with modified structures would be a useful tool for clarifying protein-substrate interactions. Furthermore, scaling up the synthesis of small proteins by our procedure is feasible because the protected segments needed can be rapidly prepared on the oxime resin.

REFERENCES AND NOTES

1. W. J. Ghering, *Science* **236**, 1245 (1987); ——— and Y. Hiromi, *Annu. Rev. Genet.* **20**, 147 (1986).
2. W. McGinnis, M. S. Levine, E. Hafen, A. Kuroiwa, W. J. Gehring, *Nature* **308**, 428 (1984).
3. M. P. Scott and A. J. Weiner, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 4115 (1984).
4. G. Desplan, J. Theis, P. H. O'Farrell, *Nature* **318**, 630 (1985).
5. R. A. H. White and M. Wilcox, *Cell* **39**, 163 (1984).
6. E. Hafen, A. Kuroiwa, W. J. Gehring, *ibid.* **37**, 833 (1984).
7. K. Harding, C. Wedeen, W. McGinnis, M. Levine, *Science* **229**, 1236 (1985).
8. S. B. Carroll and M. P. Scott, *Cell* **43**, 47 (1985).
9. P. W. Ingham and A. Martinez-Arias, *Nature* **324**, 592 (1986).
10. S. B. Carroll *et al.*, *Cell* **47**, 113 (1986).
11. A. Laughon and M. P. Scott, *Nature* **310**, 25 (1984).
12. J. C. W. Shepherd, W. McGinnis, A. E. Carrasco, E. M. DeRobertis, W. J. Gehring, *ibid.*, p. 70.
13. C. O. Pabo and R. T. Sauer, *Annu. Rev. Biochem.* **53**, 293 (1984); M. Ptashne, in *A Genetic Switch: Gene Control and Phage Lambda* (Cell and Blackwell, Cambridge, 1987), chap. 2.
14. W. F. DeGrado and E. T. Kaiser, *J. Org. Chem.* **47**, 3258 (1982); S. H. Nakagawa and E. T. Kaiser, *ibid.* **48**, 678 (1984); S. H. Nakagawa, H. S. H. Lau, F. J. Kézdy, E. T. Kaiser, *J. Am. Chem. Soc.* **107**, 7087 (1985).
15. R. B. Merrifield, *J. Am. Chem. Soc.* **85**, 2149 (1963); G. Barany and R. B. Merrifield, in *The Peptides*, E. Gross and J. Meinhoper, Eds. (Academic Press, New York, 1980), vol. 2, pp. 1–284.
16. B. T. Chait, W. C. Agosta, F. H. Field, *Int. J. Mass Spectrom. Ion Phys.* **39**, 339 (1981).
17. P. A. Beachy, unpublished data.
18. D. Galas and A. Schmitz, *Nucleic Acids Res.* **5**, 3157 (1978).
19. C. Desplan, J. Theis, P. O'Farrell, *Cell*, in press.
20. T. Hoey and M. Levine, *Nature* **332**, 858 (1988).
21. S. W. Provencher and J. Glöckner, *Biochemistry* **20**, 33 (1981).
22. A. M. Miller, V. L. Mackay, K. A. Nasmyth, *Nature* **314**, 598 (1985).
23. J. P. Tam, W. F. Heath, R. B. Merrifield, *J. Am. Chem. Soc.* **105**, 6442 (1983).
24. U. Heberlein, B. England, R. Tjian, *Cell* **41**, 965 (1985).

25. We thank G. A. Laforet for her suggestions which aided us in the initiation of this work, P. Freimuth, D. A. Kendall, T. Sasaki, J. Taylor, and C. Desplan for helpful discussions, T. Hoey and M. S. Levine for teaching us the footprinting technique, and P. A. Beachy for providing us with the plasmid containing the 5′ upstream region of the *Antp* gene. All mass spectra were done by B. Chait at the Rockefeller University Biotechnology Mass Spectrometric Research Resource, which is supported by the Division of Research Resources, NIH grant RR 862. Peptide sequencing analysis was performed at the Yale University School of Medicine Sequencing Facility. Supported in part by a grant from Hoffmann–La Roche.

18 May 1988; accepted 7 September 1988

VirD2 Protein of *Agrobacterium tumefaciens* Very Tightly Linked to the 5′ End of T-Strand DNA

ERIC R. WARD* AND WAYNE M. BARNES†

The T-strand, a probable intermediate of *Agrobacterium* plant transformation, is bound by a nondenaturable linkage to a protein moiety at its 5′ end. The protein is shown to be the polypeptide VirD2, previously identified as a component of the T-DNA border endonuclease that initiates T-strand production. T-strands from an *Agrobacterium* strain expressing a *virD2-lacZ* fusion are bound to a protein of larger size than the wild-type protein and are immunoprecipitable by antibody to β-galactosidase.

Agrobacterium tumefaciens CAUSES THE neoplastic disease crown gall by transforming the nuclear genome of higher plants (*1*). The segment of DNA transferred (the T-DNA) resides on the large tumor-inducing (Ti) plasmid and is flanked by short direct repeats called border sequences (*2*). Six virulence (*vir*) loci, also located on the Ti plasmid, act on the T-DNA in trans to promote its transfer to plant cells (*3*). The *vir* genes are inducible by crude plant exudates and by specific phenolic compounds found therein (*4*). *Agrobacterium tumefaciens* cells grown under *vir*-inducing conditions produce single-stranded copies of the lower strand of the T-DNA called T-strands (*5–7*), which are thought to be intermediates of plant transformation. The first two open reading frames (ORFs) of *virD* (*virD1* and *virD2*) encode a site-specific endonuclease that nicks the lower strand of the border sequences (*8–10*), presumably initiating T-strand production (*5, 6, 10*).

Because mating bacterial cells also produce single strands of the DNA to be transferred (*11*), it has been suggested that *Agrobacterium*-mediated plant transformation may be similar to conjugation. Recent genetic evidence, showing that the mobilization (*mob*) genes and origin of transfer (*oriT*) of a mobilizable plasmid can substitute for T-DNA border sequences (*12*), has strengthened the case. In conjugation, one of the *mob*-encoded proteins involved in *oriT* nicking is thought to bind covalently to the 5′ end of the nick and remain associated with the conjugal plasmid strand (*11, 13*). We show that the *Agrobacterium* T-strand is bound tightly at or very near its 5′ end to a component of the border endonuclease.

We developed a method to construct Ti plasmids that bear tandem arrays of many

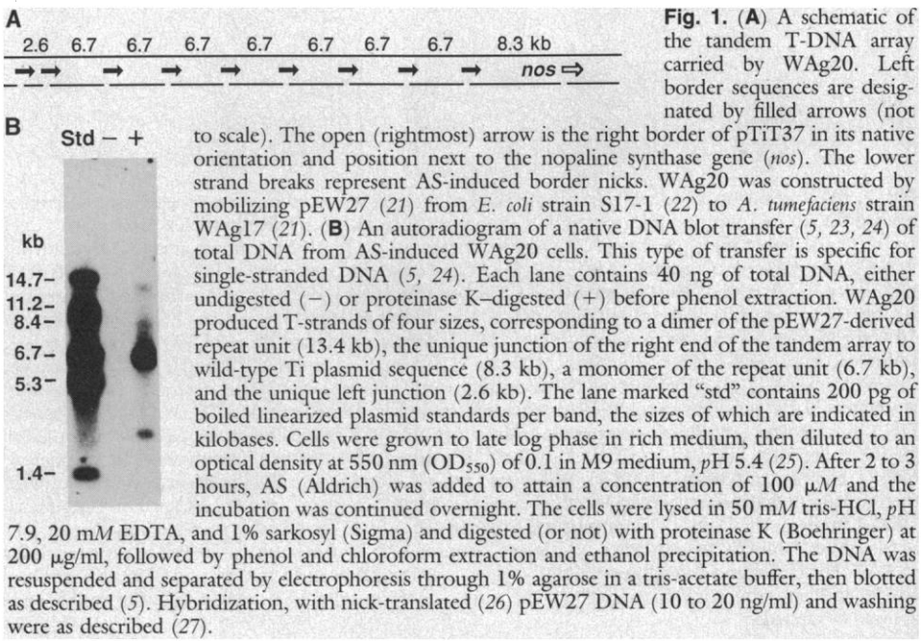

Fig. 1. (**A**) A schematic of the tandem T-DNA array carried by WAg20. Left border sequences are designated by filled arrows (not to scale). The open (rightmost) arrow is the right border of pTiT37 in its native orientation and position next to the nopaline synthase gene (*nos*). The lower strand breaks represent AS-induced border nicks. WAg20 was constructed by mobilizing pEW27 (*21*) from *E. coli* strain S17-1 (*22*) to *A. tumefaciens* strain WAg17 (*21*). (**B**) An autoradiogram of a native DNA blot transfer (*5, 23, 24*) of total DNA from AS-induced WAg20 cells. This type of transfer is specific for single-stranded DNA (*5, 24*). Each lane contains 40 ng of total DNA, either undigested (−) or proteinase K–digested (+) before phenol extraction. WAg20 produced T-strands of four sizes, corresponding to a dimer of the pEW27-derived repeat unit (13.4 kb), the unique junction of the right end of the tandem array to wild-type Ti plasmid sequence (8.3 kb), a monomer of the repeat unit (6.7 kb), and the unique left junction (2.6 kb). The lane marked "std" contains 200 pg of boiled linearized plasmid standards per band, the sizes of which are indicated in kilobases. Cells were grown to late log phase in rich medium, then diluted to an optical density at 550 nm (OD_{550}) of 0.1 in M9 medium, *p*H 5.4 (*25*). After 2 to 3 hours, AS (Aldrich) was added to attain a concentration of 100 μ*M* and the incubation was continued overnight. The cells were lysed in 50 m*M* tris-HCl, *p*H 7.9, 20 m*M* EDTA, and 1% sarkosyl (Sigma) and digested (or not) with proteinase K (Boehringer) at 200 μg/ml, followed by phenol and chloroform extraction and ethanol precipitation. The DNA was resuspended and separated by electrophoresis through 1% agarose in a tris-acetate buffer, then blotted as described (*5*). Hybridization, with nick-translated (*26*) pEW27 DNA (10 to 20 ng/ml) and washing were as described (*27*).

Department of Biological Chemistry, Washington University School of Medicine, St. Louis, MO 63110.

*Present address: CIBA-GEIGY Biotechnology Research, Box 12257, Research Triangle Park, NC, 27709.
†To whom correspondence should be addressed.

copies of a small integrative plasmid carrying a T-DNA border sequence (Fig. 1A). WAg20, an *Agrobacterium* host harboring such a Ti plasmid, produces a high level of T-strands (approximately 1% of the total cellular DNA) when induced with the plant phenolic acetosyringone (AS) (Fig. 1B). These overproduced T-strands are apparently identical in structure and biological activity to T-strands from wild-type strains (*14*). When DNA was prepared from AS-induced WAg20, failure to proteolyze the cell lysate before extraction with phenol resulted in loss of the T-strands from the aqueous phase (Fig. 1B). Extraction of DNA into the phenol phase, first demonstrated for the adenovirus genome (*15*), is highly suggestive of a covalent linkage between the DNA and protein.

We next determined where on single-stranded T-DNA the protein is attached. A crude T-strand–containing lysate was digested with nuclease S1. Portions of the reaction mixture were removed over time and heated to 70°C in the presence of SDS, EDTA, and β-mercaptoethanol at neutral *p*H. This step stops the S1 reaction and denatures most (if not all) proteins. The digests were then annealed to radiolabeled, end-specific oligonucleotide probes (*16*), separated by electrophoresis through an SDS-polyacrylamide gel, and transferred to nitrocellulose. The blotting conditions allowed binding of only protein (*17*), so that any DNA signal must have resulted from a protein-DNA association that resisted the denaturing treatment.

Because the conditions used in this case do not melt double-stranded DNA (*18*), only native single strands (that is, T-strands) were accessible to the probes. If the protein is bound to the T-strand at the end recognized by a given probe, the assay should detect products of S1 digestion. If, however, the protein is bound away from the ends or to the end opposite a given probe, S1 cleavage separates the probed end from the protein-bound portion and no signal should be detectable in the assay.

The result of such an experiment is shown in Fig. 2. The 5′ probe labeled a faint smear of many bands within 1 min after the start of the reaction. The fastest migrating of these, at approximately 75 to 80 kD, increased greatly in intensity as the reaction progressed and decreased only if tenfold more S1 enzyme was added to the reaction. Appearance of this band was AS-dependent (see Fig. 3A). No signal was detectable with the 3′ probe. This result indicates nondenaturable binding of a protein to the 5′ end of single-stranded T-DNA. The 75- to 80-kD band represents a complex of the protein moiety and an S1-resistant DNA fragment of unknown length bound to it. The end-bound protein apparently protected a short span of adjacent DNA, accounting for both the sharpness of the S1-generated signal and its relative resistance to further nucleolytic degradation.

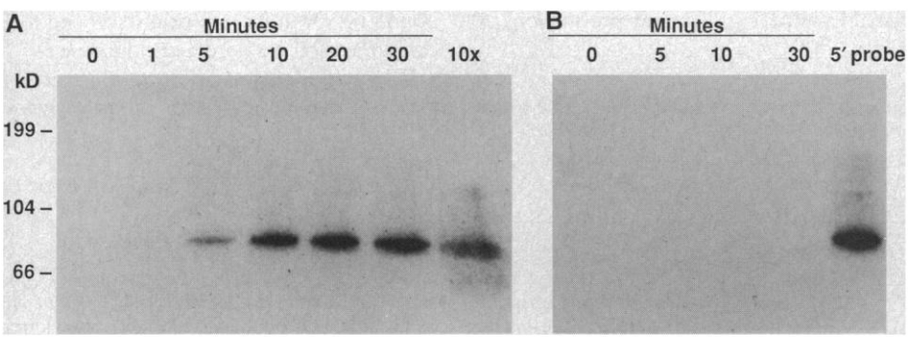

Fig. 2. (**A**) A covalent protein-DNA complex blot of samples annealed to the 5′ probe (*16*). A sample reacted for 30 min with tenfold more S1 is shown (10×). (**B**) A covalent complex blot of the same samples annealed to the 3′ probe (*16*). A sample annealed to the 5′ probe is shown as a positive control. WAg20 was induced as described for Fig. 1. Harvested cells were resuspended in 20% sucrose, 50 m*M* tris-HCl, *p*H 7.9, 50 m*M* EDTA, and 50 m*M* NaCl to a final OD_{550} of 9. After digestion with lysozyme (Cooper) at 1 mg per milliliter for 20 min on ice, sarkosyl was added to 1%, and the lysate was heated to 70°C for 15 min. The S1 reaction consisted of 180 μl of the above lysate, 45 μl of 10× S1 buffer (0.5*M* sodium acetate, *p*H 5.6, 45 m*M* $ZnSO_4$), 9 μl of 1*M* $ZnSO_4$, 9 μl of 5*M* NaCl, and 9 μl of S1 nuclease (concentration 4 U/μl) (Boehringer). S1 nuclease is relatively resistant to ionic detergent (*28*), and the amount used was sufficient to completely degrade T-strands within 30 min at 37°C, as assayed by native DNA blot transfer (*18*). Portions removed at the times indicated were added to an equal volume of 2× loading buffer [4% SDS, 0.1*M* tris-HCl, *p*H 8.8, 10% β-mercaptoethanol, 30% glycerol, bromophenol blue (1 mg/ml), and 50 m*M* EDTA], and heated to 70°C for 15 min. To the denatured samples was then added one-half volume of 5 n*M* ^{32}P-labeled oligonucleotide probe (in 1*M* NaCl, 10 m*M* tris-HCl, *p*H 7.9, and 1 m*M* EDTA). The samples were hybridized at 50°C for 2 hours, then separated on an 8% SDS-polyacrylamide gel essentially according to Laemmli (*29*). The gel was transferred to nitrocellulose (BA-85; Schleicher and Schuell) overnight between graphite plates in 25 m*M* tris, and 192 m*M* glycine at 150 mA constant (6 V/cm initially). The filter was then agitated briefly in blotting buffer and autoradiographed.

We speculated, by analogy to bacterial conjugation, that the bound protein might be identical to a component of the *virD*-encoded border endonuclease. The apparent size of the S1-generated complex suggested that it included at least the VirD2 protein and perhaps also VirD1 (47 and 16 kD, respectively) (*8, 9*). Rather than purify the complex and directly characterize its protein components, we constructed a fusion of *virD2* to the *lacZ* gene of *Escherichia coli* (*19*). As a control, we also made an unfused version, not in reading frame at the *virD2-lac* junction (*19*).

Tandem arrays of many copies of the fusion gene and the out-of-frame control, interspersed with T-DNA border sequences, were constructed in *A. tumefaciens* (*19*). Both the fusion-containing strain (WAg24) and the unfused control-containing strain (WAg25) produced low, constitutive levels of T-strands, which increased approximately 50-fold to the wild-type level on induction with AS (*18*). Sizes of the S1-digested, protein-DNA complexes present in uninduced and AS-induced WAg20, WAg24, and WAg25 were assayed by covalent complex–specific blotting (Fig. 3A). Uninduced WAg24 generated faint signals at 135 to 165 kD, the approximate mobility expected for the S1-digested complex containing a protein 60 to 90 kD larger than wild type. This result is consistent with binding of large portions of the VirD2–β-galactosidase fusion protein to the 5′ end of T-strands produced in the absence of AS induction. On induction with AS, the wild-type signal appeared at approximately 75 to 80 kD, and the putative fusion-produced signal increased severalfold in intensity (Fig. 3A). The low level of T-strands detected by agarose gel and native DNA blot analysis in uninduced WAg24 (*18*) may therefore be a consequence of a limiting AS-induced factor, rather than intrinsically lower activity or expression of the VirD2–β-galactosidase protein. No signal was detectable in uninduced WAg25, which was predicted to express a partial VirD2 protein 12 kD smaller than wild type. Uninduced WAg25 produced a level of T-strands identical to WAg24 as assayed by native DNA blot transfer (*18*), so its hybrid VirD2 protein fragment is active but may be altered in some way that changes its mobility on an SDS-polyacrylamide gel but does not affect its endonuclease activity. For instance, the protein may be cleaved by cellular proteases after nicking and covalently binding to the T-strand.

To confirm that T-strands produced by WAg24 are bound to the fusion protein, we immunoprecipitated lysates from AS-induced WAg20, WAg24, and WAg25 with

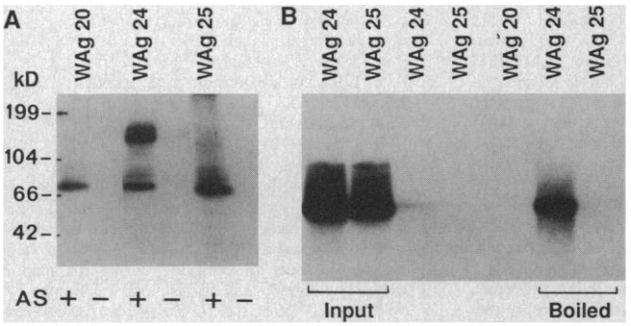

Fig. 3. (**A**) A covalent protein-DNA complex blot of lysates of uninduced and AS-induced WAg20, WAg24, and WAg25, digested with S1 for 30 min, annealed to the 5′ probe (*16*), and analyzed as described in the legend to Fig. 2. (**B**) An autoradiogram of a native DNA blot of immunoprecipitates of AS-induced WAg20, WAg24, and WAg25 lysates. WAg24 and WAg25 produce a repeated T-strand 10.6 kb in length, and a single copy 4.6-kb molecule (*18*). Only the major species is visible in this experiment. The two lanes marked "boiled" contain immunoprecipitates of lysates that were heated to 100°C for 5 min before incubation with the antibody. The lanes labeled "input" contain 10% of the amount of lysate used in the immunoprecipitations. (An equal amount of WAg20 lysate was used, but is not depicted here.) Crude AS-induced cell lysate (50 μl), prepared as described in the legend to Fig. 2, was incubated in 1 ml of immunoprecipitation buffer [50 m*M* tris-HCl, *p*H 7.9, 50 m*M* EDTA, 150 m*M* NaCl, 0.1% bovine serum albumin (BSA), 0.1% Tween 20, and 100 μg of sheared salmon testes DNA per milliliter] with 0.25 μl of mouse monoclonal antibody to β-galactosidase (Promega Biotech) at 4°C overnight. Protein A (200 μl) (*Staphylococcus aureus*, 10% suspension; Sigma) was added and the incubation continued for 2 hours. The protein A cells were pelleted and washed five times in 1.5 ml of buffer (10 m*M* tris-HCl, *p*H 7.9, 50 m*M* EDTA, 0.5*M* NaCl, and 0.5% Tween 20) before a final resuspension in 100 μl of 50 m*M* tris-HCl, *p*H 7.9, and 20 m*M* EDTA. Proteinase K was added to 400 μg/ml, and the resuspended pellets digested at 50°C for 1 hour. The deproteinized cells were pelleted and the supernatant was extracted with phenol and chloroform, and precipitated with ethanol. The resuspended DNAs were analyzed as described for Fig. 1.

antibody to β-galactosidase and analyzed precipitated DNA by native DNA blot transfer (Fig. 3B). A small fraction of T-strands from only WAg24 were immunoprecipitable, whereas T-strands from WAg20 and WAg25 did not react. Boiling the lysates in ionic detergent before incubation with the antibody increased the signal from WAg24 approximately tenfold, indicating that the bond between fusion protein and T-strand resisted this treatment. The increased intensity also suggests that double-stranded molecules, likely biogenic precursors of the T-strand (*5, 6*) that become single-stranded only on boiling, may also be bound to the fusion protein. Alternatively, boiling the protein may increase its precipitability by the antibody.

Our results show that *virD2* encodes a protein moiety that binds T-strands through a linkage that resists denaturation by phenol and by the combination of heat, detergent, and reducing agent. Unless the *virD2*-encoded protein is exceptionally resistant to standard denaturing treatments, this protein-DNA bond is almost certainly covalent. Our data do not rule out the possibility that VirD2 binds very close to and not exactly at the 5′ end of the T-strand. We have not investigated whether other proteins, especially the nicking-associated *virD1* product (*9*), may also tightly bind to the T-strand.

We propose that covalent linkage of VirD2 to the T-strand is analogous to covalent protein-DNA bonds at the *oriT* nick sites of conjugating plasmids. The biochemically best characterized example is the relaxed complex of ColEI, in which a 60-kD endonuclease-associated protein becomes tightly bound to the 5′ end of a strand-specific nick (*13*). Similar interactions are widely assumed to occur in other plasmid systems (*11*).

We envision several possible roles for the tightly attached VirD2 protein. (i) Immediately after border nicking, VirD2 may prevent ligation of the nick by cellular DNA ligase. Nonligatable *oriT* nicks have been reported for F (*20*). (ii) VirD2 may protect the T-strand from exonucleolytic degradation in either the bacterial or plant cell. (iii) The VirD2–T-strand complex may function in vectorial transport to the plant cell, by attaching to the *Agrobacterium* cell envelope (in association with membrane-bound products of other *vir* ORFs) at the point of plant cell contact. This possible role is analogous to functions postulated for the cytoplasmic TraZ and membrane-bound TraY proteins of F (*11*). (iv) Bound VirD2 may also function later in transformation, possibly even during integration of T-DNA into the plant nuclear genome.

REFERENCES AND NOTES

1. M.-D. Chilton *et al.*, *Cell* **11**, 263 (1977).
2. N. S. Yadav, J. Vanderleyden, D. R. Bennett, W. M. Barnes, M.-D. Chilton, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6322 (1982); P. Zambryski, A. Depicker, K. Kruger, H. Goodman, *J. Mol. Appl. Genet.* **1**, 361 (1982).
3. S. E. Stachel and E. W. Nester, *EMBO J.* **5**, 1445 (1986).
4. S. E. Stachel *et al.*, *Nature* **318**, 624 (1985); G. W. Bolton, E. W. Nester, M. P. Gordon, *Science* **232**, 983 (1986); S. E. Stachel, E. W. Nester, P. C. Zambryski, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 379 (1986).
5. S. E. Stachel, B. Timmerman, P. Zambryski, *Nature* **322**, 706 (1986).
6. L. M. Albright, M. F. Yanofsky, B. Leroux, D. Ma, E. W. Nester, *J. Bacteriol.* **169**, 1046 (1987).
7. K. Veluthambi *et al.*, *ibid.* **170**, 1523 (1988).
8. M. F. Yanofsky *et al.*, *Cell* **47**, 471 (1986).
9. R. K. Jayaswal, K. Veluthambi, S. B. Gelvin, J. L. Slightom, *J. Bacteriol.* **169**, 5035 (1987).
10. S. E. Stachel, B. Timmerman, P. Zambryski, *EMBO J.* **6**, 857 (1987).
11. N. Willetts and B. Wilkins, *Microbiol. Rev.* **48**, 24 (1984).
12. V. Buchanan-Wollaston, J. E. Passiatore, F. Cannon, *Nature* **328**, 172 (1987).
13. D. G. Blair and D. R. Helinski, *J. Biol. Chem.* **250**, 8785 (1975); D. G. Guiney and D. R. Helinski, *ibid.*, p. 8796; M. A. Lovett and D. R. Helinski, *ibid.*, p. 8790; G. J. Warren, A. J. Twigg, D. J. Sherratt, *Nature* **274**, 259 (1978).
14. Two facts support this assertion. First, T-strands from wild-type strains, as well as all of the T-strands produced by multiple border strains, are phenol-soluble. Second, a selectable plant marker (kanamycin resistance) placed either in the unique junction between the right T-DNA border and first repeated border, or between the tandemly repeated borders, is efficiently transferred to plant cells (*18*).
15. M. Green and M. Piña, *Proc. Natl. Acad. Sci. U.S.A.* **51**, 1251 (1964); S. V. Desiderio and T. J. Kelly, *J. Mol. Biol.* **145**, 319 (1981).
16. The probe specific for the 5′ end of the T-strand was a synthetic 19-base oligonucleotide with the sequence 5′-GGTCGGGGAGCTGTTGGCT-3′, ending 11 bases from the nick site; the 3′ probe was a 22-base oligonucleotide with the sequence 5′-GGTGTAAACAAATTGACGCTTA-3′, beginning 13 bases from the nick site. Both were derived from published pTiT37 sequence (*2*) and the previously determined position of the border nick [(*6*) and K. Wang, S. E. Stachel, B. Timmerman, M. Van Montagu, P. C. Zambryski, *Science* **235**, 587 (1987)]. Filter hybridizations against DNAs of deletion derivatives of pEW27 and M13-generated single strands showed that the oligonucleotides specifically hybridized only to the strand and sequence for which they were intended. In addition, solution hybridizations against boiled restriction fragments were performed under identical salt concentration and temperature conditions as the experiments shown in Figs. 2 and 3A. SDS–polyacrylamide gel electrophoresis showed that both the 5′ and 3′ probes detected the fragments expected (*18*). Thus, the lack of signal with the 3′ probe in Fig. 2 is not a result of simple inability of the probe to hybridize under the conditions used.
17. As a control for this type of blot, we ran [^{35}S]methionine-labeled *E. coli* proteins and ^{32}P-labeled single- and double-stranded DNA on an SDS-polyacrylamide gel. The DNA-containing lanes were heavily loaded so that the ^{32}P signal completely saturated an autoradiogram after overnight exposure. The gel was then blotted to nitrocellulose as described for Fig. 2, and the filter was autoradiographed for several days. While the ^{35}S-labeled proteins were quantitatively transferred and bound, no signal was detectable in the DNA lanes.
18. E. R. Ward and W. M. Barnes, unpublished data.
19. The fusion gene was constructed by ligating a 1.3-kb Sma I–Bam HI fragment of cosmid pSM358 [S. E. Stachel, G. An, C. Flores, E. W. Nester, *EMBO J.* **4**, 891 (1985)] between the Sma I and Bam HI sites of the *lacZ* fusion vector pJS413 [R. J. Watson, J. H. Weis, J. S. Salstrom, L. W. Enquist, *Science* **218**, 381 (1982)]. The resulting construct contains the *lac* promoter-operator, all of *virD1*, and the first 262 codons of *virD2* (31.5 kD). The *virD2* sequence is joined out of reading frame site to *lac* structural sequences. This out-of-frame construct encoded a hybrid protein with predicted size of 35 kD, comprising the *virD2* segment plus 42 additional amino acids encoded by the wrong *lac* reading frame. The *virD2* sequence was translationally fused to the *lac* codons by addition of a 10-bp synthetic linker. Expression of the gene was tested in *Escherichia coli*, where it encoded a fusion protein of the predicted size (148 kD), and specifically nicked T-DNA border-containing plasmids (*18*). A similar fusion, constructed by transposon insertion, had previously been shown to have nicking activity [see S. E. Stachel, G. An, C. Flores, E. W. Nester, above and

(*10*)]. The 4.6-kb Eco RI fragments carrying the *lac*-promoted out-of-frame construct and in-frame fusion were then inserted into pEW49 [a derivative of pEW27 (*21*) with a single Eco RI site], and the resulting plasmids (pEW80 and pEW76, respectively) mobilized to WAg17 (*21*) as described in the legend to Fig. 1. Exconjugants containing Ti plasmid–borne tandem arrays of the integrative plasmids were selected, and their structure was confirmed by restriction digestion and DNA blot analysis (*18*). *Vir D* sequences were bases 100 to 1420 of GenBank access code M14762.

20. R. Everett and N. Willetts, *J. Mol. Biol.* **136**, 129 (1980).
21. pEW27 is a 6.7-kb derivative of pBR322 [F. Bolivar *et al.*, *Gene* **2**, 95 (1977)] that confers tetracycline resistance and carries the left border of Ti plasmid pTiT37 as a 1.4-kb Eco RI fragment [N. S. Yadav, K. Postle, R. K. Saiki, M. F. Thomashow, M.-D. Chilton, *Nature* **287**, 458 (1980)]. WAg17 was constructed from *A. tumefaciens* strain A208 [D. Sciaky, A. L. Montoya, M.-D. Chilton, *Plasmid* **1**, 238 (1978)] by the method of A. J. Matzke and M.-D. Chilton [*J. Mol. Appl. Genet.* **1**, 39 (1981)]. WAg17 harbors a deleted derivative of pTiT37, having a 4.6-kb T region and sharing homology with pBR322 and pEW27. Details of the constructions are available upon request. *Agrobacterium tumefaciens* exconjugants resistant to ≥12 μg of tetracycline per milliliter always harbor Ti plasmids bearing multiple tandem insertions of the small donor plasmid, since one copy of the tetracycline resistance gene is insufficient for resistance to this level of the drug (*18*).
22. R. Simon *et al.*, *Bio/technology* **1**, 784 (1983).
23. E. M. Southern, *J. Mol. Biol.* **98**, 503 (1975).
24. L. Marsh *et al.*, *Virology* **143**, 212 (1985).
25. E. Kellenberger, K. G. Lark, A. Bolle, *Proc. Natl. Acad. Sci. U.S.A.* **48**, 1860 (1962).
26. P. W. J. Rigby, M. Dieckmann, C. Rhodes, P. Berg, *J. Mol. Biol.* **113**, 237 (1977).
27. G. M. Church and W. Gilbert, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1991 (1984).
28. V. M. Vogt, *Eur. J. Biochem.* **33**, 192 (1973).
29. U. K. Laemmli, *Nature* **277**, 680 (1970).
30. We thank M.-D. Chilton for providing pTiT37 subclones, E. Sadler (Howard Hughes Medical Institute) for providing the oligonucleotide probes, and B. Cookson for critically reading the manuscript. The initial phase of this work was funded by NSF grant PCM-8314661. E.R.W. was supported in part by an NSF graduate fellowship.

19 May 1988; accepted 30 August 1988

A Pituitary *N*-Acetylgalactosamine Transferase That Specifically Recognizes Glycoprotein Hormones

PETER L. SMITH AND JACQUES U. BAENZIGER*

The glycoprotein hormones lutropin (LH) and follitropin (FSH), which have common α-subunits but hormone-specific β-subunits, are both synthesized in the gonadotroph. However, they bear Asn-linked oligosaccharides that differ in structure. Those on LH terminate with the sequence SO_4-4GalNAcβ1→4GlcNAcβ1→2Manα, whereas those on FSH terminate with the sequence sialic acidα–Galβ1→4GlcNAcβ1→2Manα. A GalNAc-transferase was identified in bovine pituitary membranes that recognizes features of the α-subunit peptide and adds GalNAc to its oligosaccharides with an apparent Michaelis constant of 25 micromolar. The different patterns of glycosylation for LH and FSH indicate that access to the protein recognition marker on the α-subunit is modulated by the associated β-subunit. The tightly regulated synthesis of sulfated and sialylated oligosaccharides on the pituitary glycoprotein hormones suggests these oligosaccharides have an important biological role.

THE GLYCOPROTEIN HORMONES ARE a family of four closely related dimeric proteins with common α-subunits and hormone-specific β-subunits. Even though the amino acid sequences of the β-subunits of LH (lutropin), FSH (follitropin), TSH (thyrotropin), and hCG (human chorionic gonadotropin) are closely related, the Asn-linked oligosaccharides present on their α- and β-subunits differ in structure (*1–3*). Two of the hormones, LH and FSH, are synthesized within the same cell in the pituitary, the gonadotroph (*1, 4, 5*). LH bears Asn-linked oligosaccharides terminating with the sequence SO_4-4GalNAcβ1→4GlcNAcβ1→2Manα (Fig. 1), whereas FSH and other pituitary glycoproteins contain Asn-linked oligosaccharides terminating with the sequence sialic acidα–Galβ1→4GlcNAcβ1→2Manα (Fig. 1). Sulfated oligosaccharides such as S-2 have to date only been found on certain of the glycoprotein hormones (*1, 6, 7*). The presence of different oligosaccharide structures on LH and FSH indicates that glycosylation of these hormones is tightly regulated.

The steps leading to the synthesis of sulfated and sialylated oligosaccharides are illustrated in Fig. 1. The synthetic intermediates enclosed by the broken lines are precursors for both sialylated and sulfated oligosaccharides. Thus, either Gal or GalNAc can be added to the product of step 2. For FSH and the vast majority of other pituitary glycoproteins, Gal is added to the product of step 2, followed by sialic acid (steps 5 and 6). In contrast, GalNAc and sulfate are sequentially added to the identical synthetic intermediates on LH (steps 3 and 4). Since the oligosaccharide intermediates present on the glycoprotein hormones and other pituitary glycoproteins are identical, addition of GalNAc rather than Gal must involve a GalNAc-transferase that recognizes features of the α- and β-subunits of LH. Uncombined (free) α-subunits synthesized by bovine pituitaries, like LH, bear sulfated rather than sialylated oligosaccharides (*8*). This suggested that a recognition marker present on the α-subunit promotes the synthesis of sulfated oligosaccharides and that access to this recognition marker is modulated by the associated β-subunit, resulting in differential glycosylation of LH and FSH (*9*). We have identified a GalNAc-transferase in bovine pituitary membranes that displays specificity for the α-subunit of glycoprotein hormones and propose that this activity accounts for the synthesis of sulfated oligosaccharides on glycoprotein hormones and uncombined α-subunits.

The hCG α-subunit (hCGα) (*10, 11*), human transferrin, human α_1-antitrypsin, human J-chain (*12*), and bovine immunoglobulin G (IgG) (*13*) bear sialylated dibranched complex type oligosaccharides. Glycoproteins and glycopeptides bearing $GlcNAc_2Man_3GlcNAc_2$ (product of step 2 in Fig. 1) were prepared by enzymatic removal of sialic acid and Gal, and compared as substrates for GalNAc- and Gal-transferase activities present in crude bovine pituitary membrane preparations (Fig. 2 and Table 1). [^{3}H]Gal was transferred from uridine diphosphate (UDP)–Gal to agalacto-oligosaccharides on each of these substrates with apparent Michaelis constant (K_m) values between 160 and 900 μ*M*. The catalytic efficiency (V_{max}/K_m) was similar for transfer of Gal to each of these substrates (Table 1). The differences in the apparent K_m values obtained for addition of Gal to these substrates may reflect the presence of two oligosaccharide moieties on transferrin, α_1-antitrypsin, and hCGα as compared to a single oligosaccharide on J-chain and the glycopeptide.

Pituitary membranes also contained an enzyme activity that transferred [^{3}H]GalNAc to agalacto-oligosaccharides. When transfer of GalNAc and Gal were compared at similar substrate concentrations (Fig. 2), the GalNAc-transferase displayed a marked preference for hCGα. Although it was possible to detect transfer of GalNAc to each of the substrates examined, the apparent K_m values obtained were 16- to 60-fold as great as the apparent K_m for transfer of GalNAc to hCGα. The catalytic efficiency for transfer

Department of Pathology, Washington University School of Medicine, 660 South Euclid Avenue, St. Louis, MO 63110.

*To whom correspondence should be addressed.