ ID NO:6 - *Clostridium perfringens* HypA protein (Swissprot Accession No. Q46205).

5 SEQ ID NO:7 - *Borrelia burgdorferi* protein BB0228 (Swissprot Accession No. O51246).

SEQ ID NO:8 - *Caenorhabditis elegans* C05D11.1 protein (Swissprot Accession No. P48053).

SEQ ID NO:9 - *E. coli* protease III (Swissprot Accession No. P05458).

SEQ ID NO:10 - Rat NRD convertase (Swissprot Accession No. P47245).

10 SEQ ID NO:11 - Human insulysin (Swissprot Accession No. P14735).

SEQ ID NO:12 - *Arabidopsis thaliana* CPE (Genbank Accession No. T03302).

SEQ ID NO:13 - Human metalloprotease I (GenBank Accession No. AAH01150) in part, the full sequence (Swissprot Accession No. O95204).

SEQ ID NO:14 - *Bacillus subtilis* zinc protease ymxG (GenBank Accession No. Q04805).

15 SEQ ID NO:15 - *Mycobacterium tuberculosis* zinc protease Rv2782c (GenBank Accession No. O33324).

SEQ ID NO's 16 to 42 - Oligonucleotides.

SEQ ID NO's 43 to 52 - Sequences provided in Figure 12.

SEQ ID NO's 53 to 55 - Sequences provided in Figure 1.

20 SEQ ID NO's 56 to 65 - Oligonucleotides.

SEQ ID NO:66 - Consensus sequence for at least some pitrilysin proteases.

SEQ ID NO:67 - Consensus sequence for at least some pitrilysin proteases.

SEQ ID NO:68 - Consensus sequence for at least some pitrilysin proteases.

25 **Detailed Description of the Invention**

The present invention provides a host cell useful for the expression of a polypeptide, said cell being genetically manipulated in order to at least produce reduced levels of a defined metalloprotease, when compared to the parental cell. The host cell will thus be able to express a protein of interest in higher quantity due to the proteolytic action of
30 the metalloprotease has been reduced or inhibited which improves the stability of the protein of interest.

By the method of the invention, the proteolytic action of the metalloprotease has been reduced or inhibited, thereby improving the stability of the product obtained.

35

Thus, one embodiment of the present invention relates to a host cell useful for the expression of a protein of interest, wherein said cell has been genetically modified in

order to express significantly reduced levels of a metalloprotease comprising a HXXEH motif (SEQ ID NO 1) compared to the corresponding non-modified cell when cultured under identical conditions.

- 5 The metalloproteases which are to be down regulated according to the present invention do not share many regions of sequence similarity; the most noticeable is in the N-terminal section. This region includes a conserved histidine followed two residues later by a glutamate and another histidine. In pitrilysin, it has been shown that this HXXEH motif is involved in enzymatic activity; the two histidines bind zinc and the
- 10 glutamate is necessary for catalytic activity. Non active members of this family have lost from one to three of these active site residues.

The metalloprotease family which are to be down regulated according to the present invention is presently classified as member of clan ME, family M16. This family is

15 currently divided into 4 subfamilies:

M16A comprising pitrilysin

M16B comprising mitochondrial processing peptidase beta-subunit (*Saccharomyces cerevisiae*)

- 20 M16C comprising eupitrilysin (*Homo sapiens*)

M44 comprising vaccinia virus-type metalloindopeptidase (vaccinia virus).

Sequence alignments of these proteins show several sequence similarities. These sequence similarities are highly conserved and can be used to distinguish members of

25 this family from non-members.

Among such sequence similarities several individual amino acids are highly conserved and are easily recognisable in specific positions navigated from the HXXEH motif.

- 30 Thus, one embodiment of the present invention relates to a host cell, wherein the metalloprotease comprises a glutamic acid residue between 70 and 80 amino acids C-terminal of the second His residue in the HXXEH motif.

A further embodiment of the present invention relates to a host cell, wherein the

35 metalloprotease comprises a glycine residue 3 amino acids N-terminal of the first His residue in the HXXEH motif.

Another embodiment of the present invention relates to a host cell, wherein the metalloprotease comprises a glycine residue 5 amino acids C-terminal of the second His residue in the HXXEH motif.

- 5 One further embodiment of the present invention relates to a host cell wherein the metalloprotease comprises a lysine residue 8 amino acids C-terminal of the second His residue in the HXXEH motif.

- Also, one embodiment of the present invention relates to a host cell, wherein the
10 metalloprotease comprises a tyrosine residue 9 amino acids C-terminal of the second His residue in the HXXEH motif.

- Furthermore, the present invention relates to a host cell, wherein the metalloprotease comprises a proline residue 10 amino acids C-terminal of the second His residue in the
15 HXXEH motif.

- Among the sequence similarities several regions of amino acids are also highly conserved and are easily recognised. Thus, in a presently preferred embodiment the invention relates to a host cell wherein the metalloprotease comprises the consensus
20 sequence SEQ ID NO 2.

In another presently preferred embodiment, the invention relates to a host cell, wherein the metalloprotease comprises the consensus sequence SEQ ID NO 3.

- 25 In a presently most preferred embodiment, the invention relates to a host cell, wherein the metalloprotease comprises a NAXTXXXXT motif between 20 and 30 amino acids C-terminal of the second His residue in the HXXEH motif.

- In a presently another preferred embodiment, the invention relates to a host cell,
30 wherein the metalloprotease comprises the consensus sequence SEQ ID NO 66-68. One embodiment of the present invention relates to a host cell useful for the expression of a protein of interest, wherein said cell has been genetically modified in order to express significantly reduced levels of a metalloprotease which is at least 80% identical to the any of SEQ ID NO: 4-15, as compared to a parental cell.

35

In the present context, the term "protein of interest" relates to any of the numerous naturally native occurring extremely complex substances such as but not limited to proteins, enzymes and/or antibodies that consist of amino acid residues joined by

peptide bonds. It is an object of preferred embodiments of the present invention to provide such native proteins which are products of the host cell itself and/or heterologous proteins, fusion proteins, recombinant proteins, eukaryotic proteins, prokaryotic proteins, lysosomal proteins, vacuolar proteins, precursor proteins, zymogene proteins, prepro-proteins, and secreted proteins.

Preferred embodiments of the claimed method are advantageous due to the higher production of the protein of interest, thus any increase of the amount of the protein of interest when produced in a host cell modified as described herein compared to the amount produced in the corresponding non-modified cell when cultured under identical conditions are within the scope of the present invention.

One assay in which a skilled addressee could evaluate enhanced production of the protein of interest in a host cell modified as described here in and compared to the amount produced in the corresponding non-modified cell, is by culturing the two different host cells under identical condition, and measure the amount of produced protein of interest by radio-immune assay using an antibody specific for the protein of interest. One such assay is describe in more detail in the examples of the present description.

20

One embodiment of the present invention relates to a host cell, wherein the total amount of the protein of interest is increased at least 5% compared to the corresponding non-modified cell when cultured under identical conditions, such as at least 10% compared to the corresponding non-modified cell when cultured under identical conditions, such as at least 20% compared to the corresponding non-modified cell when cultured under identical conditions, such as at least 50% compared to the corresponding non-modified cell when cultured under identical conditions, such as at least 100% compared to the corresponding non-modified cell when cultured under identical conditions, such as at least 200% compared to the corresponding non-modified cell when cultured under identical conditions, or even at least 1000% or compared to the corresponding non-modified cell when cultured under identical conditions.

In the present context, the term "host cell" relates to any cell capable of producing the protein of interest. Thus, in one preferred embodiment, the host is a prokaryotic cell. In another preferred embodiment, the host cell is a eukaryotic cell, such as but not limited to a filamentous fungal cell and a non-filamentous fungal cell. Non limiting examples hereof are a strain of *Saccharomyces*, especially *Saccharomyces cerevisiae*.

All the features described herein relating to the methods of the present invention are also applicable as embodiments relating to the host cells, and vice versa.

- 5 The method described in the present application relates to the production of a protein of interest in a host cell, wherein said host cell has been genetically modified in order to express significantly reduced levels of a metalloprotease which is at least 80% identical to the SEQ ID NO: 4 as compared to a parental cell, when cultured under identical conditions, comprising

10

a) introducing into the host cell a nucleic acid sequence encoding the protein of interest;

15

b) cultivating the host cell of step (a) in a suitable growth medium for production of the protein of interest and

c) isolating the protein of interest.

One embodiment of the present invention relates to a method for the production of a
20 protein of interest in a host cell, wherein the host cell has been genetically modified by a method selected from the group comprising gene knock-out, gene disruption, random or site directed mutagenesis, introduction of dominant-negative metalloproteases, RNA interference (RNAi) using dsRNA, catalytic nucleic acids (such as ribozymes and DNAzymes), antisense nucleic acids or a combination thereof.

25

In a presently most preferred embodiment, the host cell is essentially free of any metalloprotease activity.

One preferred embodiment of the present invention relates to a method for the
30 production of a protein of interest in a host cell, in which the protein of interest is a eukaryotic protein, selected from the group comprising insulin, growth hormone, glucagon, somatostatin, interferon, adrenocorticotrophic hormones, angiotensinogen, atrial natriuretic peptides, dynorphin, endorphines, galanin, gastrin, gastrin releasing peptides, neuropeptide Y fragments, pancreastatin, pancreatic polypeptides, secretin,
35 vasoactive intestinal peptide, growth hormone releasing factor, melanocyte stimulating hormone, neurotensin, adrenal peptide, parathyroid hormone and related peptides, somatostatin and related peptides, brain natriuretic peptide, calcitonin, corticotropin releasing factor (CRF), cocaine amphetamine regulated transcript (CART), thymosin,

urotensin, glucagon and glucagon like peptides (GLP-1 and GLP-2), somatostatin, interferon, a vascular endothelial growth factor (VEGF), platelet derived growth factor (PDGF), factor VII, factor VIII, factor V, factor IX, interleukins, urokinase, erythropoietin (EPO), chymosin, tissue plasminogen activator, Positive cofactor 2
 5 glutamine/Q-rich-associated protein (PCAP), peptide tyrosine tyrosine (PYY), ghrelin, orexin, Beta-neoendorphin-dynorphin precursor, CCK or serum albumin.

Another preferred embodiment of the present invention relates to a method for the production of a protein of interest in a host cell, in which the protein of interest is a
 10 protein of fungal origin, selected from the group comprising an amylolytic enzyme, an alpha-amylase, a beta-amylase, a glyco-amylase, a alpha-galactosidase, a cellulytic enzyme, a lipolytic enzyme, a xylanolytic enzyme, a proteolytic enzyme, an oxidoreductase, a peroxidase, a laccase, a pectinase, or a cutinase.

15 A further preferred embodiment of the present invention relates to a method for the production of a protein of interest in a host cell, in which the protein of interest is a bacterial protein, selected from the group comprising an amylolytic enzyme, an alpha-amylase, a beta-amylase, a glyco-amylase, a beta-galactosidase, a cellulytic enzyme, a lipolytic enzyme, a xylanolytic enzyme, a proteolytic enzyme, an oxidoreductase, a
 20 peroxidase, a laccase, a pectinase, or a cutinase.

A special embodiment of the present invention relates to a method for production of a protein of interest in a host cell, in which the protein of interest is a precursor, i.e. a zymogen, a hybrid protein, a protein obtained as a pro sequence or pre-pro sequence,
 25 or in unmaturation form.

General Molecular Biology

Unless otherwise indicated, the recombinant DNA techniques utilised in the present
 30 invention are standard procedures, well known to those skilled in the art. Such techniques are described and explained throughout the literature in sources such as, J. Perbal, A Practical Guide to Molecular Cloning, John Wiley and Sons (1984), J. Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbour Laboratory Press (1989), T.A. Brown (editor), Essential Molecular Biology: A Practical
 35 Approach, Volumes 1 and 2, IRL Press (1991), D.M. Glover and B.D. Hames (editors), DNA Cloning: A Practical Approach, Volumes 1-4, IRL Press (1995 and 1996), F.M. Ausubel et al. (Editors), Current Protocols in Molecular Biology, Greene Pub. Associates

and Wiley-Interscience (1988, including all updates until present), Methods in Enzymology. Vol 194. Guide to Yeast Genetics and Molecular Biology. (1991) Ed Gunthrie and Fink Academic Press, Methods in Microbiology Vol. 26. Yeast Gene Analysis. (1998) Ed. Brown and Tuite. Academic Press, Miller, J. H. (1992) *A Short*
 5 *Course in Bacterial Genetics* (Manual, L., ed), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, Johnston, J. R (1994) *Molecular Genetics of Yeast* (A Practical Approach) Oxford University Press, Oxford., and Molecular Genetics of Yeast: A Practical Approach, Ed. J.R. Johnston, IRL Press (1994) and are incorporated herein by reference.

10

The % identity of a polypeptide is determined by GAP (Needleman and Wunsch, 1970) analysis (GCG program) with a gap creation penalty=5, and a gap extension penalty=0.3. The query sequence is at least 15 amino acids in length, and the GAP analysis aligns the two sequences over a region of at least 15 amino acids. More
 15 preferably, the query sequence is at least 50 amino acids in length, and the GAP analysis aligns the two sequences over a region of at least 50 amino acids. Even more preferably, the query sequence is at least 100 amino acids in length and the GAP analysis aligns the two sequences over a region of at least 100 amino acids. More preferably, the query sequence is at least 250 amino acids in length and the GAP
 20 analysis aligns the two sequences over a region of at least 250 amino acids. Even more preferably, the query sequence is at least 500 amino acids in length and the GAP analysis aligns the two sequences over a region of at least 500 amino acids.

Pitrilysin Subfamily of Metalloproteases

25

The pitrilysin subfamily of metalloproteases is characterized by the presence of a HXXEH (SEQ ID NO:1) motif. A general review of this subfamily is provided by Rawlings and Barrett (1995). Members of this family include, but are not limited to, *S. cerevisiae* Cym1p (SEQ ID NO:4) (Swissprot Accession No. P32898), *Schizosaccharomyces pombe*
 30 C119.7 (SEQ ID NO:5) (Swissprot Accession No. O42908), *Clostridium perfringens* HypA protein (SEQ ID NO:6) (Swissprot Accession No. Q46205), *Borrelia burgdorferi* protein BB0228 (SEQ ID NO:7) (Swissprot Accession No. O51246), *Caenorhabditis elegans* C05D11.1 protein (SEQ ID NO:8) (Swissprot Accession No. P48053), *E. coli* protease III (also known as pitrilysin) (SEQ ID NO:9) (Swissprot Accession No.
 35 P05458), rat NRD convertase (SEQ ID NO:10) (Swissprot Accession No. P47245), human insulysin (SEQ ID NO:11) (Swissprot Accession No. P14735), *Arabidopsis thaliana* CPE (SEQ ID NO:12) (Genbank Accession No. T03302), human

metalloprotease I (in part) (SEQ ID NO:13) (GenBank Accession No. AAH01150) (the full sequence: Swissprot Accession No. O95204), *Bacillus subtilis* zinc protease ymxG (SEQ ID NO:14) (GenBank Accession No. Q04805), and *Mycobacterium tuberculosis* zinc protease Rv2782c (SEQ ID NO:15) (GenBank Accession No. O33324). For *E. coli* protease III (SEQ ID NO:9) it has been shown that the His residues of SEQ ID NO:1, as well as Glu-169, are involved in divalent cation binding whilst the Glu residue flanked by the His residues is a catalytic residue.

A gene encoding a pitrilysin metalloprotease can readily be identified by screening by hybridization for nucleic acid sequences coding for all of, or part of, the metalloprotease, e.g. by using synthetic oligonucleotide probes, that may be prepared on the basis of a cDNA sequence, e.g. the nucleotide sequences encoding any one of the metalloproteases presented as SEQ ID NO's: 4 to 15, in accordance with standard techniques.

15

Genetic Manipulations

The host cell of the invention which is genetically manipulated in order to produce reduced levels of the defined metalloprotease may be modified using standard recombinant DNA technology known to the person skilled in the art. The gene sequence responsible for the production of the metalloprotease may be inactivated or eliminated entirely.

In a particular embodiment, the host cell of the invention is one genetically manipulated at the coding or regulatory regions of the metalloprotease gene. Known and useful techniques include, but are not limited to, gene knockout, gene disruption, random or site directed mutagenesis, introduction of dominant-negative metalloproteases, RNA interference (RNAi) using dsRNA, catalytic nucleic acids (such as ribozymes and DNAzymes), and antisense nucleic acids, or a combination thereof.

30

Mutagenesis may be performed using a suitable physical or chemical mutagenizing agent. Examples of a physical or chemical mutagenizing agent suitable for the present purpose includes ultraviolet (UV) irradiation, hydroxylamine, N-methyl-N'-nitro-N-nitrosoguanidine (MNNG), O-methyl hydroxylamine, nitrous acid, ethyl methane sulphonate (EMS), sodium bisulfite, formic acid, and nucleotide analogues. When such agents are used, the mutagenesis is typically performed by incubating the cell to be mutagenized in the presence of the mutagenizing agent of choice under suitable

35

conditions for the mutagenesis to take place, and selecting for mutated cells having a significantly reduced production of metalloprotease.

Genetic manipulation may also be accomplished by the introduction, substitution or
5 removal of one or more nucleotides in the metalloprotease coding sequence or a
regulatory element required for the transcription or translation thereof. Nucleotides
may, for example, be inserted or removed so as to result in the introduction of a stop
codon, the removal of the start codon or a change of the open reading frame. The
modification or inactivation of the structural sequence or a regulatory element may be
10 accomplished by site-directed mutagenesis or PCR generated mutagenesis in
accordance with methods known in the art.

A convenient way to inactivate or reduce the metalloprotease production of a host cell
is based on the principles of gene interruption. This method involves the use of a DNA
15 sequence corresponding to the endogenous gene or gene fragment which it is desired
to destroy. The DNA sequence is *in vitro* mutated to a defective gene and transformed
into the host cell. By homologous recombination, the defective gene replaces the
endogenous gene or gene fragment. It may be desirable that the defective gene or
gene fragment encodes a marker which may be used for selection of transformants in
20 which gene encoding the metalloprotease has been modified or destroyed.

The term "antisense" as used herein refers to nucleotide sequences which are
complementary to a specific nucleic acid sequence. Antisense molecules may be
produced by any method, including synthesis by ligating the gene(s) of interest in a
25 reverse orientation to a viral promoter which permits the synthesis of a complementary
strand. Once introduced into a host cell, this transcribed strand combines with natural
sequences, in this instance that encoding the metalloprotease, produced by the cell to
form duplexes. These duplexes then block either the further transcription or translation.
In this manner, mutant phenotypes may be generated.

30

The term "catalytic nucleic acid" refers to a DNA molecule or DNA-containing molecule
(also known in the art as a "deoxyribozyme") or an RNA or RNA-containing molecule
(also known as a "ribozyme") which specifically recognizes a distinct substrate and
catalyzes the chemical modification of this substrate. The nucleic acid bases in the
35 catalytic nucleic acid can be bases A, C, G, T and U, as well as derivatives thereof.
Derivatives of these bases are well known in the art.

Typically, the catalytic nucleic acid contains an antisense sequence for specific recognition of a target nucleic acid, and a nucleic acid cleaving enzymatic activity, also referred to herein as the "catalytic domain". The types of ribozymes that are particularly useful in this invention are the hammerhead ribozyme (Haseloff and Gerlach, 1988) and the hairpin ribozyme (Shippy et al., 1999).

Ribozymes useful for the methods of the invention, and DNA encoding the ribozymes, can be chemically synthesized using methods well known in the art. The ribozymes can also be prepared from a DNA molecule (that upon transcription yields an RNA molecule) operably linked to an RNA polymerase promoter, e.g., the promoter for T7 RNA polymerase or SP6 RNA polymerase. When the vector also contains an RNA polymerase promoter operably linked to the DNA molecule, the ribozyme can be produced *in vitro* upon incubation with RNA polymerase and nucleotides. In a separate embodiment, the DNA can be inserted into an expression cassette or transcription cassette. After synthesis, the RNA molecule can be modified by ligation to a DNA molecule having the ability to stabilize the ribozyme and make it resistant to RNase. Alternatively, the ribozyme can be modified to the phosphothio analog for use in liposome delivery systems. This modification also renders the ribozyme resistant to endonuclease activity.

dsRNA (RNAi) is particularly useful for specifically inhibiting the production of a particular protein. This technology relies on the presence of dsRNA molecules that contain a sequence that is essentially identical to the mRNA of the gene of interest, in this case a mRNA encoding the metalloprotease. Conveniently, the dsRNA is produced in a single open reading frame in a recombinant vector or host cell, where the sense and anti-sense sequences are flanked by an unrelated sequence which enables the sense and anti-sense sequences to hybridize to form the dsRNA molecule with the unrelated sequence forming a loop structure. The design and production of suitable dsRNA molecules for genetic manipulation is well known within the capacity of a person skilled in the art, particularly considering Waterhouse et al. (1998), Elbashir et al. (2001), WO 99/32619, WO 99/53050, WO 99/49029, and WO 01/34815.

Owing to the genetic manipulation, the host cell of the invention expresses significantly reduced levels of the metalloprotease. In a preferred embodiment, the level of metalloprotease expressed by the host cell is reduced more than about 25%, such as more than about 30%, such as more than about 35%, such as more than about 40%, such as more than about 45%, such as more than about 50%, such as more than about 55%, such as more than about 60%, such as more than about 65%, such as more than about 70%, such as more than about 75%, such as more than about 80%,

such as more than about 85%, such as more than about 90%, such as more than about 95%, such as more than about 98%, and such as more than about 99%.

In a presently most preferred embodiment, the product expressed by the host cell is
5 essentially free of any activity of the defined metalloprotease.

In the present context, the term "essentially free" relates to a host, wherein the metalloprotease expressed by said host cell is reduced to a level, where the function of said metalloprotease has no biologically significant reducing influence on the production
10 of the protein of interest.

Protein of Interest

The terms "polypeptide", "protein" and "peptide" are used herein interchangeably and
15 in the present context relates to any of the numerous naturally occurring extremely complex substances such as but not limited to enzymes or antibodies that consist of amino acid residues joined by peptide bonds, contain the elements carbon, hydrogen, nitrogen, oxygen, usually sulphur, and occasionally other elements such as but not limited to phosphorus or iron, that are essential constituents of all living cells, that are
20 in nature synthesised from raw materials by plants but assimilated as separate amino acids by animals, that are both acidic and basic and usually colloidal in nature although many have been crystallised, and that are hydrolyzable by acids, alkalies, proteolytic enzymes, and putrefactive bacteria to polypeptides, to simpler peptides, and ultimately to alpha-amino acids.

25

As defined herein, a "recombinant polypeptide" is a protein which is not native to the host cell, or a native polypeptide in which modifications have been made to alter the native sequence, or a native protein whose expression is quantitatively altered as a result of a manipulation of a native regulatory sequence required for the expression of
30 the native protein, such as a promoter, a ribosome binding site, etc., or other manipulation of the host cell by recombinant DNA techniques.

Owing to the absence or reduction in activity of the defined metalloprotease, at least a portion of the recombinant polypeptides expressed by the host cell may also be a
35 precursor protein, i.e. a zymogen, a hybrid protein, a protein obtained as a pro sequence or pre-pro sequence, or in unmaturation form.

In a more specific embodiment, the recombinant polypeptide is of eukaryotic origin, such as insulin, adrenocorticotrophic hormones, angiotensinogen, atrial natriuretic peptides, dynorphin, endorphines, galanin, gastrin, gastrin releasing peptides, neuropeptide Y fragments, pancreastatin, pancreatic polypeptides, secretin, vasoactive
5 intestinal peptide, growth hormone releasing factor, melanocyte stimulating hormone, neurotensin, adrenal peptide, parathyroid hormone and related peptides, somatostatin and related peptides, brain natriuretic peptide, calcitonin, corticotropin releasing factor (CRF), cocaine amphetamine regulated transcript (CART), thymosin, urotensin, glucagon and glucagon like peptides (GLP-1 and GLP-2), somatostatin, interferon, a
10 vascular endothelial growth factor (VEGF), platelet derived growth factor (PDGF), factor VII, factor VIII, factor V, factor IX, interleukins, urokinase, erythropoietin (EPO), chymosin, tissue plasminogen activator, CCK or serum albumin.

With specific regard to "glucagon and glucagon like peptides", this term as used herein
15 may refer to polypeptides of human origin or from other animals and recombinant or semisynthetic sources and include all members of the glucagon family, such as GRPP (glicentine related polypeptide), glucagon, GLP-1 (glucagon like peptide 1), and GLP-2 (glucagon like peptide 2), including truncated and/or N-terminally extended forms, such as GLP-1(7-36), and includes analogues, such as GLP-1(7-35)R36A GLP-2 F22Y,
20 GLP-2 A19T+34Y. GLP2 A2G and GLP-2 A19T, and other analogues having from 1 to 3 amino acid changes, additions and/or deletions.

Host Cells and the Expression of Recombinant Polypeptides Therefrom

25 The host cells for use in the present invention can be prokaryotic or eukaryotic. The eukaryotic host cells for use in the present invention can be, for example, fungal, mammalian, plant or insect cells. Preferably, the host cells are yeast cells.

In order to produce the desired polypeptide, the host cell of the invention comprises a
30 nucleic acid sequence encoding the recombinant polypeptide as well as regulatory sequences for directing the expression of the desired product such as regions comprising nucleotide sequences necessary or e.g. transcription, translation and termination. The genetic design of the host cell of the invention may be accomplished by the person skilled in the art, using standard recombinant DNA technology for the
35 transformation or transfection of a host cell.

Preferably, the host cell is modified by methods known in the art for the introduction of an appropriate expression cassette in, for example a plasmid or a viral vector, comprising the nucleic acid encoding the recombinant polypeptide. The expression cassette may be introduced into the host cell by a number of techniques including, but
5 not limited to, as an autonomously replicating plasmid or integrated into the chromosome.

Expression cassettes may contain regulatory sequences such as transcription control sequences, translation control sequences, origins of replication, and other regulatory
10 sequences that are compatible with the recombinant cell and that control the expression of nucleic acid molecules encoding the recombinant polypeptide. In particular, recombinant nucleic acid molecules of the present invention include transcription control sequences. Transcription control sequences are sequences which control the initiation, elongation, and termination of transcription. Particularly important
15 transcription control sequences are those which control transcription initiation, such as promoter, enhancer, operator and repressor sequences. Suitable transcription control sequences include any transcription control sequence that can function in at least one of the recombinant cells of the present invention. A variety of such transcription control sequences are known to those skilled in the art. Preferred transcription control
20 sequences include those which function in bacterial, yeast, arthropod and mammalian cells, such as, but not limited to, tac, lac, trp, trc, oxy-pro, omp/lpp, rrnB, bacteriophage lambda, bacteriophage T7, T7lac, bacteriophage T3, bacteriophage SP6, bacteriophage SP01, metallothionein, alpha-mating factor, Pichia alcohol oxidase, alphavirus subgenomic promoters (such as Sindbis virus subgenomic promoters),
25 antibiotic resistance gene, baculovirus, *Heliothis zea* insect virus, vaccinia virus, herpesvirus, raccoon poxvirus, other poxvirus, adenovirus, cytomegalovirus (such as intermediate early promoters), simian virus 40, retrovirus, actin, retroviral long terminal repeat, Rous sarcoma virus, heat shock, phosphate and nitrate transcription control sequences as well as other sequences capable of controlling gene expression in
30 prokaryotic or eukaryotic cells. Transcription control sequences of the present invention are most preferably naturally occurring transcription control sequences associated with yeast. Suitable promoters for *S. cerevisiae* include the MF α 1 promoter, galactose inducible promoters such as GAL1, GAL7 and GAL10 promoters, glycolytic enzyme promoters including TPI1 and PGK1 promoters, TRP1 promoter, CYCI promoter, CUP1
35 promoter, PHOS promoter, ADH1 promoter, and HSP promoter. A suitable promoter in the genus *Pichia* is the AOXI (methanol utilisation) promoter.

Recombinant polypeptides of the present invention may also (a) contain secretory signals to enable an expressed polypeptide to be secreted from the cell that produces the polypeptide and/or (b) contain fusion sequences which lead to the expression of fusion proteins. Examples of suitable signal segments include any signal segment
5 capable of directing the secretion of the fusion protein. Preferred signal segments include, but are not limited to, tissue plasminogen activator (t-PA), interferon, interleukin, growth hormone, histocompatibility and viral envelope glycoprotein signal segments, the leader sequence originating from the fungal amyloglucosidase (AG) gene such as galA - both 18 and 24 amino acid versions e.g. from *Aspergillus sp.*, the α -
10 factor gene of yeasts e.g. from *Saccharomyces sp.* and *Kluyveromyces sp.*, the P-factor of *Schizosaccharomyces sp.*, and the α -amylase gene from *Bacillus sp.*, as well as natural signal sequences.

The cloning vehicle may also comprise a selectable marker, e.g. a gene, the product of
15 which complements a defect in the host cell, or one which confers antibiotic resistance, such as ampicillin, kanamycin, chloramphenicol or tetracycline resistance.

The procedures used to ligate the DNA construct of the invention, the promoter, terminator and other elements, respectively, and to insert them into suitable cloning
20 vehicles containing the information necessary for replication, are well known to persons skilled in the art.

Recombinant DNA technologies can be used to improve the expression of transformed nucleic acid molecules by manipulating, for example, the number of copies of the
25 nucleic acid molecules within a host cell, the efficiency with which those nucleic acid molecules are transcribed, the efficiency with which the resultant transcripts are translated, and the efficiency of post-translational modifications. Recombinant techniques useful for increasing the expression of nucleic acid molecules useful for the methods of the present invention include, but are not limited to, operably linking the
30 nucleic acid molecule to high-copy number plasmids, integration of the nucleic acid molecule into one or more host cell chromosomes, addition of vector stability sequences to plasmids, substitutions or modifications of transcription control signals (e.g., promoters, operators, enhancers), substitutions or modifications of translational control signals (e.g., ribosome binding sites, Shine-Dalgarno sequences), modification
35 of nucleic acid molecule to correspond to the codon usage of the host cell, and the deletion of sequences that destabilize transcripts. The activity of an expressed recombinant polypeptide of the present invention may be improved by fragmenting, modifying, or derivatizing polynucleotide molecules encoding such a protein.

Methods of Producing Recombinant Polypeptides

Host cells that have been transfected or transformed with the nucleic acid encoding the
5 recombinant polypeptide are cultured for a sufficient time and under appropriate
conditions known to those skilled in the art in view of the teachings disclosed herein to
permit the production, and preferably secretion, of the polypeptide, followed by
recovery of the desired product.

10 Furthermore, owing to the reduced activity of the metalloprotease, the recombinant
polypeptide expressed by the host cell may be obtained as a precursor protein, i.e. a
zymogen, a hybrid protein, a protein obtained as a pro sequence or pre-pro sequence,
or in unmaturationed form.

15 The broth or medium used for culturing may be any conventional medium suitable for
growing the host cell in question, and may be composed according to the principles of
the prior art. The medium preferably contains carbon and nitrogen sources and other
inorganic salts. Suitable media, e.g. minimal or complex media, are available from
commercial suppliers, or may be prepared according to published protocols.

20

With regard to yeast host cells, it is often advantageous to produce heterologous
polypeptides in a diploid yeast culture, because possible genetical defects may become
phenotypically expressed in a haploid yeast culture, e.g. during continuous
fermentation in production scale, and because the yield may be higher. The production
25 of recombinant polypeptides in yeast host cell is described in Molecular Genetics of
Yeast: A Practical Approach, Ed. J.R. Johnston, IRL Press (1994) which is incorporated
herein by reference.

After cultivation, the protein is recovered by conventional methods for isolation and
30 purification proteins from a culture broth. Well-known purification procedures include
separating the cells from the medium by centrifugation or filtration, precipitating
proteinaceous components of the medium by means of a salt such as ammonium
sulphate, and chromatographic methods such as e.g. ion exchange chromatography,
gel filtration chromatography, affinity chromatography, etc.

35

The present invention is exemplified by demonstrating that the total amount of CCK
and proBNP is increased about 60% about 100%, respectively when compared to the

non-modified host cell. The examples thus demonstrate that the host cells of the present invention are able to increase the production of diverse proteins.

Further, the examples disclose that additional disruption of other proteases enhance
5 the production of the protein of interest, for example disruption of both *KEX2* and *CYM1* results in an additive effect in the yield in the production of CCK (nearly 100%).

Further, it will be understood by the skilled addressee that special amino acids within
10 some of the motifs described, particularly His, Glu, Asp or Lys, are essential to the function of the metalloprotease by functioning as metal ligands.

Furthermore, it will be recognised that the metalloproteases of the present invention
15 have been annotated widely in the literature as family members of the pitrilysin family, insulysin family, insulinase family, inverzincin family and M16 subfamily of clan ME.

Thus, it will be understood that any feature and/or aspect discussed above in connection with any of these different family annotations apply by analogy to the metalloprotease described herein, which all include the HXXEH motif.

20 However, please note that pitrilysin (without family) in itself refers to a specific member of clan ME of metalloproteases in *E. coli*.

Examples

Materials and Methods

Yeast strains and growth conditions

25

The yeast strains used are listed in Table I. Construction of strains were carried out using either the two step gene disruption technique (Rothstein, 1991) or the PCR based method by (Brachmann et al., 1998). Media were purchased from Difco, amino acids and other supplements from Sigma-Aldrich. Yeast cells were grown at 30°C in YPD (1%
30 yeast extract, 2% peptone and 2% dextrose) or synthetic complete media (SC) based on yeast nitrogen base with ammonium sulfate, succinic acid, NaOH and appropriate amino acids. Transformations with either linear DNA or plasmids were performed using the modified lithium acetate procedure as described (Gietz et al., 1995). Analysis of heterologous expressed CCK was performed from yeast growing in exponential phase
35 due to the consistency in CCK-22 biosynthesis, in contrast to the results from yeast within the stationary phase (Fig. 2). ProCCK processing was analysed from cell extract

and media of 5 A₆₀₀ units of cells per ml synthetic complete media. Cell growth was followed by the absorbance at 600 nM.

Table 1. *S. cerevisiae* strains used in this study. Null mutants of putative
5 metalloproteases are named by the ORF in the genotype and (*) represents mitochondrial proteases.

Strain	Genotype	Source
BY4705	<i>MATα ade2Δ::hisG his3Δ200 leu2Δ0 lys2Δ0 ura3Δ0</i>	(Brachmann et al., 1998)
LJY13	<i>MATα ade2Δ::hisG his3Δ200 leu2Δ0 lys2Δ0 ura3Δ0 yps1::TRP1</i>	This study
LJY14	<i>MATα ade2Δ::hisG his3Δ200 leu2Δ0 lys2Δ0 ura3Δ0 yps1::TRP1 yps3::LEU2</i>	This study
LJY15	<i>MATα ade2Δ::hisG his3Δ200 leu2Δ0 lys2Δ0 ura3Δ0 yps1::TRP1 yps3::LEU2 yps2::URA3</i>	This study
BJ2168	<i>MATα prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52 gal2</i>	(Jones, 1991)
LJY21	<i>MATα prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52 gal2 kex1::LEU2</i>	This study
LJY22	<i>MATα prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52 gal2 kex1::LEU2 kex2::TRP1</i>	This study
LJY23	<i>MATα prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52 gal2 kex2::TRP1</i>	This study
LJY122	<i>MATα ape1::KANMX ape2::LYS2 his3Δ0 leu2Δ0 lys2Δ0 ura3Δ0</i>	This study
LJY123	<i>MATα ape1::KANMX ape2::LYS2 ape3::LEU2 his3Δ0 leu2Δ0 lys2Δ0 ura3Δ0</i>	This study
LJY201	<i>MATα prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52 gal2 axl1::LEU2</i>	This study
LJY202	<i>MATα prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52 gal2 ste24::LEU2</i>	This study
LJY203	<i>MATα prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52 gal2 prd1::LEU2</i>	This study
LJY204	<i>MATα prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52 gal2 yil108w::LEU2</i>	This study
Y15298	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 ste23::KANMX</i>	Euroscarf
Y11874	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 aap1::KANMX</i>	Euroscarf
Y10148 (*)	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 afg3::KANMX</i>	Euroscarf
Y14953	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 ape1::KANMX</i>	Euroscarf
Y16224 (*)	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 rca1::KANMX</i>	Euroscarf
Y14984 (*)	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 mip1::KANMX</i>	Euroscarf
Y17144	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 yme1::KANMX</i>	Euroscarf
Y13211	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 ybr074w::KANMX</i>	Euroscarf
Y13801	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 ydl104c::KANMX</i>	Euroscarf

Y11941	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 yhr113w::KANMX</i>	Euroscarf
Y11960	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 yhr132c::KANMX</i>	Euroscarf
Y12296	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 yil137c::KANMX</i>	Euroscarf
Y15370	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 ynl045w::KANMX</i>	Euroscarf
10864B	<i>MATα ura3-Δ851 leu2-Δ1 his3Δ200 lys2Δ202 ykr035c-ykr038c::URA3</i>	Euroscarf
Y11749	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 yol057w::KANMX</i>	Euroscarf
Y16248	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 yol098c::KANMX</i>	Euroscarf
10231B	<i>MATα his3Δ200 leu2Δ1 trp1Δ63 ura3-52 yol154w(4,744)::KANMX</i>	Euroscarf
Y14266	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 ydr430c::KANMX</i>	Euroscarf
LJY430	<i>MATα prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52 gal2 ydr430c::LEU2</i>	This study
LJY432	<i>MATα prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52 gal2 kex2::TRP1 ydr430c::LEU2</i>	This study

DNA extraction and amplification

Yeast genomic DNA was isolated as described (Philippsen et al., 1991). Polymerase chain reaction (PCR) was performed using either *Pwo* polymerase or the enzyme cocktail based on *Taq*, *Pwo* and *Pfu* polymerase (Expand long range PCR kit, XL-PCR) both from Roche. All PCR products were visualised by agarose gel-electrophoresis and PCR products either purified from the gel using the gel-extraction kit (Qiagen) or from the reaction mixture by PCR purification spin columns (GENOMED). PCR based one step gene disruption was performed using 50 ng of plasmid from the pRS400 series (Brachmann et al., 1998) as template. Amplification of the marker was performed with oligonucleotides having 20 nucleotides towards the plasmid and additional 50 nucleotides flanking the target gene (Table 2). All other DNA manipulations were carried out according to standard procedures (Sambrook et al., 1989).

Plasmid constructions

Expression of proCCK was performed in pRS426 [2μ *URA3*] (Brachmann et al., 1998) using the phosphoglycerate kinase promoter (*PGK1p*) and terminator (*PGK1t*). The *PGK1* promoter was amplified with *PGK1p* 5'*Hind*III and *PGK1p* 3'*MCS* (Table 2) using 100 ng of genomic yeast DNA as template and subsequently cloned into pGEM-11 (Promega) in the *Hind*III and *Sac*I restriction enzyme sites. The terminator was amplified with *PGK1t* 5'*Bgl*II and *PGK1t* 3'*Sac*I (Table 2) and ligated into the plasmid containing the promoter at the *Sac*I and *Eco*RI restriction enzyme sites. This construct, pGEM-11 *PGK1pMCS**PGK1t* then contained the *PGK1*-promoter, a multiple cloning site (*MCS*) with the restriction enzyme sites *Eco*RI, *Bam*HI, *Xba*I and *Bgl*II followed by the *PGK1* terminator. The preproMf α 1p-proCCK fusion (Rourke et al., 1997) (Fig. 1) was subcloned into the *Eco*RI and *Xba*I sites of pGEM-11 *PGK1pMCS**PGK1t* and finally the entire gene was cloned into pRS423 as well as pRS426 to complete the yeast CCK expression constructs, pRS423 preproMf α 1p-proCCK and pRS426 preproMf α 1p-proCCK respectively. Expression of *CYM1* on a multi copy plasmid was constructed by amplification of the open reading frame (ORF) of *CYM1* and additional 926 bp at the 5' end and 703 bp at the 3' end. The amplification was carried out by XL-PCR using 100 ng of genomic yeast DNA and the oligonucleotides, *CYM1* 5'*Apa*I and

Table 2. Oligonucleotides used.

Oligo	Oligonucleotide sequence (5'-3')	Purpose
<i>PGK1p5'HindIII</i>	AATAGAAGCTTGTCTGACTGATCTATCCAAAAGT (SEQ ID NO: 16)	Expression construct
<i>PGK1p3'MCS</i>	AAAAGAGCTCGGCCAGATCTTCTAGAGGATCCAA GAATTCTGTTTTATATTTGTTGTAAAAAGTAG (SEQ ID NO: 17)	Expression construct
<i>PGK1t5'BglII</i>	TTTTGAATTCCAAGATCTCCCATGTCTCTACTGGTGG (SEQ ID NO: 18)	Expression construct
<i>PGK1t3'SacI</i>	CCCCGAGCTCGTCGACCCTTCTCGAAAGCTTTAACGAAC GC (SEQ ID NO: 19)	Expression construct
5'MF α 1- <i>EcoRI</i>	TTTTGAATTCAAAGAATGAGATTTCTTCAATTTTACTG CAG (SEQ ID NO: 20)	preproMf α 1p-proC CK
CCK3'- <i>XbaI</i>	TTTTTCTAGACTAGGAGGGGTACTCATACTCCTCGGC (SEQ ID NO: 21)	preproMf α 1p-proC CK
CCK-22 K→A (S)	CGAATGTCCATCGTTGCGAACCTGCAGAACCTG (SEQ ID NO: 22)	Lys ⁶¹ →Ala ⁶¹ mutation
CCK-22 K→A (AS)	CAGGTTCTGCAGGTTCTTAACGATGGACATTCG (SEQ ID NO: 23)	Lys ⁶¹ →Ala ⁶¹ mutation
CCK-22 K→R (S)	CGAATGTCCATCGTTAGGAACCTGCAGAACCTG (SEQ ID NO: 24)	Lys ⁶¹ →Arg ⁶¹ mutation
CCK-22 K→R (AS)	CAGGTTCTGCAGGTTCTTAACGATGGACATTCG (SEQ ID NO: 25)	Lys ⁶¹ →Arg ⁶¹ mutation
CCK-22 seq	TCGCAGAGAACGGATGGC (SEQ ID NO: 26)	Sequencing
<i>CYM15'ApaI</i>	TTTTGGGCCCTTCATGGTGATACGGTATCTCTTGGC (SEQ ID NO: 27)	Cloning of <i>CYM1</i>
<i>CYM13'XhoI</i>	TTTTCTCGAGAAGGTGGAACATACTGCCCTGGGATGG (SEQ ID NO: 28)	Cloning of <i>CYM1</i>
<i>KEX25'</i>	TTTTGAGCTCGTTTAGGAAACGTCCTTGGCGGAGATGC (SEQ ID NO: 29)	Cloning of <i>KEX2</i>
<i>KEX23'</i>	TTTTTCTAGACACTGCGAATCCATGGTATAAACCAAAACC (SEQ ID NO: 30)	Cloning of <i>KEX2</i>
<i>KEX2DC5'</i>	GTCGTTGTTCATGGACATACCTCC (SEQ ID NO: 31)	Control of $\Delta kex2$
<i>KEX2DC3'</i>	TACAAATGTTCTTCTGCCATTTCTGG (SEQ ID NO: 32)	Control of $\Delta kex2$
<i>TRP15'NdeI</i>	GGTTCATATGCGCCGGAGCTCCTCGACAGCAG (SEQ ID NO: 33)	Cloning of <i>TRP1</i>
<i>TRP13'AvrII</i>	GGTTCCTAGGATCCGCAAGTTTGATTCCATTGCGGTG	Cloning of <i>TRP1</i>

	(SEQ ID NO: 34)	
KEX15'GD400	TTAAAGAGTACCTTGGCTATAGAATACCGTAGAGATAAA GACCTGAATAGAGATTGTACTGAGAGTGCAC	KEX1 deletion
	(SEQ ID NO: 35)	
KEX13'GD400	AGGTATTATAACTATTTTTCTGTATTTTTTATATATTTTTAT TTGCCAAGCTGTGCGGTATTTACACCCG	KEX1 deletion
	(SEQ ID NO: 36)	
KEX15'DC400	CTTTGGTTAAAGAGTACCTTGGC (SEQ ID NO: 37)	Control of $\Delta kex1$
KEX13'DC400	TACTACGAAAAGCGTGTGCGAGG (SEQ ID NO: 38)	Control of $\Delta kex1$
CYM15'GD400	TAGAAGGCTACTCAAAAGAATAAAGTTACTATAAAATATA CTGCGGTATATAGATTGTACTGAGAGTGCAC	CYM1 deletion
	(SEQ ID NO: 39)	
CYM13'GD400	GATCGGCAAGAACTTTGAAGCAGTATATTTACAGGATT AAATTATATATCTGTGCGGTATTTACACCCG	CYM1 deletion
	(SEQ ID NO: 40)	
CYM15'DC400	CGGAGGGGCTCTATGATAAAGG (SEQ ID NO: 41)	Control of $\Delta cym1$
CYM13'DC400	GAGTAACTAGGGCTTCTCTTCCC (SEQ ID NO: 42)	Control of $\Delta cym1$

CYM1 3'XhoI (Table 2). The PCR product was purified on spin columns and subsequently cloned into the *Apa*I and *Xho*I restriction enzyme sites of pRS425.

5

The Lys⁶¹ residue, believed to be crucial for the proteolysis of proCCK to release CCK-22, was exchanged by Ala by site-directed mutagenesis (Horton et al., 1993). The exchange was performed by PCR using the *Pwo* polymerase (Boehringer Mannheim), where two products were amplified with the oligonucleotides sets, *PGK1p5' Hind*III / CCK-22 K→A (antisense) and CCK-22 K→A (sense) / *PGK1t3' Sac*I (Table 2) and 50 ng of pRS426 preproMf α 1p-proCCK as template to each reaction. The two products were subjected to agarose gel-electrophoresis and approximately 1 mm² of each product where cut out and used directly as template in a third PCR reaction. In this reaction the full-length cDNA encoding the fusion protein was amplified using *PGK1p5' Hind*III and *PGK1t3' Sac*I (Table 2). The PCR product was subcloned into pCR-Blunt II (Invitrogen) and sequenced with the CCK specific primer, CCK-22 seq. Finally the *PGK1p* preproMf α 1p-proCCK (K→A) *PGK1t* product was cloned into the *Hind*III and *Sac*I sites of pRS426 to construct the expression plasmid, proCCK (K→A). Substitution of Lys with Arg was performed as described above by exchanging the CCK specific primers with

CCK-22 K→R (antisense) and CCK-22 K→R (sense) to construct the proCCK (K→R) vector.

Strain construction

5

Construction of a partial *KEX2* disruption was performed in BJ2168 by amplification of the entire *KEX2* gene with 1000 bp on each site of the ORF by XL-PCR using 100 ng of genomic yeast DNA and the oligonucleotides, *KEX2* 5'*SacI* and *KEX2* 3'*XbaI* (Table 2). The PCR product was purified on spin columns and cloned into pCR-Blunt II (Invitrogen). Amplification of *TRP1* was performed by XL-PCR introducing an *NdeI* site 925 bp 5' to the ORF and an *AvrII* site 212 bp 3' to the stop codon using *TRP1* 5'*NdeI* and *TRP1* 3'*AvrII* (Table 2). The PCR product was purified and subcloned into the *NdeI* and *AvrII* sites of *KEX2* eliminating 2018 bp of *KEX2* and 170 bp of the promoter. The *kex2::TRP1* construct was excised from pCR-Blunt II using the *NotI* and *SpeI* restriction enzymes and subsequently transformed into BJ2168. Transformants were selected on SC-Trp plates followed by a colony PCR screen to test for correct integration using oligonucleotides that cover the entire marker plus an additional 1200 bp on each site of *KEX2*, *kex2* DC5' and *kex2* DC3' (Table 2). Construction of a *kex2* *kex1* strain was performed by the two step gene disruption technique (Rothstein, 1991) using the *LEU2* marker. Amplification of *LEU2* was performed by XL-PCR using 50 ng of pRS405 as template and *kex15'*GD400 and *kex13'*GD400 (Table 2). The PCR product was purified using PCR purification spin kit (GENOMED) and subsequently transformed into LJY23. Transformants were selected on SC-Leu plates and correct integration was tested by PCR-based colony screen using *kex15'*DC and *kex13'*DC (Table 2).

25

A Δ *cym1::LEU2* (LJY430) strain in a BJ2168 background was constructed by the one step gene disruption technique as described above for the Δ *kex1* strain using the oligonucleotides, *CYM15'*GD400 and *CYM13'*GD400 (Table 2) for gene disruption and *cym15'*DC and *cym13'*DC (Table 2) for disruption control. All null mutants created by this method were prepared with oligonucleotides designed towards the 50 bases adjacent to the 5' and 3' UTR with a specific 3' end to the pRS400 series of vectors containing various markers (Brachmann et al., 1998). Transformants were selected on appropriate agar plates followed by a colony PCR screen to test for correct integration using oligonucleotides that cover the entire marker plus an additional 200 bp on each site. Only the oligonucleotides that are not positioned as described above are shown in Table 2.

35

Gene deletions of *STE24*, *AXL1*, *PRD1* and *YIL108w* were made in BJ2168 using the PCR disruption technique (Brachmann et al., 1998) and pRS405 [*LEU2*] as template.

The LJY123, which contain gene deletions of *APE1*, -2 and -3, was derived from Y14953
 5 using PCR disruption technique (Brachmann et al., 1998). *APE2* was initially replaced with the *LYS2* (pRS317 [*cen*; *LYS2*]) where the PCR product was purified from agarose gel prior to transformation and *APE3* was substituted with the *LEU2* marker (pRS405 [*LEU2*]).

10 The *yps1 yps2 yps3* triple mutant (LJY15) was constructed in BY4705 using the PCR disruption technique (Brachmann et al., 1998). The ORF of *YPS1* were initially deleted by insertion of the *TRP1* locus (pRS404) to generate LJY13. This strain was then used as host for the deletion of *YPS3* by insertion of the *LEU2* marker (pRS405) and finally the *YPS2* was deleted by insertion of the *URA3* marker by amplification of pRS406
 15 [*URA3*] to construct LJY15 (Table I).

CCK and *CYM1* expression

Human proCCK was expressed as a fusion protein between the prepro leader sequence of yeast α -mating factor and proCCK (preproMf α 1p-proCCK). The fusion construct was
 20 expressed on multi-copy plasmids, with constitutively gene transcription from the phosphoglycerate kinase promoter. "ProCCK expression" refers to expression using pRS426 preproMf α 1p-proCCK, which was used in all yeast strains with exception of BY4705 and the isogenic yapsins deletion strains, where proCCK was expressed from pRS423 preproMf α 1p-proCCK. *CYM1* expression was driven by its own promoter.
 25 Plasmid constructs, and oligonucleotides used are listed in Table 2.

Enzymatic treatment

Trypsin treatment was performed using 1 mg/ml Trypsin (Worthington Biochemical Corporation) in a 50 mM sodium phosphate buffer (pH 7.5) for 30 min at RT and
 30 terminated by immersion into boiling water for 10 min. Carboxypeptidase B (Boehringer Mannheim) treatment with a final concentration of 4 μ g/ml was performed in 0.1 mM sodium phosphate buffer (pH 7.5) at room temperature for 30 min. The reaction was terminated by immersion into boiling water for 10 min.

Gel chromatography

Yeast transformants grown to late exponential phase were centrifuged at 15000 *g* to
5 collect the cells and 500 μ l of the medium was loaded directly onto a Sephadex G-50
superfine (Pharmacia) column (1 $\times$ 100 cm) at 4°C. The sample was eluted in VBA buffer
(20 mM barbital buffer, 0.11% bovine serum albumin and 0.6 mM thiomersal) at a flow
rate of 3.5 ml/h and fractions were collected every 17 min. Calibrations were performed
by including ^{125} I-albumin (V_0) and ^{22} NaCl (V_t). The elution constants K_d of peaks eluting
10 at V_e are calculated as $K_d = (V_e - V_0) / (V_t - V_0)$.

Radio-immunoassay

Two different antisera were used to determine the amount of processed
15 cholecystokinin. Ab 89009 (Paloheimo et al., 1994) is specific for the N-terminus of
CCK-22 and Ab 7270 (Hilsted et al., 1986) is specific for Gly-extended CCK. The
fraction of CCK processed to CCK-22 is calculated by division of the immuno-reactivity
measured with Ab 89009 with the amount measured with the same antibody after the
sample was treated with trypsin to measure the total amount of N-terminal extended
20 CCK-22.

Yeast extract and protease assay

Ten A_{600} units of yeast cells growing in exponential phase were sedimented by
25 centrifugation at 3000 *g* for 5 min, washed once in 25 ml H₂O and transferred to a 2 ml
Eppendorf tube. An equal amount of acid washed glass beads (Sigma-Aldrich) was
added followed by 200 μ l of 0.1 M NaH₂PO₄ (pH 4.5) including various inhibitors (150
 μ M Bestatin, 30 μ M E-64, 10 μ M Leupeptin, 1 μ M Pepstatin A, 1 mM
phenylmethylsulfonyl fluoride (PMSF), 1 mM EDTA and 1 mM 1,10-orthophenanthroline
30 or 1 tablet complete inhibitor with or without EDTA per 2.5 ml 0.1 M NaH₂PO₄
(Boehringer Mannheim)). The cells were broken by vortexing 3 $\times$ 20 sec and the extracts
were clarified by centrifugation at 15000 *g* for 10 min. All steps were carried out at 4°C.
The protease assay was performed using 20 pmol synthetic amidated CCK-33
(Peninsula Laboratorie Europe, Merseyside, England) or Ac-CCK-33-Gly (Cambridge
35 Research Biochemicals, Stockton, England) as substrate, 20 μ l yeast extract, various
inhibitors and activators in a total volume of 30 μ l. The mixture was incubated at 30°C

for 1 h and the reaction terminated by adding 500 μ l VBA buffer followed by immediate immersion into a boiling water bath for 10 min.

Protease assay using metalloprotease deficient strains

5

The assay was performed as described above, but with addition of 1 mM Bestatin and 1 mM Mn^{2+} to decrease N-terminal degradation.

Protease assay using intact yeast cells

10

Five A_{600} units of exponential growing cells were sedimented, washed once in 5 ml H_2O and once in SC media (pH 6.0), before the cells were resuspended in 25 μ l SC media. The protease assay was performed by addition of 20 pmol synthetic Ac-CCK-33-Gly as substrate and the mixture incubated with gentle shaking at 30°C for 1 h. The reaction
15 was terminated by addition of 500 μ l VBA and the cells removed by centrifugation before the supernatant was immersed into boiling water for 10 min.

Analysis of secreted CCK by MALDI-TOF

20 Fifty A_{600} units of CCK transformed yeast cells were subjected to 25 ml of fresh media, followed by inoculation for 3 h. Cells were removed by centrifugation at 15000 g for 10 min and 500 μ l of media was concentrated and desalted by reverse phase using a ZipTip C_{18} column (Millipore). The peptides were eluted with 10 μ l 50% acetonitrile. The purified peptides were analysed in a Matrix Assisted Laser Desorption/Ionization time-
25 of-flight mass spectrometer (Biflex, Bruker-Franzen, Bremen, Germany) operated in the reflected mode using time lag focusing (delayed extraction). For analysis, 0.5 μ l of the sample was mixed with 0.5 μ l matrix solution (α -cyano-4-hydroxycinnamic acid in acetonitrile/methanol, Hewlett Packard). Then 0.5 μ l of the mixture was applied to the probe and allowed to dry before introduction into the mass spectrometer.

30

Statistical analysis

Statistical calculations were performed using an unpaired students t-test to analyse whether the change in proCCK expression or the fraction of mature CCK-22 between
35 wild type yeast expressing proCCK and mutants isogenic to the wild type strain can be considered to be statistically significant. The P-value calculated for CCK-22 processing between yapsin mutants are comparisons of BY4705 and each mutant, whereas the

brackets represent comparisons between BY4705 *yps1* and BY4705 *yps1 yps3* and, BY4705 *yps1 yps3* and BY4705 *yps1 yps2 yps3*, respectively.

Expression of proBNP in *Saccharomyces cerevisiae* – Construction of the *yps1* mutant

5 Cloning of preproBNP

Messenger RNA was isolated from a 500 mg Biopsy from human heart using the Quickprep Micro mRNA purification Kit (Amersham Pharmacia Biotech). First strand cDNA was prepared from 2 µg mRNA in a reaction containing, 2.5 µl 10× first strand
10 buffer (Promega), 2.5 µl 100 mM DTT, 2.5 µl 10 mM dNTP, 2.5 µl Na pyrophosphate, 10 pmol Oligo(dT)₁₈, 10 units reverse transcriptase, AMV (Promega), and H₂O to 25 µl. Messenger RNA and Oligo(dT)₁₈ was heated to 70°C for 5 min cooled on ice for 5 min prior to cDNA synthesis. The first strand cDNA synthesis was performed at 42°C for 60 min.

15

The cDNA encoding preproBNP was amplified using Pwo polymerase (Roche), 1 µl 1. strand cDNA, 5 µl 10× Pwo buffer included MgCl₂ (Roche), 5 µl 2.5 mM dNTP, 30 pmol of each primer (BNP5'*EcoRI* and BNP3'*XbaI*). The PCR product encoding preproBNP was cloned in pBluescript II (Stratagene). All subsequent PCR reactions were performed as
20 described above.

The fusion between the cDNA's encoding the preprosequence of the α-mating factor and proBNP was performed using overhang extension PCR, where two PCR reactions were set up. One using 50 ng of pRS426 preproMfα1p-proCCK as template,
25 MFα15'*EcoRI* and MF1BNP (AS) as primers and a second with 50 ng of preproBNP cloned in pBluescript and the primers, MF1BNP (S) and BNP3'*XbaI*. In the third PCR reaction, approximately 50 ng of each PCR products were purified from agarosegel from the the two initial PCR using the gel-extraction kit (Qiagen) and used as template with the two oligonucleotides, MFα15'*EcoRI* and BNP3'*XbaI*. The preproMfα1p-proBNP
30 encoding construct was subcloned in pCR-Blunt II (Invitrogen) and sequenced with vector specific oligonucleotides prior to subcloning into the *EcoRI* and *XbaI* sites of pGEM-11 *PGK1pMCSPGK1t*. Finally the entire gene was cloned into pRS426 to complete the yeast proBNP expression constructs, pRS426 preproMfα1p-proBNP.

35 Furthermore, two additional constructs have been made, in which the proBNP fragment (1-76) has been removed. These constructs are similar to do the preproMfα1p-proBNP, but do only synthesise BNP-32. In the first construct, the Kex2p cleavage site and the

spacer peptide of the preproMf α 1p has been sustained (KREAEA)(Figure 14B), whereas in the other construct, the spacer peptide has been removed (Figure 14C). Analysis of the BNP-32 expression from wild type yeast and the isogenic *CYM1* disruptant will be analysed by RIA's using the Shionoria-BNP system from Electra-Box Diagnostica ApS.

5 This assay is specific for BNP-32.

Expression of proCCK, proBNP and Cym1p

Human proCCK was expressed as a fusion protein between the prepro leader sequence
10 of yeast α -mating factor and proCCK (preproMf α 1p-proCCK). The fusion construct was expressed on multi-copy plasmids, with constitutively gene transcription from the phosphoglycerate kinase promoter. "ProCCK expression" refers to expression using pRS426 preproMf α 1p-proCCK, which was used in all yeast strains with exception of BY4705 and the isogenic yapsins deletion strains, where proCCK was expressed from
15 pRS423 preproMf α 1p-proCCK. Human proBNP was also expressed as a fusion protein between the prepro leader sequence of yeast α -mating factor and proBNP (preproMf α 1p-proBNP) (Figure 14A). *CYM1* expression was driven by its own promoter. Plasmid constructs, and oligonucleotides used are listed in Table 2.

20 *BNP radioimmunoassay*

Antibody 98192 is specific for the N-terminus of proBNP (Gøtze et al., 2002).

Chromatography

25 FPLC chromatography was performed on a Superdex 200 column on a Äkta purifier (Amersham Pharmacia Biotech). In the 50 mM Na-phosphate buffer, 100 mM NaCl and 6 M Guanidin were included.

Strain construction

30

A $\Delta yps1::TRP1$ (LJY440) and a $\Delta cym1::LEU2 \Delta yps1::TRP1$ (LJY431) strain in a BJ2168 background were constructed by the one step gene disruption technique as described above using the oligonucleotides, *YPS15'*GD400 and *YPS13'*GD400 (Table 2) for gene disruption. The PCR product was transformed into BJ2168 and LJY430. Verification of
35 the correct integration of the disruption cassette was analysed by PCR using *yps15'*DC and *yps13'*DC (Table 2).

Identification of the gene encoding the Cym1 orthologue in *Pichia pastoris* or *Pichia methanolica*

- Identification of the unknown genes from *Pichia pastoris* and *Pichia methanolica*
- 5 encoding the Cym1p orthologues of *Saccharomyces cerevisiae* will be carried out in similar manner, using the same set of degenerated primers mentioned below. *Pichia pastoris* and *Pichia methanolica* will be referred to as *Pichia* in the following text.

- By alignment of the orthologous Cym1 proteins of *Saccharomyces kluyveri* and
- 10 *Schizosaccharomyces pombe* to Cym1p from *Saccharomyces cerevisiae*, there was identified a number of identical amino acid sequences. From these sequences it is possible to synthesize degenerated oligonucleotides (Table 3) that will bind to the complementary DNA strands of *CYM1* in all three species, and thus to the *CYM1* gene of *Pichia*. Amplification of the genomic sequence will initially be carried out by using high
- 15 quality genomic DNA as template, *Pichia-CYM1-Ia* and *Pichia-CYM1-Ib* and the *Pwo* polymerase (Roche). The amplified sequence with an expected size of approximately 525 bp will be cloned in pBlunt or a similar vector and sequenced with vector specific primers. If no band appear from the initial amplification, a second round of PCR will be performed with the two nested primers, *Pichia-CYM1-IIa* and *Pichia-CYM1-IIb* using 1 µl
- 20 of the first PCR reaction as template. The expected product is approximately 270 bp and will be cloned in pBlunt and sequenced with M13 forward and M13 reverse primers. From the obtained sequence there will be synthesized sequence specific primers, two nested sense and two nested antisense specific primers. Using one of the sense primers it is possible to obtain a PCR product with *Pichia-CYM1-IIIb* using high
- 25 quality genomic DNA as template. This product of ~2100 bp will be cloned and sequenced. If it fails to produce a band of ~2100 bp, it would be necessary to isolate mRNA from *Pichia*, produce double stranded cDNA and ligate adaptors to the ends as described by the manual to the Clontech Marathon cDNA Amplification Kit (BD (Becton, Dickinson and Company)). Using the two sequence specific sense primers it is possible
- 30 to obtain the 3' end of the mRNA of approximately 2600 bp and the sequence specific antisense primers to amplify the 5' end including the sequence encoding the hypothetical active site, HXXEH motif. Synthesis of sequence specific oligonucleotides from the 5' and 3' untranslated region, full-length cDNA encoding the Cym1 orthologue in *Pichia* can be cloned.

Table 3

	Amino acid sequence	K Y P V R D P
	Oligo <i>Pichia</i> -CYM1-Ia	5' AARTAYCCXGTXMGXGAYCC 3'
5		
	Amino acid sequence	H P S N A K
	Oligo <i>Pichia</i> -CYM1-Ib	3' GTRGGXWSXTTRCGXTTY 5'
	Amino acid sequence	D P F F K M
10	Oligo <i>Pichia</i> -CYM1-IIa	5' GAYCCXTTYTTYAARATG 3'
	Amino acid sequence	G V V Y N E M
	Oligo <i>Pichia</i> -CYM1-IIb	3' CCXCAXCAXATRTTRCTYTAC 5'
15	Amino acid sequence	E K G G A Y G
	Oligo <i>Pichia</i> -CYM1-IIIb	3' CTYTTYCCXCCXCGXATRCC 5'

X-Inosine, degenerated oligonucleotides follow the International Union of Biochemistry (<http://www.chem.qmul.ac.uk/iubmb/misc/naseq.html>).

20

Genedisruption of Cym1 orthologue in *Pichia pastoris* or *Pichia methanolica*

The sequence encoding the Cym1 orthologue in *Pichia* should be cloned in a vector like pBluescript-II in *Hind*III and *Sac*I or a similar vector, if these are not present in the

25 *ORF*. Insertion of the *ORF* in *Hind*III and *Sac*I sites removes most of the multiple cloning sites from the vector, which ease the possibility to find restriction enzyme sites that are only present in the *ORF*. Cloning of the *KanMX* cassette within the *ORF*, preferentially so that 1000 bp of the *Pichia CYM1* are present on each site of the *KanMX* cassette creates the *Pichia CYM1* disruption cassette. This construct can then be

30 amplified by PCR, using primers at specific for the 5' and 3' end of the *Pichia CYM1* gene. Transformation of the PCR product into strains of *Pichia* followed by selection of transformants on YPD plates containing 100 µg/ml geneticin (G-418). Verification of the correct integration into the *Pichia* genome should be tested by colony PCR, using *Pichia* sequence specific *CYM1* primers that binds 5' and 3' to the *KanMX* cassette. From the

35 size of the PCR product it is possible to distinguish whether the integration event is correct.

Expression of foreign proteins and peptides in *Pichia pastoris*

For expression of peptides one could use the expression vector, pPIC α (inducible expression) or pGAPZ α (constitutive expression) both from Invitrogen. Both these
5 vectors use the preprosequence of the α -mating factor from *Saccharomyces cerevisiae* to direct the fusion peptide through the secretory pathway. Within the Golgi apparatus the preprosequence of the α -mating factor is removed and the peptide of interest released to the media. If it's proteins that should be expressed, both vectors mentioned
10 above can be used without the preprosequence of the α -mating factor (pPICZ and pGAPZ), where the heterologus expressed protein is cytosolic located and can be isolated from intact cells.

Expression of foreign proteins and peptides in *Pichia methanolica*

15 Expression of proteins and peptides in *Pichia methanolica* is performed in a similar manner as in *Pichia pastoris*, where plasmids are available both for *intracellular* expression and for secretion to the media. *Intracellular* expressed proteins can be cloned into pMET (Invitrogen) and for secretion in pMET α (Invitrogen). The pMET α contain the preprosequence of the α -mating factor from *Saccharomyces cerevisiae* as
20 used for expression in *Pichia pastoris*. Expression is induced by methanol in this system.

Results

The influence of growth conditions on the CCK-22 processing

25

The intra- and extracellular fraction of CCK-22 was measured from BJ2168 expressing proCCK. The intracellular fraction remained unaltered whether the cells were in exponential growth or had reached stationary phase (Fig. 2). However, the relative amount of secreted CCK-22 changed dramatically when the cells reached stationary
30 phase. During exponential growth the fraction of CCK-22 was 23%, but in the stationary phase (after 270 min) the fraction increased to 37% (Fig. 2). Hence, for the experiments described herein only exponentially growing cells were used.

The significance of the Lys residue in the release of CCK-22

To evaluate the role of the Lys residue in proteolysis, transformants of BJ2168 with the two expression constructs, proCCK and proCCK (K→A) were grown to late exponential phase and the culture media collected. The media from each strain was subjected to gel chromatography and the content of Gly-extended CCK in the collected fractions were measured with Ab 7270. CCK from the wild type media eluted in two major peaks at $K_d = 0.8$ and 1.1 (Fig. 3 A) in accordance with the previously established elution positions for CCK-22-Gly and CCK-8-Gly (Cantor et al., 1987; Rourke et al., 1997), while the proCCK (K→A) only gave rise to CCK-8-Gly and a larger form eluting at a $K_d = 0.6$ (Fig. 3 C). The two peaks eluting at $K_d = 0.7$ and 0.8 for the wt construct (Fig. 3 B) correspond to C-terminally extended CCK-22 and CCK-22-Gly, respectively (Rourke et al., 1997), whereas no CCK-22 immuno-reactivity was observed in these positions for the proCCK (K→A) construct (Fig. 3 D). However, a small peak of immuno-reactivity was seen at $K_d = 0.55$ which may be due to the slight cross reactivity of Ab 89009 with a larger unprocessed fragment (Paloheimo et al., 1994). To investigate the effect of substituting Arg for Lys, proCCK (K→R) was transformed into BJ2168. Media from transformants were analysed before and after tryptic cleavage. The fraction of proCCK processed to CCK-22 was similar to that seen for wild type CCK (Fig. 11).

Analysis of secreted CCK peptides by mass spectrometry

Media collected from BJ2168 transformed with proCCK were analysed by mass spectrometry (Figure 12). The fragments obtained correspond to the processing leading to CCK-39, CCK-22 and CCK-8. Two peptides were identified N-terminal of Lys⁶¹ (Tyr⁴⁵-Val⁶⁰, 1805.0 Da and Tyr⁴⁵-Lys⁶¹, 1932.2 Da). It appeared likely that the former was a carboxypeptidase degradation product of the latter. To elucidate this question and in an attempt to identify the C-terminal extended CCKs, the present inventors produced a disruption strain in which both *KEX2*, encoding the serine protease responsible for the processing to CCK-8, and the carboxypeptidase encoded by *KEX1* were mutated. Following transformation of proCCK into this *kex2 kex1* strain (LJY22) and subsequent analysis of the secreted peptides the inventors found only the peak corresponding to Tyr⁴⁵-Lys⁶¹. The same pattern, with only the peak corresponding to Tyr⁴⁵-Lys⁶¹ was seen using single gene disruption of *KEX1* and *KEX2* to express proCCK (data not shown). Thus the Tyr⁴⁵-Val⁶⁰ must be a degradation product in accordance with CCK-22 arising from cleavage after Lys⁶¹. Additional fragments were discovered by CCK

expression in the *kex2 kex1* strain corresponding to processing leading to CCK-61 (not identified in mammals), CCK-58, C-terminal extended CCK-39 and C-terminal extended CCK-22 (Figure 12), whereas none of the peptides corresponding to CCK-8 could be identified, in accordance with our previous work showing that Kex2p is responsible for
5 this processing (Rourke et al., 1997).

Kex2p is involved in the biosynthesis of CCK-22

Previous analysis of CCK peptides secreted from a *kex2* strain as well as the results
10 obtained by mass spectrometry indicate that the cleavage at Lys⁶¹ releasing CCK-22 can occur without the involvement of Kex2p. However, the *kex2* strain shows a decrease in CCK-22 concentration. ProCCK was expressed both in the vacuole protease deficient and the isogenic *kex2* strain (BJ2168 and LJY23) and the processed intra- and extracellular fractions of CCK-22 from exponentially growing cells were measured.
15 Approximately 28% of the intracellular CCK content was processed after Lys⁶¹ in BJ2168, whereas only 6% was processed within the *kex2* strain. Analysis of secreted CCK peptides showed that the media collected from BJ2168 expressing proCCK contained approximately 20% CCK-22, whereas from the *kex2* mutant, the amount was reduced to 5%. These results indicate that Kex2p is involved in the processing leading
20 to CCK-22. However, there are other proteases that can perform the cleavage at Lys⁶¹.

In vitro assay of Lys⁶¹ cleavage

To investigate the nature of the protease(s) in addition to Kex2p that are able to
25 perform the endoproteolytic cleavage after the single Lys⁶¹ residue of proCCK, an *in vitro* assay was established using crude preparations from of *S. cerevisiae* and synthetic CCK-33 as substrate.

Using extract from the vacuole protease deficient strain, BJ2168, there was an
30 extensive N-terminal degradation, and the recovery of measurable CCK was less than 10% of the control without yeast extract. Because the assay depends on the intact N-terminus of CCK-22 for the antibody to bind, the inventors created a strain where some of the known *S. cerevisiae* aminopeptidases were deleted. The Y14953 strain (*ape1*) was used as parental strain in which the *APE2* and *APE3* genes were also
35 deleted. Using this LJY123 strain to prepare the cell extract there was a 2-3 fold better recovery of immuno-reactivity compared to the recovery seen with BJ2168.

Processing to CCK-22 depends on metal ions

The nature of the protease performing the cleavage of synthetic human CCK-33 to CCK-22 was analysed by inclusion of a number of different inhibitors with the extract
5 from LJY123. The results showed that only the addition of a metal chelating agent inhibited proteolysis of CCK-33 to CCK-22 (Fig. 4 A).

The metal dependency of the protease was tested *in vitro*, after the activity initially was inhibited by addition of 1 mM EDTA. Reconstitution of the activity leading to maturation
10 of CCK-22 was tested by addition of different divalent cations in 0.2 mM surplus. Addition of Zn^{2+} , Co^{2+} and Mn^{2+} could reestablish the protease activity, whereas Ca^{2+} , Cu^{2+} or Mg^{2+} had no effect (Fig. 4 B) in accordance with the properties of known metalloproteases, which are only activated by Zn^{2+} , Co^{2+} and Mn^{2+} . Reactivation using increasing Zn^{2+} concentrations showed a biphasic pattern, with Zn^{2+} acting inhibitory at
15 concentrations above 5 mM (data not shown).

The time course of CCK-cleavage by Zn^{2+} and Mn^{2+} reactivated metalloproteases was analysed using cell extract from LJY123, after initial inhibition with 1 mM EDTA. Reactivation was performed by addition of 1.2 mM Zn^{2+} or Mn^{2+} followed by incubation
20 for 30, 60 and 120 min. In this assay and the following *in vitro* protease assays the inventors used the N-terminal acetylated CCK-33-Gly (Ac-CCK-33-Gly) as substrate, which resulted in much slower non-specific degradation. Measurement of the CCK-22 immuno-reactivity before and after tryptic cleavage using Ab 89009 showed no difference in the activation potency between Zn^{2+} and Mn^{2+} at 30 and 60 min, however
25 after 120 min 10% more CCK-22 immuno-reactivity was measured using Mn^{2+} as activator compared to Zn^{2+} (Fig. 5). This increase in immuno-reactivity is probably due to an inhibition of degradation following addition of Mn^{2+} here as well as to the yeast cell extracts used in Table 4.

30 Table 4. Metalloproteases in *Saccharomyces cerevisiae*. Search performed in Swiss-Prot Sequence Retrieval System (SRS) <http://www.expasy.ch/>. Protease assay performed in two independent assays (A and B) using extracts from the metalloprotease deficient strains. The amount of CCK-22 is measured with Ab 89009 and the total amount of CCK is measured after tryptic cleavage with Ab 89009. Putative metalloproteases are marked
35 with *.

Name	Swiss-Prot acc #	ORF	CCK-22 [nM]	Total CCK [nM]	Fraction CCK-22
------	---------------------	-----	----------------	----------------	--------------------

			A₁	B₁	A₂	B₂	A₁/ A₂	B₁/B₂
AAP1	<u>P37898</u>	YHR047c	3.2	3.2	36	32	0.09	0.10
AFG3	<u>P39925</u>	YER017c	2.8	2.4	23	24	0.12	0.10
APE1	<u>P14904</u>	YKL103c	4.4	3.7	34	28	0.13	0.13
APE2	<u>P32454</u>	YKL157w						
APE3	<u>P37302</u>	YBR286w						
DPP3	<u>Q08225</u>	YOL057w	4.3	4.6	31	35	0.14	0.13
LTA4	<u>Q10740</u>	YNL045w	3.9	4.0	32	29	0.12	0.13
MIP1	<u>P35999</u>	YKL134c	3.4	3.3	34	33	0.10	0.10
PRD1	<u>P25375</u>	YCL057w	2.4	2.8	28	26	0.09	0.11
QRI7*	<u>P43122</u>	YDL104c	2.8	2.9	23	24	0.12	0.12
RCA1	<u>P40341</u>	YMR089c	4.2	3.5	31	27	0.14	0.13
STE23	<u>Q06010</u>	YLR389c	2.6	2.3	25	20	0.10	0.12
STE24	<u>P47154</u>	YJR117w	3.4	2.8	19	17	0.18	0.16
YBS4*	<u>P38244</u>	YBR074w	2.6	2.7	27	29	0.10	0.09
YHR3*	<u>P38821</u>	YHR113w	2.5	2.4	26	26	0.10	0.09
YHT2*	<u>P38836</u>	YHR132c	3.8	3.3	34	32	0.11	0.10
YIK8*	<u>P40483</u>	YIL108w	5.7	5.2	43	39	0.13	0.13
YIN7*	<u>P40462</u>	YIL137c	3.5	2.8	23	20	0.15	0.14
YK18*	<u>P36132</u>	YKR038c	2.9	2.1	23	20	0.13	0.11
YME1	<u>P32795</u>	YPR024w	2.3	2.7	25	25	0.09	0.11
MAS2	<u>P11914</u>	YHR024c	ND, Lethal genes					
MAS1	<u>P10507</u>	YLR163c						
AXL1	<u>P40851</u>	YPR122w	3.5	3.4	31	30	0.11	0.11
CYM1*	<u>P32898</u>	YDR430c	0.6	0.4	42	39	0.01	0.01
YOJ8*	<u>Q12496</u>	YOL098c	3.6	3.8	35	34	0.10	0.12

Extracellular yapsin activity

To investigate whether any protease activity is secreted or attached extracellularly to the plasma membrane, the protease activity was assayed in media and with intact
5 yeast cells. No degradation of CCK-33 occurred after 1 h of incubation at 30°C using media from exponential growing LJY123 cells in accordance with earlier observations (Rourke et al., 1997). During incubation with intact yeast cells, cleavage to expose the N-terminus of CCK-22 could be measured (Fig. 6) however, this protease activity could not be abolished by the inhibitors investigated (data not shown). By using intact cells
10 containing gene disruptions of *YPS1*, *YPS2* and *YPS3* (Table I) the fraction of processed CCK-22 decreases by deletion of each of the three aspartyl proteases compared to wild type cells (Fig. 6). These data show that the three proteases all have extracellular protease activity, which can cleave at Lys⁶¹ in proCCK. Preliminary results indicate that gene disruption of *YPS7* decreases extracellular Lys⁶¹ processing in amounts
15 comparable to the *YPS1* deletion (unpublished results).

Expression of proCCK in metalloprotease deficient strains

Based on previously described metalloproteases in *S. cerevisiae* with endoproteolytic
20 activity (Adames et al., 1995; Schmidt et al., 2000), gene deletion strains of *AXL1* (LJY201) and *STE24* (LJY202) were initially prepared in BJ2168. ProCCK expression in these strains showed that proteolysis after Lys⁶¹ was unchanged compared to wild type, and it was decided to test the remaining metalloprotease deficient strains LJY123, LJY203, LJY204 and the metalloprotease deficient strains obtained through Euroscarf
25 (Table I) for their ability to secrete CCK-22 (mitochondrial peptidases were not included). The CCK-22 immuno-reactivity did not change significantly among the CCK producing metalloprotease deficient strains (data not shown), and no protease responsible for the processing of heterologous expressed proCCK to CCK-22 was identified by this approach.

30

CYM1 encodes a protease that can release the free N-terminus of CCK-22

Cell extracts were prepared from each of the viable metalloprotease deficient strains and tested in the *in vitro* protease assay to investigate whether any reduction in
35 proteolysis was measurable. In this assay 1 mM Mn²⁺ and 1 mM bestatin were included prior to the addition of Ac-CCK-33-Gly, since it was found that the recovery was 80-90% compared to 30% without addition of these aminopeptidase inhibitors (data not

shown). Deletion of *CYM1* almost abolished the protease activity, whereas none of the other metalloprotease deficient strains showed a significant change in the biosynthesis of CCK-22 (Table 4).

- 5 Expression of *CYM1* on a multicopy plasmid increases the fraction of matured CCK-22 *in vitro*

To determine whether the amount of synthesized CCK-22 correlates with the amount of Cym1p *in vitro*, Cym1p was expressed on a multicopy plasmid and the fraction of
 10 synthesized CCK-22 analysed over time. Cell extract from BJ2168 transformed with pRS425 *CYM1* and the control transformed with the empty pRS425 vector were used in the *in vitro* protease assay with 1 mM Mn²⁺ in which the reactions were terminated after 15, 30, 45 and 60 min. The CCK-22 immuno-reactivity was measured with Ab 89009 and the remaining CCK-33 was measured with the same antibody after tryptic
 15 cleavage (Fig. 7). Expression of *CYM1* on a multicopy plasmid enhanced the rate of CCK-22 production several fold. However, the inventors also observed an increased degradation of CCK-33 and CCK-22 (Fig. 7 B). When the same experiment was performed at pH 6.0 and pH 7.5, there was a dramatically increased degradation and after 30 min incubation the CCK immuno-reactivity was undetectable at pH 6.0 (data
 20 not shown). These results show that the Lys-specific cleavage in CCK-22 maturation *in vitro* is dependent on the amount of Cym1p.

Expression of proCCK in *cym1* mutant strain enhances CCK secretion

- 25 To elucidate the role of *CYM1* in the biosynthesis of CCK-22 *in vivo*, gene deletions of *CYM1* were prepared in the vacuole protease deficient strain, BJ2168, and isogenic *kex2* strain. Deletion of *CYM1* resulted in an approximately 40% increase in the total amount of proCCK within the cells (Fig. 8 A) accompanied by a similar decrease in CCK-22 independent of *KEX2* disruption (Fig. 8 C). Also the secreted amount of total
 30 CCK in the *cym1* strains increased with more than 60% (Fig. 8 B), but unlike the fractional decrease in intracellular CCK-22 there was an increase in the extracellular fractions of CCK-22 compared to vacuole protease deficient strain and the isogenic *kex2* strain (Fig. 8 D).

Expression of CCK K→A mutant leads to intracellular CCK accumulation comparable to the accumulation of wild type CCK in a *cym1* strain

- 5 The observations that a gene disruption of *CYM1* causes an increase in intracellular concentrations of CCK (Fig. 8 A) raise the question whether the proteolytic activity of Cym1p leads to degradation of CCK-22 prior to translocation into the ER. Therefore, the inventors examined the intracellular CCK content in strains expressing CCK where the maturation of CCK-22 has been eliminated by using the Lys⁶¹ → Ala⁶¹ mutant.
- 10 Transformants of this CCK mutant in the vacuole protease deficient strain, BJ2168 and the isogenic *cym1* strain were analysed using Ab 7270 after trypsin and carboxypeptidase B treatment and there was an increase in the intracellular CCK immuno-reactivity for this construct compared to expression of wild type CCK (Fig. 9). Mutant CCK (K→A) and wild type CCK transformants resulted in an increase in the
- 15 intracellular proCCK concentration when expressed in BJ2168 and the *CYM1* disruption strain, respectively. The increase in intracellular proCCK was not additive showing that proteolytic activity of Cym1p leads to degradation of CCK-22 prior to translocation into the ER.
- 20 Expression of proCCK in aspartyl protease deficient strains
- The Lys⁶¹-specific cleavage of proCCK was analysed in null mutants of *YPS1*, -2 and -3, where the intra- and extracellular amount of CCK-22 was measured from exponentially growing cells of wild type yeast, BY4705 and the isogenic aspartyl protease deficient
- 25 strains transformed LJY13, -14 and -15 with proCCK. Both intra- and extracellular CCK immuno-reactivity of BY4705 was lowered more than 10 fold compared to the vacuolar protease deficient strain, BJ2168 (data not shown). The intracellular fraction of CCK-22 decreased significantly from approximately 28% in wild type cells to 17% in the *yps1* strain, whereas no additional decrease could be measured by gene disruptions of *YPS2*
- 30 and *YPS3* (Fig. 10 A). The extracellular fraction of CCK-22 did however show that Yps1p, Yps2p and Yps3p all are involved in the biosynthesis of CCK-22 and that the triple mutant reduced the fraction of CCK-22 to 2/3 compared to wild type yeast (Fig. 10 B).
- 35 *CYM1* disruption leads to a two-fold increase in the total amount of secreted wild type CCK as well as the CCK K→R mutant

To elucidate whether Cym1p cleaves C-terminally to a single Arg residue, the CCK (K→R) mutant was expressed in the vacuolar protease deficient strain, BJ2168 and the isogenic *cym1* strain. The concentration of both intra- and extracellular CCK was compared to wild type CCK expressed in these strains. The total amount of the mutated CCK (K→R) was increased both intra- and extracellular comparable to wild type CCK (Fig. 11). Both wild type CCK and the Lys⁶¹→Arg⁶¹ mutant showed more than a two-fold increase in the measurable amount of extracellular CCK when expressed in the *cym1* strain (Fig. 11 B).

10 Usage of Cym1p activity in the synthesis of peptides

Another aspect of the invention is to use the activity from Cym1p, either expressed from its own promoter or from a strong constitutive promoter such as *PGK1*, *ADH1* or *TPI1*, or the induceable *GAL1* promoter to produce an increased amount cytosolic Cym1p activity. As previously mentioned, the synthesis of CCK-22 is significantly increased when *CYM1* is transcribed from its own promoter on a 2μ plasmid (Fig. 7). Transcription can either be performed from a plasmid containing the promoter, *CYM1* and a terminator, or by introducing the desired promoter into the genome by heterologous recombination to substitute the endogenous promoter of *CYM1*.

20

The activity can be used intracellularly to generate peptides that do not require post-translational modifications from the secretory pathway, such as disulfide bond formation, *N*- and *O*-glycosylation or exoprotease activity.

25 The role of Cym1p cytosolic activity in intracellular peptide synthesis is shown in the biosynthesis of CCK-22 in wild type cells compared to the isogenic strain with a *CYM1* disruption, which shows a significant increase in the amount of CCK-22 (Fig. 8C). Synthesis of the peptide of interest should be performed in such a way that translocation into the endoplasmatic reticulum (ER) is avoided. This can be performed either by removal of the hydrophobic amino-terminal signal sequence from proteins that enter the ER post-translationally, or by expression in a temperature sensitive secretory mutant such as *sec61*, which abolishes translocation of secretory peptides into the ER when the temperature is elevated to 37°C.

35 The propeptide or prepropeptide of interest will then be cytosolically located and a potential substrate for Cym1p. Release of the peptide from its precursor will be carried out by the Cym1p activity by introduction of the cleavage site seen from proCCK, which results in the release of Gly-extended CCK-22 after endoproteolytic processing C-

terminal to Lys⁶¹ (Ser-Ile-Val-Lys⁶¹ ↓) (Fig. 13A). If the peptide of interest is GLP1, synthesis can be performed as a fusion to a Cym1p cleavage site, which could be part of proCCK (Fig. 13B). The peptide of interest will then accumulate in the cytosol and can be purified from sedimented cells after lysis.

5

Expression of proBNP in *cym1* mutant strain enhance proBNP secretion

To elucidate the role of *CYM1* in the biosynthesis of proBNP *in vivo*, proBNP was expressed in the the vacuole deficient strain, BJ2168 and the three protease deficient
 10 isogenic strains, $\Delta cym1::LEU2$ (LJY430), $\Delta yps1::TRP1$ (LJY440) and a $\Delta cym1::LEU2$ $\Delta yps1::TRP1$ (LJY431). Deletion of *CYM1* resulted in an approximately 100% increase in the total amount of secreted proBNP, whereas the proBNP secretion was independent on disruption of the gene encoding the aspartyl protease, Yps1p and was thereby similar to the wildtype strain (Fig. 15A). Disruption of both Cym1 and Yps1 was as
 15 expected similar to the secreted amount in a *cym1* mutant (Fig. 15A).

Analysis of the proBNP in secreted from a *cym1* mutant

To analyse the proBNP expressed in *Saccharomyces cerevisiae*, media from a *cym1*
 20 mutant was applied to FPLC chromatography and analysed by RIA using Ab. 98192. The peak eluting from fraction 34-39 corresponds to intact proBNP, whereas the peak eluting in fraction 53-62 is a processed form of proBNP, most likely the proBNP fragment 1-76 (Fig. 15B). The release of fragment 1-76 and BNP-32 from proBNP, is due to cleavage after a single Arg residue and is probably due to either Kex2 or Yps1
 25 activity.

Discussion

The secreted polypeptides varies with the growth conditions, the fraction of CCK-22 increasing when the culture reaches stationary phase, while the intracellularly processed fraction remains unaltered under stress conditions. The increase in
 30 extracellular cleavage to CCK-22 as the cells enter stationary phase could indicate that extracellular endoproteases with the ability to process proCCK to CCK-22 are secreted or expressed on the cell membrane. It is known that the aspartyl proteases, Yps1p and Yps2p, exhibit cell surface activity (Komano et al., 1998). In addition, it has previously been shown that heterologous peptide expression in a *yps1* strain improved the
 35 recovery of proteins and peptides like albumin, glucagon, GLP1, GLP2 and CART by inhibiting proteolysis C-terminal to mono-basic residues (Egel-Mitani et al., 2000;

Kerry-Williams et al., 1998). Thus, recent studies ((Egel-Mitani et al., 2000; Kerry-Williams et al., 1998) and those of the present inventors) show the importance of collecting secreted peptides during exponential growth in order to avoid additional extracellular processing.

5

ProCCK expressed in a vacuole protease deficient strain showed 30% intracellular processing at Lys⁶¹ in proCCK. The fraction of extracellular Lys⁶¹-processing is, however, decreased to 2/3 of the observed fraction within intact yeast cells, which reveals an intracellular degradation of CCK-22 prior to secretion. The increase in
10 extracellular proteolysis under limited nutrient resources is probably due to an activation or upregulation in transcription of the extracellular proteases under limited nutrient resources as seen with the upregulation of *YPS1* transcription during stationary phase (Gasch et al., 2000). Part of the cell surface activity can be assigned to the yapsins, Yps1p, Yps2p and Yps3p, but some extracellular activity was sustained even in
15 the triple mutant.

In the present study, the inventors have shown that deletion of *KEX2* causes a 5 fold reduction in both the intracellular and extracellular Lys⁶¹-cleavage. The *kex2* strain expressing proCCK do not only alter the cleavage of Lys⁶¹ in proCCK, it also changes
20 the intracellular retention time of CCK as the intracellular concentration of CCK peptides is reduced with more than 60%, while the extracellular CCK concentration is increased by almost 60% compared to wild type yeast. Moreover, analysis of the secreted CCK peptides from the *kex1 kex2* double mutant and the *kex2* mutant showed disappearance of the Tyr⁴⁵-Val⁶⁰ degradation product. Thus, the removal of Lys⁶¹ by
25 Kex1p was abolished in a *kex2* strain indicating an enhanced secretion rate through the *trans*-Golgi network. These results and the observations on the rapid secretion of proCCK suggest that it may be the intracellular retention caused by Kex2p that leads to an increased synthesis of CCK-22 in wild type yeast by Yps1p and probably to some extent by Kex2p.

30

The type of protease responsible for the intracellular maturation of CCK-22 was investigated in an *in vitro* protease assay using a crude extract of *S. cerevisiae* to analyse the processing of synthetic human CCK-33 to CCK-22 in the presence of different inhibitors. By not including detergents in extraction of protease activity,
35 activity from Kex2p as well as the GPI-anchored yapsins was avoided (Azaryan et al., 1993; Fuller et al., 1989; Komano et al., 1999). Of the inhibitors tested, the proteolysis was only inhibited by EDTA and 1,10 ortho-phenanthroline, and the activity could be

restored by addition of the divalent cations Zn^{2+} , Co^{2+} and Mn^{2+} . This indicated that a metalloprotease participates in the maturation of CCK-22.

None of the candidate metalloproteases contain an obvious signal peptide to direct the protein into ER. Therefore, the inventors investigated strains deficient in each of the metalloproteases with the exception of mitochondrial proteases. Expression of proCCK in each of the strains resulted in unaltered maturation of CCK-22 similar to that seen in wild type yeast. However, by using the *in vitro* protease assay the inventors identified Cym1p as an endoprotease performing post-Lys cleavage of CCK-33. That Cym1p can cleave Lys⁶¹ in proCCK was verified by overexpression studies, showing a several fold increase in enzyme activity.

Intracellular synthesis of CCK-22 was decreased in a *cym1* strain accompanied by an increased concentration of total proCCK. In contrast, the fraction of extracellular CCK-22 was increased compared to wild type yeast with a parallel increase in total CCK. These findings are in accordance with a cytosolic location of the Cym1p activity like most insulin-degrading enzymes (Bai et al., 1996) and show that it acts on the preproMf α 1p-proCCK construct prior to translocation into the endoplasmic reticulum. Thus, the pre-translocational degradation of proCCK is decreased by *CYM1* disruption and the total production increased.

Expression of proBNP as a fusionpeptide to the preproMf α 1p sequence in a *cym1 Δ* mutant shows a two-fold increase of the extracellular proBNP content compared to the wild type strain. Analysis of the secreted proBNP by chromatography disclosed that two major forms were present. One is the entire proBNP, whereas the other is the proBNP fragment 1-76, thus the biologically active BNP-32 is also synthesised, though not detectable in the present assay. The release of proBNP fragment 1-76, most likely depends on the Kex2p activity, however this could not be tested in the present assay, since the release of proBNP depends on both Kex2p and Kex1p.

All publications discussed above are incorporated herein in their entirety.

It will be appreciated by persons skilled in the art that numerous variations and/or modifications may be made to the invention as shown in the specific embodiments without departing from the spirit or scope of the invention as broadly described. The present embodiments are, therefore, to be considered in all respects as illustrative and not restrictive.

References

- Adames, N., Blundell, K., Ashby, M.N., and Boone, C. (1995). *Science* 270:464-67.
- Azaryan, A.V., Wong, M., Friedman, T.C., Cawley, N.X., Estivariz, F.E., Chen, H.C., and Loh, Y.P. (1993). *J. Biol. Chem.* 268:11968-75.
- Bai, J.P., Hong, H.J., Rothenberger, D.A., Wong, W.D., and Buls, J.G. (1996). *J. Pharm. Pharmacol.* 48:1180-84.
- 10 Becker, A.B and Roth, R.A. (1992) *PNAS* 89: 3835-3839.
- Brachmann, C.B., Davies, A., Cost, G.J., Caputo, E., Li, J., Hieter, P., and Boeke, J.D. (1998). *Yeast* 14:115-132.
- 15 Cantor, P. and Rehfeld, J.F. (1987). *Clin. Chim. Acta* 168:153-58.
- Egel-Mitani, M., Andersen, A.S., Diers, I., Hach, M., Thim, L., Hastrup, S., and Vad, K. (2000). *Enzyme Microb. Technol.* 26:671-77.
- 20 Egel-Mitani, M., Flygenring, H.P., and Hansen, M.T. (1990). *Yeast* 6:127-137.
- Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001). *Nature* 411, 494-98.
- 25 Fuller, R.S., Brake, A., and Thorner, J. (1989). *Proc. Natl. Acad. Sci. USA.* 86:1434-38.
- Gasch, A.P., Spellman, P.T., Kao, C.M., Carmel-Harel, O., Eisen, M.B., Storz, G., Botstein, D., and Brown, P.O. (2000). *Mol. Biol. Cell* 11:4241-57.
- 30 Gietz, R.D., Schiestl, R.H., Willems, A.R., and Woods, R.A. (1995). *Yeast* 11:355-60.
- Gøtze JP, Kastrup J, Pedersen F and Rehfeld JF. (2002) Quantification of pro-B-type natriuretic peptide and its products in human plasma by use of an analysis independent of precursor processing. *Clin Chem.* Jul;48(7):1035-42).
- 35 Haseloff, J. and Gerlach, W.L. (1988). *Nature*, 334: 585-91.

Hilsted, L. and Rehfeld, J.F. (1986). *Anal. Biochem.* 152:119-26.

Hooper N.M. (1994) *FEBS Letters* 354: 1-6.

- 5 Horton, R.M., Ho, S.N., Pullen, J.K., Hunt, H.D., Cai, Z., and Pease, L.R. (1993). *Methods Enzymol.* 217:270-79.

Jones, E.W. (1991). *Methods Enzymol.* 194:428-53.

- 10 Julius, D., Schekman, R., and Thorner, J. (1984). *Cell* 36:309-18.

Kerry-Williams, S.M., Gilbert, S.C., Evans, L.R., and Balance, D.J. (1998). *Yeast* 14:161-69.

- 15 Komano, H., Rockwell, N., Wang, G.T., Krafft, G.A., and Fuller, R.S. (1999). *J. Biol. Chem.* 274:24431-37.

Komano, H., Seeger, M., Gandy, S., Wang, G.T., Krafft, G.A., and Fuller, R.S. (1998). *J. Biol. Chem.* 273:31648-51.

20

Needleman, S.B. and Wunsch, C.D. (1970). *J. Mol. Biol.* 48: 443-53.

Paloheimo, L.I. and Rehfeld, J.F. (1994). *Clin. Chim. Acta* 229:49-65.

- 25 Philippsen, P., Stotz, A., and Scherf, C. (1991). *Methods Enzymol.* 194:169-82.

Rawlings, N.D. and Barrett, A.J. (1995). *Meth. Enzymol.* 248:183-228.

Rothstein, R. (1991). *Methods Enzymol.* 194:281-301.

30

Rourke, I.J., Johnsen, A.H., Din, N., Petersen, J.G.L., and Rehfeld, J.F. (1997). *J. Biol. Chem.* 272:9720-27.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory*

- 35 *Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Schmidt, W.K., Tam, A., and Michaelis, S. (2000). *J. Biol. Chem.* 275:6227-33.

Shippy, R., Lockner, R., Farnsworth, M. and Hampel, A. (1999). *Mol. Biotech.* 12: 117-29.

Waterhouse, P.M., Graham, M.W., and Wang, M.-B. (1998). Proc Natl Acad Sci U S A, 95: 13959-64.

SEQUENCE LISTING

<110> Cym1p A/S
Jønson, Lars
5 Rehfeld, Jens F.
Johnsen, Anders H.

<120> Methods for increasing the production of
a recombinant polypeptide from a host cell
10

<130> P32077PC01

<150> DK/PA 2002 01391
15 <151> 2002-09-20

<160> 68

<170> FastSEQ for Windows Version 4.0
20

<210> 1
<211> 5
<212> PRT
25 <213> Artificial Sequence

<220>
<223> Pitrilysin consensus sequence

30 <220>
<221> MISC_FEATURE
<222> (2)..(2)
<223> X = any amino acid

35 <220>
<221> MISC_FEATURE
<222> (3)..(3)
<223> X = any amino acid

40
<400> 1

His Xaa Xaa Glu His
1 5
45

<210> 2
<211> 46
<212> PRT
50 <213> Artificial Sequence

<220>
<223> Pitrilysin consensus sequence

55 <220>
<221> MISC_FEATURE
<222> (2)..(2)

2/67

<223> X = any amino acid

<220>
5 <221> MISC_FEATURE
<222> (3)..(3)
<223> X = any amino acid
<220>
<221> MISC_FEATURE
10 <222> (5)..(5)
<223> X = any amino acid

<220>
15 <221> MISC_FEATURE
<222> (6)..(6)
<223> X = any amino acid

20 <220>
<221> MISC_FEATURE
<222> (9)..(9)
<223> X = any amino acid

25
<220>
<221> MISC_FEATURE
<222> (10)..(10)
<223> X = any amino acid
30

<220>
<221> MISC_FEATURE
<222> (11)..(11)
35 <223> X = any amino acid

<220>
<221> MISC_FEATURE
40 <222> (12)..(12)
<223> X = any amino acid

<220>
45 <221> MISC_FEATURE
<222> (14)..(14)
<223> X = any amino acid

50 <220>
<221> MISC_FEATURE
<222> (15)..(15)
<223> X = any amino acid

55
<220>
<221> MISC_FEATURE
<222> (16)..(16)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
5 <222> (17)..(17)
<223> X = any amino acid
<220>
<221> MISC_FEATURE
<222> (18)..(18)
10 <223> X = any amino acid

<220>
<221> MISC_FEATURE
15 <222> (19)..(19)
<223> X = any amino acid

<220>
20 <221> MISC_FEATURE
<222> (20)..(20)
<223> X = any amino acid

25 <220>
<221> MISC_FEATURE
<222> (21)..(21)
<223> X = any amino acid

30
<220>
<221> MISC_FEATURE
<222> (22)..(22)
<223> X = any amino acid
35

<220>
<221> MISC_FEATURE
<222> (23)..(23)
40 <223> X = any amino acid

<220>
<221> MISC_FEATURE
45 <222> (24)..(24)
<223> X = any amino acid

<220>
50 <221> MISC_FEATURE
<222> (25)..(25)
<223> X = any amino acid

55 <220>
<221> MISC_FEATURE
<222> (26)..(26)
<223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (27)..(27)
5 <223> X = any amino acid
<220>
<221> MISC_FEATURE
<222> (28)..(28)
10 <223> X = any amino acid

<220>
<221> MISC_FEATURE
<222> (29)..(29)
15 <223> X = any amino acid

<220>
<221> MISC_FEATURE
20 <222> (30)..(30)
<223> X = any amino acid

<220>
25 <221> MISC_FEATURE
<222> (31)..(31)
<223> X = any amino acid

30 <220>
<221> MISC_FEATURE
<222> (32)..(32)
<223> X = any amino acid

35
<220>
<221> MISC_FEATURE
<222> (33)..(33)
<223> X = any amino acid
40

<220>
<221> MISC_FEATURE
<222> (34)..(34)
45 <223> X = any amino acid

<220>
<221> MISC_FEATURE
50 <222> (35)..(35)
<223> X = any amino acid

<220>
55 <221> MISC_FEATURE
<222> (36)..(36)
<223> X = any amino acid or is absent

5/67

```

    <220>
    <221> MISC_FEATURE
    <222> (37)..(37)
    <223> X = any amino acid or is absent
5  <220>
    <221> MISC_FEATURE
    <222> (40)..(40)
    <223> X = any amino acid

10
    <220>
    <221> MISC_FEATURE
    <222> (42)..(42)
    <223> X = any amino acid
15
    <220>
    <221> MISC_FEATURE
    <222> (43)..(43)
20 <223> X = any amino acid

    <220>
    <221> MISC_FEATURE
25 <222> (44)..(44)
    <223> X = any amino acid

    <220>
30 <221> MISC_FEATURE
    <222> (45)..(45)
    <223> X = any amino acid

35 <400> 2

    Gly Xaa Xaa His Xaa Xaa Glu His Xaa Xaa Xaa Xaa Gly Xaa Xaa Xaa
    1              5              10              15

40
    Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
    20              25              30

45 Xaa Xaa Xaa Xaa Xaa Asn Ala Xaa Thr Xaa Xaa Xaa Xaa Thr
    35              40              45

    <210> 3
50 <211> 46
    <212> PRT
    <213> Artificial Sequence

55 <220>
    <223> Pitrilysin consensus sequence

    <220>
    <221> MISC_FEATURE

```

6/67

<222> (2)..(2)
<223> X = any amino acid

<220>
5 <221> MISC_FEATURE
<222> (3)..(3)
<223> X = any amino acid

10 <220>
<221> MISC_FEATURE
<222> (5)..(5)
<223> X = any amino acid

15
<220>
<221> MISC_FEATURE
<222> (6)..(6)
<223> X = any amino acid

20
<220>
<221> MISC_FEATURE
<222> (9)..(9)
25 <223> X = any amino acid

<220>
<221> MISC_FEATURE
30 <222> (10)..(10)
<223> X = any amino acid

<220>
35 <221> MISC_FEATURE
<222> (11)..(11)
<223> X = any amino acid

40 <220>
<221> MISC_FEATURE
<222> (12)..(12)
<223> X = any amino acid

45
<220>
<221> MISC_FEATURE
<222> (14)..(14)
<223> X = S or T

50
<220>
<221> MISC_FEATURE
<222> (15)..(15)
55 <223> X = any amino acid

<220>
<221> MISC_FEATURE

7/67

<222> (19)..(19)
<223> X = any amino acid
<220>
<221> MISC_FEATURE
5 <222> (20)..(20)
<223> X = any amino acid

<220>
10 <221> MISC_FEATURE
<222> (21)..(21)
<223> X = any amino acid

<220>
15 <221> MISC_FEATURE
<222> (22)..(22)
<223> X = any amino acid

20
<220>
<221> MISC_FEATURE
<222> (23)..(23)
<223> X = any amino acid
25

<220>
<221> MISC_FEATURE
<222> (24)..(24)
30 <223> X = any amino acid

<220>
<221> MISC_FEATURE
35 <222> (25)..(25)
<223> X = any amino acid

<220>
40 <221> MISC_FEATURE
<222> (26)..(26)
<223> X = any amino acid

45 <220>
<221> MISC_FEATURE
<222> (27)..(27)
<223> X = any amino acid or is absent

50
<220>
<221> MISC_FEATURE
<222> (28)..(28)
<223> X = any amino acid or is absent
55

<220>
<221> MISC_FEATURE
<222> (30)..(30)

8/67

- <223> X = any amino acid
<220>
<221> MISC_FEATURE
<222> (31)..(31)
5 <223> X = any amino acid
- <220>
<221> MISC_FEATURE
10 <222> (32)..(32)
<223> X = any amino acid
- <220>
15 <221> MISC_FEATURE
<222> (33)..(33)
<223> X = any amino acid
- 20 <220>
<221> MISC_FEATURE
<222> (34)..(34)
<223> X = any amino acid
- 25
<220>
<221> MISC_FEATURE
<222> (35)..(35)
<223> X = any amino acid
30
- <220>
<221> MISC_FEATURE
<222> (36)..(36)
35 <223> X = any amino acid
- <220>
<221> MISC_FEATURE
40 <222> (37)..(37)
<223> X = any amino acid or is absent
- <220>
45 <221> MISC_FEATURE
<222> (40)..(40)
<223> X = any amino acid
- 50 <220>
<221> MISC_FEATURE
<222> (42)..(42)
<223> X = any amino acid
- 55
<220>
<221> MISC_FEATURE
<222> (43)..(43)
<223> X = any amino acid

9/67

<220>
 <221> MISC_FEATURE
 <222> (44)..(44)
 <223> X = D or E

5

<220>
 <221> MISC_FEATURE
 <222> (45)..(45)
 10 <223> X = any amino acid

<400> 3

15 Gly Xaa Xaa His Xaa Xaa Glu His Xaa Xaa Xaa Xaa Gly Xaa Xaa Lys
 1 5 10 15

20 Tyr Pro Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Xaa Xaa Xaa
 20 25 30

Xaa Xaa Xaa Xaa Xaa Asn Ala Xaa Thr Xaa Xaa Xaa Xaa Thr
 35 40 45
 25

<210> 4
 <211> 989
 <212> PRT
 30 <213> Saccharomyces cerevisiae

<400> 4

Met Leu Arg Phe Gln Arg Phe Ala Ser Ser Tyr Ala Gln Ala Gln Ala
 35 1 5 10 15

Val Arg Lys Tyr Pro Val Gly Gly Ile Phe His Gly Tyr Glu Val Arg
 20 25 30
 40

Arg Ile Leu Pro Val Pro Glu Leu Arg Leu Thr Ala Val Asp Leu Val
 35 40 45

45 His Ser Gln Thr Gly Ala Glu His Leu His Ile Asp Arg Asp Asp Lys
 50 55 60

50 Asn Asn Val Phe Ser Ile Ala Phe Lys Thr Asn Pro Pro Asp Ser Thr
 65 70 75 80

Gly Val Pro His Ile Leu Glu His Thr Thr Leu Cys Gly Ser Val Lys
 85 90 95
 55

Tyr Pro Val Arg Asp Pro Phe Phe Lys Met Leu Asn Lys Ser Leu Ala
 100 105 110

10/67

5	Asn	Phe	Met	Asn	Ala	Met	Thr	Gly	Pro	Asp	Tyr	Thr	Phe	Phe	Pro	Phe
			115					120					125			
10	Ser	Thr	Thr	Asn	Pro	Gln	Asp	Phe	Ala	Asn	Leu	Arg	Gly	Val	Tyr	Leu
		130					135					140				
15	Asp	Ser	Thr	Leu	Asn	Pro	Leu	Leu	Lys	Gln	Glu	Asp	Phe	Asp	Gln	Glu
	145					150					155					160
20	Gly	Trp	Arg	Leu	Glu	His	Lys	Asn	Ile	Thr	Asp	Pro	Glu	Ser	Asn	Ile
					165					170					175	
25	Val	Phe	Lys	Gly	Val	Val	Tyr	Asn	Glu	Met	Lys	Gly	Gln	Ile	Ser	Asn
				180					185					190		
30	Ala	Asn	Tyr	Tyr	Phe	Trp	Ser	Lys	Phe	Gln	Gln	Ser	Ile	Tyr	Pro	Ser
			195					200					205			
35	Leu	Asn	Asn	Ser	Gly	Gly	Asp	Pro	Met	Lys	Ile	Thr	Asp	Leu	Arg	Tyr
		210					215					220				
40	Gly	Asp	Leu	Leu	Asp	Phe	His	His	Lys	Asn	Tyr	His	Pro	Ser	Asn	Ala
	225					230					235					240
45	Lys	Thr	Phe	Thr	Tyr	Gly	Asn	Leu	Pro	Leu	Val	Asp	Thr	Leu	Lys	Gln
					245					250					255	
50	Leu	Asn	Glu	Gln	Phe	Ser	Gly	Tyr	Gly	Lys	Arg	Ala	Arg	Lys	Asp	Lys
			260						265					270		
55	Leu	Leu	Met	Pro	Ile	Asp	Leu	Lys	Lys	Asp	Ile	Asp	Val	Lys	Leu	Leu
			275					280					285			
60	Gly	Gln	Ile	Asp	Thr	Met	Leu	Pro	Pro	Glu	Lys	Gln	Thr	Lys	Ala	Ser
		290					295					300				
65	Met	Thr	Trp	Ile	Cys	Gly	Ala	Pro	Gln	Asp	Thr	Tyr	Asp	Thr	Phe	Leu
	305					310					315					320
70	Leu	Lys	Val	Leu	Gly	Asn	Leu	Leu	Met	Asp	Gly	His	Ser	Ser	Val	Met
					325					330					335	
75	Tyr	Gln	Lys	Leu	Ile	Glu	Ser	Gly	Ile	Gly	Leu	Glu	Phe	Ser	Val	Asn
					340				345					350		

11/67

	Ser	Gly	Val	Glu	Pro	Thr	Thr	Ala	Val	Asn	Leu	Leu	Thr	Val	Gly	Ile	
			355					360					365				
5	Gln	Gly	Val	Ser	Asp	Ile	Glu	Ile	Phe	Lys	Asp	Thr	Val	Asn	Asn	Ile	
		370					375					380					
10	Phe	Gln	Asn	Leu	Leu	Glu	Thr	Glu	His	Pro	Phe	Asp	Arg	Lys	Arg	Ile	
	385					390					395					400	
15	Asp	Ala	Ile	Ile	Glu	Gln	Leu	Glu	Leu	Ser	Lys	Lys	Asp	Gln	Lys	Ala	
					405					410					415		
20	Asp	Phe	Gly	Leu	Gln	Leu	Leu	Tyr	Ser	Ile	Leu	Pro	Gly	Trp	Thr	Asn	
				420					425					430			
	Lys	Ile	Asp	Pro	Phe	Glu	Ser	Leu	Leu	Phe	Glu	Asp	Val	Leu	Gln	Arg	
			435					440					445				
25	Phe	Arg	Gly	Asp	Leu	Glu	Thr	Lys	Gly	Asp	Thr	Leu	Phe	Gln	Asp	Leu	
	450						455					460					
30	Ile	Arg	Lys	Tyr	Ile	Val	His	Lys	Pro	Cys	Phe	Thr	Phe	Ser	Ile	Gln	
	465					470					475					480	
35	Gly	Ser	Glu	Glu	Phe	Ser	Lys	Ser	Leu	Asp	Asp	Glu	Glu	Gln	Thr	Arg	
					485					490					495		
40	Leu	Arg	Glu	Lys	Ile	Thr	Ala	Leu	Asp	Glu	Gln	Asp	Lys	Lys	Asn	Ile	
				500					505					510			
	Phe	Lys	Arg	Gly	Ile	Leu	Leu	Gln	Glu	Lys	Gln	Asn	Glu	Lys	Glu	Asp	
			515					520					525				
45	Leu	Ser	Cys	Leu	Pro	Thr	Leu	Gln	Ile	Lys	Asp	Ile	Pro	Arg	Ala	Gly	
	530						535					540					
50	Asp	Lys	Tyr	Ser	Ile	Glu	Gln	Lys	Asn	Asn	Thr	Met	Ser	Arg	Ile	Thr	
	545					550					555					560	
55	Asp	Thr	Asn	Gly	Ile	Thr	Tyr	Val	Arg	Gly	Lys	Arg	Leu	Leu	Asn	Asp	
					565					570					575		
	Ile	Ile	Pro	Phe	Glu	Leu	Phe	Pro	Tyr	Leu	Pro	Leu	Phe	Ala	Glu	Ser	
					580				585					590			

12/67

Leu Thr Asn Leu Gly Thr Thr Thr Glu Ser Phe Ser Glu Ile Glu Asp
 595 600 605

5 Gln Ile Lys Leu His Thr Gly Gly Ile Ser Thr His Val Glu Val Thr
 610 615 620

10 Ser Asp Pro Asn Thr Thr Glu Pro Arg Leu Ile Phe Gly Phe Asp Gly
 625 630 635 640

15 Trp Ser Leu Asn Ser Lys Thr Asp His Ile Phe Glu Phe Trp Ser Lys
 645 650 655

Ile Leu Leu Glu Thr Asp Phe His Lys Asn Ser Asp Lys Leu Lys Val
 660 665 670

20 Leu Ile Arg Leu Leu Ala Ser Ser Asn Thr Ser Ser Val Ala Asp Ala
 675 680 685

25 Gly His Ala Phe Ala Arg Gly Tyr Ser Ala Ala His Tyr Arg Ser Ser
 690 695 700

30 Gly Ala Ile Asn Glu Thr Leu Asn Gly Ile Glu Gln Leu Gln Phe Ile
 705 710 715 720

35 Asn Arg Leu His Ser Leu Leu Asp Asn Glu Glu Thr Phe Gln Arg Glu
 725 730 735

Val Val Asp Lys Leu Thr Glu Leu Gln Lys Tyr Ile Val Asp Thr Asn
 740 745 750

40 Asn Met Asn Phe Phe Ile Thr Ser Asp Ser Asp Val Gln Ala Lys Thr
 755 760 765

45 Val Glu Ser Gln Ile Ser Lys Phe Met Glu Arg Leu Pro His Gly Ser
 770 775 780

50 Cys Leu Pro Asn Gly Pro Lys Thr Ser Asp Tyr Pro Leu Ile Gly Ser
 785 790 795 800

Lys Cys Lys His Thr Leu Ile Lys Phe Pro Phe Gln Val His Tyr Thr
 805 810 815

55 Ser Gln Ala Leu Leu Gly Val Pro Tyr Thr His Lys Asp Gly Ser Ala
 820 825 830

Leu Gln Val Met Ser Asn Met Leu Thr Phe Lys His Leu His Arg Glu

13/67

835

840

845

5 Val Arg Glu Lys Gly Gly Ala Tyr Gly Gly Gly Ala Ser Tyr Ser Ala
850 855 860

10 Leu Ala Gly Ile Phe Ser Phe Tyr Ser Tyr Arg Asp Pro Gln Pro Leu
865 870 875 880

15 Lys Ser Leu Glu Thr Phe Lys Asn Ser Gly Arg Tyr Ile Leu Asn Asp
885 890 895

Ala Lys Trp Gly Val Thr Asp Leu Asp Glu Ala Lys Leu Thr Ile Phe
900 905 910

20 Gln Gln Val Asp Ala Pro Lys Ser Pro Lys Gly Glu Gly Val Thr Tyr
915 920 925

25 Phe Met Ser Gly Val Thr Asp Asp Met Lys Gln Ala Arg Arg Glu Gln
930 935 940

30 Leu Leu Asp Val Ser Leu Leu Asp Val His Arg Val Ala Glu Lys Tyr
945 950 955 960

Leu Leu Asn Lys Glu Gly Val Ser Thr Val Ile Gly Pro Gly Ile Glu
965 970 975

35 Gly Lys Thr Val Ser Pro Asn Trp Glu Val Lys Glu Leu
980 985

40 <210> 5
<211> 882
<212> PRT
<213> Schizosaccharomyces pombe

45 <400> 5

Met Asn Tyr Ala Lys Leu Ser Ile Ala Phe Ser Lys Lys Thr Ile Lys
1 5 10 15

50 Thr His Asn Cys Arg Leu Phe Gln Arg Trp Leu His Val Gly Asp Lys
20 25 30

55 Val His Asp Phe Arg Val Val Asp Thr Lys Lys Val Pro Glu Leu Gln
35 40 45

Leu Asn Tyr Thr Arg Leu Lys His Glu Pro Thr Asn Ala Asp Met Ile
50 55 60

14/67

5 His Leu Asp Arg Glu Asp Pro Asn Ser Val Phe Ser Ile Gly Phe Gln
 65 70 75 80

10 Thr Pro Ala Glu Asn Asp Glu Gly Ile Pro His Ile Leu Glu His Thr
 85 90 95

15 Thr Leu Cys Gly Ser Asn Lys Tyr Pro Val Arg Asp Pro Phe Phe Lys
 100 105 110

20 Met Leu Asn Arg Ser Leu Ala Thr Phe Met Asn Ala Phe Thr Ala Ser
 115 120 125

25 Asp Phe Thr Phe Tyr Pro Phe Ala Thr Val Asn Thr Thr Asp Tyr Lys
 130 135 140

30 Asn Leu Arg Asp Val Tyr Leu Asp Ala Thr Leu Phe Pro Lys Leu Arg
 145 150 155 160

35 Lys Leu Asp Phe Leu Gln Glu Gly Trp Arg Phe Glu His Ala Asp Val
 165 170 175

40 Asn Asp Lys Lys Ser Pro Ile Ile Phe Asn Gly Val Val Tyr Asn Glu
 180 185 190

45 Met Lys Gly Gln Val Ser Asp Ser Ser Tyr Ile Phe Tyr Met Leu Phe
 195 200 205

50 Gln Gln His Leu Phe Gln Gly Thr Ala Tyr Gly Phe Asn Ser Gly Gly
 210 215 220

55 Asp Pro Leu Ala Ile Pro Asp Leu Lys Tyr Glu Glu Leu Val Lys Phe
 225 230 235 240

60 His Arg Ser His Tyr His Pro Ser Asn Ala Lys Ile Leu Ser Tyr Gly
 245 250 255

65 Ser Phe Pro Leu Glu Asp Asn Leu Ser Ala Leu Ser Glu Thr Phe Arg
 260 265 270

70 Pro Phe Ser Lys Arg Glu Leu Asn Leu Pro Asn Thr Phe Leu Lys Glu
 275 280 285

75 Phe Asp Gln Glu Lys Arg Val Val Glu Tyr Gly Pro Leu Asp Pro Val
 290 295 300

15/67

Met Ala Pro Gly Arg Gln Val Lys Thr Ser Ile Ser Phe Leu Ala Asn
 305 310 315 320

5
 Asp Thr Ser Asn Val Tyr Glu Thr Phe Ala Leu Lys Val Leu Ser Lys
 325 330 335

10 Leu Cys Phe Asp Gly Phe Ser Ser Pro Phe Tyr Lys Ala Leu Ile Glu
 340 345 350

15 Ser Gly Leu Gly Thr Asp Phe Ala Pro Asn Ser Gly Tyr Asp Ser Thr
 355 360 365

20 Thr Lys Arg Gly Ile Phe Ser Val Gly Leu Glu Gly Ala Ser Glu Glu
 370 375 380

Ser Leu Ala Lys Ile Glu Asn Leu Val Tyr Ser Ile Phe Asn Asp Leu
 385 390 395 400

25 Ala Leu Lys Gly Phe Glu Asn Glu Lys Leu Glu Ala Ile Leu His Gln
 405 410 415

30 Met Glu Ile Ser Leu Lys His Lys Ser Ala His Phe Gly Ile Gly Leu
 420 425 430

35 Ala Gln Ser Leu Pro Phe Asn Trp Phe Asn Gly Ala Asp Pro Ala Asp
 435 440 445

40 Trp Leu Ser Phe Asn Lys Gln Ile Glu Trp Leu Lys Gln Lys Asn Ser
 450 455 460

Asp Gly Lys Leu Phe Gln Lys Leu Ile Lys Lys Tyr Ile Leu Glu Asn
 465 470 475 480

45 Lys Ser Arg Phe Val Phe Thr Met Leu Pro Ser Ser Thr Phe Pro Gln
 485 490 495

50 Arg Leu Gln Glu Ala Glu Ala Lys Lys Leu Gln Glu Arg Thr Ser Lys
 500 505 510
 Leu Thr Asp Glu Asp Ile Ala Glu Ile Glu Lys Thr Ser Val Lys Leu
 515 520 525

55 Leu Glu Ala Gln Ser Thr Pro Ala Asp Thr Ser Cys Leu Pro Thr Leu
 530 535 540

16/67

Ser Val Ser Asp Ile Pro Glu Thr Ile Asp Glu Thr Lys Leu Lys Phe
 545 550 555 560

5 Leu Asp Ile Ala Gly Met Lys Ala Gln Trp Tyr Asp Leu Ala Ala Gly
 565 570 575

10 Leu Thr Tyr Ile Arg Leu Leu Leu Pro Leu Lys Asn Phe Pro Glu Ser
 580 585 590

15 Leu Ile Pro Tyr Leu Pro Val Tyr Cys Asp Ala Cys Leu Asn Leu Gly
 595 600 605

20 Thr His Ser Glu Ser Ile Gly Asp Leu Glu His Gln Ile Arg Arg Tyr
 610 615 620

Thr Gly Gly Ile Ser Ile Ser Pro Ser Ala Val Thr Asn Asn Ser Asp
 625 630 635 640

25 Val Ser Lys Tyr Glu Leu Gly Ile Ala Ile Ser Gly Tyr Ala Leu Asp
 645 650 655

30 Lys Asn Val Gly Lys Leu Val Glu Leu Ile Asn Lys Ala Phe Trp Asn
 660 665 670

35 Thr Asn Leu Ser Asn Thr Asp Lys Leu Ala Ile Met Leu Lys Thr Ser
 675 680 685

Val Ser Gly Ile Thr Asp Gly Ile Ala Glu Lys Gly His Ser Phe Ala
 690 695 700

40 Lys Val Ser Ser Ala Ser Gly Leu Thr Glu Lys Thr Ser Ile Thr Glu
 705 710 715 720

45 Gln Leu Gly Gly Leu Thr Gln Val Lys Leu Leu Ser Gln Leu Ser Arg
 725 730 735

50 Glu Glu Ser Phe Gly Pro Leu Val Glu Lys Leu Thr Ala Ile Arg Glu
 740 745 750
 Ile Leu Arg Gly Thr Ser Gly Phe Lys Ala Ala Ile Asn Ala Ser Pro
 755 760 765

55 Thr Gln His Glu Val Val Glu Lys Ala Leu Gln Lys Phe Met Lys Ser
 770 775 780

Arg Gly Val Asn Gln Gln Thr Gln Thr Lys Ser Thr Ser Lys Glu Arg

17/67

	785		790		795		800
5	Asn Gly Ile Asn Ser Ile Lys Thr Tyr His Glu Leu Pro Phe Gln Thr	805		810		815	
10	Tyr Phe Ala Ala Lys Ser Cys Leu Gly Val Pro Tyr Thr His Pro Asp	820		825		830	
15	Gly Ala Pro Leu Gln Ile Leu Ser Ser Leu Leu Thr His Lys Tyr Leu	835		840		845	
20	His Gly Glu Ile Arg Glu Lys Gly Gly Ala Tyr Gly Ala Gly Leu Ser	850		855		860	
25	Tyr Ser Gly Ile Asp Gly Val Leu Ser Phe Phe Thr Tyr Arg Asp Ser	865		870		875	880
30	Asp Pro						
	<210> 6						
	<211> 973						
	<212> PRT						
	<213> Clostridium perfringens						
	<400> 6						
35	Met Asn Phe Lys Glu Asn Asn Ile Tyr Ser Gly Phe Lys Leu Leu Asn	1	5	10		15	
40	Ile Glu Asn Leu Asn Glu Ile Gly Gly Val Gly Leu Arg Phe Glu His	20		25		30	
45	Glu Lys Thr Lys Ala Lys Leu Ile Lys Ile Leu Ser Glu Asp Asp Asn	35		40		45	
50	Lys Cys Phe Ala Ile Gly Phe Arg Thr Pro Pro Glu Asn Ser Thr Gly	50		55		60	
55	Val Pro His Ile Leu Glu His Ser Val Leu Cys Gly Ser Arg Lys Phe	65	70	75		80	
	Asn Thr Lys Glu Pro Phe Val Glu Leu Leu Lys Gly Ser Leu Asn Thr	85		90		95	
	Phe Leu Asn Ala Met Thr Tyr Pro Asp Lys Thr Ile Tyr Pro Val Ala	100		105		110	

18/67

Ser Arg Asn Glu Lys Asp Phe Met Asn Leu Met Asp Val Tyr Leu Asp
 115 120 125
 5

Ala Val Leu Tyr Pro Asn Ile Tyr Lys His Lys Glu Ile Phe Met Gln
 130 135 140
 10

Glu Gly Trp His Tyr Tyr Ile Glu Asn Lys Glu Asp Glu Leu Lys Tyr
 145 150 155 160

15 Asn Gly Val Val Tyr Asn Glu Met Lys Gly Ala Tyr Ser Ser Pro Asp
 165 170 175

20 Ser Ile Leu Tyr Arg Lys Ile Pro Gln Thr Ile Tyr Pro Asp Thr Cys
 180 185 190

Tyr Ala Leu Ser Ser Gly Gly Asp Pro Asp Glu Ile Pro Asn Leu Thr
 195 200 205
 25

Tyr Glu Glu Phe Val Glu Phe His Lys Lys Tyr Tyr His Pro Ser Asn
 210 215 220
 30

Ser Tyr Ile Phe Leu Tyr Gly Asn Gly Asp Thr Glu Lys Glu Leu Glu
 225 230 235 240

35 Phe Ile Asn Glu Glu Tyr Leu Lys Asn Phe Glu Tyr Lys Glu Ile Asp
 245 250 255

40 Ser Glu Ile Lys Glu Gln Lys Ser Phe Glu Ser Met Lys Glu Glu Ser
 260 265 270

Phe Thr Tyr Gly Ile Ala Glu Ser Glu Asp Leu Asn His Lys Ser Tyr
 275 280 285
 45

Tyr Ser Leu Asn Phe Val Ile Gly Asp Ala Thr Asp Gly Glu Lys Gly
 290 295 300
 Leu Ala Phe Asp Val Leu Ala Tyr Leu Leu Thr Arg Ser Thr Ala Ala
 50 305 310 315 320

Pro Leu Lys Lys Ala Leu Ile Asp Ala Gly Ile Gly Lys Ala Val Ser
 325 330 335
 55

Gly Asp Phe Asp Asn Ser Thr Lys Gln Ser Ala Phe Thr Val Leu Val
 340 345 350

19/67

Lys Asn Ala Glu Leu Asn Lys Glu Glu Glu Phe Lys Lys Val Val Met
 355 360 365

5

Asp Thr Leu Lys Asp Leu Val Glu Asn Gly Ile Asp Lys Glu Leu Ile
 370 375 380

10

Glu Ala Ser Ile Asn Arg Val Glu Phe Glu Leu Arg Glu Gly Asp Tyr
 385 390 395 400

15

Gly Ser Tyr Pro Asn Gly Leu Ile Tyr Tyr Leu Lys Val Met Asp Ser
 405 410 415

20

Trp Leu Tyr Asp Gly Asp Pro Tyr Val His Leu Glu Tyr Glu Lys Asn
 420 425 430

25

Leu Glu Lys Ile Lys Ser Ala Leu Thr Ser Asn Tyr Phe Glu Asp Leu
 435 440 445

30

Ile Glu Arg Tyr Met Ile Asn Asn Thr His Ser Ser Leu Val Ser Leu
 450 455 460

35

His Pro Glu Lys Gly Ile Asn Glu Lys Lys Ser Ala Glu Leu Lys Lys
 465 470 475 480

40

Lys Leu Glu Glu Ile Lys Asn Ser Phe Asp Glu Lys Thr Leu Asn Glu
 485 490 495

45

Ile Ile Asp Asn Cys Lys Lys Leu Lys Glu Arg Gln Ser Thr Pro Asp
 500 505 510

50

Lys Lys Glu Asp Leu Glu Ser Ile Pro Met Leu Ser Leu Glu Asp Ile
 515 520 525

55

Asp Lys Glu Ala Thr Lys Ile Pro Thr Glu Glu Lys Glu Ile Asp Gly
 530 535 540

Ile Thr Thr Leu His His Asp Phe His Thr Asn Lys Ile Asp Tyr Val
 545 550 555 560

60

Asn Phe Phe Phe Asn Thr Asn Ser Val Pro Glu Asp Leu Ile Pro Tyr
 565 570 575

65

Val Gly Leu Leu Cys Asp Ile Leu Gly Lys Cys Gly Thr Glu Asn Tyr
 580 585 590

20/67

Asp Tyr Ser Lys Leu Ser Asn Ala Ile Asn Ile Ser Thr Gly Gly Ile
 595 600 605

5 Ser Phe Gly Ala Ile Thr Phe Ala Asn Leu Lys Lys Asn Asn Glu Phe
 610 615 620

10 Arg Pro Tyr Leu Glu Ile Ser Tyr Lys Ala Leu Ser Ser Lys Thr Asn
 625 630 635 640

15 Lys Ala Ile Glu Leu Val Asp Glu Ile Val Asn His Thr Asp Leu Asp
 645 650 655

Asp Met Asp Arg Ile Met Gln Ile Ile Arg Glu Lys Arg Ala Arg Leu
 660 665 670

20 Glu Gly Ala Ile Phe Asp Ser Gly His Arg Ile Ala Met Lys Lys Val
 675 680 685

25 Leu Ser Tyr Ser Thr Asn Arg Gly Ala Tyr Asp Glu Lys Ile Ser Gly
 690 695 700

30 Leu Asp Tyr Tyr Asp Phe Leu Val Asn Ile Glu Lys Glu Asp Lys Lys
 705 710 715 720

35 Ser Thr Ile Ser Asp Ser Leu Lys Lys Val Arg Asp Leu Ile Phe Asn
 725 730 735

Lys Gly Asn Met Leu Ile Ser Tyr Ser Gly Lys Glu Glu Glu Tyr Glu
 740 745 750

40 Asn Phe Lys Glu Lys Val Lys Tyr Leu Ile Ser Lys Thr Asn Asn Asn
 755 760 765

45 Asp Phe Glu Lys Glu Glu Tyr Asn Phe Glu Leu Gly Lys Lys Asn Glu
 770 775 780
 Gly Leu Leu Thr Gln Gly Asn Val Gln Tyr Val Ala Lys Gly Gly Asn
 785 790 795 800

50 Tyr Lys Thr His Gly Tyr Lys Tyr Ser Gly Ala Leu Ser Leu Leu Glu
 805 810 815

55 Ser Ile Leu Gly Phe Asp Tyr Leu Trp Asn Ala Val Arg Val Lys Gly
 820 825 830

Gly Ala Tyr Gly Val Phe Ser Asn Phe Arg Arg Asp Gly Gly Ala Tyr

21/67

835

840

845

5 Ile Val Ser Tyr Arg Asp Pro Asn Ile Lys Ser Thr Leu Glu Ala Tyr
850 855 860

10 Asp Asn Ile Pro Lys Tyr Leu Asn Asp Phe Glu Ala Asp Glu Arg Glu
865 870 875 880

15 Met Thr Lys Tyr Ile Ile Gly Thr Ile Arg Lys Tyr Asp Gln Pro Ile
885 890 895

Ser Asn Gly Ile Lys Gly Asp Ile Ala Val Ser Tyr Tyr Leu Ser Asn
900 905 910

20 Phe Thr Tyr Glu Asp Leu Gln Lys Glu Arg Glu Glu Ile Ile Asn Ala
915 920 925

25 Asp Val Glu Lys Ile Lys Ser Phe Ala Pro Met Ile Lys Asp Leu Met
930 935 940

30 Lys Glu Asp Tyr Ile Cys Val Leu Gly Asn Glu Glu Lys Ile Lys Glu
945 950 955 960

Asn Lys Asp Leu Phe Asn Asn Ile Lys Ser Val Ile Lys
965 970

35

<210> 7

<211> 971

<212> PRT

<213> Borrelia burgdorferi

40

<400> 7

45 Met Lys Lys Lys Lys Ile Phe Lys Leu Ile Ser Lys Thr Tyr Leu Glu
1 5 10 15

Glu His Asp Ala Glu Gly Tyr Tyr Phe Lys His Glu Ser Gly Leu Glu
20 25 30

50 Val Phe His Leu Lys Ser Asp Ser Phe Lys Glu Asn Ala Phe Cys Ile
35 40 45

55 Ala Phe Lys Thr Ile Pro Ser Asn Asn Thr Gly Val Ala His Val Leu
50 55 60

Glu His Thr Ile Phe Cys Gly Ser Ser Lys Tyr Lys Ile Lys Asp Pro
65 70 75 80

22/67

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

23/67

Asp Ile Ala His Ile Ser Gly Ile Asn Thr Asp Leu Lys Glu Ser Ile
 325 330 335

5

Phe Ser Phe Gly Leu Gln Asn Val Val Glu Asn Lys Glu Lys Glu Phe
 340 345 350

10

Lys Asn Leu Val Phe Ser Glu Leu Lys Asn Leu Val Lys Asn Lys Ile
 355 360 365

15

Pro Lys Glu Leu Ile Lys Gly Ile Leu Phe Gly Tyr Glu Phe Ala Leu
 370 375 380

20

Lys Glu Glu Lys Gly Gln Asn Phe Pro Ile Ala Leu Met Ile Lys Ser
 385 390 395 400

Phe Lys Gly Trp Leu Asn Gly Leu His Pro Ile Lys Thr Leu Gln Thr
 405 410 415

25

Ser Tyr Tyr Ile Asn Glu Ile Thr Asn Lys Leu Glu Lys Gly Ile Tyr
 420 425 430

30

Tyr Phe Glu Asn Leu Ile Glu Lys Tyr Leu Ile Phe Asn Asn His Tyr
 435 440 445

35

Thr Leu Ile Ser Phe Ile Pro Ser His Asp Thr Glu Lys Glu Met Glu
 450 455 460

40

Glu Glu Ile Glu Lys Lys Leu Met Ala Arg Glu Ile Glu Ile Lys Gln
 465 470 475 480

Asn Pro Glu Glu Phe Leu Gln Phe Lys Lys Asp Tyr Asn Gln Phe Lys
 485 490 495

45

Lys Tyr Gln Asn Lys Lys Asp Ser Lys Ala Asp Ile Ala Lys Leu Pro
 500 505 510

50

Leu Leu Lys Ile Glu Asp Leu Pro Lys Gln Ile Glu Lys Ser Leu Asp
 515 520 525

Leu Asn Glu Ile Lys Glu Leu Asn Leu His Ser Phe Lys Phe Lys Ser
 530 535 540

55

Asn Asn Ile Phe Asn Val Asn Leu Phe Phe Lys Leu Asp Phe Leu Glu
 545 550 555 560

24/67

Lys Glu Asp Tyr Ile Tyr Leu Ser Leu Phe Lys Arg Ala Leu Gln Asp
 565 570 575

5 Leu Ser Thr Lys Asn Tyr Ser Tyr Ile Asn Ile Asn Asn Lys Ile Gln
 580 585 590

10 Asn Thr Leu Gly Gln Ile Asn Ile Ser Glu Ser Tyr Asp Glu Asp Ile
 595 600 605

15 Asp Gly Asn Ile Leu Asn Ser Phe Asn Ile Ser Phe Lys Ser Phe Asn
 610 615 620

Asn Lys Val Lys Glu Ser Phe Glu Leu Ile Lys Glu Ile Leu Ile Asn
 625 630 635 640

20 Ile Asn Phe His Asp Tyr Glu Arg Leu Lys Glu Ile Thr Leu Ser Leu
 645 650 655

25 Lys Asn Asp Phe Lys Ser Leu Leu Ile Pro Lys Gly His Leu Leu Ala
 660 665 670

30 Met Leu Arg Ser Lys Ser Lys Leu Lys Leu Asn Glu Tyr Leu Lys Glu
 675 680 685

35 Leu Gln Asn Gly Ile Thr Gly Arg Glu Phe Trp Gln Lys Ala Lys Thr
 690 695 700

Asp Thr Glu Ser Leu Lys Glu Ile Ala Asn Lys Leu Asp Asn Leu Lys
 705 710 715 720

40 Asn Lys Ile Ile Leu Lys Asn Asn Leu Ser Ala Leu Ile Met Gly Asn
 725 730 735

Thr Asp Asp Ile Leu Lys Asn Leu Glu Asn Glu Phe Phe Asn Leu Lys
 740 745 750

45 Glu Ser Leu Glu Glu Ser Asn His Tyr Asn Gly Leu Leu Asn Leu Asp
 755 760 765

50 Ala Asn Ser Lys Ala Leu Arg Glu Ile Ile Ile Ile Gln Ser Lys Val
 770 775 780

55 Ala Phe Asn Ala Ile Cys Phe Pro Ser Tyr Lys Ile Asn Asp Glu Asn
 785 790 795 800

Tyr Pro Lys Ala Asn Phe Leu Glu His Val Leu Arg Ser Gly Ile Phe

25/67

805

810

815

5 Trp Glu Lys Ile Arg Val Met Gly Gly Ala Tyr Gly Ala Ser Ala Ser
820 825 830

10 Ile Ala Asn Gly Ile Phe Ser Phe Ala Ser Tyr Arg Asp Pro Asn Phe
835 840 845

15 Thr Lys Thr Tyr Gln Ala Phe Glu Lys Ser Leu Glu Glu Leu Ala Asn
850 855 860

Asn Lys Met Thr Asp Asp Glu Ile Tyr Thr Tyr Leu Ile Gly Leu Ile
865 870 875 880

20 Gly Thr Asn Ile Tyr Val Lys Thr Lys Ala Thr Glu Ala Leu Gln Ser
885 890 895

25 Tyr Arg Arg Lys Met Leu Asn Ile Ser Asp Ser Leu Arg Gln Asp Ile
900 905 910

30 Arg Asn Ala Tyr Phe Thr Ile Thr Pro Gln Asp Ile Lys Glu Ile Ser
915 920 925

Thr Lys Ile Leu Thr Gln Ile Arg Gln His Asn Ser Ile Ala Ser Leu
930 935 940

35 Val Asn Asn Gln Ile Tyr Glu Glu Glu Lys Asn Asn Leu Glu Lys Leu
945 950 955 960

40 Ile Gly Lys Glu Tyr Ser Leu Lys Lys Ile Tyr
965 970

<210> 8

<211> 995

<212> PRT

45 <213> Caenorhabditis elegans

<400> 8

50 Met Ser Ala Ser Lys Leu Trp Ser Cys Thr Glu Thr Val Leu Asn Gly
1 5 10 15

Gly Ile Lys Leu Phe Leu Tyr Ser Ser Lys Asn Thr Lys Leu Arg Val
20 25 30

55

Ala Ile Gly Glu Val Pro Gly Pro Met Val His Gly Ala Val Ser Phe
35 40 45

26/67

Val Thr Glu Ala Asp Ser Asp Asp Gly Leu Pro His Thr Leu Glu His
 50 55 60

5

Leu Val Phe Met Gly Ser Lys Lys Tyr Pro Phe Lys Gly Val Leu Asp
 65 70 75 80

10

Val Ile Ala Asn Arg Cys Leu Ala Asp Gly Thr Asn Ala Trp Thr Asp
 85 90 95

15

Thr Asp His Thr Ala Tyr Thr Leu Ser Thr Val Gly Ser Asp Gly Phe
 100 105 110

20

Leu Lys Val Leu Pro Val Tyr Ile Asn His Leu Leu Thr Pro Met Leu
 115 120 125

25

Thr Ala Ser Gln Phe Ala Thr Glu Val His His Ile Thr Gly Glu Gly
 130 135 140

30

Asn Asp Ala Gly Val Val Tyr Ser Glu Met Gln Asp His Glu Ser Glu
 145 150 155 160

35

Met Glu Ser Ile Met Asp Arg Lys Thr Lys Glu Val Ile Tyr Pro Pro
 165 170 175

40

Phe Asn Pro Tyr Ala Val Asp Thr Gly Gly Arg Leu Lys Asn Leu Arg
 180 185 190

45

Glu Ser Cys Thr Leu Glu Lys Val Arg Asp Tyr His Lys Lys Phe Tyr
 195 200 205

50

His Leu Ser Asn Met Val Val Thr Val Cys Gly Met Val Asp His Asp
 210 215 220

55

Gln Val Leu Glu Ile Met Asn Asn Val Glu Asn Glu His Met Ser Thr
 225 230 235 240

Val Pro Asp His Phe Pro Lys Pro Phe Ser Phe Ala Leu Ser Asp Ile
 245 250 255

Lys Glu Ser Thr Val His Arg Val Glu Cys Pro Thr Asp Asp Ala Ser
 260 265 270

Arg Gly Ala Val Glu Val Ala Trp Phe Ala His Ser Pro Ser Asp Leu
 275 280 285

27/67

Glu Thr His Ser Ser Leu His Val Leu Phe Asp Tyr Leu Ser Asn Thr
 290 295 300

5

Ser Val Ala Pro Leu Gln Lys Asp Phe Ile Leu Leu Glu Asp Pro Leu
 305 310 315 320

10

Ala Ser Ser Val Ser Phe His Ile Ala Glu Gly Val Arg Cys Asp Leu
 325 330 335

15

Arg Leu Asn Phe Ala Gly Val Pro Val Glu Lys Leu Asp Glu Cys Ala
 340 345 350

20

Pro Lys Phe Phe Asp Lys Thr Val Arg Glu His Leu Glu Glu Ala Asn
 355 360 365

25

Phe Asp Met Glu Arg Met Gly Tyr Leu Ile Asp Gln Thr Ile Leu Asn
 370 375 380

30

Glu Leu Val Lys Leu Glu Thr Asn Ala Pro Lys Asp Ile Met Ser His
 385 390 395 400

35

Ile Ile Gly His Gln Leu Phe Asp Asn Glu Asp Glu Glu Leu Phe Lys
 405 410 415

40

Lys Arg Thr Asn Glu Ile Asp Phe Leu Lys Lys Leu Lys Ser Glu Pro
 420 425 430

45

Ala Ser Tyr Trp Val Gln Leu Val Asn Lys Tyr Phe Thr Ala Pro Ser
 435 440 445

Ala Thr Val Ile Gly Val Pro Asn Glu Glu Leu Val Asp Lys Ile Ala
 450 455 460

50

Glu Glu Glu Glu Lys Arg Ile Ala Ala Gln Cys Glu Lys Leu Gly Lys
 465 470 475 480

Lys Gly Leu Glu Glu Ala Gly Lys Ser Leu Glu Ala Ala Ile Leu Glu
 485 490 495

55

Asn Thr Ala Asn His Pro Ser Ala Glu Leu Leu Asp Gln Leu Ile Val
 500 505 510

Lys Asp Leu Glu Ala Phe Asp Arg Phe Pro Val Gln Ser Leu Thr Ser
 515 520 525

28/67

	Asn	Ser	Pro	Ser	Leu	Thr	Pro	Gln	Gln	Ser	Thr	Phe	Leu	Ala	Gln	Phe	
	530						535					540					
5	Pro	Phe	His	Ala	Asn	Leu	His	Asn	Cys	Pro	Thr	Lys	Phe	Val	Glu	Ile	
	545					550					555					560	
10	Phe	Phe	Leu	Leu	Asp	Ser	Ser	Asn	Leu	Ser	Ile	Glu	Asp	Arg	Ser	Tyr	
					565					570					575		
15	Leu	Phe	Leu	Tyr	Thr	Asp	Leu	Leu	Phe	Glu	Ser	Pro	Ala	Met	Ile	Asp	
				580					585					590			
20	Gly	Val	Leu	Thr	Ser	Ala	Asp	Asp	Val	Ala	Lys	His	Phe	Thr	Lys	Asp	
			595					600					605				
25	Leu	Ile	Asp	His	Ser	Ile	Gln	Val	Gly	Val	Ser	Gly	Leu	Tyr	Asp	Arg	
	610						615					620					
30	Phe	Val	Asn	Leu	Arg	Ile	Lys	Val	Gly	Ala	Asp	Lys	Tyr	Pro	Leu	Leu	
	625					630					635					640	
35	Ala	Lys	Trp	Ala	Gln	Ile	Phe	Thr	Gln	Gly	Val	Val	Phe	Asp	Pro	Ser	
					645					650					655		
40	Arg	Ile	His	Gln	Cys	Ala	Gln	Lys	Leu	Ala	Gly	Glu	Ala	Arg	Asp	Arg	
				660					665					670			
45	Lys	Arg	Asp	Gly	Cys	Thr	Val	Ala	Ser	Thr	Ala	Val	Ala	Ser	Met	Val	
			675					680					685				
50	Tyr	Gly	Lys	Asn	Thr	Asn	Cys	Ile	Leu	Phe	Asp	Glu	Leu	Val	Leu	Glu	
	690						695					700					
55	Lys	Leu	His	Glu	Lys	Ile	Ser	Lys	Asp	Val	Met	Lys	Asn	Pro	Glu	Ala	
	705					710					715				720		
60	Val	Leu	Glu	Lys	Leu	Glu	Gln	Val	Arg	Ser	Ala	Leu	Phe	Ser	Asn	Gly	
					725					730					735		
65	Val	Asn	Ala	His	Phe	Val	Ala	Asp	Val	Asp	Ser	Ile	Asp	Pro	Lys	Met	
				740					745					750			
70	Leu	Ser	Ser	Asp	Leu	Trp	Thr	Trp	Val	Gln	Ala	Asp	Pro	Arg	Phe	Gly	
			755					760					765				
75	Pro	Gly	His	Gln	Phe	Ser	Ala	Glu	Ala	Gly	Glu	Asn	Val	Ser	Leu	Glu	

29/67

	770		775		780
5	Leu Gly Lys Glu Leu Leu Ile Gly Val Gly Gly Ser Glu Ser Ser Phe				
	785		790		795 800
10	Ile Tyr Gln Thr Ser Phe Leu Asp Ala Asn Trp Asn Ser Glu Glu Leu				
			805		810 815
15	Ile Pro Ala Met Ile Phe Gly Gln Tyr Leu Ser Gln Cys Glu Gly Pro				
			820		825 830
20	Leu Trp Arg Ala Ile Arg Gly Asp Gly Leu Ala Tyr Gly Ala Asn Val				
			835		840 845
25	Phe Val Lys Pro Asp Arg Lys Gln Ile Thr Leu Ser Leu Tyr Arg Cys				
			850		855 860
30	Ala Gln Pro Ala Val Ala Tyr Glu Arg Thr Arg Asp Ile Ile Arg Lys				
			865		870 875 880
35	Ile Val Glu Ser Gly Glu Ile Ser Lys Ala Glu Phe Glu Gly Ala Lys				
			885		890 895
40	Arg Ser Thr Val Phe Glu Met Met Lys Arg Glu Gly Thr Val Ser Gly				
			900		905 910
45	Ala Ala Lys Ile Ser Ile Leu Asn Asn Phe Arg Gln Thr Pro His Pro				
			915		920 925
50	Phe Asn Ile Asp Leu Cys Arg Arg Ile Trp Asn Leu Thr Ser Glu Glu				
			930		935 940
55	Met Val Lys Ile Gly Gly Pro Pro Leu Ala Arg Leu Phe Asp Glu Lys				
			945		950 955 960
60	Cys Phe Val Arg Ser Ile Ala Val His Pro Ser Lys Leu Asn Glu Met				
			965		970 975
65	Lys Lys Ala Phe Pro Gly Ser Ser Lys Ile Lys Ile Ser Asp Leu Gln				
			980		985 990
70	Phe Ala Cys				
			995		

<210> 9

<211> 962

30/67

<212> PRT

<213> Escherichia coli

<400> 9

5

Met Pro Arg Ser Thr Trp Phe Lys Ala Leu Leu Leu Leu Val Ala Leu
 1 5 10 15

10 Trp Ala Pro Leu Ser Gln Ala Glu Thr Gly Trp Gln Pro Ile Gln Glu
 20 25 30

15 Thr Ile Arg Lys Ser Asp Lys Asp Asn Arg Gln Tyr Gln Ala Ile Arg
 35 40 45

20 Leu Asp Asn Gly Met Val Val Leu Leu Val Ser Asp Pro Gln Ala Val
 50 55 60

Lys Ser Leu Ser Ala Leu Val Val Pro Val Gly Ser Leu Glu Asp Pro
 65 70 75 80

25

Glu Ala Tyr Gln Gly Leu Ala His Tyr Leu Glu His Met Ser Leu Met
 85 90 95

30 Gly Ser Lys Lys Tyr Pro Gln Ala Asp Ser Leu Ala Glu Tyr Leu Lys
 100 105 110

35 Met His Gly Gly Ser His Asn Ala Ser Thr Ala Pro Tyr Arg Thr Ala
 115 120 125

Phe Tyr Leu Glu Val Glu Asn Asp Ala Leu Pro Gly Ala Val Asp Arg
 130 135 140

40

Leu Ala Asp Ala Ile Ala Glu Pro Leu Leu Asp Lys Lys Tyr Ala Glu
 145 150 155 160

45 Arg Glu Arg Asn Ala Val Asn Ala Glu Leu Thr Met Ala Arg Thr Arg
 165 170 175

50 Asp Gly Met Arg Met Ala Gln Val Ser Ala Glu Thr Ile Asn Pro Ala
 180 185 190

55 His Pro Gly Ser Lys Phe Ser Gly Gly Asn Leu Glu Thr Leu Ser Asp
 195 200 205

Lys Pro Gly Asn Pro Val Gln Gln Ala Leu Lys Asp Phe His Glu Lys
 210 215 220

31/67

Tyr Tyr Ser Ala Asn Leu Met Lys Ala Val Ile Tyr Ser Asn Lys Pro
 225 230 235 240

5

Leu Pro Glu Leu Ala Lys Met Ala Ala Asp Thr Phe Gly Arg Val Pro
 245 250 255

10

Asn Lys Glu Ser Lys Lys Pro Glu Ile Thr Val Pro Val Val Thr Asp
 260 265 270

15

Ala Gln Lys Gly Ile Ile Ile His Tyr Val Pro Ala Leu Pro Arg Lys
 275 280 285

20

Val Leu Arg Val Glu Phe Arg Ile Asp Asn Asn Ser Ala Lys Phe Arg
 290 295 300

Ser Lys Thr Asp Glu Leu Ile Thr Tyr Leu Ile Gly Asn Arg Ser Pro
 305 310 315 320

25

Gly Thr Leu Ser Asp Trp Leu Gln Lys Gln Gly Leu Val Glu Gly Ile
 325 330 335

30

Ser Ala Asn Ser Asp Pro Ile Val Asn Gly Asn Ser Gly Val Leu Ala
 340 345 350

35

Ile Ser Ala Ser Leu Thr Asp Lys Gly Leu Ala Asn Arg Asp Gln Val
 355 360 365
 Val Ala Ala Ile Phe Ser Tyr Leu Asn Leu Leu Arg Glu Lys Gly Ile
 370 375 380

40

Asp Lys Gln Tyr Phe Asp Glu Leu Ala Asn Val Leu Asp Ile Asp Phe
 385 390 395 400

45

Arg Tyr Pro Ser Ile Thr Arg Asp Met Asp Tyr Val Glu Trp Leu Ala
 405 410 415

50

Asp Thr Met Ile Arg Val Pro Val Glu His Thr Leu Asp Ala Val Asn
 420 425 430

Ile Ala Asp Arg Tyr Asp Ala Lys Ala Val Lys Glu Arg Leu Ala Met
 435 440 445

55

Met Thr Pro Gln Asn Ala Arg Ile Trp Tyr Ile Ser Pro Lys Glu Pro
 450 455 460

32/67

	His	Asn	Lys	Thr	Ala	Tyr	Phe	Val	Asp	Ala	Pro	Tyr	Gln	Val	Asp	Lys	
	465					470					475					480	
5	Ile	Ser	Ala	Gln	Thr	Phe	Ala	Asp	Trp	Gln	Lys	Lys	Ala	Ala	Asp	Ile	
					485					490					495		
10	Ala	Leu	Ser	Leu	Pro	Glu	Leu	Asn	Pro	Tyr	Ile	Pro	Asp	Asp	Phe	Ser	
				500					505					510			
15	Leu	Ile	Lys	Ser	Glu	Lys	Lys	Tyr	Asp	His	Pro	Glu	Leu	Ile	Val	Asp	
			515					520					525				
20	Glu	Ser	Asn	Leu	Arg	Val	Val	Tyr	Ala	Pro	Ser	Arg	Tyr	Phe	Ala	Ser	
		530					535					540					
25	Glu	Pro	Lys	Ala	Asp	Val	Ser	Leu	Ile	Leu	Arg	Asn	Pro	Lys	Ala	Met	
	545					550					555					560	
30	Asp	Ser	Ala	Arg	Asn	Gln	Val	Met	Phe	Ala	Leu	Asn	Asp	Tyr	Leu	Ala	
					565					570					575		
35	Gly	Leu	Ala	Leu	Asp	Gln	Leu	Ser	Asn	Gln	Ala	Ser	Val	Gly	Gly	Ile	
				580					585					590			
40	Ser	Phe	Ser	Thr	Asn	Ala	Asn	Asn	Gly	Leu	Met	Val	Asn	Ala	Asn	Gly	
			595				600					605					
45	Tyr	Thr	Gln	Arg	Leu	Pro	Gln	Leu	Phe	Gln	Ala	Leu	Leu	Glu	Gly	Tyr	
		610				615					620						
50	Phe	Ser	Tyr	Thr	Ala	Thr	Glu	Asp	Gln	Leu	Glu	Gln	Ala	Lys	Ser	Trp	
	625					630					635					640	
55	Tyr	Asn	Gln	Met	Met	Asp	Ser	Ala	Glu	Lys	Gly	Lys	Ala	Phe	Glu	Gln	
				645						650					655		
60	Ala	Ile	Met	Pro	Ala	Gln	Met	Leu	Ser	Gln	Val	Pro	Tyr	Phe	Ser	Arg	
				660					665					670			
65	Asp	Glu	Arg	Arg	Lys	Ile	Leu	Pro	Ser	Ile	Thr	Leu	Lys	Glu	Val	Leu	
			675					680					685				
70	Ala	Tyr	Arg	Asp	Ala	Leu	Lys	Ser	Gly	Ala	Arg	Pro	Glu	Phe	Met	Val	
		690					695					700					
75	Ile	Gly	Asn	Met	Thr	Glu	Ala	Gln	Ala	Thr	Thr	Leu	Ala	Arg	Asp	Val	

33/67

	705		710		715		720									
5	Gln	Lys	Gln	Leu	Gly	Ala	Asp	Gly	Ser	Glu	Trp	Cys	Arg	Asn	Lys	Asp
				725						730						735
10	Val	Val	Val	Asp	Lys	Lys	Gln	Ser	Val	Ile	Phe	Glu	Lys	Ala	Gly	Asn
				740					745						750	
15	Ser	Thr	Asp	Ser	Ala	Leu	Ala	Ala	Val	Phe	Val	Pro	Thr	Gly	Tyr	Asp
			755					760						765		
20	Glu	Tyr	Thr	Ser	Ser	Ala	Tyr	Ser	Ser	Leu	Leu	Gly	Gln	Ile	Val	Gln
		770					775					780				
25	Val	Phe	Ala	Phe	Pro	Met	Ser	Val	Gly	Arg	Gln	Trp	Gly	Met	Gly	Phe
					805					810					815	
30	Leu	Leu	Gln	Ser	Asn	Asp	Lys	Gln	Pro	Ser	Phe	Leu	Trp	Glu	Arg	Tyr
				820					825					830		
35	Lys	Ala	Phe	Phe	Pro	Thr	Ala	Glu	Ala	Lys	Leu	Arg	Ala	Met	Lys	Pro
			835					840					845			
	Asp	Glu	Phe	Ala	Gln	Ile	Gln	Gln	Ala	Val	Ile	Thr	Gln	Met	Leu	Gln
		850				855						860				
40	Ala	Pro	Gln	Thr	Leu	Gly	Glu	Glu	Ala	Ser	Lys	Leu	Ser	Lys	Asp	Phe
	865					870					875				880	
45	Asp	Arg	Gly	Asn	Met	Arg	Phe	Asp	Ser	Arg	Asp	Lys	Ile	Val	Ala	Gln
				885						890					895	
50	Ile	Lys	Leu	Leu	Thr	Pro	Gln	Lys	Leu	Ala	Asp	Phe	Phe	His	Gln	Ala
			900						905					910		
55	Gln	Asn	Gly	Lys	Ala	Glu	Tyr	Val	His	Pro	Glu	Gly	Trp	Lys	Val	Trp
		930					935					940				
	Glu	Asn	Val	Ser	Ala	Leu	Gln	Gln	Thr	Met	Pro	Leu	Met	Ser	Glu	Lys
	945					950					955				960	

34/67

Asn Glu

5

<210> 10

<211> 1161

<212> PRT

10 <213> Homo sapiens

<400> 10

Met Leu Arg Arg Val Ala Val Ala Ala Val Phe Ala Thr Gly Arg Lys
 15 1 5 10 15

Leu Arg Cys Glu Ala Gly Arg Asp Val Thr Ala Val Gly Arg Ile Glu
 20 20 25 30

Ala Arg Gly Leu Cys Glu Glu Ser Ala Lys Pro Phe Pro Thr Leu Thr
 35 40 45

25

Met Pro Gly Arg Asn Lys Ala Lys Ser Thr Cys Ser Cys Pro Asp Leu
 50 55 60

Gln Pro Asn Gly Gln Asp Leu Gly Glu Ser Gly Arg Val Ala Arg Leu
 65 70 75 80
 Gly Ala Asp Glu Ser Glu Glu Glu Gly Arg Ser Leu Ser Asn Val Gly
 85 90 95

35

Asp Pro Glu Ile Ile Lys Ser Pro Ser Asp Pro Lys Gln Tyr Arg Tyr
 100 105 110

Ile Lys Leu Gln Asn Gly Leu Gln Ala Leu Leu Ile Ser Asp Leu Ser
 115 120 125

Asn Val Glu Gly Lys Thr Gly Asn Ala Thr Asp Glu Glu Glu Glu Glu
 45 130 135 140

Glu Glu Glu Glu Glu Glu Gly Glu Glu Glu Glu Glu Glu Glu Asp
 145 150 155 160

50

Asp Asp Asp Asp Asp Asp Glu Asp Ser Gly Ala Glu Ile Gln Asp Asp
 165 170 175

55

Asp Glu Glu Gly Phe Asp Asp Glu Glu Glu Phe Asp Asp Asp Glu His
 180 185 190

35/67

Asp Asp Asp Asp Leu Asp Asn Glu Glu Asn Glu Leu Glu Glu Leu Glu
 195 200 205

5 Glu Arg Val Glu Ala Arg Lys Lys Thr Thr Glu Lys Gln Ser Ala Ala
 210 215 220

10 Ala Leu Cys Val Gly Val Gly Ser Phe Ala Asp Pro Asp Asp Leu Pro
 225 230 235 240

15 Gly Leu Ala His Phe Leu Glu His Met Val Phe Met Gly Ser Leu Lys
 245 250 255

Tyr Pro Asp Glu Asn Gly Phe Asp Ala Phe Leu Lys Lys His Gly Gly
 260 265 270

20 Ser Asp Asn Ala Ser Thr Asp Cys Glu Arg Thr Val Phe Gln Phe Asp
 275 280 285

25 Val Gln Arg Lys Tyr Phe Lys Glu Ala Leu Asp Arg Trp Ala Gln Phe
 290 295 300

30 Phe Ile His Pro Leu Met Ile Arg Asp Ala Ile Asp Arg Glu Val Glu
 305 310 315 320
 Ala Val Asp Ser Glu Tyr Gln Leu Ala Arg Pro Ser Asp Ala Asn Arg
 325 330 335

35 Lys Glu Met Leu Phe Gly Ser Leu Ala Arg Pro Gly His Pro Met Gly
 340 345 350

40 Lys Phe Phe Trp Gly Asn Ala Glu Thr Leu Lys His Glu Pro Lys Lys
 355 360 365

45 Asn Asn Ile Asp Thr His Ala Arg Leu Arg Glu Phe Trp Met Arg Tyr
 370 375 380

Tyr Ser Ala His Tyr Met Thr Leu Val Val Gln Ser Lys Glu Thr Leu
 385 390 395 400

50 Asp Thr Leu Glu Lys Trp Val Thr Glu Ile Phe Ser Gln Ile Pro Asn
 405 410 415

55 Asn Gly Leu Pro Lys Pro Asn Phe Ser His Leu Thr Asp Pro Phe Asp
 420 425 430

Thr Pro Ala Phe Asn Lys Leu Tyr Arg Val Val Pro Ile Arg Lys Ile

36/67

435

440

445

5 His Ala Leu Thr Ile Thr Trp Ala Leu Pro Pro Gln Gln Gln His Tyr
 450 455 460

10 Arg Val Lys Pro Leu His Tyr Ile Ser Trp Leu Val Gly His Glu Gly
 465 470 475 480

15 Lys Gly Ser Ile Leu Ser Tyr Leu Arg Lys Lys Cys Trp Ala Leu Ala
 485 490 495

20 Leu Phe Gly Gly Asn Gly Glu Thr Gly Phe Glu Gln Asn Ser Thr Tyr
 500 505 510

25 Ser Val Phe Ser Ile Ser Ile Thr Leu Thr Asp Glu Gly Tyr Glu His
 515 520 525

30 Phe Tyr Glu Val Ala His Thr Val Phe Gln Tyr Leu Lys Met Leu Gln
 530 535 540

35 Lys Leu Gly Pro Glu Lys Arg Val Phe Glu Glu Ile Gln Lys Ile Glu
 545 550 555 560

40 Asp Asn Glu Phe His Tyr Gln Glu Gln Thr Asp Pro Val Glu Tyr Val
 565 570 575

45 Glu Asn Met Cys Glu Asn Met Gln Leu Tyr Pro Arg Gln Asp Phe Leu
 580 585 590

50 Thr Gly Asp Gln Leu Leu Phe Glu Tyr Lys Pro Glu Val Ile Ala Glu
 595 600 605

55 Ala Leu Asn Gln Leu Val Pro Gln Lys Ala Asn Leu Val Leu Leu Ser
 610 615 620

60 Gly Ala Asn Glu Gly Arg Cys Asp Leu Lys Glu Lys Trp Phe Gly Thr
 625 630 635 640

65 Gln Tyr Ser Ile Glu Asp Ile Glu Asn Ser Trp Thr Glu Leu Trp Lys
 645 650 655

70 Ser Asn Phe Asp Leu Asn Ser Asp Leu His Leu Pro Ala Glu Asn Lys
 660 665 670

75 Tyr Ile Ala Thr Asp Phe Thr Leu Lys Ala Phe Asp Cys Pro Glu Thr
 675 680 685

5	Glu 690	Tyr	Pro	Ala	Lys	Ile	Val 695	Asn	Thr	Pro	Gln	Gly 700	Cys	Leu	Trp	Tyr
10	Lys 705	Lys	Asp	Asn	Lys	Phe 710	Lys	Ile	Pro	Lys	Ala 715	Tyr	Ile	Arg	Phe	His 720
15	Leu	Ile	Ser	Pro	Leu 725	Ile	Gln	Lys	Ser	Ala 730	Ala	Asn	Val	Val	Leu	Phe 735
20	Asp	Ile	Phe	Val 740	Asn	Ile	Leu	Thr	His 745	Asn	Leu	Ala	Glu	Pro 750	Ala	Tyr
25	Glu	Ala	Asp 755	Val	Ala	Gln	Leu	Glu 760	Tyr	Lys	Leu	Val	Ala 765	Gly	Glu	His
30	Gly 770	Leu	Ile	Ile	Arg	Val	Lys 775	Gly	Phe	Asn	His	Lys 780	Leu	Pro	Leu	Leu
35	Phe 785	Gln	Leu	Ile	Ile	Asp 790	Tyr	Leu	Thr	Glu	Phe 795	Ser	Ser	Thr	Pro	Ala 800
40	Val	Phe	Thr	Met	Ile	Thr	Glu	Gln	Leu	Lys 810	Lys	Thr	Tyr	Phe	Asn	Ile 815
45	Leu	Ile	Lys	Pro 820	Glu	Thr	Leu	Ala	Lys 825	Asp	Val	Arg	Leu	Leu	Ile	Leu
50	Glu	Tyr	Ser 835	Arg	Trp	Ser	Met	Ile	Asp	Lys	Tyr	Arg	Ala 845	Leu	Met	Asp
55	Gly 850	Leu	Ser	Leu	Glu	Ser	Leu 855	Leu	Asn	Phe	Val	Lys 860	Asp	Phe	Lys	Ser
60	Gln 865	Leu	Phe	Val	Glu	Gly 870	Leu	Val	Gln	Gly	Asn 875	Val	Thr	Ser	Thr	Glu 880
65	Ser	Met	Asp	Phe	Leu 885	Arg	Tyr	Val	Val	Asp 890	Lys	Leu	Asn	Phe	Val	Pro 895
70	Leu	Glu	Arg	Glu 900	Met	Pro	Val	Gln	Phe 905	Gln	Val	Val	Glu	Leu	Pro	Ser
75	Gly	His	His 915	Leu	Cys	Lys	Val	Arg 920	Ala	Leu	Asn	Lys	Gly 925	Asp	Ala	Asn

38/67

Ser Glu Val Thr Val Tyr Tyr Gln Ser Gly Thr Arg Ser Leu Arg Glu
 930 935 940

5

Tyr Thr Leu Met Glu Leu Leu Val Met His Met Glu Glu Pro Cys Phe
 945 950 955 960

10 Asp Phe Leu Arg Thr Lys Gln Thr Leu Gly Tyr His Val Tyr Pro Thr
 965 970 975

15 Cys Arg Asn Thr Ser Gly Ile Leu Gly Phe Ser Val Thr Val Gly Thr
 980 985 990

20 Gln Ala Thr Lys Tyr Asn Ser Glu Thr Val Asp Lys Lys Ile Glu Glu
 995 1000 1005

Phe Leu Ser Ser Phe Glu Glu Lys Ile Glu Asn Leu Thr Glu Asp
 1010 1015 1020

25

Ala Phe Asn Thr Gln Val Thr Ala Leu Ile Lys Leu Lys Glu Cys
 1025 1030 1035

Glu Asp Thr His Leu Gly Glu Glu Val Asp Arg Asn Trp Asn Glu
 1040 1045 1050

30

Val Val Thr Gln Gln Tyr Leu Phe Asp Arg Leu Ala His Glu Ile
 1055 1060 1065

35

Glu Ala Leu Lys Ser Phe Ser Lys Ser Asp Leu Val Ser Trp Phe
 1070 1075 1080

40 Lys Ala His Arg Gly Pro Gly Ser Lys Met Leu Ser Val His Val
 1085 1090 1095

45 Val Gly Tyr Gly Lys Tyr Glu Leu Glu Glu Asp Gly Ala Pro Val
 1100 1105 1110

50 Cys Glu Asp Pro Asn Ser Arg Glu Gly Met Gln Leu Ile Tyr Leu
 1115 1120 1125

Pro Pro Ser Pro Leu Leu Ala Glu Ser Thr Thr Pro Ile Thr Asp
 1130 1135 1140

55

Ile Arg Ala Phe Thr Ala Thr Leu Ser Leu Phe Pro Tyr His Lys
 1145 1150 1155

39/67

Ile Val Lys
1160

5 <210> 11
<211> 1019
<212> PRT
<213> Homo sapiens

10 <400> 11

Met Arg Tyr Arg Leu Ala Trp Leu Leu His Pro Ala Leu Pro Ser Thr
1 5 10 15

15

Phe Arg Ser Val Leu Gly Ala Arg Leu Pro Pro Pro Glu Arg Leu Cys
20 25 30

20 Gly Phe Gln Lys Lys Thr Tyr Ser Lys Met Asn Asn Pro Ala Ile Lys
35 40 45

Arg Ile Gly Asn His Ile Thr Lys Ser Pro Glu Asp Lys Arg Glu Tyr
25 50 55 60
Arg Gly Leu Glu Leu Ala Asn Gly Ile Lys Val Leu Leu Met Ser Asp
65 70 75 80

30 Pro Thr Thr Asp Lys Ser Ser Ala Ala Leu Asp Val His Ile Gly Ser
85 90 95

Leu Ser Asp Pro Pro Asn Ile Ala Gly Leu Ser His Phe Cys Glu His
35 100 105 110

Met Leu Phe Leu Gly Thr Lys Lys Tyr Pro Lys Glu Asn Glu Tyr Ser
115 120 125
40

Gln Phe Leu Ser Glu His Ala Gly Ser Ser Asn Ala Phe Thr Ser Gly
130 135 140

45

Glu His Thr Asn Tyr Tyr Phe Asp Val Ser His Glu His Leu Glu Gly
145 150 155 160

50 Ala Leu Asp Arg Phe Ala Gln Phe Phe Leu Cys Pro Leu Phe Asp Glu
165 170 175

Ser Cys Lys Asp Arg Glu Val Asn Ala Val Asp Ser Glu His Glu Lys
55 180 185 190

Asn Val Met Asn Asp Ala Trp Arg Leu Phe Gln Leu Glu Lys Ala Thr
195 200 205

40/67

Gly Asn Pro Lys His Pro Phe Ser Lys Phe Gly Thr Gly Asn Lys Tyr
 210 215 220
 5

Thr Leu Glu Thr Arg Pro Asn Gln Glu Gly Ile Asp Val Arg Gln Glu
 225 230 235 240

10

Leu Leu Lys Phe His Ser Ala Tyr Tyr Ser Ser Asn Leu Met Ala Val
 245 250 255

15 Cys Val Leu Gly Arg Glu Ser Leu Asp Asp Leu Thr Asn Leu Val Val
 260 265 270

20 Lys Leu Phe Ser Glu Val Glu Asn Lys Asn Val Pro Leu Pro Glu Phe
 275 280 285

Pro Glu His Pro Phe Gln Glu Glu His Leu Lys Gln Leu Tyr Lys Ile
 290 295 300

25 Val Pro Ile Lys Asp Ile Arg Asn Leu Tyr Val Thr Phe Pro Ile Pro
 305 310 315 320

30 Asp Leu Gln Lys Tyr Tyr Lys Ser Asn Pro Gly His Tyr Leu Gly His
 325 330 335

Leu Ile Gly His Glu Gly Pro Gly Ser Leu Leu Ser Glu Leu Lys Ser
 340 345 350

35

Lys Gly Trp Val Asn Thr Leu Val Gly Gly Gln Lys Glu Gly Ala Arg
 355 360 365

40

Gly Phe Met Phe Phe Ile Ile Asn Val Asp Leu Thr Glu Glu Gly Leu
 370 375 380

45 Leu His Val Glu Asp Ile Ile Leu His Met Phe Gln Tyr Ile Gln Lys
 385 390 395 400

Leu Arg Ala Glu Gly Pro Gln Glu Trp Val Phe Gln Glu Cys Lys Asp
 405 410 415

50

Leu Asn Ala Val Ala Phe Arg Phe Lys Asp Lys Glu Arg Pro Arg Gly
 420 425 430

55

Tyr Thr Ser Lys Ile Ala Gly Ile Leu His Tyr Tyr Pro Leu Glu Glu
 435 440 445

41/67

Val Leu Thr Ala Glu Tyr Leu Leu Glu Glu Phe Arg Pro Asp Leu Ile
 450 455 460

5

Glu Met Val Leu Asp Lys Leu Arg Pro Glu Asn Val Arg Val Ala Ile
 465 470 475 480

10 Val Ser Lys Ser Phe Glu Gly Lys Thr Asp Arg Thr Glu Glu Trp Tyr
 485 490 495

15 Gly Thr Gln Tyr Lys Gln Glu Ala Ile Pro Asp Glu Val Ile Lys Lys
 500 505 510

20 Trp Gln Asn Ala Asp Leu Asn Gly Lys Phe Lys Leu Pro Thr Lys Asn
 515 520 525

Glu Phe Ile Pro Thr Asn Phe Glu Ile Leu Pro Leu Glu Lys Glu Ala
 530 535 540
 Thr Pro Tyr Pro Ala Leu Ile Lys Asp Thr Val Met Ser Lys Leu Trp
 25 545 550 555 560

30 Phe Lys Gln Asp Asp Lys Lys Lys Lys Pro Lys Ala Cys Leu Asn Phe
 565 570 575

Glu Phe Phe Ser Pro Phe Ala Tyr Val Asp Pro Leu His Cys Asn Met
 580 585 590

35

Ala Tyr Leu Tyr Leu Glu Leu Leu Lys Asp Ser Leu Asn Glu Tyr Ala
 595 600 605

40 Tyr Ala Ala Glu Leu Ala Gly Leu Ser Tyr Asp Leu Gln Asn Thr Ile
 610 615 620

45 Tyr Gly Met Tyr Leu Ser Val Lys Gly Tyr Asn Asp Lys Gln Pro Ile
 625 630 635 640

50 Leu Leu Lys Lys Ile Ile Glu Lys Met Ala Thr Phe Glu Ile Asp Glu
 645 650 655

Lys Arg Phe Glu Ile Ile Lys Glu Ala Tyr Met Arg Ser Leu Asn Asn
 660 665 670

55

Phe Arg Ala Glu Gln Pro His Gln His Ala Met Tyr Tyr Leu Arg Leu
 675 680 685

42/67

	Leu	Met	Thr	Glu	Val	Ala	Trp	Thr	Lys	Asp	Glu	Leu	Lys	Glu	Ala	Leu	
	690						695					700					
5	Asp	Asp	Val	Thr	Leu	Pro	Arg	Leu	Lys	Ala	Phe	Ile	Pro	Gln	Leu	Leu	
	705					710					715					720	
10	Ser	Arg	Leu	His	Ile	Glu	Ala	Leu	Leu	His	Gly	Asn	Ile	Thr	Lys	Gln	
					725					730					735		
15	Ala	Ala	Leu	Gly	Ile	Met	Gln	Met	Val	Glu	Asp	Thr	Leu	Ile	Glu	His	
				740					745					750			
20	Ala	His	Thr	Lys	Pro	Leu	Leu	Pro	Ser	Gln	Leu	Val	Arg	Tyr	Arg	Glu	
			755					760					765				
25	Val	Gln	Leu	Pro	Asp	Arg	Gly	Trp	Phe	Val	Tyr	Gln	Gln	Arg	Asn	Glu	
	770						775					780					
	Val	His	Asn	Asn	Cys	Gly	Ile	Glu	Ile	Tyr	Tyr	Gln	Thr	Asp	Met	Gln	
	785				790					795						800	
30	Ser	Thr	Ser	Glu	Asn	Met	Phe	Leu	Glu	Leu	Phe	Cys	Gln	Ile	Ile	Ser	
					805					810					815		
35	Glu	Pro	Cys	Phe	Asn	Thr	Leu	Arg	Thr	Lys	Glu	Gln	Leu	Gly	Tyr	Ile	
				820					825					830			
40	Val	Phe	Ser	Gly	Pro	Arg	Arg	Ala	Asn	Gly	Ile	Gln	Ser	Leu	Arg	Phe	
			835					840					845				
45	Ile	Ile	Gln	Ser	Glu	Lys	Pro	Pro	His	Tyr	Leu	Glu	Ser	Arg	Val	Glu	
	850						855					860					
50	Ala	Phe	Leu	Ile	Thr	Met	Glu	Lys	Ser	Ile	Glu	Asp	Met	Thr	Glu	Glu	
	865					870					875					880	
55	Ala	Phe	Gln	Lys	His	Ile	Gln	Ala	Leu	Ala	Ile	Arg	Arg	Leu	Asp	Lys	
					885					890					895		
60	Pro	Lys	Lys	Leu	Ser	Ala	Glu	Cys	Ala	Lys	Tyr	Trp	Gly	Glu	Ile	Ile	
				900					905					910			
65	Ser	Gln	Gln	Tyr	Asn	Phe	Asp	Arg	Asp	Asn	Thr	Glu	Val	Ala	Tyr	Leu	
		915						920					925				
70	Lys	Thr	Leu	Thr	Lys	Glu	Asp	Ile	Ile	Lys	Phe	Tyr	Lys	Glu	Met	Leu	

43/67

	930		935		940
5	Ala Val Asp Ala Pro Arg Arg His Lys Val Ser Val His Val Leu Ala				
	945		950		955 960
10	Arg Glu Met Asp Ser Cys Pro Val Val Gly Glu Phe Pro Cys Gln Asn				
			965		970 975
15	Asp Ile Asn Leu Ser Gln Ala Pro Ala Leu Pro Gln Pro Glu Val Ile				
			980		985 990
20	Gln Asn Met Thr Glu Phe Lys Arg Gly Leu Pro Leu Phe Pro Leu Val				
			995		1000 1005
25	Lys Pro His Ile Asn Phe Met Ala Ala Lys Leu				
			1010		1015
	<210>		12		
	<211>		1265		
	<212>		PRT		
30	<213>		Arabidopsis thaliana		
	<400>		12		
35	Met Ala Ser Ser Ser Ser Ser Ile Phe Thr Gly Val Lys Phe Ser Pro				
	1		5		10 15
40	Ile Leu Ala Pro Phe Asn Ser Gly Asp Ser Arg Arg Ser Arg Tyr Leu				
			20		25 30
45	Lys Asp Ser Arg Asn Lys Val Arg Phe Asn Pro Ser Ser Pro Arg Leu				
			35		40 45
50	Thr Pro His Arg Val Arg Val Glu Ala Pro Ser Leu Ile Pro Tyr Asn				
			50		55 60
55	Gly Leu Trp Ala Ala Gln Pro Asn Ser His Lys Gly Arg Leu Lys Arg				
			65		70 75 80
60	Asn Ile Val Ser Gly Lys Glu Ala Thr Gly Ile Ser Leu Ser Gln Gly				
			85		90 95
65	Arg Asn Phe Cys Leu Thr Cys Lys Arg Asn Gln Ala Gly Ile Arg Arg				
			100		105 110
70	Ala Leu Pro Ser Ala Phe Val Asp Arg Thr Ala Phe Ser Leu Ser Arg				
			115		120 125

44/67

Ser Ser Leu Thr Ser Ser Leu Arg Lys His Ser Gln Ile Val Asn Ala
 130 135 140

5

Thr Leu Gly Pro Asp Glu Pro His Ala Ala Gly Thr Ala Trp Pro Asp
 145 150 155 160

10

Gly Ile Val Ala Glu Arg Gln Asp Leu Asp Leu Leu Pro Pro Glu Ile
 165 170 175

15

Asp Ser Ala Glu Leu Glu Ala Phe Leu Gly Cys Glu Leu Pro Ser His
 180 185 190

20

Pro Lys Leu His Arg Gly Gln Leu Lys Asn Gly Leu Arg Tyr Leu Ile
 195 200 205

Leu Pro Asn Lys Val Pro Pro Asn Arg Phe Glu Ala His Met Glu Val
 210 215 220

25

His Val Gly Ser Ile Asp Glu Glu Glu Asp Glu Gln Gly Ile Ala His
 225 230 235 240

30

Met Ile Glu His Val Ala Phe Leu Gly Ser Lys Lys Arg Glu Lys Leu
 245 250 255

35

Leu Gly Thr Gly Ala Arg Ser Asn Ala Tyr Thr Asp Phe His His Thr
 260 265 270

40

Val Phe His Ile His Ser Pro Thr His Thr Lys Asp Ser Glu Asp Asp
 275 280 285

Leu Phe Pro Ser Val Leu Asp Ala Leu Asn Glu Ile Ala Phe His Pro
 290 295 300

45

Lys Phe Leu Ser Ser Arg Val Glu Lys Glu Arg Arg Ala Ile Leu Ser
 305 310 315 320

50

Glu Leu Gln Met Met Asn Thr Ile Glu Tyr Arg Val Asp Cys Gln Leu
 325 330 335

55

Leu Gln His Leu His Ser Glu Asn Lys Leu Gly Arg Arg Phe Pro Ile
 340 345 350

Gly Leu Glu Glu Gln Ile Lys Lys Trp Asp Val Asp Lys Ile Arg Lys
 355 360 365

45/67

Phe His Glu Arg Trp Tyr Phe Pro Ala Asn Ala Thr Leu Tyr Ile Val
 370 375 380

5
 Gly Asp Ile Asp Asn Ile Pro Arg Ile Val His Asn Ile Glu Ala Val
 385 390 395 400

10 Phe Gly Lys Asn Gly Leu Asp Asn Glu Ser Thr Pro Ser Ser Pro Ser
 405 410 415

15 Pro Gly Ala Phe Gly Ala Met Ala Asn Phe Leu Val Pro Lys Leu Pro
 420 425 430

20 Ala Gly Leu Gly Gly Thr Phe Ser Asn Glu Lys Thr Asn Thr Ala Asp
 435 440 445
 Gln Ser Lys Met Ile Lys Arg Glu Arg His Ala Ile Arg Pro Pro Val
 450 455 460

25 Glu His Asn Trp Ser Leu Pro Gly Thr Ser Val Asp Leu Lys Pro Pro
 465 470 475 480

30 Gln Ile Phe Lys His Glu Leu Leu Gln Asn Phe Ala Ile Asn Met Phe
 485 490 495

Cys Lys Ile Pro Val Ser Lys Val Gln Thr Phe Gly Asp Leu Arg Asn
 500 505 510

35 Val Leu Met Lys Arg Ile Phe Leu Ser Ala Leu His Phe Arg Ile Asn
 515 520 525

40 Thr Arg Tyr Lys Ser Ser Asn Pro Pro Phe Thr Ser Val Glu Leu Asp
 530 535 540

45 His Ser Asp Ser Gly Arg Glu Gly Cys Thr Val Thr Thr Leu Thr Val
 545 550 555 560

50 Thr Ala Glu Pro Gln Asn Trp Gln Asn Ala Val Lys Val Ala Val Gln
 565 570 575

Glu Val Arg Arg Leu Lys Glu Phe Gly Val Thr Arg Gly Glu Leu Thr
 580 585 590

55 Arg Tyr Met Asp Ala Leu Leu Lys Asp Ser Glu His Leu Ala Ala Met
 595 600 605

46/67

Ile Asp Asn Val Ser Ser Val Asp Asn Leu Asp Phe Ile Met Glu Ser
 610 615 620

5 Asp Ala Leu Ser His Thr Val Met Asp Gln Thr Gln Gly His Glu Thr
 625 630 635 640

10 Leu Val Ala Val Ala Gly Thr Val Thr Leu Glu Glu Val Asn Thr Val
 645 650 655

15 Gly Ala Lys Val Leu Glu Phe Ile Ser Asp Phe Gly Arg Pro Thr Ala
 660 665 670

Leu Leu Pro Ala Ala Ile Val Ala Cys Val Pro Thr Lys Val His Val
 675 680 685

20 Asp Gly Val Gly Glu Ser Asp Phe Asn Ile Ser Pro Asp Glu Ile Ile
 690 695 700

Glu Ser Val Lys Ser Gly Leu Leu Ala Pro Ile Glu Ala Glu Pro Glu
 705 710 715 720

25 Leu Glu Val Pro Lys Glu Leu Ile Ser Gln Ser Gln Leu Lys Glu Leu
 725 730 735

30 Thr Leu Gln Arg Asn Pro Cys Phe Val Pro Ile Pro Gly Ser Gly Leu
 740 745 750

35 Thr Lys Leu His Asp Lys Glu Thr Gly Ile Thr Gln Leu Arg Leu Ser
 755 760 765

40 Asn Gly Ile Ala Val Asn Tyr Lys Lys Ser Thr Thr Glu Ser Arg Ala
 770 775 780

45 Gly Val Met Arg Leu Ile Val Gly Gly Gly Arg Ala Ala Glu Thr Ser
 785 790 795 800

Asp Ser Lys Gly Ala Val Val Val Gly Val Arg Thr Leu Ser Glu Gly
 805 810 815

50 Gly Arg Val Gly Asp Phe Ser Arg Glu Gln Val Glu Leu Phe Cys Val
 820 825 830

55 Asn His Leu Ile Asn Cys Ser Leu Glu Ser Thr Glu Glu Phe Ile Ala
 835 840 845

Met Glu Phe Arg Phe Thr Leu Arg Asp Asn Gly Met Gln Ala Ala Phe

47/67

	850		855		860											
5	Gln 865	Leu 865	Leu 865	His 865	Met 865	Val 870	Leu 870	Glu 870	Arg 870	Ser 875	Val 875	Trp 875	Leu 875	Glu 875	Asp 880	Ala 880
10	Phe 885	Asp 885	Arg 885	Ala 885	Arg 885	Gln 885	Leu 885	Tyr 885	Leu 890	Ser 890	Tyr 890	Phe 890	Arg 890	Ser 895	Ile 895	Pro 895
15	Lys 900	Ser 900	Leu 900	Glu 900	Arg 900	Ala 900	Thr 900	Ala 905	His 905	Lys 905	Leu 905	Met 910	Ile 910	Ala 910	Met 910	Leu 910
20	Asn 915	Gly 915	Asp 915	Glu 915	Arg 915	Phe 920	Val 920	Glu 920	Pro 920	Thr 925	Pro 925	Lys 925	Ser 925	Leu 925	Gln 925	Ser 925
	Leu 930	Asn 930	Leu 930	Glu 930	Ser 930	Val 935	Lys 935	Asp 935	Ala 935	Val 940	Met 940	Ser 940	His 940	Phe 940	Val 940	Gly 940
25	Asp 945	Asn 945	Met 945	Glu 945	Val 945	Ser 950	Ile 950	Val 950	Gly 950	Asp 955	Phe 955	Ser 955	Glu 955	Glu 955	Glu 960	Ile 960
	Glu 965	Arg 965	Cys 965	Ile 965	Leu 965	Asp 965	Tyr 965	Leu 970	Gly 970	Thr 970	Val 970	Lys 970	Ala 970	Ser 970	His 975	Asp 975
30	Ser 980	Ala 980	Lys 980	Pro 980	Pro 980	Gly 980	Ser 980	Glu 985	Pro 985	Ile 985	Leu 985	Phe 985	Arg 985	Gln 990	Pro 990	Thr 990
35	Ala 995	Gly 995	Leu 995	Gln 995	Phe 995	Gln 1000	Gln 1000	Val 1000	Phe 1000	Leu 1000	Lys 1000	Asp 1005	Thr 1005	Asp 1005	Glu 1005	Arg 1005
40	Ala 1010	Cys 1010	Ala 1010	Tyr 1010	Ile 1010	Ala 1015	Gly 1015	Pro 1015	Ala 1015	Pro 1015	Asn 1020	Arg 1020	Trp 1020	Gly 1020	Phe 1020	
45	Thr 1025	Val 1025	Asp 1025	Gly 1025	Asp 1025	Asp 1025	Leu 1030	Phe 1030	Gln 1030	Ser 1030	Val 1030	Ser 1035	Lys 1035	Leu 1035	Pro 1035	
	Val 1040	Ala 1040	His 1040	Asp 1040	Gly 1040	Leu 1045	Leu 1045	Lys 1045	Ser 1045	Glu 1045	Glu 1045	Gln 1050	Leu 1050	Leu 1050	Glu 1050	
50	Gly 1055	Gly 1055	Asp 1055	Arg 1055	Glu 1055	Leu 1055	Gln 1060	Lys 1060	Lys 1060	Leu 1060	Arg 1060	Ala 1065	His 1065	Pro 1065	Leu 1065	
55	Phe 1070	Phe 1070	Gly 1070	Val 1070	Thr 1070	Met 1070	Gly 1075	Leu 1075	Leu 1075	Ala 1075	Glu 1075	Ile 1080	Ile 1080	Asn 1080	Ser 1080	
	Arg 1085	Leu 1085	Phe 1085	Thr 1085	Thr 1085	Val 1085	Arg 1090	Asp 1090	Ser 1090	Leu 1090	Gly 1090	Leu 1095	Thr 1095	Tyr 1095	Asp 1095	

48/67

5 Val Ser Phe Glu Leu Asn Leu Phe Asp Arg Leu Lys Leu Gly Trp
 1100 1105 1110
 Tyr Val Ile Ser Val Thr Ser Thr Pro Gly Lys Val Tyr Lys Ala
 1115 1120 1125
 10 Val Asp Ala Cys Lys Asn Val Leu Arg Gly Leu His Ser Asn Gln
 1130 1135 1140
 15 Ile Ala Pro Arg Glu Leu Asp Arg Ala Lys Arg Thr Leu Leu Met
 1145 1150 1155
 Arg His Glu Ala Glu Leu Lys Ser Asn Ala Tyr Trp Leu Asn Leu
 1160 1165 1170
 20 Leu Ala His Leu Gln Ala Ser Ser Val Gln Arg Lys Glu Leu Ser
 1175 1180 1185
 25 Cys Ile Lys Glu Leu Val Ser Leu Tyr Glu Ala Ala Ser Ile Glu
 1190 1195 1200
 30 Asp Ile Tyr Leu Ala Tyr Asn Gln Leu Arg Val Asp Glu Asp Ser
 1205 1210 1215
 35 Leu Tyr Ser Cys Ile Gly Ile Ala Gly Ala Gln Ala Gly Glu Glu
 1220 1225 1230
 Ile Thr Val Leu Ser Glu Glu Glu Glu Pro Glu Asp Val Phe Ser
 1235 1240 1245
 40 Gly Val Val Pro Val Gly Arg Gly Ser Ser Met Thr Thr Arg Pro
 1250 1255 1260
 45 Thr Thr
 1265
 50 <210> 13
 <211> 534
 <212> PRT
 <213> Homo sapiens
 55 <400> 13
 Met Arg Pro Asp Asp Lys Tyr His Glu Lys Gln Ala Gln Val Glu Ala
 1 5 10 15

49/67

Thr Lys Leu Lys Gln Lys Val Glu Ala Leu Ser Pro Gly Asp Arg Gln
 20 25 30

5 Gln Ile Tyr Glu Lys Gly Leu Glu Leu Arg Ser Gln Gln Ser Lys Pro
 35 40 45

10 Gln Asp Ala Ser Cys Leu Pro Ala Leu Lys Val Ser Asp Ile Glu Pro
 50 55 60

Thr Ile Pro Val Thr Glu Leu Asp Val Val Leu Thr Ala Gly Asp Ile
 65 70 75 80

15 Pro Val Gln Tyr Cys Ala Gln Pro Thr Asn Gly Met Val Tyr Phe Arg
 85 90 95

20 Ala Phe Ser Ser Leu Asn Thr Leu Pro Glu Glu Leu Arg Pro Tyr Val
 100 105 110

25 Pro Leu Phe Cys Ser Val Leu Thr Lys Leu Gly Cys Gly Leu Leu Asp
 115 120 125

Tyr Arg Glu Gln Ala Gln Gln Ile Glu Leu Lys Thr Gly Gly Met Ser
 130 135 140

30 Ala Ser Pro His Val Leu Pro Asp Asp Ser His Met Asp Thr Tyr Glu
 145 150 155 160

35 Gln Gly Val Leu Phe Ser Ser Leu Cys Leu Asp Arg Asn Leu Pro Asp
 165 170 175

40 Met Met Gln Leu Trp Ser Glu Ile Phe Asn Asn Pro Cys Phe Glu Glu
 180 185 190

45 Glu Glu His Phe Lys Val Leu Val Lys Met Thr Ala Gln Glu Leu Ala
 195 200 205

Asn Gly Ile Pro Asp Ser Gly His Leu Tyr Ala Ser Ile Arg Ala Gly
 210 215 220

50 Arg Thr Leu Thr Pro Ala Gly Asp Leu Gln Glu Thr Phe Ser Gly Met
 225 230 235 240

55 Asp Gln Val Arg Leu Met Lys Arg Ile Ala Glu Met Thr Asp Ile Lys
 245 250 255

50/67

Pro Ile Leu Arg Lys Leu Pro Arg Ile Lys Lys His Leu Leu Asn Gly
 260 265 270

5 Asp Asn Met Arg Cys Ser Val Asn Ala Thr Pro Gln Gln Met Pro Gln
 275 280 285

10 Thr Glu Lys Ala Val Glu Asp Phe Leu Arg Ser Ile Gly Arg Ser Lys
 290 295 300

Lys Glu Arg Arg Pro Val Arg Pro His Thr Val Glu Lys Pro Val Pro
 305 310 315 320

15 Ser Ser Ser Gly Gly Asp Ala His Val Pro His Gly Ser Gln Val Ile
 325 330 335

20 Arg Lys Leu Val Met Glu Pro Thr Phe Lys Pro Trp Gln Met Lys Thr
 340 345 350

His Phe Leu Met Pro Phe Pro Val Asn Tyr Val Gly Glu Cys Ile Arg
 355 360 365

25 Thr Val Pro Tyr Thr Asp Pro Asp His Ala Ser Leu Lys Ile Leu Ala
 370 375 380

30 Arg Leu Met Thr Ala Lys Phe Leu His Thr Glu Ile Arg Glu Lys Gly
 385 390 395 400

35 Gly Ala Tyr Gly Gly Gly Ala Lys Leu Ser His Asn Gly Ile Phe Thr
 405 410 415

40 Leu Tyr Ser Tyr Arg Asp Pro Asn Thr Ile Glu Thr Leu Gln Ser Phe
 420 425 430

Gly Lys Ala Val Asp Trp Ala Lys Ser Gly Lys Phe Thr Gln Gln Asp
 435 440 445

45 Ile Asp Glu Ala Lys Leu Ser Val Phe Ser Thr Val Asp Ala Pro Val
 450 455 460

50 Ala Pro Ser Asp Lys Gly Met Asp His Phe Leu Tyr Gly Leu Ser Asp
 465 470 475 480

55 Glu Met Lys Gln Ala His Arg Glu Gln Leu Phe Ala Val Ser His Asp
 485 490 495

Lys Leu Leu Ala Val Ser Asp Arg Tyr Leu Gly Thr Gly Lys Ser Thr

51/67

500

505

510

5 His Gly Leu Ala Ile Leu Gly Pro Glu Asn Pro Lys Ile Ala Lys Asp
 515 520 525

Pro Ser Trp Ile Ile Arg
 530
 10

<210> 14
 <211> 409
 15 <212> PRT
 <213> Bacillus subtilis

 <400> 14

20 Met Ile Lys Arg Tyr Thr Cys Pro Asn Gly Val Arg Ile Val Leu Glu
 1 5 10 15

25 Asn Asn Pro Thr Val Arg Ser Val Ala Ile Gly Val Trp Ile Gly Thr
 20 25 30

Gly Ser Arg His Glu Thr Pro Glu Ile Asn Gly Ile Ser His Phe Leu
 35 40 45
 30

Glu His Met Phe Phe Lys Gly Thr Ser Thr Lys Ser Ala Arg Glu Ile
 50 55 60

35 Ala Glu Ser Phe Asp Arg Ile Gly Gly Gln Val Asn Ala Phe Thr Ser
 65 70 75 80

40 Lys Glu Tyr Thr Cys Tyr Tyr Ala Lys Val Leu Asp Glu His Ala Asn
 85 90 95

45 Tyr Ala Leu Asp Val Leu Ala Asp Met Phe Phe His Ser Thr Phe Asp
 100 105 110

Glu Asn Glu Leu Lys Lys Glu Lys Asn Val Val Tyr Glu Glu Ile Lys
 115 120 125
 50

Met Tyr Glu Asp Ala Pro Asp Asp Ile Val His Asp Leu Leu Ser Lys
 130 135 140

55 Ala Thr Tyr Gly Asn His Ser Leu Gly Tyr Pro Ile Leu Gly Thr Glu
 145 150 155 160

52/67

	Glu	Thr	Leu	Ala	Ser	Phe	Asn	Gly	Asp	Ser	Leu	Arg	Gln	Tyr	Met	His	
					165					170					175		
5	Asp	Tyr	Tyr	Thr	Pro	Asp	Arg	Val	Val	Ile	Ser	Val	Ala	Gly	Asn	Ile	
				180					185					190			
10	Ser	Asp	Ser	Phe	Ile	Lys	Asp	Val	Glu	Lys	Trp	Phe	Gly	Ser	Tyr	Glu	
			195					200					205				
	Ala	Lys	Gly	Lys	Ala	Thr	Gly	Leu	Glu	Lys	Pro	Glu	Phe	His	Thr	Glu	
		210					215				220						
15	Lys	Leu	Thr	Arg	Lys	Lys	Glu	Thr	Glu	Gln	Ala	His	Leu	Cys	Leu	Gly	
	225					230					235					240	
20	Phe	Lys	Gly	Leu	Glu	Val	Gly	His	Glu	Arg	Ile	Tyr	Asp	Leu	Ile	Val	
					245					250					255		
25	Leu	Asn	Asn	Val	Leu	Gly	Gly	Ser	Met	Ser	Ser	Arg	Leu	Phe	Gln	Asp	
				260					265					270			
30	Val	Arg	Glu	Asp	Lys	Gly	Leu	Ala	Tyr	Ser	Val	Tyr	Ser	Tyr	His	Ser	
			275					280					285				
	Ser	Tyr	Glu	Asp	Ser	Gly	Met	Leu	Thr	Ile	Tyr	Gly	Gly	Thr	Gly	Ala	
		290					295					300					
35	Asn	Gln	Leu	Gln	Gln	Leu	Ser	Glu	Thr	Ile	Gln	Glu	Thr	Leu	Ala	Thr	
	305					310					315					320	
40	Leu	Lys	Arg	Asp	Gly	Ile	Thr	Ser	Lys	Glu	Leu	Glu	Asn	Ser	Lys	Glu	
					325					330					335		
45	Gln	Met	Lys	Gly	Ser	Leu	Met	Leu	Ser	Leu	Glu	Ser	Thr	Asn	Ser	Lys	
				340					345					350			
50	Met	Ser	Arg	Asn	Gly	Lys	Asn	Glu	Leu	Leu	Leu	Gly	Lys	His	Lys	Thr	
			355					360					365				
	Leu	Asp	Glu	Ile	Ile	Asn	Glu	Leu	Asn	Ala	Val	Asn	Leu	Glu	Arg	Val	
		370					375					380					
55	Asn	Gly	Leu	Ala	Arg	Gln	Leu	Phe	Thr	Glu	Asp	Tyr	Ala	Leu	Ala	Leu	
	385					390					395					400	

53/67

Ile Ser Pro Ser Gly Asn Met Pro Ser
405

5 <210> 15
<211> 438
<212> PRT
<213> Mycobacterium tuberculosis

10 <400> 15
Met Pro Arg Arg Ser Pro Ala Asp Pro Ala Ala Ala Leu Ala Pro Arg
1 5 10 15

15 Arg Thr Thr Leu Pro Gly Gly Leu Arg Val Val Thr Glu Phe Leu Pro
20 25 30

20 Ala Val His Ser Ala Ser Val Gly Val Trp Val Gly Val Gly Ser Arg
35 40 45

25 Asp Glu Gly Ala Thr Val Ala Gly Ala Ala His Phe Leu Glu His Leu
50 55 60

Leu Phe Lys Ser Thr Pro Thr Arg Ser Ala Val Asp Ile Ala Gln Ala
65 70 75 80

30 Met Asp Ala Val Gly Gly Glu Leu Asn Ala Phe Thr Ala Lys Glu His
85 90 95

35 Thr Cys Tyr Tyr Ala His Val Leu Gly Ser Asp Leu Pro Leu Ala Val
100 105 110

40 Asp Leu Val Ala Asp Val Val Leu Asn Gly Arg Cys Ala Ala Asp Asp
115 120 125

45 Val Glu Val Glu Arg Asp Val Val Leu Glu Glu Ile Ala Met Arg Asp
130 135 140

Asp Asp Pro Glu Asp Ala Leu Ala Asp Met Phe Leu Ala Ala Leu Phe
145 150 155 160

50 Gly Asp His Pro Val Gly Arg Pro Val Ile Gly Ser Ala Gln Ser Val
165 170 175

55 Ser Val Met Thr Arg Ala Gln Leu Gln Ser Phe His Leu Arg Arg Tyr
180 185 190

Thr Pro Glu Arg Met Val Val Ala Ala Ala Gly Asn Val Asp His Asp

54/67

195

200

205

5 Gly Leu Val Ala Leu Val Arg Glu His Phe Gly Ser Arg Leu Val Arg
210 215 220

10 Gly Arg Arg Pro Val Ala Pro Arg Lys Gly Thr Gly Arg Val Asn Gly
225 230 235 240
Ser Pro Arg Leu Thr Leu Val Ser Arg Asp Ala Glu Gln Thr His Val
245 250 255

15 Ser Leu Gly Ile Arg Thr Pro Gly Arg Gly Trp Glu His Arg Trp Ala
260 265 270

20 Leu Ser Val Leu His Thr Ala Leu Gly Gly Gly Leu Ser Ser Arg Leu
275 280 285

Phe Gln Glu Val Arg Glu Thr Arg Gly Leu Ala Tyr Ser Val Tyr Ser
290 295 300

25 Ala Leu Asp Leu Phe Ala Asp Ser Gly Ala Leu Ser Val Tyr Ala Ala
305 310 315 320

30 Cys Leu Pro Glu Arg Phe Ala Asp Val Met Arg Val Thr Ala Asp Val
325 330 335

35 Leu Glu Ser Val Ala Arg Asp Gly Ile Thr Glu Ala Glu Cys Gly Ile
340 345 350

40 Ala Lys Gly Ser Leu Arg Gly Gly Leu Val Leu Gly Leu Glu Asp Ser
355 360 365

Ser Ser Arg Met Ser Arg Leu Gly Arg Ser Glu Leu Asn Tyr Gly Lys
370 375 380

45 His Arg Ser Ile Glu His Thr Leu Arg Gln Ile Glu Gln Val Thr Val
385 390 395 400

50 Glu Glu Val Asn Ala Val Ala Arg His Leu Leu Ser Arg Arg Tyr Gly
405 410 415

55 Ala Ala Val Leu Gly Pro His Gly Ser Lys Arg Ser Leu Pro Gln Gln
420 425 430

Leu Arg Ala Met Val Gly
435

55/67

<210> 16
 <211> 34
 5 <212> DNA
 <213> Artificial Sequence

<220>
 10 <223> Oligonucleotide

<400> 16
 aatagaagct tgtcgactga tctatccaaa actg 34

15
 <210> 17
 <211> 66
 <212> DNA
 <213> Artificial Sequence

20
 <220>
 <223> Oligonucleotide

<400> 17
 25 aaaagagctc ggccagatct tctagaggat ccaagaattc tgttttatat ttgttgtaaa 60
 aagtag 66

30 <210> 18
 <211> 37
 <212> DNA
 <213> Artificial Sequence

35
 <220>
 <223> Oligonucleotide

<400> 18
 40 ttttgaattc caagatctcc catgtctcta ctggtgg 37

<210> 19
 <211> 41
 45 <212> DNA
 <213> Artificial Sequence

<220>
 50 <223> Oligonucleotide

<400> 19
 ccccgagctc gtcgaccctt ctcgaaagct ttaacgaacg c 41

55
 <210> 20
 <211> 43
 <212> DNA
 <213> Artificial Sequence

56/67

<220>
<223> Oligonucleotide

5 <400> 20
ttttgaattc aaagaatgag atttccttca atttttactg cag 43

<210> 21
<211> 37
10 <212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

15 <400> 21
tttttctaga ctaggagggg tactcatact cctcggc 37

<210> 22
20 <211> 33
<212> DNA
<213> Artificial Sequence

<220>
25 <223> Oligonucleotide

<400> 22
cgaatgtcca tcgttgcgaa cctgcagaac ctg 33

30 <210> 23
<211> 33
<212> DNA
<213> Artificial Sequence

35 <220>
<223> Oligonucleotide

<400> 23
caggttctgc aggttcctaa cgatggacat tcg 33

40 <210> 24
<211> 33
<212> DNA
<213> Artificial Sequence

45 <220>
<223> Oligonucleotide

<400> 24
50 cgaatgtcca tcgttaggaa cctgcagaac ctg 33

<210> 25
<211> 33
55 <212> DNA
<213> Artificial Sequence

<220>

57/67

<223> Oligonucleotide
 <400> 25
 caggttctgc aggttcctaa cgatggacat tcg 33
 5 <210> 26
 <211> 18
 <212> DNA
 <213> Artificial Sequence
 10 <220>
 <223> Oligonucleotide
 <400> 26
 tcgcagagaa cggatggc 18
 15
 <210> 27
 <211> 36
 <212> DNA
 20 <213> Artificial Sequence
 <220>
 <223> Oligonucleotide
 25 <400> 27
 ttttgggccc ttcatgggtga tacggtatct cttggc 36
 <210> 28
 30 <211> 37
 <212> DNA
 <213> Artificial Sequence
 <220>
 35 <223> Oligonucleotide
 <400> 28
 ttttctcgag aaggtggaac atactgccct gggatgg 37
 40
 <210> 29
 <211> 38
 <212> DNA
 <213> Artificial Sequence
 45
 <220>
 <223> Oligonucleotide
 <400> 29
 50 ttttgagctc gtttaggaaa cgtccttggc ggagatgc 38
 <210> 30
 <211> 40
 55 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Oligonucleotide

58/67

<400> 30
 tttttctaga cactgcgaat ccatgggtata aaccaaaacc 40
 <210> 31
 5 <211> 24
 <212> DNA
 <213> Artificial Sequence
 <220>
 10 <223> Oligonucleotide
 <400> 31
 gtcgttggtc atggacatac ctcc 24
 15
 <210> 32
 <211> 26
 <212> DNA
 <213> Artificial Sequence
 20
 <220>
 <223> Oligonucleotide
 <400> 32
 25 tacaaatggt cttctgcat ttctgg 26
 <210> 33
 <211> 32
 30 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Oligonucleotide
 35
 <400> 33
 gggtcatatg cgccggagct cctcgacagc ag 32
 40 <210> 34
 <211> 37
 <212> DNA
 <213> Artificial Sequence
 45 <220>
 <223> Oligonucleotide
 <400> 34
 gggtcctagg atccgcaagt ttgattccat tgcggtg 37
 50
 <210> 35
 <211> 70
 <212> DNA
 55 <213> Artificial Sequence
 <220>
 <223> Oligonucleotide

59/67

<400> 35
 ttaaagagta ccttggctat agaataccgt agagataaag acctgaatag agattgtact 60
 gagagtgcac 70

5
 <210> 36
 <211> 70
 <212> DNA
 <213> Artificial Sequence

10
 <220>
 <223> Oligonucleotide

<400> 36
 15 aggtattata actatcttttc tgtatcttttt atatatctttt atttgccaag ctgtgctgta 60
 tttcacaccg 70

20 <210> 37
 <211> 23
 <212> DNA
 <213> Artificial Sequence

25 <220>
 <223> Oligonucleotide

<400> 37
 30 ctttgggttaa agagtacctt ggc 23

<210> 38
 <211> 23
 <212> DNA
 35 <213> Artificial Sequence

<220>
 <223> Oligonucleotide

40 <400> 38
 tactacgaaa agcgtgtgctg agg 23

<210> 39
 45 <211> 71
 <212> DNA
 <213> Artificial Sequence

<220>
 50 <223> Oligonucleotide

<400> 39
 tagaaggcta ctcaaaagaa taaagttact ataaaatata ctgctggtata tagattgtac 60

55 tgagagtgcac c 71

60/67

<210> 40
 <211> 70
 <212> DNA
 <213> Artificial Sequence
 5
 <220>
 <223> Oligonucleotide
 <400> 40
 10 gatcggcaag aaactttgaa gcagtatatatt tacaggatta aattatatat ctgtgcggta 60
 tttcacaccg 70

15 <210> 41
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 20 <220>
 <223> Oligonucleotide
 <400> 41
 25 cggaggggct ctatgataaa gg 22

<210> 42
 <211> 23
 <212> DNA
 30 <213> Artificial Sequence
 <220>
 <223> Oligonucleotide
 35 <400> 42
 gagtaactag ggcttctctt ccc 23

<210> 43
 40 <211> 85
 <212> PRT
 <213> Homo sapiens
 <400> 43
 45 Ser Gly Leu Gln Arg Ala Glu Glu Ala Pro Arg Arg Gln Leu Arg Val
 1 5 10 15

50 Ser Gln Arg Thr Asp Gly Glu Ser Arg Ala His Leu Gly Ala Leu Leu
 20 25 30

Ala Arg Tyr Ile Gln Gln Ala Arg Lys Ala Pro Ser Gly Arg Met Ser
 55 35 40 45

Ile Val Lys Asn Leu Gln Asn Leu Asp Pro Ser His Arg Ile Ser Asp
 50 55 60

61/67

Arg Asp Tyr Met Gly Trp Met Asp Phe Gly Arg Arg Ser Ala Glu Glu
 65 70 75 80

5

Tyr Glu Tyr Pro Ser
 85

10

<210> 44
 <211> 22
 <212> PRT
 <213> Homo sapiens

15

<400> 44

Gln Leu Arg Val Ser Gln Arg Thr Asp Gly Glu Ser Arg Ala His Leu
 20 1 5 10 15

Gly Ala Leu Leu Ala Arg
 20

25

<210> 45
 <211> 19
 <212> PRT
 30 <213> Homo sapiens

<400> 45

35 Val Ser Gln Arg Thr Asp Gly Glu Ser Arg Ala His Leu Gly Ala Leu
 1 5 10 15

Leu Ala Arg

40

<210> 46
 <211> 51
 45 <212> PRT
 <213> Homo sapiens

<400> 46

50

Tyr Ile Gln Gln Ala Arg Lys Ala Pro Ser Gly Arg Met Ser Ile Val
 1 5 10 15

55 Lys Asn Leu Gln Asn Leu Asp Pro Ser His Arg Ile Ser Asp Arg Asp
 20 25 30
 Tyr Met Gly Trp Met Asp Phe Gly Arg Arg Ser Ala Glu Glu Tyr Glu
 35 40 45

62/67

Tyr Pro Ser
50

5 <210> 47
<211> 17
<212> PRT
<213> Homo sapiens

10
<400> 47

Tyr Ile Gln Gln Ala Arg Lys Ala Pro Ser Gly Arg Met Ser Ile Val
1 5 10 15

15

Lys

20 <210> 48
<211> 16
<212> PRT
<213> Homo sapiens

25 <400> 48

Tyr Ile Gln Gln Ala Arg Lys Ala Pro Ser Gly Arg Met Ser Ile Val
1 5 10 15

30 <210> 49
<211> 13
<212> PRT
<213> Homo sapiens

35
<400> 49

Asn Leu Gln Asn Leu Asp Pro Ser His Arg Ile Ser Asp
1 5 10

40

<210> 50
<211> 23
<212> PRT
45 <213> Homo sapiens

<400> 50

50 Asn Leu Gln Asn Leu Asp Pro Ser His Arg Ile Ser Asp Arg Asp Tyr
1 5 10 15

Met Gly Trp Met Asp Phe Gly
55 20

<210> 51
<211> 34
<212> PRT
<213> Homo sapiens

63/67

<400> 51

Asn Leu Gln Asn Leu Asp Pro Ser His Arg Ile Ser Asp Arg Asp Tyr
5 1 5 10 15

Met Gly Trp Met Asp Phe Gly Arg Arg Ser Ala Glu Glu Tyr Glu Tyr
20 25 30
10

Pro Ser

15

<210> 53
<211> 12
<212> PRT
<213> Artificial Sequence

20

<220>
<223> Portion of fusion protein

<400> 53

25

Lys Arg Glu Ala Glu Ala Ser Gly Leu Gln Arg Ala
1 5 10

30

<210> 54
<211> 9
<212> PRT
<213> Homo sapiens

35

<400> 54

Arg Met Ser Ile Val Lys Asn Leu Gln
1 5

40

<210> 55
<211> 13
<212> PRT
45 <213> Homo sapiens

<400> 55

Asp Arg Asp Tyr Met Gly Trp Met Asp Phe Gly Arg Arg
50 1 5 10

<210> 56
<211> 48
55 <212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Mfa1BNP (S)

64/67

<400> 56
5 ggataaaaga gaggctgaag ctcacccgct gggcagcccc ggttcagc 48
<210> 57
<211> 48
<212> DNA
<213> Artificial Sequence
10
<220>
<223> MFlaBNP (AS)

<400> 57
15 gctgaaccgg ggctgcccag cgggtgagct tcagcctctc ttttatcc 48
<210> 58
<211> 34
<212> DNA
20 <213> Artificial Sequence

<220>
<223> Oligonucleotide BNP5'EcoRI

25
<400> 58
ttttgaattc atggatcccc agacagcacc ttcc 34
<210> 59
30 <211> 33
<212> DNA
<213> Artificial Sequence

<220>
35 <223> Oligonucleotide BNP3'XbaI

<400> 59
40 tttttctaga ttaatgccgc ctcagcactt tgc 33
<210> 60
<211> 48
<212> DNA
<213> Artificial Sequence
45
<220>
<223> Oligonucleotide MFlBNP (S)

50 <400> 60
ggataaaaga gaggctgaag ctcacccgct gggcagcccc ggttcagc 48
<210> 61
<211> 48
55 <212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide MFlBNP (AS)

65/67

<400> 61
 gctgaaccgg ggctgcccag cgggtgagct tcagcctctc ttttatcc 48
 5
 <210> 62
 <211> 69
 <212> DNA
 <213> Artificial Sequence
 10
 <220>
 <223> Oligonucleotide YPS15'GD400
 <400> 62
 15 aaaaagataa ggtgaacacc aagcatatag tataatatta cctaccacat gattgtactg 60
 agagtgcac 69
 <210> 63
 <211> 70
 20 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Oligonucleotide YPS13'GD400
 25
 <400> 63
 aactccaact ggcttggaga tgtgaatgtc taaactttgt gcaacggttt ctgtgcggta 60
 tttcacaccg 70
 30 <210> 64
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 35 <220>
 <223> Oligonucleotide YPS15'DC
 <400> 64
 tcgtttcact gatgtgtccg 20
 40
 <210> 65
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 45
 <220>
 <223> Oligonucleotide YPS15'DC
 <400> 65
 50 gattataggc catatcccag g 21
 <210> 66
 <211> 13
 <212> PRT
 55 <213> Artificial Sequence
 <220>
 <223> Pitrilysin consensus sequence

66/67

<221> VARIANT
 <222> (1)...(13)
 <223> Xaa = Any Amino Acid
 5
 <221> VARIANT
 <222> (1)...(13)
 <223> Xaa = Any Amino Acid
 10 <221> VARIANT
 <222> (1)...(13)
 <223> Xaa = Any Amino Acid
 <221> UNSURE
 15 <222> (0)...(0)
 <400> 11
 Gly Xaa Xaa His Xaa Xaa Glu His Xaa Xaa Xaa Xaa Gly
 1 5 10
 20
 <210> 67
 <211> 44
 <212> PRT
 25 <213> Artificial Sequence
 <220>
 <223> Pitrilysin consensus sequence
 30
 <221> VARIANT
 <222> (1)...(44)
 <223> Xaa = Any Amino Acid
 35 <221> VARIANT
 <222> (1)...(44)
 <223> Xaa = Any Amino Acid
 <221> VARIANT
 40 <222> (1)...(44)
 <223> Xaa = Any Amino Acid
 <221> VARIANT
 <222> (34)...(34)
 45 <223> Xaa = any amino acid or absent
 <221> VARIANT
 <222> (35)...(35)
 <223> Xaa = any amino acid or absent
 50
 <400> 12
 Gly Xaa Xaa His Xaa Xaa Glu His Xaa Xaa Xaa Xaa Gly Xaa Xaa Xaa
 1 5 10 15
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 20 25 30
 55 Xaa Xaa Xaa Asn Ala Xaa Thr Xaa Xaa Xaa Xaa Thr
 35 40

67/67

<210> 68
 <211> 44
 <212> PRT
 <213> Artificial Sequence
 5
 <220>
 <223> Pitrilysin consensus sequence

 10 <221> VARIANT
 <222> (1)...(44)
 <223> Xaa = Any Amino Acid

 <221> VARIANT
 15 <222> (1)...(44)
 <223> Xaa = Any Amino Acid

 <221> VARIANT
 <222> (1)...(44)
 20 <223> Xaa = Any Amino Acid

 <221> VARIANT
 <222> (34)...(34)
 <223> Xaa = any amino acid or absent
 25
 <221> VARIANT
 <222> (35)...(35)
 <223> Xaa = any amino acid or absent

 30 <400> 13
 Gly Xaa Xaa His Xaa Xaa Glu His Xaa Xaa Xaa Xaa Gly Xaa Xaa Lys
 1 5 10 15
 Tyr Pro Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Xaa Xaa Xaa Xaa Xaa
 20 25 30
 35 Xaa Xaa Xaa Asn Ala Xaa Thr Xaa Xaa Xaa Xaa Thr
 35 40

Claims

1. A method for producing a protein of interest in a host cell, wherein said host cell has been genetically modified in order to express significantly reduced levels of a metalloprotease comprising a HXXEH motif (SEQ ID NO 1), compared to the corresponding
5 non-modified cell when cultured under identical conditions, the method comprising
 - a) introducing into the host cell a nucleic acid sequence encoding the protein of interest,
 - 10 b) cultivating the host cell of step (a) in a suitable growth medium for production of the protein of interest, and
 - c) isolating the protein of interest.
- 15 2. A method according to claim 1, wherein the metalloprotease further comprises a glutamic acid residue between 70 and 80 amino acids C-terminal of the second His residue in the HXXEH motif.
3. A method according to any of the preceeding claims, wherein the metalloprotease
20 further comprises a glycine residue 3 amino acids N-terminal of the first His residue in the HXXEH motif.
4. A method according to any of the preceeding claims, wherein the metalloprotease further comprises a glycine residue 5 amino acids C-terminal of the second His residue in
25 the HXXEH motif.
5. A method according to any of the preceeding claims, wherein the metalloprotease further comprises a lysine residue 8 amino acids C-terminal of the second His residue in the HXXEH motif.
30
6. A method according to any of the preceeding claims, wherein the metalloprotease further comprises a tyrosine residue 9 amino acids C-terminal of the second His residue in the HXXEH motif.
- 35 7. A method according to any of the preceeding claims, wherein the metalloprotease further comprises a proline residue 10 amino acids C-terminal of the second His residue in the HXXEH motif.

8. A method according to any of the preceeding claims, wherein the metalloprotease further comprises the consensus sequence SEQ ID NO 2.

9. A method according to any of the preceeding claims wherein the metalloprotease further
5 comprises the consensus sequence SEQ ID NO 3.

10. A method according to any of the preceeding claims, wherein the metalloprotease further comprises a NAXTXXXXT motif between 20 and 30 amino acids C-terminal of the second His residue in the HXXEH motif.

10

11. A method according to any of the preceeding claims, wherein the metalloprotease is selected from:

i) any one of the group consisting of SEQ ID NO's 4 to 15, and

15

ii) a sequence which is at least 80% identical to any one of SEQ ID NO's 4 to 15.

12. A method according to any of the preceeding claims, wherein the metalloprotease is at least 80% identical to the SEQ ID NO: 4.

20

13. A method according to any of the preceeding claims, wherein the total amount of the protein of interest is increased at least 5% compared the corresponding non-modified cell when cultured under identical conditions.

25 14. A method according to any of the preceeding claims, wherein the total amount of the protein of interest is increased at least 50% more than the corresponding non-modified cell when cultured under identical conditions.

15. The method according to any of the preceeding claims, in which the host cell is a
30 prokaryotic cell.

16. The method according to any of claims 1-14, in which the host cell is a eukaryotic cell.

17. The method according to claim 16, in which the host cell is a non-filamentous fungal
35 cell.

18. The method according to claim 16, in which the host cell is a filamentous fungal cell.

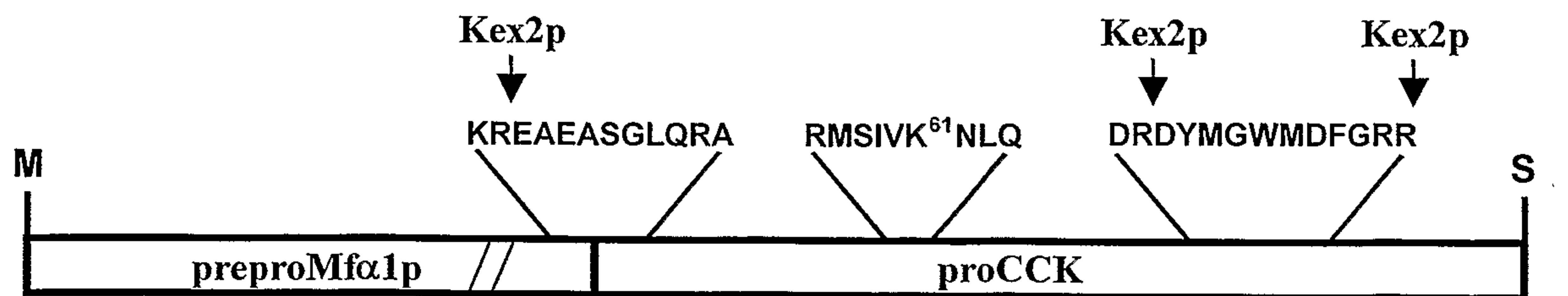
19. The method according to claim 17, in which the host cell is a strain of *Saccharomyces*.

20. The method according to claim 19, in which the host cell is *Saccharomyces cerevisiae*.

21. A host cell useful for the expression of a protein of interest, wherein said cell has been
5 genetically modified in order to express significantly reduced levels of a metalloprotease
comprising a HXXEH motif (SEQ ID NO 1) than the corresponding non-modified cell when
cultured under identical conditions.

22. A host cell according to claim 21, wherein the metalloprotease further comprises the
10 consensus sequence SEQ ID NO 3.

1/17



↓ CCK processing products secreted from *Saccharomyces cerevisiae*

C-terminally extended CCK-22



CCK-22-Gly



C-terminally extended CCK-8



Fig. 1

CCK-8-Gly

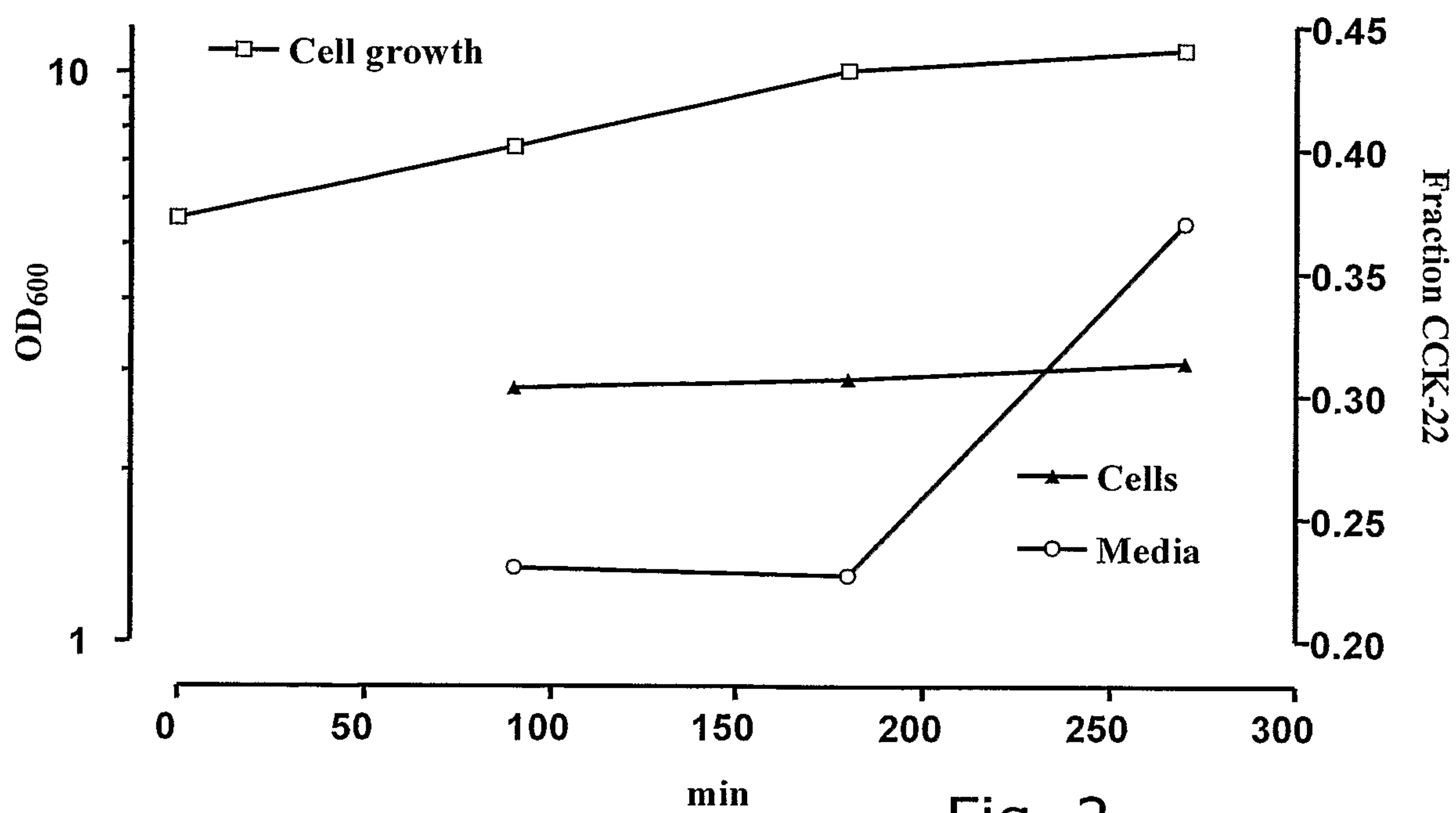


Fig. 2

2/17

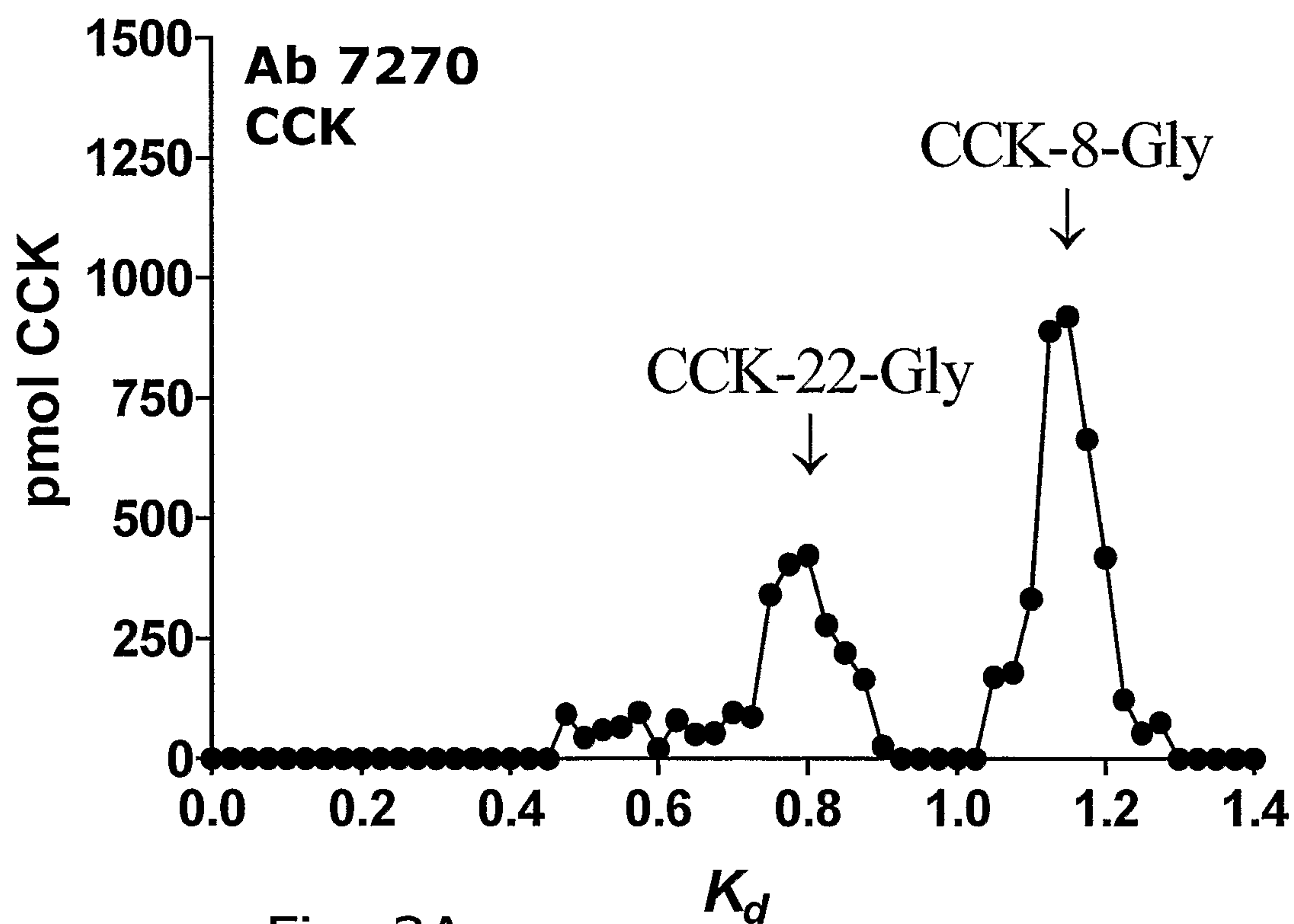


Fig. 3A

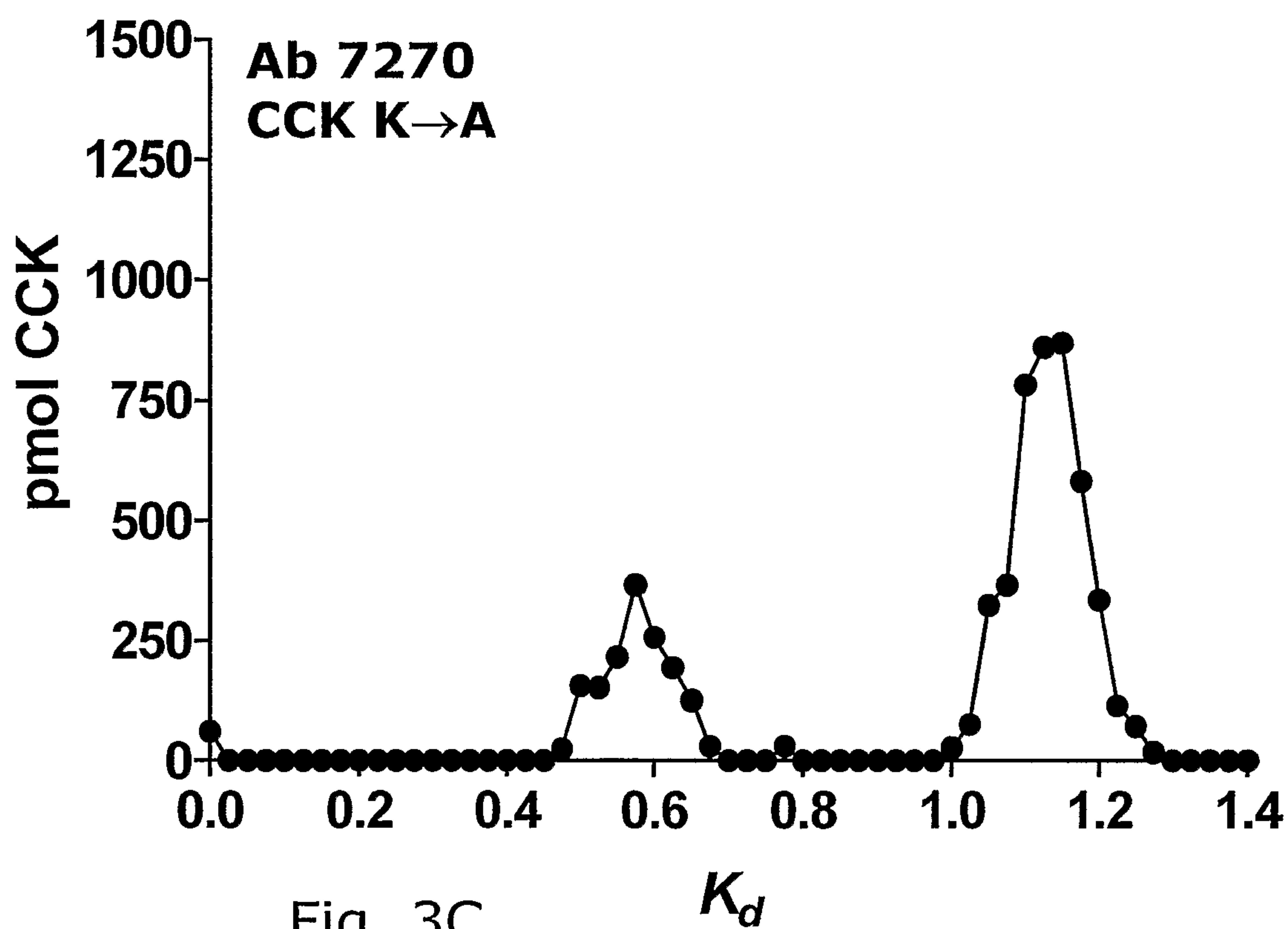


Fig. 3C

3/17

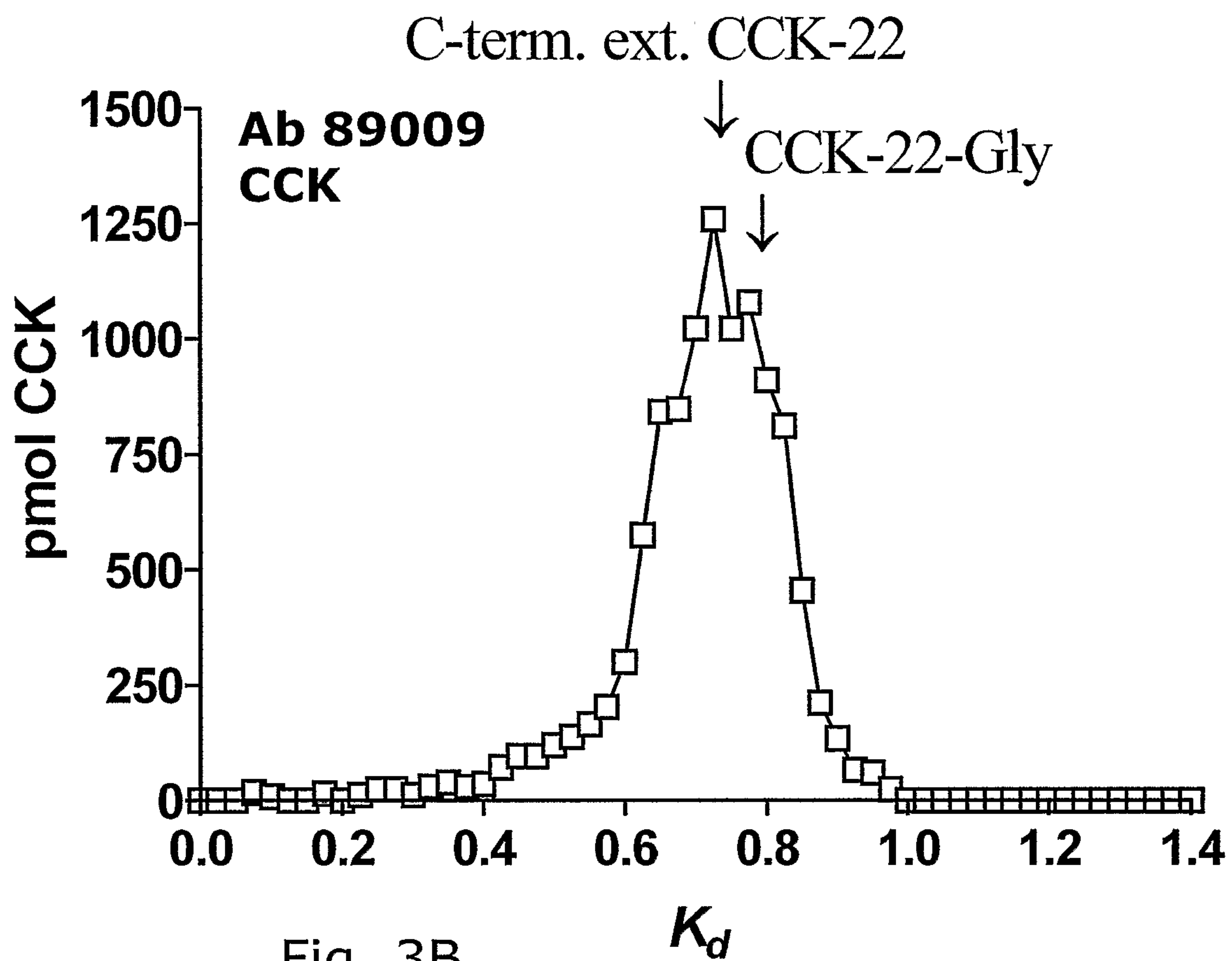


Fig. 3B

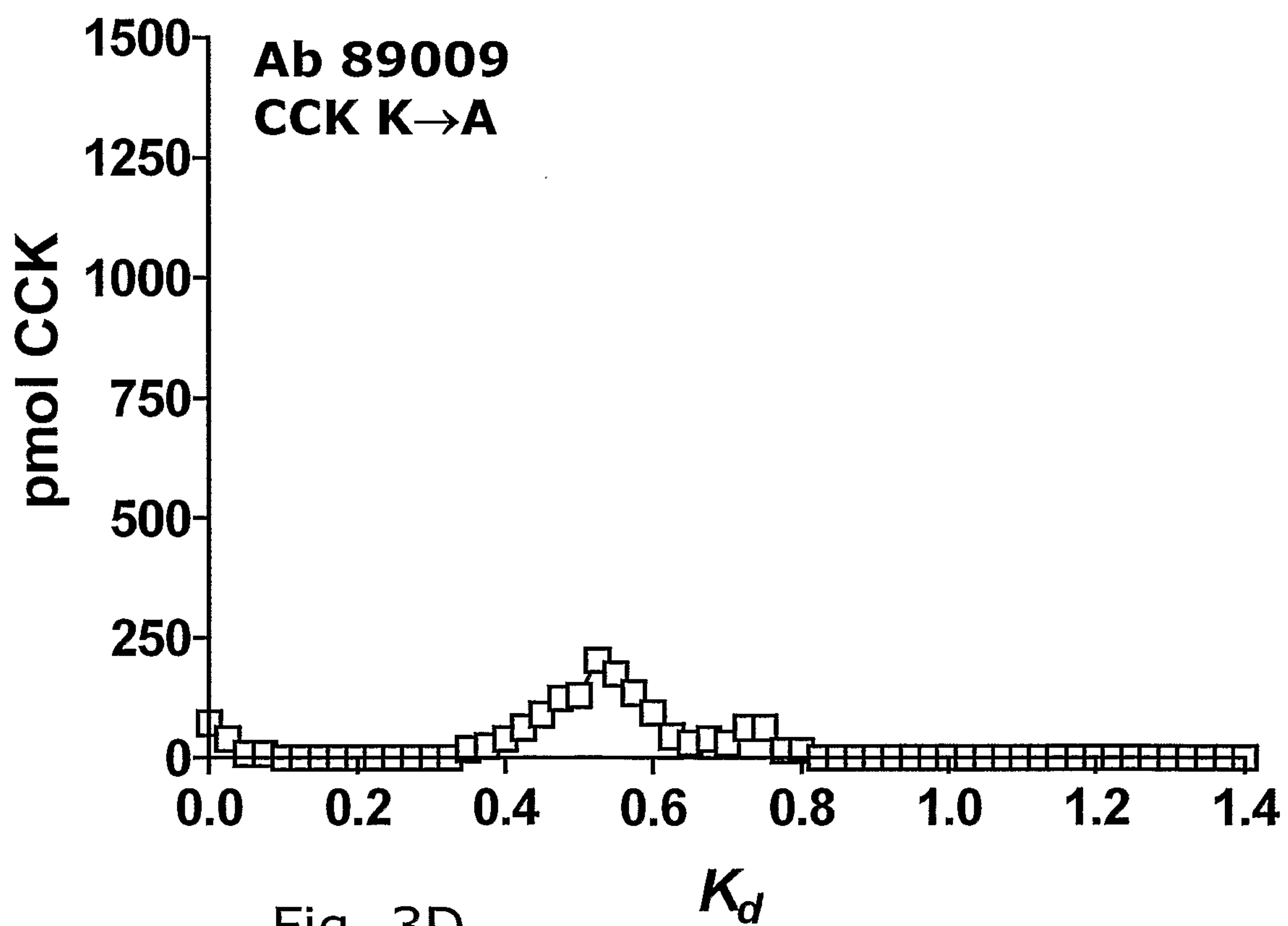
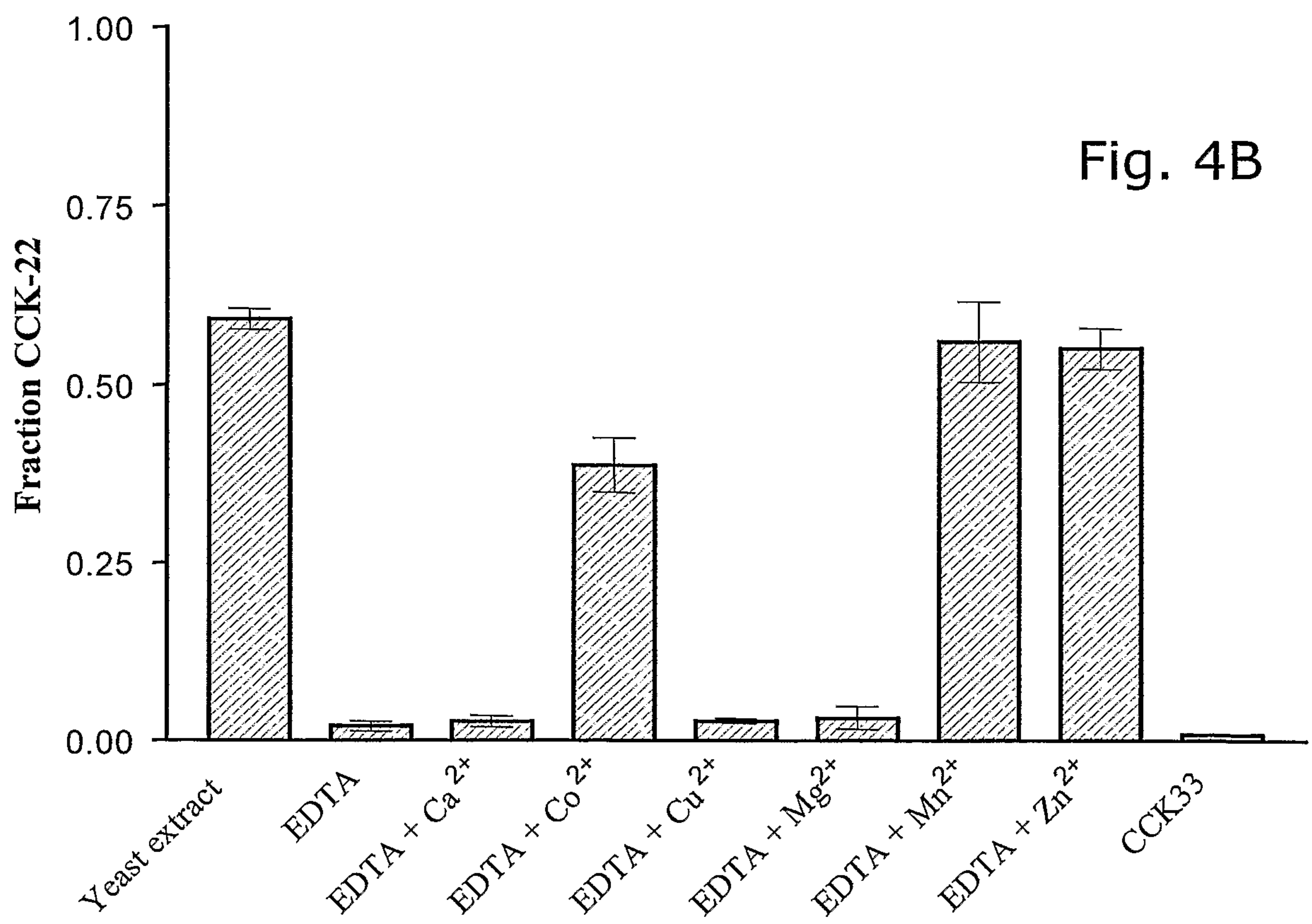
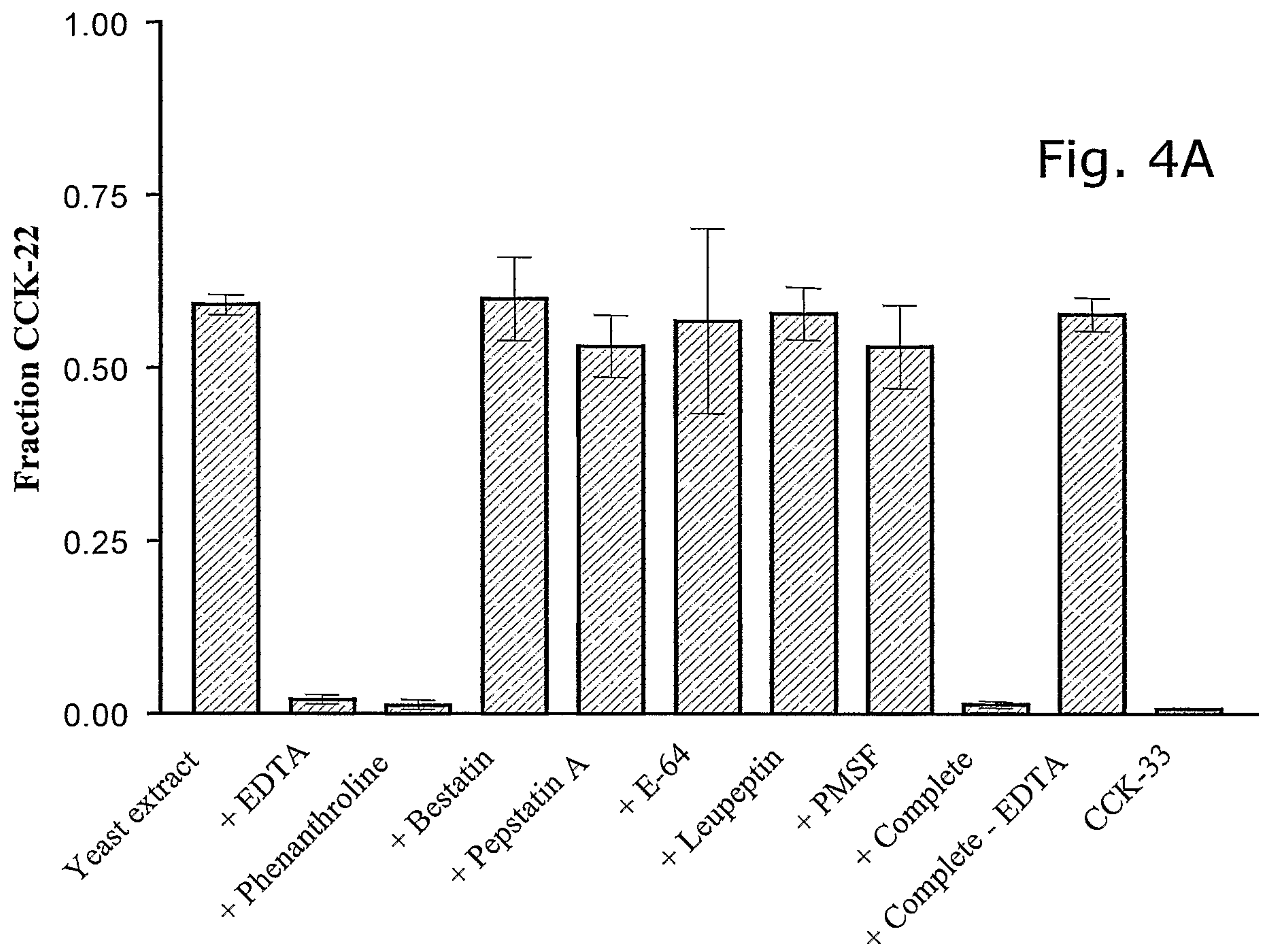


Fig. 3D

4/17



5/17

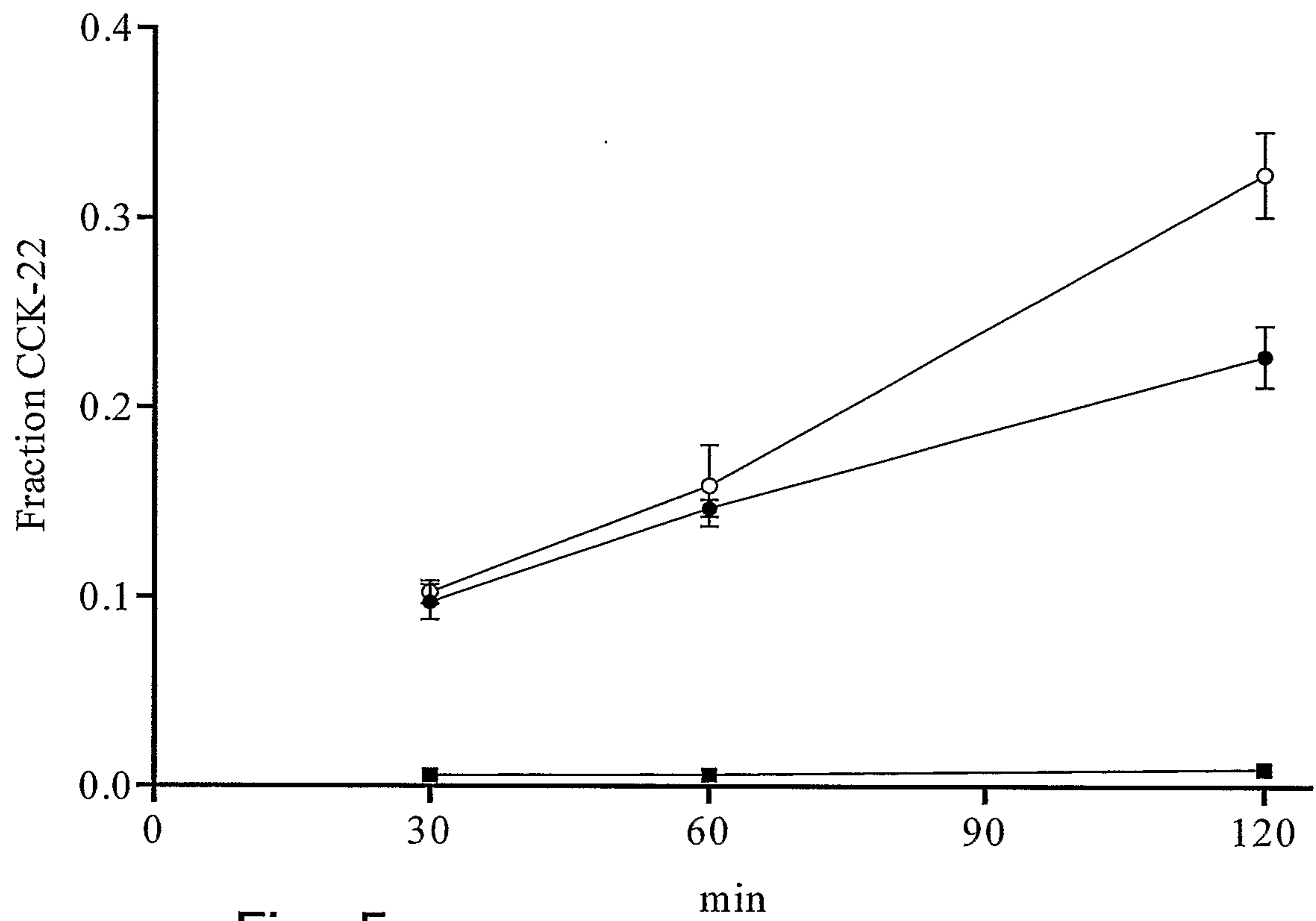


Fig. 5

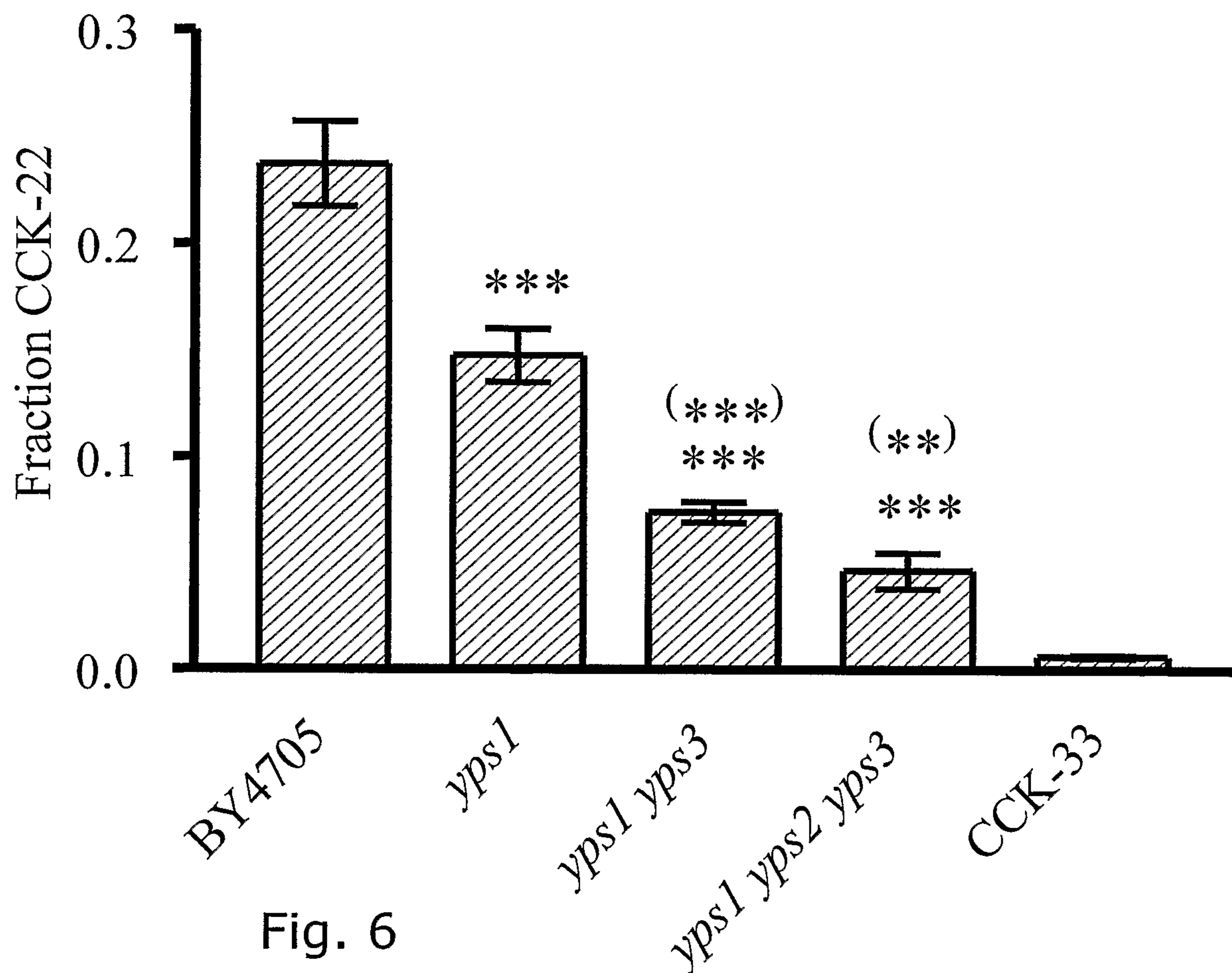
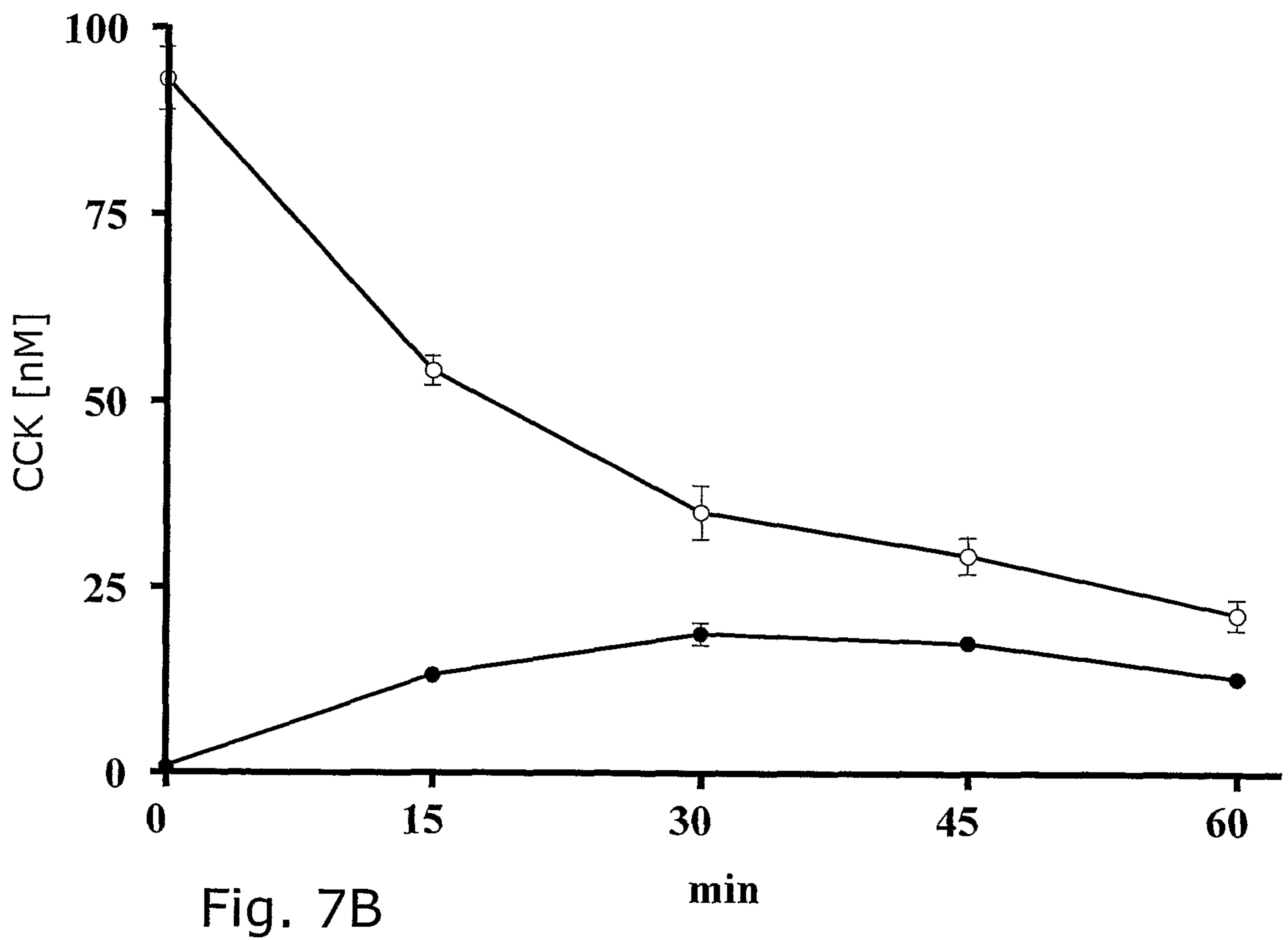
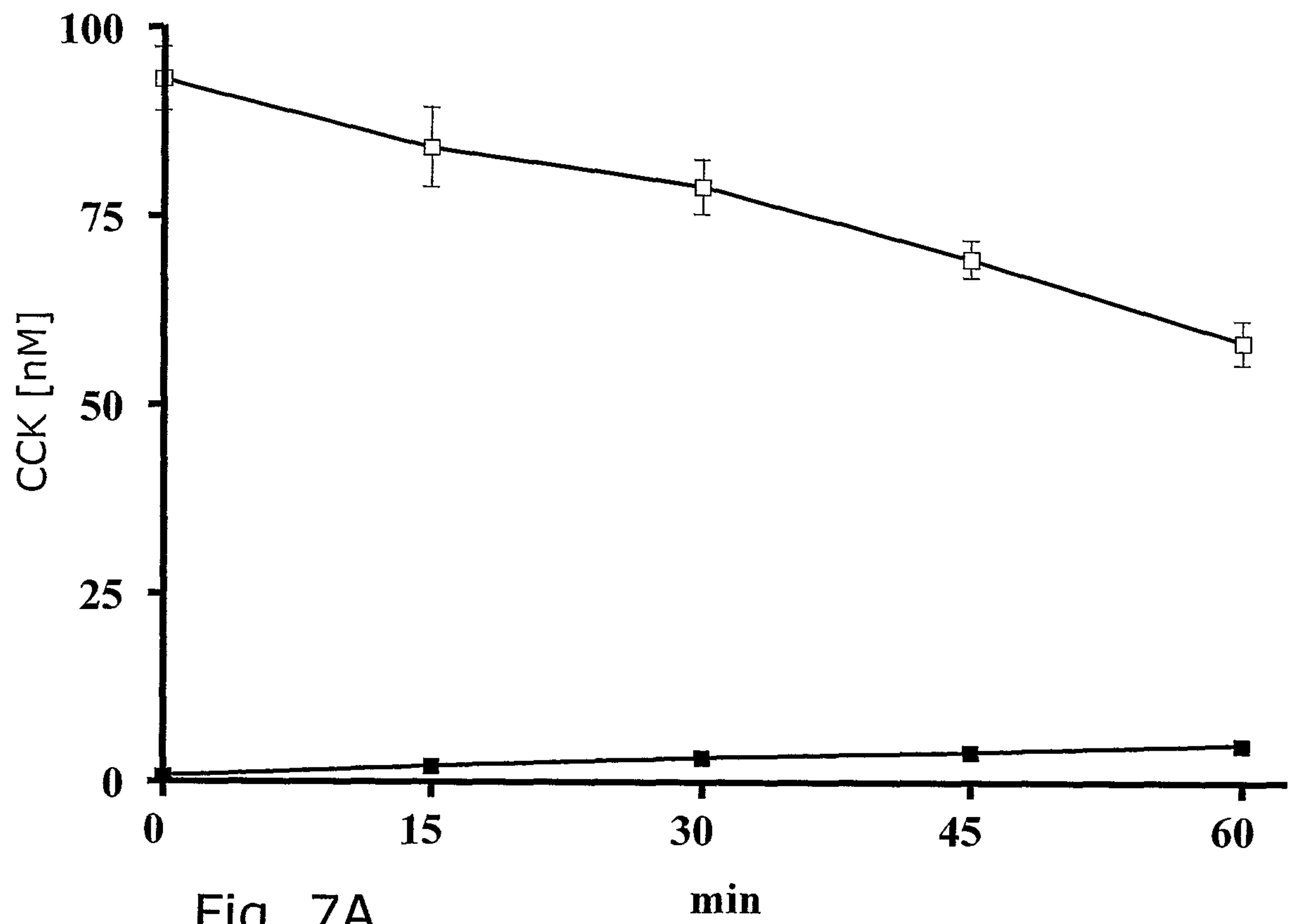


Fig. 6

6/17



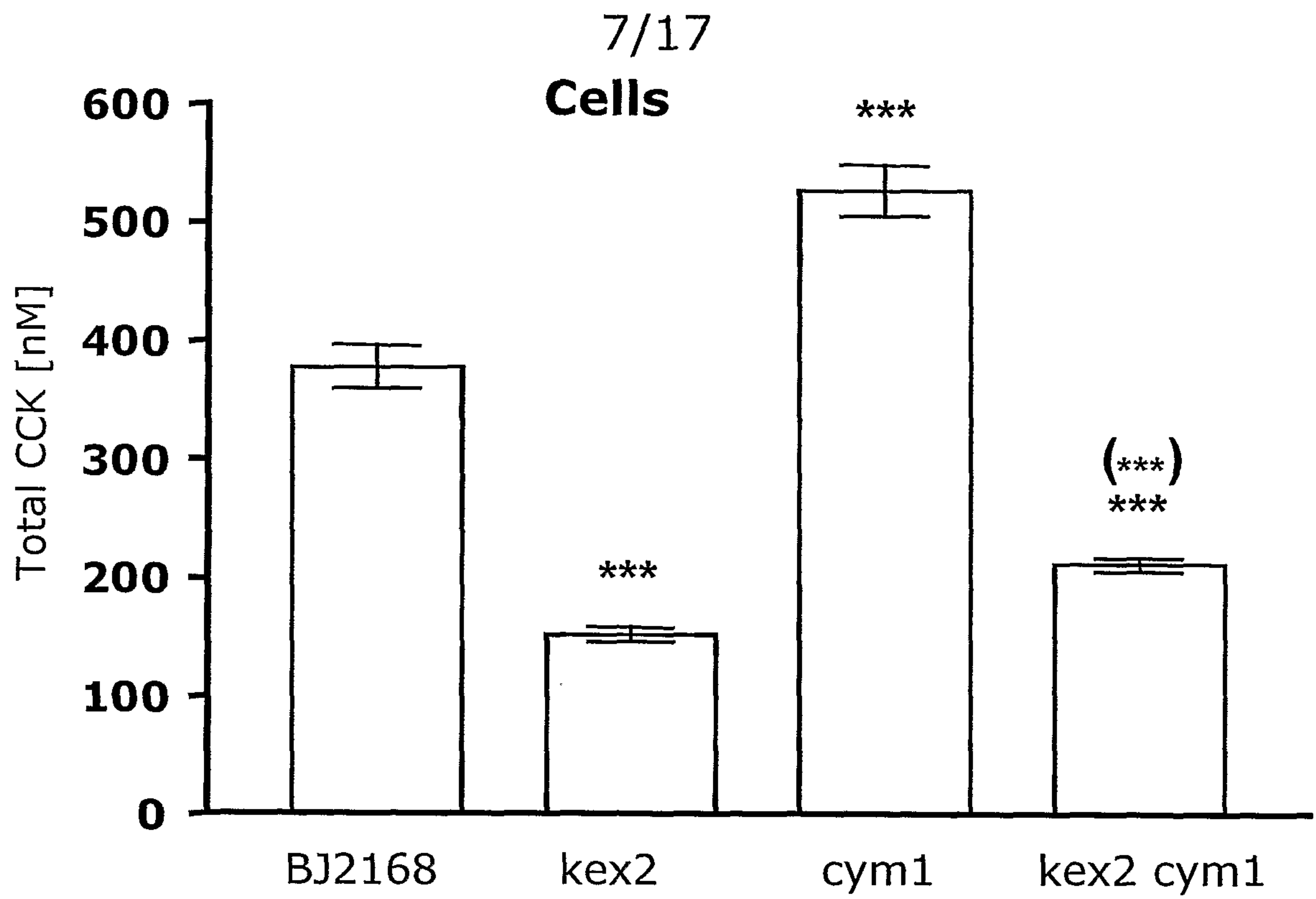


Fig. 8A

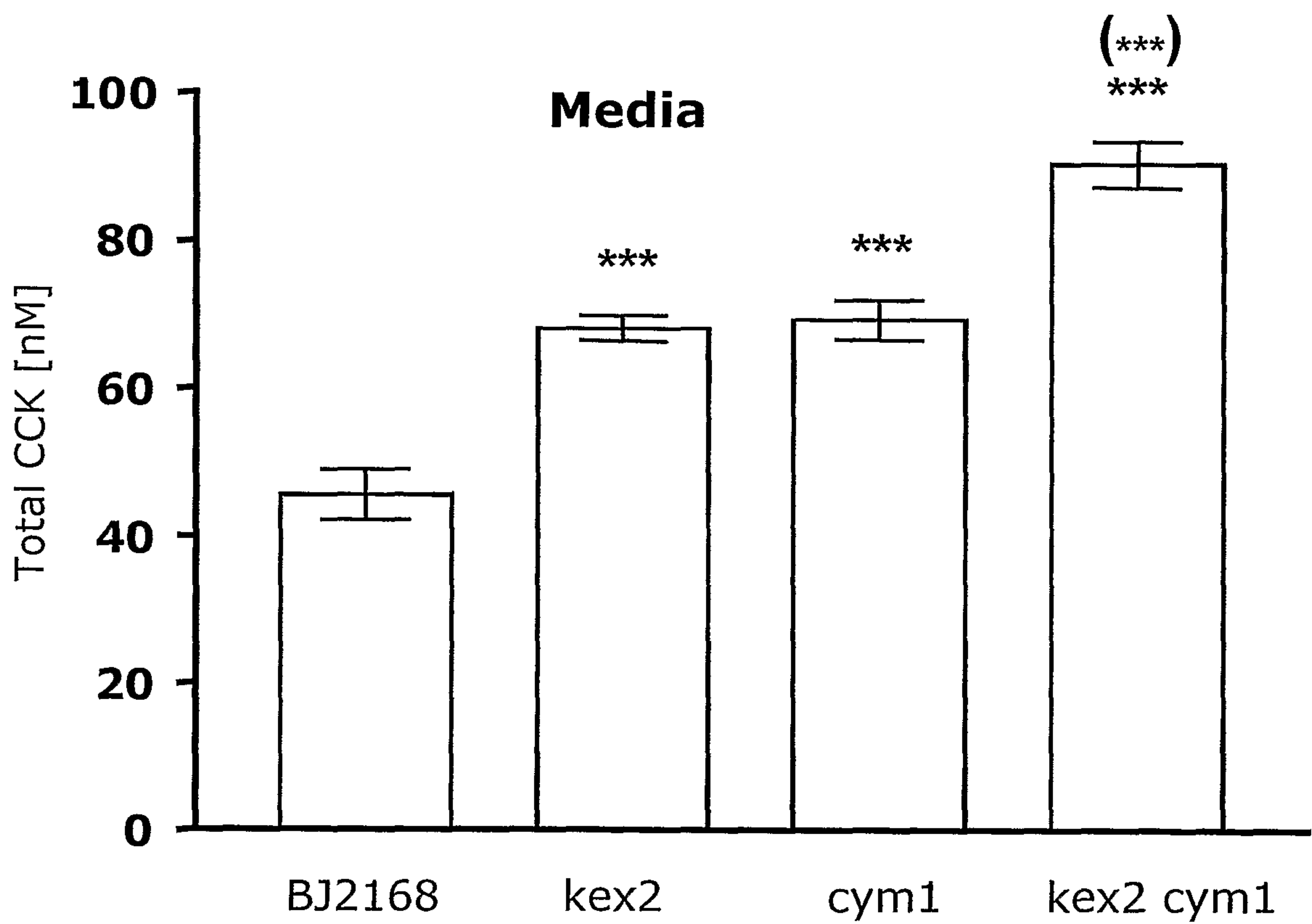


Fig. 8B

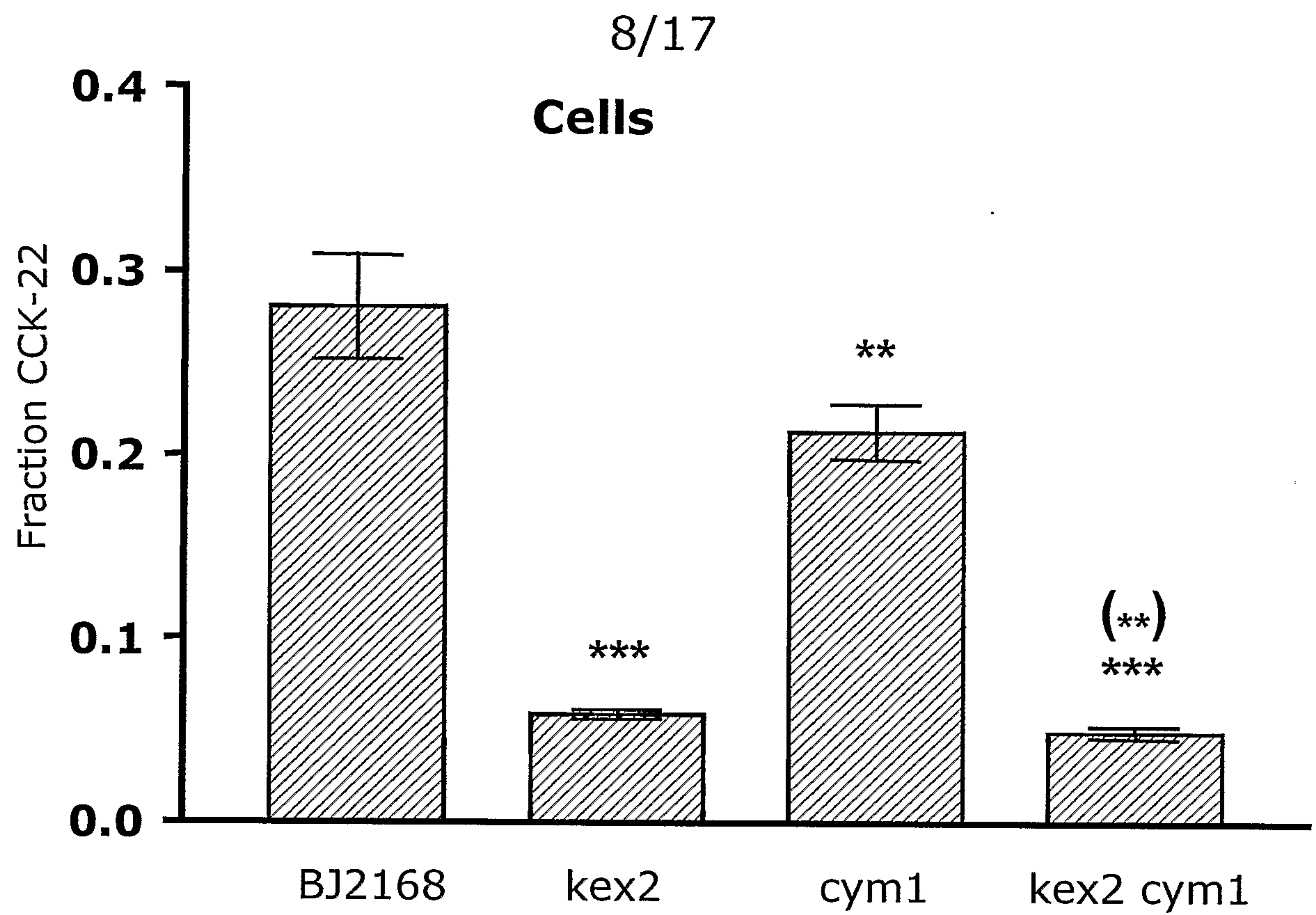


Fig. 8C

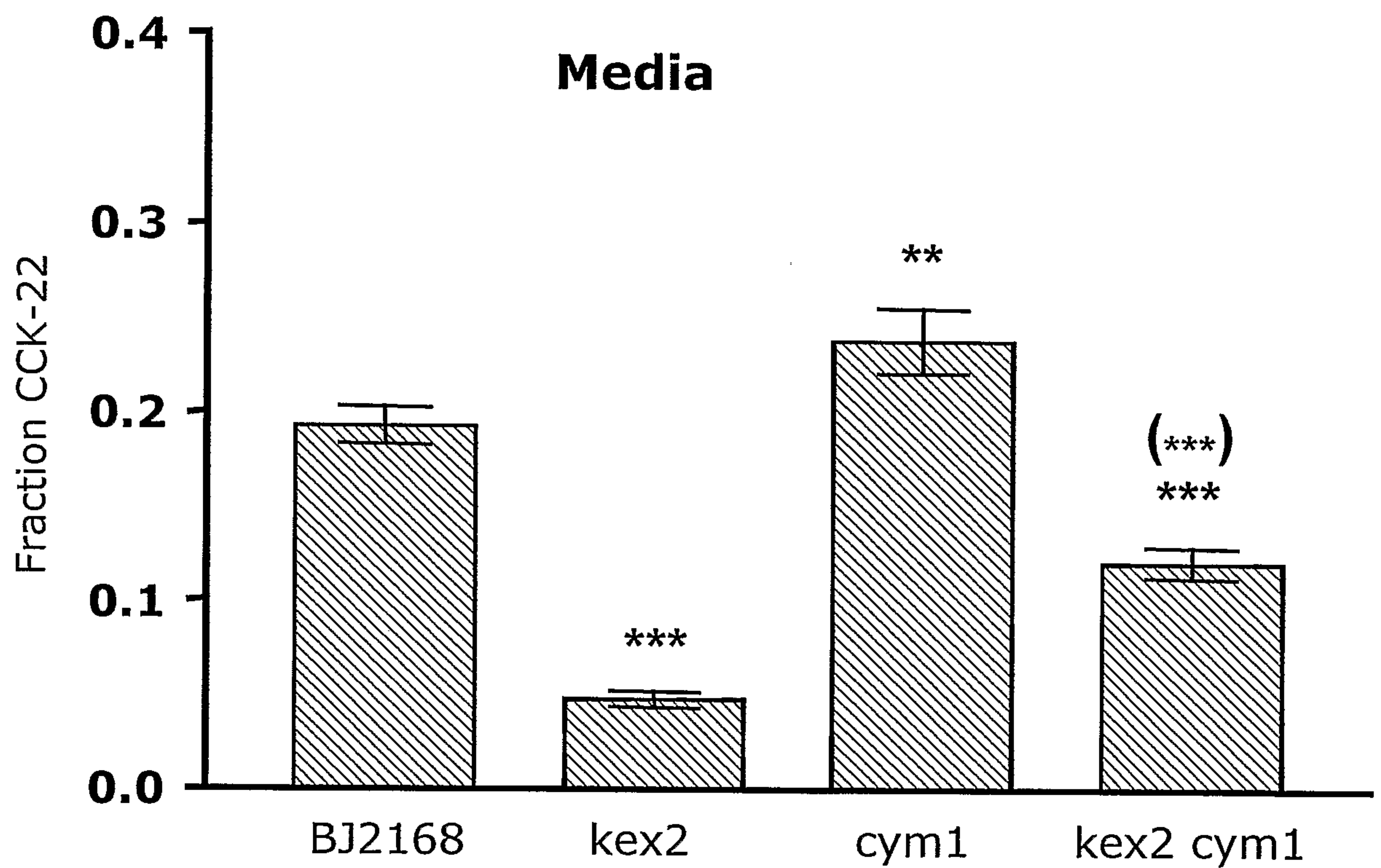


Fig. 8D

9/17

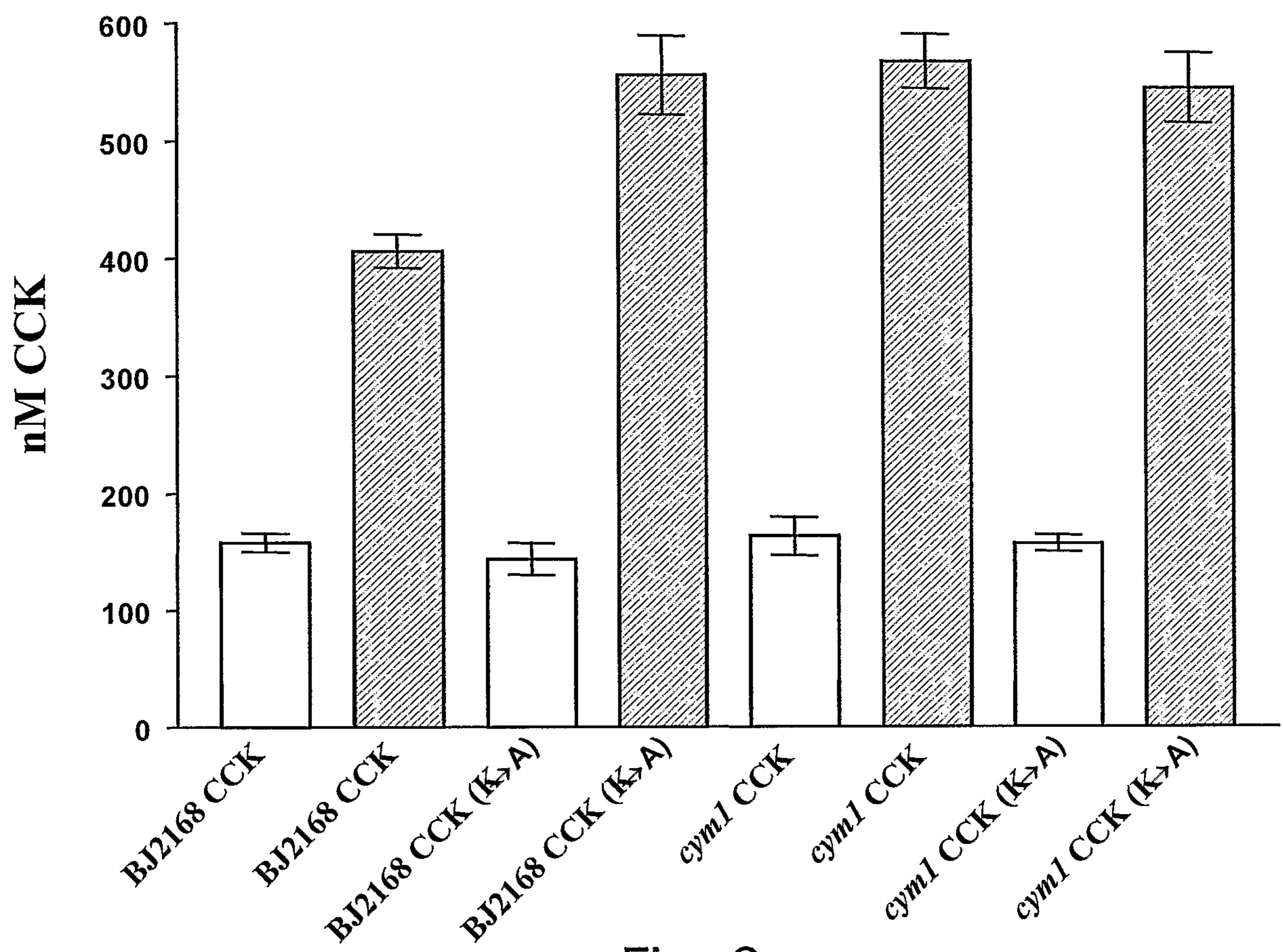
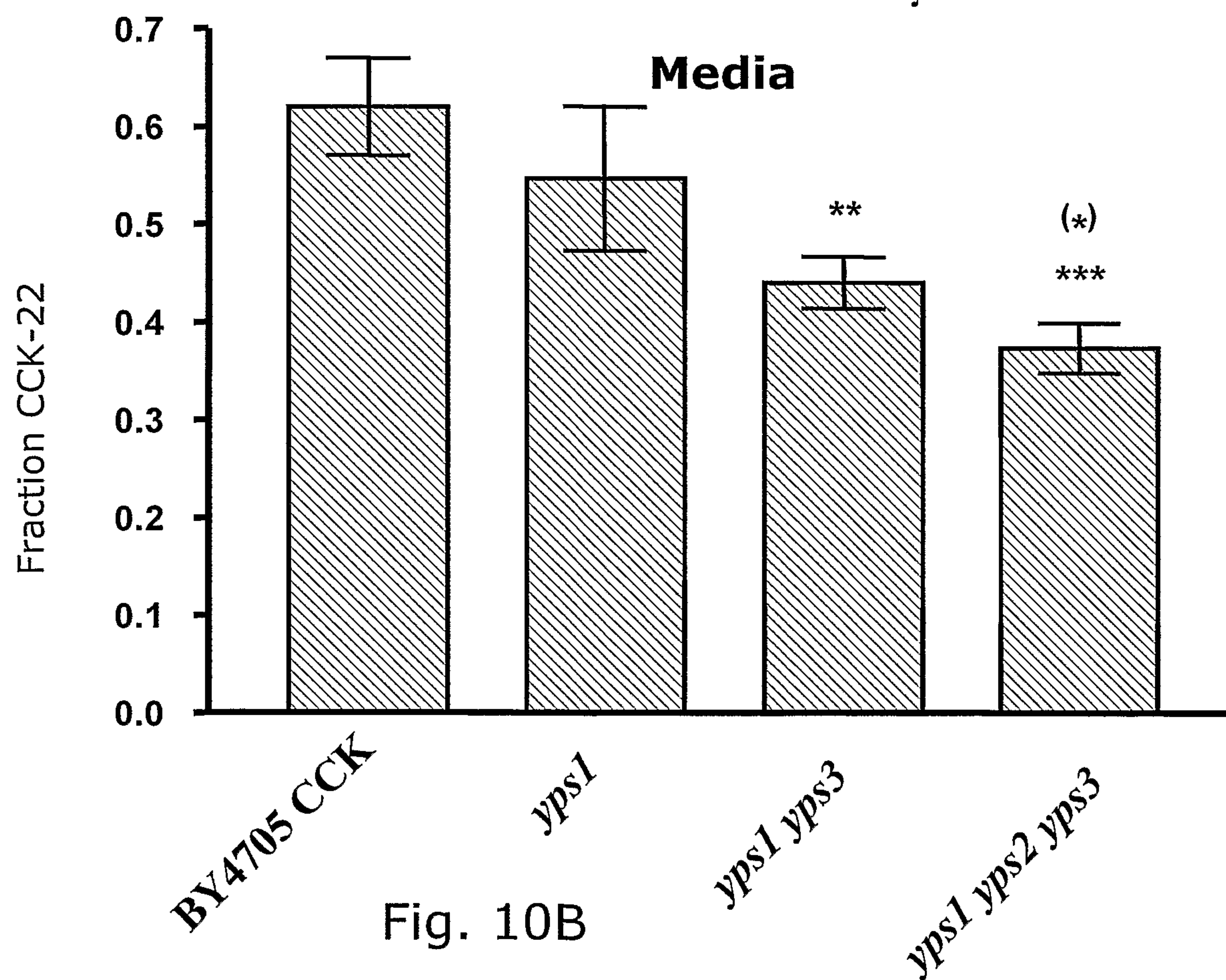
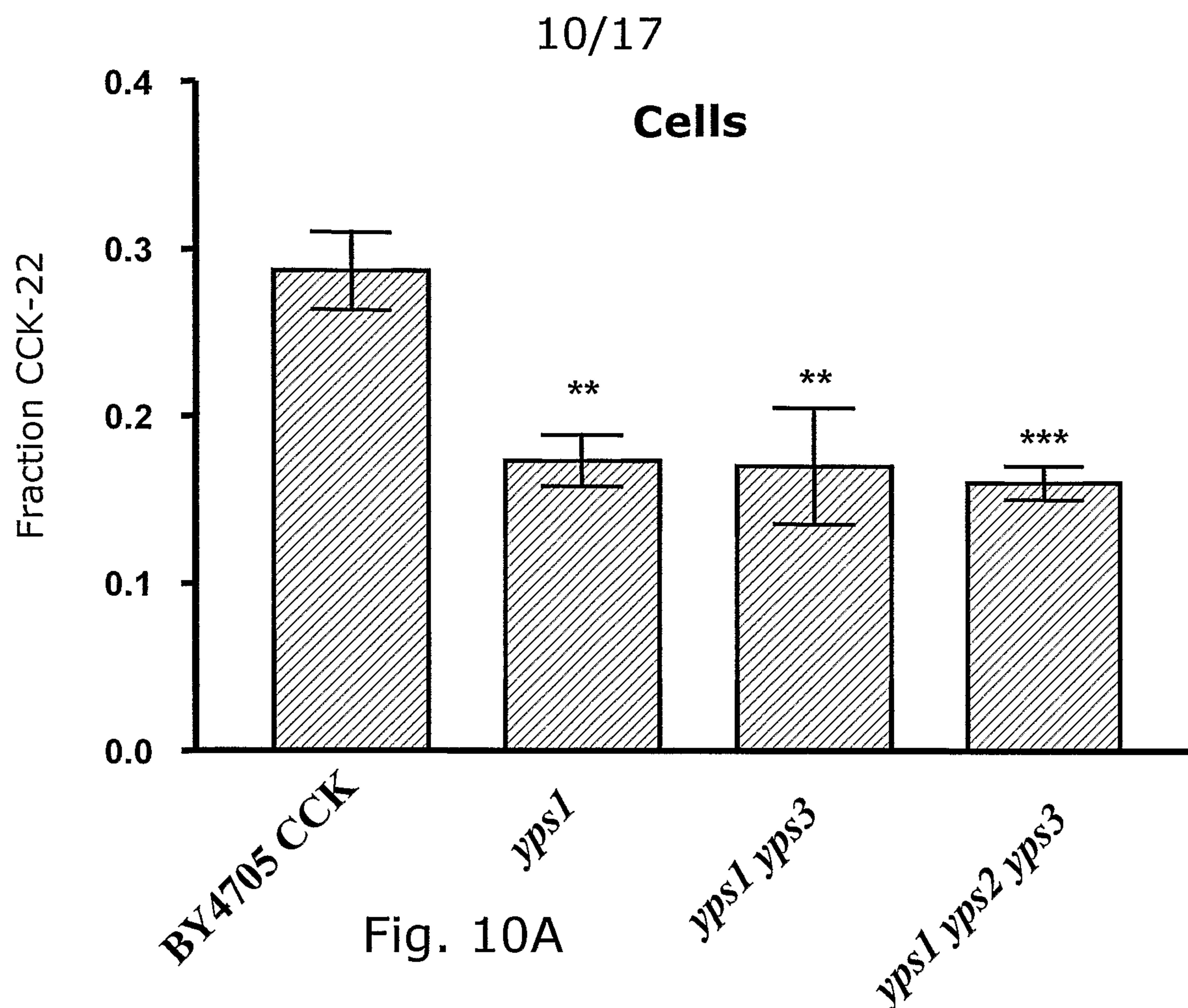
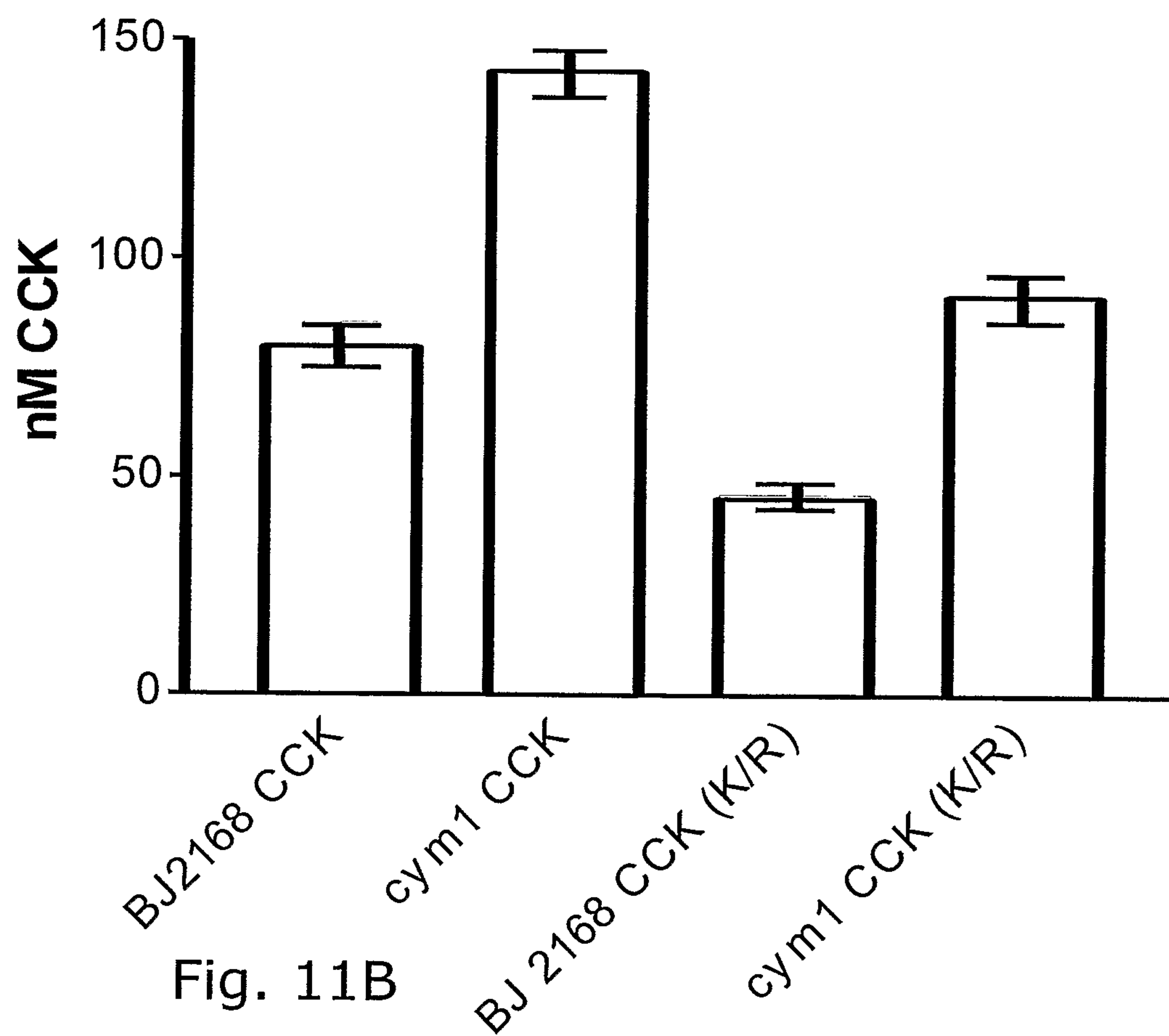
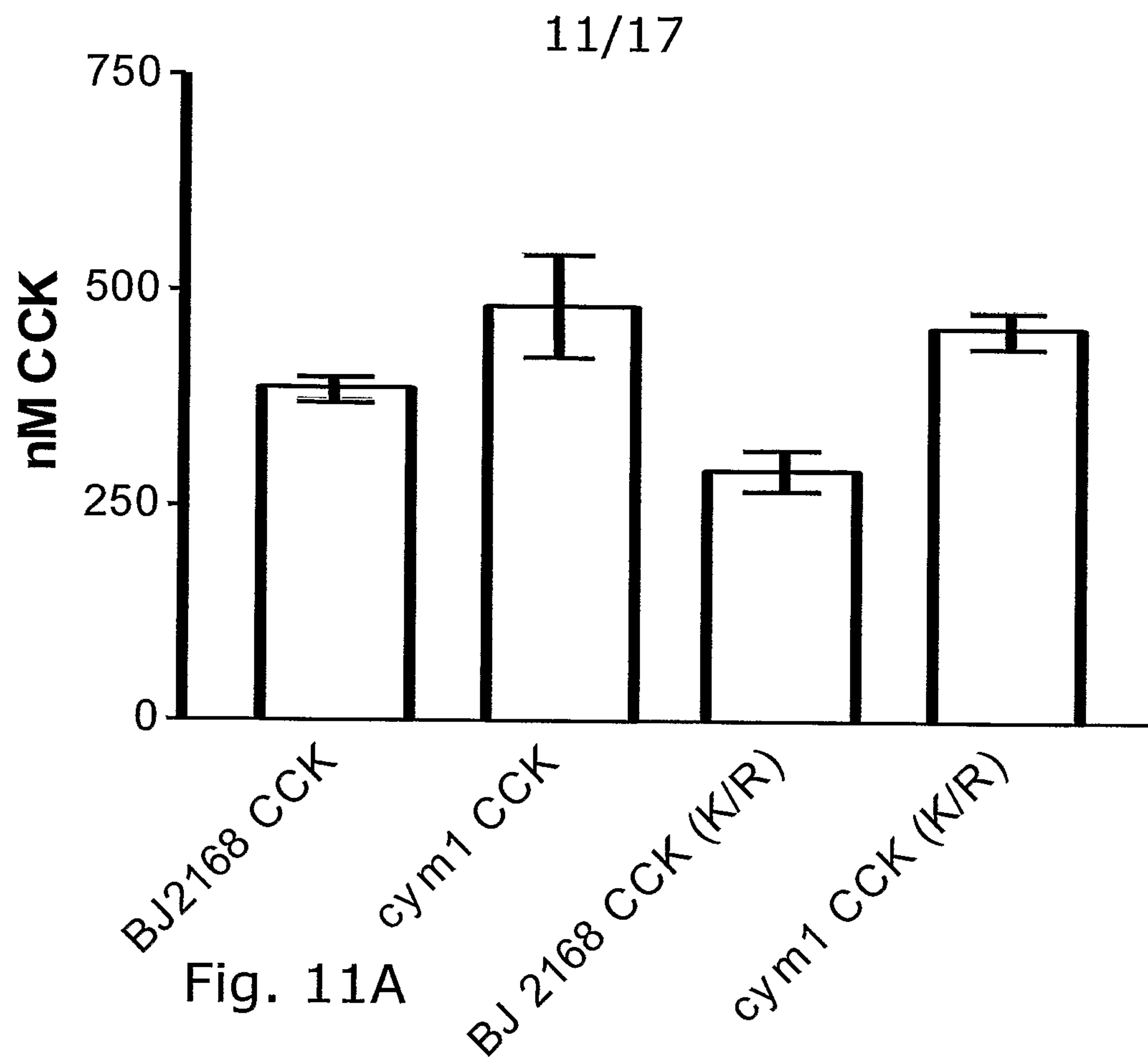


Fig. 9





ProCCK				<u>Molecular mass</u>		<u>Strain</u>	<u>SEQ ID</u>
							<u>No.</u>
PreproMfα1p // .SGLQRAEEAPRRQLRVSQRTDGESRAHLGALLARYIQQARKAPSGRMSIVKNLQNLDPSHRISDRDYMGWMDFGRRSAEEYEYPS	CCK-61	CCK-58	CCK-39	CCK-22	CCK-8		43
	QLRVSQRTDGESRAHLGALLAR						44
	VSQRTDGESRAHLGALLAR						45
	YIQQARKAPSGRMSIVKNLQNLDPSHRISDRDYMGWMDFGRRSAEEYEYPS						46
	YIQQARKAPSGRMSIVK						47
	YIQQARKAPSGRMSIV						48
	NLQNLDPSHRISD						49
	NLQNLDPSHRISDRDYMGWMDFG						50
	NLQNLDPSHRISDRDYMGWMDFGRRSAEEYEYPS						51
	DYMGWMDFGRRSAEEYEYPS						52

Fig. 12

13/17

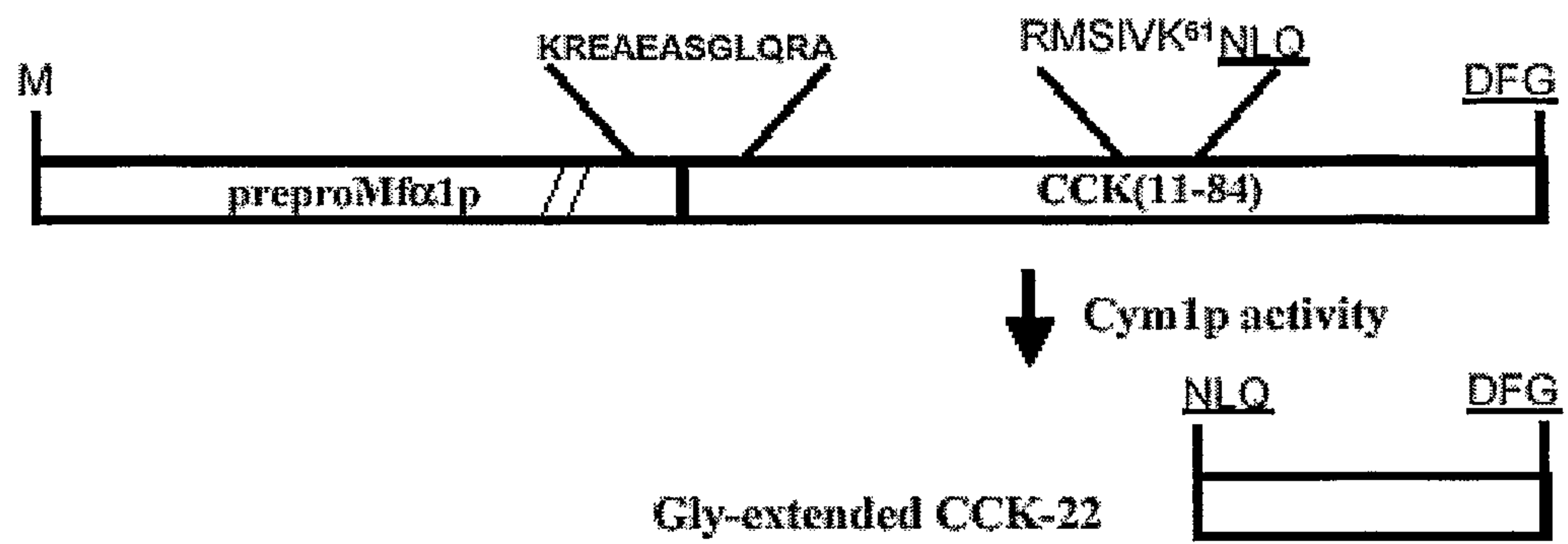


Fig. 13A

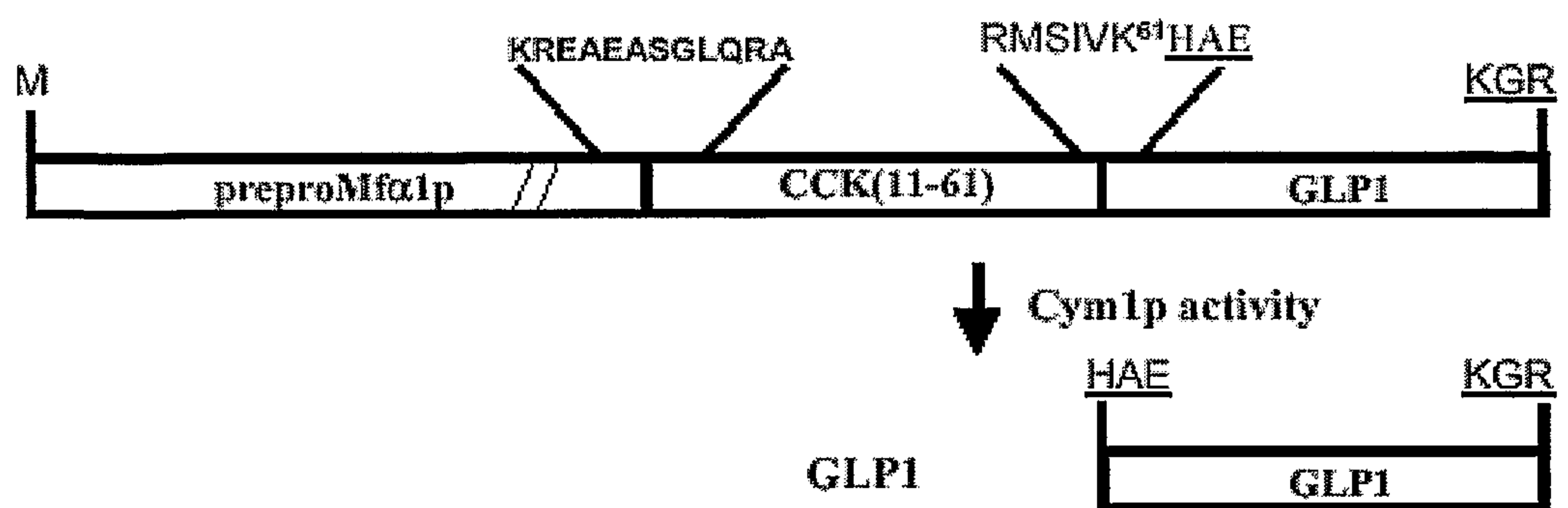


Fig. 13B

14/17

ATGAGATTTCTTCAATTTTACTGCAGTTTATTCGCAGCATCCTCCGCATTAGCTGCT
1 -----+-----+-----+-----+-----+-----+ 60
TACTCTAAAGGAAGTTAAAAATGACGTCAAATAAGCGTCGTAGGAGGCGTAATCGACGA
M R F P S I F T A V L F A A S S A L A A -
CCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTCCGGCTGAAGCTGTCATCGGT
61 -----+-----+-----+-----+-----+-----+ 120
GGTCAGTTGTGATGTTGTCTTCTACTTTGCCGTGTTTAAGGCCGACTTCGACAGTAGCCA
P V N T T T E D E T A Q I P A E A V I G -
TACTTAGATTTAGAAAGGGGATTTTCGATGTTGCTGTTTTGCCATTTTCCAACAGCACAAAT
121 -----+-----+-----+-----+-----+-----+ 180
ATGAATCTAAATCTTCCCCTAAAGCTACAACGACAAAACGGTAAAAGGTTGTCGTGTTTA
Y L D L E G D F D V A V L P F S N S T N -
AACGGGTTATTGTTTATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTA
181 -----+-----+-----+-----+-----+-----+ 240
TTGCCCAATAACAAATATTTATGATGATAACGGTCGTAACGACGATTTCTTCTTCCCAT
N G L L F I N T T I A S I A A K E E G V -
TCTTTGGATAAAAAGAGAGGCTGAAGCTCACCCGCTGGGCAGCCCCGGTTCAGCCTCGGAC
241 -----+-----+-----+-----+-----+-----+ 300
AGAAACCTATTTTCTCTCCGACTTCGAGTGGGCGACCCGTCGGGGCCAAGTCGGAGCCTG
S L D K R E A E A H P L G S P G S A S D -
TTGGAAACGTCCGGGTTACAGGAGCAGCGCAACCATTTGCAGGGCAAACCTGTCGGAGCTG
301 -----+-----+-----+-----+-----+-----+ 360
AACCTTTGCAGGCCCAATGTCCTCGTCGCGTTGGTAAACGTCCCGTTTGACAGCCTCGAC
L E T S G L Q E Q R N H L Q G K L S E L -
CAGGTGGAGCAGACATCCCTGGAGCCCCTCCAGGAGAGCCCCCGTCCCACAGGTGTCTGG
361 -----+-----+-----+-----+-----+-----+ 420
GTCCACCTCGTCTGTAGGGACCTCGGGGAGGTCCTCTCGGGGGCAGGGTGTCCACAGACC
Q V E Q T S L E P L Q E S P R P T G V W -
AAGTCCCGGGAGGTAGCCACCGAGGGCATCCGTGGGCACCGCAAATGGTCCTCTACACC
421 -----+-----+-----+-----+-----+-----+ 480
TTCAGGGCCCTCCATCGGTGGCTCCCGTAGGCACCCGTGGCGTTTTACCAGGAGATGTGG
K S R E V A T E G I R G H R K M V L Y T -
CTGCGGGCACCACGAAGCCCCAAGATGGTGCAAGGGTCTGGCTGCTTTGGGAGGAAGATG
481 -----+-----+-----+-----+-----+-----+ 540
GACGCCCGTGGTGCTTCGGGGTTCTACCACGTTCCCAGACCGACGAAACCCTCCTTCTAC
L R A P R S P K M V Q G S G C F G R K M -
GACCGGATCAGCTCCTCCAGTGGCCTGGGCTGCAAAGTGCTGAGGCGGCATTAA
541 -----+-----+-----+-----+-----+----- 594
CTGGCCTAGTCGAGGAGGTACCGGACCCGACGTTTCACGACTCCGCCGTAATT
D R I S S S S G L G C K V L R R H * -

Fig. 14A

15/17

ATGAGATTTCTTCAATTTTACTGCAGTTTATTTCGCAGCATCCTCCGCATTAGCTGCT
1 -----+-----+-----+-----+-----+-----+ 60
TACTCTAAAGGAAGTTAAAAATGACGTCAAATAAGCGTCGTAGGAGGCGTAATCGACGA
M R F P S I F T A V L F A A S S A L A A -
CCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTCCGGCTGAAGCTGTCATCGGT
61 -----+-----+-----+-----+-----+-----+ 120
GGTCAGTTGTGATGTTGTCTTCTACTTTGCCGTGTTTAAGGCCGACTTCGACAGTAGCCA
P V N T T T E D E T A Q I P A E A V I G -
TACTTAGATTTAGAAGGGGATTTTCGATGTTGCTGTTTTGCCATTTTCCAACAGCACAAAT
121 -----+-----+-----+-----+-----+-----+ 180
ATGAATCTAAATCTTCCCCTAAAGCTACAACGACAAAACGGTAAAAGGTTGTCGTGTTTA
Y L D L E G D F D V A V L P F S N S T N -
AACGGGTTATTGTTTATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTA
181 -----+-----+-----+-----+-----+-----+ 240
TTGCCCAATAACAAATATTTATGATGATAACGGTCGTAACGACGATTTCTTCTTCCCAT
N G L L F I N T T I A S I A A K E E G V -
TCTTTGGATAAAAGAGAGGCTGAAGCTAGCCCCAAGATGGTGCAAGGGTCTGGCTGCTTT
241 -----+-----+-----+-----+-----+-----+ 300
AGAAACCTATTTTCTCTCCGACTTCGATCGGGGTTCTACCACGTTCCCAGACCGACGAAA
S L D K R E A E A S P K M V Q G S G C F -
GGGAGGAAGATGGACCGGATCAGCTCCTCCAGTGGCCTGGGCTGCAAAGTGCTGAGGCGG
301 -----+-----+-----+-----+-----+-----+ 360
CCCTCCTTCTACCTGGCCTAGTCGAGGAGGTCACCGGACCCGACGTTTCACGACTCCGCC
G R K M D R I S S S S G L G C K V L R R -
CATTAA
361 ----- 366
GTAATT
H * -

Fig. 14B

16/17

1 ATGAGATTCCTTCAATTTTACTGCAGTTTATTTCGCAGCATCCTCCGCATTAGCTGCT
-----+-----+-----+-----+-----+-----+ 60
TACTCTAAAGGAAGTTAAAAATGACGTCAAATAAGCGTCGTAGGAGGCGTAATCGACGA
M R F P S I F T A V L F A A S S A L A A -
61 CCAGTCAACACTACAACAGAAGATGAAACGGCACAATTCGGGCTGAAGCTGTCATCGGT
-----+-----+-----+-----+-----+-----+ 120
GGTCAGTTGTGATGTTGTCTTCTACTTTGCCGTGTTTAAGGCCGACTTCGACAGTAGCCA
P V N T T T E D E T A Q I P A E A V I G -
121 TACTTAGATTTAGAAGGGGATTTTCGATGTTGCTGTTTTGCCATTTTCCAACAGCACAAAT
-----+-----+-----+-----+-----+-----+ 180
ATGAATCTAAATCTTCCCCTAAAGCTACAACGACAAAACGGTAAAAGGTTGTCGTGTTTA
Y L D L E G D F D V A V L P F S N S T N -
181 AACGGGTTATTGTTTATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTA
-----+-----+-----+-----+-----+-----+ 240
TTGCCCAATAACAAATATTTATGATGATAACGGTCGTAACGACGATTTCTTCTTCCCAT
N G L L F I N T T I A S I A A K E E G V -
241 TCTTTGGATAAAAGAAGCCCCAAGATGGTGCAAGGGTCTGGCTGCTTTGGGAGGAAGATG
-----+-----+-----+-----+-----+-----+ 300
AGAAACCTATTTTCTTCGGGGTTCTACCACGTTCCCAGACCGACGAAACCCTCCTTCTAC
S L D K R S P K M V Q G S G C F G R K M -
301 GACCGGATCAGCTCCTCCAGTGGCCTGGGCTGCAAAGTGCTGAGGCGGCATTAA
-----+-----+-----+-----+-----+----- 354
CTGGCCTAGTCGAGGAGGTCACCGGACCCGACGTTTCACGACTCCGCCGTAATT
D R I S S S S G L G C K V L R R H * -

Fig. 14C

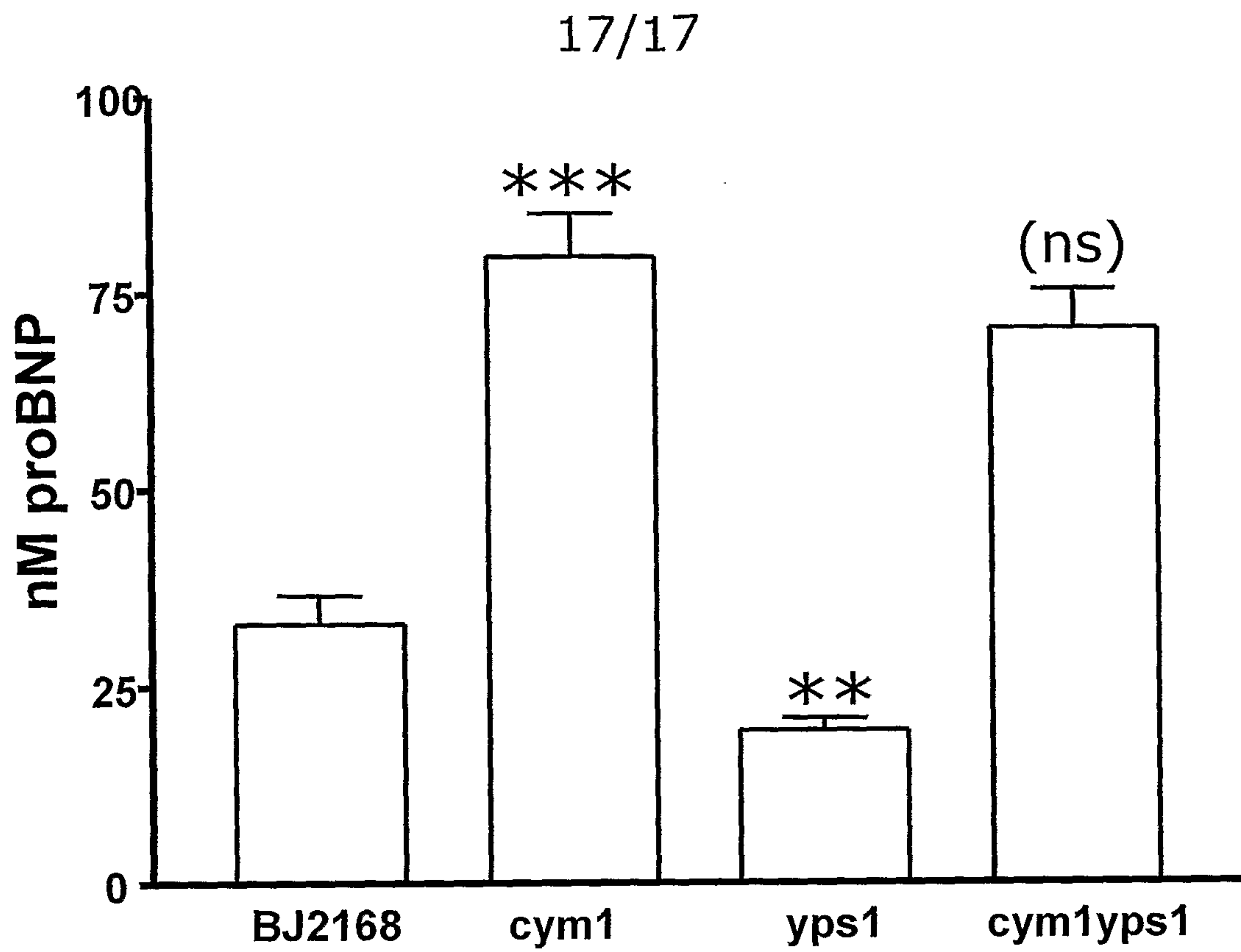


Fig. 15A

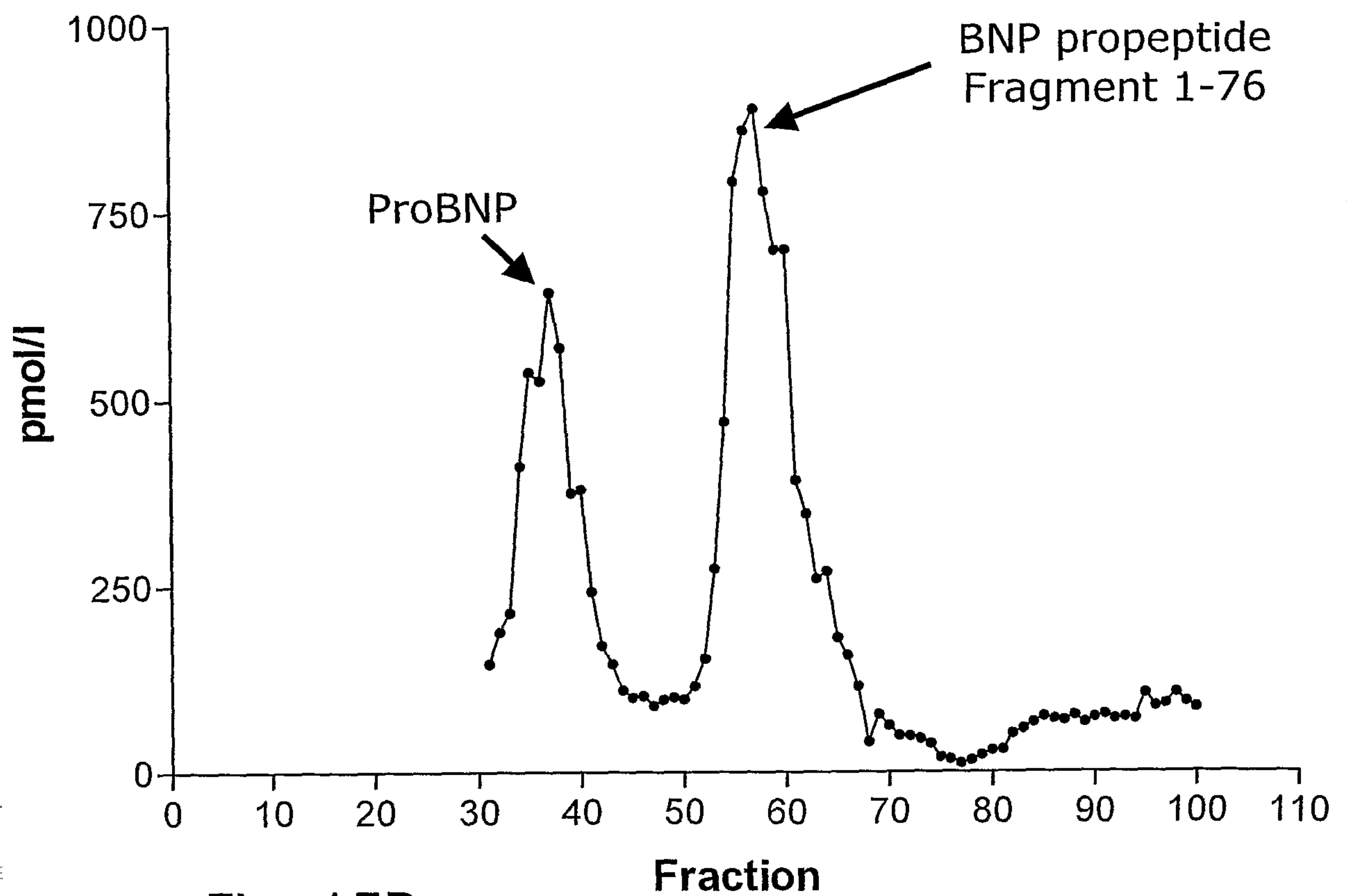


Fig. 15B

