Abstract:

Agent: E. BLUM & CO.; Vorderberg 11, CH-8044 Zurich (CH).

Inventors/Applicants (for US only): COTTIER, Valérie [CH/CH]; Bollenhofstrasse 3, CH-8953 Dietikon (CH).

LUTHI, Urs [CH/CH]; Hotzenwegstrasse 43, CH-8330 Pfaffikon (CH). BARBERIS, Alcide [CH/CH]; Allenmoosstrasse 64, CH-8057 Zurich (CH).

Inventors (for all designated States except US): ON-CALIS AG [CH/CH]; Wagistrasse 21, CH-8952 Schlieren (CH).

The present invention concerns a tester protein for identifying and/or monitoring protease activity in a cellular assay suitable for high throughput screenings by growth selection, wherein the tester polypeptide is a non-regulatory protein carrying a protease cleavage sequence. Upon co-expression of the protease recognizing said cleavage sequence the tester protein is inactivated, which influences the growth and/or survival of the host cells under the chosen conditions. However, in the presence of protease inhibitor the growth phenotype is reversed. The system can be used to identify proteases, protease inhibitors, and protease cleavage sites.
Genetic selection system to identify proteases, protease substrates and protease inhibitors

Technical Field

The present invention relates to a non-regulatory tester protein comprising a protease cleavage site, a nucleic acid encoding said tester protein and a cell expressing said tester protein; the invention also relates to the use of said tester protein in an assay for identifying and monitoring the activity of cellular proteases, for selecting inhibitors of said proteases based on cell proliferation of a suitable tester strain, and for identifying protease cleavage sequences.

Background Art

Proteases are enzymes which catalyse the splitting of interior peptide bonds in a protein. Many proteases are extracellular for the purpose of the degradation of proteins to amino acids. Other proteases are used during protein targeting, in particular secretion, whereby polypeptide precursors are cleaved specifically to yield the mature forms. For example, a membrane-bound protein can be converted to a soluble form or an inactive precursor molecule can be activated by a functional protease. Such proteases can also be found in organellar compartments or are associated with membranes.

Besides the proteasome, which is a proteolytic enzyme complex that degrades cytosolic and nuclear proteins, there are specific cytosolic proteases which specifically process polypeptides, well known are the caspases that are activated during apoptosis.

Proteases are also essential for the replication cycle of many viruses. Retroviruses, picornaviruses and herpesviruses for example encode proteins that are
synthesised as polyprotein precursors and that are later proteolytically processed to mature viral proteins (Tong 2002). Proteases have also been shown to be physiologically important for bacterial pathogens and are thus implicated in infectious diseases.

Since proteases play a critical role in the regulation of many biological processes, failures in their functioning can lead to severe diseases. Therefore, in the last decades, the pharmaceutical industry has recognised the potential of proteases as targets for drug development. Treatments against cancer, inflammatory, respiratory, cardiovascular and neurodegenerative diseases are being developed on the basis of protease inhibition (Luthi 2002). To cure hypertension a panel of angiotensin-converting enzyme (ACE) inhibitors have been identified by rational drug design and are nowadays widely prescribed (Hilleman 2000). In the same way, as indeed several viruses depend on the proteolysis of primary polypeptide precursors for their replication, viral proteases are prime therapeutic targets for the treatment of viral diseases, as highlighted by the success story of drugs against human immunodeficiency virus (HIV) (Chrusciel and Strohbach 2004; Randolph and DeGoey 2004).

Besides the HIV protease, many other viral proteases are targets for inhibitor screenings. The human cytomegalovirus (CMV), a member of the herpes virus family, is an opportunistic pathogen that can cause severe illness or death of immunocompromised individuals, such as AIDS patients or recipients of organ and bone marrow transplants (Holwerda 1997; Waxman and Darke 2000). Like the other herpes viruses, it encodes a protease that is essential for the production of infectious virus and that functions during the assembly and maturation of the capsid (Welch, Woods et al. 1991; Sheaffer, Newcomb et al. 2000; Gibson ; Trang, Kim et al. 2003). The protease itself is released from the 75 kDa precursor protein upon autoproteolytic cleavage at the maturational (M) and re-
lease (R) sites (Baum, Bebernitz et al. 1993). M-type cleavage removes the carboxy-terminal tail, whereas cleavage at the R-site releases the proteolytic domain, also called assemblin. The mature protease contains 256 amino acids, and its catalytic site is formed by the unusual triad His-Ser-His as opposed to classical serine proteases that function with the His-Ser-Asp/Glu triad (Chen, Tsuge et al. 1996; Shieh, Kurumbail et al. 1996). Remarkably, dimerisation is a prerequisite for enzymatic activity (Margosiak, Vanderpool et al. 1996) even though the two catalytic sites have been shown to act in an independent manner (Batra 2001). All herpesvirus protease enzymology and inhibition studies to date have been performed with the 28 kDa mature form (Pinko, Margosiak et al. 1995; Bonneau, Grand-Maitre et al. 1997; Hoog, Smith et al. 1997; Khayat, Batra et al. 2003) though the 75 kDa precursor has been demonstrated to be catalytically active as well (Lawler and Snyder 1999; Wittwer, Funckes-Shippy et al. 2002).

Besides herpesvirus proteases, other viral proteases such as Hepatitis C virus NS3 protease and rhinovirus 3C protease, both of which can be expressed as functional enzymes in yeast, are of interest. In addition, human soluble proteases like caspases, cathepsins (involved in different cancers: (Fehrenbacher and Jaattela 2005)), calpains (responsible for endothelial dysfunction and vascular inflammation: (Stalker, Gong et al. 2005)), or dipeptidyl peptidase IV (main cause of diabetes: (McIntosh, DeKuth et al. 2005) are targets for protease inhibitor screens.

Successful application of protease inhibitors in human therapy requires defined properties of drugs, such as membrane permeability, stability and lack of toxicity (Barberis 2002). Most high throughput screening (HTS) campaigns are performed with enzymatic in vitro assays, where compounds are tested exclusively with respect to their potential to inhibit proteolytic activity.
Cellular screening systems provide a promising alternative to screen or select directly for compounds with additional features that are essential for their use as drugs in a cellular context. Indeed, compounds are identified as hits at the condition that they not only inhibit proteolytic activity, but are also stable within the cell, capable of penetrating biological membranes, and exert no or only limited toxic effects on the cell.

Cell-based assays have notable advantages over in vitro assays. First, no purification of enzyme is required, avoiding a time consuming and costly process to obtain an active target. Second, target conformation and activity are examined in a cellular context, closer to natural physiological state than in an in vitro assay. Several cell-based assays have already been used to screen for protease inhibitors. Most of them rely on a reporter protein that allows a gradual read-out paralleling intracellular protease activity levels. Examples of such reporter proteins are GFP (green fluorescence protein; Lindsten, Uhlikova et al. 2001; Belkhiri, Lytvyn et al. 2002) or SEAP (secreted alkaline phosphatase; Lee, Shih et al. 2003; Mao, Lan et al. 2003; Oh, Kim et al. 2003). However, such systems have the disadvantage, that every toxic compound will also decrease the amount of reporter protein or signal in the medium, just by decreasing the number of cells producing it. By consequence, a high number of false positives will be obtained, which have to be further evaluated at costs of time and resources.

The yeast transcription factor Gal4p has been exploited in different detection systems for protease inhibitors due to its two-domain structural property by inserting the protease target site between the two domains. Protease activity separates the DNA-binding domain from the activation domain, causing stop of transcription of a Gal4p regulated reporter gene, e.g. lacZ. Protease in-
hibitors prevent cleavage and therefore inactivation of the Gal4p transcription factor, restoring transcriptional expression. Such systems have been developed for protease 3C from coxsackievirus (Dasmahapatra, DiDomenico et al. 1992) and for cytomegalovirus protease (Lawler and Snyder 1999). In a similar way the herpesvirus transcription factor VP16 was used in combination with a lacZ reporter gene to detect CMV protease activity. Other hybrid regulatory protein/reporter gene combinations have been used in various ways (US5721133; US2004042961; US6117639; US6699702).

Recently discovered protease inhibitors are among the more promising antiviral drugs; yet, there is still a need for more and alternative protease inhibitors, and thus for HTS systems enabling the rapid and efficient identification of new antiviral drugs. Whereas primarily mammalian or insect cells have been used in past screening campaigns (Johnston 2002; Kemnitzer, Drewe et al. 2004; Zuck, Murray et al. 2004), yeast cells provide an alternative model with several technical advantages. The fast and inexpensive cultivation, the easy genetic manipulation and the high degree of conservation of basic molecular mechanisms make this eukaryotic organism a valuable tool for drug screening (Botstein, Chervitz et al. 1997; Munder and Hinnen 1999; Brenner 2000; Hughes 2002). In addition, yeast provide a heterologous, yet eukaryotic environment, suitable for preventing redundant processes and for supplying a null background for the expression of several human targets. Of course, despite the high degree of similarity of basic cellular processes between yeast and human cells, yeast show some differences that might impair attempts to reproduce the activity of some target proteases. However, as long as the appropriate controls are respected, the employment of yeast in cell-based assays has many advantages, in particular for HTS.
Another improvement in the search of antiviral compounds would be to have a selection rather than a screening procedure, wherein only those cells survive that are exposed to an inhibitor. Such a selection system has been developed in yeast by using the Gal4p carrying a tobaccho etch virus (TEV) protease cleavage sequence between its two domains and measuring the lack of Gal4 regulatory function upon cleavage by the TEV protease as the lack of growth on the suicide substrate 2-deoxygalactose (Smith, T.A. and Kohorn, B.D., 1991). This system allows for the positive selection of inhibitors. However, the system has two further disadvantages: (i) it requires the addition of a toxic compound to the medium, and (ii) it uses a transcriptional regulatory protein, which only indirectly, i.e. by control of transcription of other genes leads to the desired phenotype, thus increasing the possibility to identify false positives.

A drug that inhibits a viral protease can be used to prevent production of new infectious viral particles. However, the efficacy of such drugs, when they are prescribed in monotherapy and especially in low dose therapy, is often limited by the rapid emergence of drug resistant strains. In the case of HIV, mutations at several key amino acid residues of the protease, which abolish protease inhibition by already marketed drugs, have been described. The occurrence of drug resistant strains is increasing, and the phenomenon of cross-resistance is gaining importance. Therefore, new drugs against such proteases, with different modes of action, are needed. Currently, most protease inhibitors are complex peptidomimetic compounds with poor aqueous solubility, low bioavailability and short plasma half-lives. The complexity of these agents not only contributes to their high cost but also increases the potential for unwanted drug interactions. There is a need for novel compounds working as protease inhibitors in the context of many widely spread diseases. In order to find such drugs, there is
also a need for biological systems, in particular selection rather than screening systems allowing by simple and reliable in vivo tests to select for protease inhibitors in high throughput screenings.

Disclosure of the Invention

Hence, it is a general object of the invention to provide a non-regulatory tester polypeptide for monitoring protease activity, which can be used in a protease inhibitor selection system and for the identification of proteases and protease cleavage sequences.

Now, in order to implement these and still further objects of the invention, which will become more readily apparent as the description proceeds, the tester polypeptide is manifested by the features that it
- comprises the sequence of a marker protein whose activity can be detected by positive and/or negative growth selection and an additional sequence, whereby said additional sequence is inserted at a specific permissible site in a surface loop of said marker protein and comprises a cognate cleavage sequence for a protease, and
- is inactivated upon cleavage by said protease.

The tester polypeptide of the present invention comprises a marker protein with a detectable activity, modified by an insertion of a cleavage sequence for a protease which creates an in frame fusion polypeptide that is still functional. Upon cleavage of the polypeptide by the matching protease the tester polypeptide is inactivated. The tescei polypeptide as well as the marker protein of the present invention are non-regulatory, i.e. are not transcriptional regulators of gene expression.

The marker protein of the present invention can either have a metabolic enzymatic activity or can be
a structural protein. If inactivation of said marker protein causes a deficiency of cellular growth, this allows a positive selection for the presence of said marker protein. This effect can depend on the growth conditions.

The marker protein can also be a negative selection marker that has an activity leading to growth inhibition. For example, it can be an enzymatic activity catalyzing the conversion of a non-toxic substrate into a toxic product. Cells comprising said activity die, whereas cells lacking said activity survive.

In a preferred embodiment of the present invention the marker protein is a cytoplasmic protein. Preferably, it is an enzyme of a biosynthetic pathway for an essential cellular compound, for example an amino acid, nucleotide, lipid or cofactor. More preferably, it is an enzyme of an amino acid biosynthesis pathway, such as the tryptophan biosynthesis pathway. Most preferably, the essential protein is the Trplp of yeast, which catalyses the isomerisation of N-5'-phosphoribosyl-anthranilate in the biosynthesis of the amino acid tryptophan that is essential for cell proliferation. Under tryptophan deficient growth conditions cells therefore can survive only if all of their enzymes involved in tryptophan biosynthesis, including Trplp or the Trplp derived tester protein, are functionally active.

In another preferred embodiment of the present invention the marker can be used for both positive and negative selection. This is possible, for example, for the preferred marker of the present invention, the Trplp protein encoded by the TRPl gene. This enzyme is required for the conversion of anthranilic acid to tryptophan, and thus is a typical auxotrophy marker allowing positive selection. The antimetabolite 5-fluoroanthranilic acid (FAA) was found to be particularly effective for counter selection, as it is converted in the presence of Trplp to the toxic 5-fluorotryptophan. Therefore,
the Trplp marker can also be used for negative selection (Toyne et al. 2000).

Alternatively, the yeast URA3 gene product orotidine 5'-decarboxylase required for uracil biosynthesis can be used for positive as well as for negative selection. In a positive selection, cells are grown on media lacking uracil, which allows growth of only those cells that express a functional enzyme. A negative selection can be performed on media containing 5-fluoroorotic acid (5-FOA), because the URA3 gene product converts 5-FOA to a toxic compound. Therefore, cells expressing a functional enzyme cannot grow.

Another example is the yeast Gal4p protein galactokinase, which converts galactose to galactose-1-phosphate. This intermediate is converted by the GAL7-encoded transferase into glucose-1-phosphate, which is metabolized. The Gal4p protein is thus essential for growth of yeast with galactose as the only carbon source, allowing positive selection. In addition, in yeast cells lacking the transferase enzyme encoded by the GAL7 gene expression of GAL1 leads to accumulation of the intermediate galactose-1-phosphate to toxic levels, thus allowing a negative selection (Gunde et al. 2004).

Another preferred marker for negative selection is the CYH2 gene, encoding the ribosomal protein Rpl28. Yeast cells carrying a mutation in their endogenous CYH2 allele are resistant to the antibiotic cycloheximide, whereas cells expressing wild-type CYH2 are sensitive.

In another preferred embodiment of the present invention the protease cleavage sequence has a size of 5-39 amino acids. For inactivation of the tester polypeptide, the protease cleavage sequence and the corresponding protease recognising and cleaving said sequence must be present in the system together in the same cellular compartment. It is possible that a protease requires a minimal cleavage sequence of only a few amino acids, or
even only a single amino acid, like for example the dipeptidyl peptidase IV, which is a post-proline cleaving enzyme. However, it is also possible that in the context of another polypeptide, into which a protease cleavage sequence is inserted, a longer extension of said cleavage sequence can be cleaved more efficiently. In some cases the minimal cleavage sequence may not be known and therefore just any number of amino acids encompassing the cleavage site within a natural target polypeptide of a specific protease may be chosen.

In a more preferred embodiment the protease cleavage site has a sequence selected from the group of cleavage sequences listed in Table 1. Most preferred are the cleavage sequences SEQ. ID. NO:1 = GVVNASCRLAGG, its longer version SEQ. ID. NO:2 = PTALLSGGAKVAERAGVVNASCRLATASGSEAATAGP, SEQ. ID. NO: 3 = KVAERANAGVVQASCRLATAS, which are all recognised by human cytomegalus virus (CMV) protease. Table 1 summarises a number of known proteases and their cognate target cleavage sequences that are in the scope of the present invention; however, these pairs are only examples and in no way exclusive or anyhow limiting.

Table 1. Proteases and their cognate cleavage sequences

<table>
<thead>
<tr>
<th>Proteases</th>
<th>Cleavage sequence and site of cleavage (↓)</th>
<th>SEQ.ID.NO:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Herpes virus proteases:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CMV (human cytomegalovirus)</td>
<td>GVVNAS↓SCRLA</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>GVVNAS↓SCRLLAGG</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>KVAERANAGVVQ↓SCRLATAS</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>PTALLSGGAKVAERAGVVNAS↓S</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>CRLATASGSEAATAGP</td>
<td></td>
</tr>
<tr>
<td></td>
<td>VXA↓S; LXA↓S; IXA↓S</td>
<td>5; 6; 7</td>
</tr>
<tr>
<td>HSV-1 (herpes simplex virus type 1)</td>
<td>ALVNA↓SSAAHV</td>
<td>8</td>
</tr>
<tr>
<td>Virus</td>
<td>Peptide Sequence</td>
<td>Reference</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>VZV (varicella zoster virus)</td>
<td>QDVNA↓VEASS</td>
<td>9</td>
</tr>
<tr>
<td>EBV (Epstein-Barr virus)</td>
<td>KLVQA↓SASGVA</td>
<td>10</td>
</tr>
<tr>
<td>HHV-6 (human herpes virus 6)</td>
<td>PSILNA↓S</td>
<td>11</td>
</tr>
<tr>
<td>Other virus proteases:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HIV-1</td>
<td>SFNF↓PQIT; TLNF↓PISP</td>
<td>12;13</td>
</tr>
<tr>
<td>Hepatitis C (NS3/4A)</td>
<td>DLEVVT↓STWVL</td>
<td>14</td>
</tr>
<tr>
<td>Coxsackievirus 3C</td>
<td>GTTLEALFQ↓GPPV</td>
<td>15</td>
</tr>
<tr>
<td>Rhinovirus 3C</td>
<td>LEVLFQ↓GPLG</td>
<td>16</td>
</tr>
<tr>
<td>SARS coronavirus 3C-like proteinase</td>
<td>SAVLQ↓SGF</td>
<td>17</td>
</tr>
<tr>
<td>Caspases:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caspase-1 (ICE)</td>
<td>WFKD↓S; FEDD↓A; YVHD↓A;</td>
<td>18;19;20</td>
</tr>
<tr>
<td>Caspase-2</td>
<td>DGPD↓G; DEVD↓G</td>
<td>21;22</td>
</tr>
<tr>
<td>Caspase-3</td>
<td>DEVD↓G</td>
<td>22</td>
</tr>
<tr>
<td>Caspase-6</td>
<td>IETD↓S; DGPD↓G; DEVD↓G</td>
<td>23;21;22</td>
</tr>
<tr>
<td>Caspase-7</td>
<td>DEVD↓N; DQQD↓N; DEPD↓S</td>
<td>24;25;26</td>
</tr>
<tr>
<td>Caspase-8 (FLICE)</td>
<td>DEAD↓G; DETD↓S; DACD↓T</td>
<td>27;28;29</td>
</tr>
<tr>
<td>Caspase-9</td>
<td>DGPD↓G; DEVD↓G; VEID↓N</td>
<td>21;22;30;</td>
</tr>
<tr>
<td>Other proteases:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plasmspsin</td>
<td>ERMF↓LSFP</td>
<td>33</td>
</tr>
<tr>
<td>Thrombin</td>
<td>VPR↓SFR</td>
<td>34</td>
</tr>
<tr>
<td>ACE (Angiotensin I-converting enzyme)</td>
<td>RPPGFSFP↓FR</td>
<td>35</td>
</tr>
<tr>
<td>Cathepsin S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cathepsin K</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMP2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMP7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMP13</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GPLG↓VRGL</td>
<td>36</td>
</tr>
</tbody>
</table>
In another preferred embodiment of the present invention the protease cleavage sequence is inserted into a surface loop of the essential protein such that it does not interfere with the function of the protein, as it does not significantly affect the folding of the essential protein. The candidate surface loops of an essential protein can either be known if the structure of said protein is known, or they can be predicted if the structure of a related protein is known. In addition, they can be predicted from computer generated secondary structure predictions and hydrophobicity analysis based on the polypeptide sequence. Often ideal insertion sites are at glycine or proline residues in sequence stretches that connect alpha helices and/or beta sheets and that are hydrophilic. Once a cleavage sequence of a known protease is inserted into a putative permissible surface loop of an essential protein, the activity of the resulting tester polypeptide is compared to the activity of the corresponding unmodified essential protein by measuring cell proliferation. A permissible site of the fusion will allow cell growth under the relevant conditions when the tester polypeptide is expressed, whereas a non-permissible site will lead to lack of cell growth. In a final step of validation it has to be tested whether the protease is able to recognise and cleave the fusion protein comprising the cleavage site. Hence, in the presence of the corresponding protease the protein should be cleaved inside the cell, which can be investigated for example by Western blot analysis or by cell growth selection. This is the case for the example of the yeast TrpLP, which tolerates the insertion of a protease cleav-
age sequence after amino acid Gly194, said sequence being recognised and cleaved by its cognate protease, thus leading to cell death. Hence, the use of the insertion site after Gly194 of the yeast Trplp protein for insertion of a protease cleavage sequence is a preferred embodiment of the present invention.

It is also comprised by the present invention that single or multiple point mutations within the essential protein and/or within the protease cleavage sequence of the present invention are used to improve the system. For example, the insertion of a cleavage sequence may have some impact on the folding and/or activity of the essential protein, which might be compensated by additional mutation(s). Any mutations can be introduced as long as the function of the tester protein in cell proliferation and the susceptibility of the cleavage sequence to the protease are not disturbed. Therefore, one or more point mutations, which can also be insertion or deletion mutations, fulfilling these requirements are envisaged in a further preferred embodiment of the present invention. Most preferred are one or more point mutations in the form of altered amino acids within the natural cognate cleavage sequence of a given protease.

In another preferred embodiment the inserted sequence is the target sequence of a viral protease. Most preferably said viral protease is the human CMV protease.

In another preferred embodiment of the present invention the inserted sequence encodes an autoprotease and comprises the cleavage sequence for said autoprotease. An autoprotease is a protein that cleaves at least one site of its own sequence in a self-processing manner. Many viral precursor proteins comprise autoprotease activities. Lhd1 lead to processed produces of the precursor molecule. The preferred autoprotease of the present invention is the autoprotease 3C from coxsackievirus.
Also a subject of the present invention is a nucleic acid encoding the tester polypeptide of the present invention. Preferably, said nucleic acid is a DNA. Said DNA comprises the gene with or without a promoter for expression of said tester polypeptide.

In a preferred embodiment of the present invention said DNA is part of a recombinant vector comprising transcriptional start and termination signals in order to allow expression of said tester protein. If said promoter is a regulated promoter, it is possible to optimise expression of said tester protein in order to optimise the ratio of tester protein to protease. Regulated promoters are well known to the person skilled in the art. The use of a regulated promoter depends on the cellular system in which the tester polypeptide is expressed. For example, if a bacterial cell is used, a lac or tac promoter may be used that is inducible by addition of isopropyl-β-D-thiogalactopyranoside (IPTG), or the ara promoter that is induced by the addition of arabinose and repressed by the addition of glucose. If a yeast cell is used, a suitable regulated promoter may be the galactose inducible GAL1 promoter, the copper inducible CUP1 promoter, the PHO5 promoter inducible by phosphate starvation, the HSP70 (heat shock) promoter inducible by increase of temperature, MET promoters inducible by methionine, or the CYCl promoter that is induced by oxygen and repressed by glucose. This list of promoters is by far not complete and many other known promoters can be used as well within the scope of the present invention.

It is also possible that the DNA of the present invention is integrated into the host chromosome. In this case, the promoter must be comprised by said DNA, or it must be provided by the host DNA flanking the site of said integration.

The present invention also provides a prokaryotic or eukaryotic cell comprising the nucleic acid of the present invention and a protease. Said nucleic
acid is transformed into said cell and either propagated as an extra-chromosomal element, or integrated into the chromosome of said cell. Expression of a protease in said cell is driven by a promoter that can be constitutive or regulated. In a preferred embodiment of the present invention an inducible promoter will be used, which allows to control the amount of synthesised protease for adaptation to the amount of tester polypeptide produced by said cell. The protease is encoded on an expression plasmid that is transformed into said cell. Alternatively, a protease naturally expressed in said cell is used.

The cloning of genes coding for a tester protein or a protease is done by gene synthesis and routine techniques including PCR known to the skilled person using known sequences of said proteins.

A further aspect of the present invention is the identification of a protease inhibitor by a method comprising the steps of

- providing a cell of the present invention comprising a tester protein with a protease cleavage sequence and comprising a matching protease,
- exposing said cell to candidate inhibitor substances,
- growing said cell under conditions that are non-permissive for cell proliferation in the presence of a functional protease, but permissive for cell proliferation in the additional presence of an inhibitor of said protease, and
- selecting an inhibitor on the basis of cell proliferation.

Candidate inhibitor molecules can be members of known chemical compound libraries, molecules from a random peptide library or natural products isolated from microorganisms, fungi, plants or animals, from water, soil or any natural environment where these organisms
live. Preferably, these molecules are able to penetrate the cell wall and reach the cytosol, where they can block the protease or mask the protease cleavage site on the tester protein. Alternatively, derivatives of known protease inhibitor molecules can be tested. Preferably, the method is based on yeast cells. More preferably a yeast mutant deficient in the multi drug export systems encoded by the genes pdr5, snq2, and yorl is used as a host.

In a preferred embodiment of the present invention, cells are exposed to putative inhibitory molecules before or at the time when they are shifted to conditions that are non-permissive for cell proliferation in the presence of a functional protease. This will eliminate candidate inhibitors which are per se toxic for the cell, i.e. which block other essential cellular functions. In another preferred embodiment of the present invention the protease is provided by expressing it under the control of a regulated promoter, for example the yeast Gall promoter. This allows to chose expression levels of the protease in accordance with the concentration of inhibitor. For example, low levels of protease expression can be used when weak inhibitors are preferred, whereas high levels of protease are useful to detect strong inhibitors. Moreover, this also allows to choose inhibitor concentrations in an non-toxic range.

The inhibitor selection system of the present invention comprises the possibility to manipulate the levels of tester protein as well as the levels of protease and can therefore be optimised in various ways. A further aspect of the present invention is the use of the inhibitor selection system in high throughput (HT) assays. The output signal of the assay, i.e. the turbidity of the cell culture can be measured directly in a single step in the microtiter plate by measurement of light absorption or light scattering without the use of special equipment or the need for additional chemicals and/or additional handling.
Another aspect of the present invention is to provide a method to identify a suitable site in a non-regulatory marker protein for insertion of a protease cleavage sequence, said marker protein being suitable for positive as well as negative growth selection. In said method the protease is modulated on the one side at the level of its presence or absence or at the level of its expression or at the level of its activity, and on the other side a positive as well as a negative selection step are used in a successive given order. This leads to several alternative embodiments of the present invention:

4) if an inhibitor of said protease is available, the method comprises the steps of

- identifying putative surface loops in the marker protein,
- providing an expression vector comprising a nucleic acid encoding said marker protein,
- inserting a nucleic acid comprising a coding sequence of said protease cleavage sequence at a random position within the coding sequence of said putative surface loops, resulting in a plasmid comprising a gene encoding a candidate tester protein as defined above,
- transforming with said plasmid a yeast cell comprising a protease that is capable of cleaving said protease cleavage sequence,
- growing transformants in the presence of a specific inhibitor of said protease under conditions requiring a function of said tester protein (positive selection),
- shifting growing clones to conditions non-permissive for a function of said tester protein and lacking an inhibitor (negative selection),
- determining the nucleic acid sequence of
  the gene encoding said tester protein of a
  surviving clone.

Transformants are cells that have stably
taken up DNA during transformation. If not otherwise men-
tioned, plasmids used for transformations in the scope of
the present invention carry a selectable marker, and
transformants can be obtained under corresponding selec-
tive conditions.

b) In the absence of a known inhibitor of
said protease, the identification of a suitable insertion
site in a non-regulatory marker protein as defined above
can be achieved by a method comprising the steps of
- identifying putative surface loops in said
  marker protein,
- providing an expression vector comprising a
  nucleic acid encoding said marker protein,
- inserting a nucleic acid comprising a cod-
ing sequence for said protease cleavage se-
quence at a random position within the cod-
ing sequence of said putative surface
loops, resulting in a plasmid comprising a
gene encoding a candidate tester protein,
- transforming with said plasmid a yeast cell
  comprising a gene encoding a protease that
  is capable of cleaving said protease cleav-
age sequence, said gene being under the
  control of a tightly regulated promoter,
- growing transformants under repressing or
  non-inducing conditions with respect to
  said promoter and under conditions requir-
ing the function of the tester protein
  (positive selection),
- shifting growing cells to derepressing or
  inducing conditions with respect to said
  promoter for protease expression and to
  non-permissive conditions with respect to a
function of said tester protein (negative selection),
- determining the nucleic acid sequence of the gene encoding said tester protein of a growing cell.

In an alternative embodiment of the present invention, instead of a single cell with an inducible promoter for protease expression two cells are used for the selection, and in this case the method comprises either the steps of c)
- identifying putative surface loops in said marker protein,
- providing an expression vector comprising a nucleic acid encoding said marker protein,
- inserting a nucleic acid comprising a coding sequence for said protease cleavage sequence at a random position within the coding sequence of anyone of said putative surface loops, resulting in a plasmid comprising a gene encoding a candidate tester protein,
- providing a first yeast cell comprising a protease capable of cleaving said cleavage sequence and a second yeast cell lacking said protease,
- transforming said first yeast cell with said plasmid and growing transformants under non-permissive conditions with respect to a function of said tester protein (negative selection),
- isolating said plasmid from a surviving cell,
- transforming said second yeast cell with said isolated plasmid and growing transformants under conditions requiring a function of said tester protein (positive selection),
- determining the nucleic acid sequence of
  said gene encoding said tester protein of a
growing cell,
or it comprises the steps of d)
- identifying putative surface loops in said
  marker protein,
- providing an expression vector comprising a
  nucleic acid encoding said marker protein,
- inserting a nucleic acid comprising a cod-
ing sequence for said protease cleavage se-
quence at a random position within the cod-
ing sequence of anyone of said putative
  surface loops, resulting in a plasmid com-
  prising a gene encoding a candidate tester
  protein,
- providing a first yeast cell comprising a
  protease capable of cleaving said cleavage
  sequence and a second yeast cell lacking
  said protease,
- transforming said second yeast cell with
  said plasmid and growing transformants un-
der conditions requiring a function of said
tester protein (positive selection),
- isolating said plasmid from a growing cell,
- transforming said first cell with said iso-
  lated plasmid and growing transformants un-
der conditions non-permissive for a func-
tion of said tester protein (negative se-
lection),
- determining the nucleic acid sequence of
  said gene encoding said tester protein of a
  surviving cell.

Yet another alternative is the use of a sin-
gle cell lacking said protease and applying a positive
selection followed by the introduction of an expression
plasmid encoding said protease into the growing cell and
then applying a negative selection, i.e. a method comprising the steps of e):
- identifying putative surface loops in said marker protein,
- providing an expression vector comprising a nucleic acid encoding said marker protein,
- inserting a nucleic acid comprising a coding sequence for said protease cleavage sequence at a random position within the coding sequence of anyone of said putative surface loops, resulting in a plasmid comprising a gene encoding a candidate tester protein,
- providing a yeast cell lacking a protease capable of cleaving said cleavage sequence,
- transforming said yeast cell with said plasmid and selecting for growth under conditions requiring a function of said tester protein (positive selection), obtaining transformants,
- providing a second plasmid capable of expressing a gene encoding said protease,
- transforming said transformants with said second plasmid and selecting for growth under conditions non-permissive for a function of said tester protein (negative selection),
- determining the nucleic acid sequence of said gene encoding said tester protein of a surviving cell.

In a preferred embodiment of the present invention the marker protein is a single domain protein. However, multi domain proteins may also be used. In this case, a suitable surface loop can also be within the sequence connecting two domains.

In a similar way, if an inhibitor of the protease is known, it is also possible according to the pre-
sent invention to identify the cleavage sequence of a known protease by a method comprising the steps of a)
- providing an expression vector encoding a non-regulatory marker protein suitable for positive as well as negative selection with at least one known permissible site in a surface loop for the insertion of a sequence,
- inserting a coding sequence for about 5-39 amino acids into said site, resulting in a plasmid comprising a gene encoding a tester protein,
- transforming with said plasmid a suitable host cell comprising said protease,
- growing transformants in the presence of a specific inhibitor of said protease under conditions requiring a function of said tester protein,
- shifting growing clones to conditions non-permissive for a function of said tester protein and lacking said inhibitor,
- determining the nucleic acid sequence of the gene encoding said tester protein of a surviving clone.

However, in the absence of an inhibitor of said protease the cleavage site of said protease can be determined by one of the following four variations of the method, namely a method comprising the steps of b)
- providing an expression vector encoding a non-regulatory marker protein suitable for positive as well as negative selection with at least one known permissible site in a surface loop for the insertion of a sequence,
- inserting a coding sequence for about 5-39 amino acids into said site, resulting in a
plasmid comprising a gene encoding a tester protein,
- transforming with said plasmid a suitable host cell comprising the gene encoding said protease under a control of a tightly regulated promoter,
- growing transformants under repressing or non-inducing conditions with respect to said promoter and under conditions requiring a function of said tester protein (positive selection),
- shifting growing cells to derepressing or inducing conditions with respect to said promoter and non-permissive conditions with respect to a function of said tester protein (negative selection),
- determining the nucleic acid sequence of the gene encoding said tester protein of a surviving cell,
or a method comprising the steps of c)
- providing an expression vector encoding a non-regulatory marker protein suitable for positive as well as negative selection with at least one known permissible site in a surface loop for the insertion of a sequence,
- inserting a coding sequence for about 5-39 amino acids into said site, resulting in a plasmid comprising a gene encoding a tester protein,
- providing a first yeast cell comprising a protease capable of cleaving said cleavage sequence and a second yeast cell lacking said protease,
- transforming said first yeast cell with-said plasmid and growing transformants under non-permissive conditions with respect
to a function of said tester protein (negative selection),
- isolating said plasmid from a surviving cell,
- transforming said second cell with said isolated plasmid and growing transformants under conditions requiring a function of said tester protein (positive selection),
- determining the nucleic acid sequence of the gene encoding said tester protein of a growing cell,
or a method comprising the steps of a)
- providing an expression vector encoding a non-regulatory marker protein suitable for positive as well as negative selection with at least one known permissible site in a surface loop for the insertion of a sequence,
- inserting a coding sequence for about 5-39 amino acids into said site, resulting in a plasmid comprising a gene encoding a tester protein,
- providing a first yeast cell comprising a protease capable of cleaving said cleavage sequence and a second yeast cell lacking said protease,
- transforming said second yeast cell with said plasmid and growing transformants under conditions requiring a function of said tester protein (positive selection),
- isolating said plasmid from a growing cell,
- transforming said first yeast cell with said isolated plasmid and growing transformants under non-permissive conditions with respect to a function of said tester protein (negative selection),
- determining the nucleic acid sequence of a said gene encoding a said tester protein of a surviving cell,

or a method comprising the steps of e) providing an expression vector encoding a non-regulatory marker protein suitable for positive as well as negative selection with at least one known permissible site in a surface loop for the insertion of a sequence,

- inserting a coding sequence for about 5-39 amino acids into said site, resulting in a plasmid comprising a gene encoding a tester protein,

- providing a yeast cell lacking a protease capable of cleaving said cleavage sequence,

- transforming said yeast cell with said plasmid and selecting for growth under conditions requiring a function of said tester protein (positive selection), obtaining transformants,

- providing a second plasmid capable of expressing a gene encoding said protease,

- transforming said transformants with said second plasmid and selecting for growth under conditions non-permissive with respect to a function of said tester protein (negative selection),

- determining the nucleic acid sequence of said gene encoding said tester protein of a surviving cell.

A variation of the method to determine the cleavage site of a protease is possible if the non-regulatory marker protein is only used for positive selection. In this case, after the first step of positive selection for cells expressing a functional tester polypeptide the transformants are picked and each split into
two identical cell populations, of which one is transformed subsequently with the second plasmid expressing a gene encoding said protease, and the other one is transformed with the empty vector, i.e. the vector not comprising the gene encoding said protease. The growth of the two transformed populations is then compared under positive selection conditions, and those clones are of interest which do not grow in the presence, but do grow in the absence of said protease. However, this method is preferably only used if few clones are investigated, as it involves more handling than the method using the negative selection. This may be the case if there is already some preliminary information on the protease cleavage sequence but better knowledge is desired. For example, the validation of specific point mutations in a known cleavage sequence may be done with the positive selection method.

Another aspect of the present invention is to provide a method to identify new proteases for a known protease cleavage sequence, said method comprising the steps of:

- providing cells expressing a functional, non-regulatory tester polypeptide suitable for negative selection,
- providing an expression library comprising putative genes encoding said protease,
- transforming said cells with said expression library,
- growing transformants under non-permissive conditions with respect to a function of said tester protein (negative selection),
- identifying among surviving clones those which lack full-length tester polypeptide,
- determining from identified clones the nucleic acid sequence of the gene encoding said protease.
Preferably, said expression library expresses proteins from the same organism and/or tissue from which the cleavage sequence has been obtained. Most preferred is a human cDNA library.

As the present invention provides a system comprising a tester protein with a protease cleavage sequence on the one hand and a protease on the other hand as outlined above, this system can be further adapted to specific uses such as the engineering of improved proteases or changing the specificity of a protease. For example, a protease A with specificity for a cleavage sequence B can be co-expressed in a cell with a tester protein comprising a protease cleavage sequence C according to the present invention, and the gene encoding the protease A can be subjected to random or site specific mutagenesis to select for clones that change the protease A such that it can recognize and cleave the cleavage site C. This is possible because the system of the present invention is based on selection, in particular if the tester protein is a genetic marker that allows positive as well as negative selection.

Brief Description of the Drawings

The invention will be better understood and objects other than those set forth above will become apparent when consideration is given to the following detailed description thereof. Such description makes reference to the annexed drawings, wherein:

Fig. IA depicts a structural model of N-(S'-phosphoribosyl)-anthranilate (Trplp), a yeast protein essential for cell proliferation, showing the predicted alpha helices, beta sheets and the intervening surface loop regions. Fig. IB shows a Kyte Doolittle hydropathy plot wherein the tested sites of insertions are indicated.

Fig. 2 shows a spotting assay (Fig. 2A) for evaluating the functionality of Trplp tester proteins
comprising an inserted protease cleavage sequence and the quantified results (Fig. 2B).

Fig. 3 shows the quality of human CMV protease cleavage of a Trplp tester protein. In Fig. 3A the inserted cleavage sequences a) = wild-type 13-mer (TRP1(194)-M (short); Seq. Id. Nc:2), b) = wild-type 39-mer (TRP1(194)-M (long); Seq. Id. No:4) and c) mutant 39-mer (TRP1(194)-M (Ala->Glu) (long); Seq. Id. No. 39) are shown. Fig 3B shows the quantification of cleavage in an experiment using human CMV (HCMV) protease and Trplp tester polypeptides comprising the different cleavage sequences. Fig. 3C shows biochemical evidence for cleavage of the substrate by human CMV protease in a Western blot experiment following the disappearance of the full-length substrate TRP1(194)-M.

Fig. 4 shows the gradual, reciprocal correlation between human CMV protease expression level and cell growth measured as a result of the protease assay.

Fig. 5 illustrates the validation of the TRP1(194)-M system with known cellular protease inhibitors. Fig. 5A shows the application of the protease inhibitors BI31 and BI36 in the CMV protease inhibitor selection system. Fig. 5B shows the inhibition of cleavage of the Trp1(194)-M (long) substrate by CMV protease in a Western blot.

Fig. 6 shows the growth inhibition of TRP1(194)-2C/3A transformed RLYO7 cells by coxsackievirus 3C protease that inactivates the Trplp tester protein substrate. A comparison of active versus inactive CVB3 3C protease is shown.

**Modes for Carrying Out the Invention**

In the following a cell-based system is described, which enables monitoring protease activity and, in addition, selecting for inhibitors of given proteases.
In this assay, the protease cleavage sequence of interest is inserted into a protein essential for proliferation of yeast cells, the Trplp protein, yielding the tester protein. Co-expression of the protease with this engineered substrate reduces cell proliferation in selective medium, as it will be shown with the human cytomegalovirus (CMV) protease. In a proof-of-principle experiment, it was demonstrated that a small molecule CMV protease inhibitor prevents inactivation of the modified Trplp tester protein by blocking of said protease, thus stimulating cell proliferation.

Growth markers impose themselves as the best candidates for the choice of the essential protein for this system. Indeed, most laboratory strains are already deleted for growth markers, allowing for the application of such a system in almost any genetic background. Among these growth markers the N-(5'-phosphoribosyl)-anthranilate isomerase (Trplp) enzyme has been intensively studied (Eder and Kirschner 1992; Eder and Wilmanns 1992; Hommel, Eberhard et al. 1995; Hennig, Sterner et al. 1997), and its 3-dimensional structure from different organisms has been determined. Trplp is a small monomeric protein that catalyses the isomerisation of N-(5'-phosphoribosyl)-anthranilate in the biosynthesis of tryptophan, an essential amino acid for cell proliferation. Therefore, Trplp is essential for proliferation of yeast cells when tryptophan is not provided externally. Trplp was chosen as the essential protein, and was now modified to become the tester protein of choice such that it comprises a protease recognition and cleavage sequence at a permissible site, yet retains its function.

1. Material and methods

1.1. Yeast strains
The three major ABC transporter proteins Pdr5p, Snq2p and Yorlp were deleted in the *S. cerevisiae* JPY5 strain (MATα ura3-52 his3Δ200 leu2Δ1 trplΔ63 lys2Δ385) to generate the RLY07 strain (MATα ura3-52 his3Δ200 leu2Δ1 trplΔ63 lys2Δ385 pdr5Δ snq2Δ and yorlΔ). Fusion proteins used in this study were expressed in RLY07.

1.2. Recombinant plasmids

All TRPl-M constructs used in this study were subcloned in the CEN4-ARS1 plasmid pMH4 that contains a LEU2 auxotrophic marker and a polylinker with unique Xbal and Sail restriction sites. Expression of the subcloned TRPl-M constructs is under control of the ADH1 promoter and the GAL1 terminator. The TRPl gene was amplified by PCR from the YCplac22 plasmid (Gietz and Sugino 1988).

The human CMV protease cleavage sequence GVVVNA 4SCRLAGG (derived from the M-site), flanked by Ncol at the 5'-end and Notl at the 3'-end, is inserted after amino acids 49, 102, 132, 165 and 194 of Trplp. The longer human CMV cleavage sequence (39 amino acids surrounding the M-site) was obtained by PCR amplification of the UL80 gene and subcloned in the previously described TRPl194-M plasmid via Ncol and Notl restriction sites. The 3C cleavage sequence of coxsackievirus B3 (GTTLEALFQ4GPPV), which is located at the junction of the viral proteins 2C and 3A, was subcloned in TRPl after amino acid Gly194. HA tags have been added at both the N- and C-terminus of the TRPl (194)-M construct for Western blot analysis. The CMV protease gene encoding amino acids 1-256 of the 75 kDa precursor was obtained by PCR from CMV infected MRC5 human cells and subcloned via unique Xbal and Notl restriction sites in pMH51. pMH51 is an CEN4-ARS1 plasmid that contains a URA3 marker and a full-length GAL1 promoter (100%). For the experiment described in figure 4, the CMV protease gene was subcloned on a plasmid series, which contain modified GAL1 promoters that express the protease
with 71%, 46% and 16% protein production relative to the original full-length (100%) GAiI promoter. To subclone the 3C protease gene from the coxsackievirus strain B3, RNA was isolated from an infected HeLa cell culture with FastRNA® Kit-Red from BIO 101. A reverse transcription reaction was performed and 3C encoding DNA fragment was amplified by PCR and cloned via XbaI and NotI sites in a CEN4-ARS1 plasmid that carries a URA3 marker and controls expression by the GAL1 promoter and the GAL11 terminator.

Clonings were done using standard molecular biology techniques (Sambrook & Russell, 3rd ed. 2001, Molecular Cloning, A Laboratory Manual).

1.3. Yeast media and transformation

All media were prepared according to Burke et al. (Burke, Dawson et al. 2000). Transformation of yeast cells was performed following the lithium acetate method (Gietz, St Jean et al. 1992).

1.4. Spotting assay

RLY07 cells transformed with the different TRP1-M constructs were inoculated in 3 ml of 2% -leu glucose medium and grown overnight at 30°C to saturation. Next morning, cells were diluted in the same medium to OD600 0.25 and grown to OD600 1. Cultures were then washed with 5 ml H2O, resuspended in 2% -leu -trp glucose medium and diluted to 10⁶ cells/ml. 10 μl of serially diluted cultures were spotted on non-selective (-leu) and selective (-leu -trp) 2% glucose plates and incubated during 3 days at 30°C.

1.5. Liquid growth assays

RLY07 cells transformed with the different TRP1-M constructs and a plasmid encoding the CMV protease or the empty vector were inoculated in 3 ml of 2% galactose -ura -leu medium and grown at 30°C to OD600 1. They were washed with 5 ml H2O and resuspended in 2% galactose
-ura -leu -trp medium supplemented with 10% glycerol (-growth selection medium, glycerol promotes CMV protease dimerisation and subsequent proteolytic activity) and diluted to a start OD600 0.01. For the experiments with the coxackievirus 3C protease, preculture medium was 2% glucose -ura -leu, assay medium was 2% glucose -ura -leu -trp, and inoculation OD600 0.001. At time zero, assay cultures at the aforesaid start OD600 were distributed in 96-well microtiter plates, with a volume of 150 µl per well, and incubated without shaking at 30°C. At the time points indicated in the "results" section, plates were shaken to resuspend cells before being submitted to light scattering measurement at 595 nm in a Tecan Genios reader for determining cell density.

CMV protease inhibitors BI31 and BI36 (Boeringer Ingelheim, Quebec) were dissolved in DMSO and added to the assay cultures at time zero of the growth assay. Final DMSO concentration was 1%.

1.6. Western blot analysis

Yeast whole cell extracts were prepared as described by Burke et al. (Burke, Dawson et al. 2000). Proteins were separated by SDS-PAGE and Western blot analysis was performed according to standard procedures (Ausubel et al., 2003, Current Protocols in Molecular Biology). An HA-monoclonal antibody from Sigma (clone 3F10), was used at a concentration of 30 ng/ml to detect expression of TRP1 (194)-M.

2. Results

2.1. Insertion of the CMV protease cleavage sequence at 5 different locations in Trplp

Three conditions are critical for appropriate functioning of the above described system: i) The inserted cleavage sequence does not affect enzymatic properties of the Trpl protein, ii) The cleavage sequence is
cleaved by the protease. iii) Cleavage must result in functional inactivation of the Trp enzyme. Indeed, cleavage might occur without separating the two fragments generated and then without impairing the enzymatic function.

The Trp1 enzyme is a member of the prominent class of proteins that fold into a $(\beta/\alpha)_B$-barrel, which is the most commonly occurring fold among enzymes. The core of $\beta/\alpha$ barrel proteins consists of an eight-stranded parallel $\beta$-barrel held together by an extensive $\beta$-sheet hydrogen-bonding network. The individual $\beta$-strands are usually followed by $\alpha$-helices that form an outer ring surrounding the cylindrical surface of the central $\beta$-barrel (Eder and Wilmanns 1992) (Figure 1A). The $s$. cerevisiae Trplp structure has not yet been determined, but amino acid sequence alignments with the N- (5'-phosphoribosyl) -anthranilate isomerase from $E$. coli (ePRAI) and Thermotoga maritima (tPRAI) provide us with a reliable model. $s$. cerevisiae Trplp shares 28% identical amino acids with $E$. coli and 33% with $T$. maritima Trplp.

Alignment and modeling for $s$. cerevisiae Trplp was performed with the SWISS-MODEL protein modelling server (Guex and Peitsch 1997; Schwede, Kopp et al. 2003).

To determine suitable sites for insertion of the cleavage sequence, several constructs were designed. Since turn sequences are in general highly mutable in $(\beta/\alpha)_B$-barrels, 5 insertion sites were chosen that are located in such turns between an $\alpha$-helix and a $\beta$-sheet, after amino acids Asp (49), Asp (102), Ala (132), GIy (165) and GIy (194) (Figure 1). Respective constructs will therefore be referred to as Trpl (49) -M, Trpl (102) -M, Trpl (132) -M, Trpl (165) -M, and Trpl (194) -M. In addition, 4 out of the 5 sites are, according to Kyte-Doolittle, situated in hydrophilic regions, increasing the probability of being located at the periphery of the protein, thereby increasing the probability of the protease to access those sites. The inserted sequence consists of 13 amino acids derived from the M-site (Figure 3A, a).
site has previously been used in a viral protease assay based on Gal4p inactivation in mammalian cells (Lawler and Snyder 1999). In that assay, increasing amounts of expressed CMV protease caused a gradual reduction of reporter gene expression.

In order to evaluate functionality of the Trpl(49)-M, Trpl(102)-M, Trpl(132)-M, Trpl(165)-M and Trpl(194)-M fusion proteins, a spotting assay was performed. Tryptophan auxotrophic RLY07 cells were transformed with wild-type Trplp (positive control), empty vector (negative control), Trpl(49)-M, Trpl(102)-M, Trpl(132)-M, Trpl(165)-M and Trpl(194)-M, and serial dilutions were spotted on selective medium lacking tryptophan and incubated for 3 days at 30°C. For Trpl(132)-M, Trpl(165)-M and Trpl(194)-M expressing cells, growth was indistinguishable from cells expressing wild-type Trplp, indicating that M-site insertion did not interfere with functionality of the enzyme "(Figure 2A, lanes 1, 5, 6, 7)". This is opposed to Trpl(49)-M and Trpl(102)-M constructs that produced non-functional enzymes, as demonstrated by likewise transformed cells unable to grow on selective plate (Figure 2A, lanes 3, 4).

2.2. Site-specific cleavage of Trpl194-M by the CMV protease

Next it was investigated whether the 3 functional Trpl-M proteins were cleaved and inactivated by the CMV protease: Trpl(132)-M, Trpl(165)-M and Trpl(194)-M were co-expressed with the CMV protease in the RLY07 strain and cell proliferation was assayed by measuring OD$_{600}$ of the respective transformed cells cultured in liquid selective medium. After 36 hours, Trpl(194)M expressing cells exhibited an OD$_{600}$ reduction of 35% compared to control cells that contained an empty plasmid instead of the protease-expressing plasmid (Figure 2B, lane 4). We conclude that cleavage of the Trpl(194)-M substrate between helix $\alpha$7 and strand $\beta$8 reduces activity of the Trpl
enzyme. Importantly, this region is situated between two neighbouring loops (loops between β7/α7 and β8/α8) that have been shown to be important for binding of the substrate phosphate ion (Wilmanns, Hyde et al. 1991). A structure disruption in this region is most likely detrimental to phosphate binding of the anthranilate substrate. As opposed to Trpl(194)-M cells, Trpl(132)-M and Trpl(165)-M expressing cells did not show growth reduction despite the fact that CMV protease was expressed and active in those cells (Figure 2B, lanes 2,3). Thus, the latter 2 engineered Trpl substrates were either not cleaved or, alternatively, they were cleaved but the separated fragments still form an active enzyme.

To improve cleavage frequency at the M-site of the Trpl (194) -M substrate, the 13 amino acid target sequence was replaced by a longer sequence consisting of 39 amino acids (Fig 3A, b). Cells expressing the modified Trpl (194)-M together with the active protease showed 85% proliferation reduction (Figure 3B, lanes 3, 4) when grown in medium lacking tryptophan for 38 h as compared to the 35% proliferation reduction with the original, shorter cleavage site (Figure 3B, lanes 1, 2). This indicates that the extended recognition site is more efficiently cleaved by the CMV protease.

The CMV protease has been published to hydrolyse both the M-site and R-site between an alanine and a serine (Burck, Berg et al. 1994). To demonstrate site-specific cleavage of the Trpl (194) -M substrate at the M-site, the following experiment was performed: The alanine of the scissile bond was substituted with a glutamic acid (Fig 3A, c), a mutation known to prevent cleavage (Welch, McNally et al. 1993). As expected, proliferation of cells co-expressing the mutant Trpl (194) -M (A-E) with the protease was comparable to proliferation of cells expressing Trpl (194) -M alone, indicating that the CMV protease cleaves the Trpl (194) -M substrate in sequence-specific manner at the scissile bond (Figure 3B, lanes 5). Impor-
tantly, an inactive version of the CMV protease, harboring the S(132)A mutation at the catalytic site (Chen, Tsuge et al. 1996), was not able to cleave the Trpl (194)-M substrate (Figure 3B, lane 6).

To provide biochemical evidence for cleavage of the substrate by the CMV protease, an HA tag was cloned both to the N-terminus and C-terminus of Trpl(194)-M. The Trpl polypeptides were detected in protein extracts from cells transformed with plasmids encoding different Trpl-Mp substrates by Western blot analysis using an anti-HA antibody. The full-length substrate migrates at 33 kDa (Figure 3C, lane 1). Co-expression of active CMV protease (lane 2) causes disappearance of the full-length substrate. However, no cleaved fragments could be detected, probably due to either degradation or too low detection threshold. Indeed, since the Trpl (194)-M construct is expressed from a weak promoter (a 5' truncated version of the ADH promoter), the intracellular concentration of the fragments is most likely very low. Lane 3 provides biochemical evidence that the inactive CMV protease does not cleave the Trpl (194)-M substrate (since the full-length substrate band does not disappear), and lane 4, that active protease has no effect on the point-mutated Trpl (194)-M (A->E) substrate, as the band is also present. The use of the calmodulin antibody serves as an internal control for protein amounts.

Taken together, the above experiments show that the Trpl (194)-M substrate is cleaved in a sequence-specific manner by the CMV protease and that this cleavage results in a slow-growth phenotype.

2.3. A gradual increase of CMV protease expression level results in a gradual reduction of cell growth

The yeast-based system described in this report was developed to identify inhibitors of CMV protease
activity in HTS format. To validate sensitivity of the system to different intracellular CMV protease activity levels, the protease was cloned behind series of GAL1 promoters. Whereas CMV protease in the above experiments was expressed from the full-length (100%) GAiI promoter, we subcloned the protease on truncated GAL1 promoters, reaching 71%, 46% and 16% of the protein production as compared to the full length GAiI promoter. This plasmid series was co-expressed with the Trpl(194)-M substrate, and cell growth was measured after 36 hrs at OD690. As shown in Figure 4, a gradual increase of promoter strength, and thus of intracellular protease activity, is inversely proportional to cell proliferation. For example, a reduction of 29% of protease expression (from the 100% promoter to the 71% promoter) results in a 53% stimulation of cell proliferation. A reduction of 54% of protease expression caused a likewise stimulation of 138%. Therefore, even weak inhibitors causing only a partial reduction of CMV protease activity, should be detectable in the system.

2.4. Validated CMV protease inhibitors specifically stimulate cell growth in a HTS format

To further validate the Trpl(194)-M system we challenged it with 2 known human CMV protease inhibitors. Since yeast cells have evolved efficient mechanisms to pump out small chemical compounds, the three major ATP-binding cassette (ABC) transporters Snq2p, Pdr5p and Yorlp (Rogers, Decottignies et al. 2001) were deleted in the strain JPY5 to generate RLY07. It has been shown that deletion of these so-called drug efflux pumps increases sensitivity of yeast cells towards small molecules, allowing to perform screenings in yeast at lower concentrations.

CMV protease inhibitors BI31 (I) and BI36 (II) from Boehringer Ingelheim were applied to our selec-
Lactam derivatives have initially been published as inhibitors of classical serine proteases, such as human leukocyte elastase. Development of such scaffolds by rational design then delivered specific inhibitors of the CMV protease (Finke, Shah et al. 1995). Both compounds show IC50 values of ~1 µM in an enzymatic assay and inhibit viral replication in cell culture with EC50 values of ~80 µM (Yoakim, Ogilvie et al. 1998).

Both compounds are built on a β-lactam scaffold.

RLY07 cells co-expressing Trp194-M substrate and the CMV protease were incubated with a concentration series of BI31 and BI36 in 96-well microtiter plates and cultivated under selective conditions. After ~2 days incubation at 30°C, increasing concentrations of both BI31 and BI36 caused a dose-dependent increase of cell proliferation (Figure 5A, triangles). For BI36 an EC50 of 31 µM in the yeast assay was calculated, suggesting that the sensitivity of this assay is similar to the antiviral assay in cell culture (Yoakim, Ogilvie et al. 1998). At a concentration of 100 µM BI36, OD600 was close to OD600 of cells expressing the inactive protease (squares), meaning that the CMV protease was almost completely inhibited. It should be noted that increasing concentrations of BI31 in RLY07 cells expressing the inactive, point-mutated CMV protease (squares) causes a gradual decrease of cell pro-

![Chemical structures of BI31 and BI36](image-url)
liferation, indicating that BI31 exerts a dose-dependent toxic effect on the cells. Importantly, despite this toxicity BI31 still stimulates growth of cells expressing the active protease (triangles). For example, at 50 µM cell density is multiplied by a factor 4 despite 25% toxicity. This suggests that in a HTS screening compounds will be scored as positives even if they exert some intrinsic toxicity.

The Western blot was performed to provide biochemical evidence for inhibition of cleavage of CMV protease by compound BI36 (Fig. 5B). A 33 kDa band corresponds to the full-length Trpl (194)-M substrate upon co-expression with inactive CMV protease (lane 1). However, co-expressing the active protease instead of the inactive version causes disappearance of the 33 kDa band (lane 2), due to cleavage at the M-site. As in the Western blot on Figure 3c, unfortunately no cleavage products could be detected.

Application of a concentration series of BI36, 100 µM (lane 3), 30 µM (lane 4), and 10 µM (lane 5), prevented substrate cleavage in a dose-dependent manner. Whereas with 10 µM of BI36 proteolysis is only slightly inhibited, treatment with 100 µM of BI36 inhibits the cleavage almost completely, which is consistent with the determined EC$_{50}$ of 31 µM in the yeast-based assay.

2.5. The Trpl–M system can be applied for other intracellular proteases

In order to test whether the above described system can be adapted for other proteases, the 39 amino acid M-site in the Trpl(94)M substrate was substituted with a 13 amino acid sequence derived from the 2C/3A cleavage site of the cysteine protease 3C from coxsackievirus B3, resulting in the Trpl-2C/3A substrate. Coxsackievirus (CV) is an enterovirus from the Picornaviridae family. Its RNA genome encodes a single polyprotein of roughly 2200 amino acids that is processed by
the viral proteases 2A and 3C. Protease 3C, responsible for the majority of the cleavage events, plays a major role during the virus replication cycle. The Trp1-2C/3A substrate was co-expressed with the 3C protease in RLY07 cells, and cell proliferation was assessed in selective medium lacking tryptophan after 27 h at 30°C. Co-expression of active 3C protease reduced cell growth by 60% compared to cells expressing only the Trp1-2C/3A substrate (Figure 6), suggesting cleavage of the substrate by the protease. This experiment shows that the Trp1 selection concept can be applied to further proteases apart from the CMV protease.

While there are shown and described presently preferred embodiments of the invention, it is to be distinctly understood that the invention is not limited thereto but may be otherwise variously embodied and practised within the scope of the following claims.

References:


Bonneau, P. R., C. Grand-Maitre, et al. (1997). "Evidence of a conformational change in the human


Hennig, M., R. Sterner, et al. (1997). "Crystal structure at 2.0 Å resolution of phosphoribosylanthranilate isomerase from the hyperthermophile Thermotoga
maritima: possible determinants of protein stability."
Biochemistry 36(20): 6009-16.


cation of its gene, putative active site domain, and cleavage site." Proc Natl Acad Sci U S A 88: 10792-10796.


Claims

1. A non-regulatory tester polypeptide for monitoring protease activity, which
- comprises the sequence of a marker protein whose activity can be detected by positive and/or negative growth selection and an additional sequence, said additional sequence being inserted at a specific permissible site in a surface loop of said marker protein and comprising a cognate cleavage sequence for a protease, and
- is inactivated upon cleavage by said protease.

2. The polypeptide of claim 1 with the marker protein being a cytoplasmic protein.

3. The polypeptide of claim 1 or 2 with the marker protein being a biosynthetic enzyme for an essential cellular compound.

4. The polypeptide of anyone of the preceding claims with the marker protein being an auxotrophy marker for both positive and negative selection.

5. The polypeptide of anyone of the preceding claims with the marker protein being an enzyme of an amino acid biosynthesis pathway.

6. The polypeptide of anyone of the preceding claims with the marker protein being the yeast Trplp protein.

7. The polypeptide of claim 6 comprising a protease cleavage sequence inserted after Gly194 of Trplp.
8. The polypeptide of anyone of the preceding claims, characterised in that the cleavage sequence is between about 5-39 amino acids long.

9. The polypeptide of anyone of the preceding claims, characterised in that the protease cleavage sequence is selected from the group consisting of SEQ. ID. NO: 2 = GGVNASCRLLAGG, SEQ. ID. NO: 3 = KVAERANAGWQCASCRLATAS and SEQ. ID. NO: 4 = PTALLSSGKVAERAGVNASCRLLATASGSEAATAGP.

10. The polypeptide of anyone of the preceding claims, characterised in that it is susceptible to cleavage by a viral protease.

11. The polypeptide of claim 10 that is susceptible to CMV protease.

12. The polypeptide of anyone of claims 1-8 characterised in that the additional sequence comprising the cleavage sequence is the sequence of an autoprotease.

13. The polypeptide of anyone of claims 1-8 that is susceptible to coxsackievirus protease 3C.

14. The polypeptide of anyone of the preceding claims that is modified by one or more point mutations.

15. The polypeptide of claim 8 wherein the point mutations are within the natural, cognate cleavage sequence of a protease.

16. A nucleic acid encoding the polypeptide of anyone of the preceding claims.
17. A nucleic acid according to claim 16 comprising a promoter for expression of the tester polypeptide.

18. A recombinant vector comprising the nucleic acid of claim 16 or 17.

19. A prokaryotic or eukaryotic cell comprising the nucleic acid of claim 16 or 17 and a protease capable of cleaving the tester polypeptide within the cognate cleavage sequence for said protease.

20. A prokaryotic or eukaryotic cell comprising the vector of claim 18 and a protease capable of cleaving the tester polypeptide within the cognate cleavage sequence for said protease.

21. A cell according to claim 19 or 20, which is a yeast cell.

22. A method to identify a protease inhibitor comprising the steps of
   - providing a cell according to claims 19, 20 or 21,
   - exposing said cell to candidate inhibitor substances,
   - growing said cell under conditions that are non-permissive for cell proliferation in the presence of a functional protease, but permissive for cell proliferation in the additional presence of an inhibitor of said protease, and
   - selecting an inhibitor on the basis of cell proliferation.

23. A method to identify a suitable site in a non-regulatory marker protein for insertion of a protease
cleavage sequence, said marker protein being suitable for positive as well as negative selection, said method comprising the steps of

- identifying putative surface loops in said marker protein,
- providing an expression vector comprising a nucleic acid encoding said marker protein,
- inserting a nucleic acid comprising a coding sequence for said protease cleavage sequence at a random position within the coding sequence of said putative surface loops, resulting in a plasmid comprising a gene encoding a tester protein according to anyone of claims 1-7,
- transforming with said plasmid a yeast cell comprising a protease that is capable of cleaving said protease cleavage sequence,
- growing transformants in the presence of a specific inhibitor of said protease under conditions requiring a function of said tester protein,
- shifting growing clones to conditions non-permissive for a function of said tester protein and lacking said inhibitor,
- determining the nucleic acid sequence of the gene encoding said tester protein of a surviving clone.

24. A method to identify a suitable site in a non-regulatory marker protein for insertion of a protease cleavage sequence, said marker protein being suitable for positive as well as negative selection, said method comprising the steps of

- identifying putative surface loops in said marker protein,
- providing an expression vector comprising a nucleic acid encoding said marker protein,
- inserting a nucleic acid comprising a coding sequence for said protease cleavage sequence at a random position within the coding sequence of anyone of said putative surface loops, resulting in a plasmid comprising a gene encoding a tester protein according to anyone of claims 1-7,
- transforming with said plasmid a yeast cell comprising a gene encoding a protease that is capable of cleaving said protease cleavage sequence, said gene being under the control of a tightly regulated promoter,
- growing transformants under repressing or non-inducing conditions with respect to said promoter and under conditions requiring a function of said tester protein,
- shifting growing cells to derepressing or inducing conditions with respect to said promoter for protease expression and non-permissive conditions with respect to a function of said tester protein,
- determining the nucleic acid sequence of the gene encoding said tester protein of a growing cell.

25. A method to identify a suitable site in a non-regulatory marker protein for insertion of a protease cleavage sequence, said marker protein being suitable for positive as well as negative selection, said method comprising the steps of
- identifying putative surface loops in said marker protein,
- providing an expression vector comprising a nucleic acid encoding said marker protein,
- inserting a nucleic acid comprising a coding sequence for said protease cleavage sequence at a random position within the cod-
ing sequence of anyone of said putative surface loops, resulting in a plasmid comprising a gene encoding a tester protein according to anyone of claims 1-7,
- providing a first yeast cell comprising a protease capable of cleaving said cleavage sequence and a second yeast cell lacking said protease,
- transforming said first yeast cell with said plasmid and growing transformants under non-permissive conditions with respect to a function of said tester protein,
- isolating said plasmid from a surviving cell,
- transforming said second yeast cell with said isolated plasmid and growing transformants under conditions requiring a function of said tester protein,
- determining the nucleic acid sequence of said gene encoding said tester protein of a growing cell.

26. A method to identify a suitable site in a non-regulatory marker protein for insertion of a protease cleavage sequence, said marker protein being suitable for positive as well as negative selection, said method comprising the steps of
- identifying putative surface loops in said marker protein,
- providing an expression vector comprising a nucleic acid encoding said marker protein,
- inserting a nucleic acid comprising a coding sequence for said protease cleavage sequence at a random position within the coding sequence of anyone of said putative surface loops, resulting in a plasmid com-
prising a gene encoding a tester protein according to anyone of claims 1-7,
- providing a first yeast cell comprising a protease capable of cleaving said cleavage sequence and a second yeast cell lacking said protease,
- transforming said second yeast cell with said plasmid and growing transformants under conditions requiring a function of said tester protein,
- isolating said plasmid from a growing cell,
- transforming said first cell with said isolated plasmid and growing transformants under conditions non-permissive for a function of said tester protein,
- determining the nucleic acid sequence of said gene encoding said tester protein of a surviving cell.

27. A method to identify a suitable site in a non-regulatory marker protein for insertion of a protease cleavage sequence, said marker protein being suitable for positive as well as negative selection, said method comprising the steps of
- identifying putative surface loops in said marker protein,
- providing an expression vector comprising a nucleic acid encoding said marker protein,
- inserting a nucleic acid comprising a coding sequence for said protease cleavage sequence at a random position within the coding sequence of anyone of said putative surface loops, resulting in a plasmid comprising a gene encoding a tester protein according to anyone of claims 1-7,
- providing a yeast cell lacking a protease capable of cleaving said cleavage sequence,
- transforming said yeast cell with said plasmid and selecting for growth under conditions requiring a function of said tester protein, obtaining transformants,
- providing a second plasmid capable of expressing a gene encoding said protease,
- transforming said transformants with said second plasmid and selecting for growth under conditions non-permissive for a function of said tester protein,
- determining the nucleic acid sequence of said gene encoding said tester protein of a surviving cell.

28. A method to identify the cleavage site of a known protease comprising the steps of
- providing an expression vector encoding a non-regulatory marker protein suitable for positive as well as negative selection with at least one known permissible site in a surface loop for the insertion of a sequence,
- inserting a coding sequence for about 5-39 amino acids into said site, resulting in a plasmid encoding a tester protein according to anyone of claims 1-7,
- transforming with said plasmid a suitable host cell comprising said protease
- growing transformants in the presence of a specific inhibitor of said protease under conditions requiring a function of said tester protein,
- shifting growing clones to conditions non-permissive for a function of said tester protein and lacking said inhibitor,
- determining the nucleic acid sequence of the gene encoding said tester protein of a surviving clone.

29. A method to identify the cleavage site of a known protease comprising the steps of
- providing an expression vector encoding a non-regulatory marker protein suitable for positive as well as negative selection with at least one known permissible site in a surface loop for the insertion of a sequence,
- inserting a coding sequence for about 5-39 amino acids into said site, resulting in a plasmid comprising a gene encoding a tester protein according to anyone of claims 1-7,
- transforming with said plasmid a suitable host cell comprising the gene encoding said protease under a control of a tightly regulated promoter,
- growing transformants under repressing or non-inducing conditions with respect to said promoter and under conditions requiring a function of said tester protein,
- shifting growing cells to derepressing or inducing conditions with respect to said promoter and non-permissive conditions with respect to a function of said tester protein,
- determining the nucleic acid sequence of the gene encoding said tester protein of a surviving cell.

30. A method to identify the cleavage site of a known protease comprising the steps of
- providing an expression vector encoding a non-regulatory marker protein suitable for
positive as well as negative selection with at least one known permissible site in a surface loop for the insertion of a sequence,
- inserting a coding sequence for about 5-39 amino acids into said site, resulting in a plasmid comprising a gene encoding a tester protein according to anyone of claims 1-7,
- providing a first yeast cell comprising a protease capable of cleaving said cleavage sequence and a second yeast cell lacking said protease,
- transforming said first yeast cell with said plasmid and growing transformants under non-permissive conditions with respect to a function of said tester protein,
- isolating said plasmid from a surviving cell,
- transforming said second cell with said isolated plasmid and growing transformants under conditions requiring a function of said tester protein,
- determining the nucleic acid sequence of the gene encoding said tester protein of a growing cell.

31. A method to identify the cleavage site of a known protease comprising the steps of
- providing an expression vector encoding a non-regulatory marker protein suitable for positive as well as negative selection with at least one known permissible site in a surface loop for the insertion of a sequence,
- inserting a coding sequence for about 5-39 amino acids into said site, resulting in a
plasmid comprising a gene encoding a tester protein according to anyone of claims 1-7,
- providing a first yeast cell comprising a protease capable of cleaving said cleavage sequence and a second yeast cell lacking said protease,
- transforming said second yeast cell with said plasmid and growing transformants under conditions requiring a function of said tester protein,
- isolating said plasmid from a growing cell,
- transforming said first yeast cell with said isolated plasmid and growing transformants under non-permissive conditions with respect to a function of said tester protein,
- determining the nucleic acid sequence of said gene encoding said tester protein of a surviving cell.

32. A method to identify the cleavage site of a known protease comprising the steps of
- providing an expression vector encoding a non-regulatory marker protein suitable for positive as well as negative selection with at least one known permissible site in a surface loop for the insertion of a sequence,
- inserting a coding sequence for about 5-39 amino acids into said site, resulting in a plasmid comprising a gene encoding a tester protein according to anyone of claims 1-7,
- providing a yeast cell lacking a protease capable of cleaving said cleavage sequence,
- transforming said yeast cell with said plasmid and selecting for growth under con-
ditions requiring a function of said tester protein, obtaining transformants,
- providing a second plasmid capable of expressing a gene encoding said protease,
- transforming said transformants with said second plasmid and selecting for growth
  under conditions non-permissive with respect to a function of said tester protein,
- determining the nucleic acid sequence of said gene encoding said tester protein of a
  surviving cell.

33. A method to identify a protease showing improved activity and/or changed specificity or a deriva-
tive of said protease, comprising the steps of
- providing cells expressing a functional, non-regulatory tester polypeptide suitable for negative selection,
- providing an expression library comprising putative genes encoding said protease,
- transforming said cells with said expression library,
- growing transformants under non-permissive conditions with respect to a function of said tester protein,
- identifying among surviving clones those which lack full-length tester polypeptide,
- determining from identified clones the nucleic acid sequence of the gene encoding said protease.
Figure 1A
Figure 2

A

<table>
<thead>
<tr>
<th>Trp1 wt</th>
<th>Empty vector</th>
<th>Trp1^{49} -M</th>
<th>Trp1^{102} -M</th>
<th>Trp1^{132} -M</th>
<th>Trp1^{165} -M</th>
<th>Trp1^{194} -M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non selective</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selective</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B

![Graph showing OD_{595} values for different Trp1 variants with and without protease](image-url)
Figure 3

A
a) GGV\textsubscript{VAS}CRLAGG
b) PTALLSGGAKVAERAQAGVV\textsubscript{NAS}CRLATASGSEAATAGP
c) PTALLSGGAKVAERAQAGVV\textsubscript{NES}CRLATASGSEAATAGP

B

\begin{tabular}{|c|c|c|c|c|}
\hline
 & No protease & HCMV protease & No protease & HCMV protease & Inactive HCMV protease & HCMV protease \\
\hline
TRP\textsuperscript{194}-M (short) & 1 & & & & & \\
\hline
TRP\textsuperscript{194}-M (long) & & & 1 & & & \\
\hline
\end{tabular}

C

33 kDa - - TRP\textsuperscript{194}-M (long)

17 kDa - - calmodulin
Figure 4

![Graph showing the OD₉₉₅ values for different HCMV protease expression levels.](image)
Figure 5

A

![Graph showing the effect of BI31 concentration on OD595 for inactive and active protease.

B

Table showing the effects of different concentrations of BI36 on TRP1194-M with and without active protease.

- TRP1194-M + active protease
- DMSO
- 100 µM BI36
- 50 µM BI36
- 10 µM BI36

-protein bands at 33 kDa and 17 kDa for TRP1194-M and calmodulin respectively.

- TRP1194-M
- calmodulin
Figure 6
ESBATech AG

Genetic selection system to identify proteases and protease inhibitors

12440PC

38

PatentInn version 3.3

1

10

PRT

artificial

protease cleavage sequence

human CMV protease cleavage sequence

(1) .. (10)

Gly Val Asn Ala Ser cys Arg Leu Ala
1 5 10

Gly Gly Val Val Asn Ala Ser Cys Arg Leu Ala Gly Gly
1 5 10

Lys Val Ala Gln Arg Ala Asn Ala Gly Val Val Gln Ala Ser cys Arg
1 5 10 15
Leu Ala Thr Ala ser
20

<210> 4
<211> 39
<212> PRT
<213> artificial

<220> protease cleavage sequence

<221> human CMV protease cleavage sequence
<222> (1) .. (39)
<400> 4

Pro Thr Ala Leu Leu ser Gly Gly Ala Lys val Ala Glu Arg Ala Gin
1 5 10 15

Ala Gly val Val Asn Ala Ser Cys Arg Leu Ala Thr Ala Ser Gly Ser
20 25 30

Glu Ala Ala Thr Ala Gly Pro
35

<210> 5
<211> 4
<212> PRT
<213> artificial

<220> protease cleavage sequence

<221> human CMV protease cleavage sequence
<222> (1) .. (4)
<223> x = any amino acid
<400> 5

val xaa Ala ser
1

<210> 6
<211> 4
<212> PRT
<213> artificial

<220> protease cleavage sequence

<221> human CMV protease cleavage sequence
<222> (1) .. (4)
<223> X = any amino acid
<400> 6
Leu xaa Ala ser

1

<210> 7
<211> 4
<212> PRT
<213> artificial

<220> protease cleavage sequence

<220> human CMV protease cleavage sequence
<221> (D..C4)
<222> X = any amino acid
<400> 7

lie xaa Ala ser

1

<210> 8
<211> 11
<212> PRT
<213> artificial

<220> protease cleavage sequence

<220> HSV-I protease cleavage sequence
<221> (l)..(11)
<400> 8

Ala Leu Val Asn Ala Ser Ser Ala Ala His Val

1 5 10

<210> 9
<211> 10
<212> PRT
<213> artificial

<220> protease cleavage sequence

<220> protease cleavage sequence
<221> vZ
<222> (l)..(10)
<400> 9

Gin Asp Val Asn Ala Val Glu Ala Ser ser

1 5 10

<210> 10
<211> 11
<212> PRT
<213> artificial
<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Pro</td>
<td>Ser</td>
<td>lie</td>
<td>Leu</td>
<td>Asn</td>
</tr>
<tr>
<td>Ser</td>
<td>Phe</td>
<td>Asn</td>
<td>Phe</td>
<td>Pro</td>
</tr>
<tr>
<td>Thr</td>
<td>Leu</td>
<td>Asn</td>
<td>Phe</td>
<td>Pro</td>
</tr>
<tr>
<td>Thr</td>
<td>Leu</td>
<td>Asn</td>
<td>Phe</td>
<td>Pro</td>
</tr>
</tbody>
</table>

- **EBV** protease cleavage sequence
- **HHV-6** protease cleavage sequence
- **HIV-I** protease cleavage sequence
protease cleavage sequence

Hepatitis C virus protease
(D...C11)

Asp Leu Glu val val Thr Ser Thr Trp val Leu
1 5 10

protease cleavage sequence

Hepatitis C virus protease cleavage sequence
(1)....(13)

Coxsackievirus 3C protease cleavage sequence
(1)....(13)

Gly Thr Thr Leu Glu Ala Leu Phe Gin Gly Pro Pro val
1 5 10

protease cleavage sequence

Rhinovirus C3 protease cleavage sequence
(1)....(10)

Leu Glu val Leu Phe Gin Gly Pro Leu Gly
1 5 10

artificial
protease cleavage sequence

SARS coronavirus 3C-like protease

C1) .. (5)

Ser Ala Val Leu Gin Ser Gly Phe

Trp Phe Lys Asp Ser

Phe Glu Asp Asp Ala
Tyr  val  His  Asp  Ala
1    5

<210>  21
<211>  5
<212> PRT
<213> artificial

<220> protease cleavage sequence

<221> Caspase-1 cleavage sequence
<222> (D..C5)

<220> Caspase cleavage sequence
<221> (C1)...C5)
<222> Cl)..C5)
<400> 21

Asp Gly Pro Asp Gly
1    5

<210>  22
<211>  5
<212> PRT
<213> artificial

<220> protease cleavage sequence

<221> Caspase-2 cleavage sequence
<222> (C1)..(5)

<220> Caspase cleavage sequence
<221> CD..C5)
<222> CD..C5)
<400> 22

Asp Glu Val Asp Gly
1    5

<210>  23
<211>  5
<212> PRT
<213> artificial

<220> protease cleavage sequence

<221> caspase-3 cleavage sequence
<222> CD-..C5)
<400> 23

Glu Thr Asp ser
1    5
protease cleavage sequence

Caspase-3 cleavage sequence
Cl)...(5)

Asp Glu Val Asp Asn
1 5

Caspase-3 cleavage sequence
(D..C5)

Asp Met Gin Asp Asn
1 5

Caspase-3 cleavage sequence
(D..C5)

Asp Glu Pro Asp Ser
1 5

protease cleavage sequence
Caspase-3 cleavage sequence

(D..C5)

Asp Glu Ala Asp Gly
1 5

Asp Glu Thr Asp Ser
1 5

Asp Ala Cys Asp Thr
1 5

val Glu lie Asp Asn
1 5
<220> Caspase-6 cleavage sequence
<222> Cl.. (5)

<220> Caspase-8 cleavage sequence
<222> Cl.. (5)

<400> 31

Val Glu Thr Asp Ser
1  5

<210>  32
<211>  5
<212> PRT
<213> artificial

<220> protease cleavage sequence

<220> Caspase-8 cleavage sequence
<222> CD-- C5)

<400>  32

Leu Glu Met Asp Leu
1  5

<210>  33
<211>  8
<212> PRT
<213> artificial

<220> protease cleavage sequence

<220> Plasmepsin cleavage sequence
<222> CD-- (8)

<400>  33

Glu Arg Met Phe Leu ser Phe Pro
1  5

<210>  34
<211>  6
<212> PRT
<213> artificial

<220> protease cleavage sequence

<220> Thrombin cleavage sequence
<222> CD ..C6)

<400>  34
VaI Pro Arg Ser Phe Arg
1 5

Gly Pro Leu Gly VaI Arg Gly Leu
1 5

Leu Ala Arg Arg Lys Pro Val Leu Pro Ala Leu Thr Lie Asn Pro
1 5 10 15

pro Phe His Leu Leu VaI Tyr Ser
INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER

INV. G01N33/53

According to International Patent Classification (IPC) or to both national classification and IPC.

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

GOIN

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic database consulted during the international search (name of database and, where practical, search terms used)

EPO-Internal, BIOSIS, COMPENDEX, INSPEC, EMBASE, MEDLINE, Sequence Search, WPI Data, PAO

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No.
--- | --- | ---
X | US 6 699 702 B1 (PATEL PRAMATHESH ET AL) 2 March 2004 (2004-03-02) claims 1-24; examples 1-3 | 1-33
X | WO 00/40745 A (BRISTOL-MYERS SQUIBB COMPANY) 13 July 2000 (2000-07-13) claims 1-26; figures 2,5; example 1 | 1-33

Further documents are listed in the continuation of Box C.

* Special categories of cited documents:

*A* Document defining the general state of the art which is not considered to be of particular relevance

%E* Earlier document but published on or after the international filing date

*L* Document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

’O* Document referring to an oral disclosure, use, exhibition or other means

’P* Document published prior to the international filing date but later than the priority date claimed

Date of the actual completion of the international search

12 December 2006

Date of mailing of the international search report

04/01/2007

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel: (+31-70) 340-2040, Tx: 31-651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Stachowiak, 01af
<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>LAWLER J F ET AL: &quot;Viral protease assay based on GAL4 inactivation is applicable to high-throughput screening in mammalian cells&quot;</td>
<td>1-33</td>
</tr>
<tr>
<td></td>
<td>ANALYTICAL BIOCHEMISTRY, ACADEMIC PRESS, SAN DIEGO, CA, US,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>vol. 269, no. 1, 10 April 1999 (1999-04-10), pages 133-138,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ISSN: 0003-2697, pages 134-135; figure 1</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>US 5 891 635 A (DASMAHAPATRA ET AL)</td>
<td>1-33</td>
</tr>
<tr>
<td></td>
<td>6 April 1999 (1999-04-06) claims 1-11</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>US 5 800 981 A (BRUGGEMAN ET AL)</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>1 September 1998 (1998-09-01) sequence 12</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>WO 01/75088 A (ESBATECH AG; LUETHI, URS; BARBERIS, ALCIDE)</td>
<td>1-33</td>
</tr>
<tr>
<td></td>
<td>11 October 2001 (2001-10-11) the whole document</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>US 6 194 143 B1 (BONNEAU PIERRE R [CA])</td>
<td>1-33</td>
</tr>
<tr>
<td></td>
<td>27 February 2001 (2001-02-27) the whole document</td>
<td></td>
</tr>
<tr>
<td>Patent document cited in search report</td>
<td>Publication date</td>
<td>Patent family member(s)</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>-----------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>US 6699702 B1</td>
<td>02-03-2004</td>
<td>NONE</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>AU 2960700 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA 2360022 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP 2002534095 T</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>US 5891635 A</td>
<td>06-04-1999</td>
<td>NONE</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>US 5800981 A</td>
<td>01-09-1998</td>
<td>AT 307207 T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AU 732254 B2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AU 1876397 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA 2246802 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DE 69734387 T2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DK 0882132 T3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WO 9731117 A2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 0882132 A2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ES 2251016 T3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP 2001517072 T</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA 2401155 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 1268755 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>US 6194143 B1</td>
<td>27-02-2001</td>
<td>NONE</td>
</tr>
</tbody>
</table>