Cellulose: Molecular and Structural Biology

Selected Articles on the Synthesis, Structure, and Applications of Cellulose

Edited by

R. Malcolm Brown, Jr. and Inder M. Saxena

The University of Texas at Austin, Austin, Texas, U.S.A

Springer
# TABLE OF CONTENTS

Preface ............................................................................................................. xiii

Chapter 1: Many Paths up the Mountain: Tracking the Evolution of Cellulose Biosynthesis ................................. 1  
*David R. Nobles, Jr. and R. Malcolm Brown, Jr.*

1. Introduction ........................................................................................ 1  
2. Sequence Comparisons ................................................................... 3  
3. Eukaryotic Cellulose Synthases ....................................................... 4  
   3.1. The case for a cyanobacterial origin of plant cellulose synthases .. 4  
   3.2. Lateral transfer of cellulose synthase in the urochordates .......... 4  
   3.3. The cellulose synthase of *Dictyostelium discoideum* ................ 5  
4. Bacterial Gene Clusters ................................................................... 6  
   4.1. Introduction ........................................................................... 6  
   4.2. Characterized gene clusters ................................................... 6  
5. Novel Gene Clusters ........................................................................ 8  
   5.1. Introduction ........................................................................... 8  
   5.2. Group III .............................................................................. 8  
   5.3. Group IV ............................................................................... 10  
6. Concluding Remarks ....................................................................... 12  
References ............................................................................................ 12

Chapter 2: Evolution of the Cellulose Synthase (*CesA*) Gene Family: Insights from Green Algae and Seedless Plants ................................................................. 17  
*Alison W. Roberts and Eric Roberts*

1. Overview ........................................................................................... 18  
2. The Prokaryotic Ancestry of Eukaryotic *CesAs* ........................... 21  
3. Green Algal *CesAs* and the Evolution of Terminal Complexes ...... 23  
   4.1. Evolution of tracheary elements .......................................... 25  
   4.2. Functional specialization of CesA proteins ......................... 26
# Table of Contents

4.3. Tip growth and the function of *Cellulose synthase-like type D (CslD) Genes* ............................................................... 26

4.4. *CesA* and *CslD* genes of the moss *Physcomitrella patens* .......... 27

5. Analysis of *CesA* Function by Targeted Transformation in *P. patens* ................................................................. 28

Acknowledgments ............................................................................. 28

References ......................................................................................... 29

## Chapter 3: The Cellulose Synthase Superfamily .................................... 35

*Heather L. Youngs, Thorsten Hamann, Erin Osborne and Chris Somerville*

1. Introduction ...................................................................................... 35

2. Identification of Cellulose Synthase .................................................. 37

3. Toward a Functional Analysis of Cellulose Synthase ........................ 38

4. Identification of the Cellulose Synthase-like Genes ........................... 40

Acknowledgments ............................................................................. 45

References ......................................................................................... 46

## Chapter 4: Cellulose Synthesis in the *Arabidopsis* Secondary Cell Wall ............................................................. 49

*Neil G. Taylor and Simon R. Turner*

1. Introduction ...................................................................................... 50

2. *irx* Mutant Isolation and Characterization ........................................ 50

3. Three CesAs Are Required for Secondary Cell Wall Cellulose Synthesis ................................................................. 51

4. Function of Multiple CesA Proteins during Cellulose Synthesis ...... 52

5. Localization of CesA Proteins ........................................................... 54

6. Conservation of CesA Protein Function in other Species ............... 55

7. Other *irx* Genes Required for Secondary Cell Wall Formation .......... 55

8. Identifying Novel Genes Required for Secondary Cell Wall Formation Using Expression Profiling ................................. 57

9. Alternative Approaches to Studying Cellulose Synthesis in the Secondary Cell Wall ................................................................. 58

10. Conclusions ....................................................................................... 59

References ......................................................................................... 59

## Chapter 5: From Cellulose to Mechanical Strength: Relationship of the Cellulose Synthase Genes to Dry Matter Accumulation in Maize ............................................................. 63

*Roberto Barreiro and Kanwarpal S. Dhugga*

1. Introduction ...................................................................................... 64

2. Role of Cellulose in Stalk Strength .................................................... 65

3. Carbon Flux through Cellulose Synthase .......................................... 65

4. Alteration of Cellulose Formation in Plants ....................................... 66

5. Mass Action and Metabolic Control .................................................... 68
Table of Contents

6. The Cellulose Synthase Gene Family .................................................. 71
7. Expression Analysis of the ZmCesA Gene Family .......................... 73
8. Rationale for Future Transgenic Work ............................................ 76
9. Summary ............................................................................................ 77

References ............................................................................................. 78

Chapter 6: Cellulose Biosynthesis in Forest Trees .......................... 85
Kristina Blomqvist, Soraya Djerbi, Henrik Aspeborg, and Tuula T. Teeri

1. The Properties of Wood ................................................................. 86
   1.1. Formation of wood cells ............................................................ 86
   1.2. Reaction wood ......................................................................... 88
2. Cellulose Synthesis ........................................................................... 89
   2.1. Rosettes: the machinery of cellulose synthesis .......................... 90
   2.2. CesA and Csl ........................................................................... 90
   2.3. Other enzymes and proteins involved in cellulose synthesis ........... 96
   2.4. Other metabolic processes involved in cell wall biosynthesis .......... 98
3. In Vitro Cellulose Synthesis .............................................................. 99

Acknowledgments .................................................................................. 100

References ............................................................................................. 100

Chapter 7: Cellulose Biosynthesis in Enterobacteriaceae ............... 107
Ute Römling

1. Introduction ...................................................................................... 107
2. The Cellulose Biosynthesis Operon in Salmonella typhimurium and Escherichia coli ......................................................... 109
3. Regulation of the Expression of the bcsABZC Operon ..................... 112
4. Regulation of Cellulose Biosynthesis .............................................. 112
5. Regulation of csgD Expression ......................................................... 114
6. Function of AdrA ............................................................................ 115
7. Occurrence of the Cellulose Biosynthesis Operon among Enterobacterial Species ......................................................... 116
8. Differential Expression of Cellulose among Enterobacteriaceae ...... 118
9. Coexpression of Cellulose with Curli Fimbriae ............................... 118
10. Conclusions ..................................................................................... 119

Acknowledgments .................................................................................. 119

References ............................................................................................. 120

Chapter 8: In Vitro Synthesis and Analysis of Plant (1→3)-β-D-glucans and Cellulose: A Key Step Towards the Characterization of Glucan Synthases ...................................................... 123
Vincent Bulone

1. Introduction ...................................................................................... 124
2. In Vitro Approaches for the Study of β-glucan Synthesis............... 127
2.1. Optimization of the conditions for callose and cellulose synthesis .......................................................... 127
2.2. Structural characterization of in vitro products ......................................................................................... 132
2.3. Purification of callose and cellulose synthases ......................................................................................... 140
References ......................................................................................................................................................... 142

Chapter 9: Substrate Supply for Cellulose Synthesis and its Stress Sensitivity in the Cotton Fiber ......................... 147

Candace H. Haigler

1. Introduction ..................................................................................................................................................... 148
2. Overview of Cotton Fiber Cellulose Biogenesis ......................................................................................... 149
   2.1. The role of cellulose biogenesis in cotton fiber development ................................................................. 149
   2.2. Changes in cellulose characteristics throughout cotton fiber development ............................................ 151
   2.3. The role of the microtubules in cotton fiber cellulose synthesis ............................................................ 152
   2.4. Molecular biology of cotton fiber cellulose biogenesis ........................................................................ 152
   2.5. Biochemistry of cotton fiber cellulose biogenesis .................................................................................. 153
3. Substrate Supply for Cotton Fiber Cellulose Biogenesis .............................................................................. 154
   3.1. A role for sucrose synthase ............................................................................................................... 154
4. Intrafiber Sucrose Synthesis as a Source of Carbon for Secondary Wall Cellulose Synthesis ....................... 158
5. A Role for Sucrose Phosphate Synthase in IntraFiber Cellulose Synthesis ..................................................... 160
6. Stress Sensitivity of Cellulose Synthesis .................................................................................................... 161
   Acknowledgments ......................................................................................................................................... 163
   References ....................................................................................................................................................... 163

Chapter 10: A Perspective on the Assembly of Cellulose-Synthesizing Complexes: Possible Role of KORRIGAN and Microtubules in Cellulose Synthesis in Plants ......................................................... 169

Inder M. Saxena and R. Malcolm Brown, Jr.

1. Introduction ..................................................................................................................................................... 170
2. Structure and Composition of Cellulose-Synthesizing Complexes ............................................................. 171
3. Stages in the Assembly of the Rosette Terminal Complex in Plants ................................................................ 172
4. Possible Role of KORRIGAN in the Digestion of Glucan Chains and in the Second Stage of the Assembly of the Terminal Complex .............................................................................................................. 174
5. Role of Microtubules in Cellulose Biosynthesis ........................................................................................ 177
Chapter 11: How Cellulose Synthase Density in the Plasma Membrane may Dictate Cell Wall Texture

Anne Mie Emons, Miriam Akkerman, Michel Ebskamp, Jan Schel and Bela Mulder.

1. Textures of Cellulose Microfibrils
2. Hypotheses about Cellulose Microfibril Ordering Mechanisms
   2.1. Microtubule-directed microfibril orientation
   2.2. The liquid crystalline self-assembly hypothesis
   2.3. Templated incorporation hypothesis
3. The Geometrical Model for Cellulose Microfibril Orientation
4. A role for Cortical Microtubules in Localizing Cell Wall Deposition
5. Criticism on the Geometrical Model
6. Outlook on the Verification/Falsification of the Geometrical Theory

References

Chapter 12: Cellulose-Synthesizing Complexes of a Dinoflagellate and other Unique Algae

Kazuo Okuda and Satoko Sekida

1. Introduction
2. Assembly of Cellulose Microfibrils in Dinoflagellates
3. Occurrence of Distinct TCs in the Heterokontophyta
4. Diversification in Cellulose Microfibril Assembly

References

Chapter 13: Biogenesis and Function of Cellulose in the Tunicates

Satoshi Kimura and Takao Itoh

1. Introduction
2. Texture of the Tunic in the Ascidians
3. Cellulose-Synthesizing Terminal Complexes in the Ascidians
4. A Novel Cellulose-Synthesizing Site in the Tunicates
5. Occurrence of a Cellulose Network in the Hemocoel of Ascidians
6. Structure and Function of the Tunic Cord in the Ascidians
7. Occurrence of Highly Crystalline Cellulose in the Most Primitive Tunicate, the Appendicularians
8. Origin of Cellulose Synthase in the Tunicates
9. Summary

References
### Chapter 14: Immunogold Labeling of Cellulose-Synthesizing Terminal Complexes

Takao Itoh, Satoshi Kimura, and R. Malcolm Brown, Jr.

1. Introduction ........................................................................................ 238
2. The Cellulose-Synthesizing Machinery (Terminal Complexes) ............ 238
3. Advances in the Understanding of Cellulose Synthases ...................... 241
4. How to Prove if the Rosette or Linear TC is the Cellulose-Synthesizing Machinery? ............................................ 242
5. Labeling of Freeze Fracture Replicas .................................................. 243
6. Specific Labeling of Rosette TCs .......................................................... 247
7. Specific Labeling of Linear TCs .......................................................... 249
8. The Mechanism of Labeling of Cellulose Synthases ........................... 249
9. Future Perspectives on SDS-FRL and Research in Cellulose Biosynthesis ..................................................................... 250
Acknowledgments ............................................................................... 252
References ........................................................................................... 252

### Chapter 15: Cellulose Shapes

Alfred D. French and Glenn P. Johnson

1. Introduction ........................................................................................ 257
2. Cellulose Polymorphy and Crystal Structures ..................................... 258
   2.1. The polymorphs ......................................................................... 259
   2.2. High-resolution structure determinations .................................. 260
   2.3. The dominant twofold shape in crystals ..................................... 260
   2.4. Topological nightmare ............................................................... 262
   2.5. Interdigitation ............................................................................ 263
3. Other Cellulosic Polymers ................................................................... 264
4. Information from Small Molecules in Self-Crystals and Protein-Carbohydrate Complexes ................................................. 266
   6. Crystal Structures in φ,ψ Space ........................................................ 268
      6.1. Cellulose and its oligomers ........................................................ 268
      6.2. Small molecules ......................................................................... 269
      6.3. Protein-cellodextrin complexes .................................................. 270
      6.4. Lactose-protein complexes ......................................................... 272
7. Computerized Energy Calculations Based on Molecular Models .......................................................... 273
8. Summary ............................................................................................. 278
Acknowledgments ............................................................................... 282
References ........................................................................................... 282
Table of Contents

Chapter 16: Nematic Ordered Cellulose: Its Structure and Properties ................................................................. 285
    Tetsuo Kondo
    1. Introduction .................................................................................. 285
    2. Structure of Nematic Ordered Cellulose ............................................ 287
        2.1. What is nematic ordered cellulose? ............................................. 287
        2.2. Nematic ordered α-chitin and cellulose/α-chitin blends (Kondo et al. 2004) ........................................... 294
        2.3. Another type of nematic ordered cellulose: Honeycomb-patterned cellulose (18) .......................................... 297
    3. Properties of Nematic Ordered Cellulose ........................................... 297
        3.1. The exclusive surface property of NOC and its unique application ................................................................................. 297
    4. The Future .......................................................................................... 301
    5. Materials and Methods ....................................................................... 302
        5.1. Materials .................................................................................... 302
        5.2. Water-swollen cellulose film from the DMAc/LiCl solution ....... 302
        5.3. Preparation of NOC from water-swollen cellulose films............ 303
        5.4. Preparation of NOC template in Schramm-Hestrin (SH) medium ................................................ 303
    Acknowledgments ............................................................................... 304
    References ........................................................................................... 304

Chapter 17: Biomedical Applications of Microbial Cellulose in Burn Wound Recovery ............................................. 307
    Wojciech Czaja, Alina Krystynowicz, Marek Kawecki, Krzysztof Wysota, Stanisław Sakiel, Piotr Wróblewski, Justyna Glik, Mariusz Nowak and Stanisław Bielecki
    1. Introduction ........................................................................................ 308
    2. Experimental Design ........................................................................... 309
        2.1. Never-dried MC membranes preparation .................................. 309
        2.2. Clinical trials ............................................................................. 310
    3. Clinical Outcomes ............................................................................... 311
        3.1. High conformability, moisture donation, and faster healing ...... 311
        3.2. MC is particularly useful in the treatment of facial burns .......... 315
    4. Conclusions ......................................................................................... 319
    Acknowledgments ............................................................................... 319
    References ........................................................................................... 319

Chapter 18: Cellulose as a Smart Material ............................................ 323
    Jaehwan Kim
    1. Introduction ........................................................................................ 324
    2. Experiments ....................................................................................... 327
# Table of Contents

2.1. EAPap sample preparation ........................................................ 327  
2.2. EAPap actuator performance .................................................... 329  
2.3. EAPap actuation principle ......................................................... 331  
2.4. Mechanical test of EAPap ......................................................... 336  
3. Potential Applications ......................................................................... 339  
4. Summary ............................................................................................. 341  
   Acknowledgment .................................................................................. 342  
   References ........................................................................................... 342  

Index ........................................................................................................ 345  

Color Plates ............................................................................................ 355
Cellulose was first described by Anselme Payen in 1838 as a “resistant fibrous solid that remains behind after treatment of various plant tissues with acid and ammonia.” In its simplest form, cellulose is composed of β-1,4-linked glucan chains that can be arranged in different ways giving rise to different forms of cellulose. In nature, cellulose is produced in a hierarchical manner with the glucan chains associating with each other to form crystalline and noncrystalline regions that are assembled into higher-order structures such as the microfibril. Depending on how the glucan chains associate, different crystalline forms of cellulose may be observed within the same microfibril. In nature, cellulose is generally obtained as the cellulose I crystalline form in which the glucan chains are aligned parallel to each other. Two forms of the native crystalline polymer, cellulose, Iα, and Iβ, have been shown to be present in differing amounts obtained from different sources. Other crystalline and noncrystalline forms of cellulose have also been identified, and many of these forms can be converted from one form to the other form by chemical or physical treatments. Although much is known about the structure and properties of the different forms of cellulose and these studies are still continuing, only recently has it been possible to understand the molecular basis of cellulose biosynthesis.

The chapters in the present volume highlight the wide range of topics that deal with not only the structure and biosynthesis of cellulose, but also some of the more exciting and novel applications of cellulose. Since the first identification of genes for cellulose biosynthesis in the bacterium *Acetobacter xylinum* in 1990, significant progress has been made in identifying similar genes in a large group of organisms. Polymerization of glucose residues into the β-1,4-linked glucan chains is catalyzed by cellulose synthase, and genes encoding this protein have been identified not only in most of the cellulose-producing organisms, but also in a number of other organisms suggesting that cellulose biosynthesis may be much more widespread than previously thought. Analyses of the cellulose-synthesizing genes and specifically the cellulose synthase genes has led to interesting views on the evolution of cellulose biosynthesis and the cellulose synthase gene family, and this is discussed in the chapters by Nobles and Brown, Roberts and Roberts, and Youngs et al. In plants, a large number of genes encoding cellulose synthases have
been identified by sequence and mutant analyses. Mutant, sequence, and expression analyses have provided information on the role of specific synthases during cellulose biosynthesis in the primary and the secondary cell wall in a number of plants. The significance of these studies is discussed in the chapters by Taylor and Turner (Arabidopsis), Barreiro and Dhugga (maize), and Blomqvist et al. (Populus). Whereas, the role of cellulose is well understood in plants, it is not as well understood in the other organisms where cellulose biosynthesis genes have been identified. In bacteria, cellulose is secreted as an extracellular polysaccharide, and in some cases it is shown to be associated with other components as part of bacterial biofilms. The organization of genes and regulation of cellulose biosynthesis is well understood in bacteria belonging to the pathogenic Enterobacteriaceae, and this is summarized by Römling.

One of the major challenges in understanding cellulose biosynthesis in plants is the biochemical characterization of the cellulose synthase. Although some success has been achieved in synthesis and characterization of cellulose in vitro using membrane fractions from plants, purification and structural characterization of the cellulose synthase from plants is still far from complete. In the present volume, Bulone discusses steps in the characterization of cellulose synthase and other glycosyltransferase activities. At the same time, a number of other proteins are involved in regulating cellulose synthesis in plants; Haigler covers the role of substrate supply during cellulose synthesis in the cotton fiber.

Among the more interesting aspects of cellulose biosynthesis is the phenomenon of coupled polymerization-crystallization, and it is a matter of faith in the field that the parallel arrangement of glucan chains in crystalline cellulose I result because of an organized arrangement of cellulose-synthesizing sites in the membrane. Determining how these sites are organized is a major goal for a complete understanding of cellulose biosynthesis. A perspective on how the cellulose-synthesizing complexes may be assembled in plants and the role that the membrane-bound endoglucanase (KORRIGAN) and microtubules may have in the assembly of this complex and the cellulose microfibril is discussed by Saxena and Brown. Emons et al. review a few hypotheses and a theory to explain the assembly of cellulose microfibrils and the architecture of the cell wall in plants. While it is not clear as to how the cellulose-synthesizing sites are assembled, these complexes are being identified in many more organisms. Okuda and Sekida report on the identification of cellulose-synthesizing complexes in a dinoflagellate and some unique algae and discuss the diversity and evolution of these complexes. Among animals, tunicates are unique in synthesizing cellulose. Cellulose-synthesizing complexes have been observed in a number of organisms within this group. The cellulose-synthesizing complexes and the function of cellulose in tunicates is described by Kimura and Itoh. Although cellulose-synthesizing complexes could be identified in membranes by freeze-fracture analysis, it was not possible to identify the components of these complexes. The breakthrough that led to the localization of cellulose synthases to these complexes in plants and associated proteins in bacteria came
about with freeze-fracture labeling of replicas using antibodies. This technique and its application with respect to understanding the composition of the cellulose-synthesizing complexes is discussed by Itoh et al.

The structure of cellulose is closely tied to its synthesis, and although many of the chapters discuss the synthesis of cellulose, the nature of the cellulose product is always kept in mind. A comprehensive account of the structure of cellulose and its polymorphism is provided by French and Johnson, and the structure and properties of a novel form of cellulose (nematic-ordered cellulose) is described by Kondo. Cellulose is the most abundant biomacromolecule in nature, and it is used in a variety of applications. In almost all cases, the applications of cellulose as an industrial material are dependent on its physical and chemical properties. Two chapters discuss novel applications of cellulose. Czaja et al. describe the use of microbial cellulose for applications in wound care and Kim discusses the usefulness of cellulose as a smart material, specifically the production of cellulose-based electroactive paper.

The field of cellulose research spans the interests of molecular biologists to industrial chemists, and we are pleased to provide this collection of articles to a broad audience with a central interest in cellulose. Cellulose has always been an important product for human endeavors, and we predict that with novel approaches in molecular and structural biology, the uses of this invaluable biomaterial will diversify and grow.

Cellulose has been used for centuries as an industrial material, but for the first time, this product is being seriously considered as an alternative source of energy for biofuels. The use of plants and other sources of cellulose for producing ethanol will change not only our perspective of how we look at cellulose as a natural product but also how it can be used to satisfy humankind’s appetite for energy!

We are grateful to all the authors for their excellent contributions and their patience during the assembly of this volume.

R. Malcolm Brown, Jr. and Inder M. Saxena
August, 2006
Multiple representations of cellulose (selected from chapter articles in this book)
CHAPTER 1

MANY PATHS UP THE MOUNTAIN: TRACKING THE EVOLUTION OF CELLULOSE BIOSYNTHESIS*

DAVID R. NOBLES, JR. AND R. MALCOLM BROWN, JR.,**

Section of Molecular Genetics and Microbiology, The University of Texas at Austin, Austin, TX 78712

Abstract
Available evidence supports a common ancestry for all cellulose synthases. These enzymes appear to have been a bacterial invention acquired by various eukaryotes via multiple lateral gene transfers. However, the proteins associated with regulation of cellulose biosynthesis and polymer crystallization seem to have evolved independently. Sequence divergence of eukaryotic cellulose synthases and the presence of multiple gene clusters associated with bacterial cellulose synthases are discussed in relation to the possible evolutionary pathways of cellulose biosynthesis.

Keywords
bacterial cellulose, cellulose biosynthesis, cellulose synthase, cyanobacteria, lateral gene transfer, synteny.

1 INTRODUCTION
Cellulose biosynthesis is a phenomenon observed in bacteria (proteobacteria, firmicutes, and cyanobacteria) (Mühlethaler 1949; Deinema and Zevenhuizen 1971; Napoli et al. 1975; Brown, Jr. et al. 1976; Roberts 1991; Ross et al. 1991; Nobles et al. 2001; Zogaj et al. 2001, 2003; Spiers et al. 2002) and eukaryotes including plants (Glaucophyceae, Rhodophyceae, and Chlorophyceae) (Brown, Jr. 1985), animals (urochordates) (Kimura et al. 2001b), stramenopiles (Brown, Jr.

* Supported in part by grants from the Division of Energy Biosciences, Department of Energy Grant DE-FG03-94ER20145 and the Welch Foundation Grant F-1217.
** Author for correspondence: Section of Molecular Genetics and Microbiology, The University of Texas at Austin, Austin, TX 78712, Tel: 512 471-3364; Fax: 512 471-3573; e-mail: rmbrown@mail.utexas.edu

David R. Nobles, Jr. and R. Malcolm Brown, Jr.

e al. 1969; Yamada and Miyazaki 1976) the mycetozoan Dictyostelium discoideum (Blanton et al. 2000), and Acanthamoeba spp. (Linder et al. 2002). The economical, ecological, and biological significance of cellulose production, as well as its widespread bacterial and eukaryotic distribution, have inspired a great deal of research into mechanisms of regulation and synthesis. The subject of this chapter is a much less researched area of cellulose research: specifically, the origins and evolutionary pathways of cellulose biosynthesis. In order to address this issue, it is first necessary to consider whether the trait of cellulose production among the wide range of organisms in which the process has been observed is a result of homology (similarity due to common descent) or homoplasy (similarity due to convergent evolution from independent origins).

Biosynthesis of the most common crystalline allomorph, cellulose I, requires the distinct processes of polymerization and crystallization (Benziman et al. 1980; Brett 2000; Saxena and Brown, Jr. 2005). These processes are coupled at highly ordered, membrane spanning, multienzyme terminal complexes (TCs) (Roelofsen 1958; Preston 1974). Such complexes have been observed in vascular plants (Mueller and Brown, Jr. 1980), algae (Brown, Jr. and Montezinos 1976; Giddings et al. 1980; Tsekos 1999; Schüaler et al. 2003; Okuda et al. 2004), Gluconacetobacter xylinum (synonym Acetobacter xylinum) (Zaar 1979; Kimura et al. 2001a), urochordates (Kimura et al. 2001b), and Dictyostelium discoideum (Grimson et al. 1996). A highly organized complex is believed to be necessary to produce the metastable parallel glucan chain orientation of the crystalline cellulose I allomorph and to prevent the biosynthesis of noncrystalline material and/or the folding of the nascent glucan chains into cellulose II – the most thermodynamically stable allomorph of cellulose (Brown, Jr. 1996).

The comparative morphology of TCs has been utilized as a tool for constructing an evolutionary history of cellulose biosynthesis of vascular plants and algae (Hotchkiss and Brown, Jr. 1988; Tsekos 1999). With regard to eukaryotic terminal complexes in general, it is not possible to know the validity of such inferences since the structural components of terminal complexes have not been identified. Although comparisons between such disparate groups as plants, stramenopiles, ascidians, and Dictyostelium discoideum are questionable, this methodology may have some merit when considering the evolution of eukaryotic TCs within related groups. The structural proteins comprising prokaryotic TCs are also known. However, the TC-associated proteins responsible for export and secretion of cellulose in bacteria with gram negative cell envelope architecture almost certainly have no relationship to their functional counterparts in eukaryotes. Therefore, the biosynthesis of cellulose I in bacteria and eukaryotes is in all likelihood a result of convergent evolution.

Polymerization of the β-1,4-glucan chain is catalyzed by cellulose synthase enzymes. All known cellulose synthases are family 2 processive glycosyltransferases, a ubiquitous family of enzymes which also includes chitin synthases, hyaluronan synthases, and NodC proteins (Saxena et al. 2001). Cellulose synthase sequences share a highly conserved catalytic region containing the D, D, D,
QXXRW (associated with regions U1, U2, U3, and U4, respectively), a motif characteristic of processive β glycosyltransferases (Saxena et al. 1995). Cellulose synthases undoubtedly share a common ancestry. Therefore, unlike the process of crystallization, synthesis of the β-1,4-glucan homopolymer in bacteria and eukaryota is a homologous process that forms an evolutionary link between all cellulose producing organisms.

2 SEQUENCE COMPARISONS

With the rapid growth of sequence databases, similarity searches such as BLAST have become integral tools for molecular biology and bioinformatics research. Although such searches are not suitable for inferring phylogenetic relationships, they allow rapid identification of probable homologous sequences by utilizing pairwise alignments and give a statistical measurement of the significance of the similarity displayed by the two sequences. Pairwise alignments of cellulose synthase amino acid sequences yield interesting results. Although strong sequence similarity is displayed within related phyla (e.g., proteobacterial sequences are similar to other proteobacterial sequences and vascular plant sequences are similar to those of other vascular plants), cellulose synthases demonstrate little similarity when comparisons are made between more distantly related organisms (Blanton et al. 2000; Richmond 2000; Nobles et al. 2001). This is not necessarily surprising, as one might expect homologous sequences from Arabidopsis thaliana and Escherichia coli to display divergence. What is surprising, however, is that without exception, eukaryotic sequences display greater similarity to prokaryotic sequences than to their other eukaryotic counterparts (Table 1-1). Furthermore, in each case, the expectation values demonstrate that the similarity between eukaryotic and prokaryotic sequences is statistically significant. This is particularly noticeable in the results of pairwise sequence alignments of the

**Table 1-1. Expectation values from pairwise BLAST alignments of eukaryotic and prokaryotic cellulose synthases**

<table>
<thead>
<tr>
<th></th>
<th>IRX3</th>
<th>Ddis</th>
<th>Pram</th>
<th>Cint</th>
<th>Aory</th>
<th>N7120</th>
<th>Styp</th>
<th>Smel</th>
<th>Gxyl</th>
<th>Bthu</th>
</tr>
</thead>
<tbody>
<tr>
<td>IRX3</td>
<td>8e-09</td>
<td>0.015</td>
<td>0.27</td>
<td>NSSF</td>
<td>2e-28</td>
<td>6e-06</td>
<td>0.27</td>
<td>0.35</td>
<td>0.006</td>
<td></td>
</tr>
<tr>
<td>Ddis</td>
<td>8e-09</td>
<td>4e-24</td>
<td>9e-24</td>
<td>2e-16</td>
<td>9e-42</td>
<td>2e-22</td>
<td>9e-19</td>
<td>3e-16</td>
<td>3e-23</td>
<td></td>
</tr>
<tr>
<td>Pram</td>
<td>0.015</td>
<td>4e-24</td>
<td>2e-13</td>
<td>3e-07</td>
<td>3e-28</td>
<td>3e-13</td>
<td>7e-10</td>
<td>3e-12</td>
<td>4e-14</td>
<td></td>
</tr>
<tr>
<td>Cint</td>
<td>0.27</td>
<td>9e-24</td>
<td>2e-13</td>
<td>7e-17</td>
<td>9e-33</td>
<td>8e-32</td>
<td>9e-38</td>
<td>3e-26</td>
<td>1e-36</td>
<td></td>
</tr>
<tr>
<td>Aory</td>
<td>NSSF</td>
<td>2e-16</td>
<td>3e-07</td>
<td>7e-17</td>
<td>2e-17</td>
<td>1e-19</td>
<td>4e-24</td>
<td>4e-20</td>
<td>3e-19</td>
<td></td>
</tr>
</tbody>
</table>

Results of alignments between two eukaryotic sequences are shaded. Expectations values from alignments involving bacterial sequences are unshaded. The expectation values demonstrating greatest sequence similarity are shown in bold and italicized. IRX3 – Arabidopsis thaliana (NP_197244.1), Ddis – Dictyostelium discoideum (AAF00200.1), Cint – Ciona intestinalis (BAD10864.1), Aory – Aspergillus oryzae (BAE64416.1), N7120 – Nostoc sp. PCC 7120 (NP_487797.1), Styp – Salmonella typhimurium (CAC86199.1), Smel – Sinorhizobium meliloti (NP_436917.1), Axyl – Gluconacetobacter xylinus (CAA38487.1), Bacillus thuringiensis serovar israelensis (ZP_00741731.1).
cellulose synthase from *Nostoc* sp. PCC 7120 with IRX3 (A. thaliana CesA) and the cellulose synthase from *Dictyostelium discoideum* (DcsA) in which expectation values of $2.5 \times 10^{-20}$ and $2.25 \times 10^{-18}$ times lower than the most similar eukaryotic sequences are obtained. Although no definite conclusions can be drawn from this data alone, it suggests that a mechanism other than vertical evolution is at work in the eukaryotic acquisition of cellulose synthesis.

3 EUKARYOTIC CELLULOSE SYNTHASES

3.1 The case for a cyanobacterial origin of plant cellulose synthases

Cellulose synthase amino acid sequences from various members of the Nostocales show striking similarity to plant cellulose synthase (CesA) and cellulose synthase-like protein (Csl) sequences (Nobles et al. 2001; Nobles and Brown, Jr. 2004). When a CesA (IRX3) from *A. thaliana* is compared with sequences from *Nostoc* sp. PCC 7120 (CcsA1), DcsA (the most similar nonplant eukaryotic sequence) and *Chloroflexus aurantiacus* J-10-fl (the most similar prokaryotic noncyanobacterial sequence), expectation values of 2e-28, 8e-09 and 7e-12 respectively, are generated. The significance of this similarity is augmented by multiple alignments of cellulose synthases which demonstrate sequence conservation within catalytic domains U1, U2, U3, and U4, but also reveal a large insertion region (first identified as the plant conserved and specific region or CR-P (Delmer 1999)) present in CesA, DcsA, and CcsA1 sequences that is absent in other prokaryotic sequences (Nobles et al. 2001; Roberts and Roberts 2004). Furthermore, protein trees generated by neighbor-joining, maximum likelihood, and maximum parsimony methods all demonstrate a sister grouping of cyanobacterial and vascular plant sequences similar to that observed with chloroplasts and cyanobacteria in 16s ribosomal trees (Olsen et al. 1994; Nobles et al. 2001; Nobles and Brown, Jr. 2004; Nakashima et al. 2004).

The primary endosymbiotic capture of an ancestral cyanobacterium, its subsequent evolution into a plastid, and concomitant transfer of genes to the host nucleus provide the most parsimonious explanation for the observed results. Gene transfers from organelles occur frequently and have had a profound effect on host genomes (Archibald et al. 2003; Huang et al. 2003). Indeed, it has been estimated that approximately 18% of the protein coding genes in *A. thaliana* are of cyanobacterial origin (Martin et al. 2002). Although xenologous transfer (lateral transfer from a free living organism) cannot be dismissed, gene transfer from the ancestral plastid (synologous transfer) seems the most probable pathway for the integration of a cyanobacterial cellulose synthase into an ancestor of vascular plants.

3.2 Lateral transfer of cellulose synthase in the urochordates

The urochordates are unique in that they are the only animals known to produce cellulose. This ability is especially curious given their position as basal chordates. In order to explain this, one must make one of three assumptions: (1) An early
diverging ancestor of animals possessed the ability to produce cellulose which was subsequently lost by all animals except the urochordates; (2) Cellulose biosynthesis in urochordates is the result of convergent evolution; or (3) The ability to produce cellulose was obtained via lateral gene transfer of one or more of the components necessary for cellulose biosynthesis. The occurrence of the scenario described by the first assumption would be extraordinary indeed! So extraordinary in fact, that it can likely be dismissed as far too improbable to occur. Furthermore, the identification of cellulose synthase sequences from *Ciona intestinalis* and *Ciona savignyi* as family 2 processive glycosyltransferases (Dehel et al. 2003) suggests that the process of synthesizing a β-1,4-glucan homopolymer by urochordates is homologous to that of other cellulose synthesizing organisms. Therefore, the second assumption is also rather unlikely. Further examination of *Ciona* cellulose synthase (Ci-CesA) sequences by BLAST alignment demonstrates that they have significantly greater similarity to cellulose synthase sequences from firmicutes, cyanobacteria, and proteobacteria than to other eukaryotic cellulose synthases. Unfortunately, differences in the expectation values generated by comparisons of Ci-CesAs with sequences from these distinct bacterial phyla are equivocal. Therefore, it is not possible to identify a likely point of origin for Ci-CesA based on sequence similarities. To date, phylogenetic analyses have also been unable to demonstrate a clear relationship of Ci-CesAs to any group of bacterial cellulose synthases. Thus, based on analysis of the glycosyltransferase, the identity of the donor organism remains a mystery (Nobles and Brown, Jr. 2004; Nakashima et al. 2004).

However, there is another piece to the Ci-CesA puzzle. Ci-CesA sequences have a unique feature: the C-terminus displays sequence similarity to bacterial family 6 glycosylhydrolases. The glycosylhydrolase regions of *Ciona savignyi* and *Ciona intestinalis* cellulose synthases are degenerate and therefore, probably retain no enzymatic activity (Matthysse et al. 2004). While the presence of this region strengthens the case for lateral gene transfer, it poses a problem for the identification of the sequence donor(s). A gene fusion of this type is most likely to occur as a result of the simultaneous transfer of adjacent coding regions rather than from independent acquisition of sequences and subsequent fusion. Although some species of *Streptomyces* (Actinobacteria) possess gene clusters containing putative cellulose synthases and family 6 glycosylhydrolases (Nakashima et al. 2004), these sequences show comparatively little similarity to the N and C termini of Ci-CesA, respectively. Therefore, even though all available evidence indicates that urochordates acquired their cellulose synthase through lateral transfer(s) from bacteria, it is not possible to determine the phylum of origin at this time.

### 3.3 The cellulose synthase of *Dictyostelium discoideum*

The cellulose synthase sequence (DcsA) from *D. discoideum* is far more similar to cellulose synthase sequences from *Nostoc* spp. (CesA1) than to any other sequences in the current databases. Additionally, DcsA branches as a sister clade to CesA1 in protein trees (Nobles and Brown, Jr. 2004; Nakashima et al. 2004).
These observations suggest a lateral transfer of cellulose synthase from cyanobacteria to *D. discoideum*. However, while the primary and secondary endosymbiotic events that led to the evolution of plastids in plants and algae provide a clear mechanism for the transfer of a cyanobacterial cellulose synthase to photosynthetic organisms, such a mechanism is lacking for *D. discoideum*. Cyanobacterial genes are known to exist in eukaryotes which have secondarily lost plastids. However, there is no evidence for the existence of an endosymbiotic relationship between ancestors of *D. discoideum* and a cyanobacterium. Therefore, if a lateral transfer occurred, it was likely xenologous, possibly via a food ratchet mechanism (Doolittle 1998).

4 BACTERIAL GENE CLUSTERS

4.1 Introduction

A comprehensive phylogenetic analysis of bacterial cellulose synthases has not been performed to date. The few studies which include significant taxon sampling demonstrate that bacterial sequences included in cellulose synthase protein trees generally branch in a manner similar to that observed in species trees (Nobles and Brown, Jr. 2004; Nakashima et al. 2004). Unfortunately, phylogenetic studies have been unable to demonstrate the origin of cellulose synthase among the bacterial phyla (Nobles and Brown, Jr. 2004). However, the presence of cellulose synthases in firmicutes, actinobacteria, cyanobacteria, and proteobacteria indicates that the evolution of synthases cellulose likely predates the divergence of these groups.

The proliferation of complete genome sequences in public databases provides an additional means to track the evolution of cellulose in bacteria. Conservation of operons and/or gene clusters (synteny) can be used to trace not only the history of cellulose synthase, but also its associated proteins. The existence of a few these gene clusters has been well documented. In the sections below, I would like to give a brief review of the known gene organizations and introduce two novel ones which may be linked to the eukaryotic acquisition of cellulose biosynthesis.

4.2 Characterized gene clusters

Gene clusters responsible for cellulose biosynthesis have only been extensively characterized in proteobacteria. Three archetypal gene organizations encoding proteins required for the synthesis of cellulose in the α, β, and γ subdivisions have been identified. It should be emphasized that variations of these archetypes exist and as such, the examples given here are meant to serve as paradigms to simplify the discussion of the general characteristics of gene clusters associated with cellulose biosynthesis (for a comprehensive review of these gene organizations, see Römling 2002). Based on sequence conservation, the three archetypal organizations can be divided into two groups (Figure 1-1): the Group I cluster found in the α, β, and γ subdivisions, encodes a cellulose synthase – A (*bcsA/lacsA*), as well as the B (*bcsBlacsB*) and C (*bcsClacsC*) proteins within an
operon (Saxena et al. 1994; Zogaj et al. 2001). Additionally, a family 8 glycosylhydrolase, is encoded within or in close proximity to the cellulose synthase operon (Römling 2002). The Group II gene cluster is characteristically found in the α-Proteobacteria and has been most extensively studied in Agrobacterium tumefaciens. This organization consists of two adjacent directionally opposed operons. The first operon – celABC encodes homologs of the Group I bcsAl/acsA (celA), bcsBlacsB (celB), and family 8 glycosyl hydrolase (celC). The second operon – celDE encodes proteins with no significant sequence similarity to the Group I proteins (Matthysse et al. 1995b). It should be noted however, that celD and bcsClacsC share two conserved domains: COG3118 (thioredoxin containing proteins responsible for posttranslational modifications and protein turnover) and COG4783 (putative Zn-dependent proteases containing TPR repeats) (Marchler-Bauer et al. 2005) suggesting the possibility of a similar function.

Although all proteins encoded by Group I and II gene clusters are necessary for wild-type cellulose biosynthesis, only cellulose synthases have a known function.  

**Figure 1-1.** Characterized cellulose biosynthesis gene clusters of Group I (Gluconacetobacter xylinum ATCC 23769, G. xylinum B42, and Salmonella spp.) and Group II (Agrobacterium tumefaciens A6). Identical shades and patterns represent homologous sequences. Open reading frames are not drawn to scale. Note that acsAB comprises a single open reading frame in G. xylinum ATCC 23769 (synonym AY201). Although not shown here, fusions of acsAB are also observed in the cellulose synthesis operons of G. xylinum ATCC 53582 (synonym NQ5) (Saxena et al. 1994) and Azotobacter vinelandii AvOP  

1 Previous research suggested that AcsB/BcsB had the ability to bind c-di-GMP which led to a general consensus that the B subunit had a regulatory function in cellulose biosynthesis (Amikam and Benziman 1989; Mayer et al. 1991). However, based on recent data, a regulatory role for AcsB/BcsB is questionable (Amikam and Galperin 2006).
Organisms with the Group I gene organization are believed to carry out cellulose biosynthesis without the use of lipid-linked intermediates in a process which is upregulated by the allosteric activator cyclic diguanosine monophosphate (c-di-GMP) (Aloni et al. 1982; Ross et al. 1986; Saxena et al. 1994; Römling 2002; García et al. 2004). In the case of bacteria with Group II gene organizations, the presence of lipid linked intermediates and regulation by c-di-GMP are matters of some debate (Amikam and Benziman 1989; Matthysse et al. 1995a; Ausmees et al. 2001). Despite clear differences between the gene clusters of Groups I and II, the universal presence of a family 8 glycosylhydrolase and the positioning of the \textit{acsB/bcsB/celB} genes adjacent to cellulose synthases suggests a common ancestry for these gene clusters and important roles for these proteins that transcend possible differences in mechanisms of synthesis.

5 NOVEL GENE CLUSTERS

5.1 Introduction

Although cellulose synthase sequences within closely related bacterial groups (e.g., within \(\gamma\)-proteobacteria or actinobacteria) generally display relatively high sequence conservation, they are often divergent when compared across phyla. Examination of sequenced genomes reveals the presence of alternative gene clusters that coincide with the sequence divergence of various groups of cellulose synthases. Such novel gene organizations exist in cyanobacteria, actinobacteria, chloroflexales, as well as proteobacteria and, unlike the Group I and II gene clusters described above, gene organizations are conserved across phyla. The characterization of novel gene organizations has the potential to inform current knowledge of the components necessary for cellulose biosynthesis, broaden our definitions of what constitutes cellulose, and ultimately provide a map of routes taken by organisms to utilize the \(\beta\)-1,4-homopolymer.

5.2 Group III

The Group III gene cluster is found in orders chroococcales and nostocales of cyanobacteria and in the \(\alpha\) and \(\beta\) subdivisions of proteobacteria. These clusters have not been experimentally shown to be responsible for cellulose biosynthesis and therefore, the designation of these glycosyltransferases as cellulose synthases is a putative one. Group III gene clusters encode a membrane fusion protein (MFP) of the AcrA/EmrA/HylD family adjacent and upstream of the cellulose synthase (Figure 1-2). This organization suggests the possibility of a three component system consisting of the cellulose synthase, a MFP, and an outer membrane protein OMP. In such a system, the cellulose synthase would be linked to an OMP pore by the membrane-bound periplasmic MFP and thus, form a continuous channel for export and secretion of the glucan polymer. Alternatively, some Group III clusters contain genes encoding ATP binding
cassette (ABC) transporter domains. This type of organization is characteristic of the bacterial ABC capsular polysaccharide exporter family (CPSE) in which secretion of polysaccharides is accomplished via the concerted actions of an ABC transporter, a membrane periplasmic auxiliary protein (MPA2 – analogous to the MFP associated with Type I bacterial secretion), an outer membrane auxiliary protein (OMA), and an as yet unidentified, outer membrane protein (OMP) (Silver et al. 2001). It is important to note that although the arrangement of bacterial genes in clusters is often indicative of components of a common pathway or mechanism (Korbel et al. 2004; Guerrero et al. 2005); this is by no means universally true. As such, in the absence of experimental data, any functional designation based on sequence organization must be considered speculative.

Although the presence of this gene organization in cyanobacteria and proteobacteria may indicate retention of key synthesis components from a common ancestor, lateral gene transfer cannot be ruled out. This is particularly true in the instances of *Rhizobium etli* CFN, *Rhizobium* sp. NGR234, and *A. tumefaciens* C58 where the Group III cluster is located on megaplasmids. Megaplasmids of the Rhizobiales have a significant propensity for recombination and transposition (Streit et al. 2004; Guerrero et al. 2005). Consequently, the sequences of these replicons are mosaic in nature – frequently shaped by lateral gene transfer (González et al. 2003).

In addition to the Group III gene cluster, the linear chromosome of *A. tumefaciens* C58 encodes a functional Group II gene cluster (Matthyse et al. 2005). The pNGR234 megaplasmid of *Rhizobium* sp. NGR234 also encodes an additional cellulose synthase (Streit et al. 2004) with significant similarity to the Group II cellulose synthase of *A. tumefaciens* C58 but does not possess the other conserved regions of the Group II cluster. The lack of sequence similarity between Group II and III cellulose synthases indicates that they are unlikely to be the products of gene duplication within these organisms. Rather, significant similarity with *Nostoc punctiforme* ATCC 29133 and *Synechococcus elongatus* PCC 7942 Group III cellulose synthase sequences (expectation values of 2e-56 – 1e-89 lower respectively, than observed when Group II and III sequences within the same organism are compared) indicates a possible lateral gene transfer.

### 5.3 Group IV

The Group IV gene cluster, found in nostocales, actinobacteria, and chloroflexales, encodes a cellulose synthase, an antisigma factor antagonist, and an antisigma factor. With the exception of those found in the chloroflexales, all Group IV clusters identified also contain an antisigma factor antagonist phosphatase (Figure 1-3). The ancillary proteins surrounding the cellulose synthases in this group are homologous to SpoIIA (antisigma factor antagonist), SpoIAB (antisigma factor), and SpoIIE (antisigma factor antagonist phosphatase) present in the regulatory clusters involved in the stage II sporulation of *Bacillus subtilis*.
This regulatory system acts to confine gene transcription to the prespore during asymmetric cell division (Stragier and Losick 1996; Yudkin and Clarkson 2005). The presence of these regulatory components in close association with cellulose synthases may indicate controlled expression of cellulose synthase biosynthesis in differentiated cells during asymmetric cell division in the actinobacteria (spores) and nostocales (heterocysts and/or akinetes). With regard to this possibility in the heterocysts of the nostocales, it is interesting to note the existence of alternative sigma factors which are expressed only under nitrogen limiting conditions (Brahamsha and Haselkorn 1992).

The presence of syntenic genes in cyanobacteria and actinobacteria is consistent with phylogenetic trees demonstrating a close relationship of cyanobacteria and actinobacteria (Olsen et al. 1994; Yu et al. 2005). Interestingly, the organization of the gene cluster of *C. aurantiacus* is a chimera which contains elements of the Group IV gene cluster combined with a BcsB homolog and a diguanylate cyclase. Since the chloroflexales are generally considered to have branched prior to the divergence of cyanobacteria, gram positive bacteria, and proteobacteria (Olsen et al. 1994), this organization could represent the prototypical organization for Groups I, II, and IV. However, since *C. aurantiacus* exists primarily as a photoheterotroph living in close association with cyanobacteria, the possibility of lateral gene transfer from a cyanobacterium to members of the chloroflexales cannot be discounted.
The cellulose synthase sequences of *Arthrobacter* sp. FB24 and *Corynebacterium efficiens* YS-314 are divided into two open reading frames: the first containing domain A (U1 and U2) and the second containing domain B (U3 and U4). There is no report of cellulose biosynthesis in these bacteria and therefore, it is unknown whether this split enzyme is functional. However, if the products of these two ORFs can combine to create a functional cellulose synthase, they would be useful tools for experiments to determine substrate binding properties and catalytic function of the conserved domains.

6 CONCLUDING REMARKS

Current data suggest that cellulose biosynthesis is a bacterial invention and that eukaryotes acquired the process via multiple lateral gene transfers. Bacteria and eukaryotes have independently evolved regulatory mechanisms and molecular structures to utilize the β-1,4-homopolymer synthesized by the catalytic activity of homologous cellulose synthase enzymes. The differences in accessory enzymes probably reflect not only convergent evolution to produce a cellulose I crystalline allomorph, but also inventions of alternative products such as cellulose II, noncrystalline cellulose, or nematic ordered cellulose.

As sequence databases continue to grow, it is certain that new cellulose synthase sequences and gene clusters will be identified. To be sure, increasing the library of available sequences is essential to the development of our understanding of the origin of cellulose biosynthesis and the evolutionary pathways utilized for its distribution. However, the primary challenges for researchers will be to elucidate the function of cellulose synthase associated enzymes and to characterize the cellulosic products synthesized by organisms with disparate enzymes and gene organizations. Without a firm grasp of the relationship of synthesizing components to the characteristics of the cellulosic product, we cannot hope to understand the genesis of the varied mechanisms and product morphologies we discover nor the evolutionary context from which they arose.

REFERENCES


Tracking the Evolution of Cellulose Biosynthesis


Tracking the Evolution of Cellulose Biosynthesis