Production and preliminary characterization of a fusion protein comprising streptavidin and a cellulose-binding domain

By

Duy Khai Le

B.Sc., University of Saigon, 1978

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

Department of Microbiology

Genetics Program

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

December 1992

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

(Signature) _____

Department of Microbiology

The University of British Columbia
Vancouver, Canada

Date February 10, 1993

ABSTRACT

The cellulose-binding domain of exoglucanase Cex (CBD_{Cex}) from *Cellulomonas fimi* can be fused to heterologous proteins. The fusion proteins can be purified by affinity chromatography on cellulose or immobilized on cellulose (Ong et al., 1989b).

Streptavidin, a protein produced by *Streptomyces avidinii* , binds to the water-soluble vitamin, D-biotin, with remarkably high affinity. The strong and specific biotin-binding affinity of streptavidin offers a variety of applications in biotechnology (Wilchek and Bayer,1990).

The streptavidin gene was fused in frame to the CBD_{cex} coding sequence and the gene fusion expressed to give a chimeric protein comprising streptavidin and the cellulose-binding domain. The fusion protein was overexpresed in *E. coli* and formed inclusion bodies. The soluble renatured protein recovered from the inclusion bodies bound both biotin and cellulose. It could be used to bind biotinylated proteins to cellulose.

TABLE OF CONTENTS

TABLE OF CONTENTS continued

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATION

BME	ß-mercaptoethanol
BSA	Bovine serum albumin
CBD_{CenA}	CBD of CenA
CBD_{Cex}	CBD of Cex
CenA	*Cellulomonas fimi* endoglucanase A
Cex	*Cellulomonas fimi* exoglucanase
dH_2O	Distilled water
IPTG	Isopropyl-ß-D-thiogalactoside
LB	Luria-Bertani
pNP	p-nitrophenol
pNPP	p-nitrophenyl phosphate
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
kb	kilobase
kDa	kilodalton
CBD	Cellulose-binding domain
EDTA	Ethylenediaminetetraacetate
K_d	Dissociation constant
TAE	Tris-Acetate buffer
TE	10 mM Tris-HCl, 1mM EDTA
STA	streptavidin

ACKNOWLEDGEMENT

I would like to thank Dr. R. A. J. Warren, my supervisor, for his clever and superb supervision. I also thank Drs. R. C. Miller, D.G. Kilburn, N. Gilkes, R. McMaster and T. J. Beatty for their constant support, direction and helpful discussions. I thank Dr. J. Smit for use of the contrast phase microscope in his lab. I am grateful to my many friends in the Cellulase group and in the Department of Microbiology who have given me so many meanings of sharing, laughing, and thinking. I especially thank my wife, Kim Loi Nguyen, who makes my life more interesting, for her constant support and loving consolation. My daughter, Khai Nhu Le, is an endless driving force for me to go forward.

I lovingly dedicate this thesis to my parents, Mr. and Mrs. P.T. Le, who spent their whole life to raise and to educate their children.

1. INTRODUCTION

1.1 Cellulose

Cellulose is the most abundant natural polymer on Earth. It occurs in the cell walls of higher plants, algae and fungi. Cellulose is also synthesized by bacteria as a product of secondary metabolism.

Although cellulose was discovered nearly 150 years ago, its structure has remained a subject of widespread interest to this day. The cause of this interest has been, largely, the structural complexity of native cellulose, from its long-chain, polymeric chemical structure to the morphological features of its crystallinity and aggregation into a fiber. As a result, structural analysis of cellulose has had to be carried out on at least three separate levels: 1) chemical structure, 2) physical structure-conformation and crystalline packing, and 3) fiber structure and supermolecular morphology (Sarko, 1987). Its chemical structure was established beyond dispute long ago (Marchessault and Sarko, 1967) while studies of its crystallinity and morphology are still in progress.

Cellulose is a linear polymer of ß-D- glucopyranosyl units linked by ß-1,4-glucosidic bonds. It crystallizes in four different forms: the polymorphs known as celluloses I, II, III, IV. Of these, only cellulose I occurs naturally, accounting for practically all native cellulose structures seen in terrestrial and aquatic plants, as well as those produced by microorganisms. Cellulose II is the normal conversion product obtained through mercerization or solubilization and regeneration of native celluloses, while celluloses III and IV are obtained, respectively, by liquid ammonia and heat treatments of other polymorphs. All cellulose polymorphs are quite crystalline and exhibit a marked degree of fibre orientation of the crystalline domain, as shown by x-ray diffraction analysis.

In nature, cellulose is degraded by fungi and bacteria. The abundance of various celluloses, has led to the development of new processes for exploitation of cellulose as a food and energy source. Recently, cellulose has been used, in combination with the cellulose-binding domains (CBDs) of bacterial and fungal cellulases, for protein purification and enzyme immobilization (Ong, 1989, Greenwood, 1992).

1.2 Cellulose degradation.

Cellulases are the group of hydrolytic enzymes able to hydrolyze insoluble cellulose to glucose. They are produced by microorganisms and plants. There are three types of enzymes that traditionally have been assigned to cellulase systems : endoglucanases (endo-1,4-ß-glucanases or 1,4-ß-D-glucan 4-glucanohydrolases, EC 3.2.1.4.), cellobiohydrolases (exo-1,4-ß-glucanase or 1,4-ß-D-glucan cellobiohydrolases, EC.3.2.1.91), and cellobiases (ß-glucosidases or ß-D-glucoside glucohydrolases, EC 3.2.1.21). All of the principal enzymes in the cellulase system have been purified to homogeneity in numerous laboratories (Klyosov, 1990).

Cellulolytic enzymes isolated from various sources differ in their molecular characteristics (molecular weight, amino acid composition and sequence, isoelectric point, carbohydrate content), adsordability onto cellulose, catalytic activity, and substrate specificity. Some cellulases, particularly of bacterial origin, are known to be strongly associated with the microbial cells, in contrast to the extracellular fungal cellulases. Cellulase systems often show only two or three individual components (including at least one endoglucanase), but in a number of cases thorough resolution of a cellulase system from a single source revealed 15-20 or more individual components (Hayn & Esterbauer, 1985 ; Knowles et al., 1987; Sprey, 1988). Fungal cellobiohydrolases sometimes show a

marked tendency to form aggregates with endoglucanases, and those aggregates are extremely difficult to break up into the component parts (Wood et al., 1989).

Synergism between the individual components of a cellulase system acting on insoluble cellulose adds further complexities to the study of the mechanisms of action of cellulases. The main problem in studying this phenomenon is that the synergistic effect varies, depending on which cellulases are used in the study, on the sources of the cellulases, and on the cellulose sample (amorphous or crystalline, to mention the two extreme variants) used for the experiment. Definitive quantitative parameters regarding these points have not been found up to the present time.

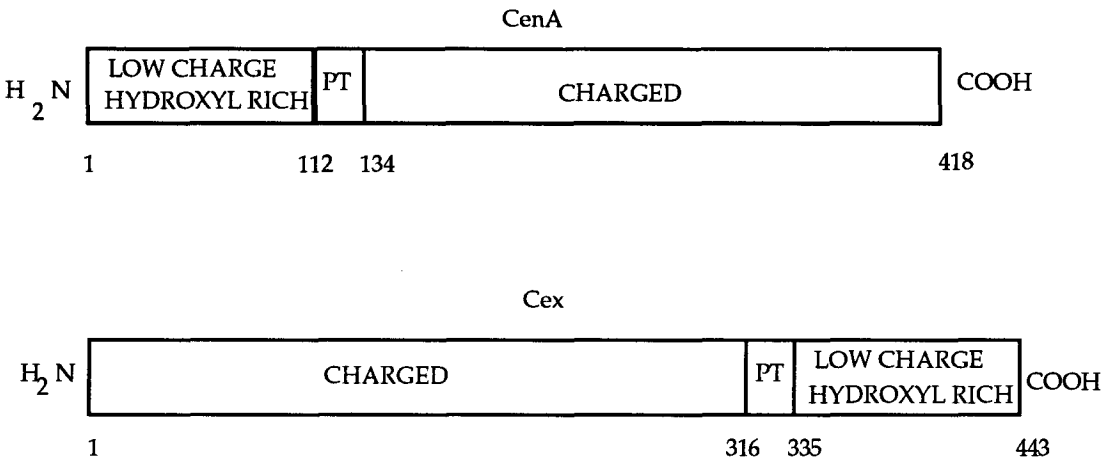
1.3 *Cellulomonas fimi* cellulases.

Cellulomonas fimi is a coryneform, Gram positive, facultatively anaerobic, rod-shaped bacterium with an optimum growth temperature of 30°C and a genome that is 72 mole % G+C (Stackebrandt and Kandler, 1979). It produces a complex array of cellulases when grown on cellulosic substrates (Beguin et al, 1977,1978, Langsford et al. 1984). Its cellulase profile varies with both the nature of the substrate and with culture age, possibly as a consequence of proteolysis and deglycosylation. An exoglucanase (Cex) and an endoglucanase (CenA) bind to the substrate in cultures grown with Avicel, a microcrystalline cellulose. This facilitates their purification to homogeneity by cellulose-affinity chromatography. Both are glycoproteins. Both enzymes hydrolyze carboxymethylcellulose (CMC), although with different kinetics (Gilkes et al, 1984); both release reducing sugar from Avicel; but only Cex hydrolyses p-nitrophenylcellobioside (pNPC) and 4-methylumbelliferylcellobioside (MUC). Both proteins are monomers of very similar size: Cex contains 443 and CenA 418 amino acids.

Each protein is composed of three discrete segments: a sequence of 20 amino acids composed of only prolyl and threonyl residues, termed the Pro-Thr box, which is almost perfectly conserved; a sequence of about 100 amino acids that is rich in hydroxyamino acids, of low charge density, and 50% conserved; and a sequence of about 300 amino acids that has a relatively high charge density, but is not conserved (Warren, 1986). The order of the segments is reversed in the two enzymes (Fig. 1).

The problems that arise by deglycosylation or proteolysis complicate the purification of native enzymes. Gene cloning simplifies the isolation and characterization of the native polypeptides because the cloned genes can be expressed in *Escherichia coli* , which is devoid of other ß-1,4-glycanases(Beguin et al., 1987). This approach led to the characterization of four ß-1,4-glycanases from *C. fimi* : an exoglucanase (Cex) and endoglucanases A, B, and C (CenA, CenB and CenC) (Whittle, 1982, Gilkes, 1984, Wong, 1986, Owolabi,1986, O'Neill, 1986, Moser, 1989). The *cex, cenA, cenB* genes were first isolated as *E. coli* clones expressing polypeptides which reacted with an antiserum to supernatant proteins from a *C. fimi* culture grown in the presence of cellulose. The *cenC* gene was isolated by taking advantage of the capacity of CenC to bind to Sephadex (Moser, 1989). Recently, the *cenD* gene has been isolated using a different approach (Meinke, 1992, personal communication). The *cenD* gene encodes endoglucanase D (CenD),which was initially identified as a cellulose-binding polypeptide from *C. fimi*, distinct from Cex, CenA, CenB and CenC (Meinke, 1992).

Figure 1. Overall structures of an exoglucanase (Cex) and an endoglucanase (CenA) from *C. fimi* . PT denotes a Pro-Thr box; the numbers refer to amino acid residues. (Kilburn et al., 1989)



1.4 Cellulase structural domains

In structural terms, the only difference between native Cex and CenA and the recombinant forms of the enzymes produced in *E. coli* is that the former are glycosylated. Glycosylation does not affect the substrate specificities of Cex and CenA; it has little effect on their catalytic activities; and it does not affect their stabilities to heat and pH (Langsford, 1987).

C. fimi secretes a serine protease that is active against the cellulases. Cleavage of Avicel-bound Cex and CenA releases catalytically active fragments from the Avicel, suggesting that Cex and CenA are organized into two independently functioning domains, a substrate-binding domain and a catalytic domain. In each case, the large fragment retains catalytic activity but does not bind to Avicel, whereas the smaller fragment is catalytically inactive but does bind to Avicel. The actual site of cleavage in both cases is at the carboxy terminus of the Pro-Thr box. It is quite clear that in each enzyme, the binding domain is separated from the catalytic domain by a Pro-Thr box. This box or linker is thought to provide spatial separation of the two domains (Gilkes et al., 1991). The reversed order of the catalytic and binding domains in the two enzymes, with the binding domain located at the N-terminus of CenA but at the C-terminus of Cex, suggests sequence shuffling occurred during the evolution of the genes encoding them (Warren et al., 1986).

Small-angle x-ray scattering analysis shows that CenA is a tadpole-shaped molecule (Pilz et al., 1990). Cex has a similar tertiary structure (Schmuck and Gilkes, unpublished results). The CBD and linker form an extended tail region. The CBD is stabilized by a disulfide bond between cysteines near each end (Ong, 1992).

CenB, CenC, and CenD have the same type of CBD as Cex and CenA but different catalytic domains. (Meinke et al., 1991, Coutinho et al., 1992, Meinke et al., 1992).

1.5 Cellulose binding

The mechanism of binding of a CBD to cellulose is poorly understood (Ong, 1992). A CBD retains its ability to bind to cellulose when separated from the rest of the protein. CBDs vary from 36 to 240 amino acids in length and can be classified into 5 different families based on sequence similarities (Coutinho et al., 1992). Similar sequences are found in many cellulases and xylanases (Gilkes et al.,1991). The amino sequences of the *C. fimi* -type bacterial CBDs are highly conserved, with low numbers of charged amino acids, high contents of hydroxyamino acids, and conserved tryptophan, asparagine and glycine residues (Ong et al., 1989, Gilkes et al., 1991). The type II CBDs of Cex and CenA are 108 and 111 amino acids long, respectively, and their sequences share more than 50% identity. Two conserved cysteine residues, which participate in disulfide bond formation in CenA and Cex, are found at the N- and C-termini in all type II CBDs but one (endoglucanase I of *Butyrivibrio fibrisolvens*) (Ong, 1992). Strict conservation of tryptophans suggests that these residues participate in binding. Tryptophans are involved in the binding of a number of polypeptides to carbohydrates (Spurlino et al, 1992, Vyas, 1991). Site-directed mutations show the importance of two tryptophan residues of CBD_{CenA} in the adsorption to bacterial microcrystalline cellulose (BMCC) (Din, unpublished observations).

Extensive studies of the CBDs of cellulases from the bacterium *Cellulomonas fimi* have been made with molecular genetic techniques. The gene fragment encoding a CBD can be expressed in *E. coli* to generate the isolated

CBD polypeptide. The CBD can be coupled chemically to other agents, e. g. dyes, antibodies or enzymes, to facilitate binding these materials to cellulose. Alternatively, the CBD gene can be fused to the gene encoding another protein. Expression of the gene fusion gives a fusion protein that binds specifically to cellulose (Kilburn et al, 1992).

Linking a CBD to either the N terminus (using CBD_{CenA} or CBD_{CenC}) or the C terminus (using CBD_{Cex}) of a protein by molecular genetic techniques provides a convenient means to immobilize the protein on cellulose . Fusion of CBD_{Cex} to the C-terminus of a ß-glucosidase from an *Agrobacterium* (Abg) gives a hybrid, Abg-CBD_{Cex}, that has ß-glucosidase activity and binds to cellulose (Ong et al, 1991). The fusion protein is as active as native Abg and retains more than 40% of its activity when bound to cellulose. The CBD functions as an affinity tag that allows the simultaneous purification and immobilization of the enzyme on cellulose. Such applications have also been demonstrated using *Escherichia coli* alkaline phosphatase (PhoA) (Greenwood et al., 1992). The purified CBDs alone can also be coupled chemically to other molecules. For example Coomassie blue, a protein binding dye, binds to CBD; the conjugate can be used to dye cotton (Kilburn et al, 1992).

1.6 Avidin-biotin technology.

Avidin was recognized as a biological factor in egg white in the late 1920s during the discovery and isolation of the vitamin biotin (Kogl & Tonnis, 1936). Of particular interest is the remarkable strength of the interaction between avidin and biotin. The binding is characterized by a dissociation constant of the order of 10^{-15}M. This value corresponds to a free energy of association of about 21 kcal/mol, a staggeringly large value for the noncovalent interaction of a protein

with a molecule as small as biotin. In general, such strong binding is found only in systems involving liganded metal ions, either as partial covalent bonds or chelates. No metal ions, partial covalent bonds or chelates, are involved in the avidin-biotin interaction. Of special importance is the very slow off-rate that accompanies such a tight association (Richards, 1990). Another distinctive feature of this system is the multiple (four) binding sites of avidin for biotin (Wilchek and Bayer, 1990).

The rationale in using the avidin-biotin system is based on the premise that if one chemically modifies any biologically active compound with biotin through its valeric side chain, the biological and physiochemical properties of the biotin-modified molecule will not be changed significantly. If a reporter group of some sort is attached to the avidin molecule, the conjugate can be used for many different purposes (Figure 2). A variety of binders can be used to label a given target site; different probes can be either conjugated with avidin (for direct interaction with the biotinylated binder) or derivatized with biotin (for complex formation with the underivatized avidin).

Some of the major advantages in using avidin-biotin technology are as follows:

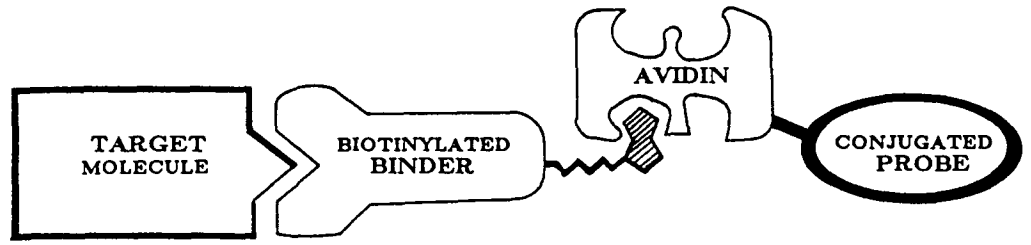
1. The exceptionally high affinity and stability of the avidin-biotin complex ensures the desired conjunction of binder and the probe.

2. Biotin can readily be attached to most binders and probes, and following biotinylation, the biological activity and physical characteristics are commonly retained.

3. The multiplicity of biotin groups per binder combined with the tetrameric structure of avidin leads to amplification of the desired signal.

4. The system is amenable to double-label and kinetics studies.

Figure 2. General scheme illustrating the essentials of avidin-biotin technology (adapted from Wilchek and Bayer, 1990).



TARGET	:	BINDER
Antigens	:	Antibodies
Antibodies	:	Antigens
Lectins	:	Glycoconjugates
Glycoconjugates	:	Lectins
Enzymes	:	Substrates, cofactors, inhibitors, etc.
Receptors	:	Hormones, effectors, toxins, etc.
Transport proteins	:	Vitamins, amino acids, sugars, etc.
Hydrophobic sites	:	Lipids, fatty acids
Membranes	:	Liposomes
Nucleic acids, genes	:	DNA/RNA probes
Phages, viruses, bacteria, subcellular organelles, cells, tissues, whole organisms	:	All of the above

AVIDIN-BIOTIN COMPLEX

PROBES
Enzymes
Radiolabels
Fluorescent agents
Chemiluminescent agents
Chromophores
Heavy metals
Colloidal gold
Ferritin
Hemocyanin
Phages
Macromolecular carriers
Liposomes
Solid supports

5. The system is extremely versatile. The versatility is further extended through the combined use of different biotinylated binders and avidin-associated probes.

6. A wide spectrum of different biotinylating reagents, biotinylated binders, and both biotinylated and avidin-containing probes is available from a variety of commercial sources.

The applications of the biotin-avidin system are numerous, from the use of avidin columns for the isolation of the target material to cytochemical localization studies. One of the most prevalent uses of avidin-biotin technology in recent years has been for immunoassays. It has also been applied to the localization of genes in chromosomes, and to other cytochemical tests.

1.7 Streptavidin from *Streptomyces avidinii.*

Streptavidin, a 60,000 dalton protein produced by *Streptomyces avidinii,* also forms a very strong and specific non-covalent complex with biotin (Chaiet et al, 1963, 1964). The protein consists of 4 identical subunits of approximate molecular weight 15,000, each of which binds a biotin molecule, and is free of carbohydrate. Avidin and streptavidin have rather different amino acids composition, but both have an unusually high content of threonine and tryptophan (Green, 1990). Avidin is a glycoprotein and its carbohydrate is characteristically heterogeneous (Bruch and White, 1982). In this respect, streptavidin, which is carbohydrate free, has proved advantageous. Its structure is now known (Weber et al, 1989, Hendrickson et al, 1989). The streptavidin gene has been cloned and sequenced (Aragana et al., 1986). The nucleotide sequence of the streptavidin gene and the deduced amino acid sequence is shown in figure 3. The deduced sequence is in good agreement with the amino acid content of the gel-purified protein, within the error of amino acid analysis. The streptavidin

monomer contains 159 amino acids compared with 128 in the avidin monomer. There are several regions of extensive homology in the two protein. Of particular interest is the homology around and including tryptophans 21, 79 and 120 of streptavidin. In avidin, the corresponding tryptophans (10, 70, 110) are protected from oxidizing agents by biotin, suggesting that these residues are implicated in the biotin-binding site of the protein (Green, 1975). A unique NH_2-group, probably one of the the three lysine residues (9, 71, 111) which are adjacent to the tryptophans, is also important for the biotin-binding activity of avidin (Green, 1975). In streptavidin, two of these three lysines (80 and 121)are conserved next to tryptophans. Both proteins show a clear structural similarity with a high preponderance of beta structure (Argarana et al., 1986). This suggests that functional and structural constraints have led to structure conservation during their evolution. It is reasonable to speculate that there is only a single type of structure which can create a binding site with such high affinity for biotin.

Vectors for the production of streptavidin-containing chimeric proteins use the T7 expression system (Sano and Cantor, 1990 and 1991). Fusions of streptavidin with a target protein can be obtained by inserting the coding sequence for a target protein into one of the unique cloning sites in a vector. Both the streptavidin and the target protein moieties are fully functional in the streptavidin-protein A and streptavidin-methallothionein chimeras (Sano et al, 1991 and 1992). Such chimeric proteins will extend the applications of the streptavidin system considerably.

Figure 3. Nucleotide sequence of the gene for streptavidin (Aragana et al., 1986). The amino acid sequence of streptavidin is shown above the nucleotide sequence. The amino acids of the signal peptide are indicated by negative numbers.

```
 1       5'  CCCTCCGTCCCCGCCGGGCAACAACTAGGGAGTATTTTTCGTGTCTCAC
                      -20                                  -10
    Met Arg Lys Ile Val Val Ala Ala Ile Ala Val Ser Leu Thr Thr
50  ATG CGC AAG ATC GTC GTT GCA GCC ATC GCC GTT TCC CTG ACC ACG
                                          1
    Val Ser Ile Thr Ala Ser Ala Ser Ala Asp Pro Ser Lys Asp Ser
95  GTC TCG ATT ACG GCC AGC GCT TCG GCA GAC CCC TCC AAG GAC TCG
                   10                                   20
    Lys Ala Gln Val Ser Ala Ala Glu Ala Gly Ile Thr Gly Thr Trp
140 AAG GCC CAG GTC TCG GCC GCC GAG GCC GGC ATC ACC GGC ACC TGG
                                       30
    Tyr Asn Gln Leu Gly Ser Thr Phe Ile Val Thr Ala Gly Ala Asp
185 TAC AAC CAG CTC GGC TCG ACC TTC ATC GTG ACC GCG GGC GCC GAC
                   40                                   50
    Gly Ala Leu Thr Gly Thr Tyr Glu Ser Ala Val Gly Asn Ala Glu
230 GGC GCC CTG ACC GGA ACC TAC GAG TCG GCC GTC GGC AAC GCC GAG
                                       60
    Ser Arg Tyr Val Leu Thr Gly Arg Tyr Asp Ser Ala Pro Ala Thr
275 AGC CGC TAC GTC CTG ACC GGT CGT TAC GAC AGC GCC CCG GCC ACC
                   70                                   80
    Asp Gly Ser Gly Thr Ala Leu Gly Trp Thr Val Ala Trp Lys Asn
320 GAC GGC AGC GGC ACC GCC CTC GGT TGG ACG GTG GCC TGG AAG AAT
                                       90
    Asn Tyr Arg Asn Ala His Ser Ala Thr Thr Trp Ser Gly Gln Tyr
365 AAC TAC CGC AAC GCC CAC TCC GCG ACC ACG TGG AGC GGC CAG TAC
                  100                                  110
    Val Gly Gly Ala Glu Ala Arg Ile Asn Thr Gln Trp Leu Leu Thr
410 GTC GGC GGC GCC GAG GCG AGG ATC AAC ACC CAG TGG CTG CTG ACC
                                      120
    Ser Gly Thr Thr Glu Ala Asn Ala Trp Lys Ser Thr Leu Val Gly
455 TCC GGC ACC ACC GAG GCC AAC GCC TGG AAG TCC ACG CTG GTC GGC
                  130                                  140
    His Asp Thr Phe Thr Lys Val Lys Pro Ser Ala Ala Ser Ile Asp
500 CAC GAC ACC TTC ACC AAG GTG AAG CCG TCC GCC GCC TCC ATC GAC
                                      150
    Ala Ala Lys Lys Ala Gly Val Asn Asn Gly Asn Pro Leu Asp Ala
545 GCG GCG AAG AAG GCC GGC GTC AAC AAC GGC AAC CCG CTC GAC GCC

    Val Gln Gln Stop
590 GTT CAG CAG TAG TCGCGTCCCGGCACCGGCGGGTGCCGGGACCTCGGCC  3'
```

1.8 Objectives of this thesis.

The general objective of this thesis was to produce a fusion protein comprising streptavidin and a cellulose-binding domain and to test it for the ability to bind biotinylated proteins to cellulose. Such a fusion protein would have many potential applications in protein purification and in enzyme immobilization.

2. MATERIALS AND METHODS

2.1 Bacterial strains and plasmids.

Escherichia coli strains JM101 (Yannish-Perron et al, 1985) and BL21(DE3)pLysS (Studier, 1990), and plasmids pTSA-18F (Sano and Cantor, 1991) and pTZEO4 (Ong, 1992) were described previously.

2.2 Media.

The medium used was LB (Maniatis et al, 1982) supplemented with 100 μg ampicillin mL^{-1} or with 25 μg chloramphenicol mL^{-1}. IPTG, final concentration 0.4 mM, was used to induce the production of fusion protein. All *E. coli* strains were grown at 37°C unless stated otherwise. Solid medium contained 15 g agar (Difco) L^{-1}.

2.3 Enzymes and Chemicals.

Restriction endonucleases, T4 DNA ligase, the Klenow fragment of DNA polymerase I (kPolI) and their buffers were from either Boehringer Mannheim (Quebec, Canada), Bethesda Research Laboratories (Ontario, Canada) or Pharmacia (Quebec, Canada). Restrictrion and modifying enzymes were used according to the suppliers' recommended procedures. DNA fragments were separated by electrophoresis through a 1.0% agarose gel using Tris -Acetate-EDTA (TAE) buffer (Maniatis et al. 1982). The desired DNA fragments were recovered from the TAE agarose gels with the Gene Clean kit (Bio/Can Scientific Inc., Toronto, Canada). The Gene Clean kit was used according to the supplier's instructions.

Standard streptavidin and biotinylated alkaline phosphatase were purchased from Sigma. Protein molecular weight standards were from BDH or Biorad. All other chemicals used were of analytical grade.

2.4 Isolation of plasmid DNA.

Plasmid DNA was isolated from *E. coli* using either the alkaline lysis procedure (Maniatis et al.,1982) or the CTAB-DNA precipitation method (Del Sal et al, 1989).

2.5 Construction of the gene fusion.

The gene fusion was constructed by standard cloning techniques (Maniatis et al., 1982). pTSA-18F was isolated and digested to completion with BamHI and HindIII. A DNA fragment of 2.7 Kb was purified as described in 2.3 above. This fragment carried most of the coding region for the mature streptavidin, under the ø10 promoter. pTZEO4 was isolated and digested to completion with BamHI and HindIII. A 0.7 Kb fragment, which had the coding region for the CBD, was purified as described in 2.3 above. The purified DNA fragments were ligated with T4 DNA ligase. The resulting plasmid was designated pTSAKL18.18.

2.6 Transformation and screening.

E. coli cells were transformed by the $CaCl_2$ method (Maniatis et al, 1982) or by electroporation with a Biorad gene pulser apparatus as instructed by the supplier. Transformants was selected initially on LB plates containing ampicillin and chloramphenicol. Positive clones containing the expected plasmid, identified by restriction endonuclease digestion, were tested for the fusion protein by Western blotting. A clone that reacted with anti-CBD_{cex} serum was chosen for further studies.

2.7 Detection of the fusion protein.

Cultures (500ml) were grown at 200 rpm and 37°C in a shaker water bath. Samples of 1mL were taken every hour for the measurement of absorbance at 600nm. IPTG was added at an O.D of ~1 to induce expression of the gene fusion. All samples were examined with a Zeiss phase-contrast microscope. Photographs of the cells were taken at a magnification of 1000 with Kodak T-max 400 film.The cells in the samples were collected by centrifugation, then lyzed with Triton X-100. The lysates were analyzed by SDS-PAGE.

2.8 Renaturation of the fusion protein.

All procedures were carried out at 4°C or on ice, unless stated otherwise . The cells were harvested by centrifugation at 3000 x g for 10 min, 4-5 hours after induction with IPTG. The cells were washed by resuspension in 100mM NaCl/1mM EDTA/10 mM Tris-Cl, pH8.0 and centrifugation as above. The cells were lysed by resuspension in 2mM EDTA/30 mM Tris-HCl, pH 8.0, containing 0.1% Triton X-100. The lysate was stored at -70°C until required.

The frozen lysate was thawed at room temperature. $MgSO_4$ (1.0M), DNase I (1mg/ml) and RNase A (1mg/ml) were added to final concentrations of 12mM, 10 μg/ml, 10 μg/ml, respectively, and the mixture was allowed to stand for 15-20 minutes at room temperature to reduce the vicosity. The suspension was centrifuged at 6000 x *g* for 15 minutes, and the pellet was washed three times with 10mM EDTA/30 mM Tris-HCl, pH8.0/1% Triton X-100. The pellet was dissolved in 6M guanidine hydrochloride, 10mM potassium phosphate pH 7.0, 1% ß-mercaptoethanol and dialyzed against the same solution. The solution was slowly diluted 100-fold in the renaturation buffer (10mM potassium phosphate pH7.0, 0.05% ß-mercaptoethanol) and stirred gently at 4°C for 16-24 hours to

renature the protein. The renaturation solution was then dialyzed against 40 volumes of 10mM potassium phosphate, 0.05% ß-mercaptoethanol to remove the guanidine. After several changes of the phosphate buffer, this solution was centrifuged at 39,000 x g for 20 minutes. The supernatant was collected and filtered through a 0.22µm low protein-binding filter. The filtrate contained the highly purified fusion protein.

2.9 Electrophoretic and Western blot analysis of the protein.

The purity of the renatured protein was assessed by SDS-PAGE, using a 12% gel (Laemmli, 1970). Gels either were stained with Coomassie blue or were used for western blotting. Native gels had the same composition except for the omission of SDS.

After electrophoresis,proteins in the gels were transferred onto a nitrocellulose membrane using a semi-dry electroblotter (Biorad semi-dry transfer cell). Rabbit anti-CBDcex serum and goat anti-rabbit IgG antibodies conjugated to alkaline phosphatase were used as the primary and secondary antibodies, respectively. The alkaline phosphatase activity was detected using 5-bromo-4-chloro-3-indolyl phosphate and p-nitroblue tetrazolium. The immunodetection was also done with the ECL (Enhanced Chemoluminescence) kit (Amersham), according to the supplier' s instructions.

2.10 Cellulose binding assay.

The renatured protein was mixed with 10mg of Avicel in 10mM potassium phosphate buffer at 4°C for 1hour. The Avicel was collected by centrifugation, then washed with 1M NaCl in 10mM potassium phosphate buffer, and then washed with 10mM potassium phosphate buffer. The washed Avicel was boiled in 1 X SDS loading buffer to release the bound protein. After

removing the Avicel by centrifugation, the protein in the supernatant was analysed by SDS-PAGE.

2.11 Biotin-binding assay.

Different amounts of the renatured protein were mixed with 10mg of Avicel in 10mM potassium phosphosphate pH 7.0, 30μg bovine serum albumin (BSA)mL^{-1} at 4°C for 1 hour. The Avicel was collected by centrifugation, then washed three times with 1M NaCl in 50mM Tris-HCl pH 7.4. The Avicel was collected by centrifugation, then washed three times with 50mM Tris-HCl pH 7.4, 30μg $BSA.mL^{-1}$ and resuspended in the same solution. Biotinylated alkaline phosphatase (~ 0.2 units) was added to the washed Avicel. The suspensions were mixed at 4°C for 1hour. After removing the Avicel by centrifugation, the supernatants were assayed for alkaline phosphatase activity, using the Sigma standard testing procedure for alkaline phosphatase. The Avicel pellets were washed with 1 M NaCl in 50mM Tris-HCl pH 7.4 and 50mM Tris-HCl pH 7.4, 30μg $BSA.mL^{-1}$, then assayed the alkaline phophatase activity.

To determine the number of biotin molecules bound per molecule of the renatured protein, biotin was added before the alkaline phosphatase step and the suspensions mixed at 4°C for 1 hour. After centrifugation, the Avicel was mixed with alkaline phosphatase at 4°C for 1hour. The alkaline phosphatase activity in the supernatants and the Avicel pellets was determined as described above.

2.12 Immobilization of biotinylated alkaline phosphatase on cellulose

Avicel or cellulose acetate was used as the support for the immoblization of the biotinylated enzymes.

Avicel was suspended in dH_2O,then washed several times with 10mM potassium phosphate buffer, pH 7.0. Approximately ~300 mg of washed Avicel

was mixed with sufficient STA-CBD, to saturate the Avicel as calculated from the previous biotin binding assays. The suspension was mixed at 4°C for 1 hour, then centrifuged. The Avicel was washed with 1 M NaCl in 50 mM Tris-HCl pH 7.4 buffer, then with 50 mM Tris-HCl pH 7.4, 30 μg BSA.mL^{-1}. Biotinylated alkaline phophatase was added. After mixing at 4°C for 1 hour, the Avicel was collected by centrifugation and washed with 1 M NaCl in 50 mM Tris-HCl pH 7.4 buffer and with 50 mM Tris-HCl pH 7.4, 30 μg BSA.mL^{-1}. A portion of the Avicel was assayed for alkaline phosphatase activity. A control was prepared by the same procedure, but omitting the STA-CBD fusion protein. The Avicel samples were packed into Biorad columns (12 x 12 mm). The columns were equilibratedwith 50 mM Tris-HCl pH 7.4. Substrate (15 mM p-nitrophenyl phosphate in the same buffer) was passed through the column at 0.1mL.h^{-1}. Samples of effluent were collected into tubes containing 3 M NaOH and the A_{405nm} measured.

A similar immobilization procedure was used with cellulose acetate membranes (0.2 x 25 mm) instead of Avicel. Five cellulose acetate membranes were saturated with STA-CBD, washed then incubated with biotinylated alkaline phosphatase as described above. The treated membranes were assembled into a Nalgene^{TM} affinity chromatography matrix holder. Substrate (15mM) was passed through the holder at 0.1mL.h^{-1}. Samples were collected and the A_{405nm} measured as described above.

3. RESULTS

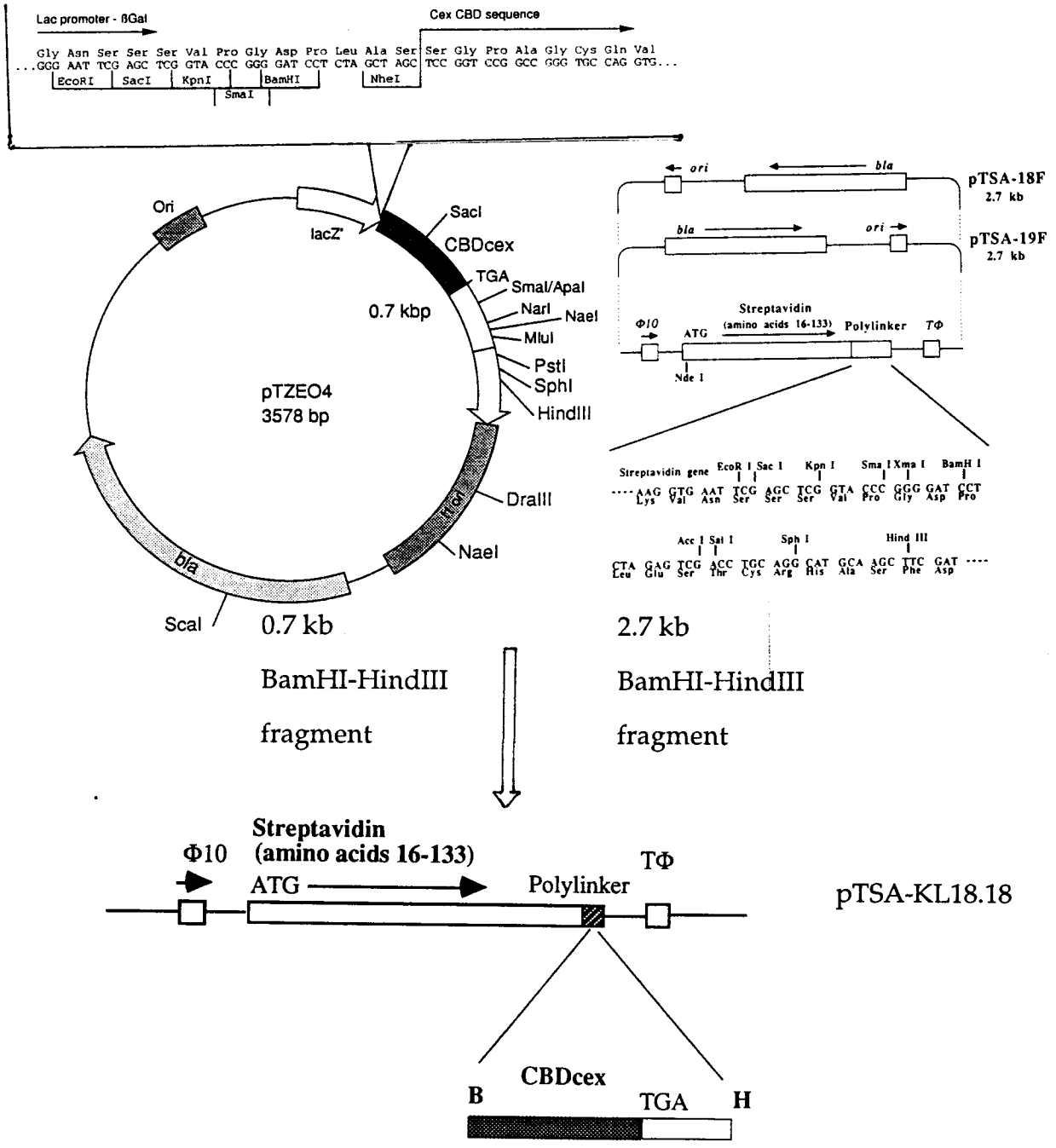
3.1 Construction of the gene fusion

The plasmid pTZEO4 contained the CBD_{Cex} coding sequence inserted downstream of the multiple cloning site of pTZ18R (Ong, 1992). It was used as a source of the CBDcex coding sequence. The mol. wt. of the CBD_{Cex} polypeptide, deduced from the DNA-sequence was 11,081 (O'Neill et al., 1986). The Mr of CBD_{Cex} as determined by SDS-PAGE was 11,000. The polypeptide reacted with anti-Cex serum (Ong, 1992).

The plasmid pTSA-18F comprised a DNA fragment carrying the T7 gene*10* promoter, the first 11 codons of the T7 gene *10* coding sequence, and the T7 transcription terminator ligated to a DNA fragment from pBR322 containing the replication origin and the *bla* gene, with the streptavidin coding sequence inserted between the T7 promoter and the T7 transcription terminator. The streptavidin coding sequence was also flanked by a translation initiation codon and a polylinker derived from pUC18, in which several unique cloning sites are found (Sano and Cantor, 1991). The predicted M_r of the encoded streptavidin was 13,400.

The 2.7 Kb BamHI-HindIII fragment from pTSA-18F was ligated to the 0.7 Kb BamHI-HindIII fragment from pTZEO4 to give pTSA-KL 18.18 (Fig. 4).

Figure 4 Diagram of the construction of the gene fusion encoding STA-CBDCex



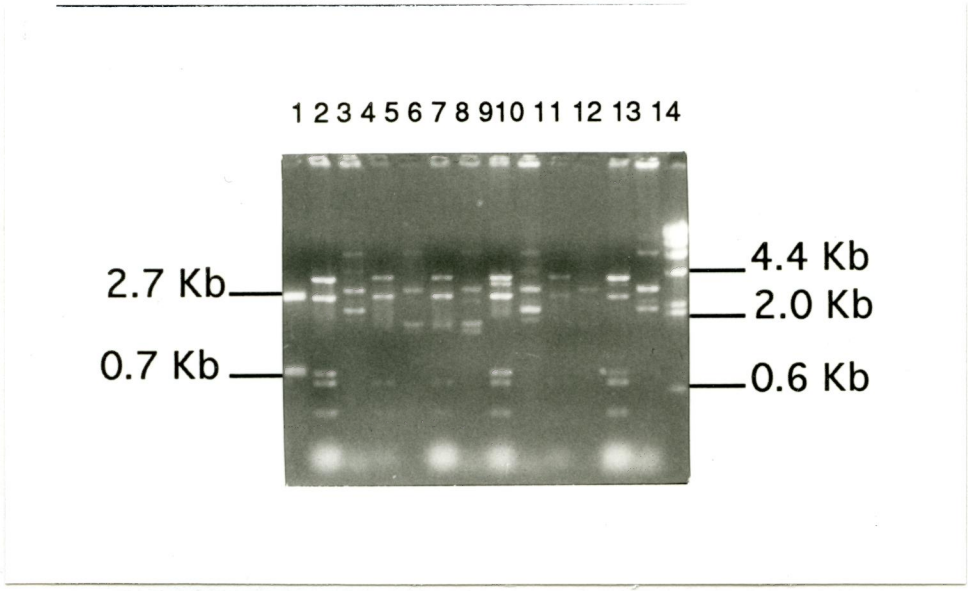
3.2 Screening

E. coli strain BL21(DE3)pLysS was transformed with pTSA-KL18.18. Lysogen BL21(DE3) carried the cloned T7 RNA polymerase gene in the chromosome under the control of the *lac* UV5 promoter. pLysS carried the T7 lysozyme gene to reduce the basal level of T7 RNA polymerase activity (Studier et al., 1990).

The transformants were selected initially on LB plus ampicillin and chloramphenicol. Restriction digests of the plasmids from ampicillin - and chloramphenicol - resistant clones showed clearly the presence of the CBD_{Cex}-encoding insert (~ 0.7 Kb) and the streptavidin - encoding fragment (~ 2.7Kb) (Fig. 5).

Clones containing such plasmids produced a polypeptide of M_r 25,000 which reacted with anti-Cex serum. The expected size of the streptavidin-CBD_{Cex} fusion protein (STA-CBD_{Cex}) was 25 kDa.

Figure 5 Analysis of plasmids in transformants. Plasmid DNA was isolated by the alkaline method, then digested with restriction endonucleases BamHI and HindIII. Lane 1, Streptavidin-containing vector (upper band, 2.7Kb) and CBD_{Cex}-containing sequence (lower band, 0.7Kb); lane 2, 4, 6, 8, 10, 12, plasmid DNA from transformants digested with BamHI and HindIII; lane 3, 5, 7, 9, 11, 13, undigested DNAs corresponding to lanes 2, 4, 6, 8, 10, 12. The positive clones contain the 2.7Kb and 0.7Kb. Lane 14, mol. wt. markers.



Control from BL21(DE3)pLysS. Lane 1, mol. wt.markers,; lane 2,3, DNA minipreps undigested and digested with BamHI and HindIII, respectively.

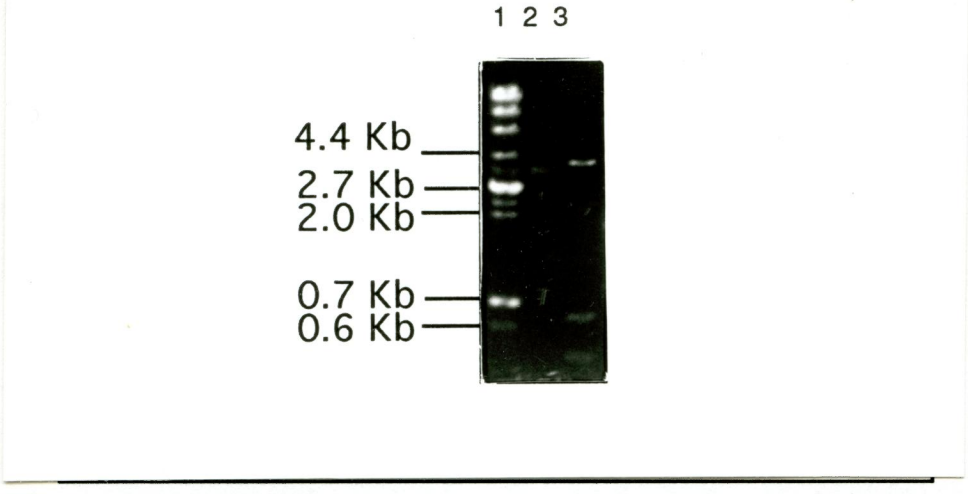
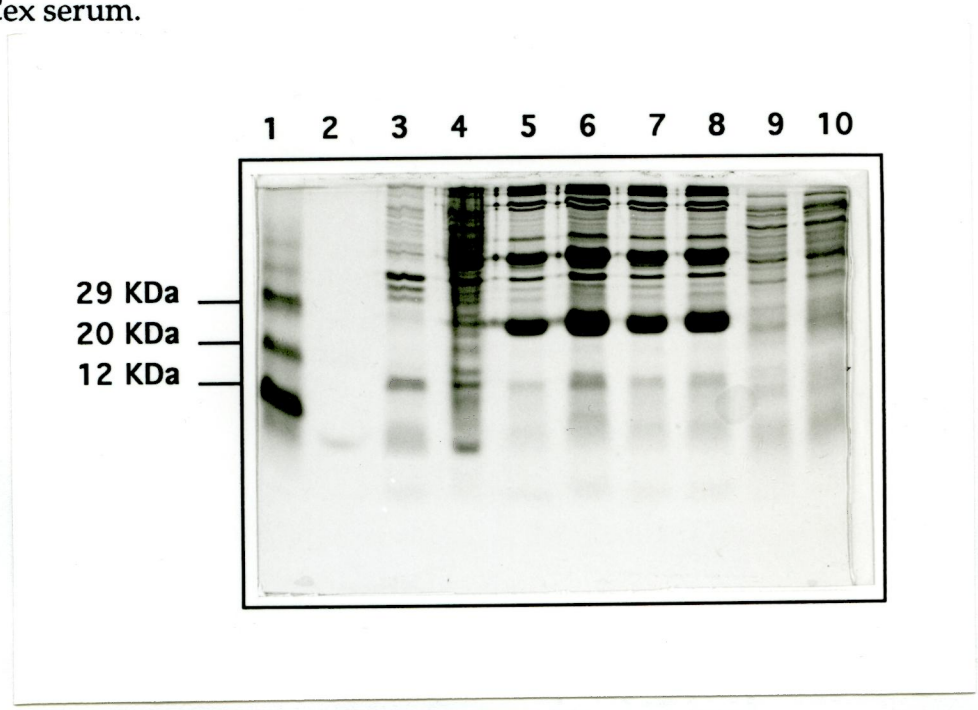
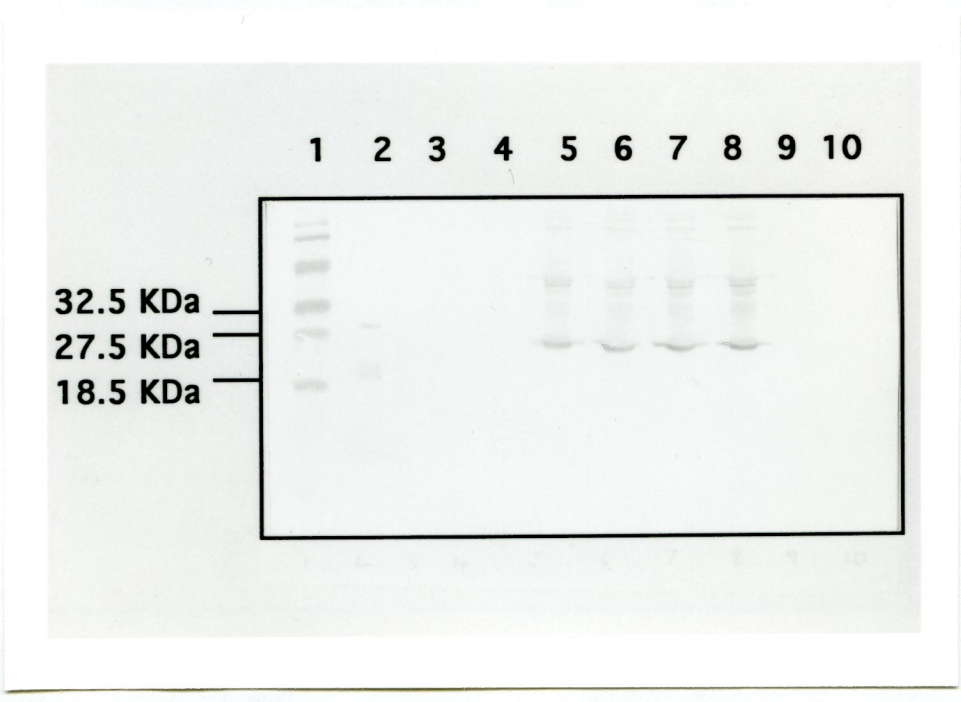


Figure 6 SDS-PAGE analysis and Western blotting of positive clones. A. Lane 1, mol. wt. markers; lane 2, CBDcex polypeptide; lane 3, lysate of the host cell without the fusion plasmid; lane 4, lysate of the cell containing the streptavidin vector pTSA-18F; lane 5, 6, 7, 8, pellets from the lysates of the positive clones pTSAKL18.10, 18.18, 18.31, and18.36, respectively; lane 9, 10, supernatants from the lysates of clones pTSAKL18.10 and 18.18, respectively. B. Western blot with anti-Cex serum.

A.



B.



3.3 Growth of BL21(DE3)pLysS/pTSA-KL18.18.

The presence of pTSA-KL18.18 and expression of the gene fusion did not have affect on the growth of the host cell BL21(DE3)pLysS (Fig. 7.a). Maximal expression of the fusion gene was seen 4-5 hours after induction by IPTG (Fig. 7.b). The cells had formed "snakes" and contained inclusion bodies 2-4 hours after induction (Fig. 8).

3.4 Renaturation of the fusion protein from the inclusion bodies.

The protein in the inclusion bodies was renatured as described as in Materials and Methods. The solubilized protein was pure enough (Fig. 9) to be used for the cellulose binding and biotin binding assays. The renaturation yield was ~ 4% of the protein in the cell lysate (lane 2, figure 9), estimated by densitometry of the gel (Bio-Scan). The overall yield of the renatured protein was estimated as ~4 mg per litre of culture.

The yield of renatured protein was much lower if, after solubilization in 6M guanidine HCl, the fusion protein was diluted rapidly rather than slowly with renaturation buffer (data not shown).

3.6 M_r of the fusion protein

The fusion protein had an M_r of approximately ~ 25,000 (figure 10). The molecular weight calculated from its amino acid sequence was 24,460. SDS-PAGE and Western blotting analysis did not reveal significant levels of degradation products.

Figure 7a Growth of BL21(DE3)pLysS with (¤) and without (o) pTSA-KL18.18. Cultures were induced with IPTG at O.D.600nm of ~1.0.

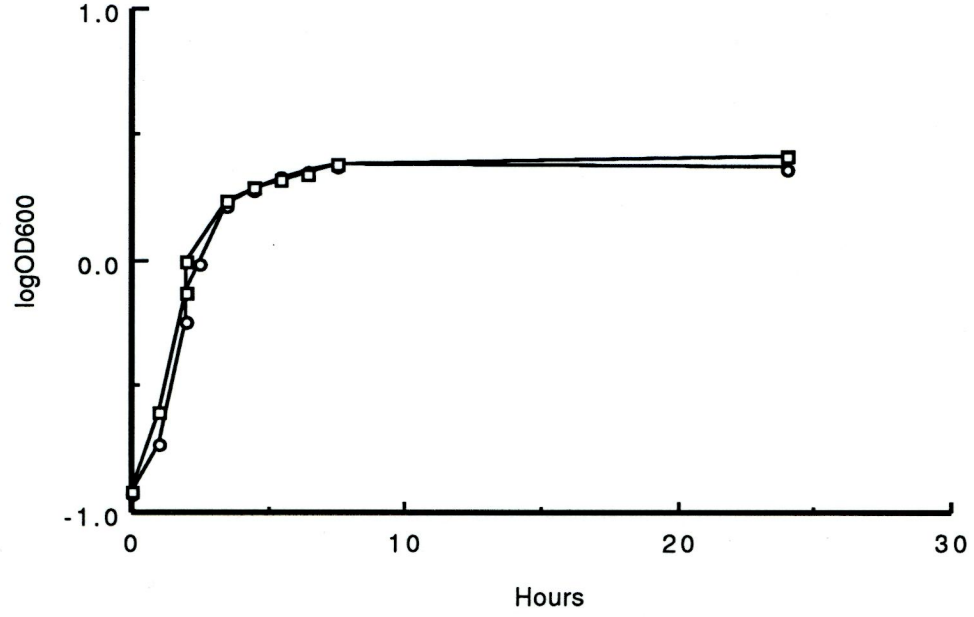


Figure 7b Induction of the gene fusion. The cells in the samples shown in Fig. 7a were analyzed by SDS-PAGE. Lane 1, mol.wt. markers; lane 2, before induction with IPTG ; lane 3-11, 1-9 hours after induction; lane 12, ~ 24 hours after induction. The culture was induced at O.D.600nm ~1.0.

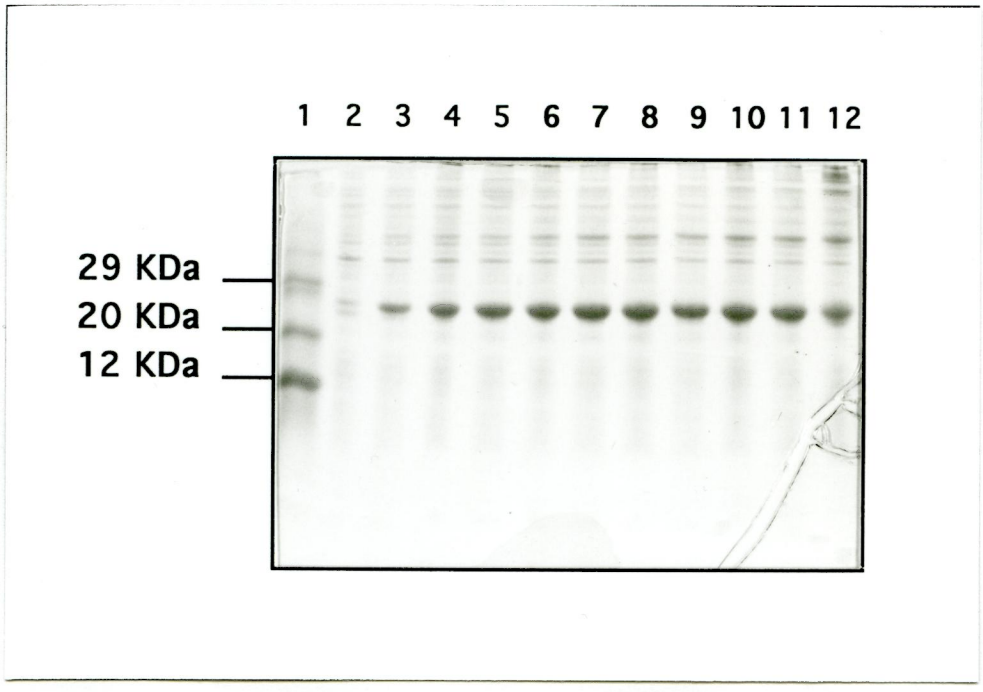
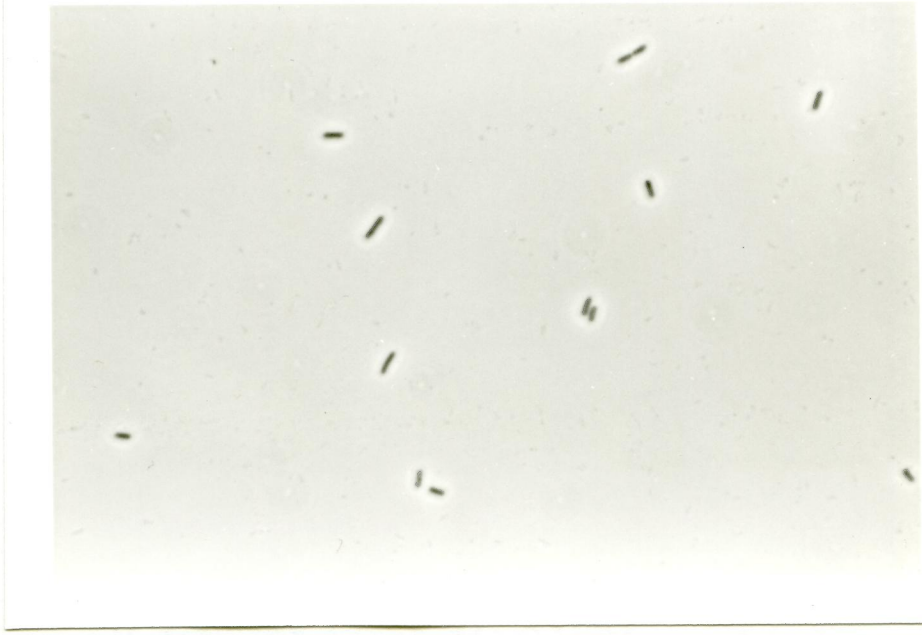


Figure 8 Formation of inclusion bodies by BL21(DE3)pLysS /pTSA-KL18.18

pTSAKL18.18 cells before induction with IPTG



2hours after induction

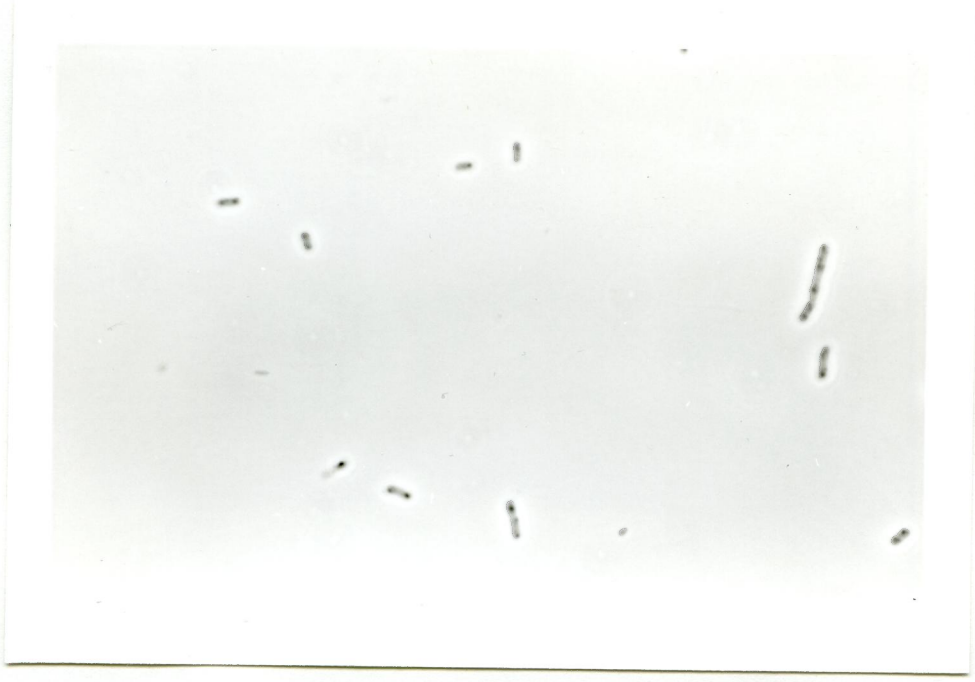
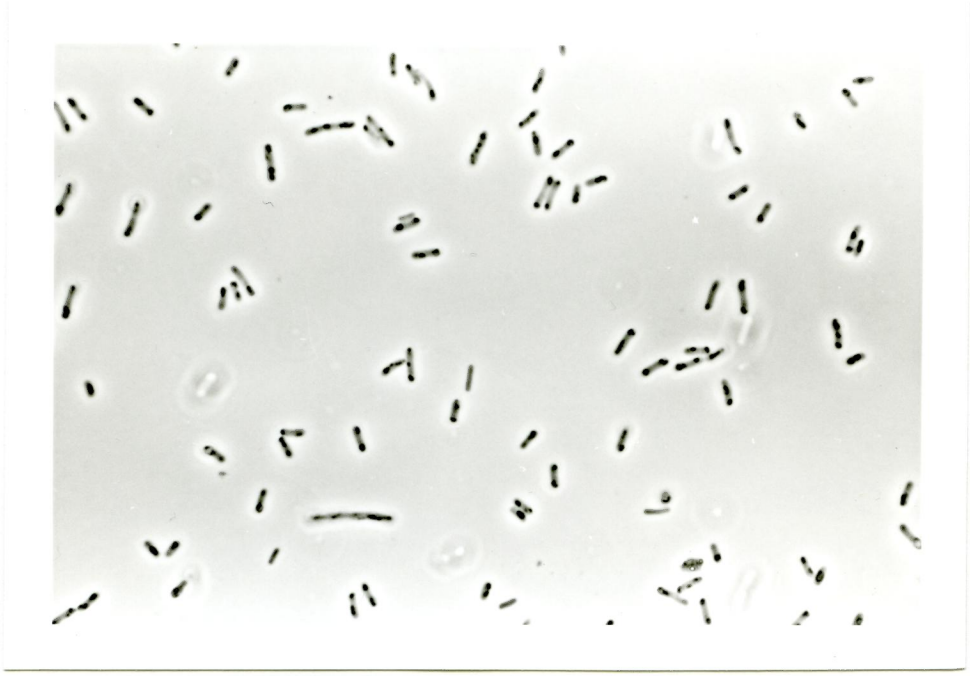


Figure 8 cont' d

4 hours after induction



BL21(DE3)pLysS cells 4 hours after the addition of IPTG

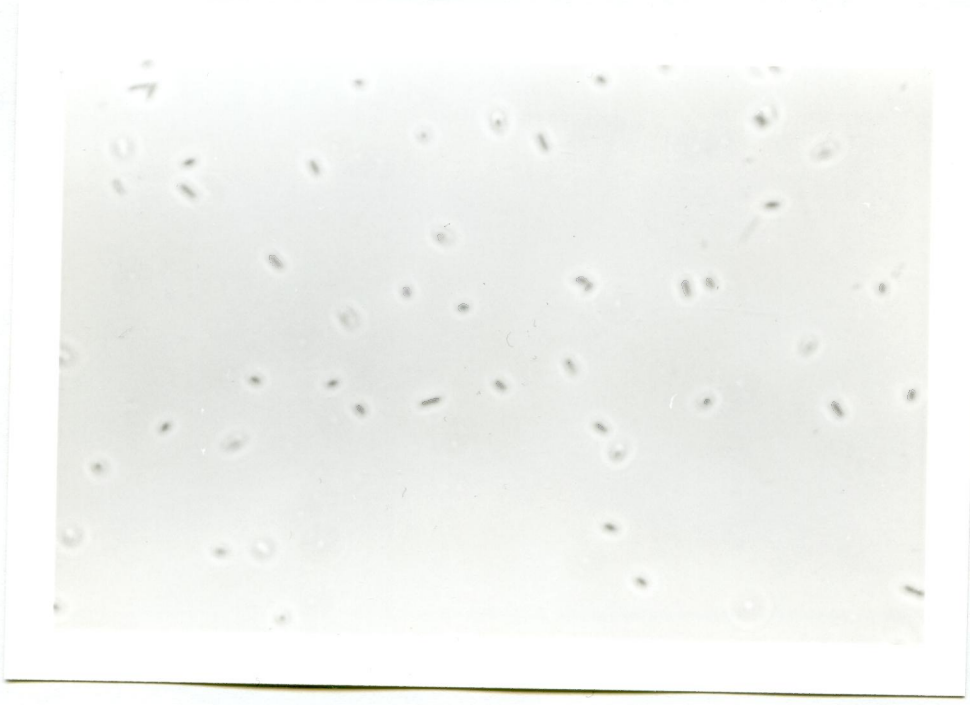
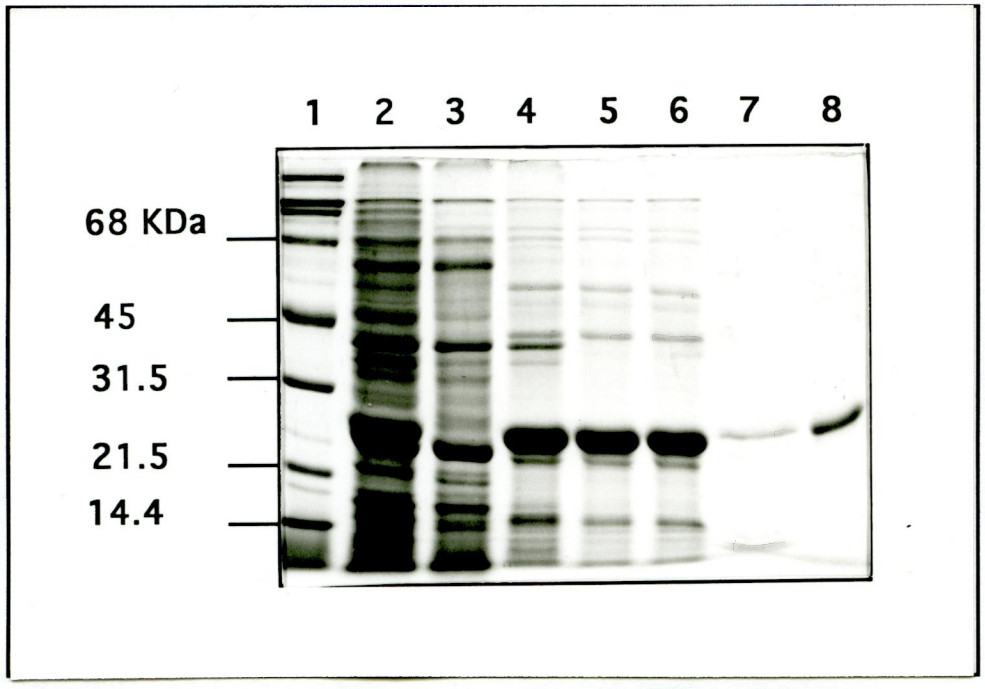


Figure 9 Renaturation of the fusion protein in inclusion bodies. Lane 1, mol. wt. markers; lane 2, cell lysate ; lanes 3 and 4, supernatant and pellet after centrifugation of the cell lysate ; lanes 5 and 6, pellet after one and two washes with Triton X-100; lane 7, soluble material after renaturation ; lane 8, insoluble material after renaturation. Experimental details are given in Materials and Methods. Lanes were loaded with material equivalent to equal volumes of the culture at harvest.



3.6 Binding of the renatured protein to cellulose.

The renatured protein bound to cellulose (Fig. 10). Unlike Cex and some fusion proteins containing CBD_{Cex} (Ong, 1982), the fusion protein could not be desorbed from the cellulose with water. It could be desorbed with 6M guanidine-HCl , 6M urea, or 2% SDS (data not shown).

3.7 Binding of biotin by the renatured protein

Biotinylated alkaline phosphatase was bound to the Avicel by the cellulose-bound fusion protein. The alkaline phosphatase activity remaining in solution decreased as the amount of the fusion protein mixed with Avicel was increased (Fig. 11). The biotinylated alkaline phosphatase did not bind to the Avicel in the absence of the fusion protein nor in the presence of streptavidin (data not shown). The addition of excess biotin before the biotinylated alkaline phosphatase blocked the binding of the latter to the cellulose-bound fusion polypeptide (Table 1).

Figure 10 Binding of the fusion protein to cellulose. The fusion protein was adsorbed to Avicel as described in Materials and Methods. The solution before the addition of Avicel (lane 3), the supernatant after adsorption to Avicel (lane 2) and the protein bound to Avicel (lane 4) were analyzed by SDS-PAGE. Lane 1, mol. wt. markers.

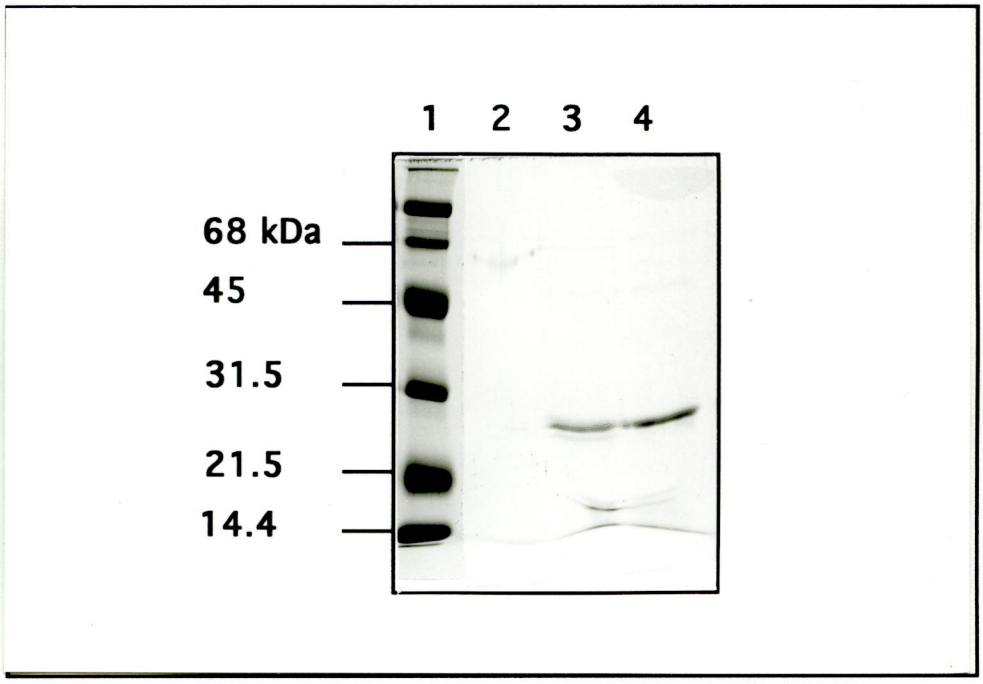


Figure 11 Binding of biotinylated alkaline phosphatase by the cellulose-bound fusion protein. Each assay tube contained 10mg Avicel. Increasing quantities of the fusion protein were added to the tubes, followed by equal amounts of biotinylated alkaline phosphatase. After incubation at 4°C for 1 hour, the Avicel was pelleted by centrifugation and the alkaline phosphatase activity measured in the supernatant (o). Then the alkaline phosphatase activity in the Avicel pellets (□) was measured as described in Materials and Methods.

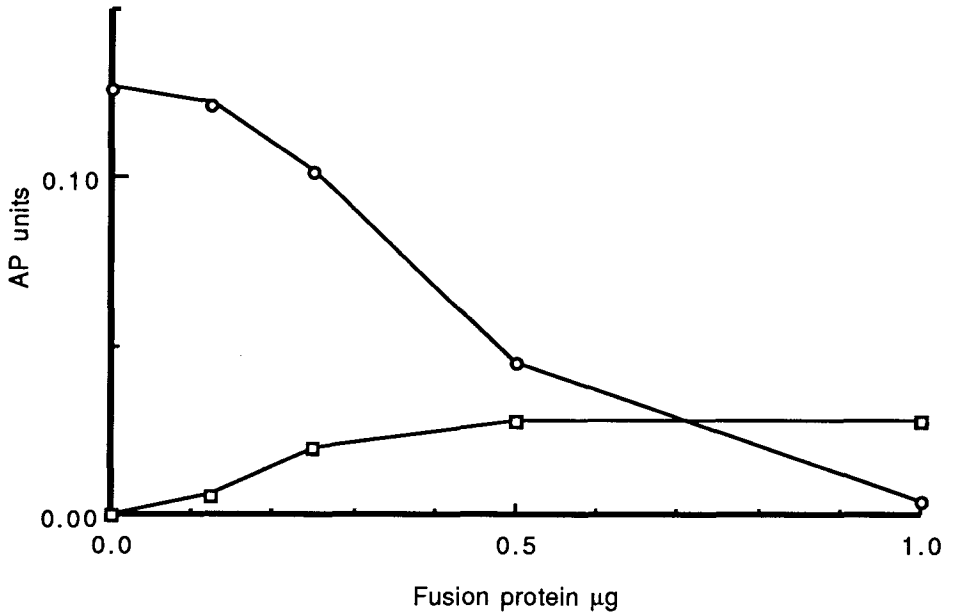


Table 1 Blocking by biotin of the binding of biotinylated alkaline phosphatase to the fusion protein.

Tube #	Avicel (mg)	nmol of STA in STA-CBD	nmol of biotin added	APunits in SN (%)	AP units in P (%)
1	10	0.000	0.000	100	0
2	10	0.016	0.004	37	>24 *
3	10	0.016	0.008	75	>24 *
4	10	0.016	0.016	89	9
5	10	0.016	0.032	99	0.4
6	10	0.016	0.064	99	0.4
7	10	0.016	0.128	99	0.4

* Alkaline phosphatase activity in pellet was at the saturation level as observed with 10 mg of Avicel in the assays for the biotin binding activity (see figure 11)

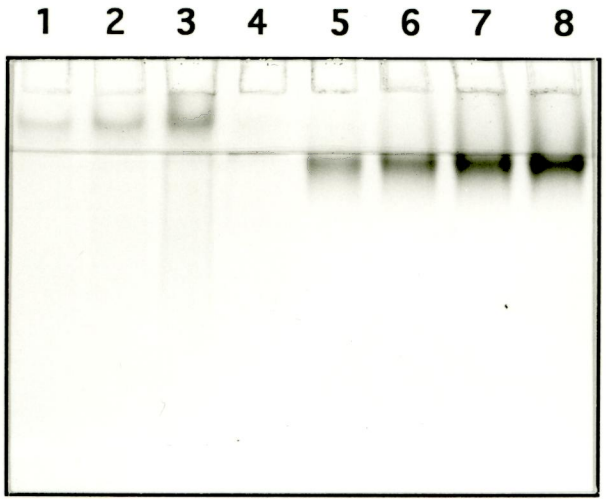
SN : supernatant

P : pellet

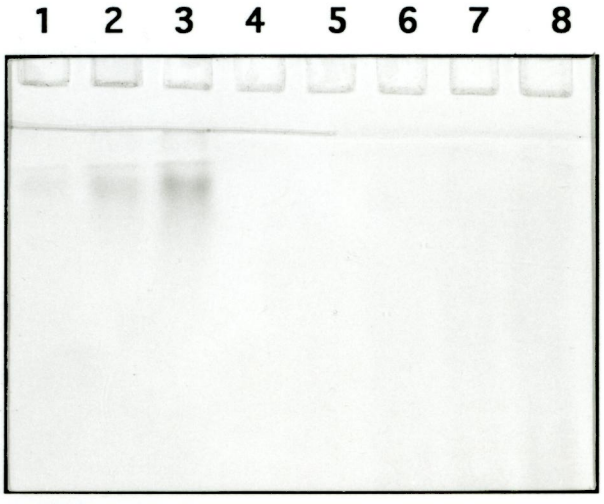
It was possible that the fusion polypeptide aggregated after renaturation. It might then sediment under the conditions used to recover the Avicel in the binding assays. If so, the apparent binding of the biotinylated alkaline phosphatase to the fusion protein-Avicel complex would be artefactual, simply the consequence of sedimentation with the Avicel. However, the fusion protein entered a 12% polyacrylamide gel during electrophoresis under non-denaturing conditions (Fig. 12), showing that it was not aggregated.

Figure 12 Non-denaturing PAGE analysis of the renatured protein. Lane 1, 2, 3, 0.25 μg, 0.5 μg, 1 μg of the renatured protein, respectively; lane 5, 6, 7, 8, 0.5 μg, 1μg, 1.5 μg, 2 μg of streptavidin, respectively.

after 1.5 hours at 100v.



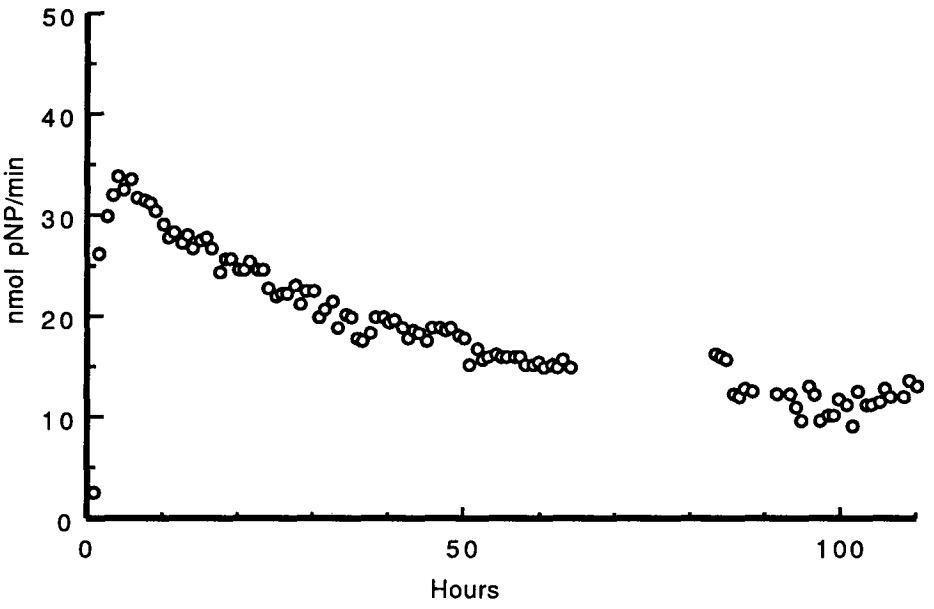
after ~6 hours at 100v.



3.8 Immobilization of biotinylated alkaline phosphatase on cellulose via the fusion protein.

A ß-glucosidase-CBD_{Cex} fusion protein can be used as an immobilized enzyme when bound to an appropriate cellulose matrix (Ong et al., 1989a). The affinity of avidin for biotin is several orders of magnitude greater than the affinity of CBD_{Cex} for cellulose. Thus, it was very likely that the STA-CBD_{Cex} could be used to immobilize biotinylated alkaline phosphatase on cellulose, and, in fact, the phosphatase retained more than 50% of its activity after continuous operation at room temperature for 5 days (Fig. 13).

Figure 13 Immobilization of biotinylated alkaline phosphatase. STA-CBD_{Cex} was adsorbed to cellulose acetate membranes, then biotinylated alkaline phosphatase was aded to the washed membranes. The membranes were assembled into a column, and pNPP (15mM) was passed through the column (0.1 mL h^{-1}) at room temperature. The A_{405nm} of the effluent was measued after adding 1/2 volume of 3M NaOH.



4. DISCUSSION

The streptavidin-CBD_{Cex} fusion protein can be produced in high levels in *E. coli.* Like streptavidin itself, it forms inclusion bodies. However, the recovery of active renatured fusion protein from inclusion bodies is much more difficult than the recovery of streptavidin and and other streptavidin-containing fusion proteins.

This may reflect the properties of the CBD rather than of streptavidin. Fusion proteins comprising factor X and CBD_{Cex} or CBD_{CenA} and interleukin-2 also form inclusion bodies which are very difficult to renature (Shen and Greenwood, personal communications).

Several hypotheses have been proposed to explained the insolubility, including the accumulation of partially folded forms that will aggregate by hydrophobic interactions, the induction of the heat shock response, which may create an aberrant environment for protein folding, and in the case of the cytosolic expression of membrane-associated proteins, the absence of a maturation pathway. These hypotheses for insolubility are not mutually exclusive and could occur in parallel (Frankel et al., 1991).

Slow rather than fast dilution of the unfolding agent appears to be the critical step in renaturation. If the rate of dilution of the denatured protein from 6 M guanidine-HCl is less than or equal to its rate of renaturation, the actual concentration of denatured protein in solution can be maintain at low levels (Fischer et al., 1992). In this way the re-aggregation during renaturation caused by hydrophobic interactions of partially refolded material is minimised. Slow dilution enhanced the recovery of lysozyme from inclusion bodies (Fischer et al., 1992). It also enhanced the recovery of STA-CBD_{Cex}.

The yields of the renatured proteins do not appear to depend on the sizes of the proteins or the expression levels. The yields most likely reflect the intrinsic properties of proteins with regard to their stability and ability to refold during renaturation under the conditions used (Lin and Cheng, 1991). The recovery of active STA-CBD_{Cex} was estimated to be about 4%, comparable to the yields of 4- 30% obtained for inclusion bodies of other proteins (Lin and Cheng, 1991).

The ultimate goal of the construction of STA-CBD_{Cex} is the binding of the biotinylated proteins to cellulose. Biotinylation of proteins is a relatively simple process which usually does not affect the biological activities of the proteins. It could offer a more convenient means for the purification and/or immobilization of proteins for which it is difficult to obtain fusion derivatives by genetic engineering.

Biotinylated alkaline phosphatase was chosen as a test protein because it is relatively stable, easily assayed and commercially available. Biotinylated alkaline phosphatase could be bound to cellulose by STA-CBD_{Cex}. The binding was specific because it could be blocked by biotin. The binding was stable during operation of the bound alkaline phosphatase as an immobilized enzyme. Clearly, STA-CBD_{Cex} possessed the expected properties. It has great potential for the immobilization and recovery of biotinylated proteins.

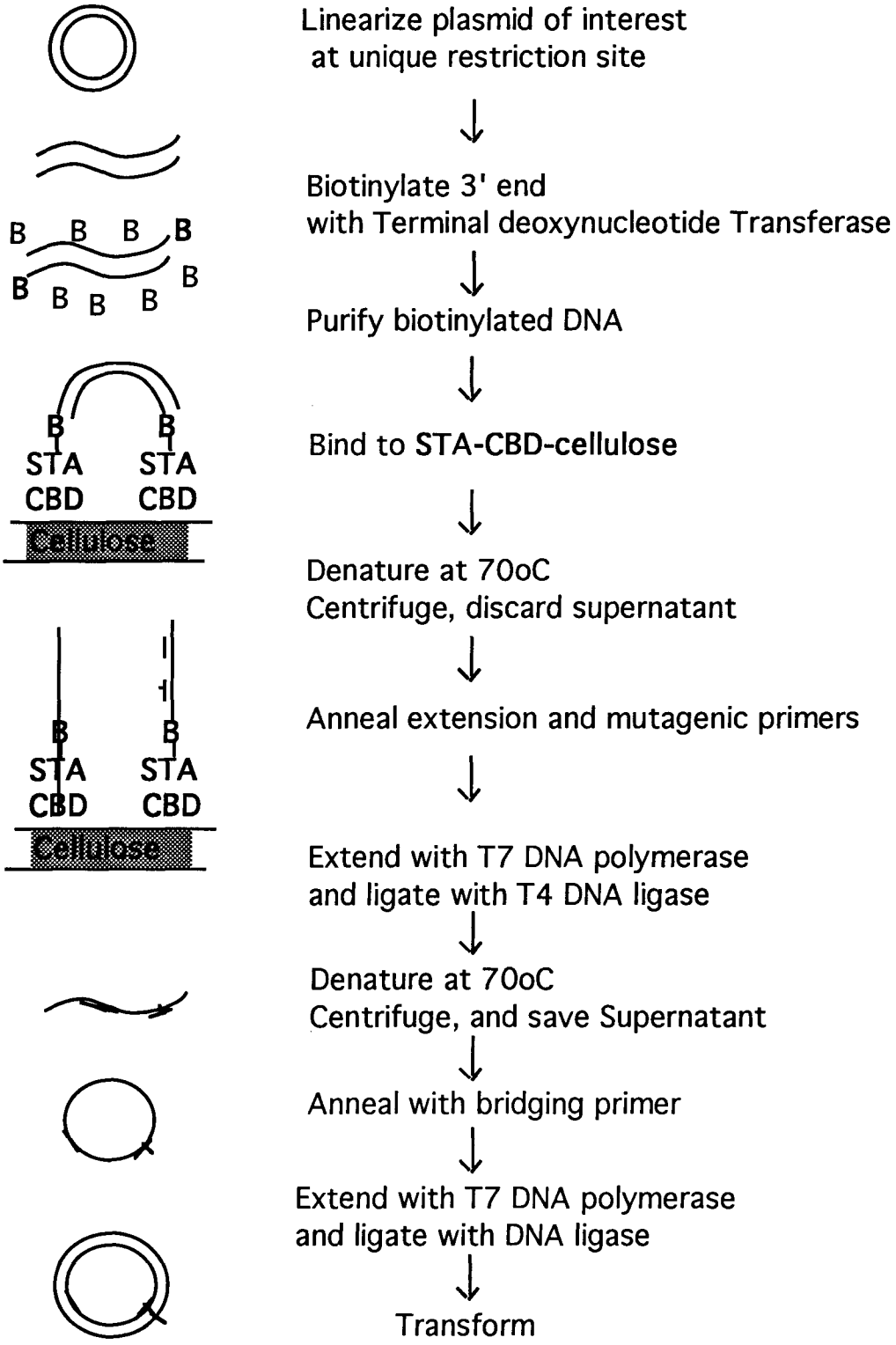
A major application of the biotin-avidin technology is in gene probing. Today, most of the work in this area deals with probe design and labeling procedures. Biotin can be introduced into DNA chain through the use of a biotinylated nucleoside triphosphate during nick-translation. Alternatively, reactive biotin-containing derivatives can be used for the direct labelling of DNA. Synthetic oligonucleotides (20 to 30 bases in length), to which biotin residues have been introduced either on the 3' or 5' terminus, have also been used

successfully applied(Wilchek and Bayer, 1990). Many biotinylated DNA products are available commercially.

STA-CBD_{Cex} may have applications in DNA technology. Any biotinylated DNA sequence should bind to a STA-CBD_{Cex}-cellulose complex. Any protein, e. g. transcription factors, that interacts with a specific DNA sequence could then be isolated through its interaction with the immobilized DNA.

An example of using this fusion protein in DNA technology can be suggested with the techniques of DNA doubled-stranded site-directed mutagenesis as illustrated in figure 14.

Figure 14 Illustration of the use of STA-CBD_{Cex} in site-directed mutagenesis of DNA



6. REFERENCES

Aragana, C. E., Kunts, I. D., Birken, S., Axel, P. and Cantor C. R. 1986. Molecular cloning and nucleotide sequence of streptavidin. Nucleic Acids Res. 14: 1871-1882.

Bayer, E. A., Benhur, Y., Hiller, Y., Wilchek, M. 1989. Postsecretory modifications of streptavidin. Biochem. J., 259: 369-376.

Beguin P., Gilkes, N.R., Kilburn, D.G., Miller, R.C., Jr., O'Neill, G. P., and Warren, R.A.J., 1987. Cloning of cellulase gene. CRC Critical Reviews in Biotechnol. 6: 129-162.

Beguin, P., Eisen H., and Rouspas, A. 1977. Free and cellulose-bound cellulases in a *Cellulomonas* species. J. Gen. Microbiol. 101: 191-196.

Beguin, P. and Eisen H. 1978. Purification and partial characterization of three extracellular cellulases from *Cellulomonas* species. Eur. J. Biochem. 87: 525-531.

Bruch, R. C. and White, H. B. 1982. Compositional and structural homogeneity of avidin polypeptides. Biochemistry, 21: 5334-5341.

Chaiet, L., and Wolf, F. J. 1964. The properties of streptavidin, a biotin-binding protein produced by *Streptomyces*. Arch Biochem. Biophys., 106: 1-5.

Chaiet, L., Miller, T. W., Tausing, F., and Wolf, F. J. 1963. Antibiotic MSD-235 II. Separation and purification of synergistic components. Antimicrobial Agents Chemother. 3: 28-32.

Coughlan, M. P. 1985. The properties of fungal and bacterial cellulases with comment on their production and application. Biotechnology and Genetic Engineering Reviews. 3: 39-109.

Coutinho, J. B. 1992. Ph. D. Thesis, University of British Columbia.

Coutinho, J.B., Gilkes, N.R., Warren, R.A.J., Kilburn, D.G. and Miller, R.C., Jr. 1992.The binding of *Cellulomonas fimi* endoglucanase C (CenC) to cellulose and Sephadex is mediated by the N-terminal repeats. Molec. Microbiol., 6: 1243-1252.

De Sal, G., Manfiole, H., and Schneider, C. 1989. The CTAB-DNA precipitation method : a common miniscale preparation of template DNA from phagemids, phages or plasmids suitable for sequencing. Biotechniques, 7: 514-517.

Fischer, B., Perry, B., Summer, I., and Goodenough, P. 1992. A novel sequential procedure to enhance the renaturation of recombinant protein from *Escherichia coli* inclusion bodies. Protein Engineering, 6: 593-596.

Frankel, S., Sonh, R., and Leinward, L. 1991. The use of sarkosyl in generating soluble protein after bacterial expression. Proc. Natl. Acad. Sci. USA., 88: 1192-1196.

Gilkes, N.R., Henrissat B., Kilburn, D.G., Miller, R.C., Jr and Warren, R.A.J. 1991. Domains in microbial ß-1, 4-glucanases : sequence, conservation, function, and enzyme families. Microbial. Rev., 55: 303-315.

Gilkes, N.R., Kilburn, D.G., Langsford, M.L., Miller, R.C., Jr., Wakarchuk, W.W., Warren, R.A.J., Whittle, D.J. and W.K.R. Wong, Isolation and characterization of *Escherichia coli* clones expressing cellulase genes from *Cellulomonas fimi*. J. Gen. Microbiol., 1984. 130: 1377-1384.

Gilkes, N.R., Langsford, M.L., Kilburn, D.G., Miller, R.C., Jr. and Warren, R.A.J., Mode of action and substrate specifities of cellulases from cloned bacterial genes. J. Biol. Chem., 1984. 259: 10455-10459.

Gilkes, N.R., Meinke A., Coutinho, J. B., Ong, E., Greenwood, J. M., Miller, R.C., Jr, Kilburn, D.G., and Warren, R.A.J. 1991. *Cellulomonas fimi* ß, 1, 4-glucanases. ACS Symposium series 460, 349-360.

Green, N. M. 1975. Avidin. Advances in Protein Chemistry. 29: 85-113.

Greenwood, J. M., Ong, E., Gilkes, N.R., Miller, R.C., Jr., Warren R. A. J. and Kilburn, D.G. 1992. Cellulose-binding domains : potential for purification of complex proteins. Protein Engineering, 5 : 361-365.

Hayn, M. and Esterbauer, H. 1985. Separation and partial chatacterization of *Trichoderma reesei* cellulase by fast chromatofucusing. J. Chromatogr., 327: 387-397.

Hendrickson, W. A., Pahler, A., Smith, J. L., Satow, Y., Merritte, E. A. and Phizackerly, R. P. 1989. Crystal structure of core streptavidin determined from multiwavelength anomalous diffraction of synchroton radiation. Proc. Natl. Acad. Sci. USA., 86: 2190-2194.

Inouye, M., Arnheim N. and Sternglanz, R. 1973. Bacteriophage T7-lysozyme is an N-acetylmuramyl-L-alanine amidase. J. Biol. Chem. 248: 7247-7252.

Kilburn, D.G., Gilkes, N.R., Miller, R.C., Jr and Warren, R.A.J. 1989. Cellulases of *Cellulomonas fimi*. The enzymes and their interactions with substrate. ACS Symposium series 349, 587-596.

Kilburn, D.G., Turner, R. F. B., Coutinho, J. B., Din, N., Gilkes, N.R., Greenwood, J. M., Hobbs, J. B., Miller, R.C., Jr, Ong, E., Phelps M. R., Ramirez, C. and Warren, R.A.J. 1992. Cellulose-biding domains: applications in biotechnology. Cellucon ' 92. In press.

Klausmeister, W. H. 1986. Economic analysis of an innovative process for cellulose acetate production. In R.A. young and R. M. Rowell, eds. Cellulose: structure, modification and hydrolysis, Wiley, N. Y., pp. 187-201.

Klyosov, A. A., 1990. Trends in biochemistry and enzymology of cellulose degradation. Biochemistry, 29 : 10577-10585.

Knowles, J., Lehtovaara, P., and Teeri, T. 1987. Cellulase families and their genes. Trends Biotechnol., 5: 255-261.

Kogl, F., and Tonnis, Z. 1936. Description of crystallized biotin from yolk. Z. Physiol. Chem., 242: 43.

Laemmli, K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature, 227: 680-685.

Langsford, M.L., Gilkes, N.R., Singh, B., Moser, B., Miller, R.C., Jr., Warren, R.A.J. and Kilburn, D.G. 1987. Glycosylation of bacterial cellulases prevents proteolytic cleavage between functional domains. FEBS Lett., 225: 163-167.

Langsford, M.L., Gilkes, N.R., Wakarchuk, W.W., Kilburn, D.G., Miller, R.C., Jr. and R.A.J. Warren. 1984. The cellulase system of *Cellulomonas fimi*. J. Gen. Microbiol., 130: 1367-1376.

Lin, K. and Cheng, S. 1991. An efficient method to purify active eukaryotic proteins from the inclusion bodies in *Escherichia coli*.. BioTechniques, 11: 748-751.

Maniatis, T., Fritsch, E. F. and Sambrook, J. 1982. Molecular Cloning, A Laboratory Manual. Cold Spring Harbor laboratory, Cold Spring Harbor, New York, p.68.

Marchessault, R. and Sarko, A.,1967. X-ray structure of polysaccharides. Adv. Carbohydr. Chem. 22: 421-482.

Marshall, K. and Sixsmith, D. 1974. Some physical characteristics of the microcrystalline cellulose I powders for pharmaceutical use. Drug Devel. Commun. 1: 51-71.

Meinke, A., Gilkes, N.R., Kilburn, D.G., Miller, R.C., Jr., and Warren, R.A.J. 1991. Multiple domains in endoglucanase B (CenB) from *Cellulomonas fimi*: functions and relatedness to domains in other polypeptides. J. Bacteriol., 173: 7126-7135.

Moffatt, B. A. and Studier, F. W. 1987. T7 lysozyme inhibits transcription by T7 RNA polymerase. Cell, 49: 221-227.

Moser, B., Gilkes, N.R., Kilburn, D.G., Warren, R.A.J., and Miller, R.C., Jr. 1989. Purification and characterization of endoglucanase C of *Cellulomonas fimi*, the cloning of its gene, and analysis of *in vivo* transcripts of the gene. Appl. Environ. Microbiol., 55 : 2480-2487.

O'Neill, G., Goh, S.H., Warren, R.A.J., Kilburn, D.G. and Miller, R.C., Jr.1986. Structure of the gene encoding the exoglucanase of *Cellulomonas fimi*. Gene, 44: 325-330.

Ong, E. 1992. Ph. D. Thesis, University of British Columbia.

Ong, E., Gilkes, N.R., Miller, R.C., Jr., Warren R.A.J. and Kilburn, D.G. 1991. Enzyme immobilization using a cellulose-binding domain: Properties of a ß-glucosidase fusion protein. Enzyme Microb. Technol., 13: 59-65.

Ong, E., Gilkes, N.R., Warren, R.A.J., Miller, R.C., Jr., and Kilburn, D.G. 1989a. Enzyme immobilization using the cellulose-binding domain of a *Cellulomonas fimi* exoglucanase. Bio/Technol., 7: 604-607.

Ong, E., Greenwood, J. M., Gilkes, N.R., Miller, R.C., Jr, , Warren, R.A.J and Kilburn, D.G.. 1989b. The cellulose binding domains of cellulases : tools for biotechnology. Trends Biotechnol. 7: 239-243.

Ong, E., Greenwood, J. M., Gilkes, N.R., Miller, R.C., Jr, , Warren, R.A.J and Kilburn, D.G.. 1992. Proteins designed for adherence to cellulose. In Biocatalyst design for stability and specificity. Himmel M. E. and Georgiu, G. Eds. ACS in press.

Ooshima, H., Sakata, M., Harano, M. 1983. Adsorption of cellulase from *Trichoderma viride* on cellulose. Biotechnol. Bioengineering, 24: 3130-3114.

Owolabi, J.B., Béguin, P., Kilburn, D.G., Miller, R.C., Jr. and Warren, R.A.J. 1988. Expression in *Escherichia coli* of the *Cellulomonas fimi* structural gene for endoglucanase B. Appl. Environ. Microbiol., 54: 518-523.

Pahler, A., Hendrickson, W. A., Kolks, M. A. G., Argarana, C. F., and Cantor, C.R. 1987. Characterization and crystallization of core streptavidin. J. Biol. Chem. 262: 13933-13937.

Pilz, I., Schwarz, E., Kilburn, D.G., Miller, R.C., Jr., Warren, R.A.J., and Gilkes, N.R. 1990. The tertiary structure of a bacterial cellulase determined by small-angle X-ray-scattering analysis. Biochem. J., 271: 277-280.

Richards, F. M. 1990. Reflections. Methods Enzymol., 184, 3-5.

Sano, T. and Cantor, C. R. 1990. Expression of a cloned streptavidin gene in *E. coli*.. Proc. Natl. Acad. Sci. USA., 87: 142-146.

Sano, T. and Cantor, C. R. 1991. A streptavidin-protein A chimera that allows one-step production of a variety of specific antibody conjugates. Biotechnol., 9: 1378-1381

Sano, T. and Cantor, C. R. 1991. Expression vector for streptavidin-containing chimeric proteins. Biochem. biophys. Res. Commun., 7 : 571-577.

Sano, T., Glazer, A. N. and Cantor, C. R. 1992. A streptavidin-methanlothionein chimera that allows specific labeling of biological materials with many different heavy metal ions. Proc. Natl. Acad. Sci. USA., 89: 1534-1538.

Sarko, A., 1987. Cellulose. How much do we know about its structure. In Woods and Cellulosics, Industrial Utilization, Biotechnology, Structure and Properties. J. F. Kennedy, G. Phillips, P. A. Williams eds. Ellis Horwod, pp 55-59.

Sprey, B. 1988. Characterrization of cellulases in an extracellular fraction from *Trichoderma reesei* under conditions of electrofucusing (IEF). FEMS Microbiol. Lett. 55: 283-294.

Spurlino, J. C., Rodseth, L. E., and Quiocho, F. A. 1992. Atomic interactions in protein-carbohydrate complexes. Tryptophan residues in the periplasmic

maltodextrin receptor for active transport and chemotaxis. J. Mol. Biol. 226: 15-22.

Stackerbrandt, E., and Kandler, O. 1979. Taxonomy of the genus *Cellulomonas* based on phenotypic characters and DNA-DNA homology, and proposal of seven neotype strains. Intl. J. Sys. Bacteriol., 29: 273-282.

Studier, F. W., Rosenberg, A. H., Dunn, J. J., Dubebdorf, J. W. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol., 185: 61-89.

Vyas, N. K. 1991. Comparison of the periplasmic receptors for L-arabinose, D-glucose/D-galactose and D-ribose. Structural and functional similarity. Curr. Opinion Struct. Biol. 1: 732-740

Warren, R.A.J., Beck, C.R., Gilkes, N.R., Kilburn, D.G., Langsford, M.L., Miller, R.C., Jr., O'Neill, G.P., Scheufens, M. and Wong W.K.R.1986. Sequence conservation and region shuffling in an endoglucanase and an exoglucanase from *Cellulomonas fimi*. Proteins, 1: 335-341.

Weber, P. C., Olendorf, D. H., Wendoloski, J. J., and Salemme, F. R. 1989. Structural origins of high-affinity biotin binding to streptavidin. Science, 243: 85-88

Wei, R. D. 1970. Assay of avidin. Methods Enzymol., 18A: 424-427.

Whittle, D.J., Kilburn, D.G., Warren, R.A.J. and Miller, R.C., Jr. 1982. Molecular cloning of a *Cellulomonas fimi* cellulase gene in *Escherichia coli*. Gene, 17: 139-145.

Wilchek, M. and Bayer, 1990. Introduction to avidin-biotin technology. Methods Enzymol., 184: 5-13.

Wong, W.K.R., Gerhard, B., Guo, Z.M., Kilburn, D.G., Warren, R.A.J. and Miller, R.C., Jr. 1986. Characterization and structure of an endoglucanase gene cenA of *Cellulomonas fimi*. Gene, 44: 315-324.

Wood, T. M., Mc Crae, S. I. and Bhat, K., M. 1989. The mechanism of fungal cellulase action-synergism between enzyme components of *Penicillium pinophilum* cellulases in solubilizing hydrogen bond-ordered cellulose. Biochem. J. 260: 37-43.

Yannish-Perron, C., Vieira, J., and Messing, J. 1985. Improved m13 phage cloning vectors and host strains : nucleotide sequence of the m13, m18 and pUC19 vectors. Gene 33: 103-109.