Available online at www.sciencedirect.com
SCIENCE DIRECT®
Protein Expression and Purification 39 (2005) 71–81
Protein Expression & Purification
www.elsevier.com/locate/yprep

# A genetically encodable microtag for chemo-enzymatic derivatization and purification of recombinant proteins

Beatrice-Elena Dursina[a], Reinhard Reents[a,1], Anca Niculae[a], Alexei Veligodsky[a,2], Reinhard Breitling[b], Konstantin Pyatkov[c], Herbert Waldmann[a], Roger S. Goody[a], Kirill Alexandrov[a,*]

[a] *Max-Planck-Institute for Molecular Physiology, Otto-Hahn-Str. 11, 44227 Dortmund, Germany*
[b] *JenaBioscience GmbH, Löbstedter Str. 78, 07749 Jena, Germany*
[c] *Institute of Biochemistry and Physiology of Microorganisms, Prospect Nauki 5, Pushchino 142292, Russia*

Received 14 July 2004, and in revised form 21 September 2004
Available online 5 November 2004

**Abstract**

Efficient separation of recombinant polypeptides from proteins of the expression host and their subsequent derivatisation with functional chemical groups is essential for the success of many biological applications. Numerous tag systems have been developed to facilitate the purification procedure but only limited progress has been made in development of generic methods for targeted modification of proteins with functional groups. In this work, we present a novel 6 amino acid long C-terminal protein tag that can be selectively modified with functionalized derivatives of farnesyl isoprenoids by protein farnesyltransferase. The reaction could be performed in complex protein mixtures without detectable unspecific labeling. We demonstrate that this modification can be used to purify the target protein by over 800-fold in a single purification step using phase partitioning. Moreover, we show that the fluorescent group could be used to monitor the interaction of the derivatized proteins with other polypeptides.


*Keywords:* Protein prenylation; Phase partitioning; Protein purification tags

Production of recombinant proteins followed by their subsequent purification and/or derivatization comprise one of the most common and most critical procedures in life sciences and biotechnology. Not surprisingly, a broad spectrum of methods was developed to separate recombinant proteins from the proteins of the host. However, the sequencing of several genomes and ensuing advent of functional genomics and proteomics has dramatically increased the need for efficient protein purification and derivatization methods [1]. This is due to the necessity to express thousands of open reading frames and to isolate and derivatize the resulting polypeptides rapidly and efficiently. The latter may involve labeling of the protein with isotopes, fluorophores, spin labels or other groups for subsequent analysis. The huge number of proteins to be studied poses two strict requirements on the experimental procedures for post-genomic applications: namely their amenability to miniaturization and automation. This in turn requires a high degree of standardization and efficiency of the methods used. Protein purification is the first step that needs to be further optimized since the levels of expression vary quite

* Corresponding author. Fax: +49 231 1331651.
*E-mail address:* kirill.alexandrov@mpi-dortmund.mpg.de (K. Alexandrov).
[1] Present address: Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, MA 02138, United States.
[2] Present address: Institute of Biochemistry, ETH-Zürich, Schafmattstrasse 18, CH-8903, Zürich, Switzerland.

1046-5928/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.pep.2004.09.015

significantly in all existing systems and the need to separate the protein of interest from thousands of other polypeptides is nontrivial.

Genetically encoded protein tags dramatically facilitate protein purification and are currently broadly used in a spectrum of applications [2]. The specificity of the genetically encoded tag for the affinity matrix is typically derived from a highly selective biological interaction such as that of an enzyme:substrate or antibody:antigen interaction. Therefore, with the exception of the poly-histidine and poly-arginine tags, all other tag systems require the presence of a specificity-conferring protein domain either linked to the matrix or fused to a target protein. Highly specific tags that allow one-step purification of proteins are derived from functional protein domains such as GST, maltose-binding protein, chitin and cellulose-binding domains significantly increase the size of the target protein and compromise its functionality. In microtags such as FLAG, Strep or calmodulin-binding peptide the high selectivity is defined by the respective enzyme or antibody bound to the solid support [2].

Traditional reliance of the tag based purification methods on the matrix-bound active substances introduces a number of additional variables into the purification schemes due to the varying density of the active substances on the matrix, limited physical stability, and chemical compatibility with reducing and chaotropic agents.

A number of efforts have been undertaken to alleviate the problem of matrix-based applications by introducing fluid phase approaches. The most prominent development in this respect relies on the hydrophobic phase partitioning of aqueous protein solutions upon reaching the cloud point. Such partition can be mediated by a number of polymer substances such as PEG, Triton X-114, or other polymers [3,4]. Following the partitioning the polymer phase still contains over 75% of water and is suitable for keeping the biological molecules in active form. This approach is potentially very promising due to the cheapness of the media used, scalability, and availability of advanced automatic techniques of liquid handing. However, use of protein hydrophobicity as the primary separation parameter makes the general selectivity of the method rather low while the development of selective protocols must be performed on a case to case basis and remain largely empirical. An incremental improvement was achieved by genetically engineering a hydrophobic tag onto the proteins, which improves their partitioning into the polymer phase [5]. Yet such tags were shown to influence the physical properties and activity of the proteins without providing a dramatic improvement to the method's selectivity and therefore are unlikely to find broad application in practice. In an alternative approach, an affinity dye was used to selectively facilitate partitioning of a target protein into one of the phases [6,7].

In this work, we present a chemo-enzymatic approach for controlling the hydrophobicity and phase partitioning via covalent attachment of a functional group to a genetically encoded microtag. We propose that the approach has general applicability and can be used for derivatization and purification of both cytosolic and secreted proteins.

## Materials and methods

### *Synthesis of fluorescent phosphoisoprenoids*

Tris-ammonium{3,7-dimethyl-8-(7-nitro-benzo[1,2,5]oxadiazol-4-ylamino)-octa-2,6-diene-1}pyrophosphate (NBD-GPP) and tris-ammonium{3,7,11-trimethyl-12-(7-nitro-benzo[1,2,5]oxadiazo-4-ylamino)-dodeca-2,6,10-trien-1}pyrophosphate (NBD-FPP) synthesized as described elsewhere (Dursina et al., unpublished).

### *Construction of expression vectors*

*Escherichia coli* expression vectors for 6his-GST-FTase, 6his GST-GGTase-I, 6his-GST-RhoA, and 6his-GST-Ki-Ras fusion proteins will be described elsewhere (Dursina et al., unpublished). CFP-TKCVIM and dsRed-TKCVIM were constructed by PCR amplification of the respective open reading frames with primers containing *Nde*I in forward primer and nucleotides corresponding to the CAAX tag and *Xho*I restriction sites in the reverse primer. The resulting PCR-products were cloned into pET19mod vector using *Nde*I and *Xho*I restriction sites [8]. The constructs for MRS6 fusion with the C-terminus of RhoA, Rab7 fusion with C-terminus of RhoA, and for YPT7 fusion with the TEV protease cleavage site and C-terminus of Ki-Ras were prepared by the same procedure but in the latter case the PCR product was cloned into the pET30a vector. For construction of Rab7-TKCVIM fusion tagged with an intein domain, we PCR amplified the coding sequence of Rab7 with the forward primer containing the *Nde*I restriction site and with the reverse primer containing nucleotides corresponding to the CAAX tag and a *Sap*I restriction site. The resulting PCR product was digested with *Nde*I and *Sap*I restriction enzymes and cloned into the pTWIN-2 expression vector in-frame with the intein-CBD assembly.

### *Expression and purification of recombinant proteins*

#### *Expression of prenyltransferases*

Expression and purification of FTase and GGTase-I were performed essentially as described for RabGGTase [9]. In some cases the 6his-GST tag was retained since it did not appear to influence the activity of prenyltransferases.