Understanding the Biological Rationale for the Diversity of Cellulose-directed Carbohydrate-binding Modules in Prokaryotic Enzymes*

Anthony W. Blake‡, Lesley McCartney‡, James E. Flint§, David N. Bolam¶, Alisdair B. Boraston**, Harry J. Gilbert‡, and J. Paul Knox†

From the ‡Centre for Plant Sciences, University of Leeds, Leeds LS2 9JT, United Kingdom, the §Institute for Cell and Molecular Biosciences, University of Newcastle-upon-Tyne, Newcastle-upon-Tyne NE2 4HH, United Kingdom, and the ¶Department of Biochemistry and Microbiology, University of Victoria, Victoria, British Columbia V8W 3P6, Canada

Plant cell walls are degraded by glycoside hydrolases that often contain noncatalytic carbohydrate-binding modules (CBMs), which potentiate degradation. There are currently 11 sequence-based cellulose-directed CBM families; however, the biological significance of the structural diversity displayed by these protein modules is uncertain. Here we interrogate the capacity of eight cellulose-binding CBMs to bind to cell walls. These modules target crystalline cellulose (type A) and are located in families 1, 2a, 3a, and 10 (CBM1, CBM2a, CBM3a, and CBM10, respectively); internal regions of amorphous cellulose (type B; CBM4-1, CBM17, CBM28); and the ends of cellulose chains (type C; CBM9-2). Type A CBMs bound particularly effectively to secondary cell walls, although they also recognized primary cell walls. Type A CBM2a and CBM10, derived from the same enzyme, displayed differential binding to cell walls depending upon cell type, tissue, and taxon of origin. Type B CBMs and the type C CBM displayed much weaker binding to cell walls than type A CBMs. CBM17 bound more extensively to cell walls than CBM4-1, even though these type B modules display similar binding to amorphous cellulose in vitro. The thickened primary cell walls of celery collenchyma showed significant binding by some type B modules, indicating that in these walls the cellulose chains do not form highly ordered crystalline structures. Pectate lyase treatment of sections resulted in an increased binding of cellulose-directed CBMs, demonstrating that decloaking cellulose microfibrils of pectic polymers can increase CBM access. The differential recognition of cell walls of diverse origin provides a biological rationale for the diversity of cellulose-directed CBMs that occur in cell wall hydrolases and conversely reveals the variety of cellulose microstructures in primary and secondary cell walls.

Cellulose, a major component of plant cell walls and of the biomass of the earth, is a chemically invariant polymer comprising of up to 10,000 β-1,4-linked-glucosyl residues. In cell walls cellulose is found in the form of linear insoluble microfibrils that result from self-association of cellulose chains subsequent to synthesis by plasma membrane-located complexes. Biophysical studies have shown that cellulose microfibrils can exist in highly ordered crystalline, semi-ordered para-crystalline, and disordered noncrystalline (amorphous) states (1). Cellulose microfibrils extensively encircle plant cells and are intimately associated with other polysaccharide networks in dense composites that comprise primary and secondary cell walls. The orientation of cellulose microfibrils and their connections with other wall polysaccharide networks are key factors in the generation of cell and organ shapes and mechanical strength during plant growth and development (2).

The interactions between cellulose chains both within microfibrils and with matrix polysaccharides within the plant cell wall restrict their accessibility to enzyme attack. To overcome this problem glycoside hydrolases that degrade cell walls often have a complex molecular architecture comprising both catalytic domains and noncatalytic carbohydrate-binding modules (CBMs).2 By binding to polysaccharides, CBMs bring appended catalytic domains into intimate contact with target substrates and thus potentiate catalysis (3–7). CBMs with the capacity to bind cellulose are associated with enzymes that hydrolyze both cellulose and other cell wall polymers such as xylan, mannan, pectin, and noncellulosic β-glucans (8–12). CBMs are grouped into sequence-based families and are named after the family in which they are located (e.g. a family 4 CBM is designated CBM4; see Ref. 13 for details of the nomenclature). Currently, cellulose-binding CBMs are located in 11 of the 45 sequence based CBM families in the CAZY data base (amfb.cnrs-mrs.fr/CAZY). Family 1 CBMs are ubiquitous among fungal enzymes and are the major CBM family found in this group of organisms (14). By contrast, cellulose-binding CBMs of bacterial origin are highly diverse. CBMs from families 1, 2a (family 2 is subdivided into two subfamilies, 2a and 2b, based on an Arg/Gly polymorphism that confers specificity for cellulose and xylan, respectively) (15), 3a (family 3 is also subdivided, and only 3a modules bind to crystalline cellulose), 5, and 10 are classified as type A CBMs, which bind to crystalline polysaccha-
rides, predominantly cellulose (reviewed in Ref. 13). Although the primary sequence and three-dimensional structure of the different type A CBM families bear little similarity, the topography of the ligand recognition site is conserved presenting a flat surface comprising predominantly aromatic residues, which interact with the multiple planar cellulose chains found in crystalline cellulose (16–19). Glycosyl hydrolases can contain several type A CBMs from different families. For example, xylanase Xyn10A, the mannanase Man5B, and several cellulases from *Cellvibrio japonicus* contain a family 2a (CBM2a) and a family 10 (CBM10) CBM (20). These modules appear to have the same ligand specificity when characterized with purified polymers, although CBM10 displays ~6-fold lower affinity than CBM2a (10, 21). In contrast to the individual cell wall-degrading enzymes from aerobic bacteria such as *C. japonicus*, the enzymes from anaerobic prokaryotes, exemplified by *Clostridium* spp., associate to form a large multienzyme complex, which contains a single type A CBM from family 3a (CBM3a) (17). Type B CBMs do not bind to the planar surface of crystalline polysaccharides but recognize isolated saccharide chains. Three well characterized examples of cellulose-binding type B CBMs are found in families 4, 17, and 28. The ligand-binding sites in these protein modules comprise extended clefts or grooves that accommodate individual glycan chains in non-crystalline regions of cellulose (22–24). In *vitro*, CBM4-1, the first of two family 4 CBMs in *Cellulomonas fimis* Cel9B (25), CBM17 from *Clostridium cellulovorans* Cel5A (5), and CBM28 from *Bacillus sp.* 1139 Cel5 (5) display similar cellulose binding properties, although competition studies indicate that these modules recognize different substructures within amorphous cellulose (25). Type C or lectin-like CBMs generally recognize the termini of polysaccharides and are exemplified by CBM9-2, the second family 9 CBM in *Thermotoga maritima* xylanase 10A, which binds to the reducing ends of cellulose and xylan chains (26). The groove in the protein surface of CBM9-2 is blocked at one end and has been described as a “blind canyon” (22), explaining the specificity of this module for only the reducing ends of glycan chains.

The evolutionary rationale for multiple type A and type B cellulose-directed CBMs is currently unclear in view of the chemical homogeneity of the target ligand. It is possible that the interaction of cellulose fibers with matrix polysaccharides may alter the conformation of the surface chains in the microfibril, and thus the cellulose-binding CBMs may have evolved to recognize the distinct cellulose substructures presented by different plant cell walls. The recognition of cellulose in the context of cell walls by diverse families and types of CBM has not been systematically examined. It has been previously demonstrated that CBMs directly coupled to fluorophores or with His tag appendages can be used to investigate the localization of these proteins when bound to target polymers in the context of cell wall composites (27–29). Here, an immunofluorescence approach is used to interrogate the recognition by type A, B, and C cellulose-directed CBMs (Table 1) of polysaccharides in the context of primary and secondary cell walls in a range of species. The data show that both the type A and type B groups of cellulose-directed CBMs exhibit plant- and tissue-specific differences in cell wall recognition, providing insight into the evolutionary drivers that have led to the structural diversity displayed by these protein modules. These results not only provide a biological basis for the structural diversity of cellulose-binding CBM families in prokaryotic enzymes, but also demonstrate the utility of these protein modules in probing the microstructure of this abundant and ubiquitous glucose-based polymer in planta.

### EXPERIMENTAL PROCEDURES

*Preparation of Recombinant His-tagged CBMs*—The CBMs were produced as His-tagged recombinant proteins expressed by *Escherichia coli*. The bacterium was routinely cultured at 37 °C in Luria broth, and expression of the recombinant proteins was induced with 100 μM isopropyl-β-D-thiogalacto-

### TABLE 1

| CBM   | Type  | Source enzyme | Organism       | Reference |
|-------|-------|---------------|----------------|-----------|
| CBM1  | A     | Cellulase     | *Trichoderma reesei* | 14        |
| CBM2a | A     | Xylanase 10A  | *C. japonicus*  | 3         |
| CBM3a | A     | Scaffoldin    | *C. thermocellum* | 17        |
| CBM10 | A     | Xylanase 10A  | *C. japonicus*  | 21        |
| CBM4-1| B     | Cellulase 9B  | *C. fimi*       | 37        |
| CBM17 | B     | Cellulase 5A  | *Cellulosorbis* | 30        |
| CBM28 | B     | Cellulase 5A  | *Bacillus sp.* 1139 | 31        |
| CBM9-2| C     | Xylanase 10A  | *T. maritima*   | 26        |

*FIGURE 1.* Indirect immunofluorescence detection of His-tagged cellulose-directed CBMs binding to transverse sections of tobacco stem (sixth internode from apex of a 6-week-old plant) showing xylem vascular elements (x), cortical parenchyma (c), and epidermis (e). All of the CBMs were incubated with sections at 5 μg/ml. Crystalline cellulose-binding modules (1, 2a, 3a, and 10) bind most effectively to cell walls. There is weak recognition of secondary cell walls by amorphous cellulose-binding modules (17 and 28) but no binding by CBM4-1 or CBM9-2. The arrows point to secondary cell walls of phloem fibers. Scale bar, 100 μm.
pyranoside at temperatures and a period of induction that was CBM-specific. The CBMs, which were produced in soluble form in the cytoplasm of *E. coli*, were purified by immobilized metal ion affinity chromatography using Talon as the column matrix. Details on the expression and purification regimes have been reported previously for CBM2a (3), CBM10 (16), CBM4-1 (23), CBM9-2 (26), CBM17 (30), and CBM28 (31). To produce CBM1 and CBM3a, DNA encoding these proteins, which were a kind gift from Pierre Béguin (Pasteur Institute, Paris, France) (32) were amplified by PCR using primers that supplied NdeI and XhoI restriction sites and cloned into suitable digested pET16b to generate pGP1 and pGP2. The culture conditions used to express CBM1 and CBM3a, and the purification protocols were as described previously (3).

**Growth and Preparation of Plant Materials and Immunofluorescence Labeling by CBMs**—Tobacco, pea, and maize seedlings were grown as described previously (28, 29). Celery petals were collected locally. All of the plants were fixed in PEM buffer (50 mM PIPES, 5 mM EGTA, 5 mM MgSO₄, pH 6.9) containing 4% paraformaldehyde. Tobacco, pea, and celery sections were made following the protocol described previously (29). Prior to labeling mucilage, *Arabidopsis thaliana* seeds (C24 wild type; Lehle Seeds) were imbibed in deionized water for 1 h at room temperature before fixation.

The binding of CBMs to preparations of plant materials was assessed by a three-stage immunolabeling technique described previously (28). In some cases immunofluorescence imaging was combined with differential interference contrast imaging using an Olympus BX61 microscope equipped with differential interference contrast optics and epifluorescence irradiation. The images were captured with a Hamamatsu ORCA 285 camera and prepared with Improvision Volocity software. In all cases, unless otherwise stated, CBMs were incubated with sections at 5 μg protein/ml. The images shown are representative of observations on at least three distinct plants in each case.

In some cases the cell walls of sectioned plant materials were treated with enzyme prior to CBM labeling. Pectic polymers were hydrolyzed by incubating sections of celery petiole with 10 μg/ml of the *C. japonicus* pectate lyase CJPel10A (12) for 2 h at room temperature in 50 mM CAPS, 2 mM CaCl₂ buffer, pH 10, prior to CBM labeling.

**RESULTS**

**Type A Cellulose-directed CBMs Bind More Effectively to Cell Walls in Sections of Tobacco Stems than Type B CBMs or a Type C CBM**—Previous studies have shown that the binding of recombinant His-tagged CBMs, at a concentration range of 1–5 μg/ml, to plant cell walls can be detected using an indirect triple labeling immunofluorescence procedure (His-tagged CBM, anti-His mouse-Ig, anti-mouse Ig fluorescein isothiocyanate) (27, 28). An initial study of CBM binding to transverse sections of a tobacco stem internode (sixth internode from the apex of a 6-week-old tobacco plant) indicated that type A CBMs from families 1, 2a, 3a, and 10, which recognize crystalline cellulose, bound more effectively to cell walls than type B CBMs from families 4-1, 17, and 28, which interact with internal regions of discrete β-glucan chains, and the type C CBM from family 9-2, which recognizes the reducing ends of polysaccharides (Fig. 1). The CBMs target primarily the secondary cell walls of the xylem vessel elements and phloem fibers. The most effective binders to both secondary cell walls and primary cell walls of epidermis and cortical cells were CBM3a and CBM2a with CBM1 and CBM10 displaying some recognition of these structures, albeit more weakly. This binding to plant cell walls is in agreement with known affinities of these CBMs for purified crystalline cellulose (3, 21, 33). CBM10, while binding effectively to secondary cell walls, interacted weakly with cortical parenchyma primary cell walls where binding was localized primarily to wall regions adjacent to cell junctions (see below). Of the noncrystalline-binding CBMs, only CBM17 and CBM28 displayed binding (albeit weak) to secondary cell walls, and in addition, CBM17 interacted weakly with the primary cell walls of cortical parenchyma. CBM4-1 showed no binding to any cell walls and CBM28, which interacted with internal regions of discrete β-glucan chains, and the type C CBM from family 9-2, which recognizes the reducing ends of polysaccharides (Fig. 1). The CBMs target primarily the secondary cell walls of the xylem vessel elements and phloem fibers. The most effective binders to both secondary cell walls and primary cell walls of epidermis and cortical cells were CBM3a and CBM2a with CBM1 and CBM10 displaying some recognition of these structures, albeit more weakly. This binding to plant cell walls is in agreement with known affinities of these CBMs for purified crystalline cellulose (3, 21, 33). CBM10, while binding effectively to secondary cell walls, interacted weakly with cortical parenchyma primary cell walls where binding was localized primarily to wall regions adjacent to cell junctions (see below). Of the noncrystalline-binding CBMs, only CBM17 and CBM28 displayed binding (albeit weak) to secondary cell walls, and in addition, CBM17 interacted weakly with the primary cell walls of cortical parenchyma. CBM4-1 showed no binding to any cell walls and CBM28, which interacted with internal regions of discrete β-glucan chains, and the type C CBM from family 9-2, which recognizes the reducing ends of polysaccharides (Fig. 1).
CBM Recognition of Arabidopsis Seed Mucilage—To explore the capacity of CBMs to bind cellulose in different contexts, these protein modules were incubated with intact, aldehyde-fixed seeds of Arabidopsis in a whole mount immunolabeling technique previously described (34). Although the mucilage that is extruded from the seed coat of epidermal cells upon incubation in water is not a compact composite, cellulose is present with extensive pectin components (34). Incubation of intact seeds with 5 μg/ml CBMs indicated very strong binding of the mucilage emerging from the seed surface by CBM1, CBM2a, and CBM3a (Fig. 5). CBM2a and CBM3a produced similar patterns of binding with recognition of distinct tufts of mucilage, whereas CBM1 binding appeared to be more diffuse. CBM10 bound only weakly to a relatively thin region of mucilage-again in significant contrast to CBM2a. Of the type B CBMs families 4-1 and 17 showed no significant binding to the mucilage or seed surface. However, CBM28 bound in a manner equivalent to CBM10, as did the type C end binder CBM9-2 (Fig. 5).

Recognition of Primary Cell Walls of Celery Colchencyma by Type B CBMs and CBM9-2—When incubated at up to 500 μg/ml protein with sections of tobacco and pea internodes CBM4-1 showed no binding detectable by the indirect immunofluorescence procedure. To extend understanding of cell wall recognition by CBMs, sections of celery petiole, which exhibit extensive thickened primary cell walls of colchencyma bundles, were probed with the cellulose-directed protein modules. All type A CBMs bound effectively throughout the petiole, including the primary cell walls of the epidermis, colchencyma, and parenchyma. This binding is

most effective recognition of these cell walls was at cell junctions adjacent to intercellular spaces. Binding of CBMs 3a, 17, 28, and 9-2 are shown in Fig. 2. CBM3a bound evenly to all primary cell walls, CBM17 showed weak punctate labeling of cortical parenchyma, which was associated with cell walls at intercellular junctions. Similar labeling, albeit very weak, was also observed with CBM28 and CBM9-2 for pith parenchyma alone (Fig. 2). Increasing the concentration of these CBMs significantly above 5 μg/ml did not appreciably increase the fluorescence signal, indicating that the weak labeling pattern reflects a low number of available binding sites rather than weak affinity per se.

Equivalent histochemical analysis of CBM recognition of more immature internodes of tobacco stems indicated a general trend of reduced recognition of cell walls in younger internodes by these protein modules. This is shown by the recognition of sections in the region of xylem vascular elements of internodes 2, 4, and 6 (from the apex of a 6-week-old plant) by CBM2a, CBM10, and CBM17 (Fig. 3). CBM2a and CBM10 bound to secondary cell walls of xylem elements at all three internodes. In contrast, the binding of CBM17 to secondary cell walls increased from no binding at internode 2 to strong binding at internode 6. CBM2a binding to primary cell walls also increased from internode 2 to internode 6 and was largely restricted to regions at cell junctions at internodes 2 and 4. CBM10 only bound to primary cell walls at internode 6.

Recognition of Cell Walls by CBM2a and CBM10 Varies between Plant Species—CBMs 2a and 10 are derived from the same enzyme-xylanase Xyn10A from C. japonicus (3, 21). In vitro analysis has indicated that CBM10 has a lower affinity for crystalline cellulose than CBM2a (21). This is consistent with the observation that CBM2a displayed more extensive binding to the primary cell walls of the cortical parenchyma and epidermis in tobacco stems than CBM10 (Fig. 1). The binding of these two CBMs under equivalent conditions was explored in more detail. Analysis of equivalent sections of internode 4 indicated that CBM2a bound effectively to primary cell walls and regions of material forming the junction of adhered and unadhered cell walls at intercellular spaces (Fig. 4a), whereas, significantly, CBM10 recognized mainly material adjacent to the intercellular spaces as shown in Fig. 4b. In contrast, CBM10 and 2a displayed similar binding to all primary cell walls of pith parenchyma (Fig. 4, c and d). These observations indicate differences in primary walls of cortical and pith parenchyma cells. Extending the study to taxonomically diverse material indicated that CBM2a bound effectively to cell walls of maize coleoptile, whereas CBM10 did not (Fig. 4, e and f). However, in a section of pea stem, although both CBMs bound to the secondary cell walls of the xylem vessel elements, CBM10 bound to the primary cell walls of the surrounding parenchyma much more effectively than CBM2a (Fig. 4, g and h).

CBM Recognition of Arabidopsis Seed Mucilage—To explore the capacity of CBMs to bind cellulose in different contexts, these protein modules were incubated with intact, aldehyde-fixed seeds of Arabidopsis in a whole mount immunolabeling technique previously described (34). Although the mucilage that is extruded from the seed coat of epidermal cells upon incubation in water is not a compact composite, cellulose is present with extensive pectin components (34). Incubation of intact seeds with 5 μg/ml CBMs indicated very strong binding of the mucilage emerging from the seed surface by CBM1, CBM2a, and CBM3a (Fig. 5). CBM2a and CBM3a produced similar patterns of binding with recognition of distinct tufts of mucilage, whereas CBM1 binding appeared to be more diffuse. CBM10 bound only weakly to a relatively thin region of mucilage-again in significant contrast to CBM2a. Of the type B CBMs families 4-1 and 17 showed no significant binding to the mucilage or seed surface. However, CBM28 bound in a manner equivalent to CBM10, as did the type C end binder CBM9-2 (Fig. 5).

Recognition of Primary Cell Walls of Celery Colchencyma by Type B CBMs and CBM9-2—When incubated at up to 500 μg/ml protein with sections of tobacco and pea internodes CBM4-1 showed no binding detectable by the indirect immunofluorescence procedure. To extend understanding of cell wall recognition by CBMs, sections of celery petiole, which exhibit extensive thickened primary cell walls of colchencyma bundles, were probed with the cellulose-directed protein modules. All type A CBMs bound effectively throughout the petiole, including the primary cell walls of the epidermis, colchencyma, and parenchyma. This binding is
Cellulose-directed Carbohydrate-binding Modules

SEPTEMBER 29, 2006•VOLUME 281 • NUMBER 39
JOURNAL OF BIOLOGICAL CHEMISTRY

exemplified by the binding profile of CBM3a (Fig. 6a). Type B modules CBM4-1 and CBM17 bound effectively to the epidermal and collenchyma cell walls but not surrounding parenchyma pri-

monary cell walls (Fig. 6b). The binding of CBM4-1 to a transverse section of collenchyma bundle was not equal, and certain inner regions of the bundle showed stronger fluorescence signals (Fig. 6, b and c). Analysis of binding in conjunction with differential interference contrast microscopy indicated that CBM4-1 bound most effectively to inner regions of the thickened cell walls and that this binding was often punctate (Fig. 6c). CBM9-2 showed a distinctive pattern of recognition, binding strongly to the inner regions of epidermal cell walls and also to inner regions of certain cells of the collenchyma bundles (Fig. 6d). In this case, the cells were adjacent to the epidermis, and binding appeared as a decreasing gradient away from the epidermis, somewhat in contrast to the pattern observed with CBM4-1 binding.

Pectinase Treatment of Celery Petiole Sections Results in Increased CBM Recognition of Cell Walls—In vivo, the interaction of CBMs from saprophytes with their target ligands is a progressive process in which these modules interact with a range of cell walls that are at different stages of degradation. The appended enzymes are acting in concert with a battery of other cell wall-degrading enzymes such as pectinases. To explore the capacity of cellulose-directed CBMs to interact with partially degraded cell walls, the binding of these modules to sections pretreated with a pectate lyase to remove the major pectic homogalacturonan component was assessed. Sections of celery petiole were used for these experiments because they contain a wide range of cellulose-containing cell wall types (see above). As shown in Fig. 7 the pretreatment of sections with the pectate lyase CjPel10A resulted in the loss of the abundant JIM5 homogalacturonan epitope (35). The same treatment increased the binding of cellulose-directed CBMs but to differing extents. Of particular interest was the large increase in binding to all primary cell walls by CBM17 and CBM4-1 after pectinase treatment. The binding of CBM17 to celery collenchyma with and without a pectate lyase pretreatment are shown in Fig. 7. CBM17 bound strongly across enzyme-treated thickened collenchyma cell walls although there was some indication that regions at cell junctions/middle lamella between collenchyma cells were less strongly labeled (Fig. 7, inset). The Type B CBM28 bound to collenchyma cell walls weakly; however, after pectinase treatment there was a slight increase in the extent of binding particularly to discrete regions of collenchyma inner cell walls that was often at a cell corner (Fig. 7). CBM9-2 binding was increased slightly by treatment with pectate lyase but retained the same distinctive pattern of epidermal cells and outer collenchyma cells (data not shown).

DISCUSSION

This study shows that type A CBMs directed against crystalline cellulose bind more effectively to both primary and secondary cell walls than the corresponding type B modules, which target amorphous cellulose. Type A CBMs, however, exemplified by CBM2a and CBM10, display both similarities and differences in cell wall specificity that are plant- and tissue-specific. Although the two CBMs exhibit similar specificities for tobacco secondary cell walls and tobacco pith parenchyma primary cell walls, CBM2a has a greater capacity to bind to tobacco cortical parenchyma and maize sheath coleoptiles, whereas CBM10 displays a preference for pea stem parenchyma, indi-
cating that these two CBMs are optimized to bind to distinct features of the microstructure of cell walls. By contrast the CBM3a module displays a broader specificity for cell walls than either the family 2a or 10 modules.

The specificities of type A CBMs are consistent with their microbial origin. CBM3a is the only known type A module in the cellulosome, a powerful multienzyme plant cell wall-degrading complex produced by the anaerobic bacterium Clostridium thermocellum (36). Thus, to target the cellulosome to the array of crystalline cellulosic material that is likely to be present in the C. thermocellum environment, there will be a strong selection pressure for the type A module in this complex to display broad specificity for the microfibrils present in a range of cell walls. CBM2a and CBM10 employed in this study are derived from the C. japonicus xylanase Xyn10A, and the differences in the specificity displayed by these modules may optimize the potential of the appended catalytic module to attack cell walls derived from a range of plant species and tissues. The recent resolution of the C. japonicus genome3 has revealed that the type A CBMs produced by this organism are derived from families 2a, 10, and 5, although the latter modules are only found in chitinases and thus do not target plant cell walls. Enzymes containing CBM2a and/or CBM10 modules include mannanases, cellulases, xylan-degrading enzymes (both xylanases and side chain-cleaving enzymes), and pectinases (pectate and rhamnogalacturonan lyases). Given the high level of sequence conservation displayed by different members of family 2a and 10 modules, it is highly likely that ligand specificity is invariant in these families. This view is supported by in vitro studies that have shown that family 10 modules derived from mannanases, xylanases, and cellulases display the same ligand specificites and affinities, as do family 2a modules from the corresponding enzymes (10, 21). Thus, by recruiting different combinations of family 2a and 10 CBMs, the plant cell wall hydrolases expressed by C. japonicus will be targeted to an increased range of cell wall materials. It is interesting to note that several of the plant cell wall hydrolases produced by the marine bacterium Saccharophagus degradans 2-40, a particularly efficient degrader of plant biomass, also contain both CBM2 (most of the 19 CBM2s in this bacterium are in subfamily 2a that binds crystalline cellulose) and CBM10 (afmb.cnrs-mrs.fr/CAZY/), pointing to a generic role for these protein modules in the saccharification of these composite structures by aerobic prokaryotes. By contrast, fungal glycoside hydrolases contain only a single CBM belonging to family 1. Although it is difficult to understand with any certainty why fungal enzymes have a relatively simply molecular architecture, it is possible that as the hyphae from these organisms grow into the cell wall, the delivery of the secreted enzyme to its target substrate does not require an elaborate repertoire of CBMs.

The type B CBMs also display differences in specificity despite binding to purified amorphous cellulose with similar affinities (30, 31, 37). CBM4-1 did not bind significantly to any material examined here except collenchyma cell walls of

3 K. E. Nelson, H. J. Gilbert, L. E. Tailford, and E. F. Mongodin, manuscript in preparation.

FIGURE 6. Indirect immunofluorescence detection of CBMs binding to collenchyma bundles in transverse sections of celery petioles. a, CBM3a bound to all primary cell walls of the epidermis (e), collenchyma (c), and parenchyma (p). b, CBM4-1 bound only to epidermal and collenchyma cell walls. CBM4-1 bound particularly effectively to cells in the inner region of a collenchyma bundle as indicated by arrows. c, higher magnification micrograph of immunofluorescence combined with differential interference contrast microscopy indicated that this CBM4-1 binding was to inner regions of collenchyma cell walls. d, CBM9-2 bound strongly to inner regions of epidermal cell walls and certain subepidermal collenchyma cells as indicated by arrows. i, interior of collenchyma cells, *, thickened collenchyma cell walls. Scale bars, 100 μm.
celery petioles. CBM17 also binds similarly to cell walls of celery collenchyma tissue, but the module also showed some recognition, albeit weak, of secondary cell walls and the intercellular junctions of tobacco stem primary cell walls. CBM28, shows similar recognition of tobacco stem sections to CBM17, while also binding to Arabidopsis seed mucilage and collenchyma cell walls. These differential patterns of binding again indicate the recognition of distinct features of cellulose by these three type B CBMs.

FIGURE 7. Indirect immunofluorescence labeling of celery collenchyma tissue before and after treatment of sections with pectate lyase. The binding of the homogalacturonan antibody JIM5 was abolished by pectate lyase treatment. Conversely, enzyme treatment increased CBM17 binding to primary cell walls of collenchyma and extended binding to associated parenchyma. The inset shows a detail of binding to collenchyma cell walls with reduced fluorescence in regions of middle lamellae. CBM28 bound only weakly to collenchyma cell walls. After pectate lyase treatment labeling remained weak but with strong binding to isolated regions of collenchyma cells (arrows). The inset of immunofluorescence combined with bright field shows that points of fluorescence are associated with inner regions/ corners of collenchyma cell walls (arrows). e, epidermis; c, collenchyma bundle; p, parenchyma; i, interior of collenchyma cell. * indicates junction/middle lamellae of thickened collenchyma cell walls. Scale bars for main panels, 100 µm. Scale bars for inset images, 20 µm.

Celery-petiole-related Carbohydrate-binding Modules

of the enzyme in similar fashion to the family 2a and 10 modules discussed above. The poor recognition of plant material by CBM4-1 may point to a highly specialized role for the enzyme containing this module. It is possible, however, that in vivo, CBM4-1 actually directs the endoglucanase toward plant material generated at the latter stages of the degradative process, either isolated regions of amorphous cellulose or even cellulooligosaccharides. Because CBM9-2 recognizes the reducing end of β-glucan chains, it is surprising that binding of this protein to any of the nondegraded plant cell walls used here was observed. The interaction of this module with the epidermal cell walls of celery and the mucilage of Arabidopsis seeds indicates that they contain a significant number of free reducing ends.

Topology of the Ligand-binding Site of the Cellulose-directed CBMs—The structural basis for the variation in cell wall recognition displayed by the cellulose-binding CBMs is an intriguing question. The four type A CBMs all present a planar hydrophobic face (Fig. 8) that likely binds to the 110 face of cellulose crystals (33), with the three surface aromatic residues in these proteins stacking against the pyranose rings of the glucose molecules. It is possible, however, that the polar side chains in the binding face of these CBMs may interact with the matrix polysaccharides that are in intimate contact with the cellulose microfibrils, and these hydrogen bonds may confer the variation in specificity observed.

In CBM4-1 the target ligand is accommodated in a deep cleft (Fig. 8) and requires access to the whole width of a cellulose chain for binding to occur (23). Furthermore, the protein contains two tyrosine residues that stack against the α and β face of the glucose at the reducing end of cellopentaose (Fig. 8), and these residues have been shown to play a key role in ligand recognition (42). In view of the complex interactions between the cellulose chains that comprise microfibrils and between these structures and the matrix hemicellulose and pectin polymers of the plant cell wall, it is likely that at least one face of the glucose molecules in the β-glucan will be in close association

| Type | Binding Characteristics |
|------|--------------------------|
| CBM17 | Primary cell walls of collenchyma and secondary cell walls |
| CBM28 | Primary cell walls of collenchyma and mucilage |
| JIM5 | Pectate lyase treatment abolished binding |

Scale bars for main panels: 100 µm. Scale bars for inset images: 20 µm.
with other polysaccharides. Thus, access to cellulose in this form is difficult, whereas the polysaccharide remains in the context of a microfibril within a composite, providing an explanation for the limited interaction of the family 4 module for the cell walls investigated in this study. Indeed, the corresponding family 4 module that binds xylans also displays limited recognition of plant cell walls (29), despite displaying high affinity for the purified polysaccharide. It is possible, therefore, that glycoside hydrolases containing family 4 CBMs may be targeted toward polysaccharides that are not in intimate association with the plant cell wall but to very disordered or even soluble forms of their target polymers. The more extensive recognition of cell walls by the CBM17 is consistent with its shallow ligand-binding cleft (Fig. 8). This site is adapted to interact with cellulose chains that are in intimate contact with other polysaccharides within the cell wall, and it has been postulated that CBM17 may therefore be able to bind to regions at the transition between crystalline and noncrystalline cellulose (22). CBM28 also displays a shallow cleft, although there are subtle differences in the topography of its cellulose-binding site compared with CBM17 (Fig. 8). Thus, the distinct conformation adopted by cellulose chains that interact with CBM17 and CBM28, respectively, provides an explanation for their differential recognition of plant cell walls.

Conclusions—In terms of CBM biology, this report provides a basis for the diversity of cellulose-directed CBMs that are appended to a wide range of cell wall hydrolases. It is possible that CBMs have evolved and diverged to maximize recognition of the taxonomic diversity of cell wall polymers and the architectural contexts of cellulose with cross-linking glycans, pectins, and other cell wall constituents. This recognition must also include the varied cell wall states of disruption and degradation and the diverse microenvironments in which the host organisms exist. Several observations remain intriguing, such as the capacity to bind to material at intercellular spaces in expanded middle lamellae, an area that is often viewed as free of cellulose. The nature of this material remains uncertain, although pectinase experiments indicate that it is not pectin-based (data not shown). The availability of these proteins also greatly extends the number of probes that can be used in the analysis of tissue and taxonomic differences of plant cell walls and of modifications resulting from physiological impacts or mutational events.

Acknowledgment—We acknowledge Pierre Bégardin for supplying the plasmid encoding CBM3a.

REFERENCES

1. Tomme, P., Warren, R. A., and Gilkes, N. R. (1995) Adv. Microb. Physiol. 37, 1–81
2. Cosgrove, D. J. (2005) Nat. Rev. Mol. Cell. Biol. 6, 850–861
3. Bolam, D. N., Ciruela, A., McQueen-Mason, S., Simpson, P., Williamson, M. P., Rixon, J. E., Boraston, A., Hazlewood, G. P., and Gilbert, H. J. (1998) Biochem. J. 331, 775–781
4. Black, G. W., Rixon, J. E., Clarke, J. H., Hazlewood, G. P., Ferreira, L. M., Bolam, D. N., and Gilbert, H. J. (1997) J. Biotechnol. 57, 59–69
5. Boraston, A. B., Kwan, E., Chiu, P., Warren, R. A., and Kilburn, D. G. (2003) J. Biol. Chem. 278, 6120–6127
6. Maglione, G., Matsushita, O., Russell, J. B., and Wilson, D. B. (1992) Appl. Environ. Microbiol. 58, 3593–3597
7. Tomme, P., Van Tilbeurgh, H., Pettersson, G., Van Damme, J., Vandekerckhove, J., Knowles, J., Teeri, T., and Claeyssens, M. (1988) Eur. J. Biochem. 170, 575–581
8. Kellett, L. E., Poole, D. M., Ferreira, L. M., Durrant, A. J., Hazlewood, G. P., and Gilbert, H. J. (1990) Biochem. J. 272, 369–376
9. Gilbert, H. J., Hall, J., Hazlewood, G. P., and Ferreira, L. M. (1990) *Mol. Microbiol.* **4**, 759–767
10. Hogg, D., Pell, G., Dupree, P., Goubet, F., Martin-Orue, S. M., Armand, S., and Gilbert, H. J. (2003) *Biochem. J.* **371**, 1027–1043
11. McKie, V. A., Vincken, J. P., Voragen, A. G., van den Broek, L. A., Stimson, E., and Gilbert, H. J. (2001) *Biochem. J.* **355**, 167–177
12. Brown, I. E., Mallen, M. H., Charnock, S. J., Davies, G. J., and Black, G. W. (2001) *Biochem. J.* **355**, 155–165
13. Boraston, A. B., Bolam, D. N., Gilbert, H. J., and Davies, G. J. (2004) *Biochem. J.* **382**, 769–781
14. Reinikainen, T., Ruohonen, L., Nevanen, T., Laaksonen, L., Kraulis, P., Jones, T. A., Knowles, J. K., and Teeri, T. T. (1992) *Proteins* **14**, 475–482
15. Simpson, P. J., Xie, H., Bolam, D. N., Gilbert, H. J., and Williamson, M. P. (2000) *J. Biol. Chem.* **275**, 41137–41142
16. Raghothama, S., Simpson, P. J., Szabo, L., Nagy, T., Gilbert, H. J., and Williamson, M. P. (2000) *Biochemistry* **39**, 978–984
17. Tormo, J., Lamed, R., Chirino, A. J., Morag, E., Bayer, E. A., Shoham, Y., and Teeri, T. T. (1996) *EMBO J.* **15**, 5739–5751
18. Kraulis, J., Clore, G. M., Nilges, M., Jones, T. A., Pettersson, G., Knowles, J., and Gronenborn, A. M. (1989) *Biochemistry* **28**, 7241–7257
19. Xu, G. Y., Ong, E., Gilkes, N. R., Kilburn, D. G., Carver, J. P., Kay, L. E., and Harvey, T. S. (1995) *Biochemistry* **34**, 6993–7009
20. Ferreira, L. M., Durrant, A. J., Hall, J., Hazlewood, G. P., and Gilbert, H. J. (1990) *Biochem. J.* **269**, 261–264
21. Gill, J., Rixon, J. E., Bolam, D. N., McQueen-Mason, S., Simpson, P. J., Williamson, M. P., Hazlewood, G. P., and Gilbert, H. J. (1999) *Biochem. J.* **342**, 473–480
22. Notenboom, V., Boraston, A. B., Chiu, P., Freelove, A. C., Kilburn, D. G., and Rose, D. R. (2001) *J. Mol. Biol.* **314**, 797–806
23. Boraston, A. B., Nurizzo, D., Notenboom, V., Ducros, V., Rose, D. R., Kilburn, D. G., and Davies, G. J. (2002) *J. Mol. Biol.* **319**, 1143–1156
24. Jamal, S., Nurizzo, D., Boraston, A. B., and Davies, G. J. (2004) *J. Mol. Biol.* **339**, 253–258
25. McLean, B. W., Boraston, A. B., Brouwer, D., Sanaie, N., Pyfe, C. A., Warren, R. A., Kilburn, D. G., and Haynes, C. A. (2002) *J. Biol. Chem.* **277**, 50245–50254
26. Boraston, A. B., Creagh, A. L., Alam, M. M., Kormos, J. M., Tomme, P., Haynes, C. A., Warren, R. A., and Kilburn, D. G. (2001) *Biochemistry* **40**, 6240–6247
27. Hilden, L., Daniel, G., and Johansson, G. (2003) *Biotechnol. Lett.* **25**, 553–558
28. McCartney, L., Gilbert, H. J., Bolam, D. N., Boraston, A. B., and Knox, J. P. (2004) *Anal. Biochem.* **326**, 49–54
29. McCartney, L., Blake, A. W., Flint, J., Bolam, D. N., Boraston, A. B., Gilbert, H. J., and Knox, J. P. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103**, 4765–4770
30. Boraston, A. B., Chiu, P., Warren, R. A., and Kilburn, D. G. (2000) *Biochemistry* **39**, 11129–11136
31. Boraston, A. B., Ghaffari, M., Warren, R. A., and Kilburn, D. G. (2002) *Biochem. J.* **361**, 35–40
32. Carrard, G., Koivula, A., Soderlund, H., and Beguin, P. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97**, 10342–10347
33. Lehtio, J., Sugiyama, J., Gustavsson, M., Fransson, L., Linder, M., and Teeri, T. T. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100**, 484–489
34. Williams, G. G., McWilliam, L., and Knox, J. P. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **100**, 484–489
35. Clausen, M. H., Williams, G. G., and Knox, J. P. (2003) *Carbohydr. Res.* **338**, 1797–1800
36. Bayer, E. A., Belaich, J. P., Shoham, Y., and Lamed, R. (2004) *Annu. Rev. Microbiol.* **58**, 521–554
37. Tomme, P., Creagh, A. L., Kilburn, D. G., and Haynes, C. A. (1996) *Biochemistry* **35**, 13885–13894
38. McLean, B. W., Bray, M. R., Boraston, A. B., Gilkes, N. R., Haynes, C. A., and Kilburn, D. G. (2000) *Protein Eng.* **13**, 801–809
39. Murashima, K., Kosugi, A., and Doi, R. H. (2005) *J. Bacteriol.* **187**, 7146–7149
40. Linder, M., Mattinen, M. L., Kontteli, M., Lindeberg, G., Stahlberg, J., Drakenberg, T., Reinikainen, T., Pettersson, G., and Annila, A. (1995) *Protein Sci.* **4**, 1056–1064
41. Ponyi, T., Szabo, L., Nagy, T., Orosz, L., Simpson, P. J., Williamson, M. P., and Gilbert, H. J. (2000) *Biochemistry* **39**, 985–991
42. Kormos, J., Johnson, P. E., Brun, E., Tomme, P., McIntosh, L. P., Haynes, C. A., and Kilburn, D. G. (2000) *Biochemistry* **39**, 8844–8852