VirB/D4-dependent protein translocation from Agrobacterium into plant cells

Vergunst, A.C.; Schrammeijer, B.; Dulk-Ras, A. den; Vlaam, C.M.T. de; Regensburg-Tuink, T.J.G.; Hooykaas, P.J.J.

Table 2. Parameter values used in the baseline matrix developed for Poverty Flat index stock of SRSS chinook salmon. The corresponding population growth rate λ is 0.760 (*9*).

Parameter	Value	Reference no.
s_1	0.022	(*13*)
s_2		
z	0.729	(*14*)
s_z	0.98	(*14*)
s_d	0.202	(*14*)
s_e	0.017	(*15*)
s_3, s_4, s_5	0.8, 0.8, 0.8	(*16*)
b_3, b_4, b_5	0.013, 0.159, 1.0	(*17*)
μ		
h_{ms}	0.020	(*8*)
s_{ms}	0.794	(*14*)
h_{sb}	0	(*8*)
s_{sb}	0.9	(*8*)
m_3, m_4, m_5	3257, 4095, 5149	(*18*)

this indirect mortality were 9% or higher, then dam breaching could reverse the declining trend of SRSS chinook salmon (Fig. 5). Unfortunately, estimating the magnitude of any indirect mortality from passage through the Snake River dams is difficult; identifying fish appropriate as a "control" for the potential effects of these dams is problematic. Also, even if the Snake River dams were removed, the fish would still have to negotiate four Columbia River dams, and baseline mortality would still include any indirect mortality attributable to passage through those dams.

For the Snake River, deliberation regarding dam removal will require us to examine the effects of dams that may be manifested outside the migration corridor. Given the current uncertainty, policy-makers may have to view the decisions they make as large experiments, the outcomes of which cannot be predicted but from which we can learn a great deal pertaining to endangered salmonids worldwide.

References and Notes

1. National Research Council, *Upstream: Salmon and Society in the Pacific Northwest* (National Academy Press, Washington, DC, 1996).
2. *Compilation of Information on Salmon and Steelhead Losses in the Columbia River Basin. Appendix D* (Northwest Power Planning Council, Portland, OR, 1986).
3. R. Waples, *Am. Fish. Soc. Symp.* **17**, 8 (1995).
4. *The Lower Snake River Juvenile Salmon Migration Feasibility Report/Environmental Impact Statement* (draft) (U.S. Army Corps of Engineers, 1999) (available at http://www.nww.usace.army.mil).
5. D. Marmorek, C. Peters, I. Parnell, Eds., *Plan for Analyzing and Testing Hypotheses (PATH): Final Report for Fiscal Year 1998* (ESSA Technologies, Vancouver, Canada, 1998).
6. H. Caswell, *Matrix Population Models* (Sinauer, Sunderland, MA, 1989).
7. S. Ratner, R. Lande, B. B. Roper, *Conserv. Biol.* **11**, 879 (1997).
8. R. C. P. Beamesderfer *et al.*, in *Plan for Analyzing and Testing Hypotheses (PATH): Retrospective and Prospective Analyses of Spring/Summer Chinook Reviewed in FY 1997,* D. R. Marmorek, C. N. Peters, Eds. (ESSA Technologies, Vancouver, Canada 1998), section 2, chap. 1.
9. Supplemental Web material is available on *Science* Online at www.sciencemag.org/feature/data/1053311.shl.
10. J. Williams, S. Smith, W. Muir, *N. Am. J. Fish. Manage.*, in press.
11. J. Hall-Griswold, T. Cochnauer, *Salmon Spawning Ground Survey* (Idaho Fish and Game Department, Boise, IA, 1988); M. White, T. Cochnauet, *Salmon Spawning Ground Surveys* (Idaho Fish and Game Department, Boise, IA, 1989); T. J. Elms-Cockrom, *Salmon Spawning Ground Survey, 1997* (Idaho Fish and Game Department, Boise, IA, 1998).
12. National Marine Fisheries Service, *Summary of Research Related to Transportation of Juvenile Anadromous Salmonids Around Snake and Columbia River Dams* (2000) (available at http://www.nwfsc.noaa.gov/pubs/white/trans4-25-00.pdf).
13. Productivity of each stock, *P*, was estimated as

$$\sum_{t=1}^{n} \ln(R_t/N_t)/n, \qquad \text{where } R_t = \sum_{x=4}^{5} N_{x,t+x}$$

is the number of recruits for a particular brood year; *t*, $N_{x,t+x}$ is the number of adults of age *x* that spawn *x* years after the brood year; and *n* is the number of data years used. s_1 was found by simultaneously solving the Euler equation

$$\mu \sum_{x=1}^{5} l_x(m_x/2)b_x\lambda^{-x} = 1$$

(*17*) and $\lambda^T = e^P$, where the generation time

$$T = \mu \sum_{x=1}^{5} l_x(m_x/2)b_x\lambda^{-x} \qquad (7)$$

.
14. National Marine Fisheries Service, *DRAFT Biological Opinion: Operation of the Federal Columbia River Power System* (2000) (available at http://www.nwr.noaa.gov/1hydrop/hydroweb/docs/2000/2000Biop.htm).
15. To calculate s_e, we used annual counts of smolts for the aggregate run of SRSS chinook made at Lower Granite Dam [C. E. Petrosky, H. Schaller, in *PATH Weight of Evidence Report*, D. Marmorek, C. Peters, Eds. (ESSA Technologies, Vancouver, Canada, 1998), submission 10]. We accounted for all estimated mortality occurring below this dam until spawning and attributed remaining mortality to the period when salmon enter the estuary and nearshore ocean.
16. No direct estimates of adult survival in the ocean exist for this ESU. We set $s_3 = s_4 = s_5 = 0.8$ [W. E. Ricker, *J. Fish. Res. Board Can.* **33**, 1483 (1976)].
17. To find f_x, the fraction of spawners of age *x* for females only, we multiplied the annual age frequencies of spawners (*8*) by the proportion of age *x* fish that are female (*11*), rescaled so the frequencies summed to 1, and averaged across the time series. Because these stocks rarely breed beyond age 5, we set $b_5 = 1$. We estimated b_3 and b_4 by solving a set of simultaneous equations:

$$f_x = b_x l_x \Big/ \sum_{i=1}^{x} b_i l_i \quad \text{for } x = 1 \text{ to } 5$$

$$\text{where } l_x = \prod_{i=1}^{x} p_i$$

$p_1 = s_1$, and $p_x = (1 - b_{x-1})s_x$ for $x > 1$ (*7*).
18. *Preliminary Information Report to the Salmon and Steelhead System Planning Documentation. Monitoring and Evaluation Group, August 1, 1989* (Northwest Power Planning Council, Portland, OR, 1989).

19. We thank M. Schiewe, G. Matthews, D. Dey, R. Zabel, C. Toole, and J. Williams for comments and S. Kiefer for fecundity data. The views expressed in this paper are not the official view of the National Marine Fisheries Service or of Santa Clara University.

20 June 2000; accepted 29 September 2000

VirB/D4-Dependent Protein Translocation from *Agrobacterium* into Plant Cells

Annette C. Vergunst, Barbara Schrammeijer, Amke den Dulk-Ras, Clementine M. T. de Vlaam, Tonny J. G. Regensburg-Tuïnk, Paul J. J. Hooykaas*

The *Agrobacterium* VirB/D4 transport system mediates the transfer of a nucleoprotein T complex into plant cells, leading to crown gall disease. In addition, several Virulence proteins must somehow be transported to fulfill a function in planta. Here, we used fusions between Cre recombinase and VirE2 or VirF to directly demonstrate protein translocation into plant cells. Transport of the proteins was monitored by a Cre-mediated in planta recombination event resulting in a selectable phenotype and depended on the VirB/D4 transport system but did not require transferred DNA.

The Gram-negative soil bacterium *Agrobacterium tumefaciens* causes crown gall disease on plants. During the infection process, a segment of the bacterial tumor-inducing (Ti) plasmid, the T region, is transferred to recipient plant cells, ultimately resulting in phytohormone overproduction (*1*–*3*). Transfer of the T region occurs as a single-stranded DNA-protein complex, resembles conjugation in many ways, and is mediated by a set of Virulence (Vir) proteins, which are encoded by the Ti plasmid. Transport requires the 11 VirB proteins, which constitute the proposed channel, and the coupling factor VirD4 (*4*, *5*).

Institute of Molecular Plant Sciences, Leiden University, Clusius Laboratory, Wassenaarseweg 64, 2333 AL, Leiden, Netherlands.

*To whom correspondence should be addressed. E-mail: hooykaas@rulbim.leidenuniv.nl

Agrobacterium strains carrying mutations in *virE2* or *virF* are avirulent on (certain) plants; however, tumors are formed upon coinoculation of these *vir* mutants with an avirulent helper strain, which lacks the T region but contains a complete *vir* region (*6*). Initially, it was thought that this "extracellular" complementation was due to the secretion of an essential enzyme or metabolite in the medium. However, it later became apparent that both the single-stranded DNA binding protein VirE2 and the F-box protein VirF have a function within the plant cell during tumor formation, given that *virE2* and *virF* mutants can incite tumors on transgenic plants that produce VirE2 or VirF, respectively (*7*, *8*). Both transferred DNA (T-DNA) transport and extracellular complementation require an intact VirB/D4 transport system; therefore, we hypothesized that the transport machinery might mediate transport of proteins besides that of the nucleoprotein T complex (*7*). However, until now it could not be formally excluded that the Vir proteins only move into the plant cell as part of a complex with the T strand.

We used the site-specific recombinase Cre from bacteriophage P1 [for review, see (*9*)] to detect translocation of a functional Cre enzyme by its fusion to VirE2 or VirF from *Agrobacterium* into recipient plant cells. To this end, we selected transgenic *Arabidopsis thaliana* C24 line 3043 (*10*, *11*) (Fig. 1), in which a *lox*-flanked (*floxed*) DNA segment prevents expression of a neomycin phosphotransferease (*nptII*) marker gene. The introduction of a T-DNA coding for Cre recombinase (*11*) in this plant line led to efficient deletion of the *floxed* DNA, resulting in the fusion of the 35*S* promoter region to the *nptII* gene as visualized by resistance to kanamycin [1.3 ± 0.6 kanamycin-resistant (Km^r) calli per root explant; Fig. 2B].

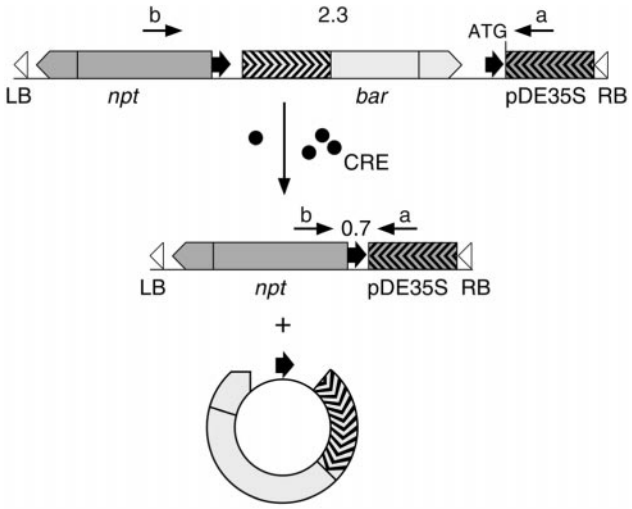

Fig. 1. Schematic representation of a Cre-mediated excision event in pSDM3043 (*11*), leading to reconstruction of a functional *lox-nptII* translational fusion. a and b are the primer binding sites. LB and RB, left and right T-DNA border sequences; *npt*, neomycin phosphotransferase; *bar*, bialaphos resistance gene; pDE35S, promoter region of the 35*S* transcript of cauliflower mosaic virus with a double enhancer sequence. Arrows indicate *lox* sites.

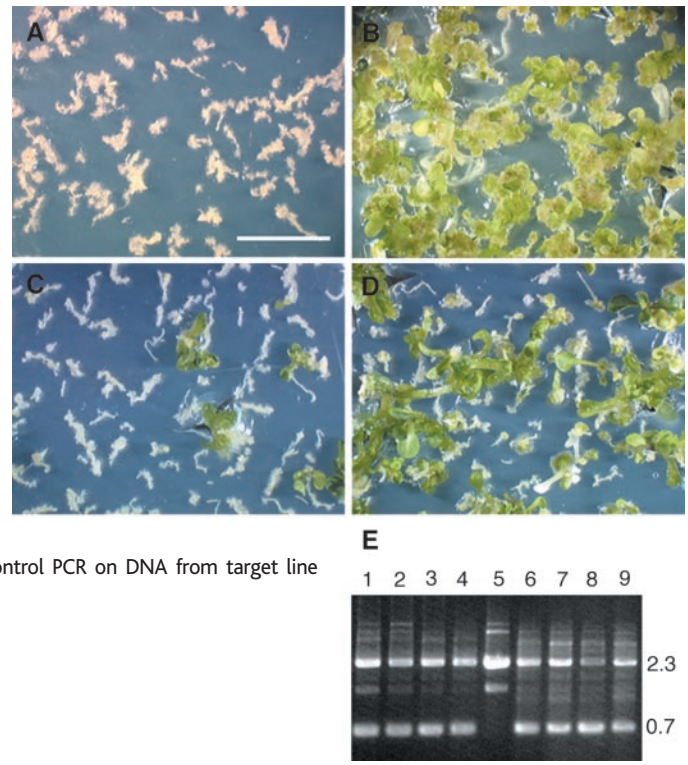

Fig. 2. Root explants of *Arabidopsis* line 3043 on medium containing kanamycin, 3 to 4 weeks after cocultivation with (**A**) LBA1100 expressing Cre, (**B**) LBA1100 harboring a T-DNA vector expressing Cre, (**C**) LBA1149 expressing the Cre::VirE2 fusion protein, or (**D**) LBA2561 expressing the NLS::Cre::VirFΔ42N fusion protein. (**E**) PCR analysis with primers a and b (Fig. 1) on Km^r shoots obtained after cocultivation with strains shown in (C) (lanes 1 to 4) and (D) (lanes 6 to 9) shows excision and original target fragments. Lane 5, control PCR on DNA from target line 3043. Scale bar, 1 cm.

Next, to assay for protein transport, Cre recombinase was expressed under control of the *vir* induction system in *Agrobacterium*, either alone or as an NH_2-terminal or COOH-terminal fusion with VirF or VirE2, respectively (*12*). Expression of the fusion proteins in *Agrobacterium* was confirmed by Western (immunoblot) analysis (*13*). Recombination activity of the fusion proteins was detected by their ability to perform an excision event in plasmid pSDM3043, which was introduced into the relevant *Agrobacterium* strains (*13*). After cocultivation of root explants (*14*) of *Arabidopsis* line 3043 with disarmed *Agrobacterium* strain LBA1100 (*15*) (wild-type *vir*, lacking T-DNA) harboring nontransmissable plasmid pRL662 (*12*) expressing the Cre recombinase alone, we detected no or sometimes a single survivor on medium containing kanamycin (on average one callus per 600 explants; Fig. 2A). This small number of Km^r background calli was also obtained upon cocultivation of plant line 3043 with *Agrobacterium* strains not expressing Cre. We can therefore conclude that bacterially expressed Cre recombinase is not transferred to plant cells from *Agrobacterium*.

However, when Cre was fused to the NH_2-terminal region of the Vir protein (Cre::VirE2; NLS::Cre::VirF), but not when fused to the COOH-terminus (VirE2::Cre; NLS::VirF::Cre), cocultivation with plant line 3043 was followed by a more efficient recovery of Km^r calli (9 ± 2 calli per 100 explants for Cre::VirF; 6 ± 2 calli per 100 explants for Cre::VirE2; Fig. 2C). There was no consistent difference when the strain expressed wild-type VirE2 or VirF protein in addition to the fusion proteins {LBA1100 compared with LBA1149 [*virE2*::Tn3*HoHo1* (*15*)] or LBA2561 [Δ*virF* (*16*)]}. Apparently, the transport channel functions such that both wild-type Vir proteins and the Cre::Vir fusion proteins are transferred efficiently and concurrently.

Polymerase chain reaction (PCR) analysis (*17*) was performed on Km^r shoots to show that resistance was indeed caused by translational fusion of the 35*S* promoter region to the *nptII* coding region because of Cre activity. A PCR reaction with primers annealing in the 35*S* promoter region (Fig. 1; primer a) and the *nptII* sequence (primer b) resulted in amplification of a 0.7-kb fragment, diagnostic for excision, whereas the expected 2.3-kb fragment was de-

tected in DNA samples from the original plant line 3043. The amplification of target DNA fragments in DNA samples from the Km^r shoots, besides the excision fragments, shows that Cre-mediated recombination occurred in cells of the original homozygous target line 3043 (Fig. 2E).

In summary, *Agrobacterium* can deliver the Cre recombinase into plant cells, resulting in detectable excision events, but only when expressed as a fusion protein attached to the NH_2-terminus of VirE2 or VirF. This implies that the COOH-termini of VirF and VirE2 need to be free to allow transport, possibly because important (transport) signals are located there. Cocultivation of 3043 root explants with *Agrobacterium* strains expressing a Cre::VirF fusion lacking 42 NH_2-terminal amino acids of VirF (*18*) (NLS::Cre::VirFΔ42N) resulted in an increase in the number of Km^r-resistant calli (54 ± 23 calli per 100 explants; Fig. 2D). This shows that the domain responsible for transport is not located in this NH_2-terminal region. In fact, this region does contain an F box (*19*), which might confer instability on the protein in plant cells or lead to retention of the protein in the cytoplasm through its binding with F-box–interacting proteins. Therefore, deletion of this domain might indirectly lead to an enhanced nuclear delivery. Additional evidence for a COOH-terminally located transport signal was obtained by using a larger 166–amino acid NH_2-terminal deletion (*18*) of VirF (NLS::Cre::VirFΔ166N). When fused to Cre and expressed in *Agrobacterium*, the remaining 37 COOH-terminally located amino acids were sufficient for obtaining Km^r calli with similar efficiency as NLS::Cre::VirFΔ42N (40 Km^r calli ± 7 per 100 root explants). Thus, we conclude that a transport signal is present in this small region. Close inspection of this area and comparison with that of VirE2 revealed the presence of a common motif of three amino acids (Arg-Pro-Arg).

Next, we examined which specific virulence functions were necessary for protein transport. Given that removal of the 42 NH_2-terminal amino acids of VirF resulted in about fivefold higher frequencies of excision after cocultivation, we transferred plasmids harboring *NLS::cre::virFΔ42N* as well as *Cre::virE2* into the *vir* mutants LBA1142 (*virA*), LBA1143 (*virB4*), LBA1144 (*virB7*), LBA1145 (*virG*), LBA1146 (*virC2*), LBA1147 (*virD2*), LBA1148 (*virD4*), and LBA1150 (*virD1*) (*15*). Additionally, *NLS::cre::virFΔ42N* was introduced into LBA1149 (*virE2*) and *Cre::virE2* into LBA2561 (*virF*). After cocultivation of 3043 root explants with strains carrying transposon insertions in *virA*, *virG*, *virB4*, *virB7*, *virD1*, *virD2*, and *virD4*, no Km^r calli were selected (Table 1), indicating that the expression of the affected genes is essential for transport of both NLS::Cre::VirFΔ42N and Cre::VirE2. In contrast, in a *virC2* mutant, protein transfer was not inhibited and calli were obtained at high efficiency. Furthermore, the VirE2 protein was apparently not essential for transport of VirF and VirF was not necessary for transport of VirE2. To rescue the distal functions of the *virD* operon in the mutants LBA1147 and LBA1150, we expressed VirD3 and VirD4 in trans in these strains. As expected, this resulted in restoration of fusion protein transport (Table 1), showing that VirD1 and VirD2 are not essential. Thus, protein translocation depends on the VirA/VirG regulatory system, necessary for the expression of the other virulence genes, and otherwise on the VirB/D4 proteins that are known to form a putative transport channel and a coupling factor, respectively.

In summary, we show directly that the *Agrobacterium* VirB/D4 transport system mediates the transfer of VirE2 and VirF proteins into plant cells independently of T-DNA transport. These data support the earlier suggestion based on extracellular complementation experiments that VirE2 and the VirD2–T-DNA nucleoprotein can be transported separately and form T complexes in the plant cell. On the basis of sequence comparison, the *Agrobacterium* VirB/D4 transport system was classified recently within a family of eubacterial transport systems, referred to as type IV secretion systems (*4*, *5*, *20*). Members include structures used by broad host range conjugative plasmids for DNA transfer, but also the Ptl transporter of the human pathogen *Bordetella pertussis* (*21*), which uses it for the secretion of proteinaceous pertussis toxins in human cells. Here, we show that a system that is involved in translocation of nucleoprotein complexes has also kept the ability to introduce monomeric proteins into recipient cells, adding a body of evidence to our earlier proposal that DNA delivery systems have evolved from protein secretion systems (*7*). This is in line with our finding that the coupling factor VirD4, of which homologs were speculated to be the interface between the relaxosome of conjugative structures and the transport apparatus, is also an essential component for protein transport.

We propose that NH_2-terminal fusions to either (parts of) VirF or VirE2 might deliver functional proteins across kingdom boundaries, for purposes in which proteins are required in recipient cells only transiently. A system based on *Agrobacterium* may be functional for plants, yeast (*22*), and fungi (*23*). The similarity between family members of eubacterial type IV secretion systems suggests that an approach similar to the one described here for *Agrobacterium* may also be used for the delivery of fusion proteins in mammalian cells by derivatives of the relevant pathogens that are attenuated in virulence.

Table 1. Efficiency of transfer of Cre::virE2 and NLS::Cre::virFΔ42N fusion proteins from different *Agrobacterium* mutants (derived from LBA 1100) in representative experiments. The number of Km^r calli was estimated 3 weeks after cocultivation with target root explants. Two Petri dishes were used per strain in each experiment. Transposon insertion mutations may affect other downstream-located genes in the operon. ND, not determined.

Bacterial strains	Transfer efficiency (number of calli per explant)			
	Cre::virFΔ42N		Cre::virE2	
	Exp. 1	Exp. 2	Exp. 1	Exp. 2
LBA1100 (wild-type *vir*)	97/190 101/270	213/310 202/330	21/300 25/325	ND
virA, *virB4*, *virB7*, *virG*, *virD4*, *virD1*	0–2/>600	ND	0–2/>600	ND
virD2	ND	0/>600	ND	0/210 0/250
virD2+*pVirD3D4*	ND	130/360 174/460	ND	29/220 19/250
virC2	79/300 131/240	ND	54/300 38/290	ND
virE2	149/435 172/335	ND	35/330 23/360	8/240 11/270
virF	157/250 147/200	242/410 242/340	25/325 26/340	8/260 2/240

References and Notes

1. P. J. J. Hooykaas, A. G. M. Beijersbergen, *Annu. Rev. Phytopathol.* **32**, 157 (1994).
2. S. B. Gelvin, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **51**, 223 (2000).
3. G. Hansen, M. D. Chilton, *Plant Biotechnol.* **240**, 21 (1999).
4. J. R. Zupan, D. Ward, P. Zambryski, *Curr. Opin. Microbiol.* **1**, 649 (1998).
5. E. Lai, C. I. Kado, *Trends Microbiol.* **8**, 361 (2000).
6. L. Otten *et al.*, *Mol. Gen. Genet.* **199**, 189 (1985).
7. A. J. G. Regensburg-Tuïnk, P. J. J. Hooykaas, *Nature* **363**, 69 (1993).
8. V. Citovsky, J. Zupan, D. Warnick, P. Zambryski, *Science* **256**, 1802 (1992).
9. D. W. Ow, S. L. Medberry, *Crit. Rev. Plant Sci.* **14**, 239 (1995).
10. *Agrobacterium*-mediated transformation of root explants was described earlier (*24*). Transgenic line

3043, segregating 3:1 on medium containing phosphinothricin, was used in further experiments.
11. A. C. Vergunst, P. J. J. Hooykaas, *Plant Mol. Biol.* **38**, 393 (1998).
12. The coding region of the *cre* gene (*25*) was fused translationally to the *virE2* and *virF* gene of pTi15955, respectively, under control of the respective *vir* promoter regions. Both NH_2-terminal and COOH-terminal fusions were made. The fusions with *virF* contained an additional NH_2-terminally located nuclear localization signal (NLS) from simian virus SV40 (*NLS*::*cre*::*virF* and *NLS*::*virF*::*cre*, respectively). The *virE2* fusions (*cre*::*virE2* and *virE2*::*cre*) were coordinately expressed with *virE1* under control of the *virE* promoter. As a control, the *cre* gene was linked directly to the *virE* promoter, at the start codon position of *virE2*. The fusion genes were cloned into broad host range, nonmobilizable plasmid pRL662. This plasmid was the result of replacing the kanamycin resistance gene and the *mob* region from pBBR1 MCS2 (*26*) by a gentamycin resistance marker.
13. A. C. Vergunst *et al.*, data not shown.
14. Roots from 10-day-old seedlings homozygous for excision locus 3043 (T3 or T4) were cocultivated for 2 days with bacteria. The explants were transferred to shoot induction medium containing kanamycin (50 mg/liter) and timentin (100 mg/liter). The number of kanamycin-resistant calli was estimated 3 weeks after cocultivation. Control cocultivations of LBA2561 (*NLS::cre::virFΔ42N*) and LBA1149 (*cre::virE2*) with wild-type *Arabidopsis* C24 root explants did not result in kanamycin-resistant calli.
15. A. G. M. Beijersbergen, A. Den Dulk-Ras, R. A. Schilperoort, P. J. J. Hooykaas, *Science* **256**, 1324 (1992).
16. B. Schrammeijer, J. Hemelaar, P. J. J. Hooykaas, *Mol. Plant Microbe Interact.* **11**, 429 (1998).
17. Chromosomal DNA was isolated from Km^r shoots (*11*). PCR reactions with primers a (5′-GAACTCGCCGTAAAGACTGGCG) and b (5′-GCGCTGACAGCCGGAACACG) were performed as follows: hot start for 5 min of 96°C, followed by 35 cycles of 1 min of 94°C, 2 min of 57°C, and 2 min of 72°C, with Ex-Taq polymerase (BioWhittaker, Belgium) according to the manufacturer.
18. A fusion was made, in which 126 5′ base pairs of *virF* were deleted, resulting in fusion protein NLS::Cre::VirFΔ42N. Deletion of the 5′ 498 base pairs of *virF* and translational fusion to *cre* resulted in fusion protein NLS::Cre::VirFΔ166N.
19. P. J. J. Hooykaas *et al.*, in *Biology of Plant-Microbe Interactions 2*, P. J. G. M. de Wit, T. Bisseling, W. J. Stiekema, Eds. (International Society for Molecular Plant-Microbe Interactions, St. Paul, MN, 2000), pp. 159–163.
20. P. J. Christie, J. P. Vogel, *Trends Microbiol.* **8**, 354 (2000).
21. A. Covacci *et al.*, *Science* **284**, 1328 (1999).
22. P. Bundock, A. Den Dulk-Ras, A. Beijersbergen, P. J. J. Hooykaas, *EMBO J.* **14**, 3206 (1995).
23. M. J. A. De Groot, P. Bundock, P. J. J. Hooykaas, A. G. M. Beijersbergen, *Nature Biotechnol.* **16**, 839 (1998).
24. A. C. Vergunst, E. C. De Waal, P. J. J. Hooykaas, *Arabidopsis Protocols*, vol. 82 of *Methods in Molecular Biology* (Humana, Totowa, NJ, 1998), pp. 227–244.
25. T. Mozo, P. J. J. Hooykaas, *Mol. Gen. Genet.* **236**, 1 (1992).
26. M. E. Kovach, R. W. Phillips, P. H. Elzer, R. M. Roop, K. M. Peterson, *Biotechnology* **16**, 800 (1994).
27. We thank our colleagues for their valuable discussions; L. Banta, E. Roine, and R. Offringa for additional comments on the manuscript; P. Hock for preparation of the figures; and J. Escudero for construction of pRL662. The investigations were supported by the Research Council for Earth and Life Sciences (ALW) with financial aid from the Netherlands Organization for Scientific Research. P. H. is also supported financially by the European Union (BI04 CT97 2028).

17 July 2000; accepted 26 September 2000

A Basal Transcription Factor That Activates or Represses Transcription

Patricia J. Willy,[1] Ryuji Kobayashi,[2] James T. Kadonaga[1]*

We have identified an activity that is required for transcription of downstream promoter element (DPE)–containing core promoters in vitro. The purified factor was found to be the *Drosophila* homolog of the transcriptional repressor known as NC2 or Dr1-Drap1. Purified recombinant dNC2 activates DPE-driven promoters and represses TATA-driven promoters. A mutant version of dNC2 can activate DPE promoters but is unable to repress TATA promoters. Thus, the activation and repression functions are distinct. These studies reveal that NC2 (Dr1-Drap1) is a bifunctional basal transcription factor that differentially regulates gene transcription through DPE or TATA box motifs.

The control of transcription by RNA polymerase II involves sequence-specific DNA binding proteins that interact with cis-acting regulatory elements, numerous coactivators and corepressors, and the structure and constitution of the chromatin template. It is important to consider, however, that the eventual target of many sequence-specific DNA binding factors and coregulators is the central machinery that mediates the basal transcription process at the core promoter. Hence, there is considerable potential for the regulation of gene activity through the basal transcription process and the core promoter.

The core promoter is generally defined as the minimal set of DNA sequences (typically about 40 base pairs) that is sufficient to direct the accurate initiation of transcription by RNA polymerase II (RNAP II) and the basal factors (*1*). There are several known core promoter motifs, which include the TATA box, the TFIIB recognition element (BRE), the initiator (Inr), and the downstream promoter element (DPE). The TATA box is an A/T-rich sequence that is located about 30 nucleotides upstream of the transcription start site, and it is bound by the TATA binding protein (TBP) subunit of TFIID. The BRE is located immediately upstream of the TATA box of some TATA-containing promoters, and increases the affinity of TFIIB for the core promoter (*2*). The Inr is a conserved sequence that encompasses the transcription start site, which functions to direct accurate transcription initiation either by itself or in conjunction with a TATA or DPE motif (*3*).

The DPE is a downstream core promoter element that is located about 30 nucleotides downstream of the transcription start site (*4–6*). The DPE is conserved from *Drosophila* to humans, and appears to be as common as the TATA box in *Drosophila* core promoters (*6*). A typical DPE-containing promoter has Inr and DPE motifs and lacks a TATA box. TFIID binds to the Inr and DPE motifs through its TAF (TBP-associated factor) subunits. The addition of a DPE in its downstream position can compensate for the loss of transcription activity that occurs upon mutation of an upstream TATA box. Thus, the activity of core promoters can be dependent on the TATA box (TATA-driven promoters) or the DPE motif (DPE-driven promoters).

We observed that high-salt nuclear extracts contain a DPE-specific activator that is deficient in low-salt nuclear extracts (Fig. 1A) (*7*). Therefore, we developed a two-template assay for the purification of this DPE-specific transcription factor (DSTF). The assay involves the simultaneous use of a DPE-driven core promoter (*Drosophila jockey*) and a TATA-driven core promoter [*Drosophila hunchback* promoter 2 (*hb*P2)] in conjunction with the low-salt nuclear extract (*8*). The presence of DSTF complements the deficiency in the low-salt nuclear extract and activates the DPE-driven promoter but not the TATA-driven promoter (Fig. 1B). In fact, the activation of the DPE-driven promoter by DSTF is accompanied by repression of the TATA-driven promoter (Fig. 1B) (*9*). By using this assay, we purified DSTF activity from the high-salt nuclear extract through seven chromatographic steps (*8*) (Fig. 1C). This purification yielded two polypeptides with apparent molecular masses of 43 kD and 22 kD that copurified with DSTF activity (Fig. 1D).

Protein microsequencing of the two DSTF

[1]Section of Molecular Biology and Center for Molecular Genetics, University of California, San Diego, La Jolla, CA 92093–0347, USA. [2]Cold Spring Harbor Laboratory, Post Office Box 100, Cold Spring Harbor, NY 11724, USA.

*To whom correspondence should be addressed. E-mail: jkadonaga@ucsd.edu