Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Biotechnology 121 (2006) 165–173
Journal of BIOTECHNOLOGY
www.elsevier.com/locate/jbiotec

Engineered proteins containing the cohesin and dockerin domains from *Clostridium thermocellum* provides a reversible, high affinity interaction for biotechnology applications

Scott J. Craig [a], Frances C. Foong [b,*], Robert Nordon [a]

[a] *Graduate School of Biomedical Engineering, University of New South Wales, NSW, Australia*
[b] *School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, NSW 2052, Australia*

Received 25 January 2005; received in revised form 23 June 2005; accepted 4 July 2005

Abstract

The cohesin–dockerin interaction, which is responsible for the formation of the cellulosome complex of cellulolytic bacteria, is a calcium-dependent, high affinity interaction. In this study, the cohesin (Cip7) and dockerin (Doc) domains of *Clostridium thermocellum* were fused to the cellulose-binding domain (CBD) of *C. cellulovorans* and the antibody-binding domain, protein LG, respectively, to form CBD-Cip7 and LG-Doc. Immobilised CBD-Cip7 was able to bind LG-Doc and subsequently antibody as determined using surface plasmon resonance. Binding was reversed by the removal of Ca^{2+} with EDTA. The dockerin containing fusion protein was affinity purified using an immobilised cohesin domain. Elution of the LG-Doc from the cohesin column was with EDTA. This affinity chromatography was repeated using an LG-dockerin column for the purification of cohesin fusion protein. The fusion proteins created in this report have shown that the properties of the cohesin and dockerin domains can be transferred to other protein domains and that the interaction between the cohesin and dockerin is specific, Ca^{2+}-dependent and reversible. We have shown that the cohesin–dockerin interaction has several properties making it suitable for use in recombinant fusion protein production and purification.

Keywords: Cohesin; Dockerin; Cellulosome; *Clostridium thermocellum*; Fusion proteins; Affinity chromatography; Purification

1. Introduction

Many cellulose degrading bacteria have evolved to produce a secreted multicomponent enzyme complex for the efficient breakdown of the resistant cellulose substrate (Schwarz, 2001). This complex, termed the cellulosome, has been most widely studied in the anaerobic Clostridia, namely *Clostridium thermocellum* (Schwarz, 2001), *C. cellulovorans* (Doi et al., 1998) and *C. cellulolyticum* (Gal et al., 1997).

The cellulosome has a common structure consisting of a central non-enzymatic scaffolding protein,

* Corresponding author. Tel.: +61 2 93853872; fax: +61 2 93136710.
E-mail address: f.foong@unsw.edu.au (F.C. Foong).

0168-1656/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbiotec.2005.07.005

which contains a cellulose-binding domain and several conserved repeated hydrophobic domains termed cohesins (Gerngross et al., 1993; Schwarz, 2001). The remaining components are glycosyl hydrolase enzymatic subunits of various specificities such as exo-glucanases, endo-glucanases, xylanases and lichenases that act synergistically to efficiently break down the cellulose substrate. Each of the cellulosomes' enzymatic subunits contains, in addition to their catalytic domains, a highly conserved duplicated sequence, which has been termed a dockerin (Doc) (Beguin and Lemaire, 1996; Gal et al., 1997). The high molecular weight cellulosome complex is formed by the attachment of the enzymatic domains to the scaffolding backbone via an interaction between the dockerin domains of the enzymes and the cohesin domains of the scaffolding. Early attempts to dissociate the *C. thermocellum* cellulosome showed that it was extremely stable even after treatment with the denaturants urea and guanidine hydrochloride (Lamed et al., 1983; Ljungdahl et al., 1988). Full dissociation of the cellulosome required treatment with sodium dodecyl sulfate (SDS), heat and reducing agents such as dithiothreitol (DTT) (Wu et al., 1988; Bhat and Wood, 1992; Mori, 1992). Subsequent studies showed that the complex could also be dissociated using a more gentle treatment of low SDS concentration, EDTA and reducing conditions (Beattie et al., 1994). It has been found that Ca^{2+} plays a key role in maintaining the structural integrity of the cellulosome complex and the activity of several components (Chauvaux et al., 1990; Choi and Ljungdahl, 1996a,b).

Studies with the cohesin domain from CipA of *C. thermocellum* have shown that the binding of dockerin containing proteins is not specific to any one cohesin (Lytle et al., 1996). Likewise in *C. cellulolyticum* the CelA endoglucanase interacts with all eight cohesin domains of the scaffolding protein CipC (Pages et al., 1999). This is despite the sequence variability between the various cohesin domains. The binding of the dockerin is however specific to cohesins of the same species (Pages et al., 1997; Mechaly et al., 2000). The dockerin–cohesin interaction is an extremely high affinity interaction with dissociation constants reported between 4.8×10^{-9} and 2.5×10^{-10} M (Pages et al., 1996, 1999; Fierobe et al., 1999).

It has been suggested that the components of the cellulosome could provide several useful tools for biotechnology applications (Bayer et al., 1994). The properties of the dockerin–cohesin interaction would be useful in the areas of affinity detection and purification. The aim of this study is to investigate the dockerin–cohesin interaction of *C. thermocellum* and to determine whether it is possible to retain the binding properties of these domains in chimeric proteins for biotechnology applications. This study will utilise a fusion protein design incorporating the dockerin domain to confirm that the properties of the dockerin are readily transferable to target proteins. This fusion protein will contain the antibody-binding protein, protein LG (Kihlberg et al., 1996) in combination with the dockerin domain of CelD (Beguin et al., 1987; Tokatlidis et al., 1993) from *C. thermocellum*. Protein LG consists of protein G from *Streptococcus* combined with protein L of *Peptostreptococcus magnus* (Kihlberg et al., 1992) and has been shown to bind a large range of human, mouse and rat antibodies and their fragments (Kihlberg et al., 1996). The 7th cohesin domain (Cip7) from CipA of *C. thermocellum* will be incorporated into a mixed species "mini-scaffolding" fusion protein consisting of a cellulose-binding domain (CBD) from *C. cellulovorans* fused to the *C. thermocellum* cohesin domain for immobilisation applications. The interaction between the various component domains of the fusion proteins will be tested using real-time biosensor analysis and affinity chromatography.

2. Materials and methods

2.1. Bacterial strains and vectors

E. coli strain XL10gold was obtained from Stratagene (La Jolla, CA, USA) and used for all DNA manipulation steps. *E. coli* strains BL21(DE3) pLysS and JM109(DE3) and pGEM-T Easy were obtained from Promega (Madison, WI). Vector pET34b+ (containing the *C. cellulovorans* CBD) and pET23d+ was obtained from Novagen (EMD Biosciences, USA).

The plasmid pCT330 (Tokatlidis et al., 1993), and strain M15(pREP4) pCip7 (Kataeva et al., 1997) were kindly provided by Dr. Pierre Beguin (Institut Pasteur, Paris, France). The plasmid pLG (Kihlberg et al., 1992) was kindly provided by Dr. Ulf Sjobring (University of Lund, Lund, Sweden).