17
77

BARD

FINAL REPORT

PROJECT NO. US-3106-99

**Molecular Analysis of Cellulosome Organization in
Ruminococcus albus and *Fibrobacter intestinalis* for
Optimization of Fiber Digestibility in Ruminants**

M. Morrison, E.A. Bayer, J. Miron, R. Lamed

2004

BARD Final Scientific Report**Date of Submission of the report:** March 8, 2004**BARD Project Number:** US-3106-99C**Project Title:** Molecular Analysis of Cellulosome Organization in *Ruminococcus albus* and *Fibrobacter intestinalis* for Optimization of Fiber Digestibility in Ruminants**Investigators****Principal Investigator (PI):** Mark Morrison**Co-Principal Investigators (Co-PI):** Edward A. Bayer
Joshua Miron**Collaborating Investigators:** Raphael Lamed**Institutions**

The Ohio State University

Weizmann Institute of Science
ARO, The Volcani Center

Tel Aviv University

Keywords *not* appearing in the title and in order of importance.*Cellulose degradation, genomics, proteomics, glycoside hydrolases.***Abbreviations** commonly used in the report, in alphabetical order:*Cel9B-family 9 glycoside hydrolase; Cel48A-family 48 glycoside hydrolase; Coh-cohesin module; Doc-dockerin module; PAA/PPA-phenylacetic and phenylpropionic acids.***Budget:** IS: \$180,000

US: \$120,000

Total: \$300,000

Signature
Principal Investigator Signature
Authorizing Official, Principal Institution

630.72
BAR-MOL

BARD Final Scientific Report

BARD Final Scientific Report

Publication Summary (numbers)

	Joint IS/US authorship	US Authors only	Israeli Authors only	Total
Refereed (published, in press, accepted)	2	1		3
Submitted, in review, in preparation	2			2
Invited review papers	1			1
Book chapters				0
Master theses	1		1	2
Abstracts	1	3		4

Postdoctoral Training: List the names and social security/identity numbers of all postdocs who received more than 50% of their funding by the grant.

Estelle Devillard, current address: Rowett Research Institute, Aberdeen, Scotland

Marco Rincon, current address: Rowett Research Institute, Aberdeen, Scotland

Cooperation Summary (numbers)

	From US to Israel	From Israel to US	Together, elsewhere	Total
Short Visits & Meetings	0	2	2	4

Description of Cooperation:

The accomplishments outlined in this report were achieved through many cooperative activities, which include: i) a 3-month visit by an Israeli graduate student to The Ohio State University; ii) the construction, sharing and use of recombinant proteins by the US and Israeli partners, supported in part by the US investigator's leadership of the *R. albus* genome sequencing project and; iii) the training of a second Israeli graduate student (from Dr. Miron's group) in Raphael Lamed's laboratory at Tel Aviv University. These cooperative activities have, so far, resulted in the coauthorship of three research manuscripts and an invited review article, as well as several abstracts and invited presentations at international symposia, given by the Israeli and US partners. Although the grant period has now concluded, additional publications and collaborative activities are planned between the US and Israeli research groups, principally in the area of cellulosome organization and module structure-function relationships in *R. albus*.

Patent Summary (numbers)

	Israeli inventor (s) only	US inventor (s) only	Joint IS/US inventors	Total
Submitted				0
Issued (allowed)				0
Licensed				0

1. Abstract:

Improving plant cell wall (fiber) degradation remains one of the highest priority research goals for all ruminant enterprises dependent on forages, hay, silage, or other fibrous byproducts as energy sources, because it governs the provision of energy-yielding nutrients to the host animal. Although the predominant species of microbes responsible for ruminal fiber degradation are culturable, the enzymology and genetics underpinning the process are poorly defined. In that context, there were two broad objectives for this proposal. **The first objective was to identify the key cellulosomal components in *Ruminococcus albus* and to characterize their structural features as well as regulation of their expression, in response to polysaccharides and (or) PAA/PPA. The second objective was to evaluate the similarities in the structure and architecture of cellulosomal components between *R. albus* and other ruminal and non-ruminal cellulolytic bacteria.** The co-operation among the investigators resulted in the identification of two glycoside hydrolases rate-limiting to cellulose degradation by *Ruminococcus albus* (Cel48A and Cel9B) and our demonstration that these enzymes possess a novel modular architecture specific to this bacterium (Devillard et al. 2004). We have now shown that the novel X-domains in Cel48A and Cel9B represent a new type of carbohydrate binding module, and the enzymes are not part of a cellulosome-like complex (CBM37, Xu et al. 2004). Both Cel48A and Cel9B are conditionally expressed in response to PAA/PPA, explaining why cellulose degradation in this bacterium is affected by the availability of these compounds, but additional studies have shown for the first time that neither PAA nor PPA influence xylan degradation by *R. albus* (Reveneau et al. 2003). Additionally, the *R. albus* genome sequencing project, led by the PI Morrison, has supported our identification of many dockerin containing proteins. However, the identification of gene(s) encoding a scaffoldin has been more elusive, and recombinant proteins encoding candidate cohesin modules are now being used in Israel to verify the existence of dockerin-cohesin interactions and cellulosome production by *R. albus*. The Israeli partners have also conducted virtually all of the studies specific to the second Objective of the proposal. Comparative blotting studies have been conducted using specific antibodies prepared against purified recombinant cohesins and X-domains, derived from cellulosomal scaffoldins of *R. flavefaciens* 17, a *Clostridium thermocellum* mutant-preabsorbed antibody preparation, or against CbpC (fimbrial protein) of *R. albus* 8. The data also suggest that additional cellulolytic bacteria including *Fibrobacter succinogenes* S85, *F. intestinalis* DR7 and *Butyrivibrio fibrisolvens* D1 may also employ cellulosomal modules similar to those of *R. flavefaciens* 17. **Collectively, our work during the grant period has shown that *R. albus* and other ruminal bacteria employ several novel mechanisms for their adhesion to plant surfaces, and produce both cellulosomal and non-cellulosomal forms of glycoside hydrolases underpinning plant fiber degradation.** These improvements in our mechanistic understanding of bacterial adhesion and enzyme regulation now offers the potential to: i) optimize ruminal and hindgut conditions by dietary additives to maximize fiber degradation (e.g. by the addition of select enzymes or PAA/PPA); ii) identify plant-borne influences on adhesion and fiber-degradation, which might be overcome (or improved) by conventional breeding or transgenic plant technologies and; iii) engineer or select microbes with improved adhesion capabilities, cellulosome assembly and fiber degradation. The potential benefits associated with this research proposal are likely to be realized in the medium term (5-10 years).

2. Achievements:

2.1 Significance of main scientific achievements or innovations. Our BARD-supported work over the last five years has resulted in the identification and characterization of several novel mechanisms for bacterial adhesion to plant surfaces, and the demonstration that *R. albus*, one of the major ruminal cellulolytic bacteria, produces cellulosomal and non-cellulosomal forms of cellulases, and employs separate and distinct mechanisms controlling the expression of cellulases and xylanases. Perhaps the most salient achievements during the current granting period was our successful use of proteomic- and genome-based analyses to identify two new processive endocellulases with a novel modular architecture (Devillard et al., *J. Bacteriol.* 2004). Both proteins lack the modules typical of cellulosomal enzymes and instead possess an “X-module” at their C-terminus, which bear similarity to the fibronectin-III binding modules present in some mammalian proteins. We have now shown these “X-modules” are a novel family of carbohydrate binding module (Xu et al., revised for publication in *FEBS Letters*), one of the first functional characterizations of “X-modules” in prokaryote cellulases. These two publications represent key collaborative and synergistic activities among the US and Israeli partners, which could not have been achieved without BARD support. We have also found that these processive endocellulases are conditionally expressed in response to micromolar amounts of phenylacetic and phenylpropionic acids (PAA/PPA) providing for the first identification of enzymes rate-limiting to ruminal cellulose degradation, and a mechanism to explain why growth and cellulose degradation by *R. albus* is conditional on the provision of PAA/PPA. The requirements for PAA/PPA in cellulose degradation were established some years ago but we have since shown that PAA/PPA does not affect xylan degradation or growth of *R. albus* on this substrate (Reveneau et al., *Appl. Environ. Microbiol.* 2003). These BARD-supported activities have provided the impetus for greater study of cellulases and cellulosomes, in all types of cellulolytic bacteria, and the Israeli and US partners are at the forefront of this research.

2.2 Agricultural and/or economic impacts of the research findings, if known: We proposed in the original application that the potential benefits associated with this research proposal are likely to be realized in the medium term (5-10 years). Up to 36% of the diet (on a dry matter basis) fed to dairy cattle is fibrous in nature, but only 50-60% is degraded and fermented by ruminal and hindgut microorganisms. Forage consumption by a US dairy cow

ranges between 13,000 and 18,000 pounds (dry matter basis) per year. There are approximately 9.2 million dairy cows in the US and using available nutrition models, a five percentage unit increase in ruminal fiber degradation will raise the Net Energy for lactation (NE_l) by 0.05 Mcals per pound of forage consumed. Assuming each cow consumes 15,000 lbs of forage a year, and 1.0 Mcal of NE_l is worth 5 cents, a five percentage unit increase in ruminal fiber degradation would be worth US\$345 million per year to the US dairy industry. Such figures do not incorporate the cost savings associated with the commensurate reduction in animal waste, which would also arise from improved fiber degradation. Despite these impressive numbers, as the human consumption of cereal grains increases, there will be an even greater need for alternative animal feeds and improved fiber digestion. Because ruminants can use crop residues and cellulose-rich byproducts, the importance of these animals in the human food chain is likely to increase. Improved fiber degradation may therefore be more relevant in maintaining the dairy industry's viability, if society invokes the use of more fibrous feeds. Our research has provided insight into both the enzymology (Cel9B and Cel48A) and adhesion mechanisms controlling ruminal fiber degradation. Future research should allow our use of this information to develop strategies (physicochemical, plant-based and (or) microbial/enzyme-based) that improve ruminal fiber degradation.

2.3 Details of cooperation: It was anticipated that a highly interactive collaboration would develop among the partners, ensuring that results from individual laboratories could be successfully incorporated into other aspects of the project. In that context, the US partners were instrumental in identifying both Cel9B and Cel48A through a combination of mutant isolation, proteomic and genome sequence analyses. On the basis of these findings, the Israeli partners, principally through the efforts of a postdoctoral scientist, Dr. Qi Xu, were able to ascribe function to the X-modules identified in Cel9B and Cel48A. The availability of the *R. albus* genome sequence further supported the collaborative and synergistic nature of our research accomplishments. For example, Dr. Bayer's identification of genes encoding tandem repeats of the X-module, and our initial findings of the relevance of tandem versus singular copies of the X-module, have already been published.

Dr. Bayer and Dr. Rincon-Torres, (a postdoctoral scientist in Dr. Morrison's lab) were able to identify no less than 20 open reading frames (ORFs) containing modules similar to the

Type I dockerins of *Clostridium thermocellum* (see Appendix). Based on sequence alignments five of these dockerin modules were selected for PCR-amplification and cloning into protein expression vectors. An Israeli graduate student, Miss Liat Levy, who was co-supervised during her MSc by Dr. Lamed and Dr. Bayer, also worked on this project during her 3-month visit to The Ohio State University. She cloned two dockerin-containing open reading frames in protein expression vectors, and examined suitable growth and induction conditions for protein overexpression. She also learned methods involved with prokaryote RNA isolation and was able to conduct a limited series of slot blot analysis of transcript abundance in *R. albus* 8 following growth in cellulose-containing media. All these various constructs of cohesin- and dockerin-modules constructed by Dr. Rincon-Torres have been provided to the Israeli partners, for further analysis of dockerin-cohesin interactions.

The Israeli investigators have a long-standing cooperative relationship, stemming from Dr. Miron's terms as a visiting scientist with Dr. Bayer and Dr. Lamed, which led to the initial isolation of adhesion-defective *R. albus* SY3 mutants and a joint publication. Both Dr. Bayer's and Dr. Lamed's groups are particularly proficient in studying the structural biology of the scaffoldins and anchoring proteins, including protein modeling, site-directed mutagenesis, and fine structure analyses of their functional domains. Dr. Miron's group is especially proficient in determining cell wall chemistry, assays of glycolytic activities and bacterial adhesion processes in ruminal bacteria. The nature of the collaboration among the Israeli partners was largely supportive, and directed towards objective 2 of this proposal (see Appendix). Additionally, a Master of Science student, Mr. Yossi Jacobovitch, worked under the tutelage of Drs. Lamed, Miron and Bayer, and he was successful in creating a variety of recombinant clones expressing candidate cohesin and dockerin modules identified from the *R. albus* genome sequence. Combined with the recombinant clones produced at The Ohio State University, we are now well positioned to dissect dockerin-cohesin interactions in *R. albus*. This research will continue between the US and Israeli partners despite the termination of this BARD award.

The most disappointing aspect of the collaboration has been the impact of the turmoil in the Middle East during the last several years, which has precluded travel to Israel by Dr. Morrison. Hopefully, this can be rectified during the first six months of 2004. Despite this limitation, Drs. Bayer, Lamed and Morrison have still coordinated visits with each other outside the USA and Israel. All three of us were invited participants at the 2003 Gordon Research

Conference on Cellulases and Cellulosomes, and also the 2003 Mie Bioforum. The net results of these interactions have been our continued commitment to the study of cellulosomal and non-cellulosomal glycoside hydrolases of *R. albus*, and new projects pertaining to cellulosome structure-function studies with *Clostridium thermocellum*.

In conclusion, the BARD award has supported productive collaborations among several Israeli groups, and provided the opportunity for Dr. Morrison to establish long-term, collaborative partnerships with Israeli scientists. The nature of these US-Israeli partnerships is reflected in the form of publications in high quality journals, the continued requests the partners receive to participate in prominent international symposia, and our desire to acquire additional grant support to further our understanding of cellulases and cellulosomes for the betterment of agriculture and society.

2.4 List of Peer Reviewed Publications:

1. Devillard, E. Goodheart D.E., Karnati, S.K.R., Bayer E.A., Lamed, R., Miron, J., Nelson, K.E., and Morrison, M. (2004) *Ruminococcus albus* 8 mutants defective in cellulose degradation are deficient in two processive endocellulases, Cel48A and Cel9B, both of which possess a novel modular architecture. *J. Bacteriol.* 186: 136-145.
2. Reveneau, C., Adams, S.E., Cotta, M.A., and Morrison, M. (2003) Phenylacetic and phenylpropionic acids do not affect xylan degradation by *Ruminococcus albus*. *Appl. Environ. Microbiol.* 69: 6954-6958.
3. Miron, J., Jacobovitch, J., Bayer E.A., Lamed, R., Morrison, M. and Ben-Ghedalia, D. (2001). Subcellular Distribution of Glycanases and Related Components in *Ruminococcus albus* SY3 and their Role in Cell Adhesion to Cellulose. *J. Appl. Microbiol.* 91: 677-685.
4. Miron, J., Ben-Ghedalia, D. and Morrison, M. (2001) Invited Review: Adhesion mechanisms of rumen cellulolytic bacteria. *J. Dairy Sci.* 84: 1294-1309.

BARD Final Scientific Report

Appendix to BARD Final Scientific Report US-3106-99C

Molecular Analysis of Cellulosome Organization in *Ruminococcus albus* and *Fibrobacter intestinalis* for Optimization of Fiber Digestibility in Ruminants

Submitted by

Mark Morrison, Joshua Miron, and Edward A. Bayer

TABLE OF CONTENTS

Summary of unpublished US studies.....	Page 9
Summary of Israeli studies.....	Page 11
Peer Reviewed Publications.....	Page 28
Submitted Manuscript	
Published Abstracts	

Summary of unpublished US studies:

Dr. Devillard and Miss Goodheart have specifically addressed the Objective in the BARD project which was to use two-dimensional protein gel electrophoresis to examine the “proteome” of *R. albus* strain 8, as affected by phenyl-substituted fatty acids, ruminal fluid, or the carbon source used for growth (cellobiose, cellulose and xylan). *R. albus* surface proteins were recovered by Sarkosyl extraction of whole cells, and concentrated by ultrafiltration (30,000 mwco). As shown with cellobiose-grown cells in Fig. 1, two proteins, were observed to increase

Fig. 2: 2D-PAGE of cell wall associated proteins recovered from *R. albus*, following growth of the bacterium in cellobiose-containing medium, either in the absence (left) or presence (right) of PAA/PPA. Analysis to date has focused on high molecular protein(s) in greater abundance following growth of the bacterium in the presence of PAA/PPA (highlighted in yellow boxes). Edman degradation and MALDI-TOF analysis of these proteins has provided sufficient evidence for these proteins to be confirmed as being Cel9B and Cel48A (Devillard Goodheart and Morrison, unpublished).

dramatically in extracts prepared from cells cultured in the presence of PAA/PPA (or ruminal fluid), as compared to cells grown in the absence of such additions. Similar results were also obtained for cellulose- and xylan-grown cells (Table 1). These two proteins were subsequently identified as Cel9B and Cel48A (Devillard et al. 2004). Miss Goodheart is currently examining *cel9B* and *cel48A* gene expression by quantitative real-time PCR, to verify the role of PAA/PPA in conditional gene expression.

In addition to the research conducted by Dr. Devillard and Miss Goodheart, Dr. Marco Rincon-Torres spent considerable time over the last 12 months cloning and overproducing candidate dockerin and cohesin modules from the *R. albus* 8 genome. In a collaborative study with the

Israeli partners, Western immunoblots with antisera, raised against either the X-domain or cohesin modules from the *R. flavefaciens* ScaA and ScaB proteins, showed that both *R. albus* strains 8 and SY3 possess epitopically similar proteins to *R. flavefaciens*. Notably, a protein of 200 kDa in *R. albus* SY3 crossreacted strongly with the anti-ScaA antiserum. All three antisera crossreacted, albeit weakly, with high-molecular-weight proteins from *R. albus* 8, and unlike strain SY3, the strongest interaction occurred with two polypeptides of less than 20 kDa, and with the anti-ScaB-X-domain antiserum (Fig. 2). Such crossreactivity suggests the presence of ScaA and ScaB X-domain homologs in *R. albus*, but perhaps, not exclusively within scaffoldin modules.

Summary of Israeli studies

The combined objective of the two Israeli research teams was to identify the cellulosomal organelles and modules involved in the adhesion mechanism(s) of the rumen bacteria *R. albus*, and to study possible similarities in the structure and architecture of cellulosomal organelles between *R. albus* and other rumen and non-rumen cellulolytic strains. The activities performed during the project included:

- Establishment of a cDNA genetic library of *R. albus* SY3. The genome was subjected to treatment by EcoRI and XhoI restriction enzymes, then inserted into phage Uni Zap Gigapack III, cloned into in E. Coli XL1-Blue and expressed. The library was amplified to a level of PFU = 4.5×10^9 cells/ml.
- Comparative blotting studies using specific antibodies: prepared against purified recombinant cohesins and X-domains, derived from cellulosomal scaffoldins of *R. flavefaciens* 17, a *Clostridium thermocellum* mutant-preabsorbed antibody preparation, or against cbpC (fimbrial protein) of *R. albus* 8. Surface proteins of *R. albus* strains SY3 and 8 were separated on SDS-PAGE and subjected to Western blotting using the above antibodies. Additional cellulolytic bacteria were examined in a similar manner. Data of the antigenic homology studies is shown in **Figure 1** and **Table 1**. The data show that a ~200-kDa *R. albus* SY3 surface protein interacted with RF-17 ScaA-specific antibodies. SY3 proteins of ~120 and ~55 kDa interacted with RF-17 ScaB antibodies; and SY3 proteins ~120 and ~90 kDa exhibit homology with RF-17 ScaA Cohesin 2. The data also suggest that additional cellulolytic bacteria including *Fibrobacter succinogenes* S85, *F. intestinalis* DR7 and *Butyrivibrio fibrisolvens* D1 may also employ cellulosomal modules similar to those of *R. flavefaciens* 17.
- According to these findings, we have designed and prepared three degenerate RNA primers, based on the amino acid sequencing of two regions on the ScaA cohesins of *R. flavefaciens* 17. The primers used were: F1= 5' GCN ATH GAY GAR TTY GAR AAY AAY3' and reverse primer R1 = 5' RTT RTT YTC RCC YTC RTC DAT NGC 3'; on the other side primer R2 = NGT NCC NCC YTC YTT RAA NAC 3'. The three primers were used for RT-PCR against genomic DNA of *R. albus* SY3. Unfortunately, these primers

failed to hybridize with SY3 DNA. This suggests possible differences in cohesin structure between *R. flavefaciens* and *R. albus*.

- We therefore concentrated further on the surface proteins of *R. albus* SY3 and separated them on 7.5% SDS-PAGE gels. The 200-kDa protein (that reacted positively with RF 17 Sca A antibodies, Table 1) was isolated and submitted to The Protein Sequencing Laboratory (Technion Haifa, Israel). The band was subjected to proteolysis with either pepsin or chymotrypsin, and selected peptides were sequenced. The peptides analysis succeeded to identify a previously recognized xylanase and additional 3-4 unidentified proteins, neither of them demonstrating any similarity to the database GenBank structures of cohesins and scaffoldin. Therefore we focused on a different approach that combined Genomic analysis of possible cohesins and dockerins from the genome project of *R. albus* 8 published on the TIGR website together with expression of possible cohesins and their functional identification.
- Establishment of a cDNA genetic library of *R. albus* 8. The genomic DNA was digested by EcoRI and XhoI restriction enzymes, then inserted into pre-digested phage using Uni Zap XR Vector kit, and packed by using a Gigapack III Gold packaging kit. The library was amplified in *E. coli* XL1- Blue to a level of 4.13×10^{13} plaque-forming units /ml.
- Identification of putative cohesin sequences in *R. albus* 8 genome. We have identified new putative sequences of cohesins from the genetic reservoir of the genome project of *R. albus* 8, published on the TIGR website. Sequences ORF-1334 and ORF-505 of Contig 2 from *R. albus* 8 showed a stop codon between the 2 ORFs and included repeated segments that were tentatively identified as cohesins. We further determined that there may be an error in the translation of the genome between these 2 ORFs, which would preclude the stop codon. Therefore, 9 cohesin sequences could be predicted for the revised ORF in Contig 2. **Figure 2** shows the sequence of the three identified possible cohesins.
- Expression and purification of the putative cohesions: In order to try to express representative forms of these putative cohesins, the following 2 primers were designed.

Primers:

NcoI

Forward 5' - ATATCCATGGAGAAATTTGAGATCGA -3'

M E K F E I E

XhoI

Reverse 5' - ATATCTCGAGCTGATCGTCATTGTTGT -3'

Q D D N N N

- Three PCR fragments of the cohesins (1.4, 1.1 and 0.7 kb, respectively) were prepared by using these two primers (**Figure 3**). The cohesin fragments were extracted and purified from an agarose gel by using a Gel Purification Kit (Roche). The selected PCR fragments were subjected to include *NcoI* and *XhoI* restriction sites, and then inserted into pET28a plasmid. The plasmids containing the cohesin inserts were transformed to competent *E. coli XL1-Blue* cells by electroporation.
- The plasmids were extracted from the host cells using the High Pure Plasmid Isolation Kit (Boehringer Mannheim) and then transformed into *E. coli* BL21 (DE3) via electroporation. The cloned cohesins were isolated and expressed. Then samples were examined by SDS-PAGE. Three different recombinant cohesin-containing proteins were thus obtained exhibiting molecular masses of about 51, 37 and 30 kDa, and the samples were provisionally termed Coh-51, Coh-37 and Coh-30.

Functional identification of the putative cohesins:

- Reaction of *R. albus* cohesins with proteins originated from bacterial envelopes: The three putative cohesins were labeled with a 6xHis-tag in order to facilitate subsequent isolation and detection. The His-tagged cohesins were then purified from a cell lysate (2 ml) and isolated on a 0.5-ml Ni-NTA column, prior to reaction with proteins from the outer bacterial envelopes of *R. albus* 8. Complexes of the His-tagged cohesins plus interacting bacterial proteins, were eluted from the column by 250 mM imidazole, separated on 15% SDS-PAGE gel and then transferred to nitrocellulose membranes, prior to reaction with anti-His-tag antibodies. All three putative cohesin samples reacted at high affinity with a group of presumed dockerin-containing proteins of M_r 35-90 kDa, obtained from the capsule and cell-wall envelope fractions of *R. albus* 8 (**Figure 4**). The His-tagged samples did not react with either the control (*E. coli* proteins) or with proteins obtained from the extra-cellular fluid or cytoplasmic fractions of *R. albus* 8.
- Reaction of *R. albus* putative cohesin Coh - 51 with seven *R. albus* 8 dockerins: Seven different cloned dockerins of *R. albus* 8 prepared by the research group of the Principal US copartner (M. Morrison) and labeled with a 6xHis-tag were delivered to our laboratory for measuring their reaction with *R. albus* 8 putative cohesin Coh - 51. These dockerins were assigned according to their Contig as: Doc 186, Doc 427, Doc 245, Doc 96, Doc 146 and Doc 8. The cloned cohesin Coh-51 was further cleaned, labeled by BIOTIN and reacted on a Ni-NTA column with the 6xHis-tag dockerins to create complexes of cohesin-dockerin. The different cohesin-dockerin complexes were elution from the columns and separated on SDS-page followed by Western blotting and labeling by the Biotin-Avidin reaction (**Figure 5**). Data show positive specific reaction between suspensions of each of these seven dockerins with Coh - 51, suggesting that these dockerins and Coh - 51 are organelles of *R. albus* 8 cellulosome.
- The affinity of the cohesin-dockerin reaction between *R. albus* 8 cohesin Coh - 51 and cloned dockerins of *Acetovibrium celluliticum*: The cloned His-tagged cohesin Coh - 51 of *R. albus* 8 was further purified (at 60°C) and reacted on a Ni-NTA column in a medium containing Ca ions with and without the presence of EDTA with three different cloned dockerins Doc A, Doc B and Doc E originated from the bacterium *Acetovibrium celluliticum*. The affinity of the dockerin-cohesin complexes was then examined by ELISA method

(Figure 6). Data show high affinity between Coh-51 and the three dockerins of *Acetovibrium celluliticum* which is not affected by the presence of Ca and EDTA. This finding point out to a phylogenetic similarity between the dockerins of *R. albus* 8 and *Acetovibrium celluliticum*.

Bioinformatics analyses: Based on the data obtained in this study and previous data available in Gene-banks, phylogenetic relationship among known cohesins from different microorganisms was examined and presented in Figure 7. Most *R. albus* cohesins (identified from the genome of *R. albus* strain 8) appear in a green cluster different in its structure from that of other microorganisms.

- We have also compiled the known sequences of *R. albus* and *R. flavefaciens* cellulosomal (dockerin-containing) and noncellulosomal proteins for comparison with the sequences obtained from the genome project. A schematic view of the known cellulosomal and non-cellulosomal enzymes from *R. albus* is illustrated in Figure 8. Genes encoding for the enzymes shown in the figure were sequenced by various different groups from different strains of *R. albus* and *R. flavefaciens* and are available in one or more of the data bases, e.g., GenBank. A brief summary of the phylogenetic diversity of *R. albus* 8 dockerins, compared to the sequences of dockerins from other cellulolytic species, is shown in Tables 2 and 3 and summarized in Figure 9. Almost 30 different dockerin species have been detected from the *R. albus* 8 genome. The majority of *R. albus* dockerins (identified from the genome of strain 8) appear in a defined cluster and their amino acids structure differ from dockerins of other bacteria, and from dockerins of egVI and egVII endoglucanases previously identified in *R. albus* F-40.

Figure 1. The reaction on western blot of capsular proteins obtained from rumen bacterial strains: *F. intestinalis* DR7 and NR9; *F. succinogenes* S85; *Butirivibrio fibrisolvens* A38 and D1; and *R. albus* SY3 and 8'; with antibodies specific to cellulosomal organelles of *R. flavefaciens* 17 (AB-ScaA; AB-ScaB and aB-CohA), *Clostridium thermocellum* (AB-Scaffoldin); and to cbpC of *R. albus* 8, and preabsorbed antibodies specific to adhesion organelles of *R. albus* SY3.

Table 1. Bacterial surface proteins (approximate MW in kDa) of some rumen cellulolytic strains that react positively with rabbit antibodies (AB) specific for cellulosome- or adhesion-related components¹ of *Ruminococcus flavefaciens* 17 (RF 17), *Clostridium thermocellum* (CT) or *Ruminococcus albus* 8 (RA 8).

Bacterial strain	AB for ScaA RF 17	AB for ScaB RF 17	AB for CohA RF 17	AB for Sca CT	AB for CbpC RA 8
<i>R. albus</i> 8	90	400, 120, 50, 30,	120, 110, 50,	400, 220, 120, 100	220, 20,
<i>R. albus</i> SY3	200, 45	120, 55	120, 90,	120, 75,	30, 40,
<i>F. succinogenes</i> S85	90	120, 60	140, 120	220, 120, 110, 90, 60, 40	240, 70,
<i>F. intestinalis</i> DR7	160	230, 160, 120, 50	220, 160, 120	220, 160, 140, 60,	240, 220, 200, 150, 95
<i>F. intestinalis</i> NR9	400	400, 60	400, 220	400, 130, 60	-
<i>B. fibrisolvens</i> D1	130, 50	170, 95, 40	130	130, 120, 100, 95, 60, 50, 40	-
<i>B. fibrisolvens</i> A38	120	300, 95	-	55	-
<i>E. coli</i>	90, 30	100, 90, 50, 40, 30	170, 50, 30	95, 75, 60, 40	95, 60, 40,

¹Cloned and expressed, recombinant components were used for preparation of cellulosome- and adhesion-related antibodies. These include: X-modules from *R. flavefaciens* (strain 17) ScaA (scaffoldin A) and ScaB, and cohesin A from ScaA; cellulosome C. *thermocellum*; and cellulose-binding protein C (CbcC, fimbrial element) of *R. albus* 8.

Figure 2. DNA sequence of a possible cohesin ORF-1334 1334 aa (marked in red) on contig_2 from the genome of *R. albus* 8 published on TIGR database.

signal peptide!

MIRNPFxHGRKISGRKRLVSGLTAGLQAASSIVGALPASVASAAGDLVDLGTfKYATQKE
LGSANKYAAFAYTYTQANHMEGIFACNNYHPYGDAGFTTNNVNVYVDSKENHIYIGENIG
GGDAELNVIEQVHQTNGINDDMLWDMILPDTITINRTANNGQGVQLEDVNGYQSGVFNSE
KDLAKKIYHVSEVTYNIDFIDALTALDEYQEWAKQDTSDVTEVTLTPDKENSQNNWVID
IECAEGGNVVTLDQAQDIKDSRLNVKAKDGVSDYSLIINVTGLNDGDHFDREIKIDDADSL
YGPQAGKVLFNMLGEDGDQYYFTKLNQGVILAPTQEVYIYQTHNGSVMGFYVENNGCEIH
QNGFRQLAVKDHNNIEISKMAVGGQKELPGARLSVTAASGNNDVDWNKVIAANDGVRPVG
GTTSAPTGITWLSGTEPRIIKGLPDGTYYLDESGADADGNVTLDGKKYQVVESSLRFTVK
EGQLTDVRSSTELSMYGVQIDGSYYTAENVQKTGYDGKLFVCNAVAEDIVSGPADLVINKF **A**
DLTGENEIAGAKLTITGTTDAGEAVSMDWTSKENETWNISLDDGVYTLSETGDIVTDDG
VFDVIPSSLNFTVKNGKVTVDKTDNGADISKDGFNKNAYDGYFTVGTNADKAEVINICDA **B**
KRTNVKLSKQSTGGDEIPNAKLTILKGV DANGRLSFENVRFADSSVTDGSIKNNVISWT
SSRTPVEVVGLPNGSYTFHEVAAPNGYQVTTDITFDVYGNVTVTSVNGKATSGTVIMTD
NRNTTVSINKFDLTGKNEVEGARLTVKGISAGVEDYEYSWVSVKGAEVLTLGLADGDYTL **DEG?**
TEEADDESGEIFDMTTGKYKVIKDTLTFSVRDGSVWNTSSNKVEIYPSTPSSAPAEYY
VAGRNNTISICDAEIKQGDVTINKFDITGQNEVEGAQLTLTGEGIDWARLAEFNPDSQYI **C**
TDKTGKVTGLQWVSEKGAEWNVTLRDGTGYVLTEATAVAKDGETEYEVIPSSISFDIVDGK
VANIDNKQTDIAGRANPNATDETDSYVVANNGVNEIYICDAKKIIVKTTDVTISKTDIT **D**
GEKEIKGAQLQVTSTDVDWTEIVAANAGLEKVTQNGKVLGLSWTSDDQSPVVLWGLTEGS
YTLTETGKPFYDGDIEVIDSKIGFTIDADGKVTVISETKPAANAESDKGYAVGKGD
ITICDASRTVITVSKKDITGEKEVIGAHLVIKDASGRIVDQWDSNGKDHIVSALTNGTYT **E**
LEETAVIDAAVAGGDFKIIDSVVFTTVKNGKITSAVAGKALYNDFNAESED SYAVVDKATG
KITICDAEKVVKTT* [stop], the ORF505 starts (the same frame!) F

ONE protein

△ >ORF-505

⇒ 4

VIINKTDITGEEELDGAAILTVTDSGKTVGESWTSKKGEKFEIELEDGTYTLTETGDTVT
NEAGEEFEIITSEIKFTVEDGKVTVTTEKLVKDTFDKDSLEGYVVVTNGNEADEITVCDA **G**

⇒ 4

KKKTKVIINKTDITGEEELDGAAILTVTDSGKTVGESWTSKKGEKFEIELEDGTYTLTE
TGDTVTNEAGEEFEIITSEIKFTVEDGKVTVTTEKLVKDDFDKDSLEGYVVVANGNDADE **H**

⇒ 4

ITVCDAKKKTKVIINKTDITGEEELDGAAILTVTDSGKTVGESWTSKKGEKFEIELEDG
TYTLTETGDTVTNEAGEEFEIITSEIKFTVEDGKVTVTTEKLVKDTFDKDSLEGYVVVTN **I/part**
GNEADEITVCDAKKKTVVDGDESSKTDVDDSSSEEQKDNESSRPGPDMSSVPEGDTT
ADSSKADTDESSATDDSSASSSTSSSASSTSTSSKATTTTSTTTNTSNPNTGARVLN **SLH**
AAEFAAAVGLCFFAVRRKKNNDDQ

← 6

Figure 3. Agarose gel of the DNA fragments cloned into *E. coli* BL21 after cutting by restriction enzymes. Lanes 1, 2 and 3 show 0.7 kb, 1.1kb and 1.4 fragments respectively. lanes 4 and 5 show markers distribution 1 - 100 bp.

Figure 4. SDS-page (D) and Western blotting showing the reaction of the three cloned cohesin proteins (Coh 30 = A, Coh 37=B, and Coh 51=C) marked by X6His-Tag with proteins containing dockerins of *R. albus* 8 envelopes. Lanes: 1) MW marker; 2, 4, 13, 16, 19) *R. albus* 8 capsular proteins; 3, 12, 15, 18) *R. albus* 8 cytoplasmic proteins; 5, 11, 14, 17) *R. albus* 8 cell wall proteins; 6, 10) *R. albus* 8 extra cellular fluid proteins; 7) Positive control = X6His-Tagged cloned Coh 51; 8, 9) Negative control = un-cloned BL21;

Figure 5. Western blotting showing the reaction of cohesin Coh 51 marked by BIOTIN with cloned dockerins of *R. albus 8* and native proteins of *R. albus 8* envelopes. Lanes: 1) Coh - 51 marked by biotin; 2) *R. albus 8* extra cellular fluid proteins; 3) Sonicated *R. albus 8* cells; 4) *R. albus 8* cell wall proteins; 5) Dockerin 186 B sup. over night; 6) Doc 427 pellet 3h ; 7) Doc 427 sup. 3h ; 8) , Dual color marker (BioRad); 9) Doc 245 pellet 3h; 10) Doc 245 total cell; 11) Doc 96 sup. over night; 12) Doc 8 pellet 3h; 13) Doc 8 sup. 3h; 14) Doc 146 sup. 3h; 15) Doc 186 A sup.

- Figure 6.** Interaction of cloned His-tagged cohesin Coh - 51 of *R. albus* 8 on a Ni-NTA column (in a medium containing Ca ions with and without EDTA) with cloned identified dockerins: Doc A, Doc B and Doc E, originated from the bacterium *Acetovibrium celluliticum*. The affinity of the dockerin-cohesin complexes was examined by ELISA method. Data point out to a phylogenetic similarity between the dockerins of *R. albus* 8 and *Acetovibrium celluliticum*.

Figure 7. Phylogenetic relationship among known cohesins from different microorganisms.

Most *R. albus* cohesins (identified from the genome of strain 8) appear in a green cluster at the left side of the figure.

Preliminary Phylogenetic Tree of *R. albus* Cohesins

0.1

Known Enzymes from *Ruminococcus albus*

Known Enzymes from *Ruminococcus flavefaciens*

Figure 8. Known enzymes in *Ruminococcus albus* and *R. flavefaciens*. Shown are the schematic modular architecture of the different enzymes that have been sequenced and are currently available in the GenBank database. Several other *R. albus* enzymes are known (EgA, EgB and Eg4) that are very similar in primary sequence and structural architecture to EgI. The dockerin-containing enzymes (3 endoglucanases from *R. albus*; 2 endoglucanases and 4 xylanases from *R. flavefaciens*) are considered to be cellulosome-related.

Table 2. Specificity of residual amino acids in places 10 and 11 in the Ca binding motif of the 1st and 2nd double segments *R. albus* dockerins analysed in this study.

Enzyme or dockerin	Protein	Residues segment ^a	in duplicated
		1st	2nd
Ruminococcal enzyme from:			
<i>R. albus</i>	Doc 8	VT	VT
	Doc 96	VT	IT
	Doc 146	VT	VT
	Doc 186A	VT	-
	Doc 186B	VT	VT
	Doc 245	?	?
	Doc 427	VT	VT

BARD Final Scientific Report

Table 3. The structural difference in residual amino acids located in places 10 and 11 in the Ca binding motif of the 1st and 2nd double segments between *R. albus* dockerins identified in this study and dockerins previously identified in other bacterial enzymes.

Enzyme or dockerin	Protein	Residues	in	duplicated
		segment ^a		
		1st	2nd	
Ruminococcal enzyme from:				
<i>R. flavefaciens</i>	EndA	LA	GD	
	EndB	LA	ND	
	XynB	LA	AD	
	XynD	LS	GD	
	CesA	SF	VA	
	XynE	SL	VA	
<i>Ruminococcus</i> sp. strain	Xyn1	SF	VA	
<i>R. albus</i>	EgV	VT	VS	
	EgVI	MA	ND	
	EgVII	DT	MA	
<i>C. thermocellum</i> enzymes	Consensus ^b	ST	ST	
Mesophilic clostridial enzymes	Consensus ^b	A(L/I)	A(I/L)	
Scaffoldin dockerins from:				
<i>R. flavefaciens</i>	ScaA	VA	AV	
<i>C. thermocellum</i>	CipA	LL	MQ	
<i>A. cellulolyticus</i>	CipV	LE	LE	
<i>B. cellulosolvens</i>	CipBc	SD	SD	

Figure 9. Phylogenetic relationship among known dockerins from different microorganisms.

The majority of *R. albus* dockerins (identified from the unfinished genome of strain 8) appear in a defined cluster at the top of the figure. A previously sequenced enzyme (EgV from strain F-40) maps within the cluster. Two other F-40 enzymes (EgVI and EgVII) are more similar to previously sequenced dockerins from *R. flavefaciens* (white circles). The *R. albus* dockerins are highlighted in yellow. Scaffoldin-borne dockerins from various species are shown as blue- or brown-colored squares.

Phylogenetic Tree of Dockerin Sequences