CHAPTER III

CONSTRUCTION, CHARACTERIZATION
AND PURIFICATION OF SK
AND ITS DERIVATIVES
3.1 Introduction

SK is constitutively produced by many species of β-hemolytic streptococci as a soluble, extra cellular protein. Biophysical studies have shown SK to be composed of three structurally homologous domains designated as α, β and γ, separated by regions of flexible structures (Parrado et al., 1996; Conejero-Lara et al., 1996; Wang et al., 1998). The molecular mechanism involved in the functioning of SK has been the subject of intense investigation in the past few decades. Two loci in the α domain and a ‘core’ region in the β domain (SK230-290), have together been implicated in playing major roles in substrate PG sequestering (Nihalani et al., 1997, 1998). Further, a variety of biochemical and physico-chemical evidence suggests the presence of high-affinity PG binding property in the β domain of SK (Nihalani and Sahni, 1995; Reed et al., 1995; Rodriguez et al., 1995; Conejero-Lara, 1996, 1998; Nihalani et al., 1998). However in contrast to these biochemical studies, the recent crystal structure of SK (Wang et al., 1998) complexed with the catalytic domain of PN (μPN) does not depict an overt interaction between the β domain and the catalytic domain of PG in the bi-molecular complex. Thus, despite the SK-μPN crystal structure offering a high-resolution view of the ‘co-factor’ role of SK in modulating the substrate preference of PN by providing a ‘valley’ for facile docking of the ‘substrate’ PG molecule, the role played by each domain in the process as well as the co-operation amongst these domains in ‘transforming’ the high affinity PG binding function/s in SK into a ‘substrate’ PG processing function is still far from clear.

In order to resolve the apparent ambiguity between the biochemical and structural studies regarding the involvement of the β domain of SK in PG binding and activating functions on the one hand, and to identify residues or loci of the β domain involved in these functions on the other, we decided to subject the β domain to detailed site-directed mutagenic studies. Further, we also sought to explore the functional significance of each domain in SK and to define their inter-domain co-operation in terms of partner and substrate interactions that eventually contribute toward the high PG activation rates observed in full-length SK. This approach required the cloning, expression and
purification of individual domains as well as their covalently contiguous bi-domain combinations as present in the native sequential orientation (αβ and βγ). The primary requirement for these molecular constructions was the cloning of the SK gene and its over-expression in a well-characterized heterologous system amenable to easy genetic manipulations and purification of SK/SK variants.

Previously, the cloning and expression of the SK gene in *E. coli* has been reported (Malke et al., 1984). Using a closely similar strategy, the SK gene from *S. equisimilus* H46A was cloned in pBR 322 and then expressed in *E. coli* as an extracellular protein (Pratap et al., 1996). However this expression system was found to be unsatisfactory for this investigation, since the full-length protein was partially found to be degraded into a smaller form of 44 kDa, which necessitated relatively complicated and protracted strategies to purify full length native 47 kDa protein. In order to develop a simple and rapid system for expression and purification, in this study, the SK gene was transferred into pET23(d) vector, which utilizes the T7 RNA polymerase promoter to drive expression, using standard methodologies, and further optimized for over-expression in *E. coli* (Nihalani et al., 1997; Yadav, 1999). The SK gene was cloned in both pET23(d) vector and pBlueScript vector. The latter was used as the template for the generation of site-directed mutants of the β domain of SK as well as for construction of expression cassettes for individual SK domains (α, β and γ) or two-domain constructs of SK (αβ and βγ). A purification strategy for the isolation of SK from the intra-cellular milieu of *E. coli* from a complex background of a vast number of other intracellular proteins was then evolved. It may be mentioned here that procedures for the purification of extracellular secreted natural SK produced by various streptococcal strains from a background of other secreted proteins, including toxins, had been evolved over the years using a combination of a number of conventional chromatographic techniques such as ion-exchange chromatography, fractionation by precipitation using salts or organic solvents, followed by high resolution techniques such as isoelectric focusing or preparative gel electrophoresis (De Renzo et al., 1967; Castellino et al., 1976; Gerlach and Kohler, 1977; Johnston and Zabriskie, 1986; Nihalani, 1997). However, much of the details regarding the processes evolved for obtaining therapeutic grade SK are not in the public domain. Affinity-based methods for the purification of SK using immobilized PG or PN (Castellino et al., 1976; Rodriguez et al., 1992; Nihalani, 1997), which were inhibited
either with reversible or irreversible inhibitors, have also been developed in recent years. However, apart from the requirement of strong denaturants for the elution of the bound SK and the consequent low yields, a troubling aspect was the frequently observed degradation of SK by small amounts of uninhibited PN or PN formed by reversible loss of the inhibitor (Castellino et al., 1976; Rodriguez et al., 1994). Therefore, a non-affinity based procedure for purification of intracellularly produced rSK from \textit{E. coli} was evolved in the present study, capable of not only purifying the native-like recombinant SK in multimilligram quantities, high purity, and good yields but also purify various site-directed mutants and the different truncated versions of the molecule without major modifications.

This chapter describes the construction, expression, purification and confirmatory analyses (Western blotting, N-terminal sequencing and CD spectropolarimeter studies) of various derivatives of SK.

### 3.2 Materials and Methods

#### 3.2.1 Reagents

Glu-HPG and chromogenic substrate, Chromozym\textsuperscript{®}PL (tosyl-Gly-Pro-Lys-anilide) that were used for HPG activation assays, were purchased from Roche Molecular Biochemicals (Mannheim, Germany). The HPG preparations contained less than 2 % Lys-HPG and less than 0.01 % HPN. All chemicals used were the best commercial grade available and were purchased from either Sigma or Merck. Thermostable DNA polymerase \textit{(pfu)} with proofreading activity and strain \textit{E. coli} XL-Blue were procured from Strategene Inc. (La Jolla, CA). All other enzymes used for genetic manipulation were obtained from New England Biolabs (Beverly, MA). Oligonucleotide primers were either synthesized in-house on an Applied Biosystems DNA synthesizer model 492 or custom-synthesized by Ransom Hill Biosciences Inc. (Ramona, CA). N-terminal protein sequencing was done on a Perkin Elmer/Applied Biosystems sequencer, model 476 A. Purification of plasmid DNA or PCR amplified fragments were routinely performed using DNA purification kits from Qiagen Inc (GMBH, Germany). Automated DNA sequencing using fluorescence-dye was carried out on Applied Biosystems / Perkin Elmer DNA sequencing system Model 310. POROS-D anion-exchange resin was procured from
Perceptive Biosystems (Farmingham, MA). DEAE-Sepharose (fast flow) and Chelating-
Sepharose were procured from Pharmacia Ltd., Uppsala, Sweden, and Phenyl-agarose for
hydrophobic interaction chromatography was procured from Affinity Chromatography
Ltd., Isle of Man, U.K. Bug-buster®, a commercial reagent for rapid bacterial cell lysis
was procured from Novagen (Madison, WI).

3.2.2 Standard methods

All the standard techniques used for DNA manipulations eg. DNA isolation,
restriction enzyme (RE) digestion, agarose gel electrophoresis, purification of DNA
fragments from agarose gels, ligation of DNA fragments, transformation using
electroporation etc., as well as standard techniques for protein manipulation, like SDS-
PAGE, Western blotting were performed according to standard published protocols
(Sambrook et al., 1989).

3.2.3 Vectors

The expression vector, pET23(d) (Fig. 1) used for the intracellular expression of
SK, site-directed mutants of SK or its individual domains in E. coli was procured from
Novagen Inc. (Madison, WI). This vector contains apart from other RE sites, an Nco I site
downstream to the RBS and a DNA sequence encoding for hexa His-tag adjacent to the
Xho I site. The cloning vector pBluescript II KS‐ was procured from Strategene Inc. (La
Jolla, CA).

3.2.4 Constructs

3.2.4.1 Native-like recombinant SK cloned in pET23(d) vector [pET23(d)-rSK]

The plasmid construct for the intracellular expression of SK in E. coli (pET23(d-
rSK) was previously constructed in our laboratory as described in Yadav, 1999. Briefly,
SK gene from S. equisimilis H46A that was previously cloned in pBR 322 (Pratap et al.,
1996), was re-cloned into pET23(d) (Fig. 2), a T7 RNA polymerase promoter based
expression vector (Studier and Moffatt, 1986). The expression levels of SK in this
Fig. 1. Diagrammatic representation of vector pET23(d). This T7 RNA polymerase based vector was used for the intracellular expression of rSK and its derivatives in E. coli BL21.
construct were initially found to be low despite complete sequence identity with the known sequence of the SK gene (Malke et al., 1984). Analysis of the upstream DNA sequence indicated the existence of a strong tendency of the 5' end of the gene to form a highly stable secondary structure (hair-pin loop). The 5' end was therefore replaced with a synthetic oligonucleotide “cassette” containing silent mutations in order to decrease the tendency for formation of secondary structure, while retaining the native amino acid sequence, but with an additional initiation codon for Methionine at the beginning of the ORF encoding SK. This construct was named pET23(d)-rSK.

3.2.4.2 Native-like recombinant SK with point mutation K59E [pET23(d)-rSKK59E]

The construction of rSK DNA cassette with a point mutation at the 59th residue (K59E) had been carried out earlier in our laboratory with a view to produce rSK with significantly enhanced stability against proteolytic digestion by PN (Wu et al., 1998; Shi et al., 1998). A brief description of the construction of SKK59E is given below. The mutation at the 59th residue was introduced into SK by a simple PCR based strategy using a mutant primer (A GGC TTA AGT CCA GAG TCA AAA CCA TTT) along with a down-stream primer (ATA GGA TCC TTA TTC CAC ATC AAC AGA TTT CGG TTG) to amplify the 5' region of the SK gene (approximately, from 57th amino acid to the 162nd amino acid of native SK sequence). The product so obtained was subjected to double digestion with restriction enzymes (REs) Afl II and BstE II and ligated into vector pET23(d)-rSK at corresponding sites (Fig. 2A). The digestion with these specific enzymes allowed docking of the amplified fragment, in frame, directly within the sequence encoding for the α domain (1-143) (Fig. 2B). Selected clones were confirmed for the specific mutation by DNA sequencing.

3.2.5 Assays for studying the activation of HPG by SK and SK mutants

A one-stage assay method was used to measure the kinetics of HPG activation by SK or its mutants (Wohl et al., 1980; Shi et al., 1994). Purified SK or mutant SK (0.5-50 nM) were added to a 100 μl quartz assay cuvette containing HPG (1 μM) in assay buffer (50 mM Tris-Cl buffer, pH 7.5) containing 1.0 mM chromogenic substrate (Boehringer-
Fig. 2. Restriction map of pET23(d)-rSK. A. The circular map shows selected, unique RE sites on the pET23(d) vector and the SK gene that were used for the construction of SK derivatives. B. Linear diagrammatic representation of the restriction map of SK gene depicting RE sites unique to the SK gene.
Mannheim). The change in absorbance at 405 nm was then measured as a function of time (t) in a Shimadzu UV-160 model spectrophotometer at 22 °C. Appropriate dilutions of standard W.H.O. S. equisimilis streptokinase were used as reference for calibration of international units/mg protein (specific activity) in the unknown preparations (Heath & Gaffney, 1990). The activator activities were obtained from the slopes of the activation progress curves, which were plotted as change in absorbance/t against t (Wohl et al., 1980).

3.2.6 Site-directed mutagenesis using 'mega primer' PCR based technique

3.2.6.1 Construction of site-directed, single charge-cluster mutants in the β-domain of SK

The different SK mutants were constructed by a PCR-based strategy, using the "megaprimer" method of site-directed mutagenesis employing one mutagenic and two (common) flanking primers (Sarkar and Sommers, 1990; Smith and Klugman, 1997) followed by cloning into plasmid vectors by standard methodologies (Sambrook et al., 1989) (Fig. 3). The sequences of the mutagenic primers used for the mutagenesis experiments are shown in Table 1. The primers were designed to be complementary to the known DNA sequence of the SK gene from S. equisimilis (Malke et al., 1984). A total of 13 SK mutants, of three distinct types were made viz., single-, double-, and triple-charge cluster sites mutated either to Ala residues (in most cases) or, in a few instances, to a complimentary side-chain to effect charge-reversal (e.g., Lys to Glu and vice-versa), with each mutant type representing the incorporation of two, four or six simultaneous point mutations, respectively. Table 1 also depicts the diagnostic RE sites introduced by translationally silent mutagenesis (Raghava and Sahni, 1994) in each primer to aid the screening of the clones. The sequence of the two flanking primers (termed the 'upstream' and 'downstream' primers) in the amplification reaction are also shown. Each of these carried an internal RE site unique to the SK ORF [as also to the plasmid pET23(d)-rSK] viz., BseR I and Bsm I to facilitate the re-ligation of the mutant PCR blocks back into the SK ORF in the expression vector (see Fig. 2).
Fig. 3. Diagrammatic representation of the PCR strategy used for the construction of site-directed mutants of SK by the 'megaprimers' method.
Briefly, the first PCRs (to generate the different mutant megaprimers) were carried out in a reaction volume of 100 µL containing (final reaction conditions are given) 200 µM dNTPs, 100 ng of template DNA [pET23(d)-rSK, the expression vector in which the full-length SK gene had been cloned], 20 pmol each of the mutagenic and downstream primers, 2.5 U of standard cloned pfu thermostable DNA polymerase, and 1 X pfu reaction buffer. The PCR cycling conditions were in accordance with the denaturation temperatures of different primers, calculated using the computer program, Oligo (version 4.0). In general, all reactions were started with a 'hot start' (94 °C, 5 min), followed by a denaturation (94 °C for 45 sec), annealing (50 °C for 1 min) and extension (72 °C, 1 min) phases. After a total of 30 cycles, a 10-min period at 72 °C was given for a final extension of partially finished daughter DNA molecules, and the reactions were terminated, and then processed through Qiagen PCR purification columns. The relevant DNA cassettes were then purified by agarose gel electrophoresis, by excising the required DNA band and further purifying by Qiagen gel extraction kit. The amplified DNAs were then used as megaprimers for a second PCR (referred to as PCR-II) with wild-type SK gene as the template. The full-length, extended DNA cassettes were then cloned back into pET23(d)-rSK expression vector at the BseRI and Bsm I sites. The PCR-II pre-mixes (100 µl) contained: 60 ng template [pET23(d)-rSK], 200 µM dNTPs, 1 X pfu buffer, and varying concentrations of megaprimers (200-600 ng). The mixes were held for 5 min at 95°C ('hot start') and the reaction initiated with addition of 5.0 U of pfu DNA polymerase per reaction. This was followed by 7 cycles of denaturation (94 °C, 45 sec) and extension (72 °C for 3 min) to ensure build-up of mutated strand DNA and to effect megaprimer extension. At the end of this phase of the PCR, 20 pmol of upstream primer was added, and thermal cycling continued as follows, for 10 cycles: 94 °C for 45 sec, 51 °C for 1 min, and 72 °C for 1 min. After the last cycle, 20 pmol of downstream primer was added, and cycling continued as before for 15 cycles, followed by a final extension at 72 °C for 10 min. The extended PCR products (504 bp) were then purified on agarose gels and the DNA bands isolated after excision of gel blocks. These were digested with BseRI I and Bsm I restriction enzymes and ligated with similarly digested pET23(d)-rSK plasmid DNA, and transformed into E. coli XL-Blue electrocompetent cells. The desired clones were selected by screening mini-prep plasmid DNAs for the diagnostic RE sites.
<table>
<thead>
<tr>
<th>Charge-cluster Mutation*</th>
<th>Oligonucleotide Sequence</th>
<th>RE** site</th>
</tr>
</thead>
<tbody>
<tr>
<td>DE 238,240 AA</td>
<td>5'-ATTTTACCcATGGGcgCAAGcGTTTACTTACC-3'</td>
<td>Nco I</td>
</tr>
<tr>
<td>RK 244,246 AA</td>
<td>5'-AAGAGTTTACGcTGTTGcAAATCGGAAC-3'</td>
<td>Bsi W1</td>
</tr>
<tr>
<td>RE 248,249 AA</td>
<td>5'-CGTGTAAAAATgctGcGAAGCTTATAGGAT-3'</td>
<td>Fsp I</td>
</tr>
<tr>
<td>KK 256,257 AA</td>
<td>5'-GCTTATAGGATcATgcAcATCTGGTCTGAATGAA-3'</td>
<td>Vsp I</td>
</tr>
<tr>
<td>KK 256,257 EE</td>
<td>5'-AGGATCAATgcAAGAATCTGGTCTcAATGAAGAAATA-3'</td>
<td>Bsm Al</td>
</tr>
<tr>
<td>EE 262,263 KK</td>
<td>5'-AAAAAAATCTGGTCTcAATcAaAATAAACACT-3'</td>
<td>Alw 26I</td>
</tr>
<tr>
<td>EE 262,263 AA</td>
<td>5'-AACATTTGAAATCTGGGcCTcAATgcAAGcATAAACACACT-3'</td>
<td>Bsp 1286I</td>
</tr>
<tr>
<td>EK 272,273 AA</td>
<td>5'-ACCTGATCTCTGGTATTACGTCTCTT-3'</td>
<td>Pst I</td>
</tr>
<tr>
<td>KK 278,279 AA</td>
<td>5'-AGAAATATTACGTcATTgcAAGcAGGGGAAGAC-3'</td>
<td>Bsa A1</td>
</tr>
<tr>
<td>EK 281,282 AA</td>
<td>5'-AAAAAGGGGcGccGGcGTATGACTCC-3'</td>
<td>Nar I</td>
</tr>
<tr>
<td>Upstream primer+</td>
<td>5'-ATTGTGAAACGTGACTCTCAATCGTC-3'</td>
<td>Bse RI</td>
</tr>
<tr>
<td>Downstream primer+</td>
<td>5'-ATAGGCTAATGATAGCTAGCATTCTCTCC-3'</td>
<td>Bsm I</td>
</tr>
</tbody>
</table>

*The mutant primers shown in this table are for the construction of single charge-cluster mutants. The nucleotides altered from the wild-type sequence are depicted by lower case letters. **Restriction sites introduced in the primers to aid screening of mutants through silent mutagenesis are underlined while codons in which the desired mutations were introduced are shown in bold face. +These flanking primers had wild-type sequences, carrying unique RE sites naturally present in the nSK gene, and were used to dock back the mutant PCR cassettes into the SK expression vector.
incorporated through the PCR primers (Table 1), and the positive clones were subjected to automated DNA sequencing to confirm the incorporation of the desired mutation, as also to rule out the presence of any unwanted mutations due to the PCR amplification.

3.2.6.2 Construction of double and triple charge-cluster mutants in the β-domain of SK by site-directed mutagenesis

For the construction of double and triple charge-cluster mutants (four or six amino acid residues mutated at one time), the same overall strategy was followed as mentioned above for construction of single site mutation (mutation of two amino acid residues at one time) except that complementary megaprimer s and templates were used for the extension step in the PCR-II step e.g., for the construction of the double charge-cluster mutant SKRE248.249AA;EK281.282AA, the megaprimer carrying the mutations RE248.249AA was used with the full-length extended PCR product carrying the mutations EK281.282AA. Similarly, for the construction of the double charge-cluster mutant EK272.273AA;EK281.282AA, the extended PCR product containing the mutations EK281.282AA was employed as the template, with the megaprimer carrying the mutations EK272.273AA used to prime the reaction at the PCR-II step. In case of the triple charge-cluster mutant viz., SKRE248.249AA;EK272.273AA;EK281.282AA, the construction strategy was as follows. First, a megaprimer containing the double charge-cluster mutations EK272.273AA and EK281.282AA was constructed using as template the purified PCR product carrying the single charge cluster mutant EK281.282AA, and carrying out another PCR using the upstream primer bearing the EK272.273AA mutation along with the downstream Bsm I-site containing primer. This double charge-cluster megaprimer was then isolated, and used along with the full-length extended PCR product with the mutation RE248.249AA (used as template in PCR-II), to generate PCR amplified DNA block containing three simultaneous charge-cluster mutations. All the purified mutant PCR DNAs were then digested with BseR I and Bsm I restriction enzymes, and cloned into similarly digested plasmid vector pET23(d)-rSK to obtain the desired mutant SK clones. All mutations were confirmed by DNA sequencing of the complete SK gene in the expression vector using the T7 promoter and terminator sequencer primers. This also established the absence of any unwanted mutation introduced in the SK gene during mutant construction particularly as a result of the PCR amplification.
3.2.7 Cloning of individual domains of SK

The primary length defining the individual domains of SK, i.e. \( \alpha \), \( \beta \) and \( \gamma \), whose exact dimensions were selected based on limited proteolysis data (Parrado et al., 1996), also cross-correlated remarkably well with the limits defined by the crystal structure of SK (Wang et al., 1998). The cDNAs that code for each domain were analyzed using the M-FOLD programme (Walter et al., 1994) for the formation of destabilizing secondary structures, especially at the 5' end, that might hinder the translational efficiency of the mRNA sequence, as well as for the compatibility of the N-terminal amino acid sequence for the production of proteolytically stable product, in accordance with the N-end rule (Varshavsky, 1996). The sequences 1-143 and 143-293 were chosen to represent the \( \alpha \) and \( \beta \) domain respectively (Parrado et al., 1996). The \( \gamma \) domain (SK293-414) was cloned from the region coding for residues 300-387, which is slightly shorter than the limits as seen in the crystal structure, as well as by proteolytic studies. The N-terminal region of \( \gamma \) domain (293-414) is composed of amino acids that are good target sites for ubiquitination if present at the N-terminus and its subsequent degradation by 26S proteasome (Varshavsky, 1996). It was therefore incompatible with the production of proteolytically stable protein. Secondly, the C-terminal end of the \( \gamma \) domain (residues 388-414) contains residues that are not critical for the functioning of full-length SK (Radek and Castellino, 1989), forming a disorganized and flexible segment (Radek and Castellino, 1989; Parrado et al., 1996; Wang et al., 1998) that undergoes rapid degradation in the presence of PN. Thus, with a view to express proteolytically stable \( \gamma \) domain, it was cloned to code for residues 300-387. The construction of individual domains of SK was carried out by PCR amplification of the desired sequence, using as template either rSK or rSKK\(_{59E}\) gene and specific upstream and downstream primers (Table 2). The cDNAs for \( \alpha \) (SK1-143) as well as the \( \gamma \) domains (SK300-387) were constructed with codons for a C-terminal hexa-His tag by the docking of the amplified cDNA at specific site, adjacent and upstream to the six Histidine codons present in the pET23(d) vector (Fig. 1). The production of these domains along with hexa-His tag facilitated their purification by affinity interaction on metal chelate matrix. The primers designed for amplification of these domains were not only complementary to the respective regions of the DNA sequence of the SK gene, but also contained specific RE sites to allow for facile, directional docking of the amplified
fragment into pET23(d) or pET23(d)-rSK (Fig. 2). In the construction of β, βγ and γ domain, whose 5’ end encoded for an amino acid different than the native N-terminus Met residue present in rSK, care was taken to use Nco I RE site for docking the 5’ end of the amplified region into the vector (Fig. 1). This allowed for the reconstruction of the Met codon required for initiation of translation. Preliminary screening for positive clones was carried out by restriction digestion of several transformants using one or more of three unique and specific REs (PshA I, BseR I and Bsm I) that digested rSK at unique sites within nucleotide sequence coding for the α, β or the γ domain (PshA I- cleaves at a unique RE site in α domain; BseR I- cleaves at a unique RE site in the β domain; and Bsm I- cleaves at a unique RE site in the γ domain). The linearization or non-linearization of a construct on independent restriction digestion with the three REs, was used as diagnostic tool for the selection of positive clone (Table 3), which were further confirmed by automated DNA sequencing.

**Table 2:** Sequence of PCR primers used for the construction of domains of SK

<table>
<thead>
<tr>
<th>Construct</th>
<th>Primer Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>α(1-143)*</td>
<td>U** AGC CAA TTA GAC GTC AGC GTT GCA GAA ACT GTT GAG G</td>
</tr>
<tr>
<td></td>
<td>D ATC TTG CTC GAG AAC GCGCAC ATG TCC ACT TAG CAA</td>
</tr>
<tr>
<td>β(143-293)</td>
<td>U TCA GCC ATG GTT AGA CCA TAT AAA</td>
</tr>
<tr>
<td></td>
<td>D ATG GGG ATC CTA TTT CAA GTG ACT GCG ATC AAA GGG</td>
</tr>
<tr>
<td>γ(300-386)</td>
<td>U AT ACC ATG GTT GAT GTC GAT ACT AAT GAA</td>
</tr>
<tr>
<td></td>
<td>D TTG CTC GAG GGC TAA ATG ATA GCT GGC ATT CTC</td>
</tr>
<tr>
<td>βγ(293-414)</td>
<td>U TCA GCC ATG GTT AGA CCA TAT AAA</td>
</tr>
<tr>
<td></td>
<td>D ATA GGC TAA ATG ATA GCT AGC ATT CTC TCC TTC</td>
</tr>
</tbody>
</table>

* Numericals within parentheses represent the amino acid stretch of SK that the recombinant domain construct will encode.

** U represents upstream primer, and D represents down stream primer.
3.2.7.1 Construction of His<sub>6</sub> tag-α domain (SK1-143)

The DNA encoding for the α domain was constructed using both rSK and rSK<sub>K59E</sub> genes as the template for the construction of recombinant α as well as α<sub>K59E</sub> domain, respectively. The PCR amplification of the α domain was carried out for 25 cycles (denaturation at 95 ºC for 45 sec, annealing for 52 ºC for 1 min and extension at 72 ºC for 1 min) using specific upstream and downstream primer (Table 2) followed by a final extension at 72 ºC for 10 min. The amplified products were double digested with Afl II and Xho I (see Fig. 2), purified after electrophoresis from agarose gel, and cloned into the T7 polymerase based expression vector, pET23(d)-rSK, at corresponding sites. After transformation of the constructs into E. coli BL21, several transformants were analyzed for positive clones by independent restriction analysis of the plasmid DNA with each of the REs- PshA I, BseR I and Bsm I, which possessed unique sites to the rSK gene (Table 3). Linearization of the plasmid with only the PshA I RE was used as a diagnostic tool for the selection of positive clones. The authenticity of the positive clones was confirmed by the release of a lower sized insert (approximately 360 bp) on double digestion of the construct with Afl II and Xho I, as compared to the release of a larger fragment (approximately 1300 bp) on similar digestion of the rSK construct. The clones were further confirmed through automated DNA sequencing, by di-deoxy method using fluorescent probes on an Applied Biosystems ABI 310 genetic analyzer.

3.2.7.2 Construction of cDNA encoding for the β domain (SK143-293)

The DNA encoding the β domain was constructed by PCR amplification of the selected area on rSK gene using pfu DNA polymerase and specific upstream and downstream primers (Table 2). The PCR amplification reaction was carried out for 25 cycles using the following cycling conditions: denaturation at 95 ºC for 45 sec, annealing at 45 ºC for 1 min and extension at 72 ºC for 1 min. This was followed by final extension at 72 ºC for 10 min. The amplified product was double digested with Nco I and BamH I, purified on agarose gel and ligated into pET23(d) vector, predigested with the same enzymes (Fig. 1). This procedure would ligate the cDNA in-frame with the initiation codon at the Nco I site and terminate at the codon corresponding to the 293<sup>rd</sup> amino acid,
after docking of the cassette at the BamH I site in the MCS. The constructs were transformed into E. coli BL-21 and the positive clones were selected by RE digestion of the plasmid minipreps with PshA I, Bsm I and BseR I (Fig. 2). The linearization of the plasmid with BseR I, and absence of sites for the other two REs were used as diagnostic tools (Table 3). The size of the insert released on double digestion of the construct with Nco I and BamH I (approximately 450 bp) was also used as a parameter for diagnosis of the positive clones, which were further confirmed by sequencing of the plasmid constructs by di-deoxy method on automated DNA sequencer.

**Table 3.** Restriction enzyme diagnostic analysis of the truncated constructs of SK

<table>
<thead>
<tr>
<th>RE</th>
<th>α (SK1-143)</th>
<th>β (SK143-293)</th>
<th>γ (SK300-387)</th>
<th>αβ (SK1-293)</th>
<th>βγ (SK143-414)</th>
<th>SK (1-414)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PshA I</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>BseR I</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Bsm I</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

+ indicates linearization of the construct with the specific RE.
- indicates absence of site for the particular RE.

**3.2.7.3 Construction of His<sub>6</sub> tag-γ domain (SK 300-387)**

The cDNA corresponding to the γ domain was cloned into pET23(d) vector after PCR amplification of the selected region from the cloned rSK template using *pfu* polymerase and specific upstream and downstream primers (Table 2). The cycling conditions used were similar to those used for the α domain i.e. 25 cycles of the following cycling conditions: denaturation at 95 °C for 45 sec, annealing at 52 °C for 1 min and extension at 72 °C for 1 min. This was followed by a final extension at 72 °C for 10 min. The amplified product was subjected to double digestion with Nco I and Xho I, purified from agarose gel and ligated into vector pET23(d) at corresponding positions (Fig. 1). The ligation of the insert at the Xho I site would allow, on expression of the
positive clone, production of the γ domain with C-terminal hexa His-tag extension (Fig. 1). The ligated product was transformed into E. coli BL21 by electroporation and the positive clones were selected by restriction analysis of the plasmid DNA. The size of the insert released by RE digestion of the construct with Nco I and Xho I (approximately 260 bp), as well as the linearization of the construct on digestion with Bsm I, and absence of sites for linearization with PshA I and BseR I (Table 3) were used as parameters for diagnosis of the positive clones, which were further confirmed by di-deoxy sequencing of the construct.

3.2.8 Cloning of two-domain constructs of SK (αβ and βγ)

The two-domain constructs containing covalently liked domains in native-like orientation (i.e. in the order as present in rSK) viz. αβ (SK1-293) and βγ(SK 143-414) were cloned in pET23(d). While the two-domain construct αβ was prepared by a single sub-cloning step, the bidomain construct βγ was cloned after employing a PCR based strategy for synthesizing its cDNA.

3.2.8.1 Construction of the covalent bi-domain combination, αβ (SK 1-293)

The cDNA encoding the two-domain construct αβ was obtained by a single sub-cloning step after double digestion of pET23(d)-rSK and pET23(d)-β using RE Bgl II and BseR I (Fig. 4). The RE Bgl II, cleaves in the vector segment upstream to the RBS and the T7 promoter region, while BseR I cleaves almost at the center of the β domain (Fig. 2). Thus, the double digestion of pET23(d)-rSK with these RE would result in the formation of two fragments: while one of the fragment is an insert segment (0.7 kb) containing sequences upstream to the SK region (the T7 RNA polymerase promoter site and RBS) and a region that codes for the N terminal portion of SK (α domain and a N-terminal portion of the β domain), the other fragment is a large segment that predominantly contains the vector sequence as well as nucleotide sequences which code for the γ domain and a C-terminal portion of the β domain (Fig. 2). The RE digestion of pET23(d)-β results in the formation of a vector segment encoding for the C-terminal region of the β domain along with a stop codon and an insert segment encoding for the
sequence of the N-terminal portion of the β domain, along with the T7 RNA polymerase promoter region and the RBS (Fig. 4). The insert released from RE digestion of pET23(d)-rSK (0.7 kb) and the vector sequence (3.76 kb) obtained after digestion of the pET23(d)-β construct were purified from agarose gel, ligated with each other and transformed into E. coli BL21 cells by electroporation. The transformants were analyzed for positive clones by restriction analyses of the mini-preps for absence of the Bsm I site (Fig. 2) and linearization of the plasmid on digestion with PshA I and BseR I (Table 3). The positive clones were further confirmed by DNA sequencing. Another variant of the bi-domain αβ, the pet23(d)-αβ<sub>K59E</sub>, containing the ‘stabilizing’ mutation at the 59<sup>th</sup> amino acid (K59E) was similarly constructed, except for the use of pET23(d)-rSK<sub>K59E</sub> in place of pET23(d)-rSK for preparing the insert fragment.

### 3.2.8.2 Construction of cDNA encoding for the bi-domain combination, βγ (SK 143-414)

The cDNA for the two-domain construct βγ (SK 143-414) was cloned by PCR amplification of the specific region, residues 143-414 of SK using <i>pfu</i> polymerase and specific primers (Table 2). The PCR was carried out for 25 cycles (denaturation at 95 °C for 45 sec, annealing at 45 °C for 1 min and extension at 72 °C for 1 min) followed by final extension at 72 °C for 10 min. The amplified product was subjected to double digestion with RE Nco I and BamH I, followed by purification on agarose gel. This was finally ligated into the larger fragment derived from pET23(d) construct previously subjected to double digestion with the same enzymes (Fig. 1). The transformants were tested for the specific construct by restriction digestion of the plasmid DNA with PshA I, BseR I and Bsm I (Table 3). Linearization of the plasmid to a size lower than obtained with digestion of plasmid containing rSK with either Bsm I or BseR I (Fig. 2) and non-linearization of the constructs with PshA I (Table 3) was used as diagnostic tool for the selection of positive clones, which were further confirmed by di-deoxy sequencing method.

### 3.2.8.3 Long-term storage and retrieval of E. coli BL-21 cells containing SK gene cloned in pET23(d) vector

45
Fig. 4. Schematic representation of the strategy adopted for the construction of bi-domain construct αβ (SK1-293). A simple sub-cloning step was utilized for the construction of αβ, involving the double digestion of pET23(d)-SK and pET23(d)-β with REs BseR I and Bgl II. The smaller fragment from the SK construct was ligated with the larger fragment from the β-domain construct to obtain a pET23(d)-αβ construct.
A single colony of *E. coli* BL21 cells containing the rSK expressing plasmid (pET23(d)-rSK) was picked up from LB.Amp agar plate (final concentration of ampicillin being 100 μg/ml) and used to seed 20 ml LB containing 100 μg/ml ampicillin (LB.Amp). It was incubated for growth at 37 °C for 12 h under moderate shaking conditions (200 rpm). This culture was used to prepare 15 % glycerol stock (v/v) (Sambrook et al., 1989) by aseptically mixing 0.5 ml inoculum with equal volume of pre-aliquoted, sterile, 30 % glycerol in cryo-vials capable of withstanding -70 °C for long periods of time. The glycerol stocks were stored at -70 °C and were found to be stable for at least 3 years. Retrieval of the clones was carried out by scraping the culture from frozen glycerol stock and streaking on an LB agar plates containing 100 μg/ml ampicillin to obtain well-separated colonies. The plate was incubated at 37 °C for 24 h to allow for colony growth and a single colony from this growth was picked up for further processing. All the culturing conditions and subsequent steps in handling the culture were performed aseptically.

3.2.9 *Standardization of procedures for purification of SK and its mutated derivatives*

SK or its site-directed mutant derivatives were purified from the respective *E. coli* clones containing the respective plasmids that encodes for SK or its different derivatives and expressed under a strong promoter (T7 RNA polymerase) for high level, intracellular production of the proteins.

3.2.9.2 *Solubility of rSK produced intracellularly in E. coli BL21*

Preinoculum was raised by inoculating a single colony (obtained after retrieval of the clones, as mentioned above- section 3.2.8.3) into a 100 ml flask containing 20 ml LB.Amp medium. This flask was incubated for 10 h at 37 °C under shaking conditions (200 rpm) and used to inoculate a 500 ml flask containing 100 ml LB.Amp medium at 4 % level (v/v).

The preinoculum raised was used to seed a 2 L flask containing 500 ml LB.Amp at 4 % v/v. The flask was incubated on a shaker (200 rpm) at 37 °C to allow for bacterial growth till mid-log phase i.e. till an absorbance value of 0.5–0.6 was detected at 600 nm.
At this stage, the culture was induced with 0.5 mM IPTG (final concentration) to initiate the expression of SK. The flask was further incubated at 37 °C and 200 rpm for varying time intervals (3, 6 and 12 h). After each time interval, cells were harvested from 100 ml of the broth by centrifugation at 6000 x g for 10 min followed by washing the cell pellet with ice-cold Sodium chloride-Tris-EDTA (STE) buffer. The pellet, usually weighing 0.8 ± 0.1g (wet weight), was thoroughly suspended in 3.2 ml of ice-cold STE buffer to obtain a suspension of 25 % w/v [(weight of the packed wet cells/volume of suspension buffer) x 100]. This was subjected to sonication (Heat System, New York) at 4 °C for 5 min using a micro-probe, under conditions of 30 sec sonic-pulses interspersed with equal periods of rest. The unlysed cells and cell debris were separated by high-speed centrifugation at 14,000 x g for 10 min and both the supernatant and the pellet so obtained were subjected to analysis on SDS-PAGE for examining the propensity of SK to be produced as soluble or insoluble protein. The supernatant (100 μl) was suspended in equal volume of double-strength SDS-sample buffer (SB), while the pellet obtained from centrifuging 100 μl sonicated material was first washed once with STE buffer followed by resuspension in 200 μl modified sample buffer (MSB) containing 0.125 M Tris, pH 6.8, 2 % SDS, 30 % glycerol, 7 % β-mercapto ethanol and 6 M urea (final concentration). Both the supernatant and the pellet were subjected to centrifugation at 10,000 x g for 15 min and analysed on SDS-PAGE (10 % w/v acrylamide) to examine whether the expressed protein remained as soluble or insoluble after intracellular expression.

3.2.9.3 Standardizing facile lysis conditions for harvesting SK

In order to test the efficacy of various methods of lysis to effect the release of intracellularly expressed SK or its site-directed mutant derivatives produced as a soluble protein, E. coli cells expressing SK were subjected to lysis using either a commercially available chemical agent (Bug Buster® agent, Novagen, WI) or by physical methods (sonication or French press) (Table 4). The chemical lysis using Bug Buster® was carried out in accordance with the manufacturer’s instructions by suspending the cell pellet obtained from 100 ml LB.Amp in the Bug Buster® reagent [final concentration of the bug buster reagent was 1 x, in 20 mM phosphate buffer (PB)] at ratios of 5 ml reagent per g wet weight of the cell pellet, in the presence of 10 μL of ‘Benzonase’ (a genetically engineered endonuclease that degrades various forms of DNA and RNA) and 1 mg/ml of
freshly prepared lysozyme. This mixture was incubated at RT for 30 min under mild shaking conditions, followed by centrifugation at 10,000 g for 10 min to separate soluble protein from unlysed cells and debris. The supernatant was quantitatively analyzed for protein content using Bradford’s reagent (Bradford, 1976) and both the supernatant and the pellet fractions were analyzed for the pattern of protein released by electrophoresis on acrylamide gels (SDS-PAGE). Physical lysis was carried out by suspending the cell pellet obtained from 100 ml LB.Amp in STE buffer, pH 8.0 at 25 % w/v (weight of the pellet/volume of the buffer used for suspension). This was subjected to sonication for 5 min at 4 °C using a mini probe with sonic pulses for 30 sec interspersed with equivalent periods of rest. The lysate was then subjected to high-speed centrifugation and both the supernatant as well as the pellet fractions were analyzed on SDS-PAGE for protein patterns. Lysis by French® Press (SLM Aminco, SLM instruments incorporation, New Jersey) was carried out with cell culture suspended in STE buffer at a final concentration of 25 % w/v. This culture suspension was subjected to lysis at 900 psi for 2 min in a minicell, pre-chilled to 4 °C. The supernatant and the pellet were analyzed, as mentioned above.

Lysis of cells using chaotropic agents (in 20 mM PB) viz., 6 M urea or 6 M GdnHCl was also attempted (Table 4). This was carried out by resuspending cells raised from 100 ml LB in the respective chaotropic agent, at 5 % weight of the cell pellet/volume of the solution of the chaotropic agent, followed by incubation at 4 °C for 30 min under mild shaking conditions. The lysates obtained were similarly analyzed for the total protein released, and visualized on polyacrylamide gels under reducing conditions as mentioned above. Further optimization of related parameters like the shortest time required to effect maximum lysis, as well as the least reagent needed for cell lysis were also carried out using 6 M GdnHCl as the lytic agent. Cells obtained from a liter of growth in LB.Amp medium were pelleted into six equal portions and the wet weights of the pellets were measured. These were suspended at varying concentrations (w/v) in 6 M GdnHCl (5, 10, 15, 25, 35% and 50 %), incubated at 4 °C for 30 min under mild shaking condition, followed by clarification of the supernatant to estimate the amount of total protein released. The next parameter, the minimum time required for optimum lysis of the cells was standardized by suspending the cells raised from 100 ml culture in 6 M GdnHCl at 25 % w/v. This mixture was incubated at 4 °C under mild shaking conditions for varying time intervals (15, 30, 45 and 60 min), and at each time interval, the total protein released, as
well as the total activity of SK released was followed by Bradford's method of estimation and estimating for PG activation capacity in the cell lysate, respectively.

**TABLE 4. Lysis of E. coli cells using physical and chemical agents**

<table>
<thead>
<tr>
<th>Method of lysis</th>
<th>Protein released (mg / 100 ml) of cell culture</th>
<th>Activity released (I.U.) / 100 ml of cell culture</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Physical</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ultra-sonication</td>
<td>17.6</td>
<td>$2.5 \times 10^5$</td>
</tr>
<tr>
<td>French press</td>
<td>20.5</td>
<td>$3.0 \times 10^5$</td>
</tr>
<tr>
<td><strong>Chemical</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bug Buster®</td>
<td>11.2</td>
<td>$1.2 \times 10^5$</td>
</tr>
<tr>
<td>6 M Urea</td>
<td>14.0</td>
<td>$2.2 \times 10^5$</td>
</tr>
<tr>
<td>6 M GdnHCl</td>
<td>17.0</td>
<td>$2.6 \times 10^5$</td>
</tr>
</tbody>
</table>

**3.2.9.4 Purification of SK after release from intracellular milieu**

Purification of SK was carried out from cell lysates produced by different procedures of lysis viz., sonication, Bug Buster® and the chaotropic agent, 6 M GdnHCl. We attempted purification of intracellularly produced SK by lysis of the cells followed by purification using a two-step chromatographic procedure – an HIC step on phenyl-agarose, followed by anion-exchange on DEAE-Sepharose fast-flow attached to an automated liquid chromatography work-station (BioCad SPRINT™, Perfusion Chromatography System, Perseptive Biosystems, MA) capable of formation of a pre-defined gradient for elution, continuous monitoring at dual wavelengths and real-time monitoring of conductivity and pH.
3.2.9.4.1 **Raising of culture for purification**

The pre-inoculum was raised from a 15% glycerol stock of *E. coli* BL21 cells containing pET23(d)-rSK as mentioned in the previous section. This was used to seed 2 L of LB.Amp (four 2 L flasks containing 500 ml LB.Amp each) at levels of 4% v/v. The flasks were incubated at 37°C on a rotary shaker (200 rpm) till a mid-log phase of growth as indicated by an OD of 0.5-0.6 at 605 nm (approximately 2 h) was reached. The culture was immediately induced for SK production by adding sterile IPTG to a final concentration of 0.5 mM and further incubating, for 3 h or overnight, under shaking condition for over-expression of this protein.

3.2.9.4.2 **Harvesting and lysis of the cells**

The cells from the culture media were harvested by centrifugation at 6000 x g for 15 min at 4°C. The supernatant was discarded and the cell pellets were first subjected to washing with ice-cold STE buffer followed by careful and thorough resuspension in 60 ml ice-cold lysis buffer containing 6 M GdnHCl in 20 mM PB, pH 7.2 (suspension concentration of 25% wet weight of cell pellet/volume of lysis buffer). This suspension was incubated for 45 min at 4°C under mild shaking conditions (150 rpm) to effect complete lysis of cells, followed by clarification of the supernatant containing rSK by high speed centrifugation at 14,000 x g for 15 min at 4°C. In the experiment where Bug Buster® was used as the lytic agent, the cell pellet obtained from 2 L culture was thoroughly resuspended in 75 ml 1x Bug Buster® agent, along with 100 μl ‘Benzonase’ and 1 mg/ml freshly prepared lysozyme. This suspension was incubated at room temperature for 30 min on a rotatory shaker (150 rpm) followed by high-speed centrifugation at 14,000 x g for 15 min at 4°C to clarify the supernatant. The supernatant so obtained was taken for further processing on the chromatographic matrices. Lysis by sonication was carried out by suspending the cells at 25% w/v levels in 60 ml STE buffer, pH 8.0 and sonicated, as described previously. The lysate was clarified from debris and unlysed cells, by high-speed centrifugation at 4°C for 15 min at 14,000 x g.

3.2.9.4.3 **Purification of SK on Phenyl-agarose, followed by DEAE-Sepahrose chromatography**
All the steps following lysis were carried out at low temperatures between 4-8 °C, on chromatographic matrices maintained at the same temperature and using buffers and solutions pre-chilled to 4 °C. The clarified supernatant obtained from the preceding step above, containing around 300 mg protein was processed further for purification on a 100 ml bed volume of Phenyl-agarose 6 XL (Affinity Chromatography Ltd., U.K.), packed either in an axial glass column (6 cm x 2.3 cm), or packed in a 100 ml acrylic radial column (Sepragen Corp., CA, USA) to achieve high flow rates. The supernatant from the cells chemically lysed using 6 M GdnHCl, was diluted 10-fold in diluting buffer (containing aliquots of 0.5 M PB stock, pH 7.2, and aliquots of stock 5 M NaCl solution) such that, the diluted lysate finally contained 20 mM PB, 0.5 M NaCl and 0.6 M GdnHCl. This was gently swirled for a few minutes and loaded onto the Phenyl-agarose matrix pre-equilibrated in equilibration buffer (20 mM PB containing 0.5 M NaCl), packed in an axial column or a radial column at 300 or 600 ml/h, respectively. The flow-through was collected, and the column was washed at the same flow rate with 5 bed-volumes of equilibration buffer, followed by another washing step with similar volume of equilibration buffer without NaCl. These effluents were collected in fractions of 50 ml each. The bound SK was finally eluted with double distilled water at a slower flow rate (100 ml/h for axial column and 200 ml/h for radial column). This elute was collected in fractions of 25 ml and the SK activity as well as the amount of protein was measured in each fraction. More than 70 % of the activity initially loaded could be routinely recovered by elution with double distilled water and this eluted protein was approximately 80-85 % pure, as judged on SDS-PAGE (10 % gel). At this stage a 3-4 fold improvement in specific activity was recorded. The use of the radial column allowed the processing of the HIC step more rapidly. This decreased the chance of possible degradation of the desired protein, due to proteases present in the cell extract. However, SK eluted from the radial column with larger volumes of water as compared to elution from the axial column, resulting in increased time of processing prior to the next chromatographic stage.

The protein obtained at HIC was subjected to further purification by anion-exchange chromatography on a DEAE-Sepharose column. The protein-rich water-elute fractions, predominantly containing SK were pooled together and diluted two-fold in buffer to finally contain 20 mM Tris buffer, pH 7.2. This was loaded at 300 ml/h onto an axial column (5.6 cm x 1.5 cm) packed with DEAE-Sepharose Fast Flow matrix
Construction and Purification

(Pharmacia Ltd, Uppsala, Sweden) pre-equilibrated with 20 mM Tris buffer, pH 7.2. The column was extensively washed with the equilibration buffer and the bound protein was eluted using a linear gradient of salt (0-0.6 M NaCl) in the equilibration buffer, at 120 ml/h. The effluents were collected as 10 ml fractions, and each was subjected to protein as well as SK activity estimation. Aliquots of the fractions containing protein were analyzed by SDS-PAGE to examine the relative purity of the eluted protein. The eluted SK was then sub- aliquoted and stored at -70 °C till further use.

3.2.9.4.4 A rapid one-chromatography analytical method for the purification of SK mutants

A quick method for the functional analyses of the various full-length site-directed mutants of SK was perceived to be essential in order to screen for changes in activity, and thus rapidly evaluate the need for their further characterization by more elaborate kinetic studies using highly pure preparations. For this purpose, a quick and effective method of purification was developed by obviating the HIC step on phenyl-agarose. A salting-out step, coupled to ion-exchange chromatography was effectively evolved for purifying small amounts of protein for rapid, but accurate initial screening purposes. The E. coli cells containing the mutant constructs were raised in 500 ml LB.Amp medium, harvested, lysed chemically with 6 M GdnHCl at 25 % w/v and clarified by centrifugation exactly as described previously. This supernatant was then subjected to ammonium sulphate precipitation by drop wise addition of saturated, cold stock solution of the salt (3.9 M ammonium sulphate, pH 7.2) under conditions of moderate and continuous stirring at 4 °C over 1 h, till a final concentration of 60 % was achieved. The precipitated protein was then separated by centrifugation (14,000 x g) for 20 min. The pellet so obtained, containing SK, was either taken in for further processing immediately or saved in mother liquor (in which it was found to be stable at -70 °C) for future use. Pellet derived from 250 ml cell culture was processed for the next step of purification by dissolving in 20 ml of 20 mM Tris, pH 7.2. After thorough resuspension by gentle vortexing, the solution was centrifuged to remove insoluble material (this material had very little SK but several background proteins); the supernatant so obtained was carefully removed, diluted to 200 ml in 20 mM Tris, pH 7.2, and analyzed for protein content and PG activating activity. It was then loaded onto a 1.44 ml anion-exchange column, containing dimethyl amino alkyl groups
immobilized onto perfusion chromatographic beads (POROS-D). This column, attached to the automated chromatographic workstation (BioCAD SPRINT™) was operated at 6 ml/min. After washing the column with 20 bed volumes of equilibrating buffer (20 mM Tris, pH 7.2) at the same flow rate for 5 min, the bound protein was eluted using a linear gradient of salt (0-0.5 M) over 20 min at 1 ml/min. The elute was collected as 1 ml fractions and the fractions containing protein, as revealed by the online UV monitor at 230 and 280 nm were used for estimation of SK activity as well as purity on SDS-PAGE; the relevant fractions were stored as convenient aliquots, at −70 °C, till further use.

3.2.10 Purification of His6-tag truncated SK derivatives (α and γ domains) using metal affinity chromatography

The single-domain constructs, α (SK1-143), αK59E, as well as γ (SK300-387) were expressed as proteins with His6-tag extensions at their C-terminii, to aid their purification by chromatography on Ni2+-immobilized metal affinity chromatography (IMAC). Identical conditions were used for the purification of both of these domains by metal affinity chromatography, as detailed below. The affinity chromatography on Chelating-Sepharose immobilized with Ni2+ ions i.e. the Ni2+-IMAC beads, is based on the principle that certain amino acids, including Histidine residues, can interact with transition metals like Ni2+, Zn2+, Fe3+ through co-ordinate bond formation. The presence of six to eight tandem Histidine residues results in the formation of strong co-ordinate interactions with these transition elements (Porath et al., 1975). Consequently, gene constructs designed with additional nucleotides at their N- or C-terminal ends encoding for 6-8 Histidine residues in the expressed protein would confer on these molecules very high affinity to the Ni2+-Sepharose beads. This high binding strength is capitalized in selective binding of the protein, as well as in washing away loosely bound, unwanted proteins using ‘stringent’ buffers during IMAC procedures (Sulkowski, 1995). The final elution is usually accomplished by washing with buffers containing imidazole, a Histidine analogue, which competes with Histidine for binding to Ni2+ ions on the matrix, resulting in elution of pure protein.

Once hyper-expressed, the α domain was found predominantly in the form of
inclusion bodies (IB), whereas the γ domain was mainly found as soluble protein. Hence, IBs of the α domain obtained after lyses of the cells by sonication, or soluble protein obtained after lyses of the cells (for γ domain), were processed further for the purification of each of these proteins by affinity chromatography on IMAC matrices.

3.2.10.1 **Preparation of Ni²⁺-IMAC matrix**

Chelating-Sepharose (Pharmacia, Uppsala, Sweden) was thoroughly washed with double distilled water to remove the preservative (ethanol). The matrix (10 ml) was then incubated with 40 ml of 100 mM nickel sulphate solution for 10 min at room temperature, followed successively by thorough washing with water and equilibrating buffer (50 mM PB containing 250 mM NaCl and 10 mM imidazole, pH 7.2). The matrix was packed into a small glass column of 3.5 cm x 0.95 cm dimensions, prior to chromatography.

3.2.10.2 **Raising culture for purification**

The *E. coli* cells over-expressing the α domain were grown in 1L LB.Amp, inoculated at 4 % v/v level with preinoculum raised from a single colony, essentially as described for rSK. However, the *E. coli* cells elaborating the γ domain at very high levels (approximately 40 % of the total cellular protein) were found to be unstable. It was observed that these were neither amenable to sub-passaging, nor storage as glycerol stocks at −70 °C for repeated retrieval in order to culture the organism for growing the pre-inoculum since a significant loss in expression of γ domain was observed within a single round of sub-culturing. Detailed analysis into the exact reason/s behind this loss in expression level was not undertaken in the present study. However, we devised a simple way to consistently express the γ domain by always raising the pre-inoculum from clones obtained from freshly transformed *E. coli* cells. The rest of the procedure for raising cells was essentially the same as previously described for rSK.

3.2.10.3 **Preparation of cell lysates for purification**

The IBs containing α domain were obtained by harvesting the cell pellet from 1 L LB.Amp medium. This was washed once with ST buffer (STE buffer without EDTA),
followed by thorough resuspension of the cells in 32 ml ST buffer (25 g cell pellet/100 ml ST buffer). The cell suspension was subjected to sonication (as mentioned before), using a medium sized probe. The lysate was centrifuged at 14,000 x g for 15 min at 4 °C and the pellet obtained (containing the $\alpha$ domain IBs) was washed once with ST buffer, followed by dissolution in 5 ml of 20 mM PB containing 6 M urea, pH 7.2, under mild shaking conditions for 30 min at 4 °C. The insoluble materials were separated by another round of high-speed centrifugation at 23,000 x g for 20 min at 4 °C and the supernatant so obtained was carefully withdrawn and processed further for purification.

The purification of structurally and functionally intact $\gamma$ domain required the use of gentle lytic procedures, like lysis of cells with Bug Buster® treatment. The cells producing $\gamma$ domain were harvested from 1 L LB.Amp medium and subjected to washing with ice-cold ST buffer, followed by lysis using 72 ml Bug Buster® reagent, along with 100 µl Benzonase® and 1 mg/ml lysozyme, as mentioned previously. The lysate obtained was subjected to high-speed centrifugation (14,000 x g) for 15 min at 4 °C to separate the supernatant containing the protein from the debris and unlysed cells, and it was processed further for purification.

The supernatant containing the His-tag protein (either $\alpha$ or $\gamma$ domain) was diluted 10-fold with the diluting buffer (50 mM PB, pH 7.2, containing 250 mM NaCl) containing 10 mM imidazole, to finally contain 0.6 M urea, 50 mM PB, 10 mM imidazole and 0.25 M NaCl, pH 7.2 and gently swirled for 15 min at 4 °C. This was passaged through a 10 ml column pre-packed with pre-equilibrated Ni$^{2+}$-IMAC beads, at 100 ml/h to ensure maximum adsorption of the His-tag protein. The matrix was subsequently washed with 5 bed volumes of the diluting buffer containing 20 mM imidazole, followed by another washing step with 5 bed volumes of diluting buffer containing 60 mM imidazole at 100 ml/h. The bound protein was finally eluted with 50 mM PB containing 150 mM imidazole at 60 ml/h. This was collected as fractions of 2.5 ml each. Each of these was analyzed for the total protein content and also analyzed by SDS-PAGE for purity. Both the His-tagged proteins eluted as single symmetrical peaks, with 95-98 % homogeneity.

The protein obtained from IMAC were subjected to a final desalting step on a 5 ml DEAE-Sepharose column (pre-equilibrated with 20 mM Tris, pH 7.2) as described below,
to remove the unwanted imidazole, which was found to interfere in PG activation assays. Control experiments using 0.1-1.0 mM imidazole showed inhibition of PG activation by rSK. The eluted protein pool was first diluted 10-fold in 20 mM Tris, pH 7.2 and then loaded onto the ion-exchange column pre-equilibrated with 20 mM Tris, pH 7.2. After extensive washing of the matrix with the equilibration buffer (20 mM Tris, pH 7.2), the bound protein was eluted with equilibrating buffer containing 0.25 M NaCl. This was analyzed on SDS-PAGE to check for proteolytic integrity as well as purity, and stored in convenient sized aliquots at −70 °C, till further use.

3.2.11 Purification of β domain of SK (143-293)

The β domain of SK (residues 143-293) was expressed as a soluble protein at a level approximately 30% of the total soluble protein fraction. The cultures for purification were raised in 1 L LB. Amp medium, cells harvested by centrifugation at 6000 x g and sonicated as described before. The supernatant was clarified by centrifugation at 23,000 x g for 15 min at 4 °C to separate it from unlysed cells and other debris. The supernatant so obtained was carefully aspirated and diluted 6-fold in 20 mM Tris, pH 7.2, to around 1 mg/ml protein. This was loaded onto a 40 ml DEAE-Sepharose column, pre-equilibrated with 20 mM Tris, pH 7.2. The β domain eluted unbound in the flow through, which was collected as 10 ml fractions. These protein rich fractions were pooled together and taken over to the next step for further processing. Simultaneously, aliquots of stock solutions of PB (0.5 M), pH 7.2 and ammonium sulphate (3.9 M saturated solution, taken as 100%) were added to this pooled fraction to obtain final concentrations of 50 mM PB and 2 M ammonium sulphate in the diluted protein solution. This solution was loaded onto a 20 ml pre-equilibrated phenyl-agarose column (equilibrating buffer contained 50 mM PB and 2 M ammonium sulphate) at a flow rate of 200 ml/h, followed by washing with 5 bed volumes of equilibrating buffer and 5 bed volumes of equilibrating buffer containing 1 M ammonium sulphate. The bound β was eluted from the matrix at 60 ml/h with 50 mM PB containing 0.5 M ammonium sulphate and collected as fractions of 5 ml each. The protein-rich fractions were pooled together and immediately subjected to desalting on a 3 kDA cut-off ultramembrane ‘Centricon’ concentrator (Millipore Corporation, MA), by centrifugation at 7000 rpm for 2 h at 4 °C in a SS-34 rotor (Sorval), according to instructions of the manufacturer. The protein solution was initially subjected to a 5-fold
concentration, followed by in situ dilution of the concentrated solution by 5-fold with 10 mM PB, pH 7.2. Following four such cycles of concentration and dilution, the concentrated protein was estimated for protein content and stored as aliquots at -70 °C, till further use.

3.2.12 Purification of the two-domain construct αβ (SK 1-293)

The covalently contiguous two-domain protein, αβ, was expressed as a soluble protein in *E. coli* BL-21 cells. The cells for purification were raised from 1 L LB.Amp and lysed by sonication as described earlier for full-length intracellular rSK. The supernatant containing the soluble protein was clarified from debris and unlysed cells as previously described. The supernatant, so obtained, was diluted in 20 mM PB containing 0.5 M ammonium sulphate to 1 mg/ml protein, and this was loaded at 200 ml/h onto a 20 ml Phenyl-agarose column (8 cm x 0.9 cm) pre-equilibrated with equilibration buffer containing 0.5 M ammonium sulphate and 20 mM PB, pH 7.2. The matrix was subjected to washing with 5 bed-volumes of equilibrating buffer, followed by washing with equilibrating buffer containing 0.2 M ammonium sulphate. The bound protein was eluted with 20 mM PB at 60 ml/h and collected as fractions. These fractions were analyzed for total protein content and the relevant fractions containing the protein were pooled together and taken over to the next purification step on an ion-exchange matrix. The HIC-eluted protein was diluted in 20 mM Tris-Cl, pH 7.2, followed by loading onto a 5 ml DEAE-Sepharose column (1.5 cm x 1 cm) pre-equilibrated with the same buffer. The extent of dilution of the protein eluted from HIC was a critical step to ensure binding of αβ onto the DEAE-Sepharose column. Generally, 15-fold dilution in 20 mM Tris, pH 7.2, which showed a conductivity of less than 2.0 millisiemen (mS), was observed to be sufficient to allow binding of αβ onto the DEAE-Sepharose matrix. After loading the protein, the column was subjected to extensive washing with 20 mM Tris, and this was followed by the elution of the bound protein using a linear salt gradient (0 - 0.3 M NaCl), prepared in the running buffer, at 60 ml/h. The eluted fractions were analyzed for protein pattern on SDS-PAGE to assess purity, and stored at -70 °C for future use after sub- aliquoting the relevant fractions into suitable small portions.
3.2.13 **Purification of the two-domain construct, βγ(SK 143-414)**

This two-domain protein was expressed predominantly as inclusion bodies in *E. coli* BL-21. Therefore, the strategy adopted for purification of βγ involved the lysis of induced cells by sonication, harvesting of IBs and dissolution of the IBs in a chaotropic agent (8 M urea), as mentioned previously for the purification of the α domain. The dissolved protein was diluted 10-fold to approximately 1 mg/ml in 50 mM PB, also containing 0.5 M ammonium sulphate, pH 7.2. It was then loaded onto a 20 ml Phenylagarose column, pre-equilibrated with the loading buffer (50 mM PB containing 0.5 M ammonium sulphate, pH 7.2) at 100 ml/h. The matrix was similarly processed as mentioned above for αβ purification for removing loosely bound and unwanted contaminating proteins, followed by elution of the bound βγ with 50 mM PB, pH 7.2 at 60 ml/h. The eluted protein was collected as 10 ml fractions, and each was assayed for total protein content. Relevant fractions were pooled together, diluted 10 fold with 20 mM Tris, pH 7.2, and processed further for purification on a pre-equilibrated DEAE-Sepharose, fast-flow column (5 ml), as described for the purification of intracellular rSK. The bound protein was eluted using a linear salt gradient (0-0.3 M NaCl) in equilibrating buffer, and the effluent was collected as 2 ml fractions. Following estimation of the total protein content in each of the fractions, the protein rich fractions were sub-aliquoted and saved at -70 °C till further use.

3.2.14 **CD measurements**

In order to check the secondary-structure features of the intracellularly expressed and purified SK, as well as the various SK derivatives, the far-UV CD spectra were determined using a Jasco-720 spectropolarimeter. Each protein solution, at a final concentration of approximately 0.15 mg/ml in PBS, pH 7.2 was taken in a 0.1 cm path length cuvette and its CD spectrum recorded in the range of 200-250 nm. From this, the respective buffer baseline was subtracted and the final spectrum represented is an average of 10 scans. This was then used to calculate the percentage of various secondary structural elements (α-helix, β-sheet, β-turn and unstructured region) using a reference algorithm (Yang et al., 1986) provided by the spectropolarimeter manufacturers.
3.3 **RESULTS**

The present study was aimed at identifying, analyzing, and assigning meaningful structure-functions correlations to the sequence at the 'core region' of the β domain earlier attributed to encompass strong PG binding site (Nihalani et al., 1996), and also to study the interdomain co-operation amongst the three domains of SK in terms of PG interactions leading to biological activity. Accordingly, the construction of site-directed mutants in the 'core region' (residues 230-290) of the β domain, as well as construction of single and two-domain constructs of SK, was carried out using PCR based methods. The biochemical study with these proteins required the evolution of purification procedures for obtaining large amounts of full-length native rSK protein or its site-directed mutants within a short period of time and in relatively good yields. This was successfully accomplished by the use of a conventional two-step chromatographic procedure, which essentially fully satisfied these criteria. Further, a quick, one-step method of protein purification on analytical scale for the initial screening of the various mutants was also standardized. Procedures for the purification of each of the individual domain constructs or their covalent two-domain combinations were also standardized to yield protein of very high purity that could be directly used for studies on structure-function correlations.

3.3.1 **Construction and expression of site-specifically altered SK variants and deletion mutants of SK corresponding to single and two-domain combinations**

The different mutants of SK site-specifically altered in the β domain were constructed by a PCR based strategy, using essentially the 'megaprimer' method of site directed mutagenesis (Sarkar and Sommers, 1990), as schematically presented in Fig. 3. The positive transformants were first screened by RE digestions of the plasmid DNA isolated from these clones using specific REs, whose sites had been introduced into the mutant primers used for the amplification reaction in PCR (see Table 1). A set of 13 different mutants, of three basic types was constructed viz, constructs with two point mutations, constructs with four point mutations, and a construct with six point mutations simultaneously introduced in the protein. Most of the mutations involved conversion of the ionic residues in the 'core' sequence to alanines. In a few cases, reversal of charges was also carried out viz., SK(KK256.257EE) and SK(EE262.263KK). Induction of all these
constructs with IPTG for over-expression in E. coli BL21 cells resulted in high-level expression of the soluble mutant SK forms, at levels of approximately 20% that of the total soluble cellular protein. The introduction of mutations at specific sites neither affected the solubility of the protein in the host cells, nor in most cases, the expression levels as compared to the native protein. Further, the mutations introduced in most of the cases did not perturb the proteolytic stability of the molecule as judged by preincubating the mutants with HPN for varying time intervals, followed by electrophoretic separation on SDS-PAGE to analyze their degradation. The pattern of rSK-HPN complex on SDS-PAGE was used as the internal control for comparison. However, the hexamutant SK(CFH) i.e. SK(RE248.249AA.EK272.273AA.EK281.282AA), that possessed six mutant alanine residues in the β domain was found to be proteolytically more susceptible than either the parent rSK molecule or the other site-directed mutants generated for the purpose of our study. Hence, extra care, including the application of quick purification protocol, was used to obtain full length, homogenous form of the mutant protein. The results of the structure-function analysis of the mutants are presented in Chapter IV.

The cloning of truncated derivatives of the SK gene resulting in the production of individual domains of SK (α: SK1-143, β: SK143-293, γ: 300-387) as well as the covalently contiguous bi-domain construct βγ (SK143-414) was performed by PCR amplification of the desired gene segment using flanking primers that annealed at the ends of the specific sequence to be amplified in the full-length SK gene, followed by docking of the amplified fragment in the pET23(d)/pET23(d)-rSK vector, at specific sites. The cloning of another bi-domain construct, αβ was carried out by a simple sub-cloning strategy, which involved the ligation of the insert released on double-digestion of pET23(d)-rSK with BseR I and Bgl II (Fig. 4), with the larger vector fragment obtained from pET23(d)-β subjected to double digestion with the same restriction enzymes. At a preliminary level, positive clones were selected by RE analysis of several transformants, for each of the construct (Table 3). This was followed by confirmation using automated DNA sequencing. The peptide bond between 59th and 60th amino acid in SK is the site of rapid cleavage when SK combines with PG (Wu et al., 1998; Shi et al., 1998). Hence, mutation of the Lys residue at the 59th position to an amino acid not recognized for cleavage by PN (non-Arg, non-Lys type) results in significant enhancement of the stability of SK against proteolytic digestion by PN. Hence the α and αβ domains were also
constructed with a point mutation at the lysine residue at the 59\textsuperscript{th} position (K59E), besides their native, unmutated forms, in order to decrease the possible susceptibility of this peptide bond to proteolysis by PN during their functional analysis.

Each of the individual domains, when expressed as intracellular proteins in \textit{E. coli} BL21 cells, resulted in high levels of expression of these proteins. The \(\alpha\) domain was expressed as IB, contributing approximately 20-25 \% of the total protein (Fig. 7), while \(\beta\) was expressed as soluble protein to a level approximately 25-30 \% of the total cellular protein (Fig. 8) content as judged by SDS-PAGE. In the case of the \(\gamma\) domain, the expression was at a level of 30-40 \% total cellular protein in soluble state (Fig. 9). However, the \textit{E. coli} cells expressing \(\gamma\) domain exhibited characteristics, distinctly different from those elaborating either SK or the other constructs. The induction of \textit{E. coli} cells at mid-log phase for over-expression of \(\gamma\) allowed virtually no further growth of the cells, as noted by an additional increase of barely 0.1-0.2 absorbance units at OD 600 nm, despite incubation up to 12 h, post-induction. In comparison to this, \textit{E. coli} cells expressing any of the other constructs (SK or other domain/domain combinations) achieved substantial growth after induction with IPTG, reflected as an additional increase of at least 0.6-0.7 absorbance units at OD 600 nm. Further, this clone was also highly unstable to passaging/subculturing and therefore had to be always cultured from freshly transformed cells for reproducible over-expression of the protein.

3.3.2 Development of a rapid purification procedure for rSK and its derivatives

3.3.2.1 Expression and release of rSK

The cells induced for intracellular production of rSK using the T7 RNA polymerase based expression plasmid, pET23(d), were lysed by sonication. This was followed by analysis of the soluble and pellet fractions on SDS-PAGE to examine the propensity of the over-expressed SK to either form IBs or remain within the cell as soluble protein. Analysis of the patterns on gels clearly showed that SK was over-expressed in soluble form, constituting almost 20-25 \% of the total soluble protein content (Fig. 5). Further incubation for longer durations at 37 \textdegree C (up till 12 h) did not change the solubility
**Fig. 5. Intracellular expression of rSK in E. coli.** The propensity of intracellularly expressed rSK to form inclusion bodies (IB) was analyzed by incubating the culture for varying time intervals (3h, 6h, 12h) at 37 °C, and examining the lysate for formation of IBs in the pellet after centrifugation, as explained in Materials and Methods. The lanes of the 10 % SDS-PAGE represent: 1- Standard molecular weight markers; 2- cell lysate of uninduced cells; 3- whole cell lysate of induced cells after 3 h incubation; 4,6 and 8- soluble protein fractions obtained on lysis of cells induced for 3,6 and 12 h respectively; 5,7 and 9- insoluble protein fractions obtained from the pellets of lysed cells induced for 3,6 and 12 h respectively.
characteristics of the over-expressed SK, and very little SK was found in the insoluble fraction even after overnight incubation at 37 °C.

The *E. coli* cells expressing rSK were lysed using both physical (sonication and French press) as well as chemical methods (Bug Buster®, 6 M GdnHCl and 6 M urea), followed by analyses of the lysate for total protein solubilized per unit wet cell mass, as well as total SK activity released. The results, depicted in Table 4 indicate that both the physical and chemical methods were almost equally efficient in lysing cells and releasing SK, with a slightly better edge observed for the physical lysis techniques. Lysis of cells with 6 M GdnHCl resulted in the release of cellular proteins and SK in amounts comparable to standard physical lysis techniques like ultra-sonication or French Press. This agent was further standardized for use as an effective lytic agent, due to relative ease and economic advantages over the other methods, as well as the stability of the agent at ordinary storage conditions.

Experiments to check the maximum concentration of the cell suspension that could be lysed by a fixed volume of 6 M GdnHCl revealed that it could effectively lyse cells suspended at as high a level as 25 % w/v. However, use of cell suspensions beyond this concentration resulted in inefficient lysis, retaining substantial portion of the total SK produced within the unlysed cell fraction. The time required to effect maximum extraction of protein and SK in the soluble phase, at 4 °C, was also optimized using cells suspended in 6 M GdnHCl (25 % w/v). This revealed that an incubation time of 30 to 45 min was necessary to release maximum amount of intracellular SK. No improvement in yields of either SK or total amount of protein liberated per unit cell mass was recorded on incubation of cells with the lytic agent beyond these durations.

### 3.3.2.2 Purification of rSK by HIC and ion-exchange chromatography

Purification of SK was attempted using 2 L shake flask culture as the starting material. Three different lytic procedures viz., physical lysis by sonication and chemical lysis by 6 M GdnHCl and Bug Buster®, were used for lysis of the cells. Identical chromatographic procedures were followed in processing of the cell lysates obtained by various methods of lysis. A combination of two chromatographic procedures was used for
the purification i.e. hydrophobic interaction chromatography (HIC) on Phenyl-agarose, followed by an ion-exchange step on DEAE-Sepharose fast-flow. Purification of the lysate on Phenyl-agarose resulted in approximately 4-fold purification of the protein, with an overall yield of around 70% of the initial activity and approximately 16-20 mg SK/L of the culture (Table 5). SDS-PAGE analysis of this partially purified protein, eluted with water, depicted SK to be roughly 80-85% pure (Fig. 6A). The eluted protein was then further subjected to purification on DEAE-Sepharose. The matrix was developed using a linear gradient of NaCl that resulted in elution of the bound SK at around 0.25 M NaCl, as a single, symmetrical peak, which on SDS-PAGE showed a mobility identical to native SK (purified from S. equisimilus) corresponding to a molecular weight of 47 kDa. The protein at this stage was > 95% pure (Fig. 6A). An approximately 6-fold purification of the rSK from the crude lysate, with a final yield of nearly 60%, was achieved using this two-step procedure. The fold purification, as well as the overall yields attained after the two chromatographic steps were very similar when using the lysates released by the different lytic procedures as the starting material. Western blot analysis with rabbit polyclonal antisera against native SK confirmed this protein to be SK. This preparation routinely had a specific activity of around 90-95,000 IU/mg and the gas phase automated N-terminal sequencing revealed identical N-terminal sequence as that of native SK reported for S. equisimilus H46A, except for the presence of an extra methionine in roughly 50% of the molecules at the first N-terminal position. Protein synthesis in E. coli is initiated by an N-terminal methionine codon, and the extra methionine, or formylmethionine residue, at the N-terminus of the newly synthesized protein is processed/removed \textit{in vivo} in E. coli by the enzyme methionine aminopeptidase (MAP), an enzyme with absolute specificity for N-terminal methionine residues (Sherman et al., 1985). Though the N-terminal methionine residues of intrinsic E. coli proteins are ably removed by this constitutively synthesized enzyme present in the cells, it is well documented that many of the heterologous, over-expressed proteins in E. coli retain/partially retain the N-terminal methionine residue, possibly due to the inability of the limited amount of intrinsic MAP enzyme to process the over-expressed, heterologous protein (Sherman et al., 1985).
TABLE 5: Purification of intracellular, native-like recombinant SK from E. coli cells raised in 2 L medium after lysis with 6 M GdnHCl

<table>
<thead>
<tr>
<th>Purification step</th>
<th>Streptokinase activity (I.U.)</th>
<th>Protein (mg)</th>
<th>Specific activity (I.U./mg)</th>
<th>Yield (percent)</th>
<th>Fold purification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell lysate (intracellular extract)</td>
<td>$6.7 \times 10^6$</td>
<td>350</td>
<td>$1.7 \times 10^4$</td>
<td>100</td>
<td>-</td>
</tr>
<tr>
<td>Phenyl-agarose (HIC)</td>
<td>$4.76 \times 10^6$</td>
<td>93.6</td>
<td>$5.1 \times 10^4$</td>
<td>71</td>
<td>3</td>
</tr>
<tr>
<td>DEAE-Sepharose (Ionexchange)</td>
<td>$4.19 \times 10^6$</td>
<td>44</td>
<td>$0.95 \times 10^5$</td>
<td>62.6</td>
<td>6.5</td>
</tr>
</tbody>
</table>
**TABLE 6:** Purification of intracellularly over-expressed recombinant SK by a quick two-step procedure from 500 ml cell culture

<table>
<thead>
<tr>
<th>Purification step</th>
<th>Streptokinase activity (I.U.)</th>
<th>Protein (mg)</th>
<th>Specific activity (I.U./mg)</th>
<th>Yield (percent)</th>
<th>Fold purification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell lysate (intracellular extract)</td>
<td>$1.82 \times 10^6$</td>
<td>157</td>
<td>$0.11 \times 10^5$</td>
<td>100</td>
<td>-</td>
</tr>
<tr>
<td>Ammonium sulphate precipitation (60%)</td>
<td>$1.13 \times 10^6$</td>
<td>40</td>
<td>$0.28 \times 10^5$</td>
<td>62</td>
<td>2.3</td>
</tr>
<tr>
<td>POROS-D (ion-exchange)</td>
<td>$1.02 \times 10^6$</td>
<td>12</td>
<td>$0.85 \times 10^5$</td>
<td>54</td>
<td>6.04</td>
</tr>
</tbody>
</table>
Fig. 6. Purification of rSK expressed in *E. coli*. rSK was cloned and over-expressed in *E. coli* BL21 cells as soluble protein. The induced cells were lysed with 6M GdnHCl and the lysate was processed for purification. **A. Purification by a two-step column chromatographic procedure.** Soluble rSK was first purified on HIC (Phenyl-agarose), followed by further purification on anion-exchange chromatography (DEAE-Sepharose). **B. Purification by a rapid, single-step chromatographic procedure.** The lysate was precipitated using 60% ammonium sulphate and subjected to rapid ion-exchange chromatography on POROS-D matrix. The lanes in both A and B represent: 1- Standard molecular weight markers, 2- whole cell lysate of uninduced cells, 3- whole cell lysate of induced cells, 4- A- HIC purified lysate, B- pellet obtained after precipitation with 60% ammonium sulphate, and 5- ion-exchange purified rSK.
3.3.2.3 Development of a rapid, single-step method for the purification of SK or its full-length site-directed mutants

The two-step chromatographic procedure described above for the purification of rSK from the intracellular milieu of the *E. coli* cells yielded about 95-98 % pure rSK in substantial quantity (~20 mg/L of shake flask culture) and satisfactory yields (approximately 60 %). However, the process was labor-intensive and time-consuming, requiring several hours (8-10 h) for purification, after lysis of the cells. For quick screening of site-directed mutants of SK for altered activity, especially for those mutant proteins that may show reduced stability if purified over long periods, the need was felt for a quick and effective method of purification, within minimum possible duration. Ammonium sulphate precipitation to 60 % saturation, followed by ion-exchange chromatography at analytical levels by Perfusion Chromatography (POROS-D) yielded approximately 90 % pure SK within a duration of 2-3 h (Fig. 6B). Though only a two-fold purification was achieved at the ammonium sulphate precipitation stage, substantial amounts of unwanted molecules like amino acids, DNA, lipids etc would be expected to removed by this procedure. The purified product obtained by this rapid process was only of slightly lower purity (as judged by the presence of a few lower molecular weight bands on SDS-PAGE) than the two-stage process discussed above. Also, this process yielded reasonably high amounts of protein (24 mg/L of cell culture, and a final activity yield of approximately 50 %). Thus, the development of this quick and relatively easy method of purification, summarized in Table 6, considerably eased the purification procedures adopted for the rapid screening and analyses of various site-directed mutants. This fast screening protocol served to select mutants with less than native-like activity, as a prelude to a more detailed structure-function analyses of the ‘short-listed’ mutants.

3.3.3 Purification of truncated derivatives of SK

3.3.3.1 Purification of individual domains of SK

The ‘unidomain’ molecules α, αK59E and γ were produced as His6-tagged proteins to allow their rapid, affinity-based purification from crude lysates. A single-step purification of the dissolved IB (in case of the α domain constructs) or the soluble protein
Fig. 7. Purification of individual domains of SK from E. coli. The individual domains were cloned, intracellularly over-expressed and purified from the lysate of E. coli cells. The samples at various stages of purification were analyzed by SDS-PAGE on 15% acrylamide gels. 

A. Purification of α-domain. α-domain was purified on Ni$^{2+}$-IMAC affinity column and desalted on DEAE-Sepharose. The lanes represent: 1- Standard molecular weight (MW) markers, 2- whole cell lysate of induced cells, 3- inclusion body containing α-domain, 4- purified α-domain. 

B. Purification of β-domain. The β-domain was purified by negative chromatography on DEAE-Sepharose, followed by purification on HIC (Phenyl-agarose). Lanes represent: 1- MW markers, 2- whole cell lysate of induced cells, 3- soluble fraction of lysed cells, 4- purified β-domain. 

C. Purification of γ-domain. The γ-domain was purified from soluble fraction using Ni$^{2+}$-IMAC affinity column, as mentioned above for α-domain. The lanes represents: 1- MW markers, 2- whole cell lysate of induced cells, 3- soluble fraction of lysed cells, 4- purified γ-domain.
(in case of the γ domain) on Ni$^{2+}$-IMAC (immobilized metal affinity chromatography), resulted in purification of the over-expressed protein to levels of approximately 98% homogeneity (Fig. 7A and C). Table 7 summarizes the strategies adopted, and the final yields obtained after purification of the individual domains. A final adsorption-desorption step on DEAE-Sepharose was performed to remove the high concentrations of imidazole present in the elution buffer at the Ni$^{2+}$-IMAC step. Approximately 40% yield was consistently observed for both α and the γ domains, with an overall recovery of around 16 mg/L of cell-culture for the α domain and nearly double in the case of the highly over-expressed γ domain.

In the case of the over-expressed, soluble β domain (expressed without any affinity tag), purification was achieved by a combination of negative chromatography on an ion-exchange matrix (DEAE-Sepharose), followed by purification on Phenyl-agarose and a final desalting step using 3 kDa cut-off ultramembrane concentrators. The purified β domain obtained by this method (Fig. 7B) was more than 95% pure, with a yield of approximately 30% (Table 7). The process resulted in harvesting ~16 mg β/L of the cell culture. The molecular weights of the individual domains, as analyzed by SDS-PAGE, were consistent with the expected sizes of these domains. The purified proteins were further tested by Western blot analyses using polyclonal anti-SK anti serum as well as by automated N-terminus sequencing of the first ten amino acids to confirm the authenticity of the purified individual domains.

### 3.3.3.2 Purification of bidomain constructs, αβ and βγ

The purification of all the bi-domain constructs viz., αβ, αβK59E and βγ could be carried out successfully by the two-step purification strategy developed for the purification of rSK, i.e. HIC, followed by ion-exchange chromatography, with only minor modifications, such as the use of ammonium sulphate in place of NaCl in the binding and washing buffers for processing on HIC. The change in the buffer system at HIC stage was made to induce a tighter binding of these 'partial-length' proteins to the matrix. Purification of αβ (from the soluble fraction released after sonication) or βγ (as 6 M urea-dissolved IB, obtained after sonication of induced cells) was first performed on a Phenyl-agarose matrix. The bound proteins, which were eluted in 20 mM buffer wash, were
Fig. 8. Purification of the bi-domain constructs of SK (αβ and βγ). Both the bi-domain constructs were purified from IPTG induced E. coli cells, after lysis, on HIC (Phenyl-agarose), followed by ion-exchange on DEAE-Sepharose column. The purified fractions were analyzed by 12.5% SDS-PAGE. **A. Purification of αβ.** Lanes represent: 1- whole cell lysate of uninduced cells, 2- whole cell lysate of induced cells producing αβ, 3- purified αβ after the ion-exchange stage. **B. Purification of βγ:** The lanes represent: 1-Standard molecular weight markers, 2- whole cell lysate of induced cells producing βγ, 3- IBs containing βγ and 4- purified βγ obtained after elution from DEAE-Sepharose.
immediately subjected to ion exchange chromatography on DEAE-Sepharose. In the purification of αβ, the development of the DEAE-Sepharose chromatography using a linear gradient of NaCl resulted in elution of this protein at low salt concentration (0.15 M NaCl). A total recovery of 5 mg αβ/L of cell culture, and an overall yield of approximately 20 % (Table 8) was achieved by this two-stage purification process. This two-stage chromatographic procedure resulted in obtaining an αβ preparation that was approximately 95 % pure (Fig. 10). The low recovery of this bi-domain construct mainly stems from the use of strictly the ascending fractions of the eluting peak for future use, due to the simultaneous elution of contaminating *E. coli* proteins with the same amount of salt, as αβ.1

Protein βγ, on the other hand, eluted from the ion-exchange column at an NaCl concentration of approximately 0.25 M (close to that of nSK), with a final recovery of around 20 mg/L of the starting culture, and an approximate yield of 35 % (Table 8). The final protein obtained was judged on SDS-PAGE to be over 95 % purified (Fig. 11). The purified proteins were further tested by Western blot analysis using polyclonal anti-SK antiserum, as well as by automated N-terminus sequencing of the first ten amino acids to confirm the authenticity of the purified bi-domain molecules. CD analysis of the purified individual and bi-domain molecules were also carried out to check for the presence of native-like structure.

3.3.4 Secondary structure analysis by Circular Dichroism

CD studies in the far-UV wavelength region, using dilute solutions (0.15 mg/ml concentration) of the various proteins was carried out to assess the secondary structure content in the various site-specifically altered/truncated derivatives of SK. The secondary structure content in rSK calculated from our CD data overall matched with the previously reported values for secondary structure of SK (Table 1, Chapter II), and also closely matched with the values expected on the basis of the crystal structure (Table 8). The secondary structure of each of the cloned, expressed and purified domain/domain combinations exhibited characteristic CD spectra, the analyses of which using standard algorithm (Yang et al., 1986) revealed substantial concurrence with the secondary structure content calculated from the crystal structure using the ProMotif programme (Hutchinson
**TABLE 7: Expression, purification strategy and characteristics of individual domains of SK produced intracellularly in *E.coli* BL-21 cells**

<table>
<thead>
<tr>
<th>Domain</th>
<th>Sequence</th>
<th>Type and level of expression</th>
<th>Purification strategy</th>
<th>Final yield (mg) / L of cells raised</th>
</tr>
</thead>
</table>
| α (His6-tag) | SK1-143  | Inclusion body (IB)           | • Dissolution of IB in 8 M Urea  
• Chromatography on Ni\(^{2+}\)-IMAC affinity resin  
• Desalting on DEAE-Sepharose column | 15                                   |
| β       | SK143-293| Soluble protein               | • Negative chromatography of soluble lysate fraction on DEAE-Sepharose column  
• Hydrophobic interaction chromatography  
• Desalting on ultramembrane concentrators | 16.4                                 |
| γ (His6-tag) | SK300-380| Soluble protein               | • Ni-IMAC chromatography of soluble fraction  
• Desalting on DEAE-Sepharose           | 32                                   |
**Table 8:** Expression, characteristics and purification strategy adopted for two-domain construct of Sk over-expressed intracellularly in *E. coli* BL21 cells

<table>
<thead>
<tr>
<th>Domain</th>
<th>Sequence</th>
<th>Type and level of expression</th>
<th>Purification strategy</th>
<th>Final yield (mg) / L of cells raised</th>
</tr>
</thead>
</table>
| αβ     | SK1-293  | Soluble protein ~12%         | - Hydrophobic interaction chromatography  
|        |          |                             | - Followed by ion-exchange on DEAE-Sepharose column | 5 |
| βγ     | SK143-414| Inclusion body (IB) ~35%     | - Dissolution of IB in 6 M urea  
|        |          |                             | - Hydrophobic interaction chromatography  
|        |          |                             | - Followed by ion-exchange on DEAE-Sepharose column | 24 |
Table 9: Secondary structure content detected by CD studies in the truncated derivatives of SK

<table>
<thead>
<tr>
<th>PROTEIN</th>
<th>SOURCE</th>
<th>% α-HELIX</th>
<th>% β-SHEET</th>
<th>% β-TURN</th>
<th>% OTHERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>α DOMAIN (1-143)</td>
<td>Recombinant*</td>
<td>9.4</td>
<td>51.2</td>
<td>9.4</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Crystal structure†</td>
<td>14.1</td>
<td>34.9</td>
<td>23.6</td>
<td>27.4</td>
</tr>
<tr>
<td>β DOMAIN (143-293)</td>
<td>Recombinant</td>
<td>13.8</td>
<td>53.3</td>
<td>0</td>
<td>32.9</td>
</tr>
<tr>
<td></td>
<td>(a)§</td>
<td>10.5</td>
<td>35.3</td>
<td>0</td>
<td>54.2</td>
</tr>
<tr>
<td></td>
<td>Crystal structure (b)†</td>
<td>11.2</td>
<td>41.1</td>
<td>0</td>
<td>47.7</td>
</tr>
<tr>
<td>γ DOMAIN (300-387)</td>
<td>Recombinant</td>
<td>3.9</td>
<td>67.9</td>
<td>0</td>
<td>28.2</td>
</tr>
<tr>
<td></td>
<td>Crystal structure†</td>
<td>7.6</td>
<td>40.5</td>
<td>30.4</td>
<td>21.5</td>
</tr>
<tr>
<td>βγ DOMAIN (143-414)</td>
<td>Recombinant</td>
<td>20.3</td>
<td>58.3</td>
<td>0</td>
<td>21.4</td>
</tr>
<tr>
<td></td>
<td>Crystal structure†</td>
<td>9.1</td>
<td>37.9</td>
<td>15.2</td>
<td>37.9</td>
</tr>
<tr>
<td>αβ DOMAIN (1-293)</td>
<td>Recombinant</td>
<td>7.1</td>
<td>62</td>
<td>9.6</td>
<td>21.3</td>
</tr>
<tr>
<td></td>
<td>Crystal structure†</td>
<td>12.3</td>
<td>35.1</td>
<td>11.8</td>
<td>40.8</td>
</tr>
<tr>
<td>SK</td>
<td>Native SK</td>
<td>17.6</td>
<td>15.4</td>
<td>24.3</td>
<td>42.7</td>
</tr>
<tr>
<td></td>
<td>Crystal structure†</td>
<td>10.7</td>
<td>36.9</td>
<td>18</td>
<td>34.4</td>
</tr>
</tbody>
</table>

- Secondary structure calculated from CD data of recombinant fragments using algorithm as described by Yang et al., 1986.

- Secondary structure calculated from the SK-μPN crystal structure (Wang et al., 1998) using programme ProMotif (Hutchinson and Thornton, 1996).

- Calculated from the independent crystal structure of β domain (Wang et al., 1999) using programme ProMotif.
and Thornton, 1996) (Table 9). These results, thus, strongly indicate that the procedures adopted for isolation of the heterologously expressed domains/domain combinations yielded pure and structurally intact proteins, which could be directly used for the study of structure-function aspect in the functioning of SK.

3.4 Discussion

In this chapter, protocols for the facile purification of rSK or its mutants have been described. In addition, protocols for the purification of individual subunits of SK (α, β or γ), or their two-domain combination (αβ or βγ; residues 1-293 and 143-414, respectively) have also been presented. Succinctly, these studies lay a firm foundation for a detailed structure-function investigation into the fundamental mechanistic question posed in this treatise. The methods developed herein resulted in the purification of SK and its derivatives in good yields and structurally intact molecular forms. Examining the secondary structures of the purified molecules by CD studies, revealed that rSK possessed structural features that were largely similar to those detected previously by CD studies of full-length native SK (Radek and Castellino, 1989), as well as to the secondary structure calculated from X-ray co-ordinates using ProMotif programme (Hutchinson and Thornton, 1996). However, while the secondary structure detected in the domain moieties by the CD studies revealed features in these molecules that correlated reasonably well with the secondary structures calculated from the crystal structure, the structural features of isolated domains showed a few distinct differences with the features detected in isolated domains produced by proteolytic studies and purified using reverse-phase HPLC (Parrado et al., 1996). It needs to be mentioned that the previous procedures adopted for isolation of the independent domains (proteolysis, use of reverse-phase HPLC for purification of the proteolytic fragments and lyophilization of purified and salt-free fragment), can cause irreversible damage to the structure of the protein molecules since purification of the proteolytically digested fragments involves the use of harsh acids and organic solvent (TFA and acetonitrile) that might possibly distort the original secondary structure. It is also possible that these denatured isolated domains, on reconstitution in aqueous solution, might only partially renature to the original structural forms. Hence, the use of spectroscopic studies like CD, which provides the average structural feature of the
molecular ensemble being tested (denatured as well as renatured protein), would yield an average data of structure of the denatured as well as the renatured proteins. Thus, structural determination using these proteolytically prepared fragments might not necessarily indicate the original structural feature of these fragments.

In our study, we took recourse to the use of recombinant tools for the production of individual domains, and employed aqueous chromatographic procedures for the isolation of these. Hence it was reasonable to expect the presence of native secondary structure in these molecules. It needs to be mentioned here that some of the domain molecules (α and βγ) were over-expressed as IBs, which necessitated the obligatory use of chaotropic agents for dissolution of the IBs, prior to purification. However, the deliberate use of chaotropic agents during purification in a few test cases (β, βγ), followed by the detection of similar secondary structure by CD studies, and the detection of similar functional characteristic (as presented in the subsequent chapter), in both the chaotropic treated and the non-chaotropic exposed domain molecules strongly points to the authenticity of the purification procedure adopted, and the secondary structure determined in all the domain constructs, including those exposed to 8M urea.

The purification procedure developed herein involved isolation of desired protein molecules from the intracellular milieu of E. coli. In general, this entailed purification from a complex background containing a number of intrinsic E. coli proteins and from a milieu containing many proteases, which could potentially degrade the over-expressed proteins, especially the truncated derivatives, which might be potentially more vulnerable to proteolysis. Nevertheless, the procedures developed herein resulted in purifying to homogeneity not only of full-length SK and various site-specific mutants of SK, but also its truncated derivatives, establishing the effectiveness of the purification schemes evolved. The rapid two-step chromatographic procedure standardized for the purification of intracellular SK from E. coli resulted in removal of predominantly large quantities of unwanted contaminating, cellular proteins from E. coli at the HIC step, resulting in the product being over 80% pure. The inordinately high efficiency of SK purified by this scheme resulting in 3-4 fold improvement in purity, likely results from a very strong and
selective binding of SK to the HIC matrix, and its easy elution by distilled water. Incidentally, higher homologues of HIC eg. Decyl- and Octyl-agarose also bound SK very tightly but elution from these matrices could be achieved only with chaotropic agents like 8 M urea. On the other hand, Hexyl- and Butyl-agarose bound SK too weakly, so that it eluted along with a number of background E. coli proteins, in a much cruder form. Thus ‘Phenyl’ proved to be the best ligand for the capture and release of SK. The results obtained with purification on HIC clearly demonstrate ease and effectiveness of this chromatographic procedure in the production of stable and pure SK with very high specific activity, and good recovery. In addition, the compatibility of this technique to the simple and relatively economical cell lysis procedure using 6M GdnHCl, added to the valuable use of this chromatography. Further, the successful use of HIC based technique for purification of most of the truncated derivatives of SK as well, demonstrated the power of utilizing HIC as a generalized purification tool for isolating SK and its derivatives. Additional purification on a second chromatographic step viz., chromatography on DEAE-Sepharose enhanced the purity of SK as well as it derivatives to virtual homogeneity. In the case of SK, a specific activity of $0.95 \times 10^5$ IU/mg could be achieved that was close to the highest reported for SK ($1.00 \times 10^5$ IU/mg) allowing the use of this product directly for kinetic analyses, with no further requirement for any polishing step.

Overall, the CD data supports the conclusion that the purification procedure evolved for the purification of SK/SK derivatives resulted in preparation of pure proteins in high yield, and in structurally intact form. The development of facile, rapid, purification methodologies for the isolation and characterization of SK, site-specifically altered directed mutants of SK, as well as truncated derivatives of SK significantly facilitated the study of these constructs in understanding the functional correlation of these structural entities.