

Disulfide Arrangement and Functional Domains of β -1,4-Endoglucanase E5 from *Thermomonospora fusca*[†]

Kathleen McGinnis and David B. Wilson*

Section of Biochemistry, Molecular, and Cell Biology, Cornell University, Ithaca, New York 14853

Received October 22, 1992; Revised Manuscript Received March 24, 1993

ABSTRACT: *Thermomonospora fusca* cellulase E5 contains six cysteine residues. The number and location of the disulfide bonds and the effect of reduction of the disulfides and modification of the resulting half-cystine residues on enzymatic activity were determined. No free sulfhydryl groups were found in E5. Reduction and subsequent labeling with iodoacetamide of E5 and of an enzymatically active 32-kDa proteolytic derivative of E5 (E5_{cd}) showed that one of the three disulfides is accessible to reduction under nondenatured conditions while the other two are not accessible. Full reduction of the disulfides and complete carboxymethylation of the six cysteines decrease the specific activity of E5 on CMC by more than half, but reduction of only the exposed disulfide bond does not affect enzymatic activity or binding of E5 to cellulose. A 14-kDa proteolytic fragment of E5 containing 120 amino acids from the N-terminus of the protein was shown to bind to crystalline cellulose. This confirms earlier evidence that the cellulose binding domain of E5 is located at the N-terminus of the protein. This 14-kDa fragment contains the accessible disulfide bond involving Cys93 and Cys100. The location of the two disulfide bonds in the other fragment (E5_{cd}) was determined by cleaving it with cyanogen bromide under conditions that left the disulfide bonds intact. The resulting peptides were separated under both nonreducing and reducing conditions using RP-HPLC. Amino acid analysis of peptide peaks indicated that one disulfide linkage in E5_{cd} joins Cys138 to Cys143 while the other joins Cys166 to Cys406.

Thermomonospora fusca is a thermophilic, filamentous, soil bacterium common in rotting organic material that was originally isolated from decaying wood (Bellamy, 1977). This organism, an actinomycete, grows aerobically at 55 °C. Actinomycetes appear to be the major degraders of cellulose and hemicellulose in heated plant residues, such as compost heaps (McCarty et al., 1988). Several very active and stable cellulolytic enzymes are produced and secreted by *T. fusca* (Calza et al., 1985; Wilson, 1988, 1992).

The present investigation involves *T. fusca* E5, an endocellulase that has a molecular weight of 46 300 as determined from the DNA sequence of its gene (Lao et al., 1991). E5 has high activity on both soluble and insoluble cellulose substrates (Wilson, 1992). In this paper, the organization and role of the six cysteine residues in E5 were determined. We also showed that E5 contains a cellulose binding domain which can be separated from the catalytic domain by proteolytic cleavage of the native enzyme.

MATERIALS AND METHODS

Materials. The sources of the chemicals used are listed in McGinnis and Wilson (1993).

Enzyme Preparation. Plasmid pGG74 carrying the *T. fusca* E5 gene was used to transform *Streptomyces lividans* (Ghangas & Wilson, 1987). *S. lividans* cultures were grown and harvested and the E5 enzyme was purified from the supernatant according to methods previously described (Walker et al., 1992). E5_{cd} was purified using essentially the same procedure employed for E5 except that plasmid pE5-38 was used.

Enzyme Assays. Cellulase assays and CMC overlays were carried out as described in McGinnis and Wilson (1993).

Avicel Binding. Twenty micrograms of protease-digested E5 was added to 10 mg of Avicel in 200 μ L of 0.05 M KPi, pH 6.0, to give a final concentration of 5% Avicel. This sample was incubated 18 h at room temperature on a rotator. To remove the Avicel, the sample was centrifuged 15 min in a microfuge. The supernatants were removed and analyzed by SDS-polyacrylamide gel electrophoresis.

Assay for Free Sulfhydryl Groups. E5 was assayed for sulfhydryl content using Ellman's reagent (DTNB) (Ellman, 1959). The reaction mixture contained 0.1 M Tris-HCl, pH 8.0, 0.01 M EDTA, 0.5% SDS, 50–150 μ g of protein, and 24 μ g of DTNB in a volume of 0.3 mL. The absorbance at 412 nm was read against a blank containing the above mixture without protein.

Reduction and Carboxymethylation. E5 and E5_{cd} were reduced and carboxymethylated under nondenaturing and denaturing conditions as described in McGinnis and Wilson (1993).

Cyanogen Bromide Cleavage and Reverse-Phase HPLC. Purified E5_{cd} was cleaved with cyanogen bromide under nonreducing conditions, and the peptide fragments obtained were separated by RP-HPLC as in McGinnis and Wilson (1993).

Identification of Disulfide-Linked Peptides by Peptide Mapping. The peptide map of nonreduced E5_{cd} CNBr peptides containing disulfide-linked peptides was compared to the reduced CNBr peptide map. All peaks from both peptide maps were collected and analyzed for their amino acid content to identify disulfide- and cysteine-containing peptides. The reduced E5_{cd} CNBr peptides were obtained by adding DTT to 10 mM to the peptides and incubating them 37 °C for 2–4 h prior to filtration and RP-HPLC.

Amino Acid Analysis. Amino acid analysis was performed on iodoacetamide-treated E5 samples and E5_{cd} CNBr-generated peptide fractions obtained from RP-HPLC. The

[†] This work was supported by Grant FG02-84-ER13233 from the U.S. Department of Energy.

* Correspondence should be addressed to this author.

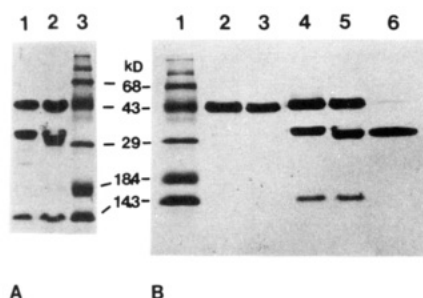


FIGURE 1: SDS-polyacrylamide gels of E5 and proteolytic fragments of E5. (Panel A) E5 proteolytic products with (lane 1) and without (lane 2) DTT. (Panel B) Purified E5 stored at -70°C (lanes 2 and 3); purified E5 stored at 4°C for 2 weeks (lanes 3 and 4); supernatant after E5 fragments are incubated with Avicel (lane 6).

samples were hydrolyzed in 6 N HCl for 95 min at 150°C , and the hydrolysates were fractionated on a Waters Pico Tag analyzer.

Other Procedures. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and protein determinations were carried out as described by Laemmli (1970) and Bradford (1976).

RESULTS

Identification of the Cellulose Binding Domain in E5. E5 purified from *S. lividans* culture supernatants contained trace amounts of one or more proteases. When purified E5 is stored above -70°C for more than 2 days, the protease begins to cleave the enzyme into two distinct fragments. Figure 1, panel B, shows a sample of purified E5 stored at -70°C (lanes 2 and 3) and the same preparation after 2 weeks at 4°C (lanes 4 and 5). Three bands were found: the intact E5 (near the 43-kDa marker), a 32-kDa fragment, and a 14-kDa fragment. It appeared that E5, MW 46 300, had been cleaved into fragments of 32 and 14 kDa. This same pattern was found in several other purified E5 preparations that had been stored at -20°C for more than 1 month. On Western blots, all three bands reacted with polyclonal antibodies raised against purified E5.

CMC overlays showed that the 32-kDa band (E5_{cd}) was catalytically active. Pure E5_{cd} has the same specific activity as E5 on CMC, but its filter paper activity is decreased. N-Terminal analysis of E5_{cd} gave a 10-residue sequence underlined in Figure 2 (residues 121–130) that was identical to that predicted from the DNA sequence. This result shows that E5_{cd} is missing the 120 most N-terminal residues of E5. E5 binds tightly to Avicel (crystalline cellulose), while E5_{cd} has lost this ability. Proteolytically cleaved E5 containing all three bands was incubated with Avicel as described under Materials and Methods. Both E5 and the 14-kDa fragment bind to Avicel, while E5_{cd} does not (Figure 1, panel B, lane 6). Thus, the 14-kDa fragment contains the cellulose binding domain of E5.

Determination of the Sulfhydryl and Disulfide Content of E5. Reaction of both native and denatured E5 with DTNB showed that no sulfhydryl groups were present, indicating that all six cysteine residues were involved in disulfide bonds. This was confirmed by the fact that amino acid analysis of denatured E5 reacted with iodoacetamide showed that no cysteine residues were modified. Cysteine residues were modified after the protein was reduced, or denatured and reduced. Amino acid analysis of these samples showed that less than half of the cysteine residues were modified in the reduced and iodoacetamide-treated protein while only (carboxymethyl)cysteine was found in denatured, reduced, and



FIGURE 2: Amino acid sequence of E5 deduced from the DNA sequence [reproduced from Lao *et al.* (1991)]. The N-terminal end of mature E5 (italicized and underlined) is preceded by a 36 amino acid signal sequence which is cleaved after the protein is excreted. The N-terminal of E5_{cd} (underlined) starts at Thr121 of E5. Cysteine residues (boldface) are numbered starting from the N-terminal of E5. The six methionine residues of E5_{cd} are underlined, and the seven E5_{cd} CNBr peptides are shown.

carboxymethylated E5. Therefore, the six cysteines in E5 are involved in three disulfide bonds.

Identification of the Disulfide Joining Cys93 and Cys100. Figure 1, panel A, shows the proteolytic fragments of E5 electrophoresed on an SDS-polyacrylamide gel with (lane 1) and without (lane 2) DTT in the solubilization buffer before electrophoresis. The relative mobilities of both E5_{cd} and the 14-kDa fragment shift upon reduction, indicating that both contain disulfide bonds. From the DNA sequence, E5_{cd} contains four cysteines, and the 14-kDa N-terminal fragment contains two cysteines (Figure 2). The two cysteines in the N-terminal fragment of E5, Cys93 and Cys100, are joined in a disulfide bond. If these two cysteine residues were disulfide-bonded to any of the other four cysteines in E5_{cd} , then without DTT, the two proteolytic fragments would be joined by disulfide bridges and would migrate at the same rate of E5 (46 kDa) on an SDS-polyacrylamide gel.

Cyanogen Bromide Cleavage of E5_{cd} . In order to determine the arrangement of the two remaining disulfides, E5_{cd} was

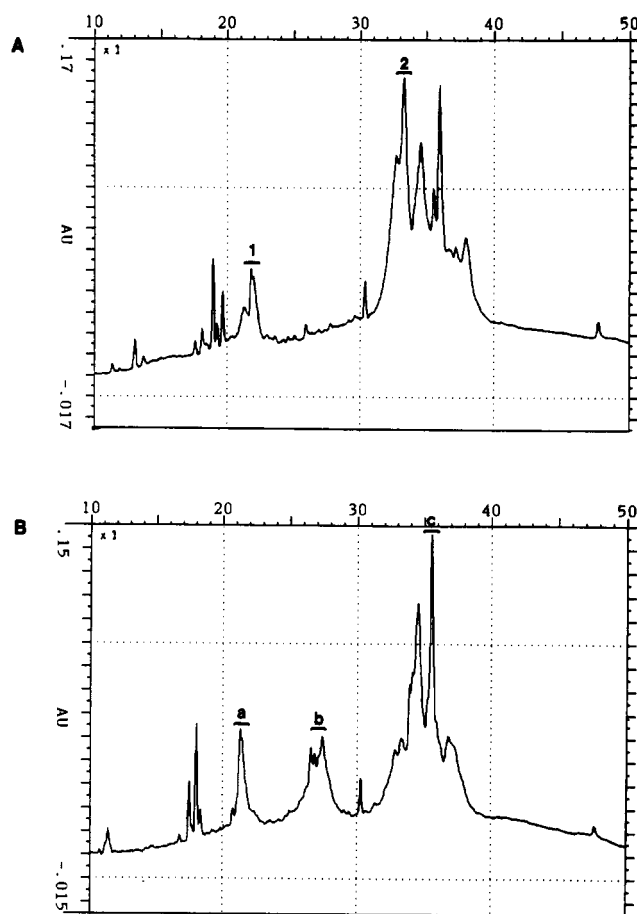


FIGURE 3: RP-HPLC CNBr peptide maps of $E5_{\alpha d}$. (Panel A) Nonreduced CNBr peptide map of disulfide-linked peptides. (Panel B) Peptide map obtained when the CNBr peptides were reduced with DTT before HPLC. The solvent system and chromatography conditions are described under Materials and Methods. Peptides were eluted with a linear gradient of 5–60% solvent B over 50 min at a flow rate of 1 mL/min. The column effluent was monitored at 210 nm.

cleaved under conditions which left them intact. $E5_{\alpha d}$ has six methionine residues; therefore, seven peptides were expected with molecular weights ($\times 10^{-3}$) of 3.5, 3.3, 2.4, 11.1, 4.7, 0.8, and 5.1, starting at the N-terminus and designated peptides 1–7, respectively (Figure 2). The MW 3500 peptide (peptide 1) contains two cysteine residues while the MW 3300 and MW 5100 peptides (peptides 2 and 7) each contain one cysteine. The locations of these peptides in the sequence are shown in Figure 2.

Disulfide Arrangement in $E5_{\alpha d}$. Peptide maps of disulfide-linked (nonreduced) (panel A) and reduced (panel B) CNBr peptides are shown in Figure 3. All peaks from both peptide maps were collected and analyzed for their amino acid composition. Those peaks that are labeled are the only ones that contained cysteine. Table I shows the expected amino acid composition of peptide 1 compared to the amino acid compositions of peak 1 from the –DTT CNBr peptide map (Figure 3, panel A) and peak a from the +DTT CNBr peptide map (Figure 3, panel B). These data indicate that peak 1 and peak a are identical and that both contain only peptide 1. No other peaks contained peptide 1. The retention time of this peptide is slightly shifted by addition of DTT. Thus, Cys138 and Cys143 in peptide 1 form one disulfide bond, and Cys166 and Cys406 form the other disulfide bond in $E5_{\alpha d}$. Peptides 2 and 7 should be linked by this disulfide. Peak 2 in the –DTT peptide map (panel A) contains these two peptides. Upon reduction, this peak disappears, and a new peak (peak

Table I: Predicted Amino Acid Composition of $E5_{\alpha d}$ CNBr Peptide 1 and the Amino Acid Composition of RP-HPLC Peak 1 from the –DTT Peptide Map (Figure 3A) and Peak a from the +DTT Peptide Map (Figure 3B)

	amino acid composition (mol %)		
	peptide 1	–DTT peak 1	+DTT peak a
Asx	5.7	5.8	3.6
Glx	17.1	13.5	11.5
Ser	0	1.2	1.3
Gly	17.1	20.4	18.8
His	2.9	2.4	2.7
Arg	5.7	5.7	5.6
Thr	11.4	9.4	10.2
Ala	0	0	0
Pro	8.6	9.9	10.6
Tyr	2.9	8.1	7.8
Val	11.4	10.0	10.8
Met ^a	2.9	0	0
Cys	5.7	1.5	2.3
Ile	0	1.0	2.9
Leu	5.7	5.7	6.0
Phe	0	0	0
Lys	2.9	2.4	2.9
Trp	0		
Hser ^a		1.8	1.5

^a Tryptophan is destroyed in the analysis, and methionine is converted to homoserine by CNBr cleavage. Tyrosine was high in all samples analyzed.

Table II: Predicted Amino Acid Composition for $E5_{\alpha d}$ CNBr Peptide 7 and the Amino Acid Composition of the RP-HPLC Fraction of Peak b from the +DTT Peptide Map (Figure 3B)

	amino acid composition (mol %)			amino acid composition (mol %)	
	peptide 7	+DTT peak b		peptide 7	+DTT peak b
Asx	5.9	4.6	Val	3.9	3.9
Glx	7.8	5.8	Met ^a	0	0
Ser	17.6	15.3	Cys	1.9	.5
Gly	13.7	15.1	Ile	1.9	4.8
His	0	0	Leu	3.9	4.9
Arg	5.9	5.8	Phe	3.9	4.0
Thr	3.9	3.3	Lys	7.8	6.0
Ala	7.8	8.4	Trp	7.8	
Pro	3.9	4.8	Hser ^a		0
Tyr	1.9	12.0			

^a Tryptophan is destroyed in the analysis. Tyrosine was high in all samples analyzed. The C-terminal CNBr peptide does not contain methionine; therefore, no homoserine is expected for this peptide. Peak b was the only peak in the +DTT map without homoserine.

b in the +DTT peptide map, panel B) appears. This was confirmed by the fact that when peak 2 was isolated, reduced, and chromatographed using the same elution conditions, two peaks which eluted in the positions of peak b and peak c in the +DTT peptide map were found. Amino acid analysis of peak b showed it contained peptide 7 (Table II), while peak c contained several peptides and thus its amino acid analysis composition did not provide definitive information.

Location of the Accessible and Inaccessible Disulfides in $E5$. $E5$ was reduced and carboxymethylated under non-denaturing and denaturing conditions. These samples were run on SDS–polyacrylamide gels with and without DTT in the solubilization buffer. With DTT, all three samples are expected to migrate to the same position, and this is what was found (Figure 4A). Without DTT, differences in the mobilities of these samples are due to the presence or absence of disulfide bonds. The differences in the relative mobilities of native $E5$, native $E5$ reduced with DTT and reacted with iodoacetamide ($E5$ D/I), and guanidine-denatured $E5$ reduced and reacted with iodoacetamide ($E5$ G/D/I) show the presence of both

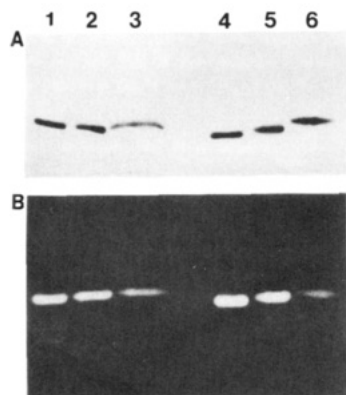


FIGURE 4: (Panel A) Lanes 1 and 4, native E5; lanes 2 and 5, partially reduced and carboxymethylated E5 (E5 D/I); lanes 3 and 6, fully reduced and carboxymethylated E5 (E5 G/D/I) run on SDS-polyacrylamide gels with (lanes 1–3) and without (lanes 4–6) DTT in the solubilization buffer. (Panel B) CMC overlay of these same samples. This overlay was incubated for 10 min at 50 °C before staining as described in McGinnis and Wilson (1993).

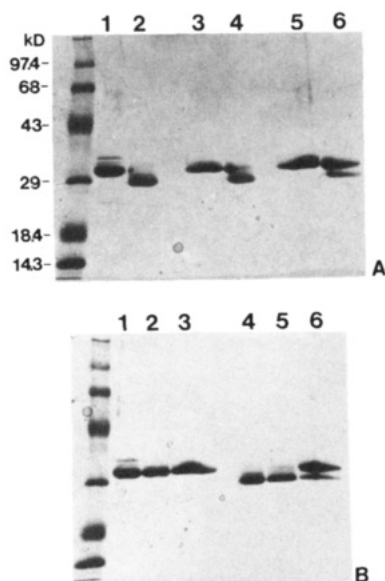


FIGURE 5: (Panel A) Coomassie-stained SDS-polyacrylamide gels of native E5_{cd}: lanes 1 and 2, native reduced and carboxymethylated E5_{cd} (E5_{cd} D/I); lanes 3 and 4, denatured, reduced, and carboxymethylated E5_{cd} (E5_{cd} G/D/I); lanes 5 and 6, samples in lanes 1, 3, and 5 contained DTT in the solubilization buffer. (Panel B) Same as panel A except lanes 1 and 4 are E5_{cd}, lanes 2 and 5 are E5_{cd} D/I, and lanes 3 and 6 are E5_{cd} G/D/I. Samples in lanes 1–3 contained DTT in the solubilization buffer.

accessible and inaccessible disulfides (Figure 4A). These three samples have increasingly lower relative mobilities. Since E5 has all three disulfide bonds intact, while E5 G/D/I has no intact disulfide bonds and E5 D/I has a mobility between these samples, one or two but not all of the E5 disulfides are reduced in the absence of denaturation.

Dithiothreitol and iodoacetamide treatment of native E5_{cd} shows that no disulfides are reduced under these conditions, since the relative mobility of E5_{cd} D/I is identical to that of E5_{cd} when both are run without DTT (Figure 5). The disulfide bonds of E5_{cd} can only be reduced when the protein is denatured so that both disulfide bonds in E5_{cd} are inaccessible. Therefore, the disulfide joining Cys93 to Cys100 is the accessible one in E5.

Effect of Reduction and Carboxymethylation on Enzymatic Activity and Binding. Reduction of all three disulfides in E5 resulted in the loss of more than 50% of the activity, while modification of the cysteines resulted in a loss of 75% of the

Table III: Effect of Reduction and Carboxymethylation on the Enzymatic Activity of Cellulase E5^a

	sp act. (units/mg)	% residual act.
E5	190	100
E5 D	175	92
E5 D/I	201	106
E5 G/D	90	47
E5 G/D/I	42	22

^a Specific activities were determined using the CMCase assay. The samples were incubated with CMC for 30 min at 50 °C. E5, native enzyme; E5 D, native enzyme assayed with 10 mM DTT present; E5 D/I, native enzyme was reduced with DTT and reacted with iodoacetamide; E5 G/D, the enzyme was denatured with guanidine hydrochloride and reduced with DTT, the guanidine was removed but not the DTT; E5 G/D/I, denatured, reduced, and fully carboxymethylated E5. Excess reagents were removed by ultrafiltration before the assays unless otherwise indicated.

activity (Table III, E5 G/D and E5 G/D/I). This loss in enzymatic activity is due to the reduction of the two disulfides located in the catalytic domain of E5, since reduction (Table III, E5 D and E5 D/I) of the accessible disulfide results in very little or no loss in activity. The activity of the modified samples can also be visualized on CMC overlays as shown in Figure 4B. The relative activities of the samples are similar to those found in the CMC assays. As the accessible disulfide was located in the 14-kDa fragment containing the binding domain of E5, the binding of E5 to Avicel in the presence of DTT was examined. Reduction of the accessible disulfide caused only a 10% decrease in the binding of E5 to crystalline cellulose. Similar results were obtained in binding assays using E5 D/I. Therefore, the disulfide linking Cys93 to Cys100 is not essential for maintaining the cellulose binding domain. The two disulfides located in the catalytic domain of E5 are important but not essential for maintaining the structure of the active site.

DISCUSSION

Evidence for separate functional domains in cellulases has been reported for several bacterial and fungal endoglucanases and exoglucanases (Van Tilbeurgh *et al.*, 1986; Gilkes *et al.*, 1988; Tomme *et al.*, 1988; Ghanges & Wilson, 1988; McGavin & Forsberg, 1989; Gilbert *et al.*, 1990; Jauris *et al.*, 1990; Durrant *et al.*, 1991). Protease treatment of purified enzymes has been used to elucidate the bifunctional organization of cellulases, and many cellulolytic bacteria excrete proteases which can cleave these enzymes *in vivo*. These cellulases appear to be composed of a noncatalytic cellulose binding domain present at either the N- or the C-terminal end of the protein linked by a hinge region rich in proline and hydroxyl amino acids to a catalytic domain. The hinge region is susceptible to proteolytic attack.

Cleavage of *T. fusca* endocellulase E5 by an *S. lividans* protease produces two degradation products: a catalytically active 32-kDa fragment (E5_{cd}), which can no longer bind tightly to crystalline cellulose, and a 14-kDa cellulose binding fragment derived from the amino terminus of E5. Amino acid sequence alignments between the binding domains of E2 (100 amino acids of the C-terminal) and E5 (120 amino acids of the N-terminal) revealed sequence similarity between the two, and in subsequent alignments, it became clear that these regions were also similar to the cellulose binding domains present in *Cellulomonas fimi* and *Pseudomonas fluorescens* (Lao *et al.*, 1991; Wilson, 1992). Thus, *T. fusca* E5 appears to exhibit the same structural organization as many other cellulases.

The three disulfide bonds of *T. fusca* E5 are apparently important in the structural stability of the enzyme as was found with E2 (McGinnis & Wilson, 1993). It appears that only the inaccessible disulfides play a role in maintaining a stable active site, since reduction of the accessible disulfide had little effect on activity. This is substantiated by the fact that the inaccessible disulfides are located in the catalytic domain of the enzyme while the accessible disulfide is found in the binding domain. The stability and active conformation of the catalytic core of CBH I are reported to be maintained by its six disulfide bridges (Bhikhabhai *et al.*, 1984). Disulfide bridges are known to be important in stabilizing proteins at high temperatures (Thornton, 1981). This would most likely be the role of disulfide bonds in the *T. fusca* enzymes since these cellulases are all thermostable.

ACKNOWLEDGMENT

We thank Diana Irwin for preparing the E5_{cd}.

REFERENCES

- Bellamy, W. D. (1977) *Dev. Ind. Microbiol.* 18, 249–254.
- Bhikhabhai, R., & Pettersson, G. (1984) *Biochem. J.* 222, 729–736.
- Bradford, M. (1976) *Anal. Biochem.* 92, 248–254.
- Calza, R. E., Irwin, D. C., & Wilson, D. B. (1985) *Biochemistry* 24, 7797–7804.
- Durrant, A. J., Hall, J., Hazelwood, G. P., & Gilbert, H. *Biochem. J.* 273, 289–293.
- Ellman, G. L. (1959) *Arch. Biochem. Biophys.* 82, 70–77.
- Ghangas, G. S., & Wilson, D. B. (1987) *Appl. Environ. Microbiol.* 53, 1470–1474.
- Ghangas, G. S., & Wilson, D. B. (1988) *Appl. Environ. Microbiol.* 54, 2521–2526.
- Gilbert, H. J., Hall, J., Hazelwood, G. P., & Ferreira, L. M. (1990) *Mol. Microbiol.* 4, 759–767.
- Gilkes, N. R., Warren, A. J., Miller, R. C., Jr., & Kilburn, D. G. (1988) *J. Biol. Chem.* 263, 10401–10407.
- Jauris, S., Rucknagel, K. P., Schwarz, W. H., Kratzsch, P., Bronnenmeier, K., & Staudenbauer, W. L. (1990) *Mol. Gen. Genet.* 223, 258–267.
- Laemmli, U. K. (1970) *Nature* 227, 680–685.
- Lao, G., Ghangas, G., Jung, E. D., & Wilson, D. B. (1991) *J. Bacteriol.* 173, 3397–3407.
- McCarty, A. J., Ball, A. S., & Bachmann, S. L. (1988) in *Biology of Actinomycetes*, pp 283–287, Japan Scientific Societies Press, Tokyo.
- McGavin, M., & Forsberg, C. W. (1989) *J. Bacteriol.* 171, 3310–3315.
- McGinnis, K., & Wilson, D. B. (1993) *Biochemistry* (second of three papers in this issue).
- Thornton, J. M. (1981) *J. Mol. Biol.* 151, 261–287.
- Tomme, P., Van Tilbeurgh, H., Pettersson, G., Van Damme, J., Vandekerckhove, J., Knowles, J., Teeri, T., & Claeyssens, M. (1988) *Eur. J. Biochem.* 170, 575–581.
- Van Tilbeurgh, H., Tomme, P., Claeyssens, M., Bhikhabhai, R., & Pettersson, G. (1986) *FEBS Lett.* 204, 223–227.
- Walker, L. P., Wilson, D. B., Irwin, D. C., McQuire, C., & Price, M. (1992) *Biotechnol. Bioeng.* 40, 1019–1026.
- Wilson, D. B. (1988) *Methods Enzymol.* 160, 314–323.
- Wilson, D. B. (1992) *Crit. Rev. Biotechnol.* 12, 45–63.