

(19) **DANMARK**



Patent- og  
Varemærkestyrelsen

(10) **DK/EP 3421597 T3**

(12) Oversættelse af  
europæisk patentskrift

- 
- (51) Int.Cl.: **C 12 N 15/00 (2006.01)** **C 12 N 15/09 (2006.01)** **C 12 P 21/02 (2006.01)**  
**C 12 R 1/645 (2006.01)**
- (45) Oversættelsen bekendtgjort den: **2021-01-25**
- (80) Dato for Den Europæiske Patentmyndigheds  
bekendtgørelse om meddelelse af patentet: **2020-10-28**
- (86) Europæisk ansøgning nr.: **18166845.0**
- (86) Europæisk indleveringsdag: **2010-03-16**
- (87) Den europæiske ansøgnings publiceringsdag: **2019-01-02**
- (30) Prioritet: **2009-03-16 EP 09003750**
- (62) Stamansøgningsnr: **10709080.5**
- (84) Designerede stater: **AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC  
MK MT NL NO PL PT RO SE SI SK SM TR**
- (73) Patenthaver: **Danisco US Inc., 925 Page Mill Road, Palo Alto, California 94304, USA**
- (72) Opfinder: **Punt, Peter J., Boekdrukkersglide 1, 3994 XT Houten, Holland  
Burlingame, Richard Paul, 181 Water Works Road, Nicholasville, KY 40356, USA  
Pynnonen, Christine M., 1019 South Kernan Avenue, Appleton, WI 54915, USA  
Olson, Phillip T., 3742 Adams Street, Two Rivers, WI 54241, USA  
WERY, Jan, Ewoltstede 22, 7213 TD Gorssel, Holland  
VISSER, Johannes Heinrich, Diemewei 4717, 6605 XK Wijchen, Holland  
EMALFARB, Mark A., 193 Spyglass Court, Jupiter, FL 33477, USA  
VISSER, Jacob, Hinkeloordseweg 5, 6703 CK Wageningen, Holland  
Verdoes, Jan Cornelis, -, deceased, Holland**
- (74) Fuldmægtig i Danmark: **Marks & Clerk (Luxembourg) LLP, 44 rue de la Vallée, B.P. 1775, L-1017 Luxembourg,  
Luxembourg**
- (54) Benævnelse: **SYSTEM TIL CHRYSOSPORIUM LUCKNOWENSE-PROTEINFREMSTILLING**
- (56) Fremdragne publikationer:  
**WO-A2-2009/018537  
US-A1- 2004 121 442  
VERDOES JAN C ET AL: "Original research: a dedicated vector for efficient library construction and high  
throughput screening in the hyphal fungus Chrysosporium lucknowense", INDUSTRIAL BIOTECHNOLOGY,  
MARY ANN LIEBERT, US LNKD- DOI:10.1089/IND.2007.3.048, vol. 3, no. 1, 1 January 2007 (2007-01-01) , pages  
48-57, XP009127268, ISSN: 1550-9087**



## DESCRIPTION

**[0001]** The present disclosure relates to a fungal host strain of *Chrysosporium lucknowense*. The disclosure relates furthermore to a method for homologous and/or heterologous production of a pure protein with a purity of higher than 75 %, to a method for production of artificial protein mixes and to a method for simplified screening of strains functionally expressing a desired enzyme. The invention relates to an isolated promoter sequence suitable for the transcriptional control of gene expression in *Chrysosporium lucknowense*. Also disclosed is a method for isolating a fungal host strain of *Chrysosporium lucknowense* wherein the protease secretion is less than 20 % of the protease secretion of *Chrysosporium lucknowense* strain UV 18-25.

**[0002]** Fungi have been proven to be excellent hosts for the production of a variety of enzymes. Strains like *Aspergillus*, *Trichoderma*, *Penicillium* and recently the fungus *Chrysosporium lucknowense* C1, have been applied in the industrial production of a wide range of enzymes. Super-producing strains have been developed that secrete up to 100 g/L or more protein in the fermentation broth (see for instance in Hans Visser et al., Abstracts, J. of Biotechnology, S211-S241 (2007)). The large protein-secreting capacity of these fungi make it preferred hosts for the targeted production of specific enzymes or enzyme mixes. However, typically, these hosts secrete a mix of many different enzymes, making the crude protein product undefined and yielding, besides the desired enzyme activity, a range of non-relevant or even contra-productive activities. This also holds true for the use of such fungal hosts for production of specific enzyme activities by over-expression of selected genes via genetic modification approaches. Also in these cases the target enzyme will only constitute a minor part of the total secreted protein.

**[0003]** A microbial production system able to secrete high amounts of a specific enzyme without the presence of high levels of other proteins would be highly desirable. It would enable simplified screening of hosts functionally expressing a desired enzyme. It would enable production of relatively pure enzyme. It would also enable simplified large scale purification of the desired enzyme. These advantages would greatly contribute to e.g. easy generation of artificial enzyme mixes tailored for different applications, e.g. plant biomass hydrolysis (biofuels and chemicals), textile finishing, applications in paper and pulp industry.

**[0004]** The relatively clean production of specific extracellular enzymes to high levels by microorganisms that do not intrinsically secrete high levels of protein would be a non-preferred approach. The limited enzyme secreting capacity of such organisms would prevent high level production of the enzyme of interest.

The object of the present comprises the isolation of mutants of a fungal strain with high secretion capacity that unexpectedly no longer produce high levels of many non-desired proteins, while maintaining good growth characteristics, and amenability to genetic modification. These mutant strains were should be able to function as a host for high level production of specific enzymes.

This disclosure provides a fungal host strain of *Chrysosporium lucknowense* wherein the endogenous cellulase secretion is less than 20 % of the endogenous cellulase secretion of *Chrysosporium lucknowense* strain UV 18-25, preferably less than 15 %, more preferably less than 10 %, specifically less than 5 %, more specifically less than 2 %, more specifically less than 1 %, most specifically less than 0,5 %, or less than 0,1%. The strain *Chrysosporium lucknowense* strain UV 18-25 has been described in the international patent application WO 0020555. Preferably, the secretion of one or more of the group consisting of endogenous protease, endogenous  $\beta$ -glucanase and endogenous cellobiohydrolase of the fungal host strain according to the invention is less than 20 % , more preferably less than 15%, most preferably less than 10%, especially less than 5%, more especially less than 1% most especially less than 0,5% or 0,1% of the secretion of endogenous protease, endogenous  $\beta$ -glucanase and endogenous cellobiohydrolase respectively of the *Chrysosporium lucknowense* strain UV 18-25. All percentages mentioned apply to the secretion of protease,  $\beta$ -glucanase and cellobiohydrolase independently.

Preferably the strains are further characterized in that secretion of endogenous cellobiohydrolase 1 (Cbh1) is absent. Most preferably the strain according to the present invention is strain W1L, deposited at the Centraal Bureau Schimmelcultures (CBS) under accession nr 122189 or W1L#100.1 deposited at the CBS under accession number 122190.

More preferably from the strains the gene encoding endochitinase 1 (*chi1*) has been disrupted. Most preferably, one or more genes selected from the group consisting of those encoding alkaline protease 1 (*alp1*), alkaline protease 2 (*alp2*), proteinase A (*pep4*),  $\beta$ -glucanase (*Gla1*), exo-chitinase (*Chi2*) and laminarinase (*Lam1*) have been disrupted. Especially, the strains according to the invention are W1L#100.1 $\Delta$ alp1 $\Delta$ pyr5 or W1L#100.1 $\Delta$ alp1 $\Delta$ chi1 $\Delta$ pyr5.

The disclosure also relates to a method for homologous and/or heterologous production of a pure protein with a purity of higher than 75 %, preferably of higher than 80%, more preferably of higher than 85, 90 or 95%, comprising expressing a gene encoding said protein in a strain according to the invention. Especially the invention provides a method for production of artificial protein mixes comprising expressing genes encoding each of said proteins of the mix in a strain according to the invention. In this way protein mixes may be prepared for different applications, e.g. plant biomass hydrolysis (bio fuels and chemicals), textile finishing, applications in paper and pulp industry). Furthermore the disclosure provides a method for simplified screening of strains functionally expressing a desired enzyme by application of strains according to the present invention.

According to the invention an isolated promoter sequence is provided suitable for the transcriptional control of gene expression in *Chrysosporium lucknowense*, selected from the group consisting of

1. a. the *chi1*(0.8) promoter comprising the nucleotide sequence of SEQ ID NO 25, and
2. b. the *chi1*(1.8) promoter comprising the nucleotide sequence of SEQ ID NO 26.

**[0005]** Also disclosed is:

- c. the *hex1* promoter comprising the nucleotide sequence of SEQ ID NO 27,

- d. the *xy/6* promoter comprising the nucleotide sequence of SEQ ID NO 28, and
- e. the *gla* promoter comprising the nucleotide sequence of SEQ ID NO 29 or a transcriptionally active part thereof.

**[0006]** Also a chimeric gene comprising said promoter sequence and a host comprising said promoter and chimeric gene are provided by the present invention. Finally a method for isolating a fungal host strain of *Chrysosporium lucknowense* wherein the cellulase and protease secretion is less than 20 % of the cellulase respectively protease secretion of *Chrysosporium lucknowense* strain UV 18-25 , comprising the steps of

1. (i) plating *Chrysosporium lucknowense* on acid swollen cellulose (ASC) plates,
2. (ii) selecting at least one colony showing a reduced cellulose clearing zone,
3. (iii) plating the strain selected in step (ii) on skim milk plates, and
4. (iv) selecting at least one colony showing a reduced protein degradation halo, has been provided. Preferably this method further comprises the steps of mutagenesis prior to steps (i) and/or (iii).

**[0007]** If in this patent hosts are defined by comparing the level of production of various enzymes with the *Chrysosporium lucknowense* strain UV 18-25 then of course the production of one and the same enzyme in the host has to be compared with that in the *Chrysosporium lucknowense* strain UV 18-25.

**[0008]** The proteins mentioned in this patent as for instance protease,  $\beta$ -glucanase, cellobiohydrolase, proteinase A,  $\beta$ -glucanase ,exo-chitinase and laminarinase have been defined as described in WO 2009/0918537 in the name of Dyadic International INC.. To emphasize the advantages provided by the present invention: it relates to the isolation of novel fungal hosts that have lost their intrinsic capability to secrete high levels variety of background proteins, while retaining the ability to secrete high levels of only few enzyme activities. The invention also relates to the use of these hosts to produce specific enzymes at high levels without co-production of high levels of non-specific proteins. In addition the invention relates to the generation of defined artificial enzyme mixes tailored for different applications. The scope of the invention is defined in the appended set of claims.

**[0009]** The invention will now be elucidated by the following non-limiting examples. Legends to the figures mentioned:

**Figure 1:** Medium samples of UV18-25, W1L wildtype, W1D wildtype and the protease mutants W1L#100.1, W1D#50.g and W1D#100.b on SDS-PAGE. These strains were cultivated in medium #1 (low density medium with cellulose) for 282 hours except for lane 8 which is a sample in medium #2 (high density). The medium sample in lane 1 was 2 times diluted and the

medium sample in lane 8 was 4 times diluted.

**Figure 2:** Shake flask culture medium samples of C1 strain W1L#100.1 and derivatives. Lane 1, W1L#100.1; lane 2, W1L#100.1 $\Delta$ chi1; lane 3, W1L#100.1 $\Delta$ alp1; lane 4, W1L#100.1 $\Delta$ alp1 $\Delta$ chi1.

**Figure 3:** Cosmid cloning vector pAopyrGcosarp1.

**Figure 4:** Plasmid pCHI4.8. The plasmid, which was isolated from E. coli clone #5 is shown. This plasmid was used in the construction of the white strain gene expression vectors.

**Figure 5:** Over-expression of chi1 via the introduction of extra chi1 gene copies in W1L#100.1. Lane 1, W1L#100.1 wild type strain (control); lane 3, W1L#100.1[chi+/pyr5]#3; lane 4, W1L#100.1[chi+/pyr5]#9; lane 5, W1L#100.1[chi+/pyr5]#17.

**Figure 6:** Plasmid Pcbh1-glaA(II)-Tcbh1. This plasmid was used in the construction of the white strain gene expression vectors.

**Figure 7:** The pPchi1(1.8)-Tcbh1 NotI gene expression vector.

**Figure 8:** A schematic map of the pCRS-pPchi1(1.8)-Tcbh1 gene expression vector.

**Figure 9:** SDS-PAGE analysis of culture supernatant samples of W1L#100.L  $\Delta$ alp1 $\Delta$ pyr5 (B2), and CL10518-expressing transformant strains of W1L#100.L  $\Delta$ alp1 $\Delta$ pyr5 (B3, B4). The arrow indicates the CL10518 position.

**Figure 10:** SDS-PAGE analysis of culture supernatant samples of W1L#100.L  $\Delta$ alp1 $\Delta$ pyr5 (D1), and of cbh2-expressing transformant strains of W1L#100.L  $\Delta$ alp1 $\Delta$ pyr5 (D3-D5). The arrow indicates the CBH2 position.

**Figure 11:** SDS-PAGE analysis of culture supernatant samples of W1L#100.L  $\Delta$ alp1 $\Delta$ chi1 $\Delta$ pyr5 transformed with pyr5 only (C), and of pgx-expressing transformant strains of W1L#100.L  $\Delta$ alp1 $\Delta$ chi1 $\Delta$ pyr5 (1, 2).

**Figure 12:** SDS-PAGE analysis of culture supernatant samples of W1L#100.L  $\Delta$ alp1 $\Delta$ chi1 $\Delta$ pyr5 transformed with xyl1 (31, 34) and with xyl1 $\Delta$ cbd. The arrows indicate the protein bands representing the respective xylanase variants.

**Figure 13:** SDS-PAGE analysis of culture supernatant samples of W1L#100.L  $\Delta$ alp1 $\Delta$ chi1 $\Delta$ pyr5 transformed with med with selection marker only (71) and with abn2 (68-70), M, marker.

**Figure 14:** Analysis of chi1 transformants. #1, control strain (transformed with selection marker pyr5 only). M, marker. #65, chi1 transformant. The arrow marks the Chi1 protein band.

**Figure 15:** Protein gel of purified heterologous *Aspergillus niger* PGII produced by C1. M, Marker proteins. The molecular weights of 3 marker proteins are indicated to the left of the gel. Lane 1, purified PGII.

**Examples****Example 1: Isolation of C1 mutants with greatly reduced cellulase activity (white strains).**

[0010] C1 strain UV 18-25 (described in WO/2000/020555) was mutated using UV light to produce the strain UV26-2 (Appendix 1 to the Examples). UV26-2 exhibited large clearing zones on ASC (acid swollen cellulose) plates, indicating overproduction of cellulase.

[0011] UV26-2 was not a stable mutant and successive streaking of UV26-2 resulted in generation of cellulase negative mutants as shown by the absence of clearing zones on ASC plates. These colonies exhibited enhanced sporulation and white color, while normal, cellulase-producing, colonies were cream colored and showed no sporulation on ASC plates.

[0012] Two white colonies (UV26-2W1 and UV26-2W2) were picked from ASC plates along with two normal colonies and cellulase production was evaluated using a shake flask screening procedure. The white colonies produced 6 and 4 U/ml of AzoCMC cellulase activity, while the two normal colonies produced 278 and 294 U/ml of cellulase activity. This confirmed the white colonies to be cellulase-negative mutants.

**Example 2: Isolation and analysis of protease deficient strains of UV26-2W1L and UV26-2W1D.**

[0013] The purification of strain UV26-2W1 on RM-ASP medium plates (Appendix 2 to the Examples) resulted in the identification of 2 types of colonies: colonies with light colored spores like UV 18-25 (UV26-2W1L further indicated as strain W1L), and colonies with dark (pink) colored spores (UV26-2W1D further indicated as strain W1D).

[0014] In additional experiments batches of spores of both W1L and W1D were irradiated with UV (Appendix 1 to the Examples) and used in a direct selection procedure for protease-deficient mutants (Braaksma et al., 2008). Positive clones were analyzed on skim milk plates for their protease activity.

[0015] After several rounds of purification and selection on skim milk plates, two mutants of W1L (W1L#50.c and W1L#100.1) and three mutants of W1D (W1D#50.g, W1D#50.n and W1D#100.b) with a reduced halo on skim milk plates were selected for cultivation for in vitro degradation assays. In a first cultivation experiment these mutants and their parent strains

were cultivated in medium #2 Appendix 1 to the Examples) for 240 hours at 35°C. Apparently, the low cellulase activity in these strains did not allow for growth in high density cellulose based medium. In following cultivation experiments W1L#50.c, W1L#100.1, W1D#50.g and W1D#100.b and their parents were grown in low (#1) and high (#2) density cellulose medium for 240 hours at 35°C. Also UV 18-25 was taken as a control. The parent strains 2W1D, 2W1L and UV 18-25 were also cultivated in medium #2. None of the W1L or W1D strains grew in high density cellulose medium #2. In medium #1 good growth could be observed for the 'white' strains and their protease-deficient mutants, although the cellulose in the medium was hardly used by the 'white' strains. Unexpectedly, it was noted that the UV26-2W1D parent strain, which showed an unstable growth phenotype on agar plates, did use the cellulose in the medium.

**[0016]** The medium samples of the W1L parent strain showed less protease activity on skim milk plates compared to medium samples of W1D parent strain and UV18-25 (Table 1). This is contrary to what was observed when the strains were grown directly on skim milk plates. In that case a large halo could be detected around the colony of UV18-25 and of W1L after 72 hours growth at 30°C, while a small halo could only be detected after 144 hours for W1D. The medium samples of protease mutant W1D #50.g showed a smaller halo on milk plates until 162 hours of cultivation. After 186 hours cultivation, halos were similar as observed for its parent strain.

**Table 1:** Medium analysis of W1L and W1D parent strains and their selected protease mutants: W1L#100.1, W1D#50.g and W1D#100.b. UV18-25 was taken as control. Protease activities of medium samples of W1L, W1L#100.1 and W1L#100.1 $\Delta$ alp1 were also determined. These strains were cultivated in medium #1 (low cellulose/lactose/pharmamedia). The pH was measured and medium was spotted on skim milk plates to determine their protease activity. The relative size of the halo is a measure for the protease activity in the medium. nd, not determined.

|                         | Relative halo size   |     |     |     |     |     |     | Protease activity |
|-------------------------|----------------------|-----|-----|-----|-----|-----|-----|-------------------|
|                         | (hrs of cultivation) |     |     |     |     |     |     |                   |
| Strain                  | 114                  | 138 | 162 | 186 | 210 | 240 | 282 | U/ml              |
| W1 L parent             | ++                   | ++  | ++  | ++  | ++  | ++  | ++  | 943               |
| W1L#100.1               | +                    | +++ | +++ | +++ | +++ | +++ | +++ | 119               |
| W1L#100.1 $\Delta$ alp1 | nd                   | nd  | nd  | nd  | nd  | nd  | nd  | 46                |
| W1D parent              | +++                  | +++ | +++ | +++ | +++ | +++ | +++ | nd                |
| W1D#50.g                | -                    | -   | -   | ++  | ++  | ++  | ++  | nd                |
| W1D#100.b               | ++                   | ++  | ++  | ++  | ++  | ++  | ++  | nd                |
| UV 18-25                | +++                  | +++ | +++ | +++ | +++ | +++ | +++ | nd                |

**[0017]** Analysis of 282 hours medium samples of these strains on SDS-PAGE gels showed that the 'white strains' produced much less protein than UV18-25 (Fig. 1). In particular the two major 50/70 kDa proteins (Cbh1) were absent in these culture supernatants. In the white

strains the 'major' proteins are 75 and 45 kDa. These proteins are present in medium of UV 18-25 as minor proteins.

[0018] From this first screening for protease-less mutants in a UV26-2W1 background the strains W1D#50.g and W1L#100.1 were selected for further analysis.

**Example 3: Comparison of extracellular enzyme activities between UV 18-25 and W1L#100.1.**

[0019] Different enzyme activities in the extracellular protein content of UV18-25 and W1L#100.1 samples were determined (Table 2). Based on these data it was concluded that W1L#100.1 secretes very little specific cellulase activity (less than 1 % of UV18-25) and has very little or no detectable protease activity when compared to UV 18-25.

**Table 2:** Specific activities of samples (U/mg of protein). Protease activities were measured at 3 different pH values.

| Activities         | UV18-25 | W1L#100.1 |
|--------------------|---------|-----------|
| CMCase (cellulase) | 6.20    | 0.04      |
| Beta-glucanase     | 10.2    | 0.53      |
| Cellobiohydrolase  | 0.72    | 0.09      |
| Protease (pH 5)    | 0.06    | 0.03      |
| Protease (pH 7)    | 0.05    | 0.00      |
| Protease (pH 9)    | 0.04    | 0.00      |

[0020] In addition, the levels of hydrolases bearing other substrate specificities (e.g. hemicellulose) were reduced as well.

**Example 4: Further reduction of protein level: identification of major proteins.**

[0021] As described above, the white strain is missing the extracellular cellulolytic enzyme spectrum when compared to its parental strain. Hence, the extracellular protein content in white strain cultures, as analyzed by SDS-PAGE, is low. This strain characteristic is beneficial with regard to protein production and purification, since the relative amount of any target protein expressed in such strain will be high. Furthermore, the (nearly) absence of cellulase activity makes the white strain an ideal host strain for testing new or modified cellulases. The same is valid for xylanases as no major xylanase activity was detectable.

[0022] To further reduce the protein background level, several major protein bands present in an SDS-PAGE gel from a W1L#100.1 and derivative strain cultures were excised and identified

by N-terminal sequencing and/or MS-MS analysis. The most abundant protein was the endochitinase Chi1 (gene identifier: CL06081, peptides MVYDYAG, MPIYGRS, and MFXEASA). Other major proteins were identified as a glucoamylase (Gla1, CL09507, peptides TGGWSVWPVLK (SEQ ID No 1) and VVGSSEL(I)GNWDTGR (SEQ ID No 2)), exo-chitinase (Chi2, CL00367, peptides TIDAMAWSK (SEQ ID No 3), NFLPVADILR (SEQ ID No 4), GAYHPSQTYSPEDVEK (SEQ ID No 5), and SWQLVYQHDPTAGLTAEAAK (SEQ ID No 6) and a laminarinase (Lam1, CL08253, peptides PQYESAGSVVPSSFLSVR (SEQ ID No 7) and VSGQVELTDFLVSTQGR (SEQ ID No 8)). Also an alkaline protease Alp1 (CL04253) has been identified in W1L#100.1 culture broth. Alp1 degrades extracellular proteins, and may degrade proteins of interest.

**Example 5: Further reduction of protein level: disruption of the *chi1*, *chi2*, *gla1* and *lam1* genes.**

[0023] The vector pChi3-4 (see Example 9, Isolation of the endochitinase 1 encoding gene) was used for the construction of the gene disruption vector. A 1.1-kb *MscI/StuI* fragment was replaced with the *amdS*-rep selection marker or the *pyr5*-rep selection marker, resulting in the vectors *pΔchi1-amdS* and *pΔchi1-pyr5*, respectively. The disruption fragment *Δchi1-amdS* was isolated from *pΔchi1-amdS* by digestion with *EcoRI*. The disruption fragment *Δchi1-pyr5* was isolated from *pΔchi1-pyr5* by digestion with *SmaI*. Transformation of strain W1L#100.1 *Δpyr5*#172-12 using the disruption fragments resulted in 215 *Δchi1-pyr5* transformants and 32 *Δchi1-amdS* transformants. All the obtained transformants were purified and analyzed with colony hybridization. Southern analysis of these transformants confirmed the isolation of one W1L #100.1 transformant with a disrupted *chi* gene (W1L#100.1 *ΔpyrSΔchi1-pyr5*#46 (*pyr5*+)).

[0024] Shake flask cultures on C1 low density medium were performed from a selection of W1L#100.1 *Δpyr5Δchi1* mutant strains. The samples were analyzed on SDS-PAGE to evaluate the protein profiles for absence of Chi1 protein (Fig. 2, lane 2 versus lane 1). As shown no 45 kDa Chi1 protein is observed in the *Δchi1* mutant strain.

[0025] The remaining most prominent extracellular proteins in white strain W1L#100.1 *Δalp1Δchi1* and derivative strains correspond to glucoamylase (Gla1), exo-chitinase (Chi2) and laminarinase (Lam1). These enzymes were purified from the culture medium. The enzymatic activities of these proteins were verified using (among others) starch, chitosan and laminarin, respectively, as substrates. Furthermore, mass spectrometry analyses data (see Example 4) combined with C1 genome sequence data revealed the corresponding genes. In order to further reduce the extracellular protein background, the Gla1, Chi2, and Lam1 encoding genes were disrupted and thereby inactivated. Disruption was based on the exchange of the gene promoter and part of the 5' coding sequence by an *amdS* selection marker via homologous recombination using approximately 1.5 kbp upstream and downstream sequences that flank these gene promoter and part of the 5' coding sequence. The gene disruption vectors therefore contained the *amdS* expression cassette plus these flanking 1.5 kb homologous gene sequences. White strains W1L#100.1 *Δalp1Δchi1* and derivative strains were

transformed with the *gla1*, *chi2* and *lam1* gene disruption vectors and transformants were screened for the correct genotype using PCR. As such, white strains with a further reduced extracellular protein composition/content were obtained. Target proteins produced by these strains were more than 80% pure in the crude cell-free culture liquid.

**Example 6: Further reduction of protease activity: targeted disruption of genes encoding proteases.**

[0026] In general, protease encoding genes were disrupted using disruption DNA fragments that contained selection markers (*amdS*, *pyr4* or *pyr5*) flanked by approximately 1.5 kb large DNA fragments homologous to regions up- and downstream of the gene to be disrupted. Upon introduction of these disruption DNA fragments into the white host, an homologous recombination exchanged the gene to be disrupted for the selection marker fragment. Corresponding transformants were selected as such. Genes that were disrupted this way either encoded disadvantageous (with regard to target protein stability) protease activities e.g. *alp1*, *alp2*, *pep4*) or significant background protein (*chi1* or were to be used as selection marker (*pyr4*, *pyr5*). Via this approach numerous white C1-strains have been constructed that can be used as hosts for target protein expression (Table 3).

**Table 3: Strain W1L and derivatives.**

|                                               |
|-----------------------------------------------|
| W1L                                           |
| W1L SUI <sup>R</sup> #S2 6.14                 |
| W1 L SUI <sup>R</sup> #S2 6S                  |
| W1L#100.1                                     |
| W1L# 100.1 Δpyr5                              |
| W1L# 100.1 Δalp1                              |
| W1L# 100.1 Δalp1 Δpyr5                        |
| W1 L# 100.1 Δpep4 Δpyr5                       |
| W1L# 100.1 Δalp1 Δpep4                        |
| W1L# 100.1 Δalp1 Δpep4 Δpyr5                  |
| W1L# 100.1 Δalp1 Aalp2 Δpyr5                  |
| W1 L#100.1 Δchi1                              |
| W1L# 100.1 Δalp1 Δchi1                        |
| W1L# 100.1 Δalp1 Δchi1 Δpyr5                  |
| W1L# 100.1 Δalp1 Δchi1 Aalp2                  |
| W1L# 100.1 Δalp1 Δchi1 Aalp2 Δpyr5            |
| W1L# 100.1 Δalp1 Δchi1 Δpep4                  |
| W1L# 100.1 Δalp1 Δchi1 Agla1 Δlam1 Δchi2Δpyr5 |

**Example 7: Identification of strong promoters for gene expression: chitinase encoding gene (*chi1*).**

**[0027]** Several major protein bands were isolated from fermentation samples of W1L#100.1 grown in low density cellulose medium in order to identify and isolate strong promoters that can be used for gene expression in the W1L strain and its derivatives. N-terminal sequencing of a mixture of peptides obtained after CNBr treatment of the major 45 kDa protein of W1L#100.1 resulted in the identification of four different peptides. Three of these peptides (MVYAG, MPIYGRS and MFXEASA) showed homology with an endochitinase of *Aphanocladium album*/*Trichoderma harzianum* (CHI\_APHAL P32470).

**[0028]** Based on 3 of these peptide sequences, primers were designed in order to obtain PCR fragments containing a part of the endochitinase encoding gene (Table 4). The PCR primers were designed based on the preferred codon usage of C1.

**Table 4:** The designed primers of putative endochitinase based on codon usage of C1.

| Primer                                                          | Region                   | Position | Deduced sequence                      |
|-----------------------------------------------------------------|--------------------------|----------|---------------------------------------|
| Endochitpep1C                                                   | MVYDYAG (SEQ ID No 9)    | 240 aa   | ATGGTSTACGACTACGCBGG (SEQ ID No 10)   |
| Endochitpep2C                                                   | MPIYGRS(S EQ ID No 11)   | 290 aa   | ATGCCSATCTACGGYCG (SEQ ID No 12)      |
| Endochitpep2revC                                                |                          |          | CGRCCGTAGATSGGCAT (SEQ ID No 13)      |
| Endochitpep3revC                                                | MFXEASA(S EQ ID No 14)   | 380 aa   | GCSSWGCCTCCCAGAACA T (SEQ ID No 15)   |
| Primer based on a conserved homologous region of endochitinase: |                          |          |                                       |
| Endochit3c                                                      | DGIDIDWEV (SEQ ID No 16) | 160 aa   | GAYGGYATCGAYRTSGAYTGGG (SEQ ID No 17) |

**[0029]** PCR reactions with these primers were carried out using chromosomal DNA of UV18-25 as template DNA. PCR fragments were cloned and sequence analysis showed that one of the cloned PCR fragments obtained with Endochitpep1c and Endochitpep2revc (173 bp) contained a part of an endochitinase-encoding gene (*chi1*). Hybridization analysis of chromosomal DNA of UV 18-25 digested *Bam*HI and *Hind*III with this *chi1* fragment as probe showed a clear hybridization signal confirming that the PCR fragment originated from C1 DNA. This fragment was used to clone the complete gene from the ordered C1-cosmid gene library. The fragment sequence (SEQ ID No 18) was as follows:

```
ATGGGCTACGACTACGCCGGCTCGTGGAGCACCGCGGGCGGGACACCAGGCCA
ACCTGTACCCGACCGCCGACGCGGGCAGGACGCCCTTCTCGACCGACAAGGC
CCTGTCCGACTACGTCCGCCGGCGTCGACCCGGCCAAGATCGTGCTCGGC
ATGCCCATCTACGGCCG
```

**Example 8: Construction of an ordered cosmid library of *Chrysosporium lucknowense* UV18-25 in *E. coli*.**

[0030] For the construction of the C1 cosmid library the non commercial cosmid cloning vector, pAOpyrGcosarp1 (Fig. 3) was used. This vector carries the *Aspergillus oryzae* pyrG selection marker allowing transformation of a wide range of fungal strains. Moreover, for highly efficient transformation of various *Aspergillus* species (for which a large collection of mutant strains is available for complementation cloning of the corresponding C1 gene) the AMA1-replicator is present in this vector. A unique *Bam*HI cloning site allows cloning of partially *Sau*3A digested genomic DNA from *Chrysosporium* strain UV 18-25.

[0031] A cosmid library of UV 18-25 was constructed in *E. coli* and stored as glycerol stocks in the -80 °C. The average insert size was 20-35 kb. In total, 6800 clones were obtained, which represents a genome coverage of about 5 fold. To order these clones in 384-wells plates, dilutions of the glycerol stocks were plated on LB-agar plates, containing ampicillin. 7680 individual colonies were manually picked and inoculated in twenty 384-wells plates. These twenty 384-wells plates represent the ordered cosmid library of UV 18-25 in *E. coli*.

[0032] The individual colonies in the twenty 384-wells plates were spotted on nylon filters (Hybond), using a Staccato 384-pintool (Zymarks). The ordered cosmid library was spotted in octuplicate (in total 160 filters). These filters were placed on LB-ampicillin medium plates and incubated at 37 °C to allow colonies to grow on the filters. Subsequently, the colonies were lysed and the cosmid DNA was bound to the filters using standard procedures.

**Example 9: Isolation of the endochitinase 1 encoding gene (*chi1*).**

[0033] The 173-bp *chi1* PCR fragment was used as a radio-active labelled probe for hybridization of duplicate filters of the cosmid library. Hybridization of the cosmid gene library using this *chi1* fragment resulted in 3 positive clones. DNA was isolated from these clones and restriction analysis followed by Southern analysis was carried out. A 4.8 kb *Hind*III fragment and a 3.4 kb *Bgl*II fragment showed hybridization using the *chi1* probe. The 4.8kb *Hind*III and 3.4 kb *Bgl*II fragments containing the *chi1* gene were isolated and subcloned in pMTL24, resulting in the vectors pCHI4.8 (Fig. 4) and pChi3-4, respectively. Subsequently, a *Bgl*II subclone was used for sequence analysis to obtain more sequence data. The sequence of the complete *chi1* gene of UV 18-25 was obtained. The size of the *chi1* gene is 1529bp in which 2 introns of 111bp and 138bp are present.

**Example 10: Overexpression of the endochitinase in W1L#100.1.**

**[0034]** For overexpression of endochitinase the 3.4kb *PgII* fragment containing the *chi1* gene was isolated from pChi3-4. Putative *chi1* multi-copy strains in W1L#100.1 $\Delta$ pyr5 were generated by co-transformation of the 3.4 kb *BglII* *chi1* fragment and a *pyr5* selection fragment and correct transformants were confirmed by colony hybridization. Shake-flask cultures on C1 low density medium were performed from a selection of W1L#100.1 *chi1*-multicopy strains. The samples were analyzed on SDS-PAGE to evaluate the protein profiles for the overproduction of Chi1 (Fig. 5). Three of the multi-copy *chi1* strains (lanes 3, 4, 5) show a stronger 45 kDa band compared to the parent strain (lane 1) indicating chitinase overexpression and the usefulness of the *chi1* promoter for high gene expression.

**[0035]** Therefore, a general C1 expression cloning vector (pPchi1(0.8)-Tcbh1 NotI) was constructed containing the *chi*-promoter (*Pchi*) to direct the overexpression of cloned genes. Initially, the *EcoRI* site upstream of *Pchi1* in pCHI#4.8 was removed by a partial *EcoRI* digestion and treatment of the linear fragment with Klenow yielding pCHI1#4.8 $\Delta$ *EcoRI*. From this vector the 1.9-kb *SacI*-*SphI* fragment was cloned in the corresponding site of *pPcbh1-glaA(II)*-*Tcbh1* (Fig. 6). In the resulting vector, *pPchi-Tcbh1* NotI#7.1, the target genes can be inserted into the *NcoI*-*EcoRI* sites. Expression cassettes for transformation into C1 strains can be isolated from these constructs as *NotI* fragments.

**[0036]** The 0.8 kb *chi1* promoter sequence in pCHI#4.8 could be more than sufficient to drive *chi1* expression. However, a longer *chi1* promoter was also generated by amplifying the a *PstI*-*HindIII* fragment (upstream of the *HindIII* site at position -775 relative to the ATG start codon), using one of the previously identified positive cosmid clone as template DNA. The resulting fragment was cloned in pGEM-T-Easy and sequenced. From this plasmid the *PstI*-*HindIII* fragment was isolated and cloned in the corresponding sites of pPchi1-*xyl1*-*Tcbh1* yielding pPchi1(1.8)-*xyl1*-*Tcbh1*, in which the promoter size is 1.8 kb. The fragment was also cloned in the corresponding site of *pPchi1-Tcbh1* NotI#7.1, yielding the general expression vector *pPchi1*(1.8)-*Tcbh1* NotI (Fig. 7).

**[0037]** The levels of gene expression directed by the extended chitinase promoter (*Pchi1*(1.8)) and by the initially used chitinase promoter (*Pchi1*(0.8)) were compared by the expression of two reporter genes, *xyl1* and *alp1*. White strain transformants were generated that either expressed *xyl1* (encoding a xylanase) or *alp1* (encoding an alkaline protease) (Table 5).

**Table 5:** Comparison of the short and extended *Pchi1* promoters in terms of reporter protein expression level. Reporter activity, xylanase activity is expressed as U/ml and alkaline protease activity as U/mg of protein. A = W1L#100.1[*Pchi*(0,8)-*alp*/pyr5]#9, B = W1L#100.1[*Pchi*1(1,8)-*alp*/pyr5]#22, C = W1L#100.1 $\Delta$ *alp1*[*Pchi*1-*xyl1*]#95, D = W1L#100.1 $\Delta$ *alp1*[*Pchi*1(1.8)-*xyl1*]#A7.

| Reporter   | 0.8 kb Pchi 1     | 1.8 kb Pchi 1     |
|------------|-------------------|-------------------|
|            | Reporter activity | Reporter activity |
| Alp1 (R19) | 0,7 (A)           | 1,9 (B)           |
| Xyl1 (R14) | 122 (C)           | 1175 (D)          |

**[0038]** Surprisingly, the reporter gene expression was higher in case of the extended *chi1* promoter (1.8 kb), which indicates the necessity of the further upstream regions. In conclusion, a *Pchi1* based expression system was developed for high level expression of genes in White C1 strains.

### Example 11: Identification of other strong promoters for gene expression.

**[0039]** A different approach for searching strong promoters was performed using the quantitative detection of messenger RNA levels from W1L or W1L#100.1 RNA. The RNA samples were isolated from mycelium which was sampled at different time points during a fed-batch fermentation process. A number of genes were identified as being strongly or stronger expressed. To verify the expression level of these genes, the RNA samples were also separated on gel, blotted and hybridized to probes specific for these genes (Table 6).

**Table 6:** Quantification of the expression signals of the different genes in controlled fed-batch fermentations. Probe hybridisation signals were quantified using a densitometer. The signal of the probe on the Northern blot was correlated with the signal of this probe on a C1 genomic DNA Southern blot. Therefore, the values in the table represent the northern hybridization signal level relative to the hybridization signal level from the Southern blot (which was set at 1). Gene sequences are given below.

[illegible]

**[0040]** The *cbh1* promoter, which is a strong promoter in UV 18.25 strains, is not active in the white strains. The *chi1* promoter was the strongest both under glucose and xylose feed conditions. The *hex1* promoter is a strong constitutive promoter during all phases of fermentation and under both sugar feed conditions. The *xyl6* promoter is highly active under the xylose feed condition only. The *pep4*, *his2a* and *bgl1* promoters are moderately active. For high level gene expression in the white strains the *chi1*, *hex1* and *xyl6* promoters are very useful. Alternative promoters that also give high expression are those of the *pep4*, *his2a* and *bgl1* genes. Additional northern experiments also indicated that the promoters of the *xyl 4* and *xyl8* genes can be used for high level gene expression in the white strains when grown on xylose. Glucoamylase (Gla1, gene identifier: CL09507) has been shown to be a major protein in white C1 strains. The *gla1* promoter is therefore also a good candidate to be used for the high level expression of genes of interest in white strains. It was shown that glucoamylase was highly abundant in a white strain grown in the presence of starch. This indicated that the *gla1* promoter is strong and inducible by starch and its degradation products, like maltose.

**[0041]** The nucleotide sequences of Pchi1(0.8), Pchi1(1.8), Phex1, Pxyl6 and Pgla1 are given below. Note that the ATG start codons of the corresponding coding regions are given in bold italics.

#### **Example 12: White strain gene expression system**

**[0042]** Two expression vectors were designed for expression of genes in W1L and derivatives: pPchi1(1.8)-Tcbh1 *NotI* was as described above. Additionally, pCRS Pchi1-Tcbh1 (Fig. 8) was constructed by placing the C1 repeat sequence in front of the *Pchi1* promoter in pPchi1(1.8)-Tcbh1 *NotI*. These vectors have been designed in such a way that they also can be combined resulting in a single vector that contains multiple expression cassettes. The multiple cassettes can be excised from the vector as a single linear DNA fragment.

**[0043]** Many genes have been cloned and expressed in W1L or derivatives. The general procedure was as follows:

- Genes were identified by purification and characterization of the gene products (reverse genetics) and / or by genome mining. The genes were amplified by PCR or synthesized chemically. Amplification of genes by PCR was performed by using proof-reading PCR DNA polymerases (Phusion or Supertaq plus). The amplified genes were cloned into PCR cloning vectors i.e. pGEM-T-Easy (Promega) or pJet1 (Fermentas) and sequenced to verify the correctness of the sequence. Subsequently, the genes were released from the PCR cloning vector and ligated into the *NcoI* and *EcoRI* sites of the expression vector(s).
- Special care was taken when designing the PCR primers. The ATG-start codon of the gene to be expressed was part of the *NcoI* restriction site in the white strain expression vectors. Therefore, the 5' (ATG) PCR primers contained restriction sites, which are compatible to the restricted *NcoI* site of the vector. These sites were i.e. *NcoI* itself

(C↓CATGG), or compatible sites that are cut within the recognition site (*Bsp*HI, T↓CATGA; *Pci*I, A↓CATGT), or compatible sites that are cut outside the recognition site (*Bsa*I, GGTCTC(1/5); *Bsp*MI, ACCTGC(4/8); *Esp*3I, CGTCTC(1/5))

- In some cases, restriction sites additional to those that were going to be used for cloning of the genes were encountered in the genes. In these cases, the genes were amplified by fusion PCR, where the fusion of the two PCR fragments was selected to take place at the undesired additional restriction site. The undesired restriction site was removed by using fusion primers containing single substitution mutations in the undesired restriction site sequence. In case the undesired restriction site was present within a protein coding region, the substituted nucleotide was selected in such a manner that the mutant codon encoded the same amino acid as the original codon.
- The expression cassettes were released from the *E. coli* DNA vector backbone by *Not*I restriction. The expression cassette was subsequently transformed in to W1L derivatives simultaneously with a selection marker i.e. *pyr5* or *amdS* in co-transformation experiments.
- Positive and high producing transformants were selected by SDS-PAGE or enzyme assay analysis of the growth medium. Best producers were applied in fermentations to produce high amount of the desired gene product.
- The following proteins have been produced using the white strain gene expression system and corresponding genes

**[0044]** In the international patent application WO 2009/018537 has been described:

Abf1, Abf2, Abn1, Axe1, Bgl1 (=Bg13A), Cbh1, Cbh2, Cbh4, Chi1, Eg2, Eg5, FaeA1, FaeA2, FaeB2, Gall (=Ga153A), Gla1 (=Gla15A), Pme1, Xyl1, Xyl1(cd), Xyl2, Xyl3, Xyl3-cbd (=xyl3(cd)), Xyl4, Xyl5, Xyl6.

**[0045]** In the international patent application WO 2009/033071 has been described:

Abf3, Abn2, Abn3, Abn4, Abn5, Abn7, Abn9, Agu1, Axe2, Axe3, Bga2, Bxl1, Bxl2, Abf5(formerly known as Bxl3), GH61 genes (gene identifiers: CL09768, CL10518, CL05022, CL04725, CL04750, CL06230, CL05366), Gln, Pgx1, Rga1, Rgx1, XglI, Xyl7, Xyl8, Xyl9, Xyl10, Xyl11.

Alp1: The Alp1 DNA sequence is given by SEQ ID NO 30. The Alp1 amino acid sequence is given in SEQ ID NO 31.

**[0046]** The *alp1* gene was expressed and Alp1 showed protease activity (Table 5). The "Protease colorimetric detection kit" (Sigma, product number PC0100) was used to determine Alp1 protease activity.

#### **Example 13: Expression of C1 GH61-family-encoding proteins genes.**

**[0047]** A GH61 protein-encoding gene (identifier CL10518) was overexpressed in C1 strains W1L#100.L  $\Delta$ alp1 $\Delta$ pyr5 and W1L#100.1 $\Delta$ alp1 $\Delta$ chi1 $\Delta$ pyr5. Culture supernatant samples were analyzed by SDS-PAGE and protein bands were stained with Coomassie brilliant blue (CBB). The CL10518 protein is  $\pm$ 26 kDa (Fig. 9). A standard fed batch fermentation was conducted, which yielded 13 g protein per liter fermentation filtrate, based on a BCA protein determination assay.

**[0048]** The functional analysis of this protein has been described in International patent application WO 2009/033071.

**Example 14: Expression of a C1 cellulase-encoding gene: cbh2.**

**[0049]** The CBH2 encoding gene (identifier CL09854) was overexpressed in C1 strain W1L#100.L  $\Delta$ alp1 $\Delta$ pyr5. Culture supernatant samples were analyzed by SDS-PAGE and protein bands were stained with coomassie brilliant blue (CBB). The CBH2 protein migrates at about 55 kDa (Fig. 10). A standard fed batch fermentation was conducted, which yielded 10 g protein per liter fermentation filtrate, based on a BCA protein determination assay.

**[0050]** The functional analysis of this protein has been described in International patent application WO 2009/018537.

**Example 15: Expression of C1 exo-polygalacturonase encoding gene pgx.**

**[0051]** The PGX encoding gene (identifier CL10389) was overexpressed in C1 strain W1L#100.L  $\Delta$ alp1 $\Delta$ chi1 $\Delta$ pyr5. Culture supernatant samples were analyzed by SDS-PAGE and protein bands were stained with coomassie brilliant blue (CBB). The PGX protein migrates at about 60 kDa (Fig. 11). A standard fed batch fermentation was conducted, which yielded 9 g protein per liter fermentation filtrate, based on a BCA protein determination assay.

**[0052]** The following assay was used to measure polygalacturonase activity. This assay measures the amount of reducing sugars released from polygalacturonic acid (PGA) by the action of a polygalacturonase. One unit of activity was defined as 1  $\mu$ mole of reducing sugars liberated per minute under the specified reaction conditions.

**Reagents**

**[0053]**

- Sodium acetate buffer (0.2 M, pH 5.0) is prepared as follows. 16.4 g of anhydrous sodium acetate or 27.2 g of sodium acetate \* 3H<sub>2</sub>O is dissolved in distilled water so that

the final volume of the solution to be 1000 mL (Solution A). In a separate flask, 12.0 g (11.44 mL) of glacial acetic acid is mixed with distilled water to make the total volume of 1000 mL (Solution B). The final 0.2 M sodium acetate buffer, pH 5.0, is prepared by mixing Solution A with Solution B until the pH of the resulting solution is equal to 5.0.

- Polygalacturonic acid (PGA) was purchased from Sigma (St. Louis, USA).
- Reagent A: 10 g of *p*-Hydroxy benzoic acid hydrazide (PAHBAH) suspended in 60 mL water. 10 mL of concentrated hydrochloric acid was added and the volume is adjusted to 200 mL. Reagent B: 24.9 g of trisodium citrate was dissolved in 500 mL of water. 2.2 g of calcium chloride was added as well as 40 g sodium hydroxide. The volume was adjusted to 2 L with water. Both reagents were stored at room temperature. Working Reagent: 10 mL of Reagent A was added to 90 mL of Reagent B. This solution was prepared freshly every day, and store on ice between uses.

**[0054]** Using the above reagents, the assay is performed as detailed below.

#### **Enzyme Sample**

**[0055]** - 50  $\mu$ L of PGA (10.0 mg/mL in 0.2 M sodium acetate buffer pH 5.0) was mixed with 30  $\mu$ L 0.2 M sodium acetate buffer pH 5.0 and 20  $\mu$ L of the enzyme sample and incubated at 40 °C for 75 minutes. To 25  $\mu$ L of this reaction mixture, 125  $\mu$ L of working solution was added. The samples were heated for 5 minutes at 99°C. After cooling down, the samples were analyzed by measuring the absorbance at 410 nm ( $A_{410}$ ) as  $A_S$  (enzyme sample).

#### **Substrate Blank**

**[0056]**

- 50  $\mu$ L of PGA (10.0 mg/mL in 0.2 M sodium acetate buffer pH 5.0) was mixed with 50  $\mu$ L 0.2 M sodium acetate buffer pH 5.0 and incubated at 40 °C for 75 minutes. To 25  $\mu$ L of this reaction mixture, 125  $\mu$ L of working solution was added. The samples were heated for 5 minutes at 99°C. After cooling down, the samples were analyzed by measuring the absorbance at 410 nm ( $A_{410}$ ) as  $A_{SB}$  (substrate blank sample).

#### **Calculation of Activity**

**[0057]**

- Activity is calculated as follows: determine polygalacturonase activity by reference to a standard curve of galacturonic acid.

$$\text{Activity (IU/ml)} = \Delta A_{410} / \text{SC} * \text{DF}$$

where  $\Delta A_{410} = A_S$  (enzyme sample) -  $A_{SB}$  (substrate blank), SC is the slope of the standard curve and DF is the enzyme dilution factor.

**[0058]** The  $\Delta A_{410}$  of Pgx1 (CL10389) was found to be 0.78 with a DF of 1 for enzyme produced in microtiter plate cultures. No standard curve was analyzed, therefore no reliable activity can be calculated. The only conclusion to be drawn is that the enzyme was found to be active towards polygalacturonic acid, and therefore it is suggested that it is a polygalacturonase.

**[0059]** The functional analysis of this protein has been described in International patent application WO 2009/033071.

**Example 16: Expression of a C1 xylanase encoding gene with and without its carbohydrate binding domain: Xyl1.**

**[0060]** The Xyl1 encoding gene (identifier CL00649) was overexpressed in C1 strain W1L#100.L  $\Delta alp1 \Delta chi1 \Delta pyr5$ . Two Xyl1 variants were produced: either full length Xyl1 or Xyl1 without its carbohydrate binding domain (cbd). Culture supernatant samples were analyzed by SDS-PAGE and protein bands were stained with coomassie brilliant blue (CBB) (Fig 12). The Xyl1 protein migrates at about 40 kDa, while its CBD-less counterpart migrates at about 30 kDa. Standard fed batch fermentations were conducted, which yielded up to 33 g protein per liter fermentation filtrate, based on a Bradford protein determination assay. Xylanase activities of these filtrates reached up to 3,500 U/mL.

**[0061]** The following assay is used to measure the xylanase activity towards AZO-wheat arabinoxylan. This substrate is insoluble in buffered solutions, but rapidly hydrates to form gel particles which are readily and rapidly hydrolysed by specific endo-xylanases releasing soluble dye-labeled fragments.

**Reagents**

**[0062]**

- Sodium acetate buffer (0.2 M, pH 5.0) is prepared as follows. 16.4 g of anhydrous sodium acetate or 27.2 g of sodium acetate \* 3H<sub>2</sub>O is dissolved in distilled water so that the final volume of the solution to be 1000 mL (Solution A). In a separate flask, 12.0 g (11.44 mL) of glacial acetic acid is mixed with distilled water to make the total volume of

1000 mL (Solution B). The final 0.2 M sodium acetate buffer, pH 5.0, is prepared by mixing Solution A with Solution B until the pH of the resulting solution is equal to 5.0.

- AZO-wheat arabinoxylan (AZO-WAX) from Megazyme (Bray, Ireland, Cat. # I-AWAXP) is used as the assay substrate. 1 g of AZO-WAX is suspended in 3 mL ethanol and adjusted to 100 mL with 0.2 M sodium acetate buffer pH 5.0 using magnetic stirrer. 96% Ethanol is used to terminate the enzymatic reaction.

**[0063]** Using the above reagents, the assay is performed as detailed below:

#### **Enzyme Sample**

**[0064]**

- 0.2 mL of 10 mg/ml AZO-WAX stock solution was preheated at 40° C for 10 minutes. This preheated stock solution was mixed with 0.2 mL of the enzyme sample (preheated at 40° C for 10 min) and incubated at 40 °C for 10 minutes. After exactly 10 minutes of incubation, 1.0 mL of 96% ethanol was added and then the absorbance at 590 nm ( $A_{590}$ ) was measured as  $A_S$  (enzyme sample).

#### **Substrate Blank**

**[0065]**

- 0.2 mL of 10 mg/ml AZO-WAX stock solution was preheated at 40° C for 10 minutes. This preheated stock solution was mixed with 200 µl of 0.2 M sodium acetate buffer pH 5.0 (preheated at 40° C for 10 min) and incubated at 40 °C for 10 minutes. After exactly 10 minutes of incubation, 1.0 mL of 96% ethanol is added and then the absorbance at 590 nm ( $A_{590}$ ) is measured as  $A_{SB}$  (substrate blank).

#### **Calculation of Activity**

**[0066]**

- Activity is calculated as follows: determine endo-xylanase activity by reference to a standard curve, produced from an endo-xylanase with known activity towards AZO-WAX.  
$$\text{Activity (IU/ml)} = \Delta A_{590} / SC * DF$$

where  $\Delta A_{S90} = A_S$  (enzyme sample) -  $A_{SB}$  (substrate blank), SC is the slope of the standard curve and DF is the enzyme dilution factor.

[0067] The functional analysis of this protein has been described in International patent application WO 2009/018537.

**Example 17: Expression of the C1 arabinase 2 encoding gene *abn2*.**

[0068] The *Abn2* encoding gene (identifier CL03602) was overexpressed in C1 strain W1L#100.L  $\Delta alp1\Delta chi1\Delta pyr5$ . Culture supernatant samples were analyzed by SDS-PAGE and protein bands were stained with coomassie brilliant blue (CBB). The *Abn2* protein migrates at about 50 kDa (Fig. 13). A standard fed batch fermentation was conducted, which yielded 7 g protein per liter fermentation filtrate, based on a BCA protein determination assay.

[0069] The functional analysis of this protein has been described in International patent application WO 2009/033071.

**Example 18: Expression of the C1 endo-chitinase encoding gene *chi1*.**

[0070] The *Chi1* encoding gene (identifier CL06081) was overexpressed in C1 strain W1L#100.L  $\Delta alp1\Delta pyr5$ . Culture supernatant samples were analyzed by SDS-PAGE and protein bands were stained with coomassie brilliant blue (CBB). The *Chi1* protein migrates at about 40 kDa (Fig. 14). A standard fed batch fermentation was conducted, which yielded 12 g protein per liter fermentation filtrate, based on a BCA protein determination assay.

[0071] The functional analysis of this protein has been described in International patent application WO 2009/018537.

**Example 19: Expression of a heterologous gene: *Aspergillus niger* poly-galacturonase II.**

[0072] The *Aspergillus niger* poly-galacturonase II encoding-gene (accession number X58893) was expressed in C1 strain W1L#100.1  $\Delta alp1\Delta chi1\Delta pyr5$ . After fermentation the enzyme was purified using ion exchange chromatography and size exclusion chromatography. The purified endo-PGII migrated at about 40 kDa on SDS-PAGE gel (Fig. 15).

[0073] This heterologous enzyme was functional as was shown by activity on polygalacturonic acid using the following assay.

**Reducing sugars assay: PAHBAH method****Stock solutions:****[0074]**

- Substrate: 1% (w/v) polygalacturonic acid in H<sub>2</sub>O.
- Reagent A: *p*-Hydroxy benzoic acid hydrazide (PAHBAH) (10 grams) is added to 60 ml of water and slurried. To this is added 10 ml of concentrated HCl and the volume is made up to 200 ml. Reagent B: Dissolve trisodium citrate (24.9 g) in 500 ml of water, add calcium chloride (2.20 g) and dissolve, add sodium hydroxide (40.0 g) and dissolve. Adjust the volume to 2 litres. The solution should be clear. Store both reagents at room temperature. Working Reagent: add 10 ml of Reagent A to 90 ml of Reagent B. Prepare this solution fresh daily, and store on ice between uses.

**Assay:****[0075]**

1. 1. 50 µl substrate
2. 2. 30 µl 0.2M HAc/NaOH pH 5.0
3. 3. 20 µl sample / enzyme (microplate undiluted; fermentor >20 x diluted)
4. 4. Incubate at 37°C for 10 minutes
5. 5. 25 µl assay mix + 125 µl Working Reagent (in PCR microplate) or 50 µl assay mix + 250 µl Working reagent (in 1.5-ml tube)
6. 6. Heat @ 99°C for 5 minutes in Thermal PCR Cycler (PCR microplate) or in boiling water (1.5-ml tube)
7. 7. Transfer 100 µl to NUNC microplate and measure extinction @ 410nm

**Example 20: Generation of artificial enzyme mixes for efficient plant biomass saccharification.**

**[0076]** An artificial enzyme mixture was created by mixing crude protein from C1 UV18-25Δalp1 with crude protein from white strains expressing C1-β-glucosidase Bgl1, C1-arabinofuranosidase Abf3 and Abn7, C1-xylanase Xyl2 and C1-β-xylosidase Bxl1, being

W1L#100.L $\Delta$ alp1 $\Delta$ pyr5[Bgl1/pyr5], W1L#100.L $\Delta$ alp1 $\Delta$ chi1 $\Delta$ pyr5[Abf3/pyr5],  
 W1L#100.L $\Delta$ alp1 $\Delta$ chi1 $\Delta$ pyr5[Abn7/pyr5], W1L#100.L $\Delta$ alp1 $\Delta$ pyr5[Xyl2/pyr5], and  
 W1L#100.L $\Delta$ alp1 $\Delta$ chi1 [Bxl1/AmdS], respectively. The ratio of the different components on a protein basis was 10 (UV18-25 $\Delta$ alp1): 1 (white strain proteins).

**[0077]** The saccharification efficiency of the crude protein from UV18-25 $\Delta$ alp1 alone was tested on wheat bran substrate and compared to the efficiency of the artificial mixture. 10 mg protein/g dry matter wheat bran was used. Conditions were as follows: temperature 50°C, pH 5.0, time 72 hours.

**[0078]** It was shown that the enzyme mixture from UV18-25 $\Delta$ alp1 alone liberated approximately 30% of the glucose, approximately 5% of the xylose and 12% of the arabinose from the wheat bran. The artificial mixture liberated at least 60% of the glucose, 60% of the xylose and 25% of the arabinose from the wheat bran.

**Example 21: Construction of gene libraries in white C1 strain W1L#100.1 $\Delta$ alp1 $\Delta$ chi1 $\Delta$ pyr5 and screening of xylanases.**

**[0079]** A gene library from genomic DNA of C1 strain UV 18-25 was constructed in C1-strain W1L#100.1 $\Delta$ alp1 $\Delta$ chi1 $\Delta$ pyr5 by methods previously described by Verdoes *et al.* (2007). The library was screened for xylanase activity as described by Example 4 in chapter 4, which yielded several positive clones that expressed different xylanases. SDS-page analysis revealed the presence of extra protein bands in the positive clones. PCR analysis using different primer combinations based on the sequence of the known C1 xylanases and the vector sequence revealed the presence of 3 different C1-xylanases. This result was confirmed by Southern analysis.

**Appendix 1 to the Examples: UV mutation procedure for C1 strains**

**[0080]**

1. 1. Spread parent strain onto PDA (potato dextrose agar) plates and incubate at 35°C for 14 days to obtain spores.
2. 2. Scrape spores into 0.9% saline and filter through cotton to remove mycelia. Dilute the resulting spore suspension to 1x10<sup>6</sup> spores/ml using saline. Remove a small aliquot of spore suspension, dilute in saline and spread plate to PDA to determine the initial viable spore count.
3. 3. Add 10 ml spore suspension to a sterile glass Petri dish containing a paper clip and stir on a magnetic stir plate. Remove the glass top and irradiate with UV light to obtain 90-99.9% kill. Use a Pen-Ray lamp as the UV light source (254 nm) and warm it up for at least 10 minutes prior to irradiating the spore suspension.
4. 4. Spread plate to ASC selective plates (Appendic 2 to the Examples) with room lights

- off, using a volume to obtain less than 30 colonies on each plate.
5. 5. Invert plates, put in red plastic bags and incubate at 30°C for 6-7 days to grow and allow clearing zones to develop.
6. 6. Determine % kill for the mutation as the difference between the initial viable plate count and a plate count on PDA after UV irradiation.

Appendix 2 to the Examples: Media

ASC Selective Agar Plates

[0081]

| Component                                       | Amount |    |
|-------------------------------------------------|--------|----|
| Deionized water                                 | 800    | ml |
| K <sub>2</sub> HPO <sub>4</sub>                 | 1.0    | g  |
| KCl                                             | 0.1    | g  |
| NaCl                                            | 0.1    | g  |
| MgSO <sub>4</sub> -7H <sub>2</sub> O            | 0.3    | g  |
| FeCl <sub>3</sub> -6H <sub>2</sub> O            | 0.016  | g  |
| (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> | 1.9    | g  |
| 20 g/l ASC                                      | 200    | ml |
| Noble Agar                                      | 15     | g  |

[0082] Adjust pH to 7.5 with HCl and sterilize 30 minutes at 121°C. After sterilization add 20 ml of 25 g/l DOC (deoxycholic acid), sterile filtered. Pour about 20 ml/ plate. Spread UV-mutated spores to ASC plates and incubate for 7-14 days to allow colony growth and cellulose clearing.

RM-ASP medium

[0083]

| Component           | Amount |    |
|---------------------|--------|----|
| Bacto Peptone       | 2      | g  |
| Bacto Yeast Extract | 1      | g  |
| 50x AspA (+N)       | 20     | mL |

| Component            | Amount                     |    |
|----------------------|----------------------------|----|
| Glucose              | 10                         | g  |
| 1000x trace elements | 1                          | mL |
| MgSO4-7 H2O          | 0.493                      | g  |
| Water                | Bring total volume to 1 L. |    |

[0084] Adjust pH to 6.5 prior to autoclaving. Sterilize glucose separately as a 50% solution.

| 50x AspA (+N)                                                           |                               |    |
|-------------------------------------------------------------------------|-------------------------------|----|
| Component                                                               | Amount                        |    |
| NaNO <sub>3</sub> (or (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ) | 300 (or 233)                  | g  |
| KCl                                                                     | 26                            | g  |
| KH <sub>2</sub> PO <sub>4</sub>                                         | 76                            | g  |
| KOH 10 N                                                                | 22.5                          | mL |
| Water                                                                   | Bring total volume to 1 L.    |    |
| 1000x trace elements                                                    |                               |    |
| Component                                                               | Amount                        |    |
| ZnSO <sub>4</sub> -7H <sub>7</sub> O                                    | 2.2                           | g  |
| H <sub>3</sub> BO <sub>4</sub>                                          | 1.1                           | g  |
| MnSO <sub>4</sub> -H <sub>2</sub> O                                     | 0.427                         | g  |
| FeSO <sub>4</sub> -7H <sub>2</sub> O                                    | 0.5                           | g  |
| COCl <sub>2</sub> -6H <sub>2</sub> O                                    | 0.17                          | g  |
| CuSO <sub>4</sub> -5H <sub>2</sub> O                                    | 0.16                          | g  |
| Na <sub>2</sub> MoO <sub>4</sub> -2H <sub>2</sub> O                     | 0.15                          | g  |
| EDTA                                                                    | 5                             | g  |
| Water                                                                   | Bring total volume to 100 mL. |    |

*Low and high density cellulose media*

[0085]

| Component (g/L)                 | Low density cellulose medium (#1) | High density cellulose medium (#2) |
|---------------------------------|-----------------------------------|------------------------------------|
| BisTris                         | 15.7                              | 15.7                               |
| K <sub>2</sub> HPO <sub>4</sub> | 0.22                              | 0.66                               |

| Component (g/L)                                 | Low density cellulose medium (#1) | High density cellulose medium (#2) |
|-------------------------------------------------|-----------------------------------|------------------------------------|
| KH <sub>2</sub> PO <sub>4</sub>                 | 0.08                              | 0.24                               |
| (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> | 4                                 | 12                                 |
| Na <sub>3</sub> Citrate·H <sub>2</sub> O        | 4                                 | 12                                 |
| MgSO <sub>4</sub> ·7H <sub>2</sub> O            | 0.03                              | 0.09                               |
| CaCl <sub>2</sub>                               | 0.4                               | 0.8                                |
| Yeast Extract                                   | 0.05                              | 0.15                               |
| Pharmamedia                                     | 5                                 | 15                                 |
| Lactose·H <sub>2</sub> O                        | 5                                 | 15                                 |
| Cellulose                                       | 20                                | 80                                 |

[0086] Adjust to pH 7.0.

[0087] chi1 sequence: see WO 2009/018537.

pep4 DNA sequence (SEQ ID No 19):

1 gctggctcaccggttatttgctcccgccaggaagtcagggtcctcctcgcagttggacaaac  
61 tctgcttcgcagcctgcaactttgactcaaggagcgccctcggcctcgtegattgggtaag  
121 acagcatgacgttggcctgccaaatgtcagcctctagaagcacactcccaetctcgttgg  
181 aaaggttcctaccccaagccacaagtaaacctcgtcogtcggcggtatctcggccttcgc  
241 atagagagtgtcgttcaattcgaatggtgtctctatcggatcagattcgccctgccacaa  
301 tcaaccgccgatcagcaccatggccgctcatcgagagtggaacgcctcgccctaccgtc  
361 ctcagcttcaaaaagcggacagcctccagcgttttccgaatgtcgggcattttgtccttt  
421 agtcccgctaccctccgctgcaggttctgtctccatgaactgggtatttccctgcacgcgcac  
481 cacgtatcagccgaacgcgcgtccgtcaaggctggatttcaatcttaaccgggagagctca  
541 cgcaatcattctcttggaaaccgacgcagcgctcggtcaacatctgctcgtgacgtgacata  
601 gtccctcgaccttgctcgacgaatggcgcatagcgaatgccacgaggattggagggtgtggc  
661 gtccctgtctcgtgcaggtgggtcagtcagcaataacagccagagtgcatatgctagaatg  
721 gcgcccgcgggggaggggaaagtttggttaccttgctgctgcttccctgtctgtgctcgcc  
781 atcttggacaaattctcacatggtgcagtggaaggatactgcaagcgactgttaaccga  
841 gccaacggagtgacgtcgggtttggtacctagtttaggtcaagccgttctcaagctgctg  
901 gccaaaaattcatggcggggtcgagtgggcagcgagggtactcctcgtagggagcaagggtg  
961 aagatgtggggtagcaggggtcgacgctacaaagtactttgtatccggattgctgtgtgg  
1021 tacgaagcgcccggtgtgttggtatgctctctgtatgtacggagtagtgcaccttctccat  
1081 gcgctgccccattctctatatttggttgcaacctgcttcgtagtgtagtacagcagta  
1141 caactatctacgacacctgcaactgactagtgcgtagaattcttttagtttctcgagtaagg  
1201 cgctaacgcttcgcgcagcaagcaccttctctgattgtgttactgtgctcaaacctcgc  
1261 cagccagctgcggtgctccacaagcccggccgtgcccacccgccatttgcatcccggtcc  
1321 catgaatctgtggacgaccocatccctctctgtacccgctcgcggtatcagcccagaatga  
1381 tagcgggaagacaaacgcagtgattcggattacgctcgaggaaatggggggagtagctt  
1441 gatagctctccacggcgaggggtgtctcaggctgaggtgtcaactagttgtatgtacactc  
1501 aqgacgaagcattctgcgtttttgaaacaccaatcttccaataaccgaaggtgtttgtatqca

1561 ggatcacttgaatatgtttgcacccattattactgtacctggatgattcggacagggcga  
 1621 gcatgattggtegcgcccgttttgtcacccgcattegcagcgteggcggaagcagccacgt  
 1681 agagcaactgccaaacgtttcaagagacaccccatatggagtaaattggagtaatctgtat  
 1741 ccttcagagccgtcaatcaaactattgtttctcagcaggatggcccggttgcctcatggggg  
 1801 atgtacctggttaggtagttcggttggtgatgacttccttggatgagcctgctgcgcatga  
 1861 aggtgcggggggcccccaggttgggtgcctaaaaactaactgtaaacagacgcacgggtggcga  
 1921 cgacgtagccgaaccggtgtagcgagctttccccggccactacgtaatcggggcgatgca  
 1981 ctgcaggaacacctcacacctgacctaccccccttcgcctccgcatccgtcccaaccogct  
 2041 tccccaacctttccatcaactacttccgagactcgacatcaccttttcgcgctcgtgtctc  
  
 2101 atcgtcgttatcatcaccatcggcgatagatttggttcgcttcgatcgtcgcacgccttg  
 2161 acttccattcgtccttcacgcccagccgaccgagaccagacagtcgccccaaaATGAAGGATG  
 2221 CTTTTTGTGACCGCAGCTGTGCTGCTCGGCTCCGCCCAGGGAGCAGTTCACAAAATGA  
 2281 AGCTGCAGAAGATCCCTCTCTCTGAGCAGCTTgtacgtctgaccccggttcaagcacgcgt  
 2341 cagcggctaactgaccttatcgcgtccagGAGGCGGTTCCTCATCAACACCCAGCTCGAGCA  
 2401 TCTCGGCCAAAAATACATGGGGTTGCGCCACGTGAATCTCAAGCCGATGCCATCTTTAA  
 2461 GGGCATGGTTGCCGACGTCAAGGGCAACCATCCTATTCCTCATCTCCAACCTTCATGAACGC  
 2521 ACAGTgtatgtgacgccactgtggtggcatggatggctcgtcctcaattcggagactgac  
 2581 actggagcacccctagACTTCTCCGAGATCACGATTGGAACACCCCCCT**CAGTCATTCAAGG**  
 2641 **TGGTCT**CGATACCGGTAGCTCCAACCTGTGGGTTCATCAGTCGAGTGCGGCTCGATTG  
 2701 CTTGTTACCTGCACTCGAAGTATGACTCATCTGCCTCGTCCACCTACAAGAAGAACGGAA  
 2761 CCTCGTTCGAGATCCGCTACGGGTACGGCAGCCTGAG**CGGGTTGTCTCTCAGGACACAG**  
 2821 TGTCATCGGCGATATCACTATCCAGGGCCAGGACTTTGCCGAGGCGACCAGCGAGCCCG  
 2881 GTCTTGCCCTTTGCCCTTTGGCCGTTTCGACGGTATCCTTGCCCTTGGCTACGACCGGATCT  
 2941 CAGTCAACGGCATCGTCCCGCCTTTTACAAGATGGTCGAGCAGAAGCTCATCGATGAGC  
 3001 CCGTCTTCGCCTTCTACCTGGCCGATACCAATGGCCAGTCTGAGGTTGTCTTTGGCGGTG  
 3061 TTGACCACGACAAGTACAAGGGCAAGATCACCACCATTCCGTTGAGGCGCAAGGCCTACT  
 3121 GGGAGGTTGACTTCGATGCCATTTCTTACGGCGACGACACTGCCGAGCTTGAGAACACTG  
 3181 GCATCATCCTGGACACCGGTACTTCTCTGATCGCTCTGCCAGCCAGCTCGCCGAGATGC  
 3241 TCAACGCTCAGATCGGCGCTAAGAAGAGCTACACTGGCCAGTACACCATCGACTGCAACA  
 3301 AGCGCGACTCCCTCAAGGATGTCACGTTCAACCTGGCTGGCTACAATTTACGCTCGGCC  
 3361 CCTACGACTACGTTCTCGAGGTCCAGGGCAGCTGCATTTCTACCTTTATGGGCATGGATT  
 3421 TCCCGGCTCCTACTGGGCCACTTGCGATCCTGGGCGATGCCTTCCTCCGGAGGTATTACT  
 3481 CCATTTATGACCTTGGCGCCGACACCGTCGGTCTGGCTGAGGCCAAGtgattgaaggatg  
 3541 ggcggcagggaaagacgatgggtaatacgggggagtcctgggaatcgggctttggactgtgg  
 3601 tctgtatctagttgctcaagagagttgtcgtttgattttgttataggatctgtctaggaa  
 3661 ccttagcaggagtgaatttttttcgtgtacgagcatcggcgggctgaagtggtttgataa  
 3721 caagtctggacttgagtacgcaggcagttgcacaatctgcttcgcgagggagagcaaagg  
 3781 cgtcctctttgaaaaagcctacctaagcgtcacaggggtataattttttgagtttgacct  
 3841 acgocctgtccataccaaccgcgtcccaatccccgtcaacccttgcaatgtcattaccc  
 3901 gtggatgtatcacgtagcagaagccgacatcccacacgcttcaaccttcctatccagaca  
 3961 atgacatggtaagctcatttttttaaaggctgcgcgtcctccctcccttcacgtgattcatt  
 4021 ttcccttgcgcccttggtggcgccatccctgacttcgatgcggtacggatcaaagggtgcaaac  
 4081 ttgccccgcacctcttttctgcccgcacatcatcaccatcategcgcgtttgtcgccctgc  
 4141 gcagcatgtagcacggaacgacgccttgctgtagtcacaaacggctcctgctcggcatcgtea  
 4201 tcatggccttcctcctgttgcggcgaggtctgttcgtcggtgcgaggtcgcgggcgag  
 4261 gcagatgtctgctgctgctgctgctgctgctgcttctgggctttcttgcgggctcgaagt  
 4321 gccttcctggcttgagccttgagttcctttgctccctttatgtctccgcttttgagccagt

4381       tgctctgccaaagagctgagcacgcttgaactcttctcgagcagccttcttggcttggttc  
 4441       tttgccctgcttggcggccttgtcatcaccacccctcaacttccctgctcgacactaggagac  
 4501       ttccgggtggtcttttgccctgcggaactatctccacccatctcgatgtcggaaactgcttcg  
 4561       gcttcggatgctgactcaacatcaacatccctagacttccgcttttcgaccagccttcaga  
 4621       gtgaaaccttcttcttcgagaacagggagacccttggtgtcttgttcagcgacacgcctg

**Pep4 amino acid sequence (SEQ ID No 20):**

1       MKDAFLLTAA VLLGSAQGAV HKMKLQKIPL SEQLEAVPIN TQLEHLGQKY MGLRPRESQA  
 61       DAIFKGMVAD VKGNHPIPIIS NFMNAQYFSE ITIGTPPQSF KVVLDTGSSN LWVPSVECGS  
 121       IACYLHISKYD SSASSTYKKN GTSFEIRYGS GSLSGFVSQD TVSIGDITIQ GQDFAEATSE  
 181       PGLAFAFGRE DGILGLGYDR ISVNGIVPPF YKMVEQKLID EPVFAFYLD TNGQSEVVFG  
 241       GVDHDKYKKG MTTIPLRRKA YWEVDFDAIS YGDDTAELN TGIILDTGTS LIALPSQLAE  
 301       MLNAQIGAKK SYTGQYTIDC NKRDSLKDVT FNLAGYNFTL GPYDYVLEVQ GSCISTFMGM  
 361       DFPAPTGPLA ILGDAFLRRY YSIYDLGADT VGLAEAK

**his2a DNA sequence (SEQ ID No 21):**

1   cattcatggg tttagagccc gattttgaac gtatatccta ggctatattc ggggtaagat  
 61   actggaagcg ctgggcccga tgactagcta tttcaagtgc ccaagagccc atcataccta  
 121   acttgtggcc taagatctag ccaaatcatt cattgggttac ccagactcg acgaacctga  
 181   tattcgaatc cagggcaagt caaatcgccg agtaagactt gacaaaccog gaaccaaga  
 241   actgcgcaat ctgggagcag gtttcggacc agcatggaaa caccocgatg gaaaaccac  
 301   acatacgggg atggggacta acgccggaca aatcaaaaac cctggaggat tgggtaacga  
 361   tggggaagtg cgacgggcac tcaacccttc aagcgttgca ggaccttgta cagccaagca  
 421   gaatgacgga aaccgatgag caaaccocga atctgatgat cctggaacag aatcatctgt  
 481   cttgggtacc gacgttgag tgagagtgtg caaattagca ggatcaagca actatactac  
 541   ctaaatcagg tcgatcagtt atcagccctt gcaaaccaga cttgatggag ggaagagggtg  
 601   aaagctgtga ttgagggagg aagctgagaa ttggtgggtg ttgttttgct cagccagggt  
 661   gtaggacgag aagaacgcgt tcgagatttc ggagagcagg ctgtcctaga gcattatatt  
 721   cctggccttg agcaaaetta agccagtttt tttttcccg tcgggaggga agtcgcttg  
 781   aatttgaagc ttgcgggcgc agagctcggc tccataagca tccaatcaaa tgagcctgaa  
 841   gcagtcgacc gatttttttt tatctgggtg taatcgcaac catgcacata accgttttgg  
  
 901   gactagctcc aacagctcgc atcaacaacc tgagaaaggc gcgagtgate cgtgatccca  
 961   cacccttaag cgaaaactac ttaactccca cctccccac cgcggtcaa cttcttccaa  
 1021   ctcccactca accaacttcc gttttcccat caatcactgc attcgcgct caagctcttc  
 1081   ctcgccctta caccaaccac ataacttttt tatcctttga caaggacct caatcaaaAT  
 1141   GACTGGCGGC GGCAAGTCCG GTGGCAAGGC GAGCGGTTC AAGAACGCGC AATCgtaggt  
 1201   gcccttttgc cgtcatctac ccgcgccttc gtgcagttgg gcattggtca gccttgaact  
 1261   ccagatgccg gttccggtgc tttacagtt ggctaacttt ttgtagTCGT TCATCTAAGG  
 1321   CCGGTCTTGC GTTCCCTGTC GGTCGTGTCC ACCGCCTTCT CCGGAAGGGC AACTACGCCC  
 1381   AGCGTGTGGG TGCCGGTGCT CCGTTTACC TGGCTGCCGT TCTCGAGTAT CTGCGCGTG  
 1441   AAATTCTGGA GCTGGCTGGC AACGCCGCTC GCGACAACAA GAAGACGCGT ATCATCCCGC  
 1501   GTCGCTTCCG TCGCGCTTTC TCGCTTTCG TCGCTTTCG TCGCTTTCG TCGCTTTCG

1501 GTCACCTTGA ACTCGCTATC AGGAACGATG AGGAGTTGAA CAAGCTTCTC GGGCACGTCA  
 1561 CCATCGCCCA GGGTGGTGTC CTTCCCAACA TCCACCAGAg taogttgcot taccagacga  
 1621 tctctaattgc gcaaattctaa ctttgtttcc agACCTTCTG CCGAAGAAGA CCGGCAAGAC  
 1681 CGGCAAGAAC TTGTGCGCAGG AGCTCtgatt ttgcgggttg ggtttttttg ctttattttc  
 1741 tggtcggcac gctgggttca tgatatcggg gtcacggttt cgggtcattg gttgcttttt  
 1801 gcgcgtggtt gggctgtaca ttaattccat gatgggcatg gtcattggtta tgaatgagaa  
 1861 tatcctctga acatccaaat cctgacacag ttgctcgag ttgatgtctg cattggaagc  
 1921 gactcgttga cggtaaccgcg tagagtcttg tcgcttacga aattcttgca tcgcacagat  
 1981 taccagtag tgccatagta ctctttaaga tgataagtgc atttgagccc ggcatcgcac  
 2041 agactttccc atgcottgat atatgcgaat tcctatgtac aagagattcg tcgcgaaaga  
 2101 gcccgtaaaa acttgagcgg ggggggagct gcaaaagcct gtcagctaatt tcgagtgaga  
 2161 cgcgcgaaagc aagccaactt acgatccagg tggggcgccg ggaggtttct ctcgtatttc

#### His2A amino acid sequence (SEQ ID No 22):

1 MTGGGKSGGK ASGSKNAQSR SSKAGLAFPV GRVHRLLRKG NYAQRVGAGA PVYLAHVLEY  
 61 LAEILELAG NAARDNKKTR IIPRHLQLAI RNDEELNKLL GHVTIAQGGV LPNIHQNLLP  
 121 KKTGKTGKNL SQEL

#### hex1 DNA sequence (SEQ ID No 23):

1 gtcaacttactccgagtcctcgcatcgagttcgatactgagcacggtactcacaactccgt  
 61 cattgacgttgctgagggcgagtatcgtgcccggtgtccagcccaaccacgcgaagcaagc  
  
 121 ttccgtagtcgggtaccacgcgtcaacggatcgcggttcagccacagccgcaaggccagcag  
 181 caccacctccacccacacgcgagtagacacgcgtcgatccccctagccacgcgccccgtcta  
 241 caagaaggagtcgggtgaagtcgccggtaccactggtgacccccctgctcctcggttcgac  
 301 ctaccacgagcaggtgaacattggtgaagagacgcgttgacgctcaccggttacgctcctca  
 361 acccaacaacaacaacaagATGGGCTACTACGACGAGGACGgtaagcatcttccctcccc  
 421 tttgatgttggttccttaccgcgtgacatccatcggtcgatgctttcttagccacacacaa  
 481 gtggttgtagaagtgccgtgctcagcccgatatcagGCCACTACCACTCTTTCCGCCATG  
 541 GATTGCACAAGTTGGCTGACCGTATTGCGCATCCTGAAGGCCATGACCGCGTTGAGGTGA  
 601 GCGAGGTTCTGTGAGACCCGCCGCACCCGCGCTCCGTCTTCGGAGGCGTACACGCCGAACA  
 661 CGGTCACCATTCCGTGCCACCACATCCGCCTCGGCGACATCCTGATCCTCCAGGGCCGCC  
 721 CCTGCCAGGTCATCCGTATCTCGACCTCGGCTGCCACTGGCCAGCACCGCTATCTTGGTG  
 781 TCGACCTCTTACCAAGCAGCTCCATGAGGAGTCGTCGTTCTGTCGAACCCTGCTCCCA  
 841 GCGTCGTCGTCCAGACGATGCTTGGCCCTGTTTTCAAGCAGTACCGCGTCCTCGACATGC  
 901 AGGACGGCCACATCGTCGCCATGACCGAGACGGGCGATGTCAAGCAGAACCTGCCCGTCA  
 961 TCGACCAGAGCAACCTCTGGGGCCGCCTCAAGCAGGCCTTCGAGACTGGCCGCGGCAGCG  
 1021 TCCGTGTCCTGGTCGTTTCTGACAACGGCAACGAGATGGCTGTTGACATGAAGGTCGTCC  
 1081 ACGGCTCGCGCCTCTAAgtcaagccggcaggtttcatgcaagctttggggctacgagtc  
 1141 gggcggcattgggtttgcgtttgatgcattcttggttacggcggtgatgtcatttgaagat  
 1201 tgaaagctgcgccttggtcgactcctggcgccggatggatatacatgttcctcgggagga  
 1261 tatgaaggtttcatgtcgctagtttcacgtgtatatgatgactgtaatggatggatgttt  
 1321 atgcccactttgcgattgatatcttgaacctttttctggtcgtgtgagtgaacagtga  
 1381 ttaagtgaagtgaggtatgcaccgtttatcacaagggttgcccttgatatccaccttcaa  
 1441 cggcggtggggaatcgaagtcctccctacagtaagtagcctctcttgaatgatctgaa  
 1501 acgcaaccctccgagccactaccacacctaactacgaaacaaccactttcctgttcag

1561      gaagotccagttotcccgtaccctcccctcccgcogttcaggttgtaogcttatctccc  
 1621      aacctcatcttcgagaggtctaataccgtacacacttaacagtgcattcctgacatagctaa  
 1681      ccatcatcactctagttcatttagccgtcccgcctatccgtcaattacattcccggctggt

#### Hex1 amino acid sequence (SEQ ID No 24):

1      MGYDDEGHY HSRHGLHKL ADRIAPPEGH DRVEVSEVRE TRRTRAPSSE AYPNTVTIP  
 61      CHHIRLGDIIL ILQGRPCQVI RISTSAATGQ HRYLGVDLFT KQLHEESSFV SNPAPSVVVQ  
 121      TMLGPVEKQY RVLDMQDQHI VAMTETGDVK QNLPVIDQSN LWGRLKQAFE TGRGSVRVLV  
 181      VSDNGNEMAV DMKVHGSRL

[0088] bglI: see WO 2009/018537. Note that bglI = Bgl3A.

xyl6: see WO 2009/018537.

cbh1: see WO 2009/018537. Note that cbh1 = CBH1a .

#### Pchi1(0.8) (SEQ I No 25):

AGCTTGACCCTTTCAGAGCTAGGTTTCATTAGGCCTTCGAAAACAACCCAAGGCCCCGTC  
 GCAACCATCACAAACGGCCGATAACCAGATCTCGGTAGGTCCGATAAGGATCCAAAATGG  
 TGTCGGCTGACGTTGCATGTGCCAGGCAGGAGGATGATCCCCAGGGTTGTTGCCGGCAG  
 CTCCCGCACGTCGGGGAGGGGGAGGGGGAGGGGAAAGCCCTAACTAACGTTTCGTTCTATC  
 ACGGGCCGACCGGGCCATGCTTTCGGCTTGTGAGCGGTGGGGTCAAGGGCAACAAGAAAT  
 GCTAAGTGCGGGACGAAGACACGCGGGCATGAGGTCTCAGGGTGACCTGCGCAAAACCAA  
 GTCCCACTCGCCATGCCTCCAGCAGCAACGTTGCCGTAGAAGGGTCAGGGGGTTTGTGTGT  
 AGACCCACGACCATGCTGCCGGCGAGCGGAGGGTTGGCTTGCTACAGGCGCTGAAGGGTC  
 AACTCGGTGCCCAAAGTGGCTACCAAGCGTGCCATCAAGGGAAATGAGATGATGGTGGCT  
 CGTGGGCAAAGAAAAGACAAGGGAGGTGACTCTAGAGAGATGCTCTCGAGTTCACGGGTA  
 TAAGAGCACTGTGATCGTTCACAAAGCCGGCGTACTCCTCTAGAGCATCTATCATCAACA  
 TCACCAGAAAGGTCNTAGACCAGGTGGTTGCCATATCCAGTCGCAAAAGAGCCAAAGAGC  
 GAAGGAGCACGAAAGCACAGCCCAATCATTCCCTGCTTTGCTACTTCTTCTCCACC**ATG**

#### Pchi1(1.8)(SEQ ID No 26):

GTCCCTTACCTATGGGCTCCTAGTCTCGTTCCTCTTTTTGATAGATTTGTATTTTGCAAC  
 GTTGCAAAATGAGACATTTCAATCATATGTAGCCGCCAGCTACTGTTAGCGTACTCAGCG  
 TTGCCCAAACGGCGGTTTTTCTGGGTAGCACTGTGCCGCGTGCCCTGAGCCGTGCGTCG  
 CGGAAACCCCTTAAGTAGCAAGTATGTTACCGCCGAGACCGACAATGCTGTTGGTTACC  
 TCGCTGGTCCATGATTGCAATCTAGATATCGTGCGGGGCTTTTGCAATCGGTTTTCCCTA  
 CCCACTTTCTTCTTTTGGACACTTTCTCTTTTGGAAAATGCCGAAATGATGCGGCTCGCT  
 CACGCCCCGAAGTCCCGAGCTGGGGCTAGATCCGTGATTGCAACGCGGTGCGAACGCGAC  
 TGGGGCAGACCTCGCTCAGCCTTGGTCGTGCCGGAATGGCGGGTACCTTTACCAGGTCCG  
 GATCAATTACATAGGATGCCATGTGCGTGGATTTGATTGCATCGCTGTCCCTTTTGTATG  
 TGTCCGAGAGCGAGACATCAACGCGAAAACCGGAATGCTCCCAACGTGCTCTCTGTTCA  
 TAGGGTCTTTTTTTTTCTTCTGCTCCATATCATCTGTCTTGAACCTAAGTGATCATCTGCT  
 GTCACGTCCCGCCCAATGATTGTAAAGAATGATAAGTGATGCTCGCCGGGGCCAGGCTCT

GTGAAAGTTCCCTCTTTGGTTGACGATCAGGTAGCGCCAACGTTGATTGGGCCGCCCGTA  
AAATCCGACCCTGTCTCCTTTTCGTTGCAAGTCTCCGCGAGACCGTGCCAAGCATGTTCTC  
CGGATCCCTCAATTACATAAGGTTTGGCTCCAGGGTAGGTCTGGAAGCTACCCACCTCGG  
CCAAGCAACCAATCACAACCAGACCTCGCGGCGTTTCGACCTTCCTGGTTTGTCTCAGGG  
CTGGCCAACGTCCTCCCGTGGCGGGTGCCTGGTGATCGCAGGTGCGAGGCGAGTGCCGGG  
CACGCGGAGCCCCCGTCAAAGCTTGACCCTTTCAGAGCTAGGTTTCATTAGGCCTTCGAA  
AACAAACCAAGGCCCGTCGCAACCATCACAACCGGCCGATAACCAGATCTCGGTAGGTC  
CGATAAGGATCCAAAATGGTGTGCGGCTGACGTTGCA'TGTGCCCAGGCAGGAGGATGATCC  
CCAGGGTTGTTGCCGGCAGCTCCCGCACGTGCGGGAGGGGGAGGGGGAGGGGAAAGCCCT  
AACTAACGTTTCGTTCTATCACGGGCCGACCGGGCCATGCTTTCGGCTTGTGAGCGGTGGG  
GTCAAGGGCAACAAGAAATGCTAAGTGCGGGACGAAGACACGCGGGCATGAGGTCTCAGG  
GTGACCTGCGCAAAACCAAGTCCCCTCGCCATGCCTCCAGCAGCAACGTTGCCGTAGAA  
GGGTCAGGGGGTTTGTGTTGTAGACCCACGACCATGCTGCCGGCGAGCGGAGGGTTGGCTTG  
CTACAGGCGCTGAAGGGTCAACTCGGTGCCCAAAGTGGCTACCAAGCGTGCCATCAAGGG  
AAATGAGATGATGGTGGCTCGTGGGCAAAGAAAAGACAAGGGAGGTGACTCTAGAGAGAT  
GCTCTCGAGTTCACGGGTATAAGAGCACTGTGATCGTTCACAAAGCCGGCGTACTCCTCT  
AGAGCATCTATCATCAACATCACCAGAAAGGTCNTAGACCAGGTGGTTGCCATATCCAGT  
CGCAAAGAGCCAAAGAGCGAAGGAGCACGAAAGCACAGCCCAATCATTCCCTGCTTTGC  
TACTTCTTCTCCACC**ATG**

**Phex1 (SEQ ID No 27):**

GATCCTAAGTAAGTAAACGAACCTCTCTGAAGGAGGTTCTGAGACACGCGCGATTCTTCT  
GTATATAGTTTTATTTTTCACTCTGGAGTGCTTCGCTCCACCAGTACATAAACCTTTTTT  
TTCACGTAAACAAAATGGCTTCTTTTCAGACCATGTGAACCATCTTGATGCCTTGACCTCT  
TCAGTTCTCACTTTAACGTANTTCGCGTTAGTCTGTATGTCCCAGTTGCATGTAGTTGAG  
ATAAATACCCCTGGAAGTGGGTCTGGGCCTTTGTGGGACGGAGCCCTCTTCTGTGGTCT  
GGAGAGCCCGCTCTCTACCGCCTACCTTCTTACCACAGTACACTACTCACACATTGCTGA  
ACTGACCCATCATACCGTACTTTATCCTGTTAATTCGTGGTGCTGTCGACTATTCTATTT  
GCTCAAATGGAGAGCACATTCATCGGCGCAGGGATACACGGTTTATGGACCCCAAGAGTG  
TAAGGACTATTATTAGTAATATTATATGCCTCTAGGCGCCTTAACTTCAACAGGCGAGCA  
CTACTAATCAACTTTTGGTAGACCCAATTACAAACGACCATACGTGCCGGAAATTTTGGG  
ATTCCGTCCGCTCTCCCAACCAAGCTAGAAGAGGCAACGAACAGCCAATCCCGGTGCTA  
ATTAAATTATATGGTTCATTTTTTTTAAAAAAATTTTTTCTTCCCATTTTCCTCTCGCTT  
TTCTTTTTTCGCATCGTAGTTGATCAAAGTCCAAGTCAAGCGAGCTATTTGTGCTATAGCT  
CGGTGGCTATAATCAGTACAGCTTAGAGAGGCTGTAAAGGTATGATACCACAGCAGTATT  
CGCGCTATAAGCGGCACTCCTAGACTAATTGTTACGGTCTACAGAAGTAGGTAATAAAAG  
CGTTAATTGTTCTAAATACTAGAGGCACTTAGAGAAGCTATCTAAATATATATTGACCCT  
AGCTTATTATCCCTATTAGTAAGTTAGTTAGCTCTAACCTATAGATAGATGCATGCGGCC  
GCAGGTACCAGGCAATTCGCCCTATAGTGAGTCGTATTACGCGCGCTCACTGGCCGTCGT  
TTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACA

TCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACA  
GTTGCGCAGCCTGAATGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGG  
TGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTT  
CGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCG  
GGGGCTCCCTTTAGGGTTCGATTTAGTGCTTTACGGCACCTCGACCCCCAAAAAACTTGA  
TTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTTCGCCCTTTGAC  
GTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCC  
TATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTTCGGCCTATTGGTTAAA  
AAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAAT  
TTAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATA  
CATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGA  
AAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTTCGGGCA  
TTTTGCCTTCCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGAT  
CAGTTGGGTGCACGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAG  
AGTTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC  
GCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTGCGCGCATACACTATTCT  
CAGAATGACTTGGTTGAGTACTACCAGTCACAGAAAAGCATCTTACGGATGGCATGACA  
GTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAAACTGCGGCCAACTTACTT  
CTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCAT  
GTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGT  
GACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACCTATTAACCTGGCGAACTA  
CTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGA  
CCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTTATTGCTGATAAATCTGGAGCCGGT  
GAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATC  
GTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT  
GAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATA  
CTTTAGATTGATTTAAAACCTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTT  
GATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTTCGTTCCACTGAGCGTCAGACCCC  
GTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTG  
CAAACAAAAAAACCACCGCTACCAGCGGTGGTTTTGTTTGCCGGATCAAGAGCTACCAACT  
CTTTTTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTG  
TAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTG  
CTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGAC  
TCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGCGGCTGAACGGGGGGTTTCGTGCACA  
CAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGA  
GAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTC  
GGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCT  
GTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGG  
AGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCT  
TTTGCTCACATGTTCTTTCCTGCGTTATCCCTGATTCTGTGGATAACCGTATTACCGCC

TTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGC  
GAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCAT  
TAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATT  
AATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGT  
ATGTTGTGTGGAATTGTGAGCGGATAACAATTTACACACAGGAAACAGCTATGACCATGAT  
TACGCCAAGCGCGCAATTAACCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGGT  
GGCGGCCGCTCTAGAACTAGTACGGCGTGCAAGTAGTGTCTTTCTTTGCACTCCCGCCGT  
CCCAGAAGACGCCGCAACAAGCTGAGCTTGCTGGAAGCCGAACAAAGGCGTTACAGAGCA  
CAAACATAGTGGCAGTGTAGGAACTCTAACTGGGACCAAAACTACGGGCCCCGGCAGAAAC  
GTTCCCCGCCCCGAAGCGAAGGCGAACGTCGAAAAGCAAGACCGGGACCGCTCGTCCCAG  
GATTAGCCACGAAGTTCCAGACCAAGTATAGGAGTAAACGCTCGCTCGTCAAAACAATTG  
TCACCAATCAGCACCCACATCGGCACATAACAACCGGTTGCGGAACTCGCATGTGAACAAC  
AAGCGGCTCCGGGGGAGTGATCGGCTCGGGCGGATGACCCGGACTCTTCCGCGCAGCAAC  
TCGGCGTGTTGTTGACGGCAGTACTCCGTAGTTGCCATGACAACAGTCAATGGCGTGCTT  
CACAAGGTGGAGAGCCGAGAAAGCACCTCGGCATGTACGAGTATGTAGATAGTGTATCAA  
GCAGGAAGATGGGGGTACTTTATCTCAATCAGATGCCTGTAAGCGAGAGCCGAGAGCCT  
GCCCTGTTGTTGACACAATTCTGGCCTGATACGAGTGACAAGCGCTGGGACGGCGGCTGG  
GGTCTTTTGCTCGCGGCTTCAGCTCAATTCCAATCCTGGGCCGGTGCCGAACGGCCCAAT  
CGCGAGCGCCACGAAATCGGAGGTGAGGAAAGAAGGCTGGGCGAGACGCGGCGACAAG  
CTGTGGCAAAATGGCCAATTGAGGTTCTGGGTGCGCTGGTGATCAACCATGCATTTCCCA  
GCCCCGAGATTCTCTTTCTCTCTCGTGCAGCAGCGGCACCAGCAGCAGCAGCAGCCAGGG  
GTTTGACCAACCTCTCCGCCCAGCCACCGATAGTAAAGATGCTGCCTGCGTATTCTGGGC  
TGCAGGAGTTCCAAGATCTTTCGGTCTGGCCACCAGCTGTCACGTCACCCTCCACCTTTG  
GACGACGTTGCTGGAAAATTCGAAGCCTTCACTAAGATAACTATGCCGTAGCACTTGCAG  
CCCCGGAAGCTGCAAGTTGATTCTTGGAGGGCTCTCTCCACCACCAATACGGGAGATCTG  
GCCCCGCACTTGAGGAGGCTGGAGTCTCGGATCGCCCCTTCGCGTCGCCCTGGGCCCTG  
GGCCCTGGGGTGATGGGCCCCGTTGCCGTGGTGGATGGCAGGAGCTTTTCAGCTCTCAATG  
GGCGAATGCTACTCCGTAGGTGCGAGTGGCTGGAAGCGGCGGAACGGACAGGGGGAGGTT  
GGGGAAAATGCTCCGCAGGAAGAGCAGGGAGTGGGGAGCTGCGGTGCGGCCCTGTGGAGCC  
CGTGCAGGGCCAGCTAATCCAATTCGGGCCACAATAACAAGAGAGGGCCCCACATCATG  
TAAACAGAGGCTCAGAAGCTCCTGCCACACTGGGAGGGTTTCGAAGTCTGACGACTGCCA  
ATGGACCCCAGCCATCGCGAGCACACAGCAGTTCGCACGCTCCCATTTGGGTTCTCATCA  
CGCAGTCGCTCTCCCCGCCAACCAGCGCCAGGTCCGGGAACAGCGGCGCAAATGCGTATT  
TGAGGGCGCCTCGCTCGAGCAACCTGTGCCTGACCTTCTCCTCCTCCTTCTGCACCTTGC  
ATCTCGTCGCGTCCACTCGCAGGCAACCACACATCCTCCTCCTCTCCCAAACCCCCCG  
CTTTTCTTTCCCTTGTTGGAATTCGATTGAAAAAGAAGACGGGTCCGTCTAGAGACCGC  
CTTCTCACCTTTCTCTCGACTTCTTTCTAGGAAAAGAAGCAAGAGTCATTCTTGTCC  
ACCTTCTGGTTCACGGAAGGTGAGGAGAAGATTGCCTCTGCCCCAAAGTCGCCAACCT  
GGACTTTGAAGCACGTGTTCCGGTCCCTTTCAGTGTCTTCCCGTCCTCGTACAGGGAGTC

CGAGACCGCCACCCAAACCCACTCCCACGAAGAGGTTGAGATCAAGCTCCCCCAGCTCGC  
 CGGACGGGAAGGTCAACACTCTTCATTCCAAGCCCAAGCACATCTTCCTCCCAGCGGAGA  
 GGGTCGCTTCAGAGAAGAAGAGGTCCGCATCACTCGTCAAGAGGAACATCACCGCCGTCC  
 CGGCATCCGTGAAGAGTTCGTTACCGCGAGGAGCGTCACCGGTAAGTTTAGTTTTTGT  
 TTGATTACCAACCCATTGTCTTCCCCGCCTTTTTCTTTTTCTTCCCTTGCTCTCTTGCCC  
 CTGTCTAGTGTAGGGCATTGCCAAGGCCATCTTCACACACACACACCCCCCCCCCCCCC  
 ACCCTCAGCTGGGGGGGGGGTGGCCTGGGTGACCAAGGGACGGTGAAGACTACTACTA  
 CTTGAGCCACTCAAACCCATGCATGACACAGGGTTTTCTTTTTCTTTTCTTTTCCTT  
 TAACTAACCAACCACTCCAACATTAGCCCTCAGTCAACCTACTCCGAGTCTCGCATCGAG  
 TTCGATACTGAGCACCGCACTCACAACCTCCGTCATTGACGTTGCTGAGAGCGAGTATCGT  
 GCCCGTGTCCAGCCCAACTACCGCAAGGAAGCTTCCGTAGTCGGTACCACCGTCGACGGA  
 TCCCGCTTCAGCCACAGCCGCAAGGCCAGCAGCACCACCTCCACCCACACCGACGAGTAC  
 ACCGTCGATCCCCCTAGCCACCGCCCCGTCTACAAGAAGGAGTCGGTTGAAGTCGCCGGT  
 ACCACTGTTGACCCCCCTGCTCCTCGTTTCGACCTACCACGAGCAGGTGAACATTGTTGAA  
 GAGACCGTTGACGCTCACCGTTACGCTCCTCAACCCAACAACAACACACC**ATG**

**Pxyl6 (SEQ ID No 28):**

GCGGCCGCTTCCCCATGAATGGCAACCGGGCTGATGACCTGTGTGGGAAGAAATGGGGTT  
 GGGTCGGGCAATGGGAAGAAAACGGAAAGAGGGAAGGAAACATGCCTGTAGTCGAGGCTG  
 AGAGTGACGTACGTCCGTACATTCCAGTAACCAGGCGAGAATGAGCAATGATACCCCGC  
 ATTTCTTGATAATTAACCTCGTTCCAGAGCAGACTTACGCAGCACTACTCCGTACTGTT  
 GGAGCGCTTAGCACGCTGGAACTTGGCAGCCGTCCGAAGCCGCTCGGCCCCATCCTCTC  
 GCTGGTAGCTAGTGTAGTCCCGTGCTTTACAACGCGGCTATACAGCCCGTACAGTTGTAA  
 AGTACCTACATACATGCACTACTATTATTATCCTTCTAGAGTGGGTTCGAATTCCAGGG  
 AAGATCTTCCTATGGCTATCTGGCTGAAACTTGGGGGAGGAGTGCGGAAGGGGGGAGGGG  
 AACGAGCCTCCACATTGCATACGACCGGGGAATGCGGGACCCTAAGCGAACCAGGAACCC  
 GGTATTGCACTCGGAATTGCCGCAGATCCCTGCGTTCCACCCGCTCGAACGGTCAACAT  
 TAACTAATCTGTAGTGGAGTTACTGTTGACTTTCTGACTCGTGTCACTGGTCCTCGCCCA  
 AGTTCGAAAACAGAATTGCATTTTTGTCTTTTGTTCGGAGCTTTCGAGGAATAATTCCA  
 TTGTAGGTATGGAGTAATTATGGAGTATACACGGCCCAGGGGCGCTACACACACCATCGC  
 CGAGAATGGGAGGTCGAGCTCGCGACGCTCAGGATCCCATCGATATTTTCCCTTATCCCT  
 GCTCTCACTAGCGCGCAGAGCCGCCTCCGCGCGGGGATGCCGGTTGTTGCCGGCGTGCTT  
 TTTATCCGCTGCCCTTGGTTGCTCATTTCCTGGTTCTTGGGTCGCTTGCCAAGCAGCTCC  
 GGCGGAGAAGAATACCACAGGAGGGAGCATCGGGGCGCGAAGGGCATTGCACTATGCGGA  
 CGAGATGCTTCAACACCATCATGGACCTGTCCGGAACCTCCCAAGAACAGGCGACGCCAAG  
 GACGGAGTAGACCTCCCCCGTCCGTCTTCTCTCTGCCTGGCAATTTAGCCAAAATCCGA  
 CCCGACTTGCGACGATTCCCTACCTCCTAGCGCGTGCGCGCTGAAGCAGTCGCGAGAGTCG  
 CAAGGCATGGGCCCCGAGTCTGGCTGGCATCGTCAAACGTGATCGGCCCCGTCGAGCGTGCG  
 TGTATAAATGCATCAAGGAGCGACTGCCCCCCCATCAATAACCACCCGGTTGCTTGAGTC  
 TCTCGCACTCGCGGCCCTTCTTCTCTGCTTCGCACGCATCTCGCTGTCTCGCTGTCTCG

CTGTCTCACTGTCTCGCTGTCTCACTGTCTCGCTGTCTCACTGTCTCGCTGTCTCACTGT  
CTCACTCGTCCATCAGAGCAAAACC**ATG**

**Pgl1 (SEQ ID No 29):**

TAGTAGTTGTCAACCTTGGCAGCGAGAGTCCCGAGGCGGTAGATGAGAGAAAAAAGGACC  
GATGTTGACTTCCATGCCATCGATGGCGTCGTCTCGGCTAGACGTCGTCTGGCGTTATTCT  
GGGGGAGGCAATCCCGGGTGAGGAGAGAAATAGACGCGTCGCCATCTAGCAGCCATCACT  
CAGTGGCATCACCTGCGCGTTGACTTGCCCTTCGAAGGCTCTCCTGAGCCGAGCATGTGAT  
TACGATGTATAAGACCTGCATTGAGCTCGACGTTCCCGAGCGTCGGCGCGAGCTTCCAAT  
TCGGTTGAGGCTCCGGCGGCTTCCCCCGGTTTCCTGCTGGACTAGCTGCCGTGGCGGGGG  
GACGGCAGAGCGACTCCGACGCGCCCCATGCGAGCAACGGCCCGATTTTCGATGAGATCT  
GCGGGGCGCCGGAGTGGCAGCAGTTCGTCAGCTTGGCAGGCACGGCTCCCCACCTTCTTC  
CTTCTTCCACACTAGGCCCTCCCACAAGCGACCAGATGCTTGTTAAGTACGCAGTAGTGT  
CTCGGCTCGCCCAGAGAACAATGGCACGCCGATCTGTCTAATGACCAAGAGCCACGGTTC  
GAGACCATCCATTGGACTGGAGGGCCTGCGAGGCATCACGCCGAACCCATGTCATGCTAC  
TCTTTCTGTTCACCCCCGGAGATGGCGTGAACTGCGCGTTTACTCGCGGCTCAGCATGT  
GCTCACGTTGGGTAGGTCCCGCAAAGTCAGAGGTAGGGAGGTACTTTGTAGGCACAAATC  
ATGTACACGTTTCGTACCTGAGGTAGCTATCTCGCCTCAGGCACACGAGGCCCGTTTCGACG  
AGAGAGAGGAAGAGCAACCAAGAATAGTCAAGGATATTATTACTCTTTCCCTGGTATTTC  
TGGACATTTTGTCCAGGATTTTGTTCGCCCTTTAATTTTGAACAATTATGCTCCCGTCGG  
CTCCGATCCACGCCTCTTAACCTCTCCTTTAGCCTTTCGCCTCTATTTCTTGAATTTCAA  
TTCTCCCAAGGGCCCTGCTTTCTACAGCAAAGAATCCGTACCCTACTCTCTTTTCGCGCAC  
AGAGTGAGGGAGCAACAGGGATTGCGAAATGCACAGCAGAGTTTGTGTAACCTTCGGCAGC  
TCTTCCCCACATTCAGATGCATGTTACTGGAGAATGCGGAGAAGTTATAGTCTGGGGTAG  
TAGGTATAACGCTGGTACTCCCGAGGTAGGTAGCAACCTTGGCTGACCTTGGGAAGCGAG  
GGCGCTTGTGACGCTGACGATCCAGAAGCAGCCCGCCGATAGTATACGTGGAGACGGTGC  
TTCTTGCTATAAGCGCTCAACTCCGCTACCCATGTTTACCGTCTTCCCCTTGGACGACGG  
CATCACTCCGATACCCATGTCTCCTGGGTAGCTCCGAGTAGTCGCCCCGAGCGCCCTTCTC  
CCCCCTCCCCCTTTCTCCTAATAAACGGCCGAGTCGGGCAGCCTCGACGTTGCACCGTAG  
CGTCGCAGCCTGCGTAGAAGCACGCGTAGAAGCACCGAGCTCCAAGCTCCAAGACGCCAA  
AAGCCGCCGCGAAGTGGCCGTGGGCCCTTCCCCGCATGCGCAGCTCCGGCACCAAGTCCG  
AAACGCTCCATCACCCCATATCCCAGTCAGAACAGCGGCTGCTTTCCGGATTTGGAAGTC  
TGGAGGTCGCGAATGAAGGCTCGCGTTCGACTATAATAACAGCTCCGGATGGCAGGCCTC  
GTTGCCAGCTCCAGGACCACCTCCCATCCGTAAACGGATCTGGCCTCGTCACGCCCGCC  
**ATG**

**Alp1 DNA sequence comprises (SEQ ID No 30):**

ATGCACTTCTCCACCGCTCTCCTGGCCTTCTGCCCCGCCGCCCTCGCGGCCCTACTGCCGAGACCTCGAC  
AAGCGCGCCCCGATCCTGACTGCTCGCGCTGGCCAGGTCGTCCCGGGCAAGTACATCATCAAGCTCCGCGAC  
GGAGCCAGCGACGATGTCCTTGAGGCCGCCATCGGCAAGCTCCGCTCCAAGGCCGACCAGTCTACCGCGGC  
AAGTTCAGSGGCTTTGCCGGCAAGCTCGAGGATGACGTCCTTGACGCCATCCGTCTTCTCCCCGAgtgagt  
cgcggtcccggaagaaatagagcgagcgggggagagagtggaagggcgaaaagagccggtgttttggttaaccg

cttgtcttttctttctctttgcaatagGTCGAGTACGTCGAGGAGGAGGCCATCTTCACCATCAACGCGTA  
CACCTCGCAGTCCAACGCCCCCTGGGGCCTTGCGCGCCTCTCGTCCAAGACCGCGGGCTCCACCACCTACAC  
CTACGACACCAGCGCCGGCGAGGGCACCTGTGCCTATGTGATCGACACGGGCATCTACACTAGCCACTCCgt  
atgtctcgcggttacctccccctttcggaagaaggggcatccatatgctgacccctcctgatcacagGACTTC  
GGCGGCCGTGCCACTTTCGCGCCAACTTCGTCGACAGCTCTAACACCGATGGCAACGEGCCACGGCACCCAC  
GTGCGCGGCACCATCGGCGGCACCACGTACGGTGTGCCAAGAAGACCAAGCTCTACGCCGTCAAGGTTCTC  
GGCTCCGACGGCTCTGGCACCAACgtatgectcgaccccgcgaccccgacacccgccccggcggttatcttct  
gactgacattcctctttctcctctctagTTCTGGTGTGATTGCTGGCATCAACTTCGTCGCTGACGACGCGC  
CCAAGCGCAGCTGCCCCAAGGGCGTCGTCGCCAACATGTGCTCGGCGGTAGCTACTCGGCCTCCATCAACA  
ACGCCGCCGCCGCCCTCGTCAGGTCGGGCGTCTTCCTGGCCGTGCGCGCCGGCAACGAGAACCAGAACGCCG  
CCAACCTCGTCGCCCGCCTCCGAGGCGTCCGCCTGCACCGTCGGCGCCACCGACAGGAACGACGCCAAGGCCA  
GCTACTCCAACCTACGGCAGCGTCGTCGATATCCAGGCCCCCGGCTCCAACATCCTGAGCACCTGGATCGGCA  
GCACCTCTGCTACCgtaagccccccctccccccacccacccccagcctttggcgacattcccgccccgtatt  
tatttctccggggtgggggagaaacaaaacaaaatagctaacatgagatgcactctcagAACACCATCTCGG  
GTACCTCGATGGCCTCCCCCACATTGCCGGCCTCGGTGCCTACCTCCTGGCCCTCGAGGGCTCCAAGACCC  
CTGCCGAGCTCTGCAACTACATCAAGTCGACCGGCAACGCCGCCATCACTGGCGTTCCCAGCGGCACCACCA  
ACCGCATCGCCTTCAACGGCAACCCCTCTGCCTga

#### Alp1 amino acid sequence (SEQ ID No 31):

1 MHFSTALLAF LPAALAAPTA ETLDKRAPIL TARAGQVVPK KYIIKLRDGA SDDVLEAAIG  
61 KLRSKADHVV RGKFRGFAGK LEDDVLDAIR LLPEVEYVEE EAIFTINAYT SQSNAPWGLA  
121 RLSSKTAGST TYTYDTSAGE GTCAYVIDTG IYTSHSDFGG RATFAANFVD SSNTDGNNGH  
181 THVAGTIGGT TYGVAKKTKL YAVKVLGSDG SGTTSGVIAG INFVADDAPK RSCPKGVVAN  
241 MSLGGSYSAS INNAAAALVR SGVFLAVAAG NENQNAANSS PASEASACTV GATDRNDAKA  
301 SYSNYGSVVD IQAPGSNILS TWIGSTSATN TISGTSMA SP HIAGLGAYLL ALEGSKTPAE  
361 LCNYIKSTGN AAITGVPSGT TNRIAFNGNP SA

#### List of references

[0089]

Braaksma, M. and P. J. Punt. 2008. *Aspergillus* as a cell factory for protein production: controlling protease activity in fungal production, p. 441-455. In G. H. Goldman and S. A. Osmani (ed.), *The Aspergilli: Genomics, Medical Aspects, Biotechnology, and Research Methods*, CRC Press, Boca Raton.

Verdoes, J.C., Punt, P.J., Burlingame, R., Bartels, J., van Dijk, R., Slump, E., Meens, M., Joosten, R. and Emalfarb, M., 2007. A dedicated vector for efficient library construction and high throughput screening in the hyphal fungus *Chrysosporium lucknowense*. *Industrial Biotechnology* 3, 48-57.

# REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

## Patent documents cited in the description

- [WO0020555A](#) [0004]
- [WO20090918537A](#) [0008]
- [WO2000020555A](#) [0010]
- [WO2009018537A](#) [0044] [0050] [0067] [0071] [0087] [0088] [0088] [0088]
- [WO2009033071A](#) [0045] [0048] [0059] [0069]

## Non-patent literature cited in the description

- **HANS VISSER et al.**J. of Biotechnology, 2007, S211-S241 [0002]
- Aspergillus as a cell factory for protein production: controlling protease activity in fungal production  
**BRAAKSMA, M.P. J. PUNT.**The Aspergilli: Genomics, Medical Aspects. Biotechnology, and Research MethodsCRC Press20080000441-455 [0089]
- **VERDOES, J.C.PUNT, P.J.BURLINGAME, R.BARTELS, J.VAN DIJK, R.SLUMP, E.MEENS, M.JOOSTEN, R.EMALFARB, M.**A dedicated vector for efficient library construction and high throughput screening in the hyphal fungus Chrysosporium lucknowenseIndustrial Biotechnology, 2007, vol. 3, 48-57 [0089]

SYSTEM TIL CHRYSOSPORIUM LUCKNOWENSE-PROTEINFREMSTILLING

PATENTKRAV

1. Kimært gen omfattende en promotor, der omfatter nukleotidsekvensen ifølge SEQ ID NO: 25.
2. Kimært gen følge krav 1, hvori promotoren omfatter nukleotidsekvensen ifølge SEQ ID NO: 26.
- 5 3. Værtscelle, der omfatter det kimære gen ifølge krav 1 eller 2.
4. Værtscelle ifølge krav 3, hvori værtscellen er en *Chrysosporium lucknowense*-celle.
5. Fremgangsmåde til ekspresion af det kimære gen ifølge krav 1 eller 2 i en *Chrysosporium lucknowense*-celle, hvorved genproduktet af det kimære gen fremstilles.

## DRAWINGS

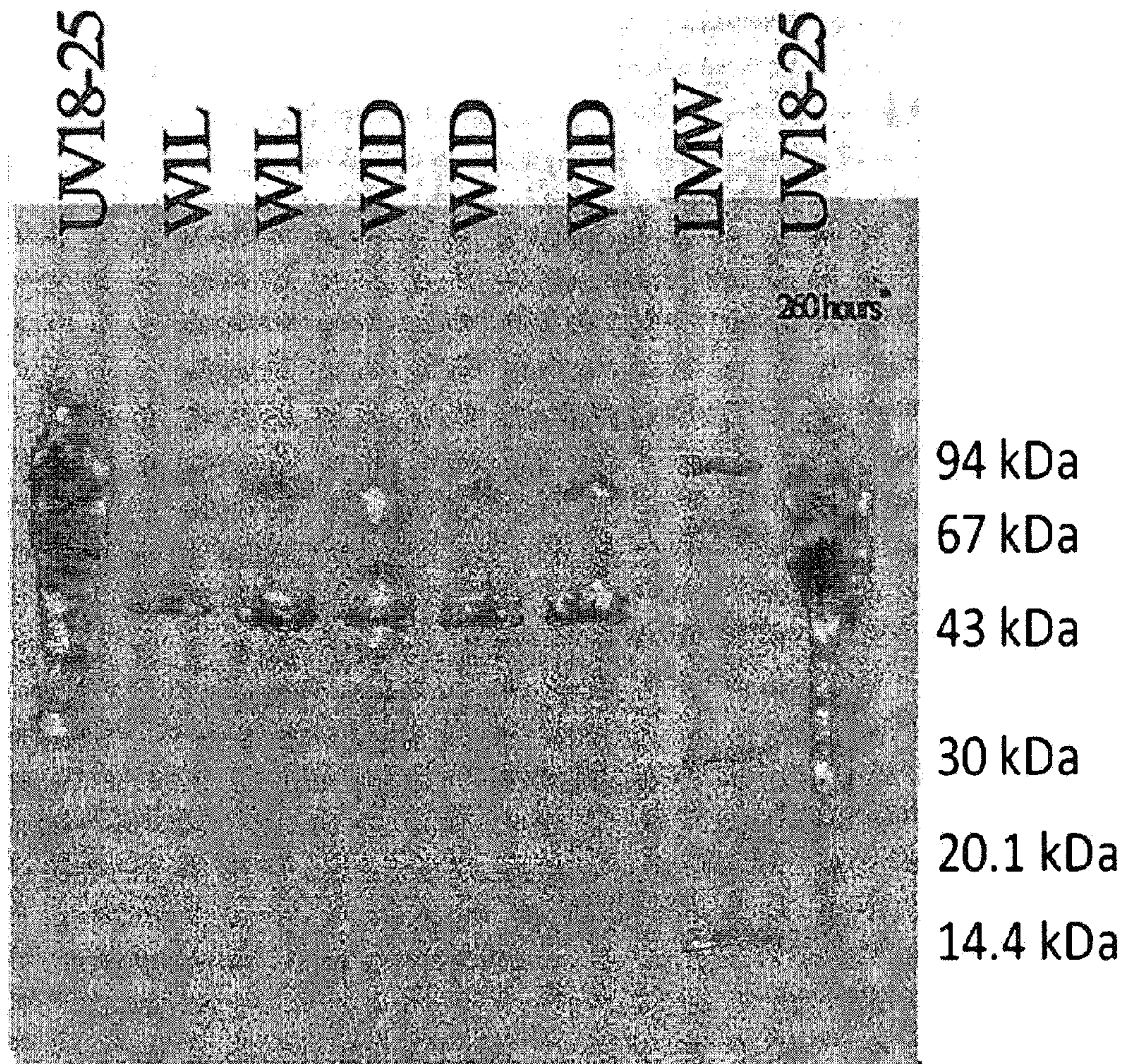


Fig. 1

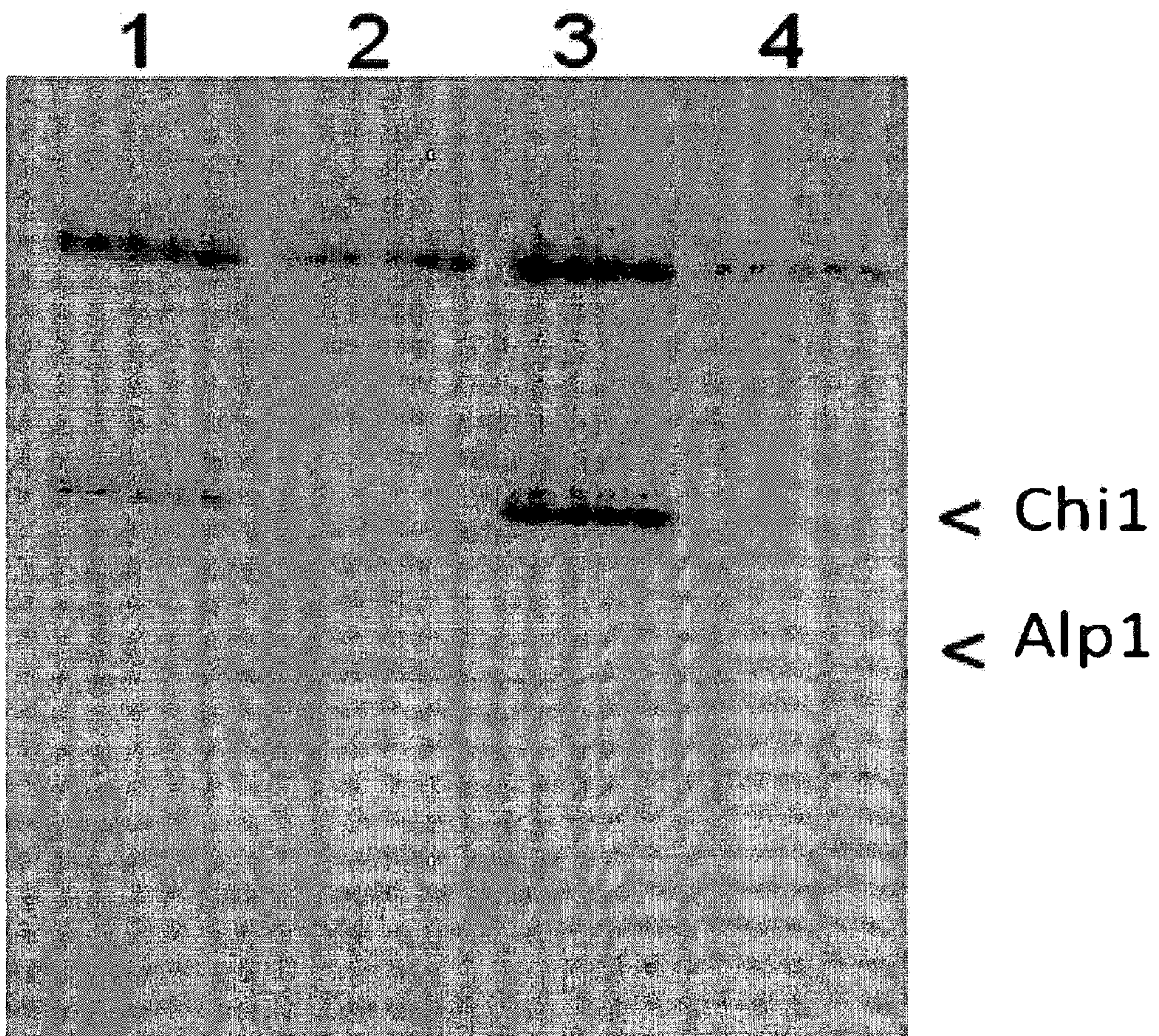


Fig. 2

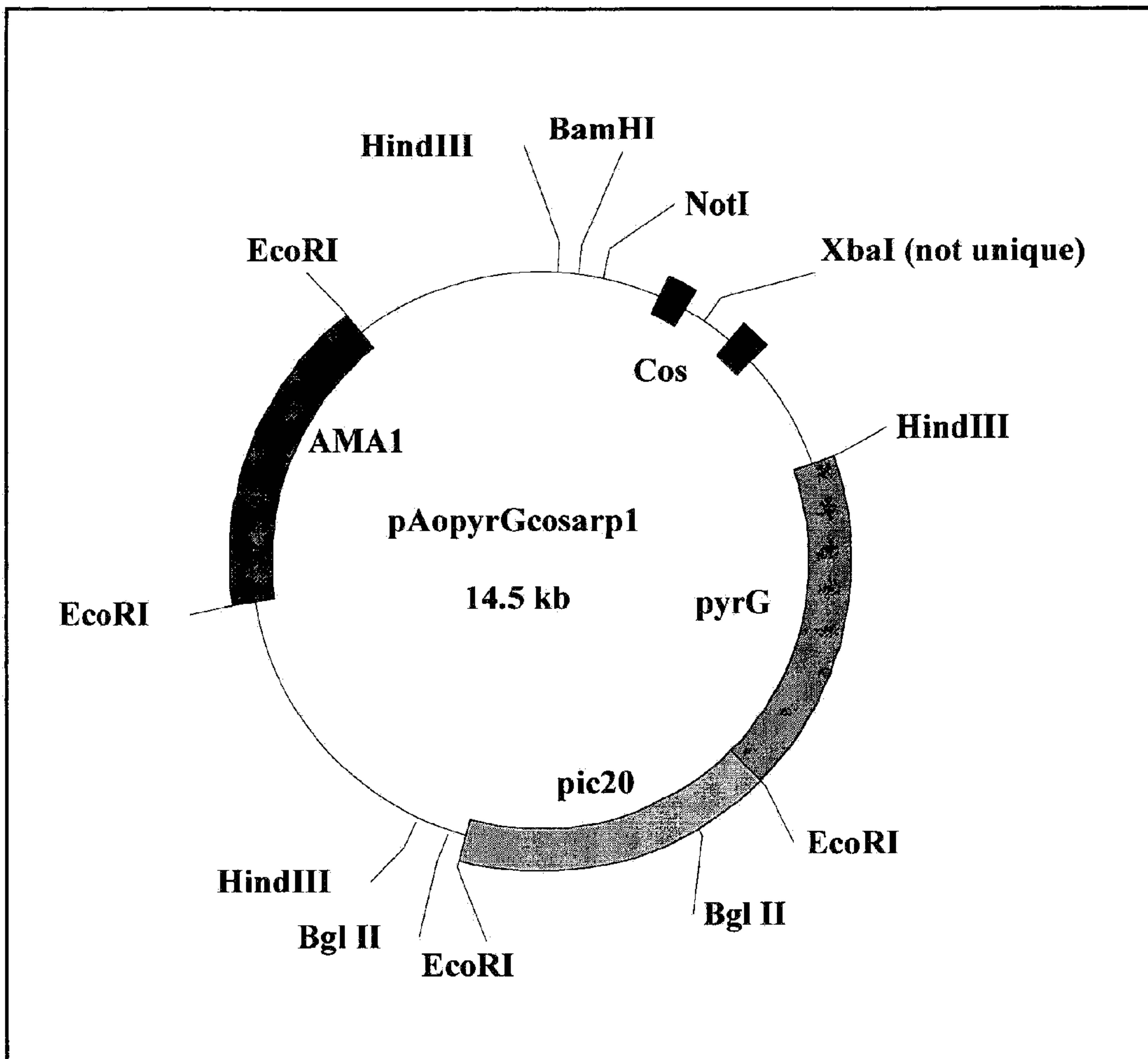


Fig. 3

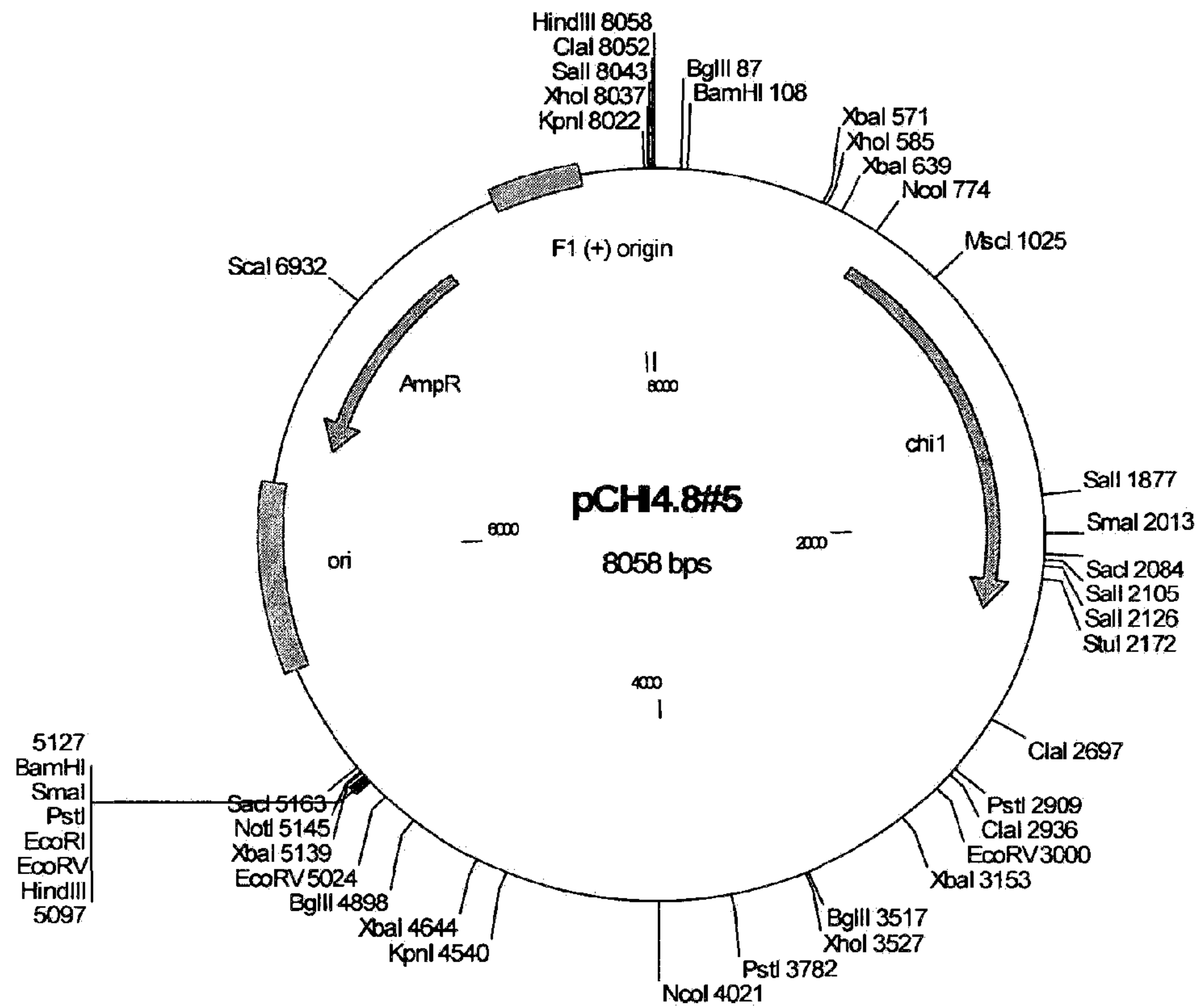
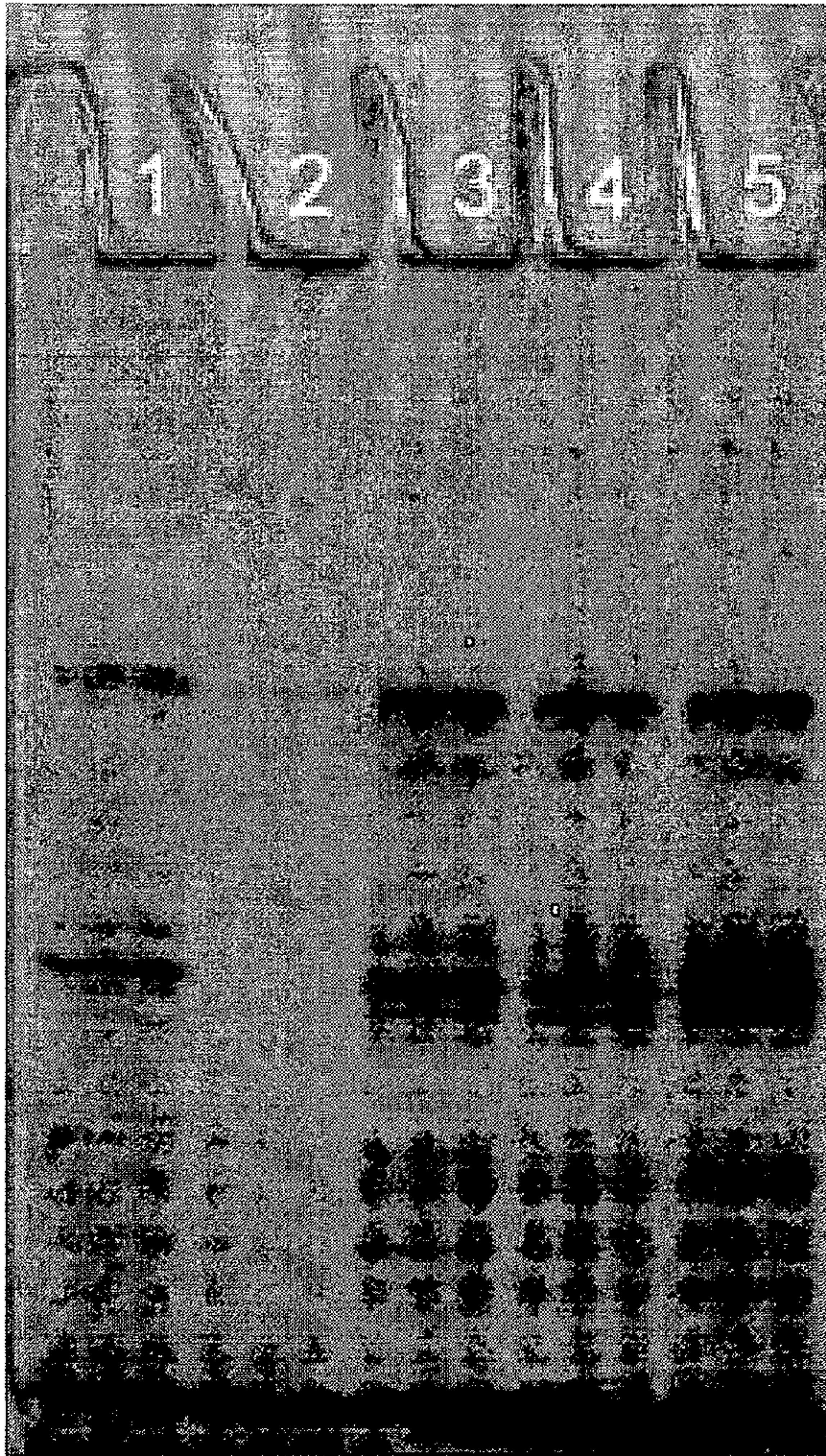


Fig. 4



← Chi1

Fig. 5

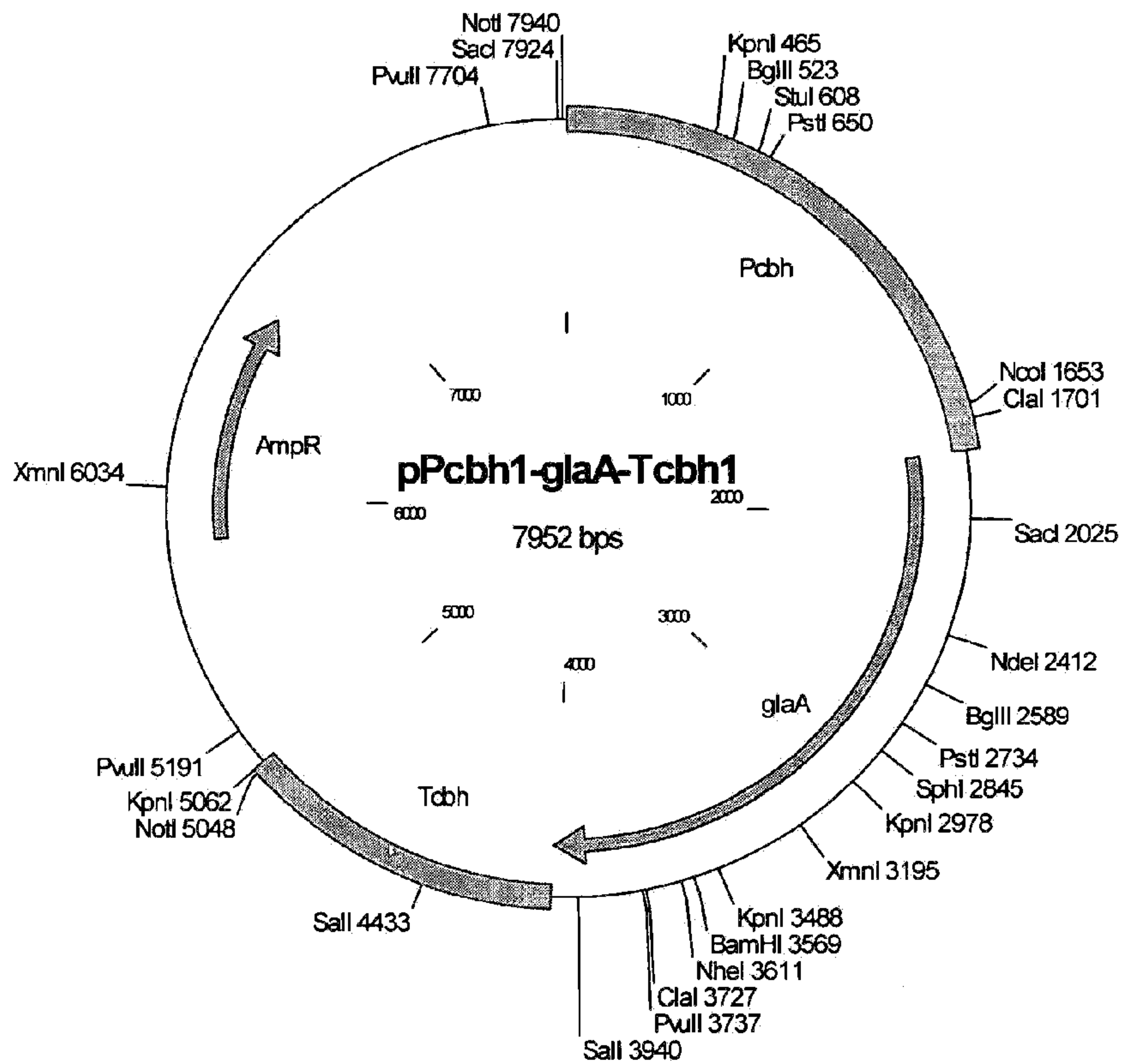


Fig. 6

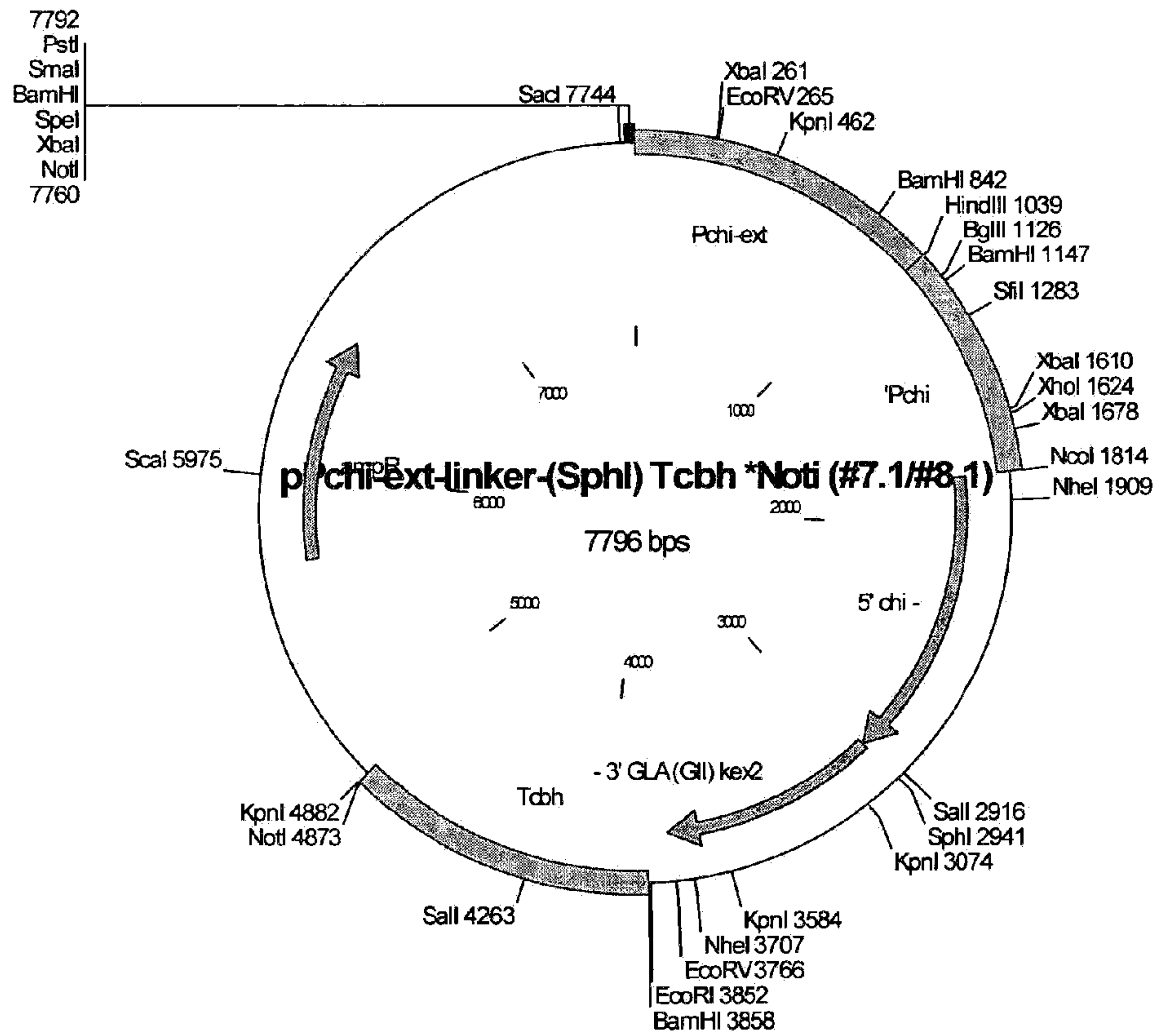


Fig. 7

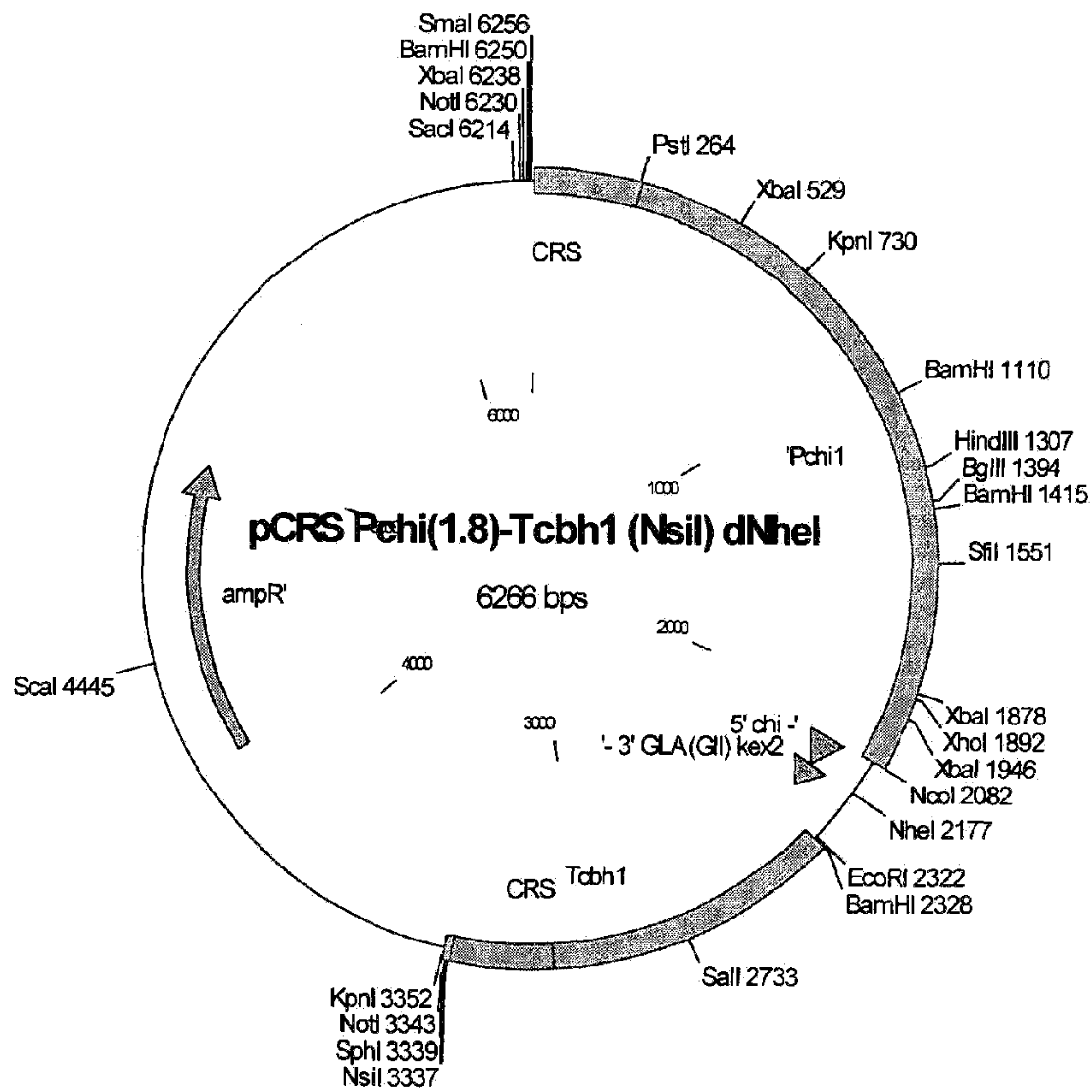


Fig. 8

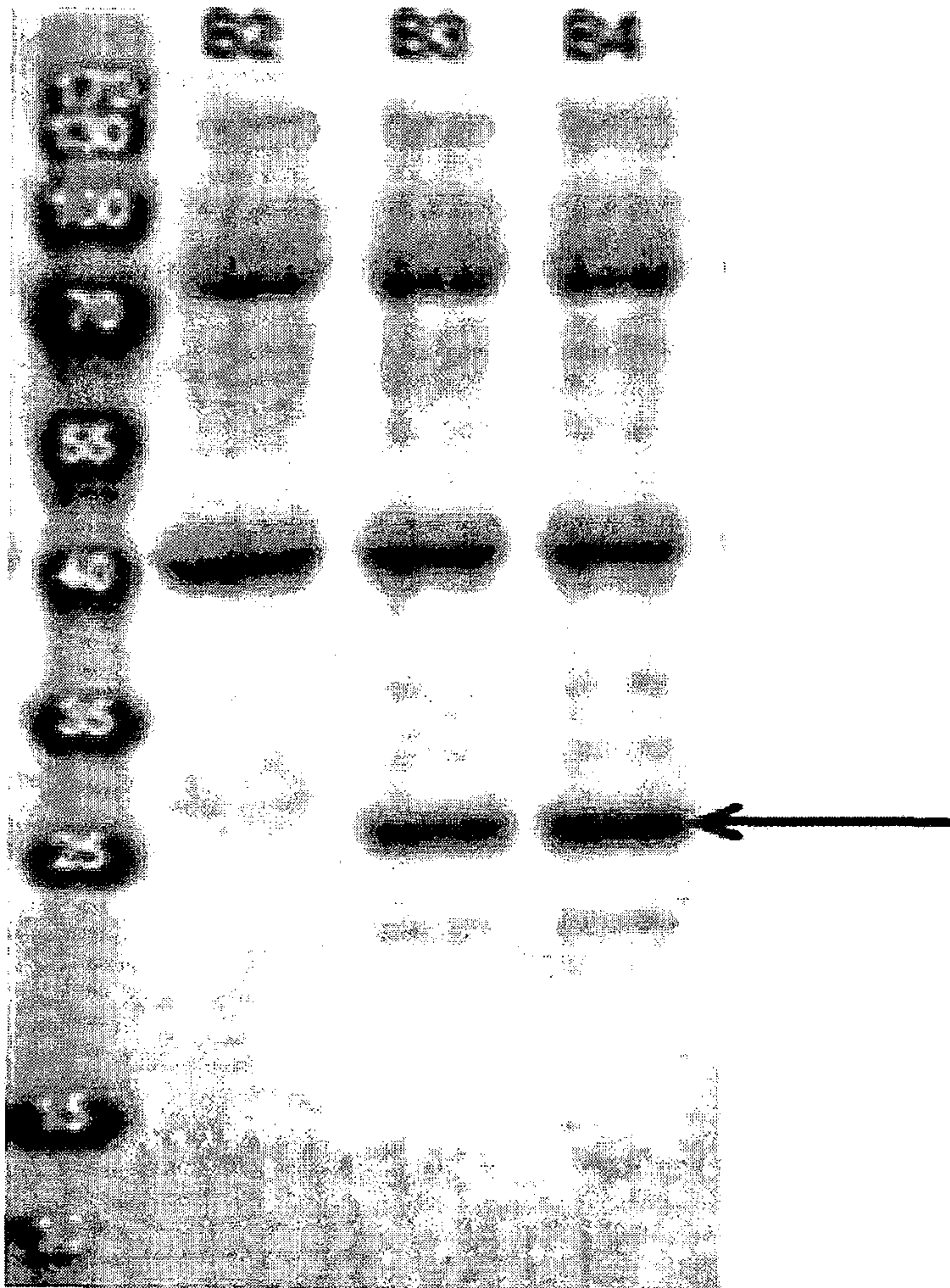


Fig. 9

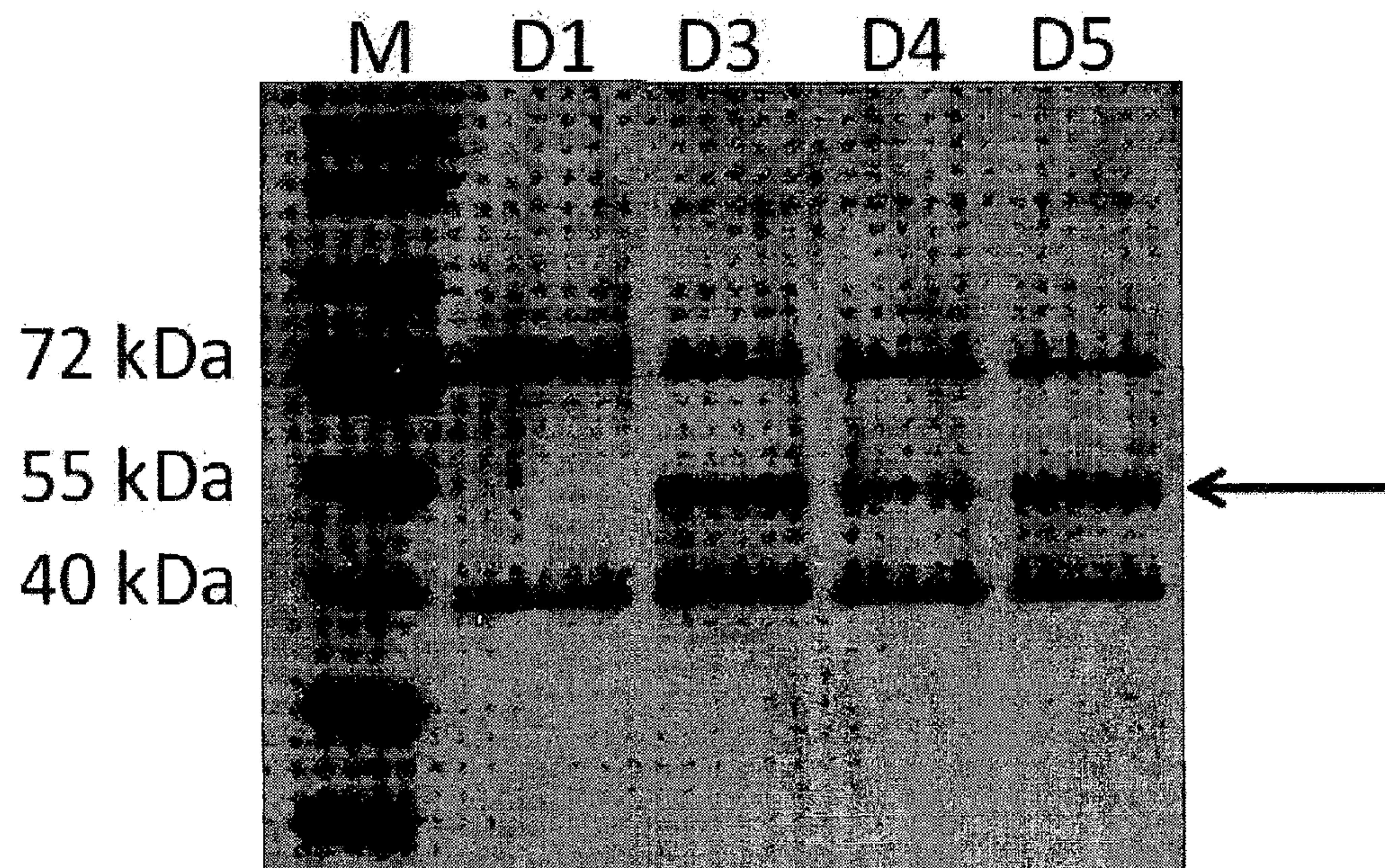


Fig. 10

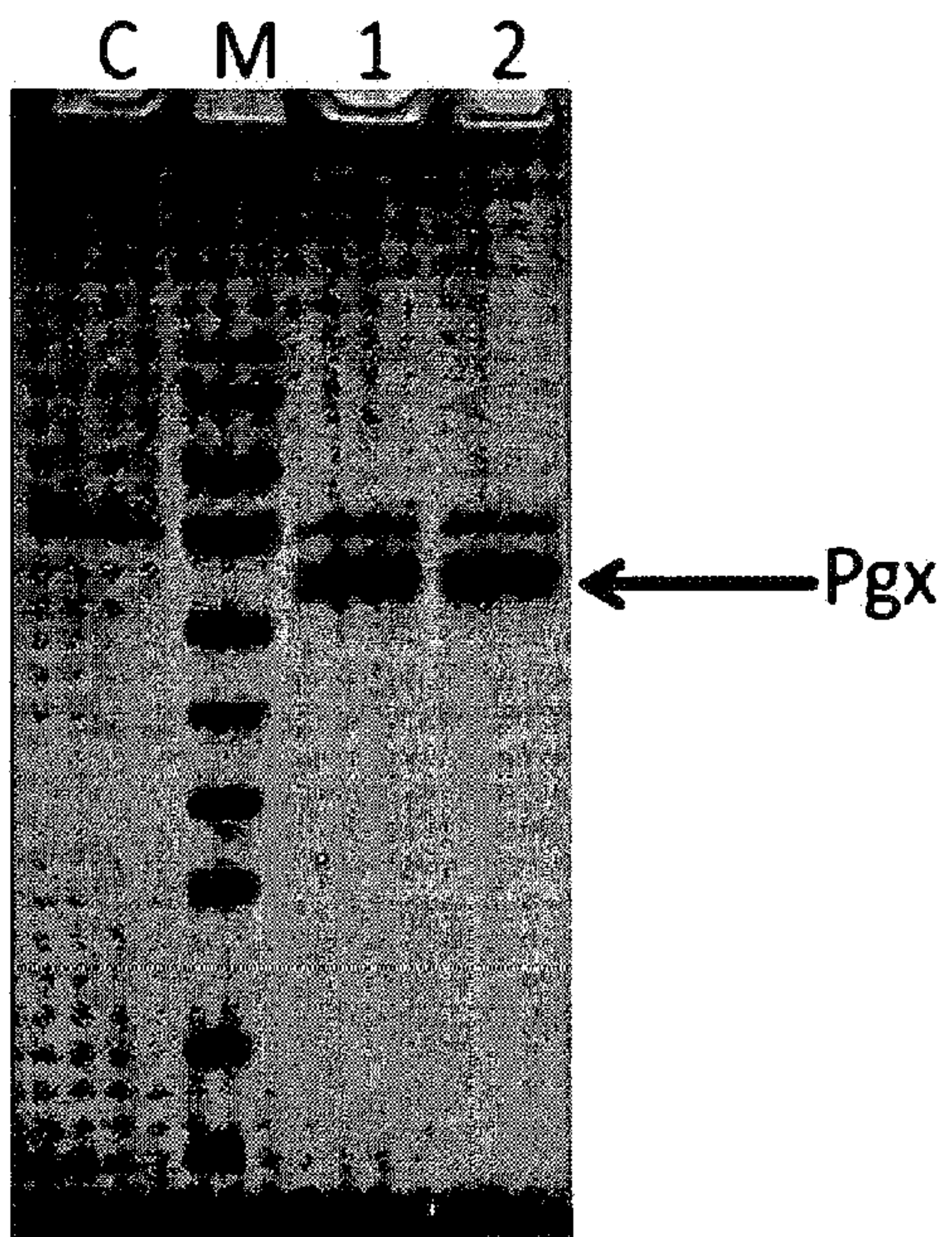


Fig. 11

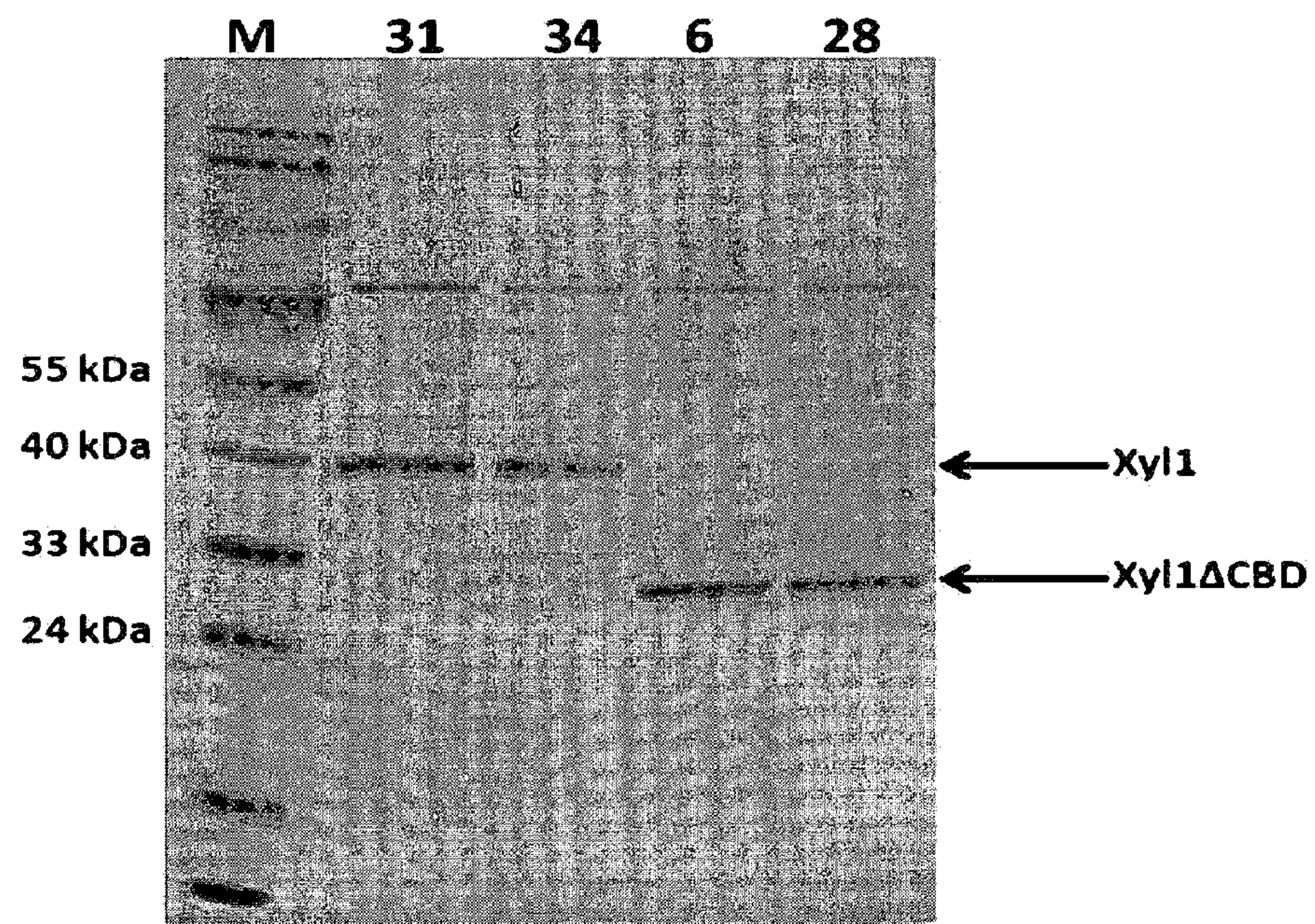


Fig. 12

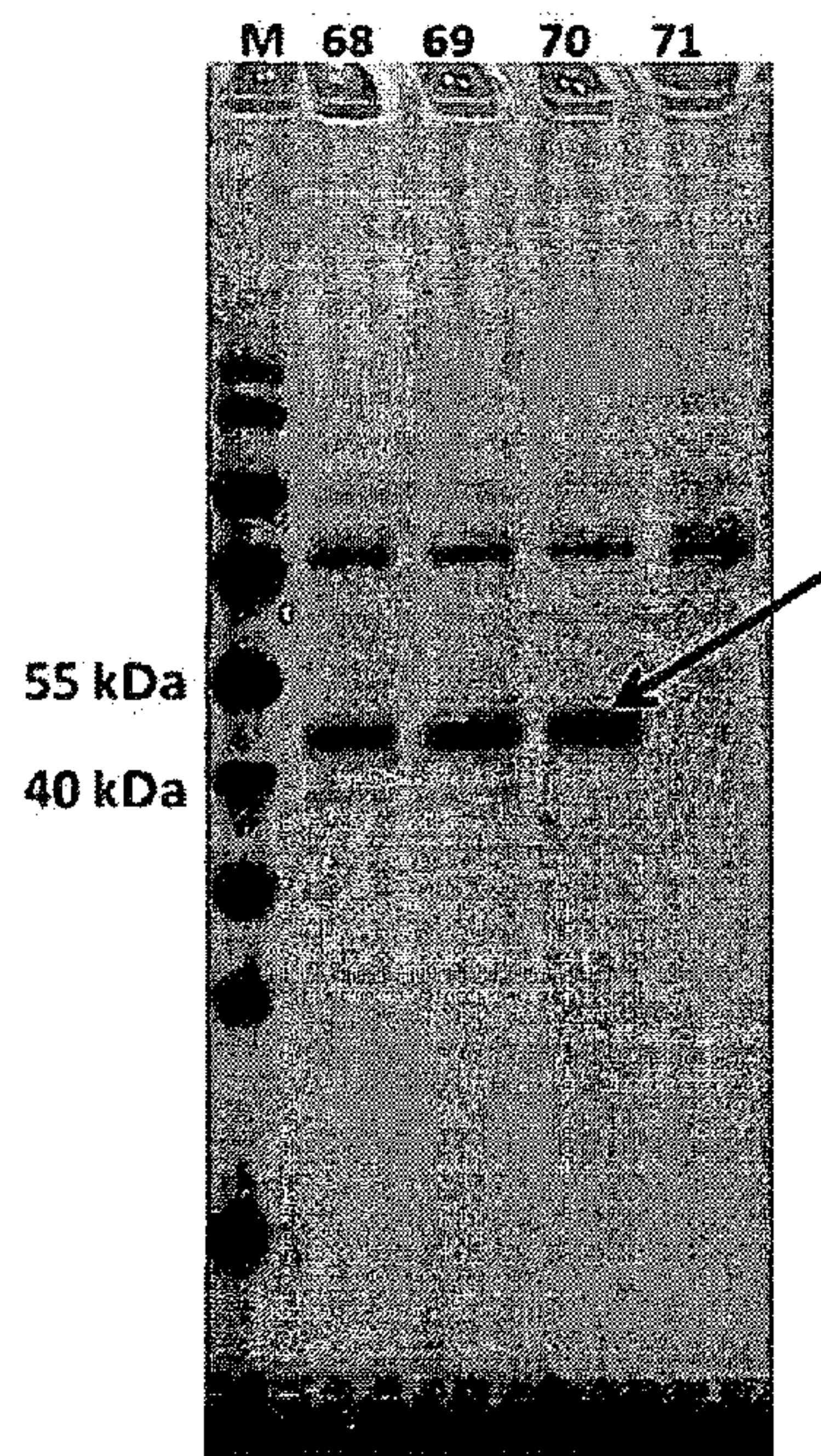


Fig. 13

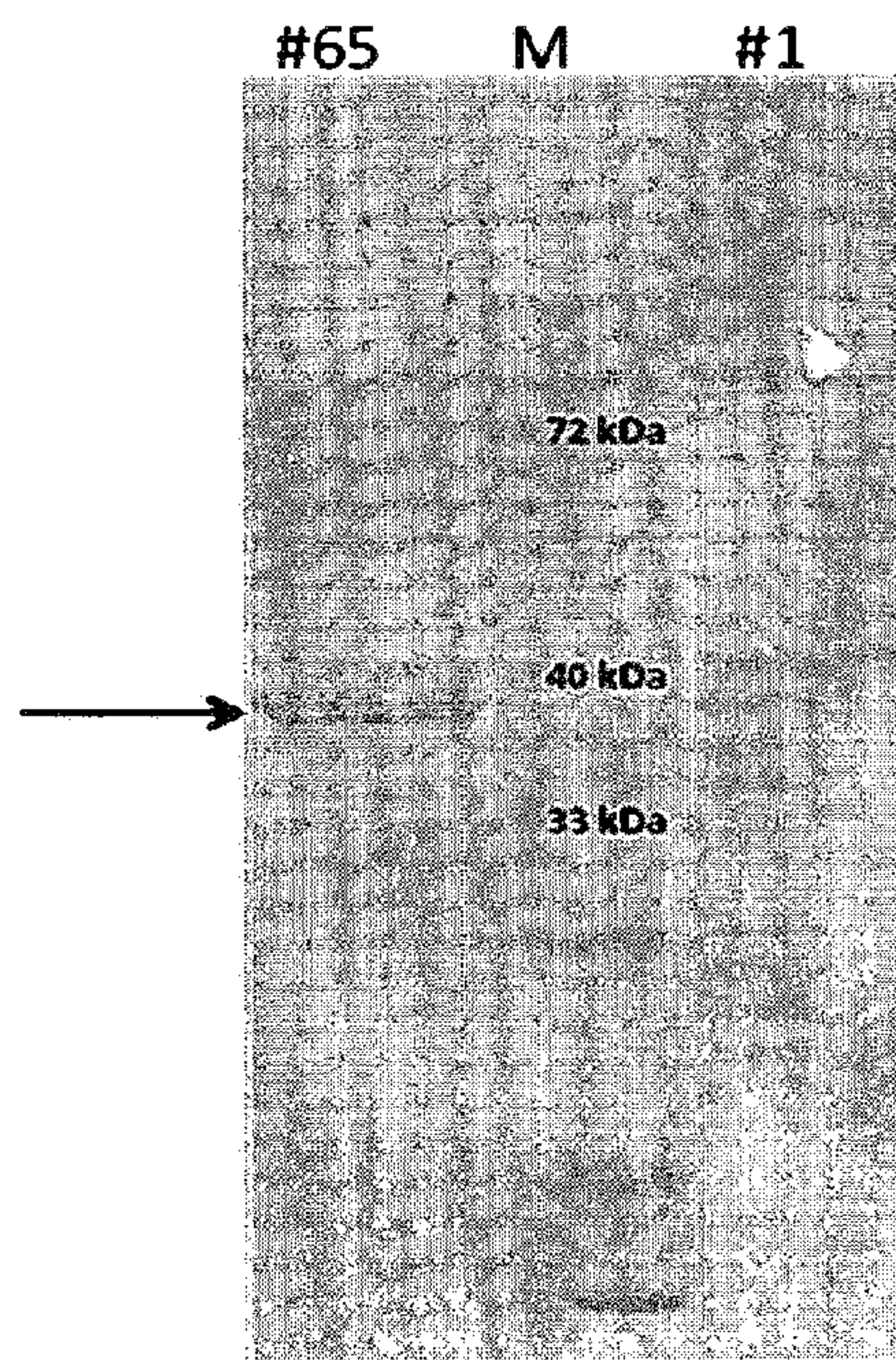


Fig. 14

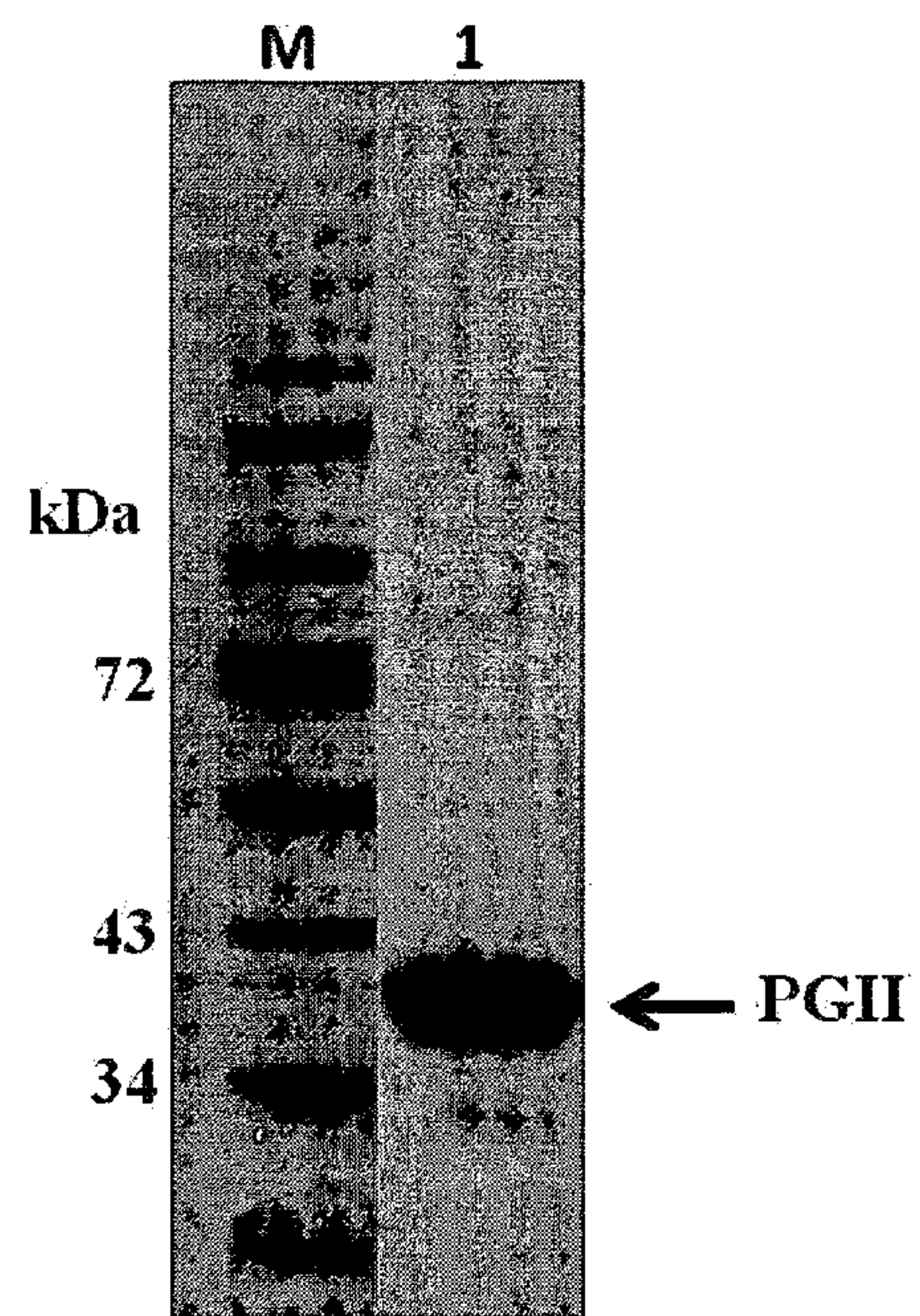


Fig.15