Available online at www.sciencedirect.com
ScienceDirect
Protein Expression and Purification 53 (2007) 80–86
Protein Expression & Purification
www.elsevier.com/locate/yprep

# The expression and purification of the N-terminal activation domain of the transcription factor c-Myc: A model substrate for exploring ERK2 docking interactions

William F. Waas [a], Kevin N. Dalby [a,b,c,d,*]

[a] Division of Medicinal Chemistry, College of Pharmacy, University of Texas at Austin, TX 78712, USA
[b] Graduate Program in Biochemistry, University of Texas, Austin, TX 78712-0252, USA
[c] Graduate Program in Molecular Biology, University of Texas, Austin, TX 78712-0252, USA
[d] Center for Molecular and Cellular Toxicology, University of Texas, Austin, TX 78712-0252, USA

Received 10 October 2006, and in revised form 5 December 2006
Available online 12 December 2006

**Abstract**

ERK2 is a mitogen-activated protein kinase (MAPK) that plays pivotal roles in cell signal transduction, where it mediates effects on proliferation and differentiation by growth factors and hormones. An important substrate of ERK2 is the transcription factor c-Myc, which mediates cell cycle progression. The phosphorylation of Ser-62 on c-Myc by ERK2 is thought to contribute to the increased stability of c-Myc during the cell cycle and is thus a critical cellular event. However, the mode of c-Myc recognition by ERK2 is not understood. Early studies by Gupta and Davis concluded that ERK2 specificity determinants are located in residues 1–100 of c-Myc, its activation domain. To pursue both structural and kinetic studies a rapid, but efficient purification method, for the production of the activation domain of c-Myc from an *Escherichia coli* source, was developed. We chose the minimal number of high-resolution steps to maximize both yield and efficiency without sacrificing purity. Thus, GST-(c-MycΔ2-99)-$His_6$ was expressed in *E. coli*, and purified using glutathione–agarose affinity chromatography. Cleavage of the GST fusion protein by thrombin and subsequent purification by nickel–agarose affinity chromatography yielded 8 mg of purified (c-MycΔ2-99)-$His_6$ from one liter of LB culture. Rigorous characterization demonstrated that under standard assay conditions (c-MycΔ2-99)-$His_6$ is phosphorylated by ERK2 with the following Michaelis parameters: $k_{cat} = 10.4\ s^{-1}$, $K_M^{c\text{-}Myc} = 57.4\ \mu M$. In summary, a rapid procedure is outlined for the preparation of (c-MycΔ2-99)-$His_6$ that will be useful for mechanistic and biophysical studies of ERK2.


*Keywords:* MAPK kinase; ERK2; c-Myc

Mitogen-activated protein kinases (MAPKs)[1] represent a specific class of enzyme, which catalyze the transfer of phosphate from ATP to protein substrates. They differ from other well-studied kinases in that their natural substrates are not small molecule acceptors, such as glycerol, but large dynamic polypeptides with many potential sites for phosphorylation. Recent evidence suggests that MAPKs take advantage of the special interactions afforded them by protein substrates, to stabilize key ternary complex arrangements that facilitate the targeting of certain serine or threonine residues that lie on the surface of these substrates [1]. However, the mechanism of how this targeting is achieved is not well-understood [2]. Early studies with ERK1/2, identified φ-χ-Ser/Thr-Pro (where φ is typically a small hydrophobic residue and χ appears to be unrestricted)

* Corresponding author. Fax: +1 512 2322606.
*E-mail address:* Dalby@mail.utexas.edu (K.N. Dalby).

[1] *Abbreviations used:* BSA, bovine serum albumin fraction V; DTNB, 5,5′-dithiobis(2-nitrobenzoic acid); DTT, dithiothreitol; EDTA, ethylene diamine tetraacetic acid; EGTA, ethylene glycol-bis[2-aminoethyl ether]-*N*,*N*,*N*′*N*′-tetraacetic acid; ERK, extracellular signal-regulated protein kinase; Hepes, *N*-(2-hydroxyethyl)-piperazine-*N*′-2-ethanesulfonic acid; IPTG, isopropyl-β-D-thiogalactopyranoside; MAPK, mitogen-activated protein kinase. (c-MycΔ2-99)-$His_6$ corresponds to the residues 2–99 of c-Myc flanked by the N-terminal sequence GSMPGRPL and the C-terminal sequence HHHHHH.

1046-5928/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.pep.2006.12.004

as a minimal consensus sequence for the phosphorylation of synthetic polypeptides on serine or threonine [3]. However, short synthetic peptides that are based on this consensus sequence display Michaelis constants that are high relative to natural substrates [4]. It is now realized that the greater specificity exhibited by ERK2 towards natural substrates is often a result of 'docking' interactions that occur between ERK2 and discrete docking sites which lie on the surface of a substrate, often several Angstroms from the phosphorylation site [5].

Two types of docking site have been characterized and consensus sequences have been proposed for them that reflect the important features of their interactions[2]. These docking sites are termed D- [6] and F-sites [7,8] and are proposed to interact with the D- and F-recruitment sites of ERK2 (and other MAPKs), respectively. The D-recruitment site of the MAPKs is centered on a groove that is formed, in part, by a reverse turn between two β-sheets (β7 and β8) (see [9]). A number of structural studies have revealed how D-site sequences often bind this groove in an extended configuration and then extend into a region characterized by some conserved acidic residues that constitute the common docking domain [10–13]. Typically, the peptide sequences that engage the groove are recognized through both their side chain and backbone atoms with none of the docking interactions being highly conserved. The F-site is thought to bind a hydrophobic pocket near the activation segment and the MAPK insert [14].

Due to the frequency that consensus sequences for these docking sites appear in proteins, it is impossible to identify a functional docking site based on the analysis of a sequence alone. Furthermore, a number of important proteins appear to interact with MAPKs using docking interactions that are currently uncharacterized. These observations in particular, underline the need for more structural information. One such protein, of high significance, is c-Myc, which is a nuclear phosphoprotein that is involved in apoptosis, cell cycle progression, and cellular transformation. It operates as a transcription factor and regulates the transcription of specific target genes. Its over-expression, mutation, rearrangement, and translocation have been linked to a variety of leukemias, lymphomas, and hematopoietic tumors [15–18]. Recently, high levels of phosphorylated c-Myc were discovered in neurons from brains affected with Alzheimer disease [19]. In quiescent cells c-Myc is extremely unstable, however, upon entry into the cell cycle following serum stimulation it becomes transiently stabilized and accumulates to higher levels [20]. This stabilization is thought to be mediated by the phosphorylation of c-Myc on Ser-62 by ERK2 [21].

Thus, a biophysical description of how ERK2 recognizes c-Myc is of high physiological significance. Early studies concluded that ERK2 specificity determinants are located in residues 1–100 of c-Myc [22]. Interestingly, this region does not contain a clearly defined D- or F-site, suggesting that a detailed structure/function analysis is necessary to elucidate the mode of interaction. To pursue both structural and kinetic studies a rapid, but efficient purification method, for production of (c-MycΔ2-99)-$His_6$ from an *Escherichia coli* source, was developed. We chose the minimal number of steps to maximize both yield and efficiency without sacrificing purity. The purified substrate is obtained in good yield (8 mg of purified substrate/liter of culture) and notably is an excellent substrate for ERK2 ($k_{cat} = 10.4\ s^{-1}$, $K_M^{c\text{-}Myc} = 57.4\ \mu M$, under assay conditions). Furthermore, it can be stored at conveniently high concentrations (1 mM) in 5 mM Hepes buffer at −80 °C for an indefinite amount of time. The procedure (outlined below) should prove valuable for reproducibly purifying large amounts of (c-MycΔ2-99)-$His_6$ for mechanistic and biophysical studies of ERK2.

## Methods

### *Buffers and reagents*

Trizma base was purchased from EM Industries (Gibbstown, NJ) and ammonium carbonate from Fisher (Fair Lawn, NJ). All other buffer components and chemicals were obtained from Sigma (St. Louis, MO). Qiagen Inc., (Santa Clarita, CA) supplied Ni–NTA agarose. Kinase assays were conducted with Roche (Indianapolis, IN) special quality adenosine triphosphate (ATP) and [γ-$^{32}$P]ATP from ICN (Costa Mesa, CA). Proteins were digested with sequencing grade trypsin from Roche (Indianapolis, IN). Activated ($His_6$-tagged) ERK2 was generated essentially as described previously [23].

### *Site directed mutagenesis*

The pGEX 2T1 vector containing DNA encoding residues 2–99 of c-Myc [24], was modified by PCR mutagenesis to give a C-terminal hexahistidine tag immediately after Gln-99. The tag was incorporated in one PCR reaction by making use of the unique *Bam*HI and *Eco*RI sites that were originally used to subclone the c-Myc DNA into the pGEX 2T-1 vector. The forward primer (5′-CTGGTT CCGCGTGGATCCATGCC-3′) contained a *Bam*HI site, while the reverse primer (5′-CTAGAATTCTAATGATG ATGATGGAGATGCTGATCGGCGGTGGAGAAGT-3′) contained an *Eco*RI site. The polymerase chain reactions (0.1 mL) contained 20 mM Tris–HCl, pH 8.4, 50 mM KCl, 1.5 mM $MgCl_2$, 0.4 mM of each deoxynucleotide triphosphate, 50 pmol of each oligonucleotide, 50 ng pGEX2T-1 c-MycΔ2-99 as template, and 5 U of *Taq* polymerase (Cetus). Conditions were: 94 °C for 4 min to denature the samples, followed by 5 cycles of 53 °C for 45 s, 72 °C for 90 s, and 94 °C for 60 s; then 5 cycles of 56 °C for 45 s, 72 °C for 90 s, and 94 °C for 60 s; then 15 cycles of 63 °C for 45 s, 72 °C for 90 s, and 94 °C for 60 s. The reactions were extracted with an equal volume of phenol/

[2] These consensus sequences should only be taken as a guide, however, as they pertain to the entire MAPK family, are generally weak and may sometimes only contribute a fraction of the binding energy of a complex.