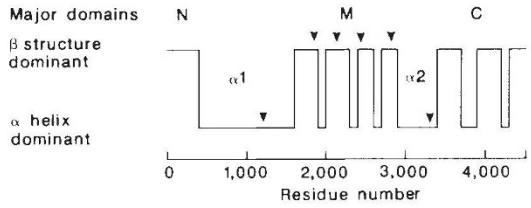

FIG. 2 Predicted secondary structure for the DβHC obtained by methods described in refs 7 and 8. These methods distinguish β-structure dominant regions and α-helix dominant regions, which are pictured here in relation to the numbered positions of the amino-acid residues deduced from the DβHC sequence. The four arrowheads on the M domain indicate the GKT, GKS1, GKS2 and SGK sites of the ATP-binding consensus sequences from the left. The two arrowheads on two long α-helix dominant regions define the trypsin cleavage sites.

of DβHC cDNA reveals the presence of four putative ATP-binding sites in the midregion of the molecule. Each of these sites may be located in one of the four predicted β-structure dominant regions of the M domain.

Does this multiplicity of putative ATP-binding sites indicate that DβHC possesses several active sites for ATP hydrolysis? Ultraviolet irradiation of DβHC in the presence of magnesium ion, ATP and vanadate cleaves the DβHC at a single site, the V1 site[14]. Because vanadate can potently inhibit dynein ATPase, probably as a result of occupying a site normally reserved for the γ-phosphate of ATP, the V1 site probably lies in the hydrolytic domain of the DβHC. As the V1 site is located 70K to the carboxy terminal side of the T1 site[14], it corresponds to the GKT site. The DβHC can be covalently modified with the hydrolysable photoaffinity ATP analogue, 8-azido adenosine 5′-triphosphate (8-N_3ATP)[15]. As 8-N_3ATP is a competitive inhibitor of dynein ATPase from *Chlamydomonas*[16], the modified portion should correspond to an ATP-hydrolysis site of DβHC. This site is located 170K to the carboxy-terminal side of the T1 site[15], suggesting that the SGK site may correspond to this 8-N_3ATP-binding site. Therefore, both the GKT and SGK sites may be functional ATP hydrolysis sites. The joint method[7] for predicting the secondary structure of DβHC makes all four ATP-binding sites equivalent, suggesting that the two GKS sites may also be functional ATP hydrolysis sites. Motor proteins such as myosin and kinesin have one ATP-binding consensus sequence in the molecule, so the presence of several ATP-binding sites, perhaps as active sites in the motor domain, may be a characteristic of the dynein superfamily.

The outer arm dynein contains two types of microtubule-binding site. The arms are thought to associate transiently with the B-subfibres of the adjacent axonemal doublet microtubles during the ATP-hydrolysis cycle, and also to have permanent attachments on the A-subfibres. The N domain may contain the microtubule-binding site for attachment to the A-subfibre. MAP2 (ref. 17) and tau[18] proteins form stable complexes with microtubules; they share a region of homologous sequences at their carboxy termini, which contains three 18-residue repeats and which has been proposed as the microtubule-binding site. These repeated sequences are not found in the N domain of the DβHC sequence. Members of the kinesin superfamily of microtubule motor proteins also associate transiently with microtubules during the ATP-hydrolytic cycle. They share a region of sequence homology extending from the ATP-binding site towards the carboxy terminal end of the molecule, which has been suggested to constitute the ATP-dependent microtubule-binding site[19]. This sequence homology was not found in the C domain or anywhere else in the DβHC sequence. DβHC seems to be a member of a new family of microtubule-binding motor proteins with unique microtubule-binding sequences which are unlike those of MAP2, tau, or members of the kinesin superfamily. □

Received 22 May; accepted 24 June 1991.

1. Vale, R. D. & Goldstein, L. S. B. *Cell* **60,** 883–885 (1990).
2. Gibbons, I. R. *J. Cell Biol.* **91,** 107s–124s (1981).
3. Ogawa, K. *et al. Eur. J. Cell Biol.* **43,** 3–9 (1987).
4. Tang, W.-J. Y. *et al. J. biol. Chem.* **257,** 508–515 (1982).
5. Sale, W. S. & Fox, L. A. *J. Cell Biol.* **107,** 1793–1797 (1988).
6. Ogawa, K. *et al. Cell Motil. Cytoskel.* **16,** 58–67 (1990).
7. Nishikawa, K. & Noguchi, T. *Meth. Enzym.* **202,** 31–44 (1991).
8. Nishikawa, K. & Ooi, T. *J. Biochem.* **100,** 1043–1046 (1986).
9. Ogawa, K. & Mohri, H. *J. biol. Chem.* **250,** 6476–6483 (1975).
10. Ow, R. A. *et al. J. biol. Chem.* **262,** 3409–3414 (1987).
11. Ogawa, K. *Proc. Japan Acad.* **67B,** 27–31 (1991).
12. Sale, W. S. *et al. J. Cell Biol.* **101,** 1400–1412 (1985).
13. Walker, J. E. *et al. EMBO J.* **1,** 945–951 (1982).
14. Gibbons, I. R. *et al. J. biol. Chem.* **262,** 2780–2786 (1987).
15. Tang, W.-J. Y. & Gibbons, I. R. *J. biol. Chem.* **262,** 17728–17734 (1987).
16. Pfister, K. K. *et al. J. biol. Chem.* **259,** 8499–8504 (1984).
17. Lewis, S. L. *et al. Science* **242,** 936–939 (1988).
18. Lee, G. *et al. Science* **239,** 285–288 (1988).
19. Yang, J. T. *et al. Cell* **56,** 879–889 (1989).
20. Kozak, M. *Cell* **44,** 283–292 (1986).
21. Matsudaira, P. *J. biol. Chem.* **262,** 10035–10038 (1987).

ACKNOWLEDGEMENTS. I thank K. Nishikawa of the Protein Engineering Research Institute for predicting the secondary structure of the DβHC; C. J. Brokaw of the California Institute of Technology for critically reading the manuscript; H. Mori and M. Hayashi for synthesising oligonucleotides; H. Kajiura for amino-acid sequencing; and K. Inaba for preparing f3 peptide from sperm 21S dynein. The sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number D01021.

'Pseudo' domains in phage-encoded DNA methyltransferases

C. Lange, A. Jugel, J. Walter, M. Noyer-Weidner & T. A. Trautner

Max-Planck-Institut für molekulare Genetik, Ihnestrasse 73, 1000 Berlin 33, Germany

5-Cytosine-DNA-methyltransferases, which are found in many organisms ranging from bacteriophages to mammals, transfer a methyl group from *S*-adenosylmethionine to the carbon-5 of a cytosine residue in specific DNA target sequences[1]. Some phage-encoded methyltransferases methylate more than one sequence: these enzymes contain several independent target-recognizing domains each responsible for recognizing a different site. The amino-acid sequences of these multispecific methyltransferases reveal that some enzymes in addition carry domains that do not contribute to the enzymes' methylation potential, but strongly resemble previously identified target-recognizing domains. Here we show that introducing defined amino-acid alterations into these inactive domains endows these enzymes with additional methylation specificities. Gel retardation analysis demonstrates that these novel methylation specificities correlate with the acquisition of additional DNA-binding potential of the proteins.

Comparisons of the primary structures of about 40 different methyltransferases revealed at least four highly conserved amino-acid sequence motifs in identical order but variable in spacing (Fig. 1). At the same relative location with respect to a core of conserved sequences, all methyltransferases have in addition an extended region that is variable in size and composition between different enzymes (Fig. 1). The conserved core sequences are thought to be responsible for common functions, such as *S*-adenosylmethionine (SAM) binding and the methyl group transfer[2–5], whereas target recognition has been assigned to the variable region. Enzymes with multiple specificities have several independent domains within the variable region of the enzymes, each of which recognizes a different DNA target[5–8]. These target-recognizing domains (TRDs) represent contiguous, consecutively arranged amino-acid segments, comprising from 38 to at least 59 amino acids (T.A.T. *et al.*, unpublished results). Although TRDs specifying the recognition of different targets vary in size and amino-acid composition, they share several conserved amino acids, providing a 'consensus' sequence (Fig. 2)[2].

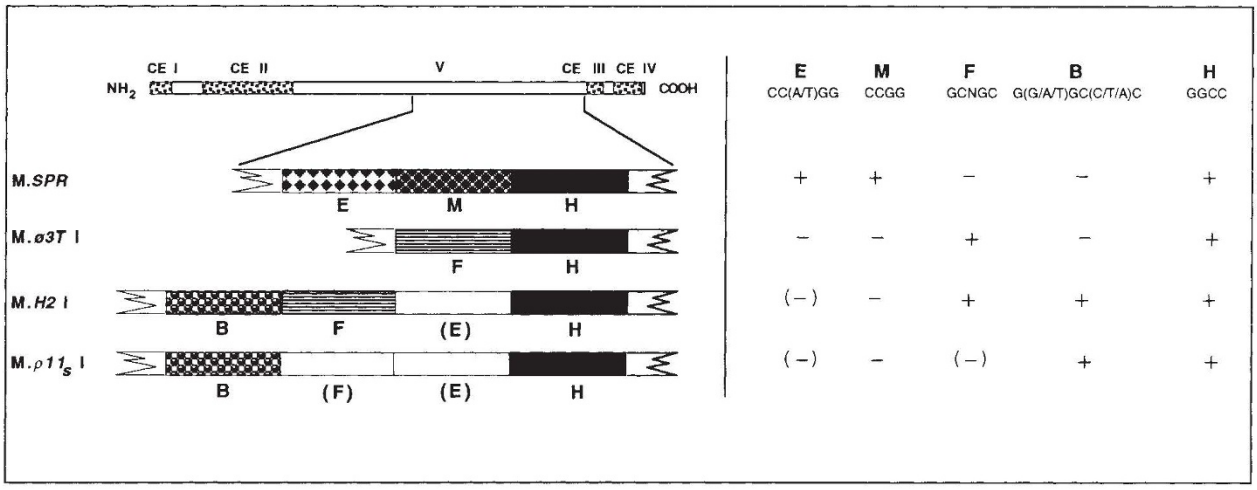

FIG. 1 Schematic presentation of the building plan of multispecific methyltransferases. The top line shows the arrangement of conserved sequence elements (CE I-CE IV), also present in monospecific methyltransferases, and the variable region (V) as described previously[2]. These enzymes each have about 400 amino acids. Enlargements of parts of the variable regions of the multispecific methyltransferases encoded by *B. subtilis* phages SPR (ref. 10), ϕ3T (ref. 11) and $\rho 11_S$ (ref. 6) and *B. amyloliquefaciens* phage H2 (ref. 9), showing the relative positions of the different TRDs (refs 5, 8), are shown below. E, M, F, B, and H identify TRDs recognizing CC($^{A}/_{T}$)GG (*Eco*RII), CCGG (*Msp*I), GCNGC (*Inu*4HT), G(G/A/T)GC(C/T/A)C (*Bsp*1286) and GGCC (*Hae*III) target sequences, respectively. Inactive TRDs are indicated by bracketed letters. The methylation capacities of the different phage methylases (M.*SPR*, M.ϕ*3T*I, M.*H2*I and M.$\rho 11_S$I) are shown in the right-hand part of the figure. (−), Activatable specificities.

With most multispecific methyltransferases, the number of different targets that can be methylated correlates with the number of TRDs. Recently, however, we have found two multispecific enzymes that contained more TRD-like motifs than could be accounted for by their methylation potential. Thus M.*H2*I contains a region between the TRDs F and H, which strongly resembles the TRD E identified in M.*SPR* (ref. 9) (see Fig. 1). There is also a TRD E-related sequence in M.ρ*11s*I, which in addition has a silent sequence motif related to the TRD F identified in M.ϕ*3T*I (ref. 6). *In vivo* or *in vitro* methylation by these enzymes does not render DNA resistant to degradation by the relevant cognate restriction endonucleases. Neither do such enzymes bind to oligonucleotides containing the targets for either *Eco*RII or *Fnu*4HI (Fig. 4) recognized by TRDs E or F. The 'inert' TRDs do not have any undetected specificity because the partially purified enzymes do not incorporate methyl groups into DNA multiply methylated *in vivo* at *Hae*III, *Bsp*1286, *Fnu*4HI and *Eco*RII target sites (data not shown).

The alignment of active and inactive F TRDs (Fig. 2*a*) suggests that the presence of Gly at location 291 is crucial for the

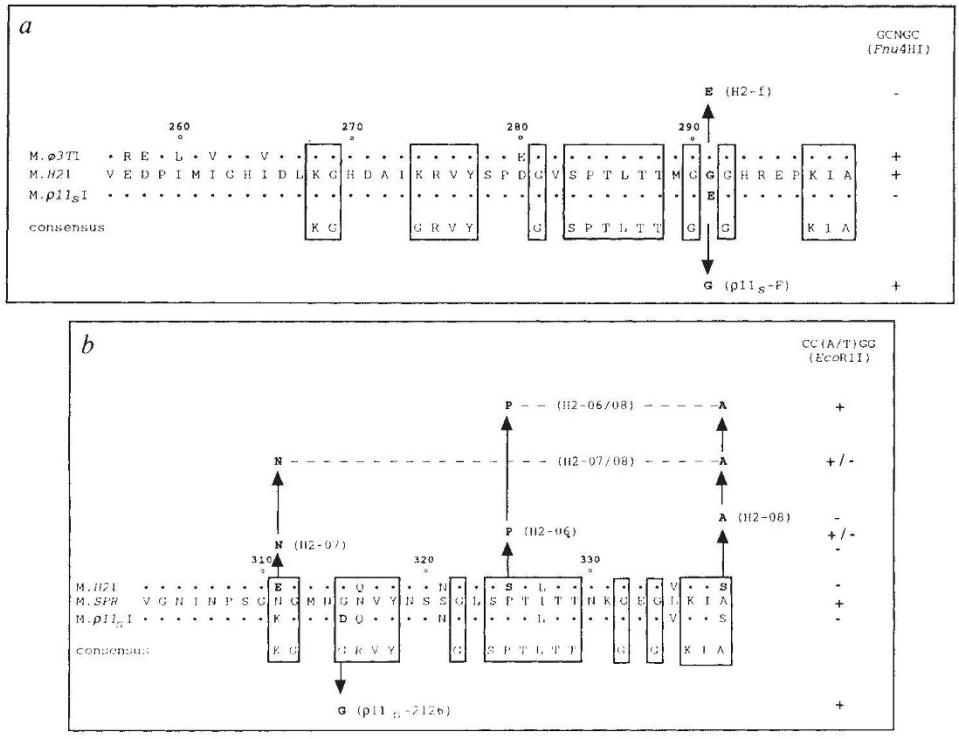

FIG. 2 Amino-acid sequence alignments of active, inactive and activated TRDs identified in multispecific methyltransferases. Amino acids of the wild-type enzymes, which were exchanged by others (indicated by arrows), are emphasized by bold letters. The numbers refer to amino-acid coordinates as described previously[6]. Activities of the wild-type and mutagenized TRDs are shown on the right-hand side, with +, +/− and − indicating high, reduced and absence of methylation activity, respectively, as defined through *in vitro* restriction analyses (Fig. 3). *a*, Amino-acid sequences of active *Fnu*4HI specific TRDs F (Fig. 1) of M.ϕ*3T*I and M.*H2*I and the highly related inactive TRD (F) of M.$\rho 11_S$I. H2-f and $\rho 11_S$-F designate the mutant enzymes, which were obtained by introducing single amino-acid exchanges into M.*H2*I and M.$\rho 11_S$I as indicated. *b*, Structure of the *Eco*RII specific TRD E (Fig. 1) of M.*SPR* and inactive TRDs (E) identified in M.*H2*I and M.$\rho 11_S$I. H2-06, -07, -08 and $\rho 11_S$-2126 are mutant enzymes with single amino-acid exchanges, H2-07/08 and H2-06/08 are double mutants as indicated in the figure. The sequences are aligned to the consensus sequence of TRDs as described previously[2]. Amino acids of the TRDs belonging to the consensus are boxed. Oligonucleotide-directed mutagenesis followed conventional procedures as described elsewhere[12] and involved appropriate 21-mer oligonucleotide primers. In all cases mutagenesis was verified by DNA sequencing[13] of the regions concerned.

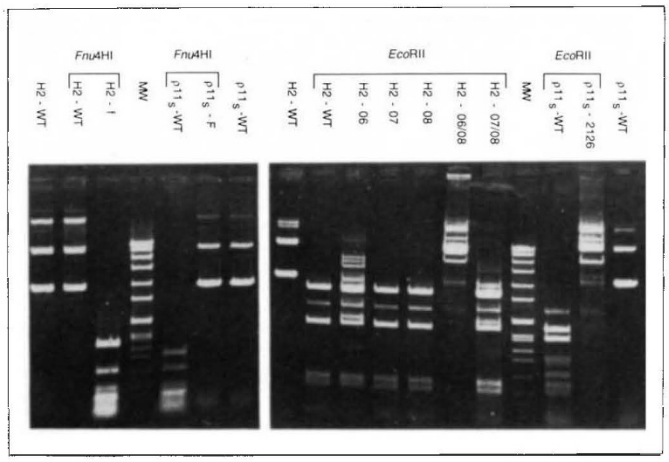

FIG. 3 Electrophoretic separation of plasmid DNA modified by M.*H2*I, M.$\rho 11_S$I and their mutant derivatives after incubation with restriction endonucleases R.*Fnu*4HI and R.*Eco*RII. Methyltransferase genes with original (WT) and mutagenized TRDs were cloned into plasmid pMS119, which is a derivative of pJF118[14], and transformed into *E. coli* K12 strain GM271 (*mcrA*$^-$, *mcrB*-1, *hsdR*-2, *dcm*-6; obtained from M. G. Marinus). Following preparation of plasmid DNA from these transformants[15], DNA (1 μg) was incubated for 1 h with an excess of the appropriate restriction enzymes and subsequently loaded on a 1% agarose gel. The proteins that were expressed by the different plasmids are indicated at the top of the figure as described in Fig. 2. SPP1 DNA digested with R.*Eco*RI served as a size marker (MW)[16].

function of this TRD, as this is the only position where the sequences of M.*H2*I and M.$\rho 11_S$I diverge. Indeed, changing Glu 291 to Gly provided M.$\rho 11_S$I with an additional *Fnu*4HI specific methylation capacity, whereas the reverse mutation introduced into M. *H2*I eliminated methylation of *Fnu*4HI sites. Apparently other amino-acid differences between M.$\phi 3T$I and M.*H2*I are irrelevant to the function of this domain.

The inactive E-like TRDs from M.$\rho 11_S$I and M.*H2*I differ from the M.*SPR* sequence at seven amino-acid positions (Fig. 2*b*). Our attempts to activate the H2-encoded domain concentrated on those amino acids found in the TRD consensus sequence. According to this rationale we introduced the following single amino-acid exchanges into M.*H2*I: Glu 311 → Asn, Ser 325 → Pro and Ser 338 → Ala (Fig. 2*b*). The conversion from Ser 325 into Pro induced an *Eco*RII specific methylation activity of the enzyme (H2-06 in Fig. 2*b*), which makes the plasmids expressing the mutagenized gene partially resistant to *Eco*RII specific restriction (Fig. 3). Introduction of Asn 311 (H2-07) and Ala 338 (H2-08) had no effect on the phenotype of M.*H2*I. In addition, double mutants with changes Ser 325 → Pro/Ser 338 → Ala (H2-06/08) and Glu 311 → Asn/Ser 338 → Ala (H2-07/08), gave full or partial resistance to *Eco*RII restriction (Fig. 3). These results also provided some rationale for the mutagenesis of the inactive domain (E) of M.$\rho 11$sI, which differs from the corresponding M.*H2*I sequence at three positions. Based on this and the TRD consensus sequence, we replaced Asp 315 of M.$\rho 11s$I with Gly present at this position in the active E TRD of M.*SPR* (Fig. 2*b*), resulting in *Eco*RII specific methylation activity in M.$\rho 11s$I($\rho 11_S - 2126$ in Figs 2*b* and 3). By gel retardation we could show that the amino-acid exchanges responsible for activating the nonfunctional TRDs also provide them with the expected new DNA binding specificities (Fig. 4).

These results support the concept of a strictly modular organization of multispecific methyltransferases. In these enzymes a highly conserved core provides the DNA methylation function, and this can be associated with various DNA recognition domains, which determine the methylation specificity. These TRDs function irrespective of their relative location to each other. Thus, the active E TRD of M.*SPR* and the activatable related domains of M.*H2*I or M.$\rho 11_S$I occupy different relative locations within the variable region of these enzymes (Fig. 1). Similarly, the activated F domain of M.$\rho 11_S$I has a location in this methyltransferase that is different from the equivalent domain in M.$\phi 3T$I.

The inert TRDs are probably 'down' mutations of active domains. This follows from the observation that the number of amino acids in inert and the matching active TRDs is the same (the TRDs identified in multispecific methyltransferases are characterized by distinct numbers of amino acids), and by their

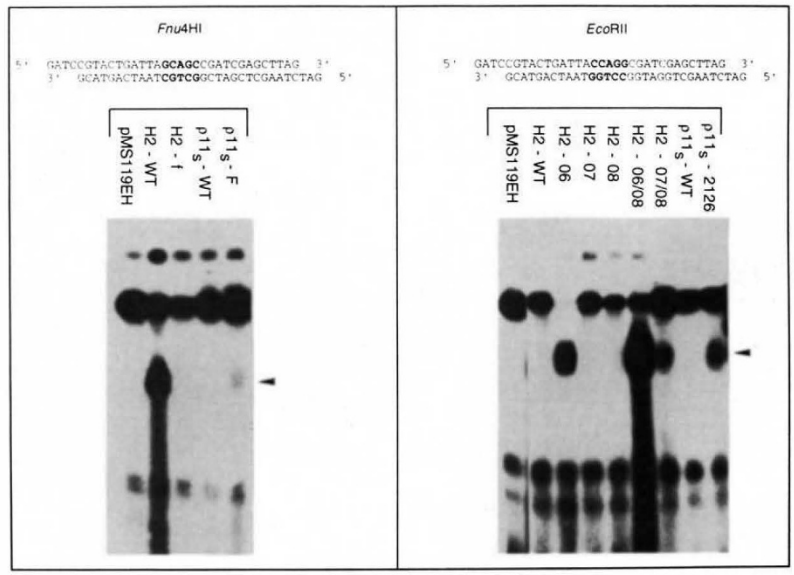

FIG. 4 Gel retardation analysis performed with M.*H2*I and M.$\rho 11_S$I and mutant derivatives. Crude extracts of *E. coli* K12 cells overproducing the various methyltransferases indicated above the gel tracks were incubated with the synthetic oligonucleotides shown, containing either a *Fnu*4HI or a *Eco*RII target site. A crude extract prepared from cells harbouring pMS119EH, the vector used for methyltransferase overproduction, served as control. DNA–protein complexes were separated by electrophoresis on nondenaturing polyacrylamide gels. In addition to a 'background' complex, detectable also with the control extract, specific methyltransferase–DNA complexes, whose position is indicated by the arrow, were formed in some of the extracts analysed. Binding of the *Fnu*4HI-specific oligonucleotide was only detectable with methyltransferases methylating *Fnu*4HI target sites (H2-WT, $\rho 11_S$-F). The same is true for the *Eco*RII-specific oligonucleotide, where the formation of specific complexes is observed only with those enzymes which contain partially (H2-06, H2-07/08) or fully active (H2-06/08, $\rho 11_S$-2126; see Fig. 2) *Eco*RII specific TRDs.

METHODS. Cultures (2 ml) of *E. coli* K12 GM271 (for genotype see legend of Fig. 3) harbouring the different methyltransferase-expressing plasmids were induced with IPTG (2 mM) at early log phase and grown for further 4 h. The cells were collected, suspended in 50 mM Tris-HCl (pH 7.5), 50 mM NaCl, 100 mM EDTA, 5 mM DTE and disrupted by sonification, as described elsewhere (A.J. *et al.*; manuscript in preparation). Binding reactions were performed by incubation of 1 nM of the ^{32}P-end-labelled oligonucleotides with 2 μl of the crude extracts in buffer containing final concentrations of 50 mM Tris-HCl (pH 7.5), 50 mM NaCl, 10 mM EDTA, 5 mM DTT and 50 μM *S*-adenosylhomocysteine. Samples were incubated for 5 min at room temperature followed by separation of protein–DNA complexes on 5% polyacrylamide gels at 5 V cm^{-1}.

great sequence similarity. The identity of N-terminal amino acids of active and inactive TRDs F of M.*H2*I and M.$\rho 11_S$I suggests that the inactive TRD of M.$\rho 11_S$I derived from M.*H2*I. The inactivating mutation must have been a recent event, as the nucleotide sequences of the two regions differ only in the base causing the Glu to Gly change at position 291. The nucleotide sequences encoding the inactive TRDs (E) of M.$\rho 11_S$I and M.*H2*I differ from each other in five positions, three of which account for the amino-acid differences and two of which are null. Comparing the inactive E TRDs of M.$\rho 11_S$I and M.*H2*I to the active E TRD of M.*SPR* (Fig. 2*b*) shows that the evolution of the inactive TRDs from an active ancestor must have followed separate avenues.

These inert TRDs represent an intriguing case of preservation of nonfunctional genetic information. This information has been preserved in the absence of positive selective pressure, which in the case of the phage-encoded modification methyltransferases would be exerted by a bacterial restriction and modification system. Negative selection, counteracting the maintenance of genetic information encoding inactive TRDs in the phage genome, has also apparently not acted.

Thus the silent TRDs represent a reserve of genetic material of general evolutionary importance as suggested, for example, by the presence of 'pseudogenes' in some bacterial and eukaryotic genomes. The inactive TRDs described here—although 'expressed' within the context of the active methyltransferases—would operationally fulfil the role of pseudogenes. For this reason we term the inactive domains 'pseudo' TRDs. Considering the modular structure of the methyltransferases discussed here it is conceivable that maintenance of these pseudo domains is significant for the generation of new specificities. □

Received 8 May; accepted 24 June 1991.

1. Razin, A., Cedar, H. & Riggs, A. D., eds. *DNA Methylation: Biochemistry and Biological Significance*, Springer Verlag, Berlin, 1984.
2. Lauster, R., Trautner, T. A. & Noyer-Weidner, M. *J. molec. Biol.* **206,** 305–312 (1989).
3. Pósfai, J., Bhagwat, A. S., Pósfai, G. & Roberts, R. J. *Nucleic Acids Res.* **17,** 2421–2435 (1989).
4. Ingrosso, D., Fowler, A. V., Bleibaum, J. & Clarke, S. *J. biol. Chem.* **264,** 20131–20139 (1989).
5. Wilke, K. *et al. EMBO J.* **7,** 2601–2609 (1988).
6. Behrens, B. *et al. EMBO J.* **6,** 1137–1142 (1987).
7. Balganesh, T. S. *et al. EMBO J.* **6,** 3543–3549 (1987).
8. Trautner, T. A., Balganesh, T. S. & Pawlek, B. *Nucleic Acids Res.* **16,** 6649–6658 (1988).
9. Lange, C., Noyer-Weidner, M., Trautner, T. A., Weiner, M. & Zahler, S. A. *Gene* **100,** 213–218 (1991).
10. Buhk, H.-J. *et al. Gene* **29,** 51–61 (1984).
11. Tran-Betcke, A., Behrens, B., Noyer-Weidner, M. & Trautner, T. A. *Gene* **42,** 89–96 (1986).
12. Kunkel, T. A. *Proc. natn. Acad. Sci. U.S.A.* **82,** 488–492 (1985).
13. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5463–5467 (1977).
14. Fürste, J. P. *et al. Gene* **48,** 119–129 (1986).
15. Sambrook, J., Fritsch, E. F. & Maniatis, T. (eds) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1989).
16. Ratcliff, S. W. *Mol. gen. Genet.* **168,** 165–172 (1979).

CORRECTION

Identification of a G1-type cyclin *puc1*$^+$ in the fission yeast *Schizosaccharomyces pombe*

Susan L. Forsburg & Paul Nurse

Nature **351,** 245–248 (1991)

WE have discovered an error in the DNA sequence of our *puc1*$^+$ gene in Fig. 1 of the above paper. The CG pair at nucleotides 710–711 should be a GC pair. The amino acids underneath should thus be Leu Gln instead of Phe Glu. However, the protein sequence in Fig. 2 is correct. □