Review Hypothesis

A common mechanism of transcriptional activation by the three positive regulators, VirG, PhoB, and OmpR

Takashi Aoyama and Atsuhiro Oka

Laboratory of Molecular Biology, Institute for Chemical Research, Kyoto University, Uji-shi, Kyoto-fu 611, Japan

Received 15 December 1989

VirG, PhoB, and OmpR are positive regulators of the virulence, phosphate, and osmolarity regulons, respectively. Their target genes are attended with one or more significant blocks of similar 6-bp sequences in their regulatory regions. The distance between the block(s) and transcriptional start site varies with the target genes, but the blocks are located in a similar manner phase-specific to the promoter sequences. On the basis of these facts, we offer a novel model of transcriptional activation common to these regulators.

Transcriptional activation; Two-component regulatory system; Cis-acting element; Phased sequence; DNA-binding protein

1. INTRODUCTION

In bacteria, a large number of transcriptions respond to specific stimuli, and in some cases, several genes are controlled by a single regulatory system and constitute a regulon. Recently, it has been revealed that regulatory systems for several different regulons have common features. These systems, called the two-component regulatory systems, consist of environmental sensors and transcriptional regulators (for review, see [1]). Their members are the *Escherichia coli phoR-phoB* activating the genes for phosphate transport and assimilation (*pho* and *pst*) in response to phosphate limitation, the *E. coli envZ-ompR* regulating the porin genes (*ompC* and *ompF*) in response to osmolarity changes, the *virA-virG* of *Agrobacterium* pRi and pTi activating the virulence genes (*vir*) in response to plant phenolic compounds, the *ntrB-ntrC* (*glnL-glnG*) of enteric bacteria activating the nitrogen assimilation genes (*nif* and *gln*) in response to nitrogen limitation, and so on. The sensors and the regulators respectively contain conserved domains. Especially, the amino acid sequences of three positive regulators, VirG, PhoB, and OmpR, are highly conserved along their entire regions [2-5]. Therefore, these three proteins appear to have common molecular mechanisms for the interactions not only with the sensor proteins but also with the DNA signals. In this report, we review briefly the sequence structures of the regulatory regions of respective target genes, argue their similarities, and present a model of transcriptional activation common to the three positive regulators.

Correspondence address: T. Aoyama, Laboratory of Molecular Biology, Institute for Chemical Research, Kyoto University, Uji-shi, Kyoto-fu 611, Japan

Abbreviations: bp, base-pair(s); DMS, dimethyl sulfate; OmpR, protein coded for by *ompR*; PhoB, protein coded for by *phoB*; pRi, hairy-root-inducing plasmid; pTi, crown-gall-inducing plasmid; VirG, protein coded for by *virG*

2. SEQUENCE STRUCTURES IN THE UPSTREAM REGIONS OF *vir*, *pho*, AND *omp* GENES

For the cis-acting element recognized by VirG, some candidates have been reported with pTi *vir* genes from the sequence analysis of the upstream regions [6,7]. Recently, we precisely determined the start point of mRNA for the pRiA4 *vir* genes, and found that a block of 5'TG(A/T)AA(C/T)3' and its homologues frequently appear in the upstream regions (boxed in Fig.1a) [8]. In the case of plural blocks existing for one *vir* gene, they are phased with an interval of 11 bp. The block(s) are always located in a nearly opposite phase to the −35 region of promoters, though the distance between the block(s) and RNA start site varies with the *vir* genes. The most upstream block in each promoter region is preceded by an inverted block. These characteristic sequence structures are conserved in the upstream regions of pTiA6 *vir* genes [6,8]. Therefore, it is very likely that the phased blocks are the cis-acting element recognized by VirG. Indeed, the binding of purified VirG specifically inhibited the methylation by DMS at the G residues in the blocks (Fig.2a; our unpublished data). We call the sequence of 5'TG(A/T)AA(C/T)3' and its homologues the vir box, and argue below the presence of similar phased sequences (vir-box-like sequences) in the upstream regions of the genes regulated by PhoB or OmpR.

The upstream sequences of *phoB* and *pstS* are shown in Fig.1b [9]. In the phosphate regulon genes, the sequence which is called the pho

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

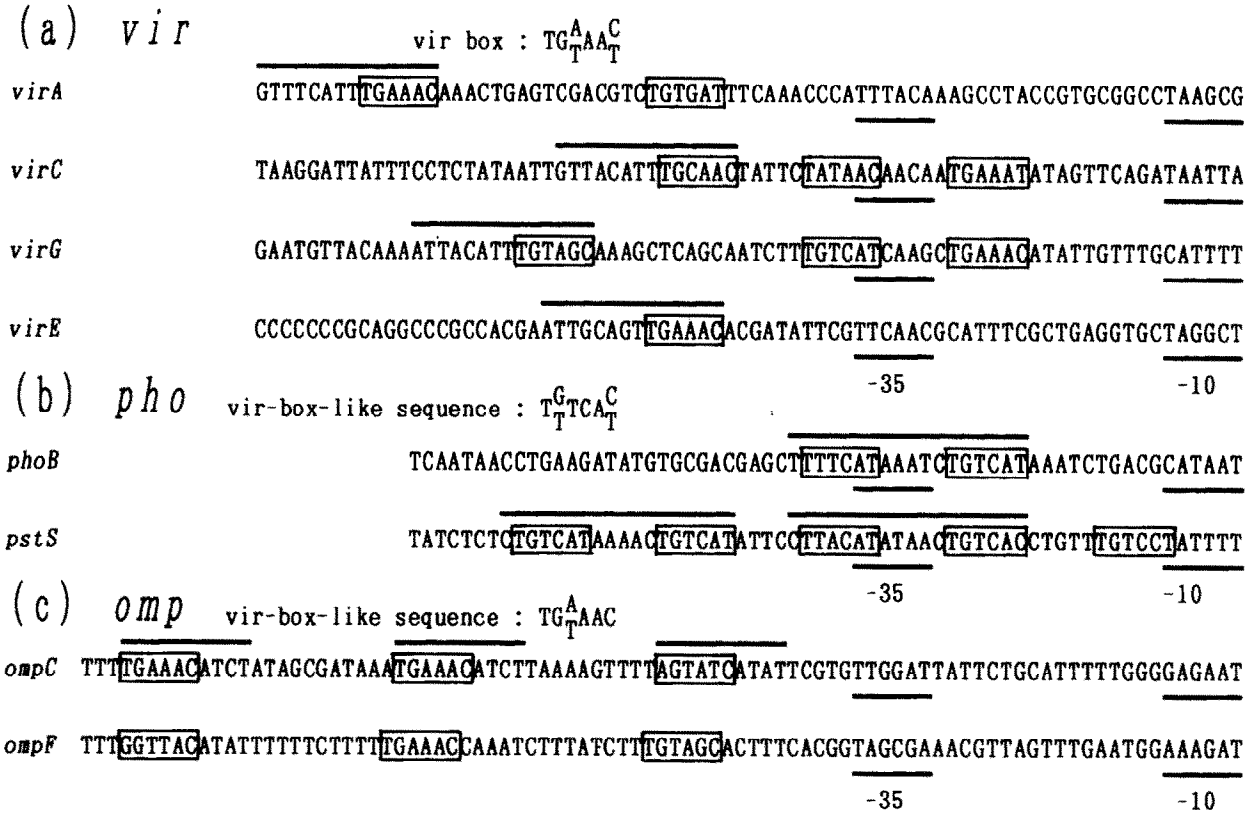

Fig.1. Nucleotide sequences of the upstream regions of *vir* [8] (a), *pho* [9] (b), and *omp* [11,12] (c). The vir boxes and the vir-box-like sequences are boxed, and the −35 and −10 regions of the promoters are underlined. The inverted vir boxes for the *vir* genes, the pho boxes for the *pho* genes, and the 10-bp repeated sequences for the *omp* genes are overlined.

box (5′CT(G/T)TCATA(A/T)A(A/T)CTGTCA(C/T)3′) has been pointed out to lie overlapping with the −35 region of the promoter [10]. The pho box is included in the actual binding region of PhoB determined by the DNase-I-footprinting experiment [9]. By scrutinizing the sequences within and surrounding the pho box, we found a vir-box-like sequence, generally similar to 5′T(G/T)TCA(C/T)3′, to repeat two times for *phoB* and five times for *pstS* at 11-bp intervals (boxed in Fig.1b). The −35 regions of the promoters are located between two vir-box-like sequences, namely in an opposite phase to the vir-box-like sequences.

Similarly, in the upstream region of *ompC*, one of the target genes of OmpR, we found another vir-box-like sequence, generally similar to 5′TG(A/T)AAC3′ (boxed in Fig.1c). This is a subset of the 10-bp sequence, 5′TGAAACATCT3′, which is repeated in the OmpR-binding region determined by the DNase-I-footprinting experiment [11]. This vir-box-like sequence and its homologues are also present at similar positions upstream from another target gene, *ompF*, though the 10-bp repeated sequences are not clear [12]. These vir-box-like sequences, unlike those for the phosphate regulon genes, are upstream apart from the −35 regions of promoters, and their intervals are 20 or 21 bp. However, it can also be said that the sequences are phased with a 10- or 11-bp interval and that the −35 region is located in an opposite phase to them. The transcription of *ompC* is partially activated in an OmpR-dependent manner when near-integral multiples of 10 bp are inserted between the OmpR-binding sites and the −35 region, but not when other numbers of base pairs are inserted [11], indicating that the opposite phase of the vir-box-like sequences to the −35 region is important for transcriptional activation.

From these facts, it was concluded that similar 6-bp sequences (the vir boxes or the vir-box-like sequences) are present in the regulatory regions of the genes regulated by VirG, PhoB, or OmpR. In all the cases, the 6-bp sequences are phased with an interval nearly equal to the pitch of the DNA helix (10.5 bp), and located in an opposite phase to the −35 region of the promoter.

3. THE MODEL FOR THE INTERACTION AMONG THE POSITIVE REGULATOR, TARGET DNA, AND RNA POLYMERASE

The nucleotide residues of which sensitivity to DMS is changed by the binding in vitro of VirG (our unpublished data) and PhoB [9] are projected on helix maps in Fig.2. Both the proteins protect the DNA molecule on the side including the major grooves of the relevant 6-bp sequences. Recently, OmpR has also been demonstrated by in vivo DMS footprinting to interact with the major grooves of the vir-box-like sequences [13]. Therefore, they are thought to line up in tandem along the one side of the DNA helix. We call such arrangement 'binding-cade'. The *E. coli* RNA polymerase with a major sigma factor approaches the promoters from one side of the DNA helix so as to make a close contact at the major grooves of the −35 and −10 regions [14]. The RNA polymerase of *Agrobacterium* probably behaves in the same manner, since the enzyme is composed of similar subunits to those of *E. coli* [15]. As shown in Fig.2, the regions thought to interact with the RNA polymerase are free from the contact of VirG and PhoB in the DNA-protein complexes. Thus, VirG or PhoB and the RNA polymerase could simultaneously bind to the promoter region without steric hindrance and face each other across the −35 region. The expected complex is shown in Fig.3a.

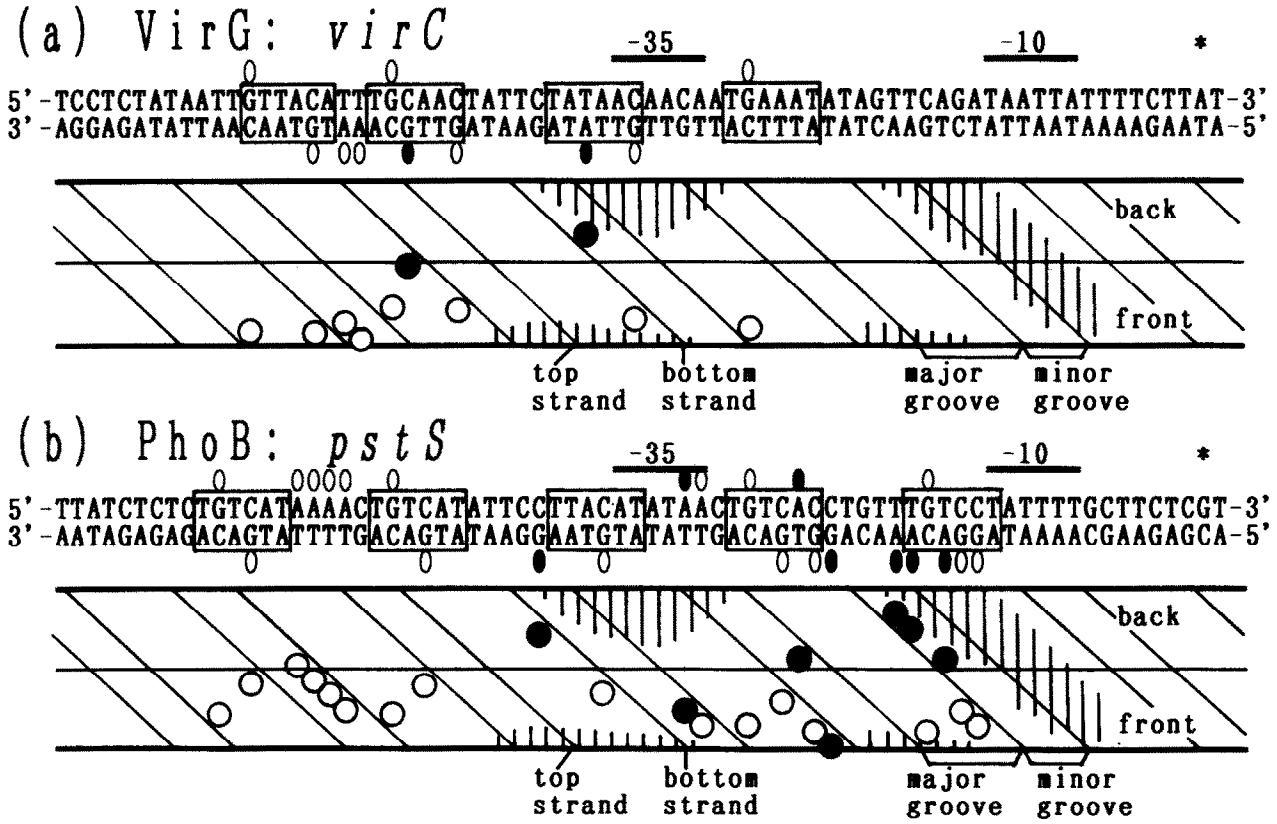

Fig.2. Helix map (10.5 bp per turn) of DMS footprinting. The nucleotide sequence is shown in alignment with each helix map. The residues changed in DMS sensitivity by VirG within the *virC* upstream region (our unpublished data) (a) and by PhoB within the *pstS* upstream region [9] (b) are shown by open circles (protected residues) and filled circles (enhanced residues). The interacting regions of phage T3 A3 promoter with *E. coli* RNA polymerase [14] (shaded with vertical lines) are superimposed on the helix maps. The vir box, inverted vir box, and vir-box-like sequence are boxed, and the −35 and −10 regions of the promoters are indicated above the sequences. RNA-start sites are indicated by asterisks.

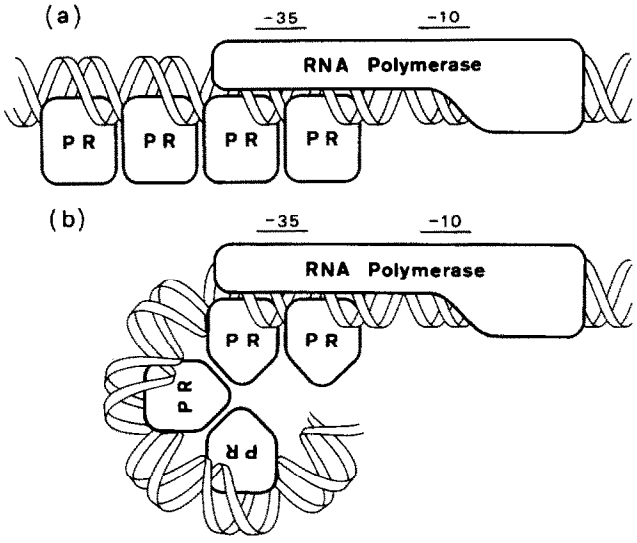

Fig.3. The binding-cade model for the complex of the positive regulator (VirG, PhoB, or OmpR), the target DNA, and the RNA polymerase. (a) Several molecules of the positive regulator (indicated by 'PR') cooperatively bind to the consecutive major grooves along the one side of the DNA helix, and face the RNA polymerase across the −35 region of the promoter. (b) Several molecules of the positive regulator cooperatively bind to the non-consecutive major grooves by the help of DNA bending.

In some *vir* genes, however, the vir box is not present at the sites flanking the −35 region. Even in such cases, VirG could bind to the sites by cooperative binding in a head-to-tail manner (Fig.3a), since the vir box is always located in an opposite phase to the −35 region. In the case that the vir boxes are not continuously arrayed, DNA bending in the intervening region might help the cooperative binding (Fig.3b). Since the arrangements of the relevant 6-bp blocks for the *omp* genes are similar to that of *virA* (see Fig.1), OmpR presumably acts in a similar manner to that of VirG. In any case, we assume that several molecules of VirG, PhoB, or OmpR are arranged along the side of the DNA helix nearly opposite to the side for RNA polymerase. We would like to call this binding model the 'binding-cade model'.

The most important assumption to be demonstrated is that VirG, PhoB, and OmpR interact with RNA polymerase at the −35 region for transcriptional activation. At present, other possibilities cannot completely be ruled out. However, if a common mechanism of transcriptional activation by these positive regulators is assumed, the model described above would be the most likely one. The binding-cade model is, at present, concerned only with the steric arrangement of the

positive regulator, the RNA polymerase, and the target DNA. The contents of molecular reactions involved should be elucidated experimentally. Recently, PhoR and EnvZ proteins have been found to phosphorylate PhoB and OmpR, respectively [16,17], but its relationship to the transcriptional activation mechanism is not clear. We believe that the binding-cade model is a helpful working model to study elementary molecular reactions for transcriptional activation.

Acknowledgements: We thank Drs K. Makino and H. Shinagawa for valuable discussion. This work was supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Science, and Culture of Japan and from the Yamada Science Foundation.

REFERENCES

[1] Ronson, C.W., Nixon, B.T. and Ausubel, F.M. (1987) Cell 49, 579–581.
[2] Aoyama, T., Hirayama, T., Tamamoto, S. and Oka, A.(1989) Gene 78, 173–178.
[3] Winans, S.C., Ebert, P.R., Stachel, S.E., Gordon, M.P. and Nester, E.W. (1986) Proc. Natl. Acad. Sci. USA 83, 8278–8282.
[4] Makino, K., Shinagawa, H., Amemura, M. and Nakata, A. (1986) J. Mol. Biol. 190, 37–44.
[5] Nara, F., Matsuyama, S., Mizuno, T. and Mizushima, S. (1986) Mol. Gen. Genet. 202, 194–199.
[6] Das, A., Stachel, S., Ebert, P., Allenza, P., Montoya, A. and Nester, E. (1986) Nucleic Acids Res. 14, 1355–1364.
[7] Winans, S.C., Jin, S., Komari, T., Leroux, B. and Nester, E.W. (1987) in: Plant Molecular Biology (Wettstein, D.V. and Chua, N.H. eds) pp. 573–582, Plenum Press, New York.
[8] Aoyama, T., Takanami, M. and Oka, A. (1989) Nucleic Acids Res. 17, 8711–8725.
[9] Makino, K., Shinagawa, H., Amemura, M., Kimura, S. Nakata, A. and Ishihama, A. (1988) J. Mol. Biol. 203, 85–95.
[10] Shinagawa, H., Makino, K., Amemura, M. and Nakata, A. (1987) in: Phosphate Metabolism and Cellular Regulation in Microorganisms (Torriani-Gorini, A., Rothman, F.G., Silver, S., Wright, A. and Yagil, E. eds) pp. 20–25, American Society for Microbiology, Washington, D.C.
[11] Maeda, S., Ozawa, Y., Mizuno, T. and Mizushima, S. (1988) J. Mol. Biol. 202, 433–441.
[12] Inokuchi, K., Mutoh, N., Matsuyama, S. and Mizushima, S. (1982) Nucleic Acids Res. 10, 6957–6968.
[13] Tsung, K., Brissette, R.E. and Inouye, M. (1989) J. Biol. Chem. 264, 10104–10109.
[14] Siebenlist, U., Simpson, R.B. and Gilbert, W. (1980) Cell 20, 269–281.
[15] Knopf, U.C. (1974) Biochem. J. 143, 511–520.
[16] Makino, K., Shinagawa, H., Amemura, M., Kawamato, T., Yamada, M. and Nakata, A. (1989) J. Mol. Biol. 210, 551–559.
[17] Aiba, H., Mizuno, T. and Mizushima, S. (1989) J. Biol. Chem. 264, 8563–8567.