Generation of a dual-labeled fluorescence biosensor for Crk-II phosphorylation using solid-phase expressed protein ligation

Graham J Cotton and Tom W Muir

Background: The site-specific chemical modification of proteins has proved to be extremely powerful for generating tools for the investigation of biological processes. Although a few elegant methods exist for engineering a recombinant protein at a unique position, these techniques cannot be easily extended to allow several different chemical probes to be specifically introduced into a target sequence. As such multiply labeled proteins could be used to study many biological processes, and in particular biomolecular interactions, we decided to investigate whether such protein reagents could be generated using an extension of the semisynthesis technique known as expressed protein ligation.

Results: A solid-phase expressed protein ligation (SPPL) technology is described that enables large semisynthetic proteins to be assembled on a solid support by the controlled sequential ligation of a series of recombinant and synthetic polypeptide building blocks. This modular approach allows multiple, different chemical modifications to be introduced site-specifically into a target protein. This process, which is analogous to solid-phase peptide synthesis, was used to dual-label the amino and carboxyl termini of the Crk-II adapter protein with the fluorescence resonance energy transfer pair tetramethylrhodamine and fluorescein, respectively. The resulting construct reports (through a fluorescence change) the phosphorylation of Crk-II by the nonreceptor protein tyrosine kinase, c-Abl, and was used to probe the protein–protein interactions that regulate this important post-translational process.

Conclusions: SPPL provides a powerful method for specifically modifying proteins at multiple sites, as was demonstrated by generating a protein-based biosensor for Crk-II phosphorylation. Such protein derivatives are extremely useful for investigating protein function *in vitro* and potentially *in vivo*. This modular approach should be applicable to many different protein systems.

Address: Laboratory of Synthetic Protein Chemistry, The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA.

Correspondence: Tom W Muir
E-mail: muirt@rockvax.rockefeller.edu

Key words: Crk-II phosphorylation, fluorescent biosensor, protein engineering, site-specific labeling

Received: **13 December 1999**
Accepted: **20 January 2000**

Published: **16 March 2000**

Chemistry & Biology 2000, **7**:253–261

1074-5521/00/$ – see front matter

Introduction

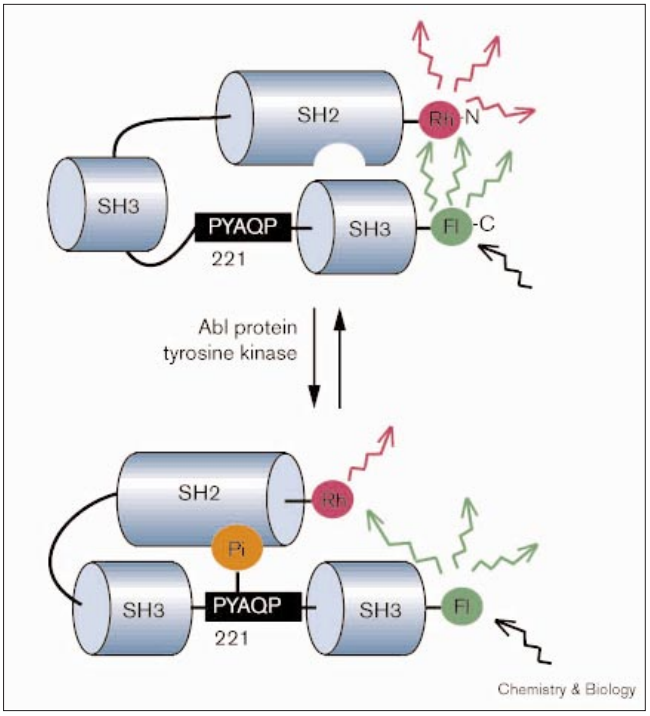
The incorporation of biophysical probes or post-translational modifications at defined positions within a target protein provides an extremely powerful way of investigating the molecular mechanisms that control complex biological processes. There are now several methods available for labeling a recombinant protein at a single defined position; in particular, unnatural amino acid mutagenesis [1] and cysteine modification [2] have been used extensively for this purpose. These approaches, however, do not offer a straightforward way of introducing, in a homogeneous fashion, multiple different modifications at specific sites within a protein. Thus, sophisticated protein-engineering strategies, which require specific combinations of biophysical/biochemical probes to be incorporated into proteins (e.g. fluorescence resonance energy transfer (FRET) pairs, isotopic labels and post-translational modifications), have proven very difficult to perform. In principle, protein total synthesis through the chemical [3] or enzymatic [4] ligation of synthetic peptide fragments provides a route to proteins that possess diverse patterns of chemical modification. Although these peptide ligation approaches have proven extremely powerful for studying small proteins [3], their practical utility decreases with the increasing size of the target protein. This is due to size constraints on the synthetic peptide building blocks [5]. The major aim of this study was, therefore, to develop a generally accessible methodology that would allow several recombinant and synthetic polypeptides to be regioselectively linked together, thereby allowing multiple different chemical probes to be site specifically incorporated into the resulting semisynthetic protein product.

The basis of our semisynthetic approach was to use sequential peptide ligation — an iterative fragment condensation strategy that allows a series of unprotected peptide building blocks to be assembled in a unidirectional, stepwise fashion. The so-called ‘native chemical ligation’ reaction was chosen as the synthetic framework for our approach because it allows two peptide fragments

to be joined together through a normal peptide bond [6], and because it has been successfully applied to the sequential ligation of multiple synthetic peptides both in solution [7,8] and, most recently, on the solid phase [9]. Importantly, recent advances in protein engineering allow the necessary reactive functionalities for native chemical ligation (namely, an amino-terminal cysteine and a carboxy-terminal thioester) to be introduced into recombinant polypeptides [10–13]. Such advances have enabled semisynthetic and fully recombinant protein constructs to be generated through ligation of the appropriate two fragments, in a procedure termed expressed protein ligation (EPL) [14] or intein-mediated protein ligation [15] (for review see [16]).

Recently, we extended EPL to permit the insertion of a synthetic peptide into a recombinant protein, through the sequential ligation, in solution, of two recombinant protein fragments to the amino and carboxyl termini of a synthetic peptide cassette [17]. Although this strategy is, theoretically, extendable to the ligation of any combination of synthetic and/or recombinant fragments, the need to perform all of the steps in solution renders the approach technically demanding; after each ligation reaction it is necessary to isolate the desired product from the reaction mixture, a process that is time-consuming and, importantly, leads to substantial handling losses. In principle, these problems should be overcome by transferring the entire process to the solid phase in a manner analogous to solid-phase peptide synthesis (SPPS) [18]. As with SPPS, this solid-phase expressed protein ligation (SPPL) approach should allow each reaction to be driven to completion by using a large excess of reagents, which can then be simply removed by washing. In addition, there would be no need to isolate intermediate ligation products, which would remain immobilized on the support. Here we describe the development of a SPPL technology and its successful application to the generation of a dual-labeled version of the ~35 kDa adapter protein Crk-II. This semisynthetic protein analog was shown to biosense specifically for a post-translational tyrosine-phosphorylation event important in the regulation of Crk-II-mediated signal transduction.

Figure 1



Biosensor for c-Abl phosphorylation of the Crk-II adapter protein. c-Abl phosphorylates Crk-II on Tyr221, which is thought to induce an intramolecular association with the SH2 domain. This rearrangement is expected to yield a net change in the distance between the termini of the protein. This change would be reported by a dual-labeled derivative of Crk-II in which the FRET pair tetramethylrhodamine (Rh) and fluorescein (Fl) are specifically incorporated at its amino and carboxyl termini, respectively.

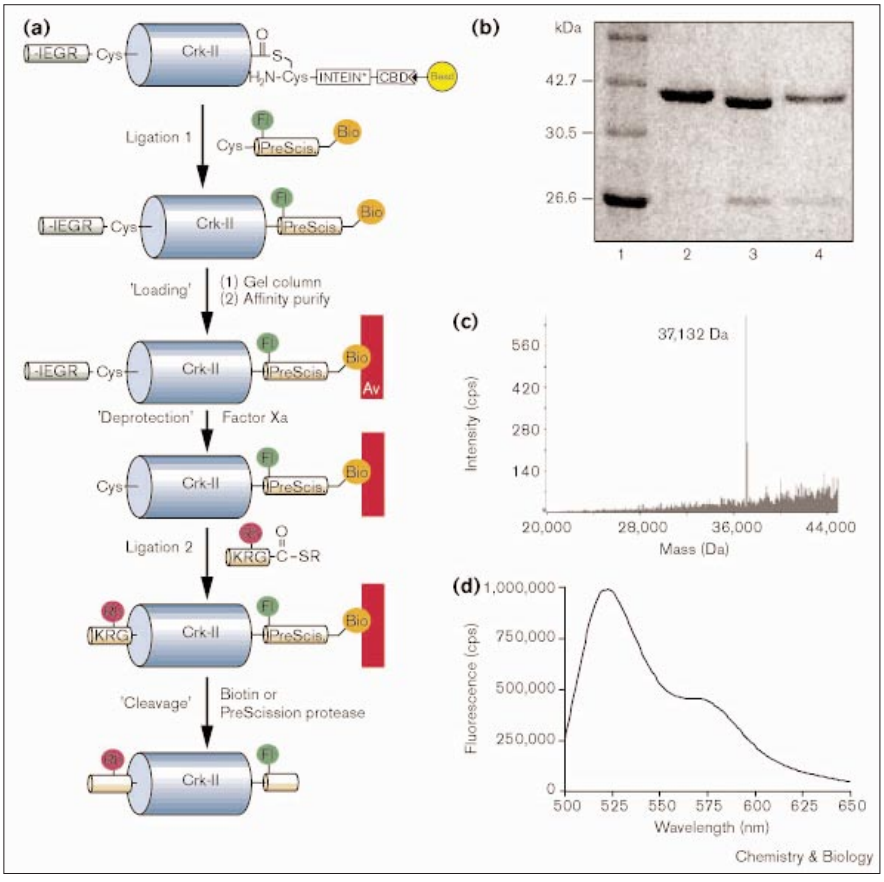
Results

To demonstrate the feasibility of SPPL, we undertook the synthesis of a semisynthetic version of the 304 amino acids adapter protein Crk-II, in which the FRET pair tetra methylrhodamine (Rh) and fluorescein (Fl) were incorporated at the amino and carboxyl termini of the protein, respectively (hereafter referred to as Rh–(Crk-II)–Fl). Crk-II has been implicated in a number of cellular signaling processes and is composed predominantly of one Src homology 2 (SH2) and two SH3 domains through which it mediates intermolecular protein–protein interactions [19,20]. Two protein tyrosine kinases, c-Abl and the epidermal growth factor receptor (EGFR), are known to phosphorylate Crk-II on a unique tyrosine residue (Tyr221) located between the SH3 domains [21,22]. Based on a number of studies, including nuclear magnetic resonance (NMR) spectroscopic analysis, this post-translational modification is thought to regulate Crk-II function by inducing an intramolecular association with the SH2 domain [21,23], which in turn inhibits certain intermolecular protein interactions [19–22]. It was anticipated that phosphorylation and subsequent intramolecular association would result in a change in distance between the termini of Crk-II, which would lead to a change in FRET between the two fluorophores in the dual-labeled analog (Figure 1). Consequently, this protein construct would directly biosense this important post-translational event.

Synthesis of Rh–(Crk-II)–Fl by SPPL

The overall scheme for the synthesis of Rh–(Crk-II)–Fl by SPPL is summarized in Figure 2a. As with SPPS, the strategy can be divided essentially into three parts: attachment of the first building block to a solid support, chain assembly in a carboxyl-to-amino direction involving

Figure 2



Solid-phase expressed protein ligation (SPPL). **(a)** Generation of Rh–(Crk-II)–Fl. Analogous to SPPS, the procedure involves a loading step followed by rounds of deprotection and ligation, and culminates in a cleavage step. Av, monomeric avidin; Bio, biotin; PreScis, sequence cleavable by the protease PreScission. (-IEGR- or Xa- is the pro-sequence removable by the protease factor Xa). **(b)** Coomassie-stained 12% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) gel of: lane 1, molecular weight markers; lane 2, monomeric-avidin beads loaded with the first ligation product Xa–Cys–(Crk-II)–Fl–PS–Biotin; lane 3, the same beads after treatment with factor Xa to yield Cys–(Crk-II)–Fl–PS–Biotin; lane 4, the beads after overnight ligation of the second synthetic peptide to generate Rh–(Crk-II)–Fl. **(c)** ESMS (expected mass = 37,123.6 Da) and **(d)** fluorescence emission spectrum (excitation 490 nm) of purified Rh–(Crk-II)–Fl.

successive deprotection and ligation steps, and cleavage of the completed polypeptide off the solid support.

In the first step, full-length mouse Crk-II was expressed as an in-frame fusion to an engineered yeast vacuolar membrane ATPase (VMA) intein (itself linked to a chitin-binding domain; CBD), which allows the subsequent generation of a reactive $^{\alpha}$thioester derivative of Crk-II. An extra glycine residue was added to the carboxyl terminus of Crk-II to improve the kinetics of the first ligation reaction [8], and the amino-terminal methionine was replaced by the sequence MASSRVDGGRSDLIEGRC to facilitate controlled sequential ligation [17]. (The pro-sequence up to but not including the Cys residue is referred to as 'Xa-' or -IEGR- hereafter.) Soluble expression of this fusion protein [Xa–Cys–(Crk-II)–Intein–CBD] was optimized using standard protocols (no *in vivo* intein cleavage of the full-length fusion could be detected) and the desired material was purified by affinity chromatography using a chitin column.

A synthetic peptide, Cys–Fl–PS–Biotin, containing both a fluorescein probe (Fl) and a biotin affinity handle separated by a linker region that contains the cleavage site for the PreScission protease [LEVLFQGP, (PS)], was chemoselectively ligated to the carboxyl terminus of recombinant Crk-II using EPL. This ligation reaction was found to be more than 95% complete after 48 hours in the presence of a large excess of peptide and the thiol cofactors, ethanethiol and 2-mercaptoethanesulfonic acid (MESNA). Gel filtration was used to separate the unreacted peptide from the desired ligation product, which was then attached to monomeric-avidin beads through its biotin functionality. Preliminary model studies had established that the monomeric-avidin–biotin complex was stable to all of the washing, deprotection and ligation steps used in SPPL, but that the interaction can be disrupted under mild conditions with exogenous biotin. Trace amounts of unreacted Crk-II protein and any remaining bacterial protein contaminants were then removed by vigorously washing the beads with high salt and detergent, both at pH 5.2 and pH 8.0 concentration. This yielded the pure protein, Xa–Cys–(Crk-II)–Fl–PS–Biotin, immobilized on a solid support (Figure 2b, lane 2).

In order to continue the solid-phase synthesis, the Xa pro-sequence must be removed from the immobilized Xa–Cys–(Crk-II)–Fl–PS–Biotin to give an amino-terminal

cysteine residue ready for ligation to the next peptide fragment. (The Xa motif acts as an N^{α} protecting group for the cysteine residue in Crk-II and prevents uncontrolled self-ligation during the first ligation step [17].) Complete enzymatic deprotection was achieved by treating the beads with the protease factor Xa for 3 hours to give Cys–(Crk-II)–Fl–PS–Biotin (Figure 2b, lane 3). A small amount (~10%) of a lower molecular weight protein contaminant was also observed (Figure 2b, lane 3, weak band, ~26 kDa) suggesting that some nonspecific cleavage had occurred during this step. The proteolysis reaction was terminated by simply washing the protease from the column; dithiothreitol (DTT) was included in this wash buffer to simultaneously reduce any disulfide bonds that may have formed during the deprotection step. The beads were then equilibrated into ligation buffer, and the newly exposed amino-terminal cysteine residue reacted with a tetramethylrhodamine containing $^{\alpha}$thioester peptide (Rh–KRG–propionamide$^{\alpha}$thioester) in a second ligation step. A large excess of synthetic peptide was again used in the reaction and MESNA was added as the sole thiol cofactor. This reaction was deemed complete after overnight incubation, as determined by SDS–PAGE analysis of the beads (Figure 2b, lane 4), thus generating the dual-labeled Crk-II derivative, Rh–(Crk-II)–Fl. The beads were then washed thoroughly to remove all unreacted tetramethylrhodamine peptide.

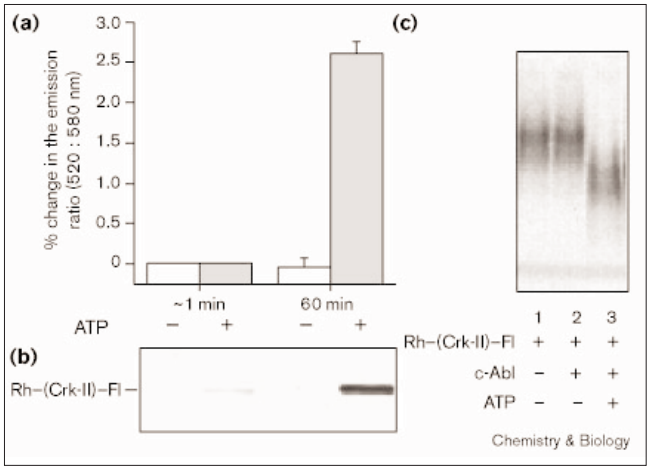
Rh–(Crk-II)–Fl was desorbed from the solid support by washing the beads with a solution containing 2 mM biotin. Approximately 55% of the immobilized material was recovered in a single washing step, although further protein could be eluted by repeating this procedure. The combined washes were passed over a gel-filtration column to remove the free biotin and to remove the protein contaminant arising from nonspecific factor Xa proteolysis. The so-purified dual-labeled Crk-II analog was characterized using electrospray mass spectrometry (Figure 2c) and fluorescence spectroscopy (Figure 2d). The analog was then shown to bind peptide ligands specific to the SH2 and central SH3 domains of Crk-II (data not shown), thus indicating that it had the same gross functional properties as the wild-type protein. We have also used tetrameric avidin as the solid support for SPPL. Because of the high affinity of this interaction, however, the completed protein cannot be competitively eluted from the column as above. In this case the beads were treated with the highly specific PreScission protease. The enzyme cleaved the construct at its recognition site, incorporated between the fluorescein and the biotinyl functionalities, and released Rh–(Crk-II)–Fl from the beads (data not shown).

Phosphorylation studies

Purified Rh–(Crk-II)–Fl was assayed for its ability to biosense Crk-II phosphorylation by the c-Abl protein tyrosine kinase. As indicated previously, phosphorylation by c-Abl leads to an intramolecular association between a phosphotyrosine motif and the Crk-II SH2 domain, which could potentially be reported by the dual-labeled Crk-II derivative (Figure 1). Rh–(Crk-II)–Fl was treated with full length recombinant c-Abl and aliquots of the reaction mixture were analyzed by fluorescence spectroscopy and western blotting at ~1 minute and 60 minute time-points. In the absence of ATP, essentially no change in FRET (i.e. the ratio of the fluorescein/tetramethylrhodamine emission intensities) was observed during the reaction (Figure 3a), and no Rh–(Crk-II)–Fl phosphorylation could be detected using an anti-phosphotyrosine monoclonal antibody (Figure 3b). In contrast, when ATP was included in the reaction mixture, a phosphorylation-dependent increase in the emission intensity ratio (a decrease in FRET) was consistently observed. Rh–(Crk-II)–Fl was completely phosphorylated after 60 minutes as determined using native PAGE mobility (Figure 3c). The fairly modest decrease in FRET (~3% after 60 minutes) suggests that the SH2–phosphotyrosine interaction, which is triggered by Rh–(Crk-II)–Fl phosphorylation, results in only a small net change in the relative distance between the amino and carboxyl termini in the protein.

An interaction between the central SH3 domain of Crk-II (N–SH3) and a proline-rich region in c-Abl (located

Figure 3

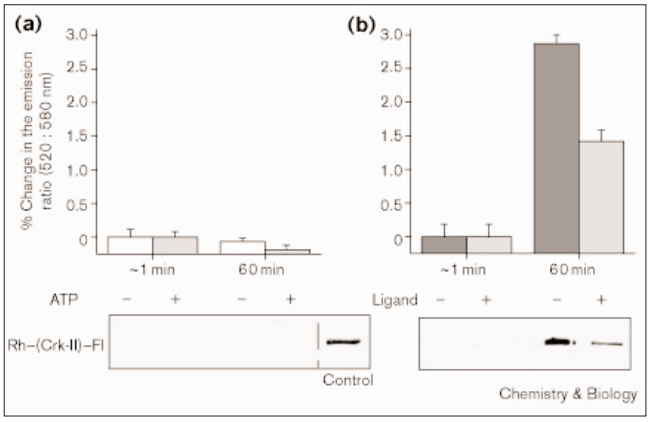


Phosphorylation of Rh–(Crk-II)–Fl by full-length c-Abl. Rh–(Crk-II)–Fl was treated with recombinant c-Abl with or without ATP. **(a)** The percentage change in the Fl:Rh fluorescence emission intensity ratio of Rh–(Crk-II)–Fl at ~1 min and 60 min timepoints. **(b)** Anti-phosphotyrosine western analysis of the corresponding Rh–(Crk-II)–Fl samples shown in (a). **(c)** 6% Native-PAGE gel of untreated Rh–(Crk-II)–Fl (lane 1), Rh–(Crk-II)–Fl after treatment with c-Abl for 60 min in the absence of ATP (lane 2), and Rh–(Crk-II)–Fl after treatment with c-Abl for 60 min in the presence of ATP (lane 3). The gel was imaged for fluorescein fluorescence using a Storm instrument (Molecular Dynamics). All experiments were performed in triplicate.

carboxy-terminal to its kinase domain) has been implicated in formation of the enzyme–substrate complex. Mutations in either this proline-rich region or in the N–SH3 domain, which are predicted to disrupt this intermolecular association, lead to impaired phosphorylation of Crk derivatives [21,22,24]. Similarly, an interaction between the SH2 domain of Crk and the SH3 domain of c-Abl might also contribute to formation of the complex [25]. A truncated version of c-Abl lacking this proline-rich region and the SH3 domain would not be expected to phosphorylate Crk-II with normal kinetics. Indeed, treatment of Rh–(Crk-II)–Fl with a recombinant c-Abl fusion consisting of only the SH2 and kinase domains did not lead to any detectable phosphorylation over 60 minutes, as indicated by fluorescence and western blotting analysis (Figure 4a). Notably, an optimized peptide substrate (EAIYAAPFAKKK [26]) was completely phosphorylated by this truncated version of the kinase after 60 minutes (data not shown).

Taken together, the above studies indicate that Rh–(Crk-II)–Fl is a fluorescence biosensor for c-Abl phosphorylation of Crk-II and confirm that regions of c-Abl outside the SH2 and kinase domains are crucial for this process. One potential use for this biosensor is in the rapid screening of c-Abl kinase inhibitors or compounds that inhibit interactions necessary for phosphorylation. As a simple illustration, the system was used to investigate whether an exogenous ligand for the N–SH3 of Crk-II can modulate Crk-II phosphorylation by inhibiting binding to c-Abl. Treatment of Rh–(Crk-II)–Fl with full length c-Abl in the presence of a saturating amount of a high-affinity N–SH3 ligand [27] resulted in a ~50% reduction in the change in FRET after 60 minutes of reaction, relative to the change in the positive control (Figure 4b). This suggests that the peptide ligand interferes with, but does not completely inhibit, phosphorylation, a conclusion substantiated by western blotting analysis (Figure 4b).

Figure 4



(a) Change in Rh–(Crk-II)–Fl fluorescence after treatment with a truncated version of c-Abl containing only the SH2 and kinase domains. Kinase reactions were performed over 60 min with or without the addition of ATP. Anti-phosphotyrosine western analysis of the corresponding Rh–(Crk-II)–Fl samples is shown below. As a positive control, an equimolar amount of Rh–(Crk-II)–Fl was treated with full length c-Abl and ATP for 60 min. **(b)** Change in Rh–(Crk-II)–Fl fluorescence after treatment with full length c-Abl in the presence of a saturating concentration of a high-affinity peptide ligand (see the Materials and methods section for sequence) for the N–SH3 domain of Crk-II. As above, the anti-phosphotyrosine western analysis of the respective reactions is shown directly below the fluorescence data. In both (a) and (b) the fluorescence values are the mean over three measurements.

Discussion

A solid-phase EPL procedure is described that allows a series of polypeptide fragments to be assembled in a manner analogous to SPPS. Importantly, the functionalities necessary for chemical ligation, amino-terminal protection and attachment to the solid support, are readily incorporated into both recombinant and synthetic polypeptides. A combination of synthetic and recombinant polypeptide building blocks can therefore be used in the protein ligation procedure.

As illustrated in Figure 2a, SPPL was used to prepare a dual-labeled version of Crk-II from three fragments; full-length recombinant Crk-II, and two small synthetic peptides. The well-established native chemical ligation reaction [6] was used to hook the polypeptides together in a stepwise fashion. In each of the two ligation reactions, a large excess of the synthetic component (>10 equivalents) was added to drive the reaction to completion. The first ligation, between the Crk-II–intein fusion and Cys–Fl–PS–Biotin, was performed directly from the chitin affinity beads and was most efficient when both ethanethiol and MESNA were included as thiol cofactors. Ethanethiol has previously been shown to cleave intein-fusions more efficiently than MESNA [28]. It is therefore likely that Crk-II is cleaved off the chitin beads predominantly as an ethyl $^{\alpha}$thioester derivative and that this derivative is then converted through transthioesterification into a more reactive MESNA $^{\alpha}$thioester derivative, *in situ*. The second ligation reaction was performed on the solid phase and therefore the excess peptide was simply removed from the resin-bound product by washing (a gel filtration step was required after the first ligation to prevent unreacted peptide from interfering with the subsequent loading step).

Attachment to the solid phase was achieved through a biotin–monomeric-avidin interaction (note, in many cases it will be possible to directly introduce a biotin group at the carboxyl terminus of the recombinant polypeptide [29]). This association was stable to the reducing conditions of ligation and was not disrupted by the 'factor Xa' deprotection step. It was also stable in a combination of

high salt and detergent at pH 5.2 and pH 8.0, which permitted stringent washing of the column; importantly, this stage allowed removal of trace amounts of bacterial protease contaminants that had been carried through from the Crk-II protein expression stage. Upon completion of the synthesis, the semisynthetic protein was eluted from the support by washing with exogenous biotin. Note that to maximize the recovery of the protein, this competitive elution procedure may have to be repeated several times. Alternatively, a proteolytic cleavage strategy could be employed that takes advantage of the recognition sequence for the PreScission protease, incorporated between the biotin and fluorescein moieties in the carboxy-terminal peptide. This latter strategy allows the use of higher capacity tetrameric-avidin beads, although in some systems it may be less specific than competitive elution with biotin.

Factor Xa induced deprotection of the immobilized intermediate, Xa–Cys–(Crk-II)–Fl–PS–Biotin, proceeded efficiently and was complete after 3 hours. A small amount of nonspecific cleavage was observed, however, leading to an unreactive lower molecular weight fragment. This fragment was easily removed by gel filtration post assembly; conceivably such side products could also be removed using an orthogonal amino-terminal affinity purification strategy. It is also worth noting that the use of alternative proteolytic deprotection strategies, based on enzymes such as enterokinase or ubiquitin hydrolase, might lead to less nonspecific cleavage than factor Xa in certain protein systems.

SPPL has allowed the synthesis of a semisynthetic Crk-II analog in which the FRET pair Rh and Fl were specifically introduced at the amino and carboxyl termini of the protein. The two fluorophores were positioned close to the natural ends of Crk-II (≤ 10 Å) in order to maximize the sensitivity to conformational change in this region. This type of chemical labeling is analogous to the incorporation of different green fluorescent protein (GFP) derivatives at the termini of recombinant proteins, by standard DNA cloning methodologies [30]. In a number of elegant studies, peptide and protein sequences have been expressed as simultaneous amino- and carboxy-terminal fusions to GFP derivatives that act as FRET pairs; the resulting constructs were used to biosense for a number of biological events both *in vitro* and *in vivo* [31,32]. Our semisynthesis approach complements this existing method by allowing a vast array of different FRET pairs, biophysical probes or chemical modifications to be introduced into different regions of a protein (not merely the ends), and thus alleviates potential problems that might arise from the incorporation of two large proteins (GFP = 27 kDa) into the target primary sequence.

Rh–(Crk-II)–Fl was found to biosense for c-Abl phosphorylation of Crk-II. Treatment with the full-length kinase induced a small, but reproducible, decrease in FRET between the two fluorophores, which was dependent upon phosphorylation as indicated by western blotting. This phosphorylation event also inhibited the binding of a known phospho–peptide ligand to the SH2 domain of Rh–(Crk-II)Fl (data not shown), which is consistent with the formation of an intramolecular SH2–phosphotyrosine association within the dual-labeled construct. Although western analysis was crucial to the initial validation of the approach, it should be stressed that FRET provides a direct (i.e. more rapid) and quantitative read-out of Crk-II phosphorylation and hence c-Abl kinase activity. Our results argue that the distance between the termini of Crk-II slightly increases after this post-translational event, implying that there is either a gross re-organization of the termini, which results in only a small net distance change, or that the conformational changes are remote from the termini.

The resonance energy transfer between the fluorophores in the unphosphorylated molecule was calculated to be 52.5%. This was as determined from both the quenching of the fluorescein emission intensity and the sensitized emission of the rhodamine acceptor (as in [33]). Assuming that both fluorophores have random orientations and using a Förster distance of 45 Å for the Fl–Rh pair [33], then the distance between the two fluorophores is ~44 Å. Interestingly, this suggests that unphosphorylated Crk-II has a somewhat compacted domain architecture, as opposed to a linear array of domains; on the basis of the primary sequence, the amino and carboxyl termini could be as much as ~200 Å apart if the interdomain linkers assume a fully extended conformation.

A truncated version of c-Abl, lacking both the proline-rich carboxy-terminal region and the SH3 domain, does not induce a FRET change in Rh–(Crk-II)–Fl, which, as expected, is due to a complete lack of phosphorylation of this protein during the time-frame of the experiment. This both substantiates the ability of the Crk-II analog to biosense phosphorylation specifically and confirms that regions outside these domains are crucial for this process. It also indicates how such a biosensor maybe used for assaying the kinase activity of c-Abl or exploring the molecular mechanisms of Crk-II phosphorylation.

The deregulation of protein tyrosine kinases, such as c-Abl, has been implicated in the development of many disease states, making these proteins important targets in the field of drug discovery [34]. Current approaches for screening small-molecule inhibitors rely mostly on the use of ^{32}P phosphotransfer assays, which are expensive and create obvious safety issues. Nonradioactive assays that enable compounds to be rapidly screened are therefore of significant value. In principle, our fluorescence-based strategy can be used for this purpose. As a simple demonstration,

our system was used to assay rapidly the effect of a high-affinity ligand for the N–SH3 of Crk on phosphorylation by c-Abl. This compound was found to partially inhibit Crk-II phosphorylation, presumably by blocking crucial interactions with the proline-rich region of c-Abl.

In summary, a solid-phase EPL technology has been developed that eliminates the handling and purification problems associated with solution-based sequential ligation procedures. This, coupled with the high efficiency of the deprotection and ligation steps, should allow several protein fragments (greater than the three component synthesis described here) to be selectively assembled by simply performing multiple rounds of ligation and deprotection. This will allow two or more different chemical modifications to be incorporated specifically into a target protein.

Significance

Proteins specifically labeled with biophysical and biochemical probes are extremely valuable tools for investigating the molecular mechanisms that regulate biological processes. To date, however, the incorporation of more than one chemical modification into a target protein sequence has proved to be difficult, which has thus hindered the development of strategies that require a protein to be specifically modified with combinations of different chemical probes. We describe here a technology, solid-phase expressed protein ligation (SPPL), that allows multiple different modifications to be specifically incorporated into large proteins. The approach was used to dual label the amino and carboxyl termini of the Crk-II adapter protein with the fluorescence resonance energy transfer (FRET) pair tetramethylrhodamine and fluorescein, respectively. The resulting construct is a fluorescent biosensor for the phosphorylation of Crk-II by the Abl protein tyrosine kinase — an event that is thought to regulate the function of Crk-II *in vivo*. The fluorescent Crk-II derivative was used to show that a ligand for the N–SH3 domain of Crk-II interferes with phosphorylation by Abl. This demonstrates the potential of this dual-labeled construct for investigating the molecular determinants that control this important post-translational process. It also indicates how this construct could be used for *in vitro* screening for Abl kinase inhibitors. It is envisaged that SPPL will be a generally applicable approach for specifically labeling proteins with multiple combinations of different biophysical and biochemical probes. This technique will enable highly sophisticated protein tools to be generated for studying a wide variety of biological processes.

Materials and methods

Protein expression; Xa–Cys–(Crk-II)–Intein–CBD

The polymerase chain reaction (5′ primer AAAAGAAAAAAAGGCGGCCGCTCGGATCTGATCGAAGGTCGTTGTGCGGGCAACTTCGACTCGG and 3′ primer GCAAACTGGCTCTTCCGCAGCCGCTGAAGTCCTCATCGGG) was used to amplify the region corresponding to full length mouse Crk-II (residues Ala2 to Ser304) from a pcDNA-mCrk vector template (gift from H. Hanafusa, Rockefeller University). After digestion with *Sap1* and *Not1*, the desired fragment was purified using gel electrophoresis and subcloned into a *Sap1–Not1* treated pTYB3 plasmid (New England Biolabs). This $pTYB3_{Xa–Cys–Crk\text{-}II}$ vector encodes a fusion protein comprising full length mouse Crk-II linked via a glycine residue to the amino terminus of the yeast VMA intein–CBD region and containing the sequence MASSRVDGGRSDLIEGRC immediately amino terminal to Ala2 of Crk-II (confirmed by DNA sequencing); the pro-sequence up to, but not including, the Cys residue is referred to as 'Xa-' or -IEGR-. *E coli* BL21 cells were transformed with this plasmid and grown in LB medium (6 l) to mid-log phase. Protein expression was then induced for 4 h at 30°C using 0.2 mM IPTG. After centrifugation the cells were resuspended in lysis buffer (0.1 mM EDTA, 250 mM NaCl, 5% glycerol, 1 mM PMSF, 25 mM HEPES, pH 7.4) and lysed using a French press. The soluble fraction was then loaded onto a chitin column (~12 ml), pre-equilibrated in wash buffer (1 mM EDTA, 250 mM NaCl, 0.1% Triton X-100, 25 mM HEPES, pH 7.0), and the column was then washed with same buffer. Typically, this procedure gave a loading of ~2 mg fusion protein per ml of chitin beads.

Peptide synthesis

Peptides were manually synthesized according to the *in situ* neutralization/HBTU activation protocol for Boc-SPPS [5]. The peptide N^{α}–tetramethylrhodamine–KRG–propionamide$^{\alpha}$thioester was assembled on a S-propionamide derivatized 4-methylbenzhydrylamine (MBHA) resin [7], whereas CGK[Dapa(Fl)]–GLEVLFQGPVRKG–[K^{ε}-(Biotin)]–G–NH_2 (Cys–Fl–PS–Biotin) was synthesized using an MBHA resin. Orthogonal sidechain NH_2 protection using the Fmoc group allowed direct solid-phase attachment of the fluorescein (Fl) and biotin groups, whilst tetramethylrhodamine (Rh) was incorporated at the amino terminus of KRG–propionamide$^{\alpha}$thioester after N^{α} deprotection of the final residue; each of the functionalities was activated as the corresponding N-hydroxysuccinimide ester. All peptides were purified by preparative HPLC and their composition characterized by ESMS.

Solid-phase protein ligation

[Rh–(Crk-II)–Fl] was prepared as follows. Note that all of the steps were performed in the dark and at 4°C unless otherwise stated.

Step1 – Loading. Purified Cys–Fl–PS–Biotin peptide (1 mM) was dissolved in ligation buffer (0.1% Triton-X 100, 200 mM NaCl, 200 mM phosphate pH 7.3) containing both MESNA (4% w/v) and ethanethiol (3% v/v) and then added to the pre-equilibrated chitin beads containing immobilized Xa–Cys–(Crk-II)–Intein–CBD (5 ml), to give a 50% slurry. The mixture was rocked for 48 h at room temperature after which time >95% of the protein (as determined using SDS–PAGE) had reacted to form the desired ligation product [Xa–Cys–(Crk-II)–Fl–PS–Biotin]: ESMS; observed mass = 38,010 ± 19 Da, expected (average isotope composition) 38,027 Da. DTT was then added to the ligation mix to give a 10 mM final concentration and the excess unreacted peptide was removed by gel filtration (HR-75 column; running buffer, 0.1% Triton X-100, 2 mM DTT, 140 mM NaCl, 50 mM Tris, pH 7.4). A portion of the isolated ligation product (typically 1–2 mg) was then incubated for 1 h at 4°C with 4 ml of monomeric-avidin beads (Pierce) which had been pre-equilibrated in gel column buffer. Unbound contaminants were then removed by washing the beads with wash buffer A (0.2% Triton X-100, 2.5 mM DTT, 400 mM NaCl, 100 mM sodium acetate buffer, pH 5.2) followed by wash buffer B (0.2% Triton X-100, 2.5 mM DTT, 400 mM NaCl, 50 mM Tris, pH 8.0), 20 column volumes each. This gave a final loading of ~0.4 mg of [Xa–Cys–(Crk-II)–Fl–PS–Biotin] per ml of monomeric-avidin beads.

Step 2 – Deprotection. The monomeric avidin beads were equilibrated into deprotection buffer (140 mM NaCl, 5 mM phosphate, pH 7.3) and then treated with factor Xa (Amersham Pharmacia, 10u/ml of beads) for 3 h at room temperature. This facilitated complete removal of the cysteine protecting pro-sequence (Xa) as determine dusing SDS–PAGE, to generate the desired material containing a free

amino-terminal cysteine [Cys–(Crk-II)–Fl–PS–Biotin]: ESMS; observed mass = 36,370 ± 18 Da, expected (av. isotope comp.) 36,369 Da. The beads were then washed thoroughly with wash buffer C (5 mM DTT, 140 mM NaCl, 5 mM phosphate, pH 7.2) to remove the protease.

Step 3 – Ligation. The beads were equilibrated into ligation buffer and a solution of purified Rh–KRG–propionamide$^{\alpha}$thioester peptide with MESNA in ligation buffer added to give a 50% slurry of beads containing 2% w/v MESNA and ~2.5 mM synthetic peptide. After rocking the mixture overnight, all of the protein had reacted (as determined using SDS–PAGE) forming the desired ligation product [Rh–(Crk-II)–Fl]. Unreacted peptide was removed by washing with ligation buffer and recycled; the beads were then further washed with ligation buffer supplemented with 2 mM DTT.

Step 4 – Cleavage. The beads were washed with cleavage buffer (1 mM DTT, 0.1% Triton X-100, 1 mM EDTA, 140 mM NaCl, 50 mM Tris, pH 7.0, 20 column volumes) and the protein was liberated from the monomeric avidin support by either: (i) competitive desorption or (ii) proteolysis. (i) To compete the protein off of the monomeric-avidin support, the beads were incubated with eight column volumes of 2 mM biotin in cleavage buffer (~55% of Rh–(Crk-II)–Fl was eluted in this step, though further material could be obtained by repeating the process). The supernatant was then passed over a gel filtration column as in Step 1 to obtain the desired pure material: ESMS; observed mass = 37,132 ± 18 Da, expected (av. isotope comp.) 37,124 Da. (ii) For proteolytic cleavage, the beads were treated overnight with 1 column volume of cleavage buffer containing the enzyme PreScission (Amersham Pharmacia, 2.5 u/ml of beads). The supernatant was then passed over a glutathione-agarose column to remove the protease and yield the desired material: ESMS; observed mass = 36,125 ± 18 Da, expected (av. isotope comp.) 36,118 Da.

Kinase assays

Purified Rh–(Crk-II)–Fl was treated with either full length recombinant (Baculovirus/SF9) mouse c-Abl (gift from T. Schindler and J. Kuriyan, Rockefeller University) or a GST fusion of mouse c-Abl containing only the SH2 and kinase domains (expressed in *E. coli* BL21 essentially as described [35]). In a typical experiment, the appropriate c-Abl construct was incubated in reaction buffer (2 mM DTT, 0.2 mg/ml BSA, 10 mM Mg^{2+}, 50 mM Tris, pH 7.4, either with or without ATP (500 μM)) for 5 min at 30°C before addition of Rh–(Crk-II)–Fl (final concentration = 0.25 μM). In order to ensure that equal amounts of active c-Abl enzyme were added to each reaction, preliminary titration experiments were carried out using an optimized peptide substrate for c-Abl (EAIYAAPFAKKK [26]). For peptide inhibition studies, Rh–(Crk-II)–Fl was pre-incubated for 30 min with a high affinity ligand for the N–SH3 domain of Crk, PPPALPPKRRR-NH_2 [27], such that the final concentration of ligand in the kinase assay was 12 μM. In all cases, aliquots of the reaction mixtures were removed at ~1 min and 60 min, quenched with EDTA (final conc. = 40 mM), and then analyzed by native PAGE and/or Western blotting, and fluorescence spectroscopy.

Western blotting

Standard procedures were used to probe for tyrosine phosphorylation using a mouse monoclonal anti-phosphotyrosine primary antibody (PY20, Santa Cruz Biotechnology) and a HPO-conjugated goat anti-mouse polyclonal secondary antibody (Amersham Pharmacia).

Fluorescence spectroscopy

Experiments were conducted at 18°C in a stirred 0.5 cm-pathlength cell using a SPEX FL3-11C fluorimeter. Samples from the reactions (50 μl) were diluted into 2 mM DTT, 0.4 mg/ml BSA, 140 mM NaCl, 50 mM Tris, pH 7.4 buffer (450 μl) for analysis. Excitation was at 490 nm with a 2.5 nm slit and the fluorescence emission was monitored at 520 nm and 580 nm through a 4 nm slit.

Acknowledgements

We thank K. Shokat (UCSF) for providing a GST–Abl–SH (21) expression plasmid and T. Schindler and J. Kuriyan (Rockefeller University) for recombinant c-Abl protein. The research was supported by a Pew Scholarship in the Biomedical Sciences (T.W.M.), the National Institutes of Health (GM55843-01, T.W.M.) and the Burroughs-Wellcome Fund (G.J.C.).

References

1. Cornish, V.W., Mendel, D. & Schultz, P.G. (1995). Probing protein structure and function with an expanded genetic code. *Angew. Chem. Int. Ed. Engl.* **34**, 621-633.
2. Chen, Y., Ebright, Y.W. & Ebright, R.H. (1994). Identification of the target of a transcription activator protein by protein–protein photocrosslinking. *Science* **265**, 90-92.
3. Wilken, J. & Kent, S.B.H. (1998). Chemical protein synthesis. *Curr. Opin. Biotechnol.* **9**, 412-426.
4. Jackson, D.Y., Burnier, J., Quan, C., Stanley, M., Tom, J. & Wells, J.A. (1994). A designed peptide ligase for total synthesis of ribonuclease A with unnatural catalytic residues. *Science* **266**, 243-247.
5. Schnölzer, M., Alewood, P., Jones, A., Alewood, D. & Kent, S.B.H. (1992). *In situ* neutralization in Boc-chemistry solid phase peptide synthesis: rapid, high yield assembly of difficult sequences. *Int. J. Pept. Protein Res.* **40**, 180-193.
6. Dawson, P.E., Muir, T.W., Clark-Lewis, I. & Kent, S.B.H. (1994). Synthesis of proteins by native chemical ligation. *Science* **266**, 776-779.
7. Camarero, J.A., Cotton, G.J., Adeva, A. & Muir, T.W. (1998). Chemical ligation of unprotected peptides directly from a solid support. *J. Pept. Res.* **51**, 303-316.
8. Hackeng, T.M., Griffin, J.H. & Dawson, P.E. (1999). Protein synthesis by native chemical ligation: expanded scope by using straightforward methodology. *Proc. Natl. Acad. Sci. USA* **96**, 10068-10073.
9. Canne, L.E., Botti, P., Simon, R.J., Chen, Y., Dennis, E.A. & Kent, S.B.H. (1999). Chemical protein synthesis by solid phase ligation of unprotected peptide segments. *J. Am. Chem. Soc.* **121**, 8720-8727.
10. Erlanson, D.A., Chytil, M. & Verdine, G.L. (1996). The leucine zipper domain controls the orientation of AP-1 in the NFAT–AP-1–DNA complex. *Chem. Biol.* **3**, 981-991.
11. Chong, S., *et al*., & Xu, M.Q. (1997). Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. *Gene* **192**, 271-281.
12. Mathys, S., *et al*., & Xu, M.Q. (1999). Characterization of a self-splicing mini-intein and its conversion into autocatalytic N- and C-terminal cleavage elements: facile production of protein building blocks for protein ligation. *Gene* **231**, 1-13.
13. Southworth, M.W., Amaya, K., Evans, T.C., Xu, M.-Q. & Perler, F.B. (1999). Purification of proteins fused to either the amino or carboxy terminus of the *Mycobacterium xenopi* Gyrase A intein. *BioTechniques* **27**, 110-120.
14. Muir, T.W., Sondhi, D. & Cole, P.A. (1998). Expressed protein ligation: a general method for protein engineering. *Proc. Natl Acad. Sci. USA* **95**, 6705-6710.
15. Evans Jr., T.C., Benner, J. & Xu, M.Q. (1998). Semisynthesis of cytotoxic proteins using a modified protein splicing element. *Protein Sci* **7**, 2256-2264.
16. Cotton, G.J. & Muir, T.W. (1999). Peptide ligation and its application to protein engineering. *Chem. Biol.* **6**, R247-R256.
17. Cotton, G., Ayers, B., Xu, R. & Muir, T.W. (1999). Insertion of a synthetic peptide into a recombinant protein framework: a protein biosensor. *J. Am. Chem. Soc.* **121**, 1100-1101.
18. Merrifield, R.B. (1963). Solid Phase Peptide Synthesis. I. The synthesis of a tetrapeptide. *J. Am. Chem. Soc.* **85**, 2149-2154.
19. Birge, R.B., Knudsen, B.S., Besser, D. & Hanafusa, H. (1996). SH2 and SH3-containing adapter proteins: redundant or independent mediators of intracellular signal transduction. *Genes Cells* **1**, 595-613.
20. Buday, L. (1999). Membrane-targeting of signalling molecules by SH2/SH3 domain-containing adaptor proteins. *Biochim. Biophys. Act* **1422**, 187-204.
21. Feller, S.M., Knudsen, B. & Hanafusa, H. (1994). c-Abl kinase regulates the protein binding activity of c-Crk. *EMBO J.* **13**, 2341-2351.
22. Hashimoto, Y., *et al*., & Matsuda, M. (1998). Phosphorylation of CrkII adaptor protein at tyrosine 221 by epidermal growth factor receptor. *J. Biol. Chem.* **273**, 17186-17191.
23. Rosen, M.K., Yamazaki, T., Gish, G.D., Kay, C.M., Pawson, T. & Kay, L.E. (1995). Direct demonstration of an intramolecular SH2-phosphotyrosine interaction in the Crk protein. *Nature* **374**, 477-479.

24. Ren, R., Ye, Z.-S. & Baltimore, D. (1994). Abl protein-tyrosine kinase selects the Crk adapter as a substrate using SH3-binding sites. *Genes Dev.* **8**, 783-795.
25. Anafi, M., Rosen, M.K., Gish, G.D., Kay, L.E. & Pawson, T. (1996). A potential SH3 domain-binding site in the Crk SH2 domain. *J. Biol. Chem.* **35**, 21365-21374.
26. Songyang, Z., *et al*., & Cantley, L.C. (1995). Catalytic specificity of protein-tyrosine kinases is critical for selective signalling. *Nature* **373**, 536-539.
27. Posern, G., *et al*., & Feller, S.M. (1998). Development of highly selective SH3 binding peptides for Crk and CRKL which disrupt Crk-complexes with DOCK180, SoS and C3G. *Oncogene* **16**, 1903-1912.
28. Ayers, B., Blaschke, U.K., Camarero, J.A., Cotton, G.J., Holford, M. & Muir, T.W. (1999). Introduction of unnatural amino acids into proteins using expressed protein ligation. *Biopolymers* **51**, 343-354.
29. Smith, P.A., Tripp, B.C., DiBlasio-Smith, E.A., Lu, Z., LaVallie, E.R. & McCoy, J.M. (1998). A plasmid expression system for quantitative in vivo biotinylation of thioredoxin fusion proteins in *Escherichia coli*. *Nucleic Acids Res.* **26**, 1414-1420.
30. Tsien, R.Y. (1998). The green fluorescent protein. *Annu. Rev. Biochem.* **67**, 509-544.
31. Miyawaki, A., *et al*., & Tsien, R.Y. (1997). Fluorescent indicators for Ca^{2+} based on green fluorescent proteins and calmodulin. *Nature* **388**, 882-886.
32. Romoser, V.A., Hinkle, P.M. & Persechini, A. (1997). Detection in living cells of Ca^{2+}-dependent changes in the fluorescence emission of an indicator composed of two green fluorescent protein variants linked by a calmodulin-binding sequence. *J. Biol. Chem.* **272**, 13270-13274.
33. Selvin, P.R. (1995). Fluorescence resonance energy transfer. *Methods Enzymol.* **246**, 300-334.
34. Cohen, P. (1999). The development and therapeutic potential of protein kinase inhibitors. *Curr. Opin. Chem. Biol.* **3**, 459-465.
35. Garcia, P., Shoelson, S.E., George, S.T., Hinds, D.A., Goldberg, A.R. & Miller, W.T. (1993). Phosphorylation of synthetic peptides containing Tyr-Met-X-Met motifs by nonreceptor tyrosine kinases *in vitro*. *J. Biol. Chem.* **33**, 25146-25151.