*Journal of Biomolecular NMR*, **26:** 317–326, 2003.
*KLUWER/ESCOM*



# Efficient expression of isotopically labeled peptides for high resolution NMR studies: Application to the Cdc42/Rac binding domains of virulent kinases in *Candida albicans*[*]

Michael J. Osborne[**,***], Zhengding Su[***], Vasanth Sridaran and Feng Ni[**]
*Biomolecular NMR and Protein Research Group, Biotechnology Research Institute, National Research Council of Canada, 6100 Royalmount Avenue, Montreal, Quebec, H4P 2R2, Canada*

Received 20 December 2002; Accepted 7 April 2003

*Key words:* isotope labeling, NMR, peptide expression

**Abstract**

The production of bioactive peptides and small protein fragments is commonly achieved via solid-phase chemical synthesis. However, such techniques become unviable and prohibitively expensive when the peptides are large (e.g., >30 amino acids) or when isotope labeling is required for NMR studies. Expression and purification of large quantities of unfolded peptides in *E. coli* have also proved to be difficult even when the desired peptides are carried by fusion proteins such as GST. We have developed a peptide expression system that utilizes a novel fusion protein (SFC120) which is highly expressed and directs the peptides to inclusion bodies, thereby minimizing in-cell proteolysis whilst maintaining high yields of peptide expression. The expressed peptides can be liberated from the carrier protein by CNBr cleavage at engineered methionine sites or through proteolysis by specific proteases for peptides containing methionine residues. In the present systems, we use CNBr, due to the absence of methionine residues in the target peptides, although other cleavage sites can be easily inserted. We report the production of six unfolded protein fragments of different composition and lengths (19 to 48 residues) derived from the virulent effector kinases, Cla4 and Ste20 of *Candida albicans*. All six peptides were produced with high yields of purified material (30–40 mg/l in LB, 15–20 mg/l in M9 medium), pointing to the general applicability of this expression system for peptide production. The enrichment of these peptides with $^{15}$N, $^{15}$N/$^{13}$C and even $^{15}$N/$^{13}$C/$^{2}$H isotopes is presented allowing speedy assignment of poorly-resolved resonances of flexible peptides.

*Abbreviations:* CNBr – cyanogen bromide; NMR – nuclear magnetic resonance; NOE – nuclear Overhauser effect; RP-HPLC – reversed-phase high pressure liquid chromatography; GST – glutathione-S-transferase; Ni-NTA – nickel nitrilotriacetic acid; PCR – polymerase chain reaction; CRIB – *C*dc42/*R*ac - *I*nteractive *B*inding; mCRIB – minimal CRIB motif; cCRIB – peptide sequence comprising the C-terminal extension to the CRIB motif; eCRIB – a peptide comprising the mCRIB and cCRIB sequences; OD – optical density; *E. coli* – *Escherichia coli*.

## Introduction

Peptides constitute a group of biomolecules for which there is an increasing demand in many fields of biological, medical and pharmaceutical research. Peptides are products of protein splicing but also have roles in translation (Gong and Yanofsky, 2002) and other important biological processes such as controlling enzymatic activity (Fox et al., 1992). Peptides have been implicated in a variety of diseases, including (among others): diabetes (Tyrberg and Levine, 2001), neurodisorders (Jacobsen, 2002), asthma (Bolin et al., 1995), renal disease (Meyer et al., 1998), certain cancers (Disis et al., 2001) and heart failure (Nathisuwan and Talbert, 2002). Interest in pharmaceutical peptides

[*]The Canadian Crown right to retain a non-exclusive royalty-free license in and to any copyright is acknowledged.
[**]To whom correspondence should be addressed.
[***]These authors contributed equally to this work.

has been longstanding, insulin being an obvious example, but problems with production, administration and bioavailability has precluded major advances in the past. Recent developments, particularly in per-oral delivery (Sood and Panchagnula, 2001) and understanding of endothelial, epilethlial and enzymatic interference, have stimulated a new model of clinical treatment, termed 'biotherapy', based on synthetic peptides and their involvement in biological events (Lauta, 2000).

The increasing popularity of large-scale genomic and proteomic programs is further likely to increase the demand for peptides in target validation. Peptides are very useful tools for investigating protein-protein interactions. Indeed, the bound conformations of peptides in complex with target proteins are commonly used as templates for drug discovery (Mazitschek et al., 2002). In addition there is increasing evidence for the existence of peptide-like, or naturally unfolded, proteins which are encoded by the genomes and endowed with critical functional activities (Wright and Dyson, 1999; Uversky, 2002; Dunker et al., 2002).

Given the increasing demand for peptides in pharmaceutical and biotechnology research, it is somewhat surprising that the main source of peptides still comes from synthetic techniques. Although solid-phase synthesis can produce good yields of peptides, the cost of synthetic peptides becomes unviable and/or prohibitively high when the desired peptide is greater than 30 residues. Moreover, uniform isotopic enrichment with $^{15}N/^{13}C$ or $^{2}H$ for NMR studies is practically impossible for larger peptide fragments by solid state synthesis. Majerle et al. (2000) have demonstrated that isotope-labeled peptides could be prepared based on an expression system first described by Kuliopulos et al. (1994) using a fusion-protein system together with CNBr cleavage for peptide release. It was shown that recombinant peptide production had potentially many advantages over the solid-phase method of peptide synthesis, especially for isotope-labeled peptides of 10 residues in size. However, there is still the question of expression yields and whether the available method is suitable for the production of peptides of larger sizes.

A number of other publications have reported alternative peptide expression systems and described their utility for the production of one or two specific peptides (Kuliopulos et al., 1994; Baker et al., 1996; Campbell et al., 1997; Kohno et al., 1998; Jones et al., 2000; Fairlie et al., 2002; Sharon et al., 2002; Lindhout et al., 2002; Sprules et al., 2003). In general, the target peptides are fused to a highly expressed carrier protein in order to overcome the problem of low yields of peptide production. In some cases a carrier protein with low solubility has been exploited to direct the peptide to the inclusion bodies, thereby minimizing proteolysis and simplifying purification (Kuliopulos et al., 1994, 1998; Majerle et al., 2000; Jones et al., 2000). In this paper, we report the use of a novel carrier protein derived from the N-terminal oligonucleotide binding domain of *S. nuclease* for peptide expression. This protein domain is small with 120 residues and is highly expressed into the inclusion bodies. We show that this new expression system generates superior or at least similar yields of purified peptides for a variety of peptide sequences of differing lengths and amino acid compositions (Figure 1). In particular, the high yields of peptide production allowed isotopic enrichment, including triple-labeling with $^{15}N/^{13}C/^{2}H$, of the targeted peptides needed for the study of peptide-protein complexes by use of NMR spectroscopy.

## Methods

### *Construction of the fusion protein expression vector*

A small fusion sequence, termed SFC120, was used as the carrier protein to express the six CRIB peptides summarized in Figure 1. SFC120 was adopted from the N-terminal oligonucleotide binding domain of *S. nuclease* which comprises 120 residues. The cDNA was amplified by standard PCR methods while the restriction enzyme site of *Nco I* was generated in the 5′-end and the two restriction enzyme sites of *EcoR I* and *BamH I* were generated in the 3′-end. The PCR product was double-digested by *Nco I* and *BamH I* and ligated into the pET15M vector, which was modified from the pET-15b vector (Novagen) by removing the *EcoR I* site. The constructed fusion protein expression vector was defined as *pTSN-6A* (Figure 2).

The DNA fragments encoding the CRIB peptides from Cla4 and Ste20 were amplified from a cDNA library by PCR or synthesized as oligonucleotides using the codon preference of *E. coli*. The DNA fragments were digested with *EcoR I* and *BamH I*, and subcloned into the *pTSN-6A* vector. The expression constructs were transformed into the BL21(DE3) expression host and confirmed by DNA sequencing. A single methionine residue was inserted between the SFC120 fusion protein and the desired peptide sequence to facilitate release of the peptides by CNBr cleavage. A His-tag

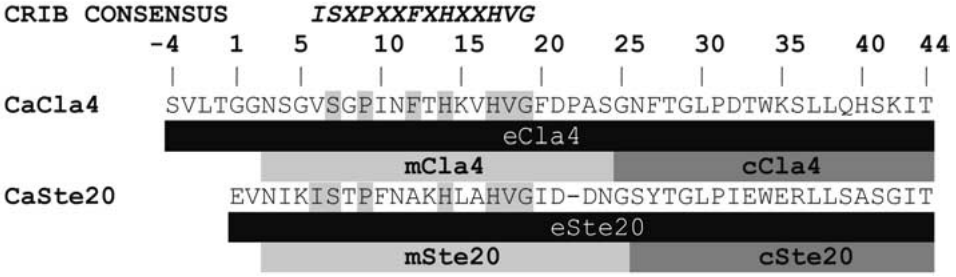


*Figure 1.* Summary of the CRIB peptide fragments of the *Candida* Ste20 and Cla4 proteins chosen for expression. The 'extended' CRIBs, eCla4 and eSte20 comprise 48 and 44 residues, respectively. The minimal CRIB fragments, mCla4 and mSte20, comprise 21 residues and contain the CRIB consensus sequence (highlighted). The C-terminal CRIB fragments (cCla4 and cSte20) are derived from the sequence segments to the C-terminus of the consensus CRIB motif.

with six histidines can be placed at the N-terminus of SFC120 to simplify purification of the fusion protein by adsorption onto a Ni-NTA agarose column (QIAGEN). In the present case, the His-tag SFC120 vector was used to express the eCRIB fragments (eSte20 and eCla4). The generic non-His-tag SFC120 was used to express the mCRIB and cCRIB fragments (mCla4, mSte20, cCla4, and cSte20).

*Expression of isotopically enriched fusion protein-CRIB fragments in* E. coli*: $^{15}N$ and $^{13}C/^{15}N$ CRIB peptides*

Expression of the peptide fragments was achieved by transformation of the appropriate plasmid into *E. coli* BL21(*DE3*) competent cells. An overnight culture grown in 2YT containing 100 μg/ml ampicillin (25 ml) was used to inoculate 1 l of M9 minimal media (100 μg/ml ampicillin) supplemented with BME vitamins solution (10 ml/l of 100x stock – SIGMA). $^{15}$N-labeled peptides (mCla4, mSte20, cCla4, cSte20, eSte20, eCla4) were expressed using $^{15}(NH_4)_2SO_4$ (2 g/l) as the sole nitrogen source. Uniformly labeled $^{15}N/^{13}C$ eSte20 was expressed using $^{15}(NH_4)_2SO_4$ (2 g/l) and $^{13}C_6$ glucose (2 g/l) as the sole nitrogen and carbon sources in the M9 media, respectively. The cells were grown at 37 °C to a cell density of $OD_{600} = 0.8$ and induced by adding IPTG to a final concentration of 1 mM. The cells were incubated for 4–12 h at 37°C and collected by centrifugation (8000 rpm for 20 min).

*Expression of $^{2}H/^{15}N/^{13}C$-labeled eCla4*

A uniformly $^{15}N/^{13}C$ enriched eCRIB peptide from Cla4 (Figure 1) was obtained with 70% random fractional deuteration by growing the cells in 99.9% $D_2O$ on minimal medium (with BME vitamins and ampicillin) containing 2 g/l $(^{15}NH_4)_2SO_4$ and 2 g/l $^{13}C_6$ glucose as the sole sources of nitrogen and carbon. Prior to growth on this medium the cells were adapted to $^{2}H$ in the following manner (100 μg/ml ampicillin was used throughout and all minimal media solutions contained $(^{15}NH_4)_2SO_4$, $^{13}C_6$ glucose and BME vitamins). First a colony was picked from a LB plate and grown in 3 ml of LB for 3 h or until $OD_{600} = 0.5$. A 50 μl aliquot was transferred to 3 ml of minimal media in 100% $H_2O$ and grown at 37 °C until $OD_{600} = 0.5$, when another 50 μl aliquot was transferred to a 3 ml culture of 50% $D_2O$ minimal media. After growth at 37 °C to $OD_{600} = 0.5$, a 200 μl aliquot was used to inoculate 25 ml of 99.9% $D_2O$ at 37 °C. At an $OD_{600} = 0.5$ this 25 ml solution was used to inoculate 1 l of minimal media in 99.9% $D_2O$, which was then induced with 1 mM IPTG at $OD_{600} = 0.7$ and harvested after 12–16 h growth at 37 °C by centrifugation. A summary of the purification protocols for the peptides are summarized in Figure 3 and described in detail in the following sections.

*Purification of non His-tag fusion peptides: mCRIB and cCRIB peptides*

Thawed cell pellets were resuspended in 6 M urea in 20 mM Tris, 100 mM NaCl buffer, pH 8.0 for 4 h and then sonicated for 45 s on ice. The solution was then centrifuged at 7 K rpm for 20 min. An equivolume of 100% cold ethanol was added to the supernatant and the solution allowed to stand at 4 °C for at least 2 h. After centrifugation, the pellet containing precipitated DNA and large proteins was discarded and another equivolume of cold ethanol was added to the collected supernatant and allowed to stand overnight. The solution was centrifuged at 8,500 rpm and the pellet containing the relatively pure fusion-peptide fragment subjected to SDS-PAGE analysis. If necessary the pel-

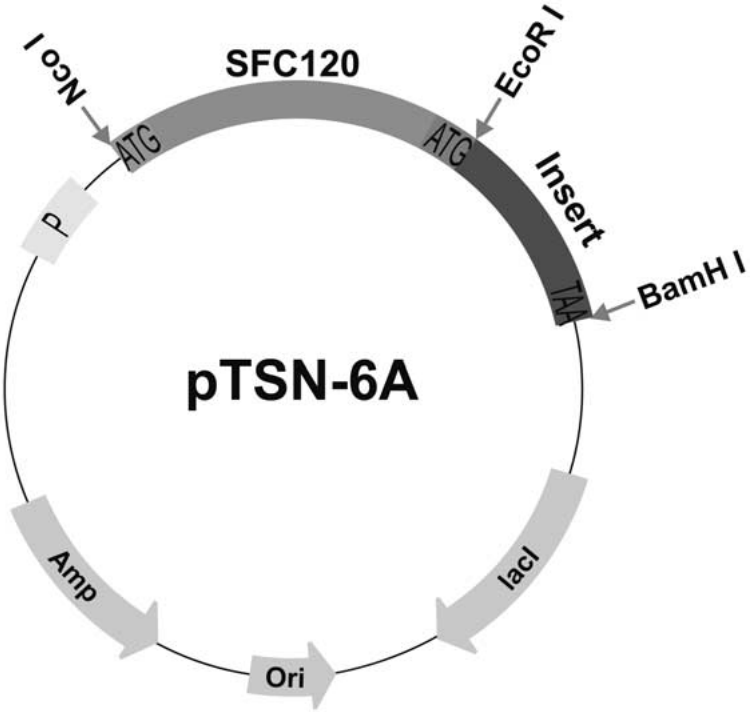


*Figure 2.* The Construction of the expression vector (*pTSN-6A*) for expression of the CRIB peptides. One unique protein domain, SFC120, was used as the carrier protein. P denotes the promoter, either *T7* or *Tac*. His-tag with six histidines can be placed at either N-terminal or C-terminal side of the SFC120 carrier protein to simplify the purification step.

let was resuspended in 6 M urea and applied to a Sep-Pak column (Waters) to remove any impurities. The fusion protein was then lyophilized.

*Purification of His-tag fusion peptides: eCRiBs*

Cell pellets were resuspended in 6 M urea in Tris-HCL buffer at pH 8.0 by gentle shaking for $\sim$ 4 h and briefly sonicated on ice. After centrifugation at 7 K rpm for 20 min the supernatant was applied to a Ni-NTA agarose column (QIAGEN) previously equilibrated with the lysis buffer. The column was then washed with $\sim$ 20 column volumes of 6 M urea in Tris buffer at pH 6.3 to eliminate non-specific binding to the column. The His-tagged fusion protein was then eluted with 6 M urea in 20 mM Tris buffer at pH 4.5. The solubilized fusion protein was then lyophilized to dryness.

*Peptide cleavage and Reversed-Phase HPLC purification*

CNBr cleavage was used to release the target peptide from the fusion protein. The fusion protein was dissolved in 70% TFA and CNBr added to a final molar ratio of 100:1 and the solution allowed to stand for $\sim$ 24 h. The samples were then diluted with water ($\times$10) and lyophilized to dryness and purified by RP-HPLC on a C18 column using an acetonitrile-water gradient containing 0.1% TFA. The peptides were lyophilized and confirmed by electrospray mass spectrometry.

*NMR sample preparation*

Free peptides were prepared for NMR by resuspending the lyophilized peptides into the appropriate NMR buffer (50 mM Phosphate, 50 mM NaCl, 5 mM $MgCl_2$, 0.02% $NaN_3$, pH 6.0, 95% $H_2O$/5% $D_2O$). Typical concentrations of the NMR samples was 1.2 mM.

The complex of unlabelled Cdc42 with $^{2}H/^{13}C/^{15}N$-eCla4 or $^{15}N$-eSte20 was formed by adding a slight excess of Cdc42 (activated with the GTP analogue GMPPCP) to the labeled peptide in $\sim$ 10 ml of NMR buffer (details of the Cdc42 purification and activation will be published elsewhere). After 3 h incubation, the complex was concentrated to $\sim$ 400 $\mu$l with a YM-10 centricon. The final concentration of the peptides in complex with Cdc42 was $\sim$ 0.7 mM.

*NMR measurements*

NMR spectra were recorded at 25 °C on a Bruker DRX 500 spectrometer equipped with a triple resonance probe and triple axis pulsed field gradients. $^{1}H$-$^{15}N$ heteronuclear single quantum coherence (HSQC) spectra (Bodenhausen and Ruben, 1980) were acquired with States-TPPI to achieve quadrature detection in the indirect dimension (Marion et al., 1989). Coherence selection was achieved using pulsed field gradients (Muhandiram and Bax, 1994). A flip back pulse and coherence selection was used to minimize the water signal (Grzesiek and Bax, 1993). Sweep widths and $t_1$ increments varied depending on the peptide. However, resolution in the indirect dimension was typically 10 Hz/point prior to processing. Typically 4–8 scans were acquired per $t_1$ increment. Data sets were zero-filled once in each dimension and apodized with a 90 ° shifted sine-squared function. $^{1}H$ chemical shifts were referenced to external DSS and the $^{15}N$ shifts referenced indirectly.

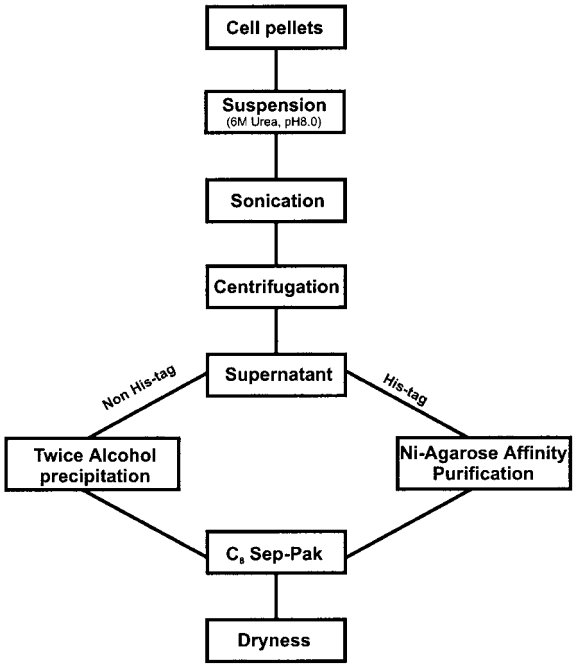


*Figure 3.* Summary of the purification protocols for the SFC-fusion protein and the CRIB fragments.

## Results and discussion

*Choice of peptides and strategy*

Figure 1 summarizes the fragments of the Ste20 and Cla4 proteins chosen for expression based on interactions with Cdc42 reported for highly homologous kinase-Cdc42 interactions in humans (Thompson et al., 1998; Zhang et al., 1997; Zhao et al., 1998; Stevens et al., 1999). The peptide sequences termed as the extended-CRIB fragments (eCla4 and eSte20, Figure 1), which comprise the CRIB motif and residues to its C-terminus, exhibit similar affinity for Cdc42 compared to the full-length kinases. The high-affinity e CRIB fragments were then separated into two fragments, comprising the minimal CRIB motif (mCla4 and mSte20) and the C-terminal fragments (cCla4 and cSte20). In total, six target peptide sequences were chosen for expression and purification using the new expression system: eSte20, mSte20, cSte20, eCla4, mCla4 and cCla4 (Figure 1).

Our previous attempts to express and isotopically enrich recombinant eCRIB fragments from Ste20 and Cla4 had limited success. Expression as a His-tag was unsuccessful, presumably due to in-cell proteolysis. Using a GST-fusion protein, as reported by other groups for human eCRIB peptides (Abdul-Manan et al., 1999; Mott et al., 1999; Morreale et al., 2000; Gizachew et al., 2000) was unsuccessful except for the eCla4 fragment; but even in this case the yield was poor such that isotopic enrichment was viable only using the isotope-labeled rich media, e.g., BIOexpress (CIL) or Celtone (Martek). Since we wanted to study a large number of small CRIB peptides by heteronulcear NMR, the use of such media is prohibitively expensive. To overcome the poor peptide expression, a problem commonly encountered in recombinant peptide production, we used an insoluble fusion protein that is expressed in high yields in *E. coli* but is directed to the inclusion bodies, thereby minimizing protease degradation in the cell whilst maintaining high yields of peptide expression (Kuliopulos et al., 1994; Majerle et al., 2000; Jones et al., 2000). Direction of the fusion peptide to the inclusion bodies also guarantees relatively pure fusion protein and thus simple purification by ethanol precipitation. A His-tag sequence can be further introduced (as demonstrated later for the eCRIB peptides) to simplify purification of the fusion protein.

*Expression and purification of SFC120-peptide fusion proteins*

Examples of the high level of expression of the SFC120-peptide fusion proteins are shown in Figures 4A and 4C, for the cCla4 (non-His tag) and eSte20

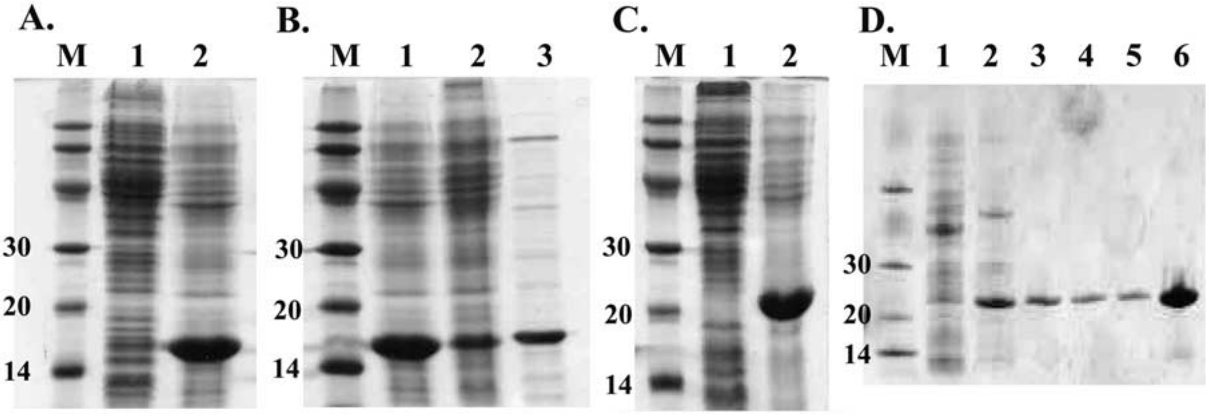


*Figure 4.* SDS-PAGE gel of the fusion protein expression. (A) expression of SFC120-mCla4. M: Molecular weight marker; 1: Uninduced; 2: ITPG induced. (B) purification of SFC120-mCla4. M: Molecular weight marker; 1: Cell lysate. 2: Pellet after 1st alcohol precipitation; 3: Pellet after 2nd alcohol precipitation. (C) Expression of His-SFC120-eSte20. M: Molecular weight marker; 1: Uninduced; 2: ITPG induced. (D) Purification of His-SFC120-eSte20 from Ni-NTA agarose column. M: Molecular weight marker; 1: Flow through; 2–5: Wash fractions, 6: Elution.

(His tag) peptides, respectively. High expression of the desired fusion protein is evident after induction with IPTG. The amount of fusion protein for both expressions can be estimated to be $\sim$ 60% of the total protein content (from SDS-PAGE). All the CRIB peptide fragments exhibited similar high expression profiles, indicating the generality of the SFC120 sequence as an efficient carrier for peptides of variable lengths. Relatively pure SFC120-peptide fusion protein is desired to eliminate side products from the CNBr cleavage. The purification protocols differed depending on the fusion protein construct (Figure 3). The His-tag fusion protein (e.g., eSte20, Figures 4C and 4D) was easily purified using the $Ni^{2+}$ affinity column (Figure 4D). A small amount of fusion protein was lost whilst washing extensively with 6 M urea pH 6.3 and can be minimized by reducing the wash volume (Figure 4D). However, the amount lost during the wash stage is minor compared to the yield obtained after elution with 6 M urea at pH 4.5 (Figure 4D, lane 6). The eluted fusion protein was $>$ 98% pure (estimated from SDS-PAGE).

Direction of the SFC120-peptide fusion protein to the inclusion bodies meant that virtually pure fusion protein for the non His-tag fusion protein (cCRIBs and mCRIBs) could be obtained after two rounds of ethanol precipitation (Figure 4B shows cCla4 purification). We found that for 1 l of cells it was best to resuspend the thawed pellets in 60–100 ml of 6 M urea to obtain pure peptide after the second ethanol precipitation. In general only a small amount of fusion protein remained in the first ethanol precipitation (see Figure 2B, lane 2) and, if necessary, can be reclaimed virtually pure by Sep-Pak purification. The majority of the fusion protein is precipitated after the second addition of ethanol, and is the major species present ($>$ 95%). The minor impurities were successfully separated from the fusion protein by application over a Sep-Pak (Waters) column. However, in this case removal did not affect quality of the final pure peptide.

*Cleavage and purification of the CRIB peptides*

The desired CRIB peptide was liberated from the SFC120 carrier protein by CNBr cleavage. As described above, only pure fusion proteins were subjected to CNBr cleavage since potential cleavage of contaminating proteins containing methionine residues may interfere with the purification of the desired peptide in subsequent HPLC protocols.

Another potential problem with CNBr cleavage can arise from the cleavage at Met sites within the carrier (SFC120) protein. In this study, the carrier protein contained two Met residues, resulting in three peptide fragments from the carrier protein alone after CNBr cleavage. With the exception of the eCla4 peptide, overlap with cleavage products of the carrier protein was not a problem. In all cases the major product of CNBr cleavage was the desired CRIB peptide, which was confirmed by electrospray mass spectroscopy. The high purity of the CRIB fragments after HPLC purification is exemplified by their signature $^{1}H$-$^{15}N$ HSQC NMR spectra (Figures 5 and 6) which exhibit no detectable impurities. The eCla4 fragment was the only

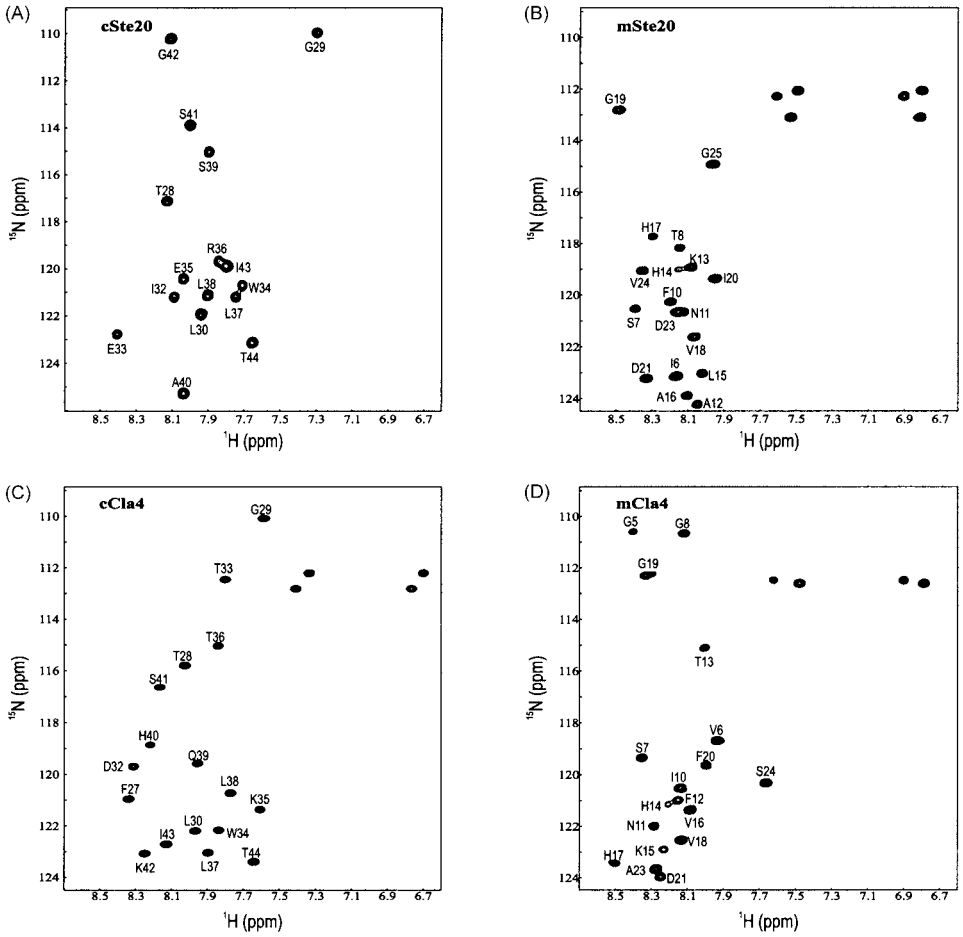


*Figure 5.* $^{1}$H-$^{15}$N HSQC spectra of $^{15}$N-labelled mCRIB and cCRIB fragments of Ste20 and Cla4. (A) cSte20, (B) mSte20, (C) cCla4, (D) mCla4. Assignments (see Figure 1 for numbering) based on $^{15}$N edited 3D TOCSY and NOESY experiments are shown.

peptide that required further purification due to overlap with CNBr cleavage products from the SFC120 carrier protein. Pure eCla4 was obtained by using His-tagged Cdc42 (activated with GMPPNP) attached to a $Ni^{2+}$ agarose column as a pulldown for eCla4 (unpublished observations). Such a technique can be employed as an alternative method for purifying peptides after CNBr cleavage if they have high affinity for a target protein.

Typical yields of the *purified* peptides chosen for expression in this study were very high, ranging from 30–40 mg/l in LB medium. Significantly, high quantities of *pure* peptides were also obtained from growths in M9 minimal media (15–20 mg/l in M9 media grown in $H_2O$) facilitating uniform isotopic enrichment of the peptides with $^{13}$C and $^{15}$N for NMR studies. Indeed, we found that enough peptide for NMR studies could be isolated from only 0.5 l of M9 minimal media, making this expression system attractive as an alternative to other systems requiring more expensive isotope labeled media. Moreover high yields, up to 12 mg/l, are obtained for growth in M9 with 99.9% $D_2O$. Previous workers have reported expression of CRIB fragments of similar length to the eCRIB's reported here (Abdul-Manan et al., 1999; Mott et al., 1999; Morreale et al., 2000; Gizachew et al., 2000). In those studies, GST was used as a fusion carrier, however, expression was significantly lower than reported here, requiring special minimal media (BIOexpress (CIL) or Celtone (Martek)) for enrichment. This is demonstrated by the yield of pure peptide we obtained for eCla4 when expressed as a GST-fusion protein with a thrombin cleavage site ($\sim$ 5 mg/l in LB medium). In addition our attempts to express and purify eSte20 in the same manner were unsuccessful – probably due to the presence of protease sites in the eSte20 sequence. No

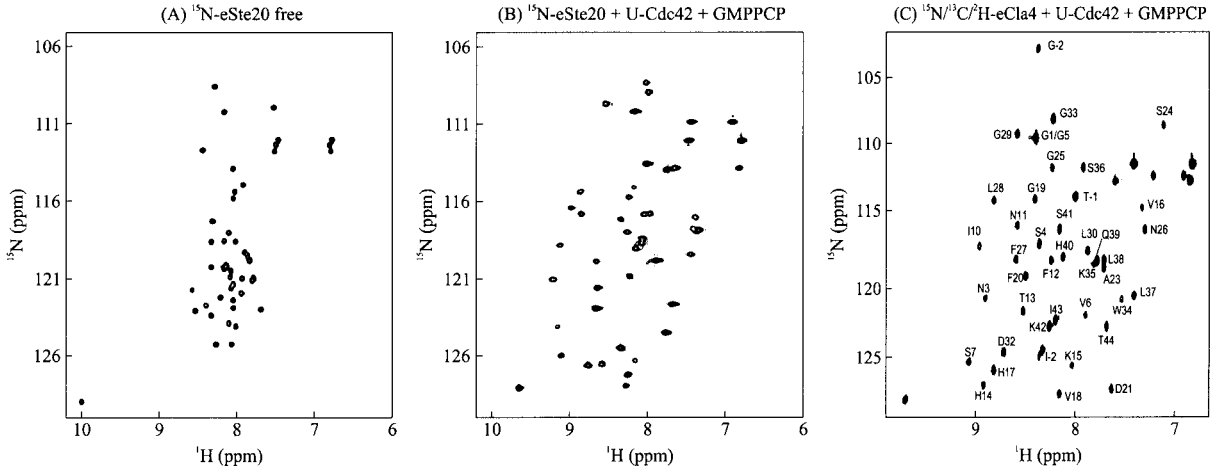


*Figure 6.* $^{1}$H-$^{15}$N HSQC spectra of eCRIB fragments. (A) $^{15}$N-eSte20, (B) $^{15}$N/$^{13}$C-eSte20 in complex with unlabelled Cdc42 activated with GMPPCP, (C) $^{2}$H/$^{15}$N/$^{13}$C-eCla4 in complex with unlabelled Cdc42 activated with GMPPCP. Note the increased spectral width and linebroadening accompanied with the formation of a complex (A and B). The incorporation of $^{2}$H at non-labile sites increases $^{1}$H relaxation rates, giving rise to sharper $^{1}$HN signals and allowing triple resonance techniques to be applied for backbone assignment of eCla4 in complex with Cdc42 (C).

attempts were made to express the m- and c-CRIB peptides in the GST-fusion vector.

*NMR characterization of the CRIB peptides*

The NMR spectra for the mCRIB and cCRIB fragments of Ste20 and Cla4 are shown in Figure 5. The NMR spectra give the expected number of signals for each peptide taking into consideration the first two residues are not visible due to solvent exchange effects. Moreover there are no minor components present, indicating the peptides produced are very pure which can be essential for certain applications. Assignments for the peptides are indicated on the spectra and were obtained in a sequential manner from homonuclear $^{1}$H-$^{1}$H 2D and 3D $^{15}$N-edited TOCSY and NOESY spectra recorded on $^{15}$N labeled samples. Analysis of all the peptides by the chemical shift index indicated that no significant secondary structure was present. Thus, our vector is capable of producing high quantities of unfolded, small peptides without substantial proteolysis.

The $^{1}$H-$^{15}$N HSQC spectra for the eCRIB fragments are shown in Figure 6. The spectrum of free $^{15}$N eSte20 (Figure 6A) exhibits poor resonance dispersion – characteristic of an unfolded polypeptide. Upon addition of unlabelled Cdc42 in excess (Figure 6B) the resonance dispersion of eSte20 dramatically increases, concomitant with increased linewidths (N.B., signals from the unlabelled Cdc42 are not detected in this experiment). The NH linewidths range from 18–20 Hz in the free form to as much as 35 Hz in the complex and are the consequence of the increased rotational tumbling time for the complex. This spectrum is consistent with the formation of a tight, irreversible complex between Cdc42 and eSte20. Enrichment of eSte20 with $^{15}$N/$^{13}$C enabled $^{13}$C/$^{15}$N backbone assignments via HNCA, HN(CO)CA triple resonance experiments. Figure 6C shows the $^{15}$N-$^{1}$H HSQC spectrum of $^{2}$H/$^{15}$N/$^{13}$C eCla4 in complex with Cdc42 (activated with GMPPCP). The effects of deuteration on the $^{1}$HN linewidths due to increased relaxation times and therefore narrow peaks is evident. Incorporation of $^{2}$H for non-labile proteins is crucial to overcoming the size limitations of NMR (generally $<$ 20 kDa for non-deuterated systems) and particularly critical for the study of peptides bound irreversibly to large target proteins by NMR.

*Comparison with other peptide expression systems*

A number of publications have described the expression of peptide for NMR studies (Campbell et al., 1997; Kohno et al., 1998; Jones et al., 2000; Majerle et al., 2000; Sharon et al., 2002; Lindhout et al., 2002; Fairlie et al., 2002). The targeted peptides were produced together with a small and highly expressed carrier protein and the peptide release was achieved via CNBr cleavage (Campbell et al., 1997; Jones et al., 2000; Majerle et al., 2000; Sharon et al., 2002; Lind-

hout et al., 2002; Fairlie et al., 2002) or by a specific protease (Kohno et al., 1998). In general, the expression yields of purified peptides are similar to, or lower than, those reported here (*vide-infra*). In addition, enrichment of peptides with the $^{15}N$ and $^{13}C$ isotopes has been reported (Sprules et al., 2003) for a GST-carried peptide with a Precission Protease (Amersham Biosciences) cleavage site, although the yields are not as high as reported here for purified peptides.

We note that much of the previous work described above have dealt with the expression of one, or at most, two target peptides using a number of different carrier proteins. In contrast, we attempt to evaluate the applicability of one carrier system for the expression of different peptides of varying lengths and sequences. Indeed, the high expression we observe is demonstrated for all six peptides, ranging from 19 to 48 residues irrespective of amino acid compositions.

Of all the expressions systems used to generate peptide fragments, only a small number have been demonstrated to work successfully with a number of different peptides. The commonly used GST-fusion (Amersham Biosciences) and ubiquitin fusion (Baker et al., 1996) systems generally have low yields of expression and are therefore not always suitable for growth in M9 and require expensive rich media. In contrast, we have clearly shown that high levels of expression are obtained for all the peptides studied herein even for growth on minimal media recipes. Thus, Lindhout et al. (2002) achieved yields of 80 mg/l in LB medium of fusion-protein using the 56-residue GB1 domain of protein G, whilst the same expression system was shown to generate high yields (10 mg/l) of pure peptide in the M9 medium (Campbell et al., 1997). Sharon et al. (2002) reported a yield of 6-12 mg/L for a 23 residue gp120 V3 peptide, but a rich media was required for isotopic labeling, and it is unclear if the achieved yield was for isotopically enriched peptides. Kohno et al. (1998) described the use of an ubiquitin fusion system, but the yields for the production of the 14 residue mastoparan-X in minimal (M9) media was very low (0.3–0.6 mg/l). Both Jones et al. (2000) and Majerle et al. (2000) use the strategy of directing the peptides to the inclusion bodies. Jones et al. (2000) reported a yield of 5–10mg/l for two relatively large peptide fragments of 50 residues. Majerle et al. (2000) however, reported substantially lower yields ($\sim$ 6 mg/l in LB) for a peptide of 11 residues than we demonstrate here for six peptides ranging from 19 to 48 residues in length. Fairlie et al. (2002) have recently described the use of a mutated SH2 domain as a carrier protein for the expression of six recombinant peptides of 14 residues in length, but comparisons are difficult to make since only growths on LB were reported and the yields were given only for the fusion-protein (25–75 mg/l), which is comparable to the yields of pure peptides in LB media achieved with our expression system.

## Conclusions

In this paper we describe a strategy for generating large quantities of purified peptides by carrying peptides into cell inclusion bodies via a new and highly expressed fusion protein. We have successfully applied this strategy for the production of six peptides of varying lengths (from 19 to 48 residues) derived from the CRIB motifs of *Candida* Cla4 and Ste20. Thus, isotopic enrichment with $^{13}C/^{15}N/^{2}H$ can be obtained at significantly lower costs than using rich media alternatives and is demonstrated herein. In the present work, the absence of methionine residues in any of the desired peptides facilitated use of CNBr cleavage for the release of the peptides by an engineered Met residue. However, a variety of cleavage sites, e.g. specific for formic acid and by specific proteases such as thrombin can be used as alternatives for peptide release (both have recently been applied in our laboratory – unpublished results). The described peptide production strategy will be particularly useful for NMR studies of peptides and their interactions with target proteins. Our expression system may also be applicable for the production of larger recombinant peptides, such as naturally unfolded proteins, which have recently been shown to have important functional activities (Wright and Dyson, 1999; Uversky, 2002; Dunker et al., 2002). Overall, the efficacy of the new peptide expression system may make it attractive for any applications that require significant quantities of pure peptides, particularly as a cost-effective alternative to solid-state synthesis.

## Acknowledgements

We thank Barbara Requesens for technical assistance. This work was supported by the Genomics and Health Research Initiatives of the National Research Council of Canada, sponsored by the Government of Canada (NRCC Publication No. 46152)

## References


Abdul-Manan, N., Aghazadeh, B., Liu, G.A., Majumdar, A., Ouerfelli, O., Siminovitch, K.A. and Rosen, M.K. (1999) *Nature*, **399**, 379–383.

Baker, R.T. (1996) *Curr. Opin. Biotechnol.*, **7**, 541–546.

Bodenhausen, G. and Ruben, D.J. (1980) *Chem. Phys. Lett.*, **69**, 185–189.

Bolin, D. R., Michalewsky, J., Wasserman, M.A. and O'Donnell, M. (1995) *Biopolymers*, **37**, 57–66.

Campbell, A.P., Bautista, D.L., Tripet, B., Wong, W.Y., Irvin, R.T., Hodges, R.S. and Sykes, B.D. (1997) *Biochemistry*, **36**, 12791–12801.

Disis, M.L., Knutson, K.L., McNeel, D.G., Davis, D. and Schiffman, K. (2001) *Crit. Rev. Immunol.*, **21**, 263–273.

Dunker, A.K., Brown, C.J., Lawson, J.D., Iakoucheva, L.M. and Obradovic, Z. (2002) *Biochemistry*, **41**, 6573–6582.

Fairlie, W.D., Uboldi, A.D., De Souza, D.P., Hemmings, G.J., Nicola, N.A. and Baca, M. (2002) *Prot. Express. Pur.*, **26**, 171–178.

Fox, T., de Miguel, E., Mort, J.S. and Storer, A.C. (1992) *Biochemistry*, **31**, 12571–12576.

Gizachew, D., Guo, W., Chohan, K.K., Sutcliffe, M.J. and Oswald, R.E. (2000) *Biochemistry*, **39**, 3963–3971.

Gong, F. and Yanofsky, C. (2002) *Science*, **297**, 1864–1867.

Grzesiek, S. and Bax, A. (1993) *J. Am. Chem. Soc.*, **115**, 12593–12594.

Jacobsen, J.S. (2002) *Curr. Top. Med. Chem.*, **2**, 343–352.

Jones, D.H., Ball, E.H., Sharpe, S., Barber, K.R. and Grant, C.W.M. (2000) *Biochemistry*, **39**, 1870–1188.

Kohno, T., Kusonoki, H., Sato, K. and Wakamatsu, K. (1998) *J. Biomol. NMR*, **12**, 109–121.

Kuliopulos, A. and Walsh, C.T. (1994) *J. Am. Chem. Soc.*, **116**, 4599–4607.

Lauta, V.M. (2000) *Fundam. Clin. Pharmacol.* , **14**, 425–442.

Leberer, E., Harcus, D., Broadbent, I.D., Clark, K.L., Dignard, D., Ziegelbauer, K., Schmidt, A., Gow, N.A., Brown, A.J. and Thomas, D.Y. (1996) *Proc. Natl. Acad. Sci. U.S.A.*, **93**, 13217–13222.

Leberer, E., Ziegelbauer, K., Schmidt, A., Harcus, D., Dignard, D., Ash, J., Johnson, L. and Thomas, D.Y. (1997) *Curr. Biol.*, **7**, 539–546.

Lindhout, D.A., Li, M.X., Schieve, D. and Sykes, B.D. (2002) *Biochemistry*, **41**, 7267–7274.

Majerle, A., Kidriè, J. and Jerala, R. (2000) *J. Biomol. NMR*, **18**, 145–151.

Marion, D., Ikura, I., Tschudin, R. and Bax, A. (1989) *J. Magn. Reson.*, **85**, 393–399.

Mazitschek, R., Baumhof, P. and Giannis, A. (2002) *Mini. Rev. Med. Chem.*, **2**, 491–506.

Meyer, M., Richter, R. and Forssmann, W.G. (1998) *Eur. J. Med. Res.*, **3**, 103–110.

Morreale, A., Venkatesan, M., Mott, H.R., Owen, D., Nietlispach, D., Lowe, P.N. and Laue, E.D. (2000) *Nat. Struct. Biol.*, **7**, 384–388.

Mott, H. R., Owen, D., Nietlispach, D., Lowe, P.N., Manser, E., Lim, L. and Laue, E.D. (1999) *Nature*, **399**, 384–388.

Muhandiram, D.R. and Kay, L.E. (1994) *J. Magn. Reson. Ser. B*, **103**, 203–216.

Nathisuwan, S. and Talbert, R.L. (2002) *Pharmacotherapy*, **22**, 27–42.

Sharon, M., Görlach, M., Levy, R., Yehezkiel, H. and Anglister, J. (2002) *Prot. Express. Pur.*, **24**, 374–383.

Sood, A. and Panchagnula, R. (2001) *Chem. Rev.*, **101**, 3275–3303.

Sprules T., Green N., Featherstone M. and Gehring K. (2003), *J. Biol. Chem.*, **278**, 1053–1058.

Stevens, W. K., Vranken, W., Goudreau, N., Xiang, H., Xu, P. and Ni, F. (1999) *Biochemistry*, **38**, 5968–5975.

Thompson, G., Owen, D., Chalk, P.A. and Lowe, P.N. (1998) *Biochemistry*, **37**, 7885–7891.

Tyrberg, B. and Levine, F. (2001) *Curr. Opin. Investig. Drugs*, **2**, 1568–1574.

Uversky, V.N. (2002) *Prot. Sci.*, **11**, 739–756.

Wright, P.E. and Dyson, H.J. (1999) *J. Mol. Biol.*, **293**, 321–331.

Zhang, B., Wang, Z. and Zheng, Y. (1997) *J. Biol. Chem.*, **272**, 21999–22007.

Zhao, Z. S., Manser, E., Chen, X.Q., Chong, C., Leung, T. and Lim, L. (1998) *Mol. Cell Biol.*, **18**, 2153–2163.