Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 4159–4162, May 1992
Biochemistry

A genetic system for studying the activity of a proteolytic enzyme

BIMALENDU DASMAHAPATRA*†, BETH DIDOMENICO‡, SHEILA DWYER*, JUN MA§, IVAN SADOWSKI§¶, AND JEROME SCHWARTZ*

*Antiviral Chemotherapy, ‡Antifungal Chemotherapy, Schering-Plough Research Institute, Bloomfield, NJ 07003; and §Department of Biochemistry and Molecular Biology, Harvard University, 7 Divinity Avenue, Cambridge, MA 02138

Communicated by Mark Ptashne, January 27, 1992 (received for review June 27, 1991)

ABSTRACT **We describe a genetic system for monitoring the activity of a specific proteolytic enzyme by taking advantage of the properties of the yeast transcriptional activator GAL4. The GAL4 protein contains two separable and functionally essential domains: the amino-terminal DNA binding domain and the carboxyl-terminal transcriptional activating domain. We constructed two hybrid proteins by inserting between the DNA binding domain and the activation domain of GAL4 either (*i*) a self-cleaving protease (3C protease of a picornavirus, coxsackievirus B3) or (*ii*) a mutant form of the protease that is unable to cleave. We show that, although the hybrid protein containing the mutant protease activates transcription of *GAL1–lacZ* reporter gene, the hybrid protein bearing the wild-type protease is proteolytically cleaved and fails to activate transcription. Our approach to monitor the proteolytic activity could be used to develop simple genetic systems to study other proteases.**

Proteases play a critical role in the regulation of many biological processes (1, 2). In particular, proteolysis of primary polypeptide precursors is essential to replication of several viruses. For example, picornaviruses, a group of small positive-strand RNA viruses, as well as retroviruses, including human immunodeficiency virus type 1 (HIV-1), encode proteins that are synthesized as large polyprotein precursors and are later proteolytically processed to mature viral proteins (3, 4). These proteases are very specific for their cognate substrates and have been studied using various techniques such as electrophoresis and chromatography (5, 6). Other biochemical methods that have been developed recently to assay the HIV-1 protease activity in multiple samples include resonance-energy transfer (7), release of radioactive products (8), and spectrophotometry of chromogenic peptide substrates (9). As an alternative, we have developed a genetic assay that could be used not only to screen large numbers of compounds for specific protease inhibitors but also to detect protease-deficient mutants.

To demonstrate our genetic system for monitoring protease activity we chose the 3C protease of coxsackievirus (CV) B3 (10), a member of the picornavirus family. Picornaviruses, a group of small positive-strand RNA viruses, produce all of their gene products by proteolytic cleavage of a single polyprotein precursor (11). The 3C protease, which is conserved among the members of picornavirus family, plays a major role during the virus replication cycle (12). The protease is contained within the polyprotein precursor and functions intramolecularly to cleave itself out of the precursor and also catalyzes other specific cleavages within the polyprotein to generate mature viral proteins (13, 14).

We have exploited the autocatalytic activity of the 3C protease to regulate the yeast transcriptional activator GAL4. GAL4, a protein of 881 amino acids, binds specifically to the upstream activating sequence (UAS_G) and stimulates transcription of the adjacent *GAL1* and *GAL10* genes (15, 16). The amino-terminal fragment (amino acids 1–147) of GAL4 binds specifically to DNA *in vivo* but does not activate transcription unless it is linked to one of the two activating regions, regions I and II, contained in the carboxyl end of the protein (16). We have fused the 3C protease coding sequence between the two functional domains of GAL4. When expressed in yeast, this modified GAL4 protein is cleaved by the autocatalytic activity of the viral protease and fails to activate transcription. An otherwise identical GAL4 hybrid protein with a mutant 3C protease inserted between the two functional domains of GAL4 is not proteolytically cleaved and therefore is capable of activating transcription from the GAL1 promoter. Thus transcriptional activation by GAL4, which is easily monitored by a reporter gene expression, provides a simple assay for the 3C protease.

MATERIALS AND METHODS

Strains and Media. The yeast strain used was GGY1::RY171 (Δ*gal4*, Δ*gal80*, *ura3*, *his3*, *leu2*; carries a *GAL1–lacZ* fusion integrated at the *URA3* locus of the genome) (17). Yeasts were grown in either YEPD-rich medium or minimal media (18) that lacked histidine and uracil, and the carbon sources of induction media were 2% galactose, 2% glycerol, and 2% ethanol instead of glucose. *Escherichia coli* strain XL1-Blue (Stratagene) was used to perform all plasmid construction, and bacteria were grown in LB medium with the appropriate antibiotic (ampicillin, 100 μg/ml) for selection of plasmids.

Plasmid Construction. The plasmid pBD1009, which expresses the GAL4/3C protease hybrid protein, was constructed as described below. An *Xba* I–*Bam*HI DNA fragment of about 1100 base pairs (bp), encoding the CV 3C protease and the flanking regions, was isolated from the plasmid pC1B1 (10). The isolated fragment was further digested with *Hin*fI and treated with Klenow fragment of *E. coli* DNA polymerase to fill in the ends. The 795-bp DNA fragment was then purified and ligated to the plasmid pMA236, which had been linearized with *Eco*RI, treated with alkaline phosphatase to remove 5′ phosphate groups, and blunt-ended with Klenow polymerase. To construct the plasmid pBD1013, we used the plasmid pC11B9 (10) containing the mutant 3C protease [which was constructed by inserting a 12-bp linker fragment (GGATCCGGATCC) within the 3C coding sequence of the plasmid pC1B1]. The mutation, which inserts four amino acids (Asp-Pro-Asp-Pro) after the Gly-166 in the protease sequence, abolishes the proteolytic activity (10, 19).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: MUG, methylumbelliferyl galactoside; HIV-1, human immunodeficiency virus type 1; UAS_G, upstream activating sequence galactose; CV, coxsackievirus.

†To whom reprint requests should be addressed.
¶Present address: Department of Biochemistry, University of British Columbia, Health Sciences Mall, Vancouver, BC, V6T 1Z3 Canada.

Yeast Transformation and Assay of β-Galactosidase. Yeast cells were made competent for transformation by treatment with lithium acetate (20). To assay β-galactosidase, 50 μl of the induced yeast culture (10^6 cells) was aliquoted into a 96-well microtiter plate and lysed by adding one drop each of 0.1% SDS and chloroform (21). Cells were incubated with 100 μl of freshly prepared MUG reagent mix [4:1 (vol/vol) Z buffer (60 mM Na_2HPO_4/40 mM NaH_2PO_4/10 mM KCl/1.0 mM $MgSO_4$/50 mM 2-mercaptoethanol) and MUG reagent (12.5 mg of methylumbelliferyl galactoside in 100 ml of distilled water)] at 30°C for 1 hr and read on the DYNATECH plate reader at 340 nm. β-Galactosidase activity for each sample was determined in quadruplicate, and the average value was expressed in fluorescence (MUG) units (22, 23).

RESULTS AND DISCUSSION

The basic strategy of our experiments is shown in Fig. 1. The native GAL4 protein, an activator of transcription, contains two functional domains (Fig. 1*a*). The amino-terminal domain binds specifically to UAS_G but fails to stimulate transcription. The carboxyl-terminal domain contains two activating regions, region I (amino acids 148–236) and region II (amino acids 768–881), either of which activates transcription when fused to the DNA binding domain (amino acids 1–147). A chimeric protein that contains an autocatalytically active 3C protease fused in frame between the two functional domains of GAL4 protein is proteolytically cleaved. Because an activator requires DNA binding and activating domains, the autoprocessing activity of the 3C protease dissociates the two domains of GAL4, thus inactivating the protein (Fig. 1*b*). However, the chimeric protein bearing an inactive protease is not proteolytically processed and is able to activate transcription by maintaining the physical association of the two functional domains of GAL4 protein (Fig. 1*c*).

The chimeric proteins that are expressed from the recombinant plasmids and are used in the experiments are shown in Fig. 2. The hybrid protein expressed from the plasmid pBD1009 comprises GAL4 (amino acids 1–147), CV B3 peptide [(amino acids 1505–1770) (24), containing the carboxyl-terminal 13 amino acids of 3A, all of 3B and 3C, the amino-terminal 47 amino acids of the 3D protein], and GAL4 (amino acids 768–881). This GAL4/CVB3 hybrid protein contains three potential cleavage sites for the 3C protease (Fig. 2). As a control, we constructed a similar fusion protein

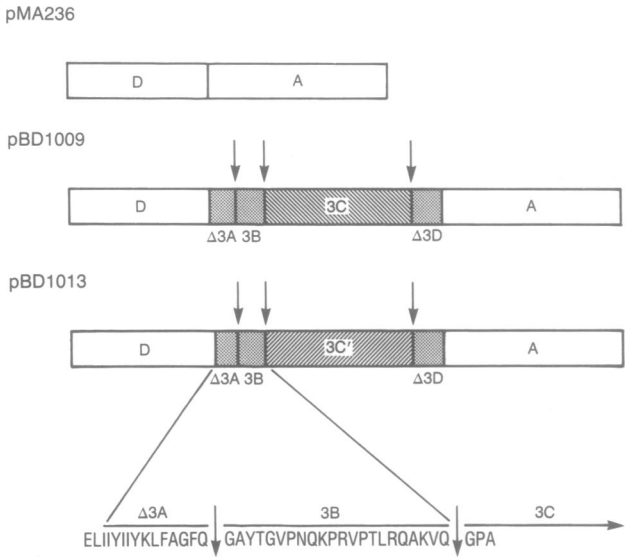

FIG. 2. Constructs expressing GAL4 derivatives. The GAL4 derivative containing the DNA binding domain, D (amino acids 1–147), fused to activation region II, A (amino acids 768–881), is expressed from the plasmid pMA236 (16). The hybrid peptide expressed from the plasmid pBD1009 provides three potential cleavage sites for the 3C protease: Δ3A/3B, 3B/3C, and 3C/Δ3D (shown by arrows). The hybrid protein expressed from the plasmid pBD1013 contains a mutant protease (3C') instead of the wild-type 3C protease flanked by the same viral and the GAL4 sequences. The amino acid sequences surrounding the cleavage sites between Δ3A and 3B and 3B and 3C are given in single-letter code.

using a mutant form of the CVB3 3C protease that is defective in the autocatalytic and trans processing activities (10). We introduced each of these plasmids into the yeast strain GGY1::RY171 (17), which is deleted for *GAL4* but contains, at the *URA3* locus, an integrated *GAL1–lacZ* fusion gene that can be activated by GAL4 protein. β-Galactosidase activity was used as a measure of the transcriptional activation function of each GAL4 hybrid protein expressed.

Results of these experiments are shown in Table 1. As expected, there was no β-galactosidase activity in the absence of any GAL4-expressing plasmid. Approximately 1800 MUG units (22) of β-galactosidase were produced in cells expressing the GAL4 derivative [GAL4(1–147)(768–881)];

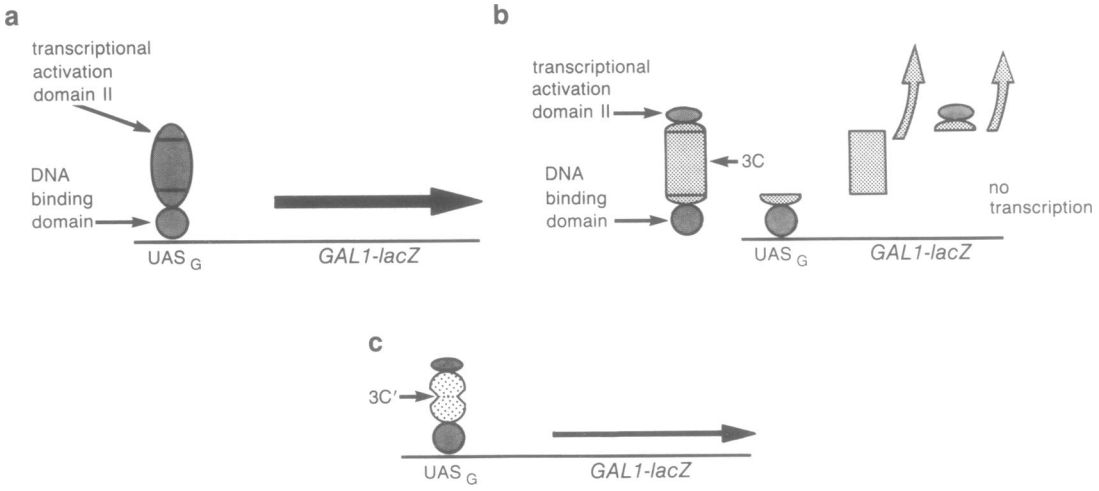

FIG. 1. Model of transcriptional activation by GAL4 derivatives. (*a*) The native GAL4 protein containing the DNA binding domain and both activating regions activates *GAL1–lacZ* transcription. (*b*) A hybrid protein containing the native 3C sequence fused in frame between the coding sequences of the GAL4 DNA binding domain and activating region II fails to activate transcription due to the autocatalytic cleavage of the 3C protease that physically separates the two GAL4 functional domains. (*c*) A hybrid protein bearing the mutant 3C protease prevents the proteolysis from occurring and by maintaining the physical association of two domains can activate transcription.

Table 1. β-Galactosidase activity resulting from activation of transcription by GAL4 derivatives

Plasmid	GAL4 derivative	β-Galactosidase activity
None		—
pMA210	GAL4(1–881)	3877
pMA236	GAL4(1–147)(768–881)	1770
pBD1009	GAL4(1–147)-3C-GAL4(768–881)	14
pBD1013	GAL4(1–147)-3C′-GAL4(768–881)	466

β-Galactosidase activity in MUG units was obtained by substracting the fluorescence values with the MUG reagent mix only.

the wild-type GAL4 (amino acids 1–881) was twice as active as the derivative. Although the GAL4 hybrid protein bearing the native viral protease 3C (expressed from the plasmid pBD1009) did not activate transcription, the GAL4 fusion protein that contained the mutant protease 3C′ (expressed from the plasmid pBD1013) activated transcription effectively (466 MUG units of β-galactosidase). These results are consistent with the idea that the autoprocessing activity of the 3C protease dissociates the activating region II from the DNA binding domain, thus inactivating GAL4 protein.

To test that the inactivation of the GAL4/3C hybrid protein is due to the endogenous proteolysis and not to degradation of the fusion protein *per se*, we have analyzed the GAL4 derivatives by immunoprecipitation using antisera to the GAL4 DNA binding domain. Results are shown in Fig. 3. Lanes A and B represent the proteins immunoprecipitated by the antisera from cells expressing GAL4 derivatives GAL4(1–147)(768–881) and GAL4(1–147), respectively. The multiple GAL4 species immunoprecipitated by the antisera (lane A) result from phosphorylation within the activation domain of GAL4 (26). Lanes C and D represent the proteins expressed in the cells bearing the GAL4 protease constructs. We observe that although the fusion protein containing the mutated 3C protease migrates slower than the 66 kDa protein (Fig. 3, lane D, single arrow), most of the fusion protein bearing the wild-type 3C protease is degraded to a peptide migrating close to the GAL4 (amino acids 1–147) (lane C, double arrow). This result is consistent with the idea that the 3C protease cleaves itself out of the hybrid precursor and catalyzes the cleavage at the Δ3A/3B site to generate a GAL4(1–147 amino acid)-Δ3A(13 amino acid) peptide. We believe that the degradation of this hybrid protein is a specific cleavage event because nuclear extracts of yeast cells expressing the GAL4/3C hybrid protein exhibit CV 3C protease activity when tested *in vitro* (data not shown). We also note that the cleavage is not complete since a small amount of the protein larger than the 66-kDa marker protein is observed (lane C). This is consistent with the observation that cells expressing the GAL4/3C hybrid protein turn slightly blue after an extended incubation (1–2 weeks) on plates containing 5-bromo-4-chloro-3-indolyl β-D-galactoside. It should be noted that the 13-amino acid peptide that remains attached to the GAL4 DNA binding domain (amino acids 1–147) fails to activate transcription; this peptide is neither rich in acidic amino acids nor capable of forming an amphipathic helix, two features that are shared by many eukaryotic transcriptional activating sequences (27).

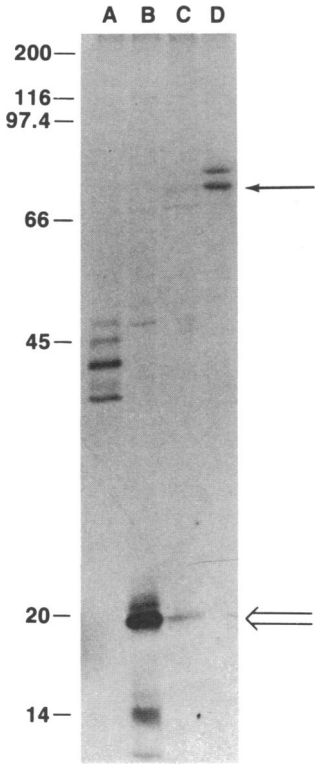

FIG. 3. Immunoprecipitation of GAL4 derivatives. Yeast cells expressing GAL4 derivatives were metabolically labeled with [^{35}S]methionine and lysed by Vortex mixing with glass beads. GAL4 proteins were isolated by immunoprecipitation as described (25) and resolved by electrophoresis on a SDS/10% polyacrylamide gel. The GAL4 derivatives are as follows: lane A, GAL4(1–147)(768–881); lane B, GAL4(1–147); lane C, pBD1009; lane D, pBD1013. The proteolytically processed amino-terminal fragment of GAL4/3C hybrid protein is indicated by the double arrow (lane C). Migration of size markers (in kDa) is indicated. The multiple species of the GAL4 derivatives (lanes A and D) result from phosphorylation within GAL4's activating region 11 (26).

Our current genetic system to assay proteolytic enzymes is not limited to the picornavirus protease. We have performed similar experiments using the HIV-1 protease. We constructed plasmids that express analogous GAL4 fusion proteins bearing the wild-type and the mutant HIV-1 protease in yeast. We found that the GAL4/HIV-1 protease hybrid protein did not activate transcription, whereas the fusion protein containing the mutant HIV-1 protease activated transcription (B. Dasmahapatra, unpublished results).

This system also provides a unique method to generate and analyze various mutants as well as revertants of important proteases. For these purposes it should be noted that the hybrid proteins are stable in yeast cells and are not rapidly degraded by endogenous yeast proteases. The genetic system may be used for studying proteases that are not autocatalytically active; in these cases, the protease cleavage site would be inserted between the two functional domains of GAL4 and the protease would be supplied in trans, as recently reported by Smith and Kohorn (28). Finally, since the GAL4 protein also activates transcription in mammalian cells (29), our approach will provide an opportunity to study the structure–function relationship of these proteases in their natural environment as well as the ability to screen a large number of compounds for specific inhibitors of these therapeutically important proteases.

We are grateful to Bill Sugden for suggesting this experiment. We thank Drs. Claude Nash, George Miller, and Roberta Hare for their support and Dr. Wesley Hung for scientific help.

1. Neurath, H. (1984) *Science* **224,** 350–357.
2. Pontremoli, S. & Melloni, E. (1986) *Annu. Rev. Biochem.* **55,** 455–481.
3. Krausslich, H. G. & Wimmer, E. (1988) *Annu. Rev. Biochem.* **57,** 701–754.
4. Wellink, J. & van Kammen, A. (1988) *Arch. Virol.* **98,** 1–26.
5. Loeb, D. D., Swanstrom, R., Everitt, L., Manchester, M., Stamper, S. E. & Hutchison, C. A. (1989) *Nature (London)* **340,** 397–400.
6. Louis, J. M., Wondrak, E. M., Copeland, T. D., Dale Smith, C. A., Mora, P. T. & Oroszlan, S. (1989) *Biochem. Biophys. Res. Commun.* **159,** 87–94.
7. Matayoshi, E. D., Wang, G. T., Krafft, G. A. & Erickson, J. (1990) *Science* **247,** 954–958.
8. Billich, A., Hammerschmid, F. & Winkler, G. (1990) *Biol. Chem. Hoppe-Seyler's* **371,** 265–272.
9. Richards, A. D., Phylip, L. H., Farmerie, W. G., Scarborough, P. E., Alvarez, A., Dunn, B. M., Hirel, P. H., Konvalinka, J., Strop, P., Pavlickova, L., Kostka, V. & Kay, J. (1990) *J. Biol. Chem.* **265,** 7733–7736.
10. Dasmahapatra, B., Rozhon, E. J., Hart, A. M., Cox, S., Tracy, S. & Schwartz, J. (1991) *Virus Res.* **20,** 237–249.

11. Rueckert, R. R. (1985) *Picornaviruses and Their Replication in Virology*, ed. Field, B. N. (Raven, New York).
12. Werner, G., Rosenwirth, B., Bauer, E., Seifert, J.-M., Werner, F.-J. & Besemer, J. (1986) *J. Virol.* **57**, 1084–1093.
13. Parks, G. D., Duke, G. M. & Palmenberg, A. C. (1986) *J. Virol.* **60**, 376–384.
14. Hanecak, R., Semler, B. L., Ariga, H., Anderson, C. W. & Wimmer, E. (1984) *Cell* **37**, 1063–1073.
15. Keegan, L., Gill, G. & Ptashne, M. (1986) *Science* **231**, 699–704.
16. Ma, J. & Ptashne, M. (1987) *Cell* **48**, 847–853.
17. Gill, G. & Ptashne, M. (1987) *Cell* **51**, 121–126.
18. Sherman, F., Fink, G. & Hicks, J. (1983) *Methods in Yeast Genetics* (Cold Spring Harbor Lab., Cold Spring Harbor, New York).
19. Windheuser, M., Dwyer, S. & Dasmahapatra, B. (1991) *Biochem. Biophys. Res. Commun.* **177**, 243–251.
20. Ito, H., Fukuda, Y., Murata, K. & Kimura, A. J. (1983) *J. Bacteriol.* **153**, 163–168.
21. Miller, J. (1972) *Experiments in Molecular Genetics* (Cold Spring Harbor Lab., Cold Spring Harbor, New York).
22. Mallon, R., Borkowski, J., Albin, R., Pepitoni, S., Schwartz, J. & Kieff, E. (1990) *J. Virol.* **64**, 6282–6285.
23. Geballe, A. P., Spaete, R. R. & Mocarski, E. S. (1986) *Cell* **46**, 865–872.
24. Lindberg, A. M., Stalhandske, P. O. K. & Pettersson, U. (1987) *Virology* **156**, 50–63.
25. Gill, G., Sadowski, I. & Ptashne, M. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 2127–2131.
26. Sadowski, I., Niedbala, D., Wood, K. & Ptashne, M. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 10510–10514.
27. Giniger, E. & Ptashne, M. (1987) *Nature (London)* **330**, 670–672.
28. Smith, T. A. & Kohorn, B. D. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 5159–5162.
29. Kakidani, H. & Ptashne, M. (1988) *Cell* **52**, 161–167.