Fig.3. Evolutionary relationships among the 3 hyaluronan synthases. The phylogenetic tree was generated by the Higgins-Sharp (Higgins and Sharp, 1988) algorithm using the DNASIS multiple alignment program. The program takes as input a dendrogram produced by applying the UPGMA method (Sneath and Sokai, 1973) to a matrix of similarity scores for all of the aligned sequences. The calculated matching percentages are indicated at each branch point of the dendrogram.

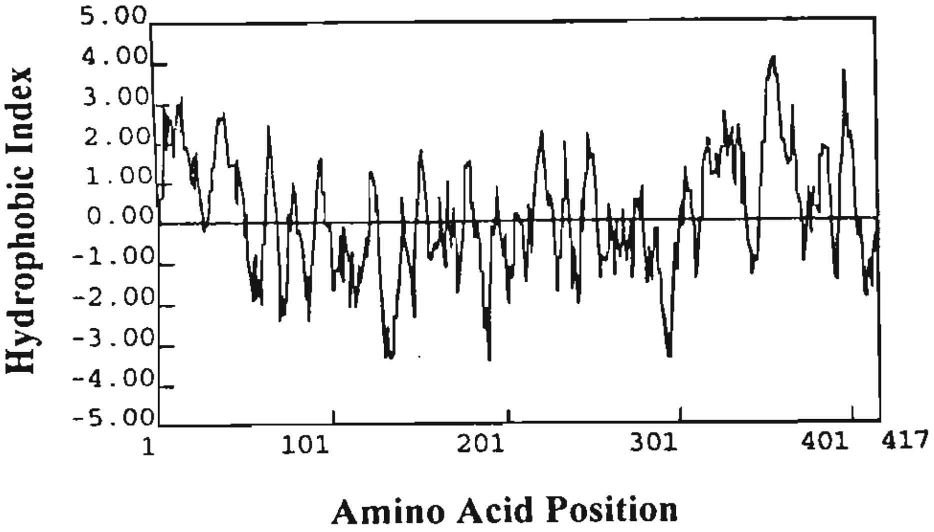


Fig. 4. Hydropathic analysis of the deduced amino acid sequence of HAS from *S. zooepidemicus*. The hydropathic indices are on the Y-axis and amino acid numbers for HAS (1-417) are on the X-axis. The hydrophobicity values were obtained according to the algorithm of Kyte and Doolittle (Kyte and Doolittle,1982). Positive values represent increased hydrophobicity.

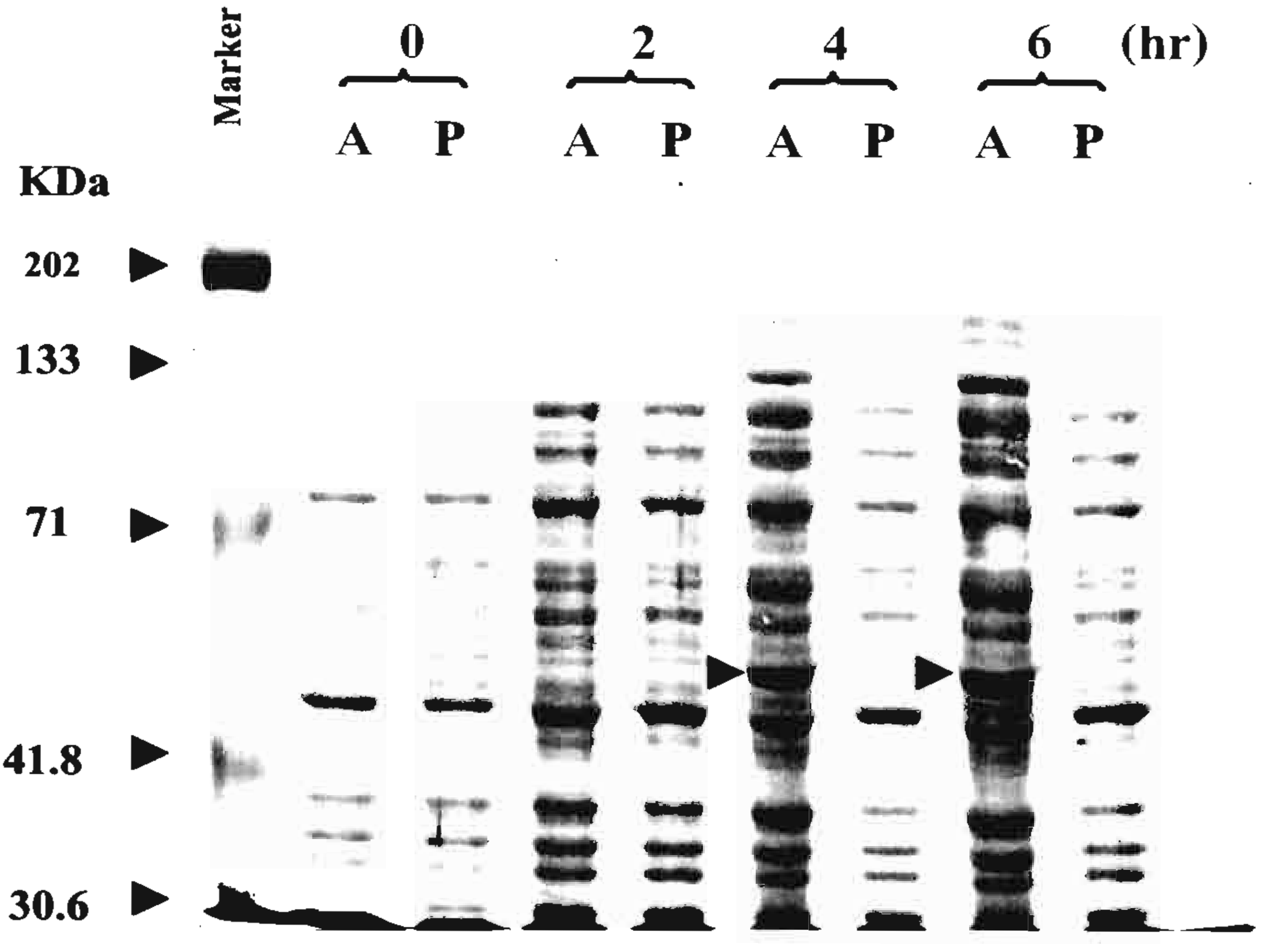


Fig. 5. SDS-PAGE analysis of overexpressed streptococcal HAS in *E. coli* HMS 174(DE3) pLysS at induction time 0, 2, 4 and 6 hrs. The gel was stained with Coomassie Blue R-250. The position of HAS is marked by the arrow.

lane A : recombinant pHAS-r (pET-28a + modified *has*A)
lane P : vector alone (pET-28a)

เอกสารอ้างอิง

Brinkmann, U., Matles, R.E. and Buekel, P (1989). High-level expression of recombinant genes in *Escherichia coli* is dependent on the availability of dna Y gene product. *Gene.* **85**, 109-114.

Caparon, M.G. and Scott, J.R. (1991) Genetic manipulation of pathogenic Streptococci. *Methods Enzymol.* **204**: 556-586.

Chen GF, Inouye M. (1990). Suppression of the negative effect of minor arginine codons on gene expression; preferential usage of minor codons within the first 25 codons of the *Escherichia coli* genes. *Nucleic Acids Res.* **18**(6):1465-1473.

DeAngelis, P.L. and Achyuthan, A. M. (1996) Yeast-derived recombinant DG42 protein of Xenopus can synthesize hyaluronan in vitro. *J. Biol. Chem.* **271**: 23657-23660.

DeAngelis, P.L. and Weigel, P.H. (1994) Immunochemical confirmation of the primary structure of Streptococcal hyaluronan synthase and synthesis of high molecular weight product by the recombinant enzyme. *Biochemistry* **33**: 9033-9039.

Dische, Z. (1947). A new specific color reaction of hyaluronic acids. *J. Biol. Chem.* **164**, 189-198.

Higgins, D.G. and Sharp, P.M. (1988) CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene **73**(1): 237-244.

Komine Y, Adachi T, Inokuchi H, Ozeki H. (1990). Genomic organization and physical mapping of the transfer RNA genes in *Escherichia coli* K12. *J Mol Biol.* **212**(4):579-698.

Kyte, J., and Doolittle, R.F. (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**: 105-132.

Nagahashi, S., Sudoh, M., Ono,N., Sawada, R.,Yamaguchi, E., Uchida, Y., Mio, T., Takagi, M., Arisawa, M. and Yamada-Okabe, H. (1995) Characterization of chitin synthase 2 of *Saccharomyces cerevisiae*. Implication of two highly conserved domains as possible catalytic sites. *J. Biol. Chem.* **270**: 13961-13967.

Sanger, F., Nicklen, S. and Coulson A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Ncad. Sci. USA.* **74**: 5463-5467.

Sneath, P.H.A. and Sokai, R.R. (1973) Numerical Taxonomy, W.H. Freeman & Co., San Francisco.

Sorensen, M.A, Kurland, C.G. and Pedersen S. (1989). Codon usage determines translation rate in *Escherichia coli*. *J. Mol. Biol.* **207**(2):365-377.

Weigel, P.H., Hascall, V.C., and Tammi, M. (1997) Hyaluronan synthase. *J. Biol. Chem.* **272**: 13997-14000.

Zhang, S.P., Zubay, G.and Goldmam, E. (1991). Low-usage codons in *Escherichia coli*, yeast, fruit fly and primates. *Gene.***105**(1):61-72.

Output จากโครงการวิจัยที่ได้รับทุนจาก สกว

1. ผลงานตีพิมพ์ในวารสารวิชาการนานาชาติ

Jongsareejit, B., Pewnim, T., Bhumiratana, A., Masaaki, M. and Kanaya, S. (2003). Molecular Cloning and Sequence of the Hyaluronan Synthase (HAS) Gene from Group C *Streptococcus zooepidemicus.* (manuscript submitted to FEMS Microbiol. Lett.)

2. การนำผลงานไปใช้ประโยชน์

ในเชิงวิชาการ : คือ มีการเริ่มต้นงานวิจัยและพัฒนาห้องปฏิบัติการวิจัยทางด้านอณูพันธุศาสตร์ขึ้นมาอย่างจริงจังในภาควิชาชีววิทยา คณะวิทยาศาสตร์ จนมีความพร้อมในระดับหนึ่งในการทำงานวิจัยด้านนี้ ซึ่งนำไปสู่การพัฒนาการเรียนการสอนวิชาพันธุศาสตร์ของจุลินทรีย์ในภาควิชาด้วย นอกจากนั้นยังเป็นส่วนสำคัญที่ผลักดันให้ผู้วิจัยดำเนินการวิจัยในสาขาที่ศึกษามา ซึ่งเท่ากับเป็นการสร้างนักวิจัยใหม่ รวมทั้งมีการติดต่อแลกเปลี่ยนความรู้กับผู้วิจัยในประเทศญี่ปุ่น และประเทศออสเตรเลียด้วย ซึ่งหากได้รับการสนับสนุนต่อไปก็คาดว่าจะนำไปสู่การสร้างเครือข่ายความร่วมมือกับต่างประเทศต่อไปได้

ในเชิงพาณิชย์ : โดยหากมีการศึกษาวิจัยอย่างต่อเนื่องต่อไป คาดว่าจะสามารถผลิตกรดไฮยาลูโรนิกเป็นการค้าได้ ซึ่งจะเป็นการลดการเสียดุลการค้าจากการนำเข้าผลิตภัณฑ์ดังกล่าวจากต่างประเทศได้

3. การเสนอผลงานในที่ประชุมวิชาการ

Jongsareejit, B., Bhumiratana, A., Masaaki, M. and Kanaya, S. (2001). Molecular Cloning and Sequence of the Hyaluronan Synthase (HAS) Gene from Group C *Streptococcus zooepidemicus.* In the Proceeding of BioThailand 2001 (From Research to Market), Gueen Sirikit National Convention Center, Bangkok, Thailand.

ภาคผนวก

บทความวิจัยสำหรับการเผยแพร่

Molecular Cloning and Sequence of the Hyaluronan Synthase (HAS) Gene from Group C *Streptococcus zooepidemicus*

Boonsri Jongsareejit(1), **Thanit Pewnim**(2), **Amaret Bhumiratana**(3)
Masaaki Morikawa(4), **Shigenori Kanaya**(4)

(1) *Department of Biology, Faculty of Science, Silpakorn University, Nakorn Pathom 73000, Thailand*
(2) *Department of Chemistry, Faculty of Science, Silpakorn University, Nakorn Pathom 73000, Thailand*
(3) *Department of Biotechnology, Faculty of Science, Mahidol University, Bangkok 10400, Thailand*
(4) *Department of Material & Life Science, Graduate school of Engineering, Osaka University, Osaka 565-0871, Japan*

The cloning of hyaluronan synthase (HAS) gene of *Streptococcus zooepidemicus* (group C Streptococcus, GCS) was performed by the strategy of the polymerase chain reaction (PCR), using the design of degenerate oligonucleotides based upon conserved sequences from other organisms. The PCR product was isolated, and then characterized by DNA sequencing. The nucleotide sequence was determined and the deduced amino acid sequence was then compared with known HAS sequences. The amino acid sequence showed high similarity to other known HAS. In order to obtain the entire *has*A sequence, gene walking was performed. Sequence analysis indicated that *has*A encode a protein; hyaluronan synthase, designated szHAS, 417-amino acids protein (calculated molecular mass, 47.77 kDa). The protein has two putative transmembrane domains at the amino terminus and 2-3 transmembrane domains at the carboxyl end. Streptococcus DNA fragment containing *has*A was amplified using PCR. The amplified fragment was incorporated into expression vector pET-28a (Invitrogen, USA). The pHAS, bearing the *has*A has been transformed into *E. coli* BL21 DE 3, *E. coli* BL 21 (codon+) and *E. coli* HMS 174 (DE3) plysS. The result showed that the *has*A could only express in *E. coli* BL 21 (codon+). However, the expression level was made in disappointingly small amounts. One factor influencing expression level might be the result of rare codon usage of some codons near the N-terminus of *has*A which could have severe effects on protein yield. Therefore, the rare codon modifications near the N-terminus of *has*A were performed. However, the *E. coli* host could not produce HA even if they contain the pHAS plasmid with the rare codon modifications and could produce hyaluronan synthase.

Keywords: Hyaluronic acid, Hyaluronan synthase, *Streptococcus* sp.,
Polysaccharide engineering, Gene cloning

Introduction

Hyaluronic acid (HA) is a naturally occurring polysaccharide composed of a β 1,4-linked repeating disaccharide of glucuronic acid and 1,3-linked to *N*-acetylglucosamine. HA is a major constituent of the vitreous humor of the eye, synovial fluid, extracellular matrices, and skin. The polysaccharide also interacts with various receptors and binding proteins that modulate cellular behavior such as migration, adhesion, and wound healing. Since the discovery of HA, much attention has focussed on the possible biomedical applications of highly purified HA fractions. Numerous studies have proven that HA isolated from various sources has an identical chemical structure. Therefore, since it is already present in the human body, exogenously applied HA of sufficient purity is highly biocompatible (i.e. it does not provoke adverse host reactions) and completely biodegradable by natural catabolic pathways.

In order to fulfill the raw material requirements for products for the biomedical applications, the identification of dependable and economically viable sources of HA has been an industrial priority. Initially, attention was focussed on the extraction of HA from animal tissues. Numerous tissue sources, including umbilical cord, skin and rooster combs, have been evaluated. Subsequently, rooster comb HA became the most widely used and traditionally accepted source both from an industrial and regulatory point of view. However, the high-molecular-weight material from animal source is difficult and costly to isolate due to the fact that the HA is complexed with proteoglycans. Therefore, attention has turned in recent years to the identification of alternative HA sources. Interestingly, HA is also found in the extracellular capsule of pathogenic group A and group C *Streptococci*. The enzyme that polymerizes the polysaccharide is HA synthase. It is now clear that, in group A streptococci, a single protein utilizes both sugar substrate; UDP-*N*-acetylglucosamine (UDP-GlcNAc) and UDP-glucuronic acid (UDP-GlcA), to synthesize HA (DeAngelis and Weigel, 1994). The abbreviation HAS, for the HA synthase, has gained widespread support for designating this class of enzymes. However, in this study we try to clone group C Streptococci which has been used for HA production in industry.

The molecular cloning of genes encoding enzymes responsible for hyaluronan biosynthesis is one of the essential steps in elucidating details of HA biosynthetic pathway in group C Streptococci. The cloning will also help in paving ways to improve enzyme function and resulting in non-pathogenic mutants for larger production of the desired hyaluronic acid which ensure the absence of other toxic impurities

Materials and Methods

Materials

Media reagents were purchased from Difco (Michigan, USA). Restriction and DNA modifying enzymes were purchased from Promega (Madison, USA). Synthetic oligonucleotides were made at the Bioservice Unit, National Science and Technology Development Agency, Thailand. All other reagents were of the highest grade available from Sigma chemical company (St. Louis, MO) except where state otherwise.

Microorganisms and plasmids

The selected strain of *Streptococcus zooepidemicus* was used through out the experiments. The *Escherichai coli* host strain TOP 10 was used as hosts for pCR®-Blunt and pBAD-TOPO® vector (Invitrogen, USA). *E.coli* strain JM 109 was used as host for pUC18/19, and PCR 2.1 cloning vector (original TA cloning vector kit) (Invitrogen, USA) and pCR® Blunt cloning vector. *E. coli* strain BL21 DE3, strain BL21 (codon+) and strain HMS 174 (DE3) pLysS were used as host for pET-28a expression vector (Invitrogen, USA). Unless otherwise noted, streptococci were grown in Brain Heart Infusion (BHI) and *E.coli* strains were grown in LB medium.

Isolation of streptococcal DNA

Genomic DNA was isolated from *S. zooepidemicus* by the method of Caparon and Scott (1991) with slight modifications.

Gene cloning by PCR

The genomic DNA was used as the template for the polymerase chain reaction. The hyaluronan synthase gene (*has*A) region was amplified with designed primers. PCR was performed in a thermocycler (Geneamp PCR System 2400, Perkin Elmer, Foster, CA) using the following program: i) 1 cycle of 95°C for 2 min, 45°C for 3 min, and 72°C for 5 min; ii) 30 cycles of 95°C for 45 sec , 45°C for 30 sec, and 72°C for 1 min 30 sec. The amplified products was separated on an agarose gel electrophoresis and purified using the Gene clean kit (BIO 101. Inc.) The DNA fragment was directly sequenced. The remaining portion of the *has*A gene was obtained by gene walking.

Southern and colony hybridizations

Southern blot analysis was performed according to standard procedures. Probe labelling hybridization and immunological detection was performed using the digoxigenin DNA labelling system (Boehringer Mannheim GmbH, Mannheim, Germany) and, as a probe, a DNA fragment amplified by PCR using appropiate primers. A fragment, which gave the strongest signal, was cut from a preparative agarose gel and ligated into the appropiate site of the pUC19 vector. *E. coli* JM109 was transformed with the recombinant plasmids and then a plasmid gene library was established.

Colony hybridization was also performed by the digoxigenin DNA labelling system (Boehringer Mannheim GmbH, Mannheim, Germany). Recombinant plasmids from colonies that gave positive signals was analyzed by restriction digestion following Southern hybridization analysis to confirm that the insert DNA originated from *S. zooepidemicus*.

DNA sequence analysis

DNA sequence analysis was performed by the dideoxynucleotide chain termination method modified from that described by Sanger et al. (1977) with fluorescent primers (Auto Read Sequencing kit, Pharmacia Biochemicals, Uppsala, Sweden) using an automated DNA sequencing system (ALF express, Pharmacia Biochemicals, Uppsala, Sweden). The nucleotide and deduced amino acid sequences were analyzed using DNASIS™ software (Hitachi softwares Co. Ltd., Tokyo, Japan).

Expression cloning of the putative hasA in E. coli

PCR primers were designed at the start and stop codon regions of szHAS to contain a *Nde*I restriction site in the sense oligonucleotide and a *Bam*HI site in the antisense oligonucleotide. Using Taq DNA polymerase (Promega, USA), these primers amplified a 1.4-kb PCR product from genomic DNA of *S. zooepidemicus*. The amplified fragment was purified by agarose gel electrophoresis, and cloned into PCR 2.1 vector. Then, the recombinant plasmid was extract and cut with *Nde*I and *Bam*HI. The *Nde*I–*Bam*HI fragment was ligated into pET-28a® (Invitrogen, USA), expression vector. The recombinant plasmid, bearing the *has*A gene from *S. zooepidemicus*, as designated pHAS-1 was transformed into *E. coli* JM 109 and *E. coli* BL 21 DE 3, respectively. Furthermore, the pHAS-1 was transformed into *E.coli* BL 21 (codon$^+$) and *E.coli* HMS 174 (DE3) pLysS cells in order to examine the expression level. Screening of the positive transformant was performed by colony PCR.

Production of HAS

E. coli strain BL21 DE3, strain BL21 (codon+) or strain HMS 174 (DE3) pLysS cells carrying the expression plasmid were grown at 37°C in LB medium containing kanamysin (50 µg/ml) until the optical density at 660 nm reached 0.4. Expression was then induced with 1 mM isopropyl-*β*-D-thiogalactopyranoside (IPTG) and cells were incubated for 4 hr at 37°C. The cells were harvested by centrifugation 8,000x g for 10 min at 4°C and washed with buffer twice. The cells were disrupted by sonication. The disrupted bacteria were centrifuged at 120,000 g for 60 min at 4°C. The membranes were suspended in appropriate buffer. The suspension were incubated for 30 min at 37°C with 1 mM-UDP-*N*-acetylglucosamine and 1 mM-UDP-glucuronic acid. The membranes was separated from the substrate buffer by ultracentrifugation at 120,000x g for 30 min. The membranes were suspended in appropiate buffer. Triton X-114 was added to a final concentration of 2%(v/v) at 4°C. The suspension was subjected to ultracentrifugation and the supernatant was used for further analysis.

Electrophoretic analyses

The subunit molecular weight of HA was determined by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

HA Production and Determination

Hyaluronic acid concentration in the culture broth was determined by the carbazole method measuring uronic acid (Dische, 1947; Bernfeld, 1955; Miller, 1958). The culture broth was centrifuged at 3,000x g for 20 min. After the cells were removed, 2 volumes of ethanol were added to 1 volume of the supernatant from the culture broth and the solution was refrigerated at 4°C for 1 h to precipitate hyaluronic acid. The precipitate was recovered, redissolved in distilled water, and analyzed for the hyaluronic acid concentration.

Results and Discussion

Cloning of hasA from the S. zooepidemicus chromosome

Comparative analysis of the amino acid sequences of spHAS (*S. pyogenes*), DG42 (*Xenopus laevis* HAS) (DeAngelis and Achyuthan, 1996) and nodC (a *Rhizobium meliloti* modulation factor) (Nagahashi et al., 1995) has revealed the existence of conserved regions. Therefore, we designed two deoxyoligonucleotide primers for PCR. The PCR product was isolated, introduced into pCR®-Blunt cloning vector, and the insert was then characterized by DNA sequencing. The nucleotide sequence was determined and the deduced amino acid sequence was then compared with known HAS sequences. The amino acid showed high similarity to other known HAS. In order to obtain the entire *has*A sequence, Southern hybridization was performed using the PCR fragment as a probe against chromosomal DNA of *S. zooepidemicus* which were digested with *Hin*dIII, *Bam*HI, *Eco*RI, *Sma*I or *Sal*I.

*Sequence analysis of has*A

HAS, the enzyme that polymerizes HA from *S. zooepidermicus* is composed of 417 amino acids with a calculated molecular mass of 47.77. Kda, and the predicted pI of 9.01. It is the same number of amino acids with *S. equisimilis.* The DNA sequence and the deduced amino acid sequences are shown in Fig. 1. Interestingly, amino acid sequence of *S. zooepidemicus* showed 99.7 % match to *S. equisimilis* and 70.4 % match to *S. pyogenes* by the homology search, as shown in Fig. 2 and 3.

The overall membrane topology predicted for szHAS (Fig. 4) was identical to that for seHAS (from *S. equisimilis*) and sp HAS (from *S. pyrogenes*) and the eukaryotic HASs reported thus far (Weigel et.al.,1997). The protein has two putative transmembrane domains at the amino terminus and 2-3 membrane-associated or transmembrane domains at the carboxyl end. The hydropathy plots for the three streptococcal enzymes are virtually identical.

Nucleotide sequence accession number

The *S. zooepidemicus has*A nucleotide sequence has been assigned Accession No. BankIt 419336 AF 414053 by GenBank.

*Expression cloning of the putative has*A *in E. coli*

The complete coding region of the *has*A gene from *S. zooepidemicus* was amplified using PCR with a set of forward and reverse primers. The recombinant plasmid, pHAS, bearing the *has*A gene from *S. zooepidemicus* was constructed and transformed into *E. coli* BL 21 DE 3, *E.coli* BL 21 (codon+) and *E.coli* HMS 174

(DE3) pLysS. The result showed that the *has*A could only express in *E.coli* BL 21 (codon+). However the expression level was made in disappointingly small amounts. One factor influencing expression level might be the result of rare codon usage near the N-terminus of *has*A which could have severe effects on protein yield (Zhang et al., 1991; Sorensen et al., 1989; Brinkmann et.al., 1989; Chen and Inouye, 1990). Therefore, the new designed primer near the N-terminus of *has*A which the Arg codon changed was from AGA to CGT and Thr codon from ACA to ACC as well as Ile codon from ATA to ATC were performed (Komine et.al., 1990). We found that the expression of modified rare codons of *has*A could only express in *E.coli* HMS 174 (DE3) pLysS as shown in Fig. 5. However, the *E. coli* host could not produce HA even if they contain the pHAS plasmid and could produce hyaluronan synthase. It might be that they require another enzyme to generate other substrates for HA production. As we know, the hyaluronic acid production in *S. pyogenes* need 2 more enzymes; UDP-glucose dehydrogenase and UDP-glucose pyrophosphorylase, encoded by *has*B and *has*C respectively. Therefore, the result suggesting that gene cloning of *has*B and/or *has*C may be required for HA production in *E. coli*

Acknowledgment

This work has been supported by grants from the Thailand Research Fund (PDF/30/2542) and Silpakorn University as well as the JSPS-NRCT/DOST/LIPI/VCC Large scale cooperative research program in the field of Biotechnology.

References

Brinkmann, U., Matles, R.E. and Buekel, P (1989). High-level expression of recombinant genes in *Escherichia coli* is dependent on the availability of dna Y gene product. *Gene.* **85**, 109-114.

Caparon, M.G. and Scott, J.R. (1991) Genetic manipulation of pathogenic Streptococci. *Methods Enzymol.* **204**: 556-586.

Chen GF, Inouye M. (1990). Suppression of the negative effect of minor arginine codons on gene expression; preferential usage of minor codons within the first 25 codons of the *Escherichia coli* genes. *Nucleic Acids Res.* **18**(6):1465-1473.

DeAngelis, P.L. and Achyuthan, A. M. (1996) Yeast-derived recombinant DG42 protein of Xenopus can synthesize hyaluronan in vitro. *J. Biol. Chem.* **271**: 23657-23660.

DeAngelis, P.L. and Weigel, P.H. (1994) Immunochemical confirmation of the primary structure of Streptococcal hyaluronan synthase and synthesis of high molecular weight product by the recombinant enzyme. *Biochemistry* **33**: 9033-9039.

Dische, Z. (1947). A new specific color reaction of hyaluronic acids. *J. Biol. Chem.* **164**, 189-198.

Higgins, D.G. and Sharp, P.M. (1988) CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene **73**(1): 237-244.

Komine Y, Adachi T, Inokuchi H, Ozeki H. (1990). Genomic organization and physical mapping of the transfer RNA genes in *Escherichia coli* K12. *J Mol Biol.* **212**(4):579-698.

Kyte, J., and Doolittle, R.F. (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**: 105-132.

Nagahashi, S., Sudoh, M., Ono,N., Sawada, R.,Yamaguchi, E., Uchida, Y., Mio, T., Takagi, M., Arisawa, M. and Yamada-Okabe, H. (1995) Characterization of chitin synthase 2 of *Saccharomyces cerevisiae*. Implication of two highly conserved domains as possible catalytic sites. *J. Biol. Chem.* **270**: 13961-13967.

Sanger, F., Nicklen, S. and Coulson A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Ncad. Sci. USA*. **74**: 5463-5467.

Sneath, P.H.A. and Sokai, R.R. (1973) Numerical Taxonomy, W.H. Freeman & Co., San Francisco.

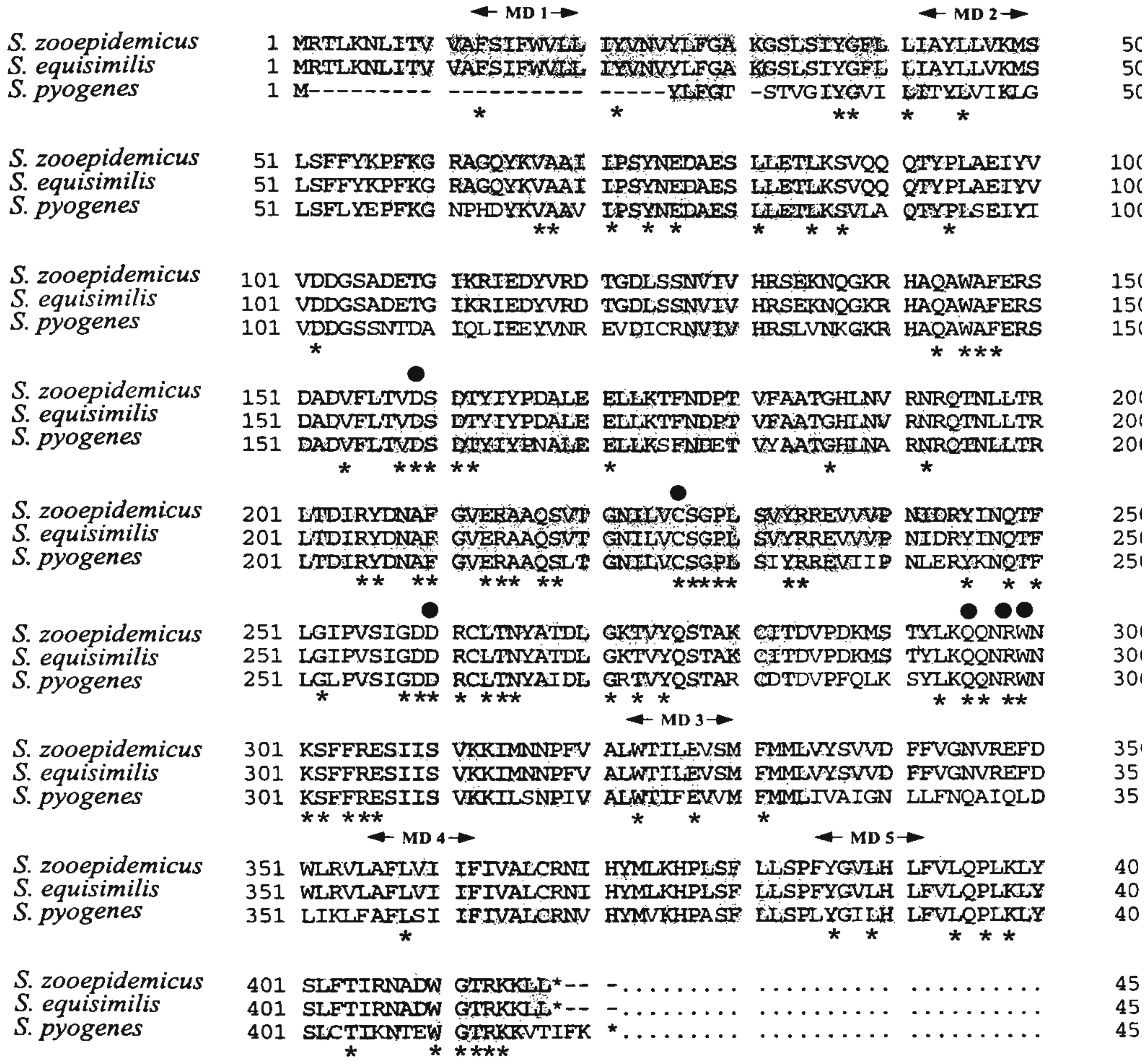
Sorensen, M.A, Kurland, C.G. and Pedersen S. (1989). Codon usage determines translation rate in *Escherichia coli*. *J. Mol. Biol.* **207**(2):365-377.

Weigel, P.H., Hascall, V.C., and Tammi, M. (1997) Hyaluronan synthase. *J. Biol. Chem.* **272**: 13997-14000.

Zhang, S.P., Zubay, G.and Goldmam, E. (1991). Low-usage codons in *Escherichia coli*, yeast, fruit fly and primates. *Gene*.**105**(1):61-72.

```
   1 ATGAGAACATTAAAAAACCTCATAACTGTTGTGGCCTTTAGTATTTTTTGGGTACTGTTG     60
   1  M  R  T  L  K  N  L  I  T  V  V  A  F  S  I  F  W  V  L  L      20

  61 ATTTACGTCAATGTTTATCTCTTTGGTGCTAAAGGAAGCTTGTCAATTTATGGCTTTTTG    120
  21  I  Y  V  N  V  Y  L  F  G  A  K  G  S  L  S  I  Y  G  F  L      40

 121 CTGATAGCTTACCTATTAGTCAAAATGTCCTTATCCTTTTTTTACAAGCCATTTAAGGGA    180
  41  L  I  A  Y  L  L  V  K  M  S  L  S  F  F  Y  K  P  F  K  G      60

 181 AGGGCTGGGCAATATAAGGTTGCAGCCATTATTCCCTCTTATAACGAAGATGCTGAGTCA    240
  61  R  A  G  Q  Y  K  V  A  A  I  I  P  S  Y  N  E  D  A  E  S      80

 241 TTGCTAGAGACCTTAAAAAGTGTTCAGCAGCAAACCTATCCCCTAGCAGAAATTTATGTT    300
  81  L  L  E  T  L  K  S  V  Q  Q  Q  T  Y  P  L  A  E  I  Y  V     100

 301 GTTGACGATGGAAGTGCTGATGAGACAGGTATTAAGCGCATTGAAGACTATGTGCGTGAC    360
 101  V  D  D  G  S  A  D  E  T  G  I  K  R  I  E  D  Y  V  R  D     120

 361 ACTGGTGACCTATCAAGCAATGTCATTGTTCACCGGTCAGAAAAAAATCAAGGAAAGCGT    420
 121  T  G  D  L  S  S  N  V  I  V  H  R  S  E  K  N  Q  G  K  R     140

 421 CATGCACAGGCCTGGGCCTTTGAAAGATCAGACGCTGATGTCTTTTTGACCGTTGACTCA    480
 141  H  A  Q  A  W  A  F  E  R  S  D  A  D  V  F  L  T  V  D  S     160

 481 GATACTTATATCTACCCTGATGCTTTAGAGGAGTTGTTAAAAACCTTTAATGACCCAACT    540
 161  D  T  Y  I  Y  P  D  A  L  E  E  L  L  K  T  F  N  D  P  T     180

 541 GTTTTTGCTGCGACGGGTCACCTTAATGTCAGAAATAGACAAACCAATCTCTTAACACGC    600
 181  V  F  A  A  T  G  H  L  N  V  R  N  R  Q  T  N  L  L  T  R     200

 601 TTGACAGATATTCGCTATGATAATGCTTTTGGCGTTGAACGAGCTGCCCAATCCGTTACA    660
 201  L  T  D  I  R  Y  D  N  A  F  G  V  E  R  A  A  Q  S  V  T     220

 661 GGTAATATTCTCGTTTGCTCAGGCCCGCTTAGCGTTTACAGACGCGAGGTGGTTGTTCCT    720
 221  G  N  I  L  V  C  S  G  P  L  S  V  Y  R  R  E  V  V  V  P     240

 721 AACATAGATAGATACATCAACCAGACCTTCCTGGGTATTCCTGTAAGTATCGGTGATGAC    780
 241  N  I  D  R  Y  I  N  Q  T  F  L  G  I  P  V  S  I  G  D  D     260

 781 AGGTGCTTGACCAACTATGCAACTGATTTAGGAAAGACTGTTTATCAATCCACTGCTAAA    840
 261  R  C  L  T  N  Y  A  T  D  L  G  K  T  V  Y  Q  S  T  A  K     280

 841 TGTATTACAGATGTTCCTGACAAGATGTCTACTTACTTGAAGCAGCAAAACCGCTGGAAC    900
 281  C  I  T  D  V  P  D  K  M  S  T  Y  L  K  Q  Q  N  R  W  N     300

 901 AAGTCCTTCTTTAGAGAGTCCATTATTTCTGTTAAGAAAATCATGAACAATCCTTTTGTA    960
 301  K  S  F  F  R  E  S  I  I  S  V  K  K  I  M  N  N  P  F  V     320

 961 GCCCTATGGACCATACTTGAGGTGTCTATGTTTATGATGCTTGTTTATTCTGTGGTGGAT   1020
 321  A  L  W  T  I  L  E  V  S  M  F  M  M  L  V  Y  S  V  V  D     340

1021 TTCTTTGTAGGCAATGTCAGAGAATTTGATTGGCTCAGGGTTTTGGCCTTTCTGGTGATT   1080
 341  F  F  V  G  N  V  R  E  F  D  W  L  R  V  L  A  F  L  V  I     360

1081 ATCTTCATTGTTGCTCTTTGTCGTAATATTCACTATATGCTTAAGCACCCGCTGTCCTTC   1140
 361  I  F  I  V  A  L  C  R  N  I  H  Y  M  L  K  H  P  L  S  F     380

1141 TTGTTATCTCCGTTTTATGGGGTACTGCATTTGTTTGTCCTACAGCCCTTGAAATTGTAT   1200
 381  L  L  S  P  F  Y  G  V  L  H  L  F  V  L  Q  P  L  K  L  Y     400

1201 TCTCTTTTTACTATTAGAAATGCTGACTGGGGAACACGTAAAAAATTATTATAA         1254
 401  S  L  F  T  I  R  N  A  D  W  G  T  R  K  K  L  L  *          418
```

Fig. 1. Nucleotide and deduced amino acid sequences of the *S. zooepidemicus* HAS ORF. The DNA sequence (top row) of the ORF encoding szHAS was cloned from genomic DNA of *S. zooepidemicus*. The encoded amino acid sequence is shown on the bottom row.



```
                                  ←MD 1→                           ←MD 2→
S. zooepidemicus   1 MRTLKNLITV VAFSIFWVLL IYVNVYLFGA KGSLSIYGFL LIAYLLVKMS   50
S. equisimilis     1 MRTLKNLITV VAFSIFWVLL IYVNVYLFGA KGSLSIYGFL LIAYLLVKMS   50
S. pyogenes        1 M--------- ---------- -----YLFGT -STVGIYGVI LITYLVIKLG   50
                                  *         *               **   *  *

S. zooepidemicus  51 LSFFYKPFKG RAGQYKVAAI IPSYNEDAES LLETLKSVQQ QTYPLAEIYV   100
S. equisimilis    51 LSFFYKPFKG RAGQYKVAAI IPSYNEDAES LLETLKSVQQ QTYPLAEIYV   100
S. pyogenes       51 LSFLYEPFKG NPHDYKVAAV IPSYNEDAES LLETLKSVLA QTYPLSEIYI   100
                                      **   *  * *     *   * *       *

S. zooepidemicus 101 VDDGSADETG IKRIEDYVRD TGDLSSNVIV HRSEKNQGKR HAQAWAFERS   150
S. equisimilis   101 VDDGSADETG IKRIEDYVRD TGDLSSNVIV HRSEKNQGKR HAQAWAFERS   150
S. pyogenes      101 VDDGSSNTDA IQLIEEYVNR EVDICRNVIV HRSLVNKGKR HAQAWAFERS   150
                      *                                            * ***
                             ●
S. zooepidemicus 151 DADVFLTVDS DTYIYPDALE ELLKTFNDPT VFAATGHLNV RNRQTNLLTR   200
S. equisimilis   151 DADVFLTVDS DTYIYPDALE ELLKTFNDPT VFAATGHLNV RNRQTNLLTR   200
S. pyogenes      151 DADVFLTVDS DTYIYPNALE ELLKSFNDET VYAATGHLNA RNRQTNLLTR   200
                        *   *** **         *               *      *
                                                ●
S. zooepidemicus 201 LTDIRYDNAF GVERAAQSVT GNILVCSGPL SVYRREVVVP NIDRYINQTF   250
S. equisimilis   201 LTDIRYDNAF GVERAAQSVT GNILVCSGPL SVYRREVVVP NIDRYINQTF   250
S. pyogenes      201 LTDIRYDNAF GVERAAQSLT GNILVCSGPL SIYRREVIIP NLERYKNQTF   250
                         **  **   *** **        *****   **           *  * *
                              ●                                      ●  ●●
S. zooepidemicus 251 LGIPVSIGDD RCLTNYATDL GKTVYQSTAK CITDVPDKMS TYLKQQNRWN   300
S. equisimilis   251 LGIPVSIGDD RCLTNYATDL GKTVYQSTAK CITDVPDKMS TYLKQQNRWN   300
S. pyogenes      251 LGLPVSIGDD RCLTNYAIDL GRTVYQSTAR CDTDVPFQLK SYLKQQNRWN   300
                      *     *** * ***      * * *                   * ** **
                                             ←MD 3→
S. zooepidemicus 301 KSFFRESIIS VKKIMNNPFV ALWTILEVSM FMMLVYSVVD FFVGNVREFD   350
S. equisimilis   301 KSFFRESIIS VKKIMNNPFV ALWTILEVSM FMMLVYSVVD FFVGNVREFD   35
S. pyogenes      301 KSFFRESIIS VKKILSNPIV ALWTIFEVVM FMMLIVAIGN LLFNQAIQLD   35
                     ** ***                  *   *    *
                          ←MD 4→                          ←MD 5→
S. zooepidemicus 351 WLRVLAFLVI IFIVALCRNI HYMLKHPLSF LLSPFYGVLH LFVLQPLKLY   40
S. equisimilis   351 WLRVLAFLVI IFIVALCRNI HYMLKHPLSF LLSPFYGVLH LFVLQPLKLY   40
S. pyogenes      351 LIKLFAFLSI IFIVALCRNV HYMVKHPASF LLSPLYGILH LFVLQPLKLY   40
                            *                              *  *     * * *

S. zooepidemicus 401 SLFTIRNADW GTRKKLL*-- -......... .......... ..........   45
S. equisimilis   401 SLFTIRNADW GTRKKLL*-- -......... .......... ..........   45
S. pyogenes      401 SLCTIKNTEW GTRKKVTIFK *......... .......... ..........   45
                        *     * ****
```

Fig. 2. Alignment of the deduced amino acid sequence of HAS from *S. zooepidemicus* with *S. equisimilis* and *S. pyogenes*. The alignment was adjusted to the degree of overall sequence similarity with *S. zooepidemicus* using the DNASIS software (Hitachi Software, Yokohama, Japan). Identical amino acids are shown on a black background. Residues in szHAS that are identical in all other HAS family members (including human HAS1, HAS2 and HAS3, which are not shown) are denoted by asterisks. The amino acids indicated by dots are conserved in all members of the larger β-glycosyltransferase family (Nagahashi et.al.,1995). The approximate midpoints of predicted membrane domains (MD) 1 though 5 are indicated.

Fig.3. Evolutionary relationships among the 3 hyaluronan synthases. The phylogenetic tree was generated by the Higgins-Sharp (Higgins and Sharp, 1988) algorithm using the DNASIS multiple alignment program. The program takes as input a dendrogram produced by applying the UPGMA method (Sneath and Sokai, 1973) to a matrix of similarity scores for all of the aligned sequences. The calculated matching percentages are indicated at each branch point of the dendrogram.

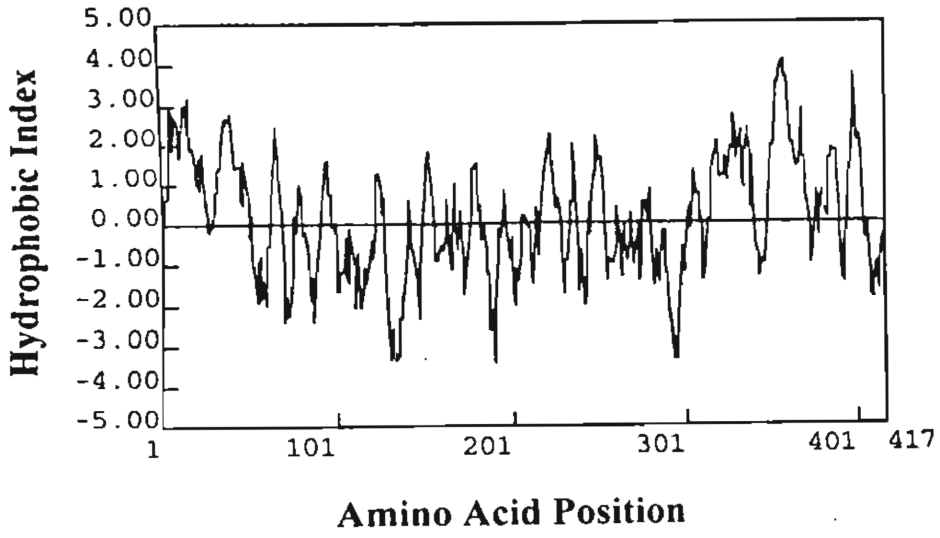


Fig. 4. Hydropathic analysis of the deduced amino acid sequence of HAS from *S. zooepidemicus.* The hydropathic indices are on the Y-axis and amino acid numbers for HAS (1-417) are on the X-axis. The hydrophobicity values were obtained according to the algorithm of Kyte and Doolittle (Kyte and Doolittle,1982). Positive values represent increased hydrophobicity.

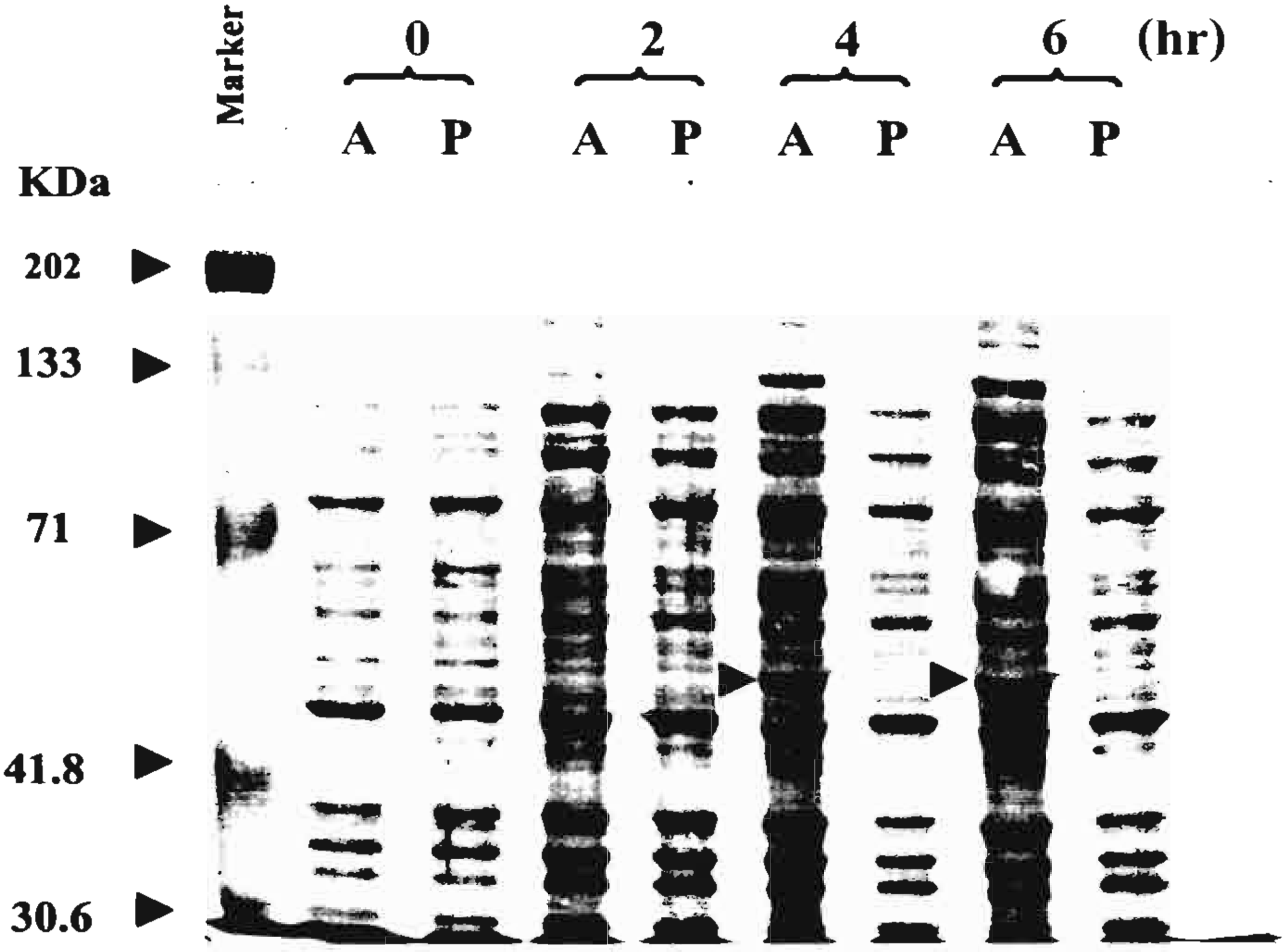


Fig. 5. SDS-PAGE analysis of overexpressed streptococcal HAS in *E. coli* HMS 174(DE3) pLysS at induction time 0, 2, 4 and 6 hrs. The gel was stained with Coomassie Blue R-250. The position of HAS is marked by the arrow.

lane A : recombinant pHAS-r (pET-28a + modified *has*A)
lane P : vector alone (pET-28a)