Expression, purification and subunit-binding properties of cohesins 2 and 3 of the *Clostridium thermocellum* cellulosome

Sima Yaron[a], Ely Morag[b], Edward A. Bayer[b], Raphael Lamed[c], Yuval Shoham[a,*]

[a]*Department of Food Engineering and Biotechnology, Technion–Israel Institute of Technology, Haifa, Israel*
[b]*Department of Biophysics, The Weizmann Institute of Science, Rehovot, Israel*
[c]*Department of Molecular Microbiology and Biotechnology, Tel Aviv University, Ramat Aviv, Israel*

Received 24 November 1994; revised version received 15 January 1995

Abstract **The enzymatic subunits of the cellulosome of *Clostridium thermocellum* are integrated into the complex by a major non-catalytic polypeptide, called scaffoldin. Its numerous functional domains include a single cellulose-binding domain (CBD) and nine subunit-binding domains, or cohesin domains. Two of the cohesin domains, together with the adjacent CBD, have been cloned and expressed in *Escherichia coli*, and the recombinant constructs were purified by affinity chromatography on a cellulosic matrix. Both cohesin domains, which differ by about 30% in their primary structure, showed a similar binding profile to the cellulosomal subunits. Calcium ions enhanced dramatically this binding. Under the conditions of the assay, only one major catalytic subunit of the cellulosome failed to bind to either cohesin domain. The results indicate a lack of selectivity in the binding of cohesin domains to the catalytic subunits and also suggest that additional mechanisms may be involved in cellulosome assembly.**

Key words: Cellulosome (*Clostridium thermocellum*); Cellulase; Multi-enzyme complex; Scaffoldin; Cohesin domain; Subunit-integrating domain

1. Introduction

In many different cellulolytic bacteria, the causative agent which leads to efficient solubilization of cellulosic substrates appears to be a multicomponent, multifunctional enzyme complex, called the cellulosome [1,2]. The cellulolytic enzymes occur as distinct subunits of the cellulosome and are held together by a novel type of non-catalytic polypeptide [3–6], termed scaffoldin [7]. The scaffoldins include a cellulose-binding domain (CBD) and a succession of subunit-binding domains, or *cohesin domains*. The enzymes themselves (endoglucanases, exoglucanases and/or xylanases) are composed of two key types of component; one or more catalytic domains, which determine the hydrolytic character and specificity of the enzyme, and a docking domain (*dockerin domain*), which binds strongly to the cohesin domains of the scaffoldin subunit. Thus, the unique cohesin–dockerin interaction appears to be of definitive importance to the construction of the cellulosome and independent of its ability to bind to the target substrate or to break it down [7].

The scaffoldin subunits of two different bacteria, *Clostridium thermocellum* (ATCC 27405) and *Clostridium cellulovorans* (ATCC 35296), have been cloned and sequenced [8,9]. Independently, we have cloned and sequenced a segment of the scaffoldin subunit from a different strain of *C. thermocellum* YS [10], which included the CBD and selected cohesin domains. The CBD was subsequently expressed and purified in an *Escherichia coli* host system [11]. In the current communication, we concentrate on the two cohesin domains immediately adjacent to the CBD. The constructs, which consisted of the two cohesin domains linked to the CBD (either separately or together), were expressed and isolated. The subunit-binding characteristics of the constructs were assessed by Western blotting.

2. Materials and methods

2.1. Bacterial strains and vectors

Escherichia coli strains BL21(DE3) and BL21(DE3)pLysS and the T7 RNA polymerase expression vector pET9d (all from Novagen, Madison, WI) were described elsewhere [12]. *E. coli* strain XL-1 Blue was obtained from Stratagene (La Jolla, CA). Plasmid pT*rc*99A was purchased from Pharmacia. *Clostridium thermocellum* YS was described earlier [1,2].

2.2. DNA manipulation

DNA was manipulated by standard procedures [13,14]. DNA transformation was performed using either the calcium chloride method [14] for strains BL21(DE3) and BL21(DE3)pLysS or by electroporation using Gene Zapper (IBI, New Haven, CT) for strain XL-1 Blue.

2.3. Cloning the cohesin-CBD constructs

Cohesins 2 and 3 were cloned together with the CBD in 3 separate configurations. For this purpose, four synthetic DNA primers were prepared with partial homology to the N-termini of cohesin 2 and CBD and to the C-termini of the CBD and cohesin 3 (see Fig. 1). The primers were based on the sequence of the scaffoldin subunit (CipA), as reported by Gerngross et al. [9].

The cohesin-CBD segments were amplified by the polymerase chain reaction (PCR) from total chromosomal DNA of *C. thermocellum* strain YS, prepared as described previously [10]. Vent DNA polymerase (New England Biolabs, Beverly, MA) was used, under conditions recommended by the manufacturer. The PCR products were separated and extracted from agarose gels with activated glass beads (Geneclean II kit, Bio 101, La Jolla, CA) and cleaved with *Nco*I and *Bam*HI. The cleaved fragments were again separated, extracted and ligated with *Nco*I–*Bam*HI linearized pET9d or pT*rc*99A plasmid DNA. The ligation mixtures were used to transform competent *E. coli* XL-1 Blue cells by electroporation. Kanamycin-resistant colonies were isolated and their plasmid DNA was subject to DNA sequencing to verify the constructs.

The resulting plasmids were designated p2CBD (containing cohesin 2 and CBD), pCBD3 (containing CBD and cohesin 3), and p2CBD3 (containing both cohesins 2 and 3 with CBD between the two). The primer for cohesin 3 contained a *Bam*HI site at its C-terminus. When cloned into the *Nco*I–*Bam*HI sites of pET9d, an additional 22-amino acid segment, originating from the vector, was added to the carboxy terminus of the corresponding gene products. Alternatively, the constructs were cloned similarly into pT*rc*99A which resulted in a shorter (3 amino acid) C-terminal extension.

*Corresponding author. Fax: (972) (4) 320 742.

0014-5793/95/$9.50 © 1995 Federation of European Biochemical Societies. All rights reserved.
SSDI 0014-5793(95)00074-7

2.4. Expression

Plasmid DNA was used to transform *E. coli* strains containing an inducible T7 polymerase BL21(DE3) or BL21(DE3)(pLysS). The transformed cells were grown on LB medium with the appropriate antibiotics (kanamycin with or without chloramphenicol) to an OD_{600} of 0.6 to 0.9. Induction was initiated by adding 0.1 to 0.4 mM of isopropyl-β-D-thiogalactoside (IPTG) and growing the cells for an additional 6 to 8 h. Alternatively, cultures were grown for 12 to 18 h without induction.

2.5. Purification of recombinant constructs by affinity chromatography

One-liter cultures, containing the appropriate plasmids (either after induction or without induction) were centrifuged, and resuspended in 50 ml of Tris buffer (Tris-HCl, 50 mM with 2 mM EDTA). The cells were then sonicated on ice and centrifuged again. Two grams of microcrystalline cellulose (Avicel Type PH-101, FMC) was added to the clear supernatant fluids, and the suspension was stirred for 1 h at 4°C. After centrifugation, the pellet was washed twice with 50 ml of phosphate buffer (50 mM, pH 7.4, containing 1 M NaCl) and twice with the same buffer without NaCl. The cohesin-CBD segments were eluted from the cellulosic matrix with 8 ml of 1% triethylamine. The eluents were neutralized with phosphate buffer (0.2 M, pH 7) and kept at 4°C.

2.6. Miscellaneous methods

The cellulosome was purified from culture broth of *C. thermocellum* using the affinity digestion procedure [15]. The cohesin-CBD constructs were biotinylated using biotinyl *N*-hydroxysuccinimide ester at a 20-fold molar ratio of reagent to protein as previously described [16]. SDS-PAGE and affinity blotting were carried out according to Morag et al. [17]. In experiments on the calcium effect, blots were incubated in the presence of either 15 mM $CaCl_2$ or 5 mM EDTA for 1 h at 23°C, prior to introduction of the desired biotinyl cohesin–CBD construct, diluted with the same reagents. Protein was determined by the Bradford method using bovine serum albumin as a standard [18].

3. Results

3.1. Cloning, expression and purification of cohesins 2 and 3

In order to clone and express cohesins 2 and 3 of the scaffoldin subunit from the *C. thermocellum* cellulosome, we chose to express them together with the adjacent cellulose-binding domain (CBD). Such a design enables the isolation of resultant constructs on cellulose and may also overcome their possible toxicity to the *E. coli* host system. The resultant constructs were termed Coh2–CBD, CBD–Coh3 and Coh2–CBD–Coh3. DNA sequencing revealed that the segment of the scaffoldin gene encoding for Coh3 of *C. thermocellum* strain YS is identical to that of ATCC strain 27405. In the earlier work on the scaffoldin gene from strain YS [10], complementary portions of Coh3 and Coh9 were fused to form a chimeric cohesin domain, presumably as a result of a major deletion event.

To estimate the levels of expression for each clone, the cells were sonicated, the soluble fraction of each extract was introduced to a cellulose resin, and the adsorbed material was analyzed by SDS-PAGE. Without induction, similar basal levels of each construct could be detected. Several methods of induction were examined to optimize the level of expression. None of these resulted in improved levels of expression (data not shown), presumably due to the toxicity of the expressed polypeptide to the host cell.

The final amount of purified protein obtained for each clone was approximately 10 mg per liter of culture. The proteins obtained were not very soluble and tended to precipitate upon storage. The SDS-PAGE profile of the cohesin-CBD constructs is shown in Fig. 2. The M_rs of the purified products were in agreement with the theoretical calculated values.

3.2. Interaction of recombinant cohesin domains with cellulosomal subunits

It was important to determine whether the two recombinant cohesin domains interacted selectively with cellulosomal subunits, and, particularly, whether a given cohesin domain interacted specifically with a single subunit or group of subunits. Moreover, we were interested to know whether there would be differences in the recognition pattern between the two cohesin domains.

In order to address these questions, biotinylated cohesin–CBD constructs were subjected to interaction with Western blots of SDS-PAGE-separated cellulosome subunits (between 0.5 and 5 μg of cellulosomal protein per sample). The labeling pattern was then developed using an avidin-complexed enzyme system. Titration of the interaction using various concentrations of each construct (between 1 ng and 5 μg protein per sample) was thus performed, yielding essentially an indistinguishable set of labeling patterns among parallel samples. At

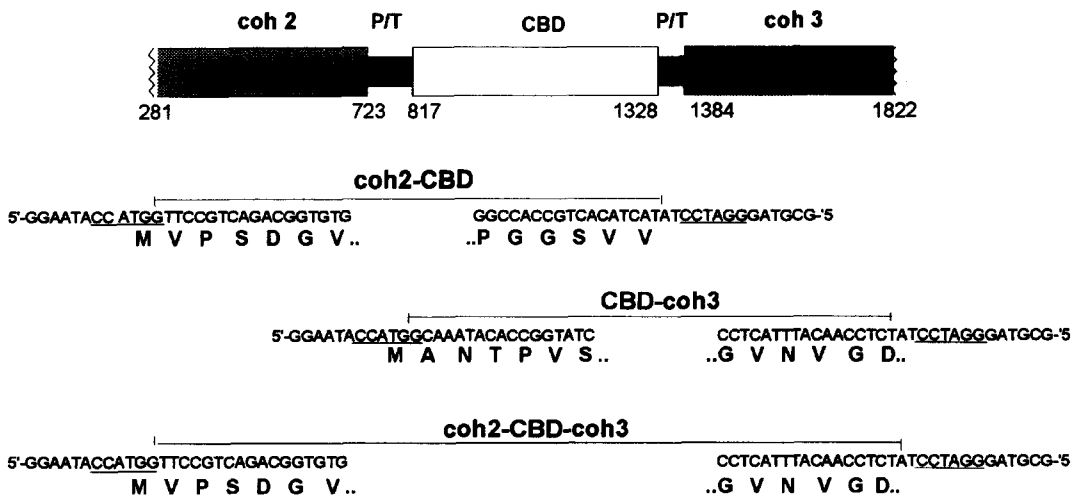

Fig. 1. Schematic representation of the amplified regions from the *cip* gene of *C. thermocellum* YS. The primers used to amplify the appropriate regions are shown together with the amino acid sequence (bold type) at the beginning and end of each segment. The *Nco*I and *Bam*HI sites which appear in the N- and C-terminal primers, respectively, are underscored.

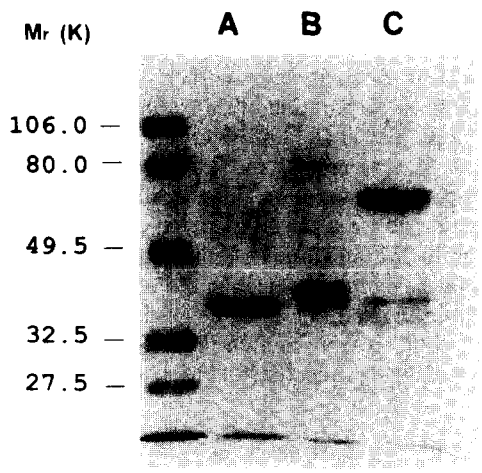

Fig. 2. SDS-PAGE of purified cohesin-CBD products. (A) Coh2–CBD; (B) CBD–Coh3; (C) Coh2–CBD–Coh3. The calculated values for the molecular weights of the three constructs are 37,655, 39,670 and 58,651, respectively.

the lowest concentrations tested, the S8 subunit (CelS) was the predominant band detected.

The results using elevated levels of the probes are shown in Fig. 3. Both cohesins 2 and 3 appeared to interact with all of the cellulosomal subunits between S3 and S13 (inclusive). In both cases, the pattern of label was almost identical – the major labeling occurred with subunits S5, S8, S9(10) and S13. Moreover, competition experiments (i.e. excess concentrations of underivatized cohesin 2 vs. biotinylated cohesin 3 and vice versa) showed a similar weakening of label in both cases (data not shown). Interestingly, cellulosomal subunit S2 was not recognized at all by either cohesin domain, and, in both cases, the interaction with subunit S3 was particularly weak. As expected, neither cohesin domain interacted with subunit S1, i.e. the scaffoldin subunit itself.

These results were confirmed using the same constructs with a different labeling procedure, which employed anti-CBD antibodies as a detection probe. The results (data not shown) revealed essentially the same labeling pattern as presented in Fig. 3, except, of course, for a strong labeling of the S1 (scaffoldin) subunit, which emanated from the interaction of the anti-CBD antibodies with the resident CBD.

3.3. *The calcium effect*

Calcium dependence of the cellulolytic activity of the cellulase system and, in particular, the cellulosome of *C. thermocellum* has been noted earlier [19–22]. It was therefore of interest to determine whether the interaction between the recombinant cohesin domain and the cellulosomal subunits would be affected by calcium. Indeed, as seen in Fig. 4, the addition of calcium ions resulted in a strong enhancement of the labeling pattern using the CBD–Coh3 construct as a probe. An identical effect was also observed with Coh2–CBD (data not shown). In both cases, 5 mM EDTA resulted in the complete or near-complete elimination of the interaction between the cohesins and the S3, S5, S11, S12 and S13 subunits. The label associated with the S4, S6, S7, S8, and S9(10) subunits were highly reduced but still visible. Added calcium increased markedly the cohesin-induced label associated with all of the subunits – save the S2 subunit which was still unlabeled.

4. Discussion

In *C. thermocellum* and various other cellulolytic bacteria, the singular cohesin–dockerin interaction (which takes place between the non-catalytic scaffoldin subunit and the different catalytic subunits) appears to dictate the formation of discrete multi-enzyme complexes known as cellulosomes [1,2]. This premise was elegantly demonstrated for *C. thermocellum* by fusing a duplicated segment (i.e. a dockerin component derived from a confirmed cellulosomal subunit) to the C-terminus of a noncellulosomal endoglucanase from the same organism [23]; the resultant chimeric protein was then shown to interact with a recombinant form of a cohesin domain of scaffoldin [24]. Likewise, an expressed cohesin domain (namely, an 'hydrophobic repeated domain'), derived from the cellulosomal scaffoldin subunit in *C. cellulovorans*, was similarly shown to interact with two cellulosomal endoglucanases [25]. Indeed, heterogeneous populations of cellulase complexes have recently been reported in yet another cellulolytic species [26].

In this communication, we investigated the cohesin–dockerin interaction further. Specifically, it was of interest to determine whether an individual cohesin domain interacts in a selective manner with a single dockerin domain located on a given catalytic subunit, or whether the cohesin–dockerin interaction is less selective in nature. In this context, lack of specificity of cellulosomal assembly may suggest that the character of individual cellulosome complexes reflects different levels of induction of the enzymatic subunits.

Of the nine cohesin domains of this organism, we chose to clone and express two, i.e. cohesins 2 and 3, for several reasons. First, both are located, in the native state, immediately adjacent to the cellulose-binding domain (CBD) of the scaffoldin subunit; and the expression of the cohesin domains together with the CBD thus provided us with a handle for their facile purification on a cellulose affinity column. Indeed, the use of the naturally fused cohesin–CBD construct provides a probe which is closer to the inherent condition. Secondly, although all of the

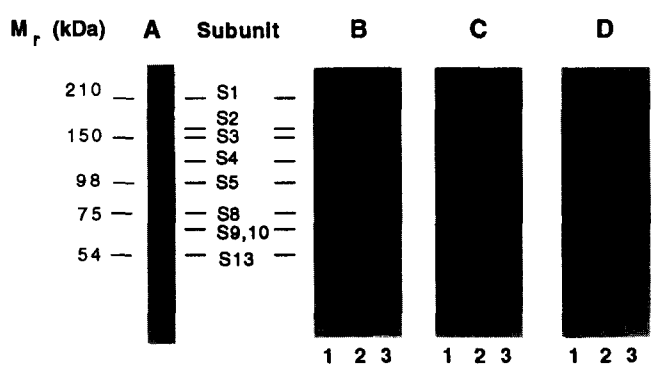

Fig. 3. Recognition of cellulosomal subunits by recombinant cohesin domains. The cellulosome preparation was subjected to SDS-PAGE, the separated subunits were blotted onto nitrocellulose and probed with biotinylated forms of the cohesin–CBD constructs (0.4 μg per sample). The blots were developed using avidin–peroxidase complexes. (A) Amido black staining of the original blot (total protein). (B–D) Affinity blotting with Coh2–CBD, CBD–Coh3 and Coh2–CBD–Coh3, respectively. (1–3) Different amounts (4.8, 1.6 and 0.48 μg, respectively) of cellulosomal protein were applied to the designated lanes of the gel. Since the recombinant cohesin domains were expressed as constructs together with the CBD, a biotinylated CBD preparation (without the cohesin domains) served as a negative control and failed to bind to any of the subunits of the SDS-PAGE-separated cellulosome (not shown).

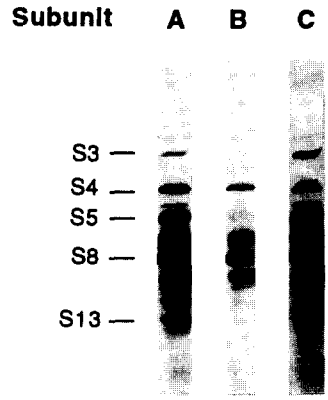

Fig. 4. The effect of calcium and EDTA on the interaction of the CBD–Coh3 construct with the cellulosomal subunits. The experiment (lane A) was carried out as described in the legend to Fig. 3 using 2.7 μg of cellulosomal protein per lane. The blot in lane B was pretreated with 5 mM EDTA, whereas the blot in lane C was pretreated with 15 mM $CaCl_2$.

nine cohesin domains on the scaffoldin subunit are quite similar to each other (the *least* similar of which still exhibit about 60% sequence identity) cohesins 2 and 3 are relatively distinct in their primary structures. We therefore considered that, were there a significant difference in the binding properties among the different cohesin domains, we may be able to detect such a difference by using these two cohesin domains as representative probes.

The results showed clearly that both cohesin 2 and cohesin 3 recognize a very similar group of subunits in a very similar manner. This suggests that the incorporation of the catalytic subunits into the cellulosome complex may not be a particularly selective process, or, if it is, the selectivity might not be a strict function of the cohesin-dockerin interaction per se. In addition, the finding in the present work that both cohesin domains recognize all of the enzymatic subunits except S2 reinforces previous observations of Tokatlidis et al. [5]. These authors showed that antibodies against the duplicated segment of endoglucanase CelD (i.e. the dockerin domain of cellulosomal subunit S11) label some of the cellulosomal components corresponding to subunits S5 to S13. As shown in the present communication, subunits S3, S4 and others can also be added to this list. Consequently, it now appears that most of the subunits indeed bear a similar type of dockerin domain and that, collectively, they appear to interact with the majority of the cohesin domains on the scaffoldin subunit.

The enhancement of the cohesin–dockerin interaction by calcium adds to the growing list of contributions to cellulosome action by this divalent cation. Calcium is known to stimulate or stabilize cellulolytic activity of the cellulase system [19], the intact cellulosome [20], and its components [17,27]. Interestingly, a conserved portion of the dockerin component of the cellulosome bears a pronounced resemblance to calcium-binding motifs in various proteins [27], but its exact role in this capacity has yet to be demonstrated. In view of the results presented herein, it would be worthwhile reinvestigating the binding of calcium to cellulosomal components, particularly with respect to the cohesin–dockerin interaction.

The question still remains as to whether cohesins 1 and 9, i.e. those exhibiting the highest number of substitutions, are more selective in their interaction with the cellulosomal subunits. Perhaps one of them binds to the largest catalytic subunit (namely S2) or, alternatively, another type of interaction is responsible for incorporation of this subunit into the complex. In this context, the scaffoldin subunit includes two additional domains at its C-terminus, and the possibility remains that additional components on the intact scaffoldin may modulate the interaction between the cohesin and dockerin domains.

In any event, the integration of cellulosomal subunits into the complex may be more intricate than the simplistic view, implicit in the cohesin–dockerin model. The elucidation of the mechanism(s) of cellulosome assembly will provide continued insight into the microbial production of discrete multienzyme complexes.

References

[1] Lamed, R., Setter, E., Kenig, R. and Bayer, E.A. (1983) Biotechnol. Bioeng. Symp. 13, 163–181.
[2] Lamed, R. and Bayer, E.A. (1988) Adv. Appl. Microbiol. 33, 1–46.
[3] Bayer, E.A., Setter, E. and Lamed, R. (1985) J. Bacteriol. 163, 552–559.
[4] Wu, J.H.D., Orme-Johnson, W.H. and Demain, A.L. (1988) Biochemistry 27, 1703–1709.
[5] Tokatlidis, K., Salamitou, S., Béguin, P., Dhurjati, P. and Aubert, J.-P. (1991) FEBS Lett. 291, 185–188.
[6] Salamitou, S., Tokatlidis, K., Béguin, P. and Aubert, J.-P. (1992) FEBS Lett. 304, 89–92.
[7] Bayer, E.A., Morag, E. and Lamed, R. (1994) Trends Biotechnol. 12, 378–386.
[8] Shoseyov, O., Takagi, M., Goldstein, M.A. and Doi, R.H. (1992) Proc. Natl. Acad. Sci. USA 89, 3283–3287.
[9] Gerngross, U.T., Romaniec, M.P.M., Kobayashi, T., Huskisson, N.S. and Demain, A.L. (1993) Mol. Microbiol. 8, 325–334.
[10] Poole, D.M., Morag, E., Lamed, R., Bayer, E.A., Hazlewood, G.P. and Gilbert, H.J. (1992) FEMS Microbiol. Lett. 99, 181–186.
[11] Morag, E., Lapidot, A., Govorko, D., Lamed, R., Wilchek, M., Bayer, E.A. and Shoham, Y., submitted.
[12] Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990) Methods Enzymol. 185, 60–89.
[13] Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1992) Wiley, New York.
[14] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratories, Cold Spring Harbor, New York.
[15] Morag, E., Bayer, E.A. and Lamed, R. (1992) Enz. Microb. Technol. 14, 289–292.
[16] Bayer, E.A. and Wilchek, M. (1990) Methods Enzymol. 184, 138–160.
[17] Morag, E., Halevy, I., Bayer, E.A. and Lamed, R. (1991) J. Bacteriol. 173, 4155–4162.
[18] Bradford, M. (1976) Anal. Biochem. 72, 248–254.
[19] Johnson, E.A., Sakojoh, M., Halliwell, G., Madia, A. and Demain, A.L. (1982) Appl. Environ. Microbiol. 43, 1125–1132.
[20] Lamed, R., Kenig, R., Setter, E. and Bayer, E.A. (1985) Enz. Microb. Technol. 7, 37–41.
[21] Hon-nami, K., Coughlan, M.P., Hon-nami, H. and Ljungdahl, L.G. (1986) Arch. Biochem. Biophys. 145, 13–19.
[22] Beattie, L., Bhat, K.M. and Wood, T.M. (1994) Appl. Microbiol. Biotechnol. 40, 740–744.
[23] Tokatlidis, K., Dhurjati, P. and Béguin, P. (1993) Protein Engineer. 6, 947–952.
[24] Salamitou, S., Raynaud, O., Lemaire, M., Coughlan, M., Béguin, P. and Aubert, J.-P. (1994) J. Bacteriol. 176, 2822–2827.
[25] Takagi, M., Hashida, S., Goldstein, M.A. and Doi, R.H. (1993) J. Bacteriol. 175, 7119–7122.
[26] Pohlschröder, M., Leschine, S.B. and Canale-Parola, E. (1994) J. Bacteriol. 176, 70–76.
[27] Chauvaux, S., Béguin, P., Aubert, J.-P., Bhat, K.M., Gow, L.A., Wood, T.M. and Bairoch, A. (1990) Biochem. J. 265, 261–265.