Pergamon

Tetrahedron, Vol. 53, No. 35, pp. 11953-11960, 1997

0040-4020/97 $17.00 + 0.00

PII: S0040-4020(97)00708-4

SITE-DIRECTED OXIDATIVE PROTEIN CROSSLINKING

David A. Fancy and Thomas Kodadek*

Department of Chemistry and Biochemistry and the Institute of Cellular and Molecular Biology, University of Texas at Austin, Austin TX 78712.

Abstract: Many important biological processes are catalyzed by large multiprotein complexes. A central goal in the study of these "protein machines" is to characterize the organization of proteins and to determine if this organization changes during the catalytic cycle. For this purpose, methodology has been developed to deliver a latent crosslinking reagent to a particular protein in complex solutions. When activated, the crosslinking of the tagged protein to its nearest neighbors occurs without large-scale covalent modifications elsewhere in the complex. Recent technical advances in this methodology are described. © 1997 Elsevier Science Ltd.

Many biological processes are catalyzed by large, multiprotein complexes. For example, protein synthesis is catalyzed by the ribosome, a complex with a molecular mass of approximately 2.7 million Daltons and comprised of 55 proteins and three RNA molecules (in Eschericia coli). In eukaryotic cells it has recently been found that mRNA synthesis is mediated by an RNA polymerase II holoenzyme[1], which has a native molecular mass of somewhere between three and five million Daltons and is comprised of almost 100 proteins. To chemical biologists who wish to probe the mechanism of action of these "protein machines"[2] in detail, the sheer size and complexity of these assemblies can be daunting. For example, a fundamental, problem in the study of protein machines is to understand the architecture of the complex and to determine whether the protein-protein contacts change during the course of the catalytic cycle. A common practice is to use genetic assays such as the two-hybrid system[3] to ask if individual polypeptides in a given complex interact with one another, but this approach, while useful, has the disadvantage of moving the proteins out of their native context. Biochemical assays using recombinant proteins often suffer from the same problem.[4] In general, it would be far better to develop methods to study the organization of intact complexes isolated in their native form.

Protein crosslinking is one method that can be applied for this purpose. However, traditional crosslinking reagents such as glutaraldehyde or bis-succinimide esters can react with any exposed residue bearing the appropriate functional group. This leads to massive chemical modification of the complex and raises the spectre of artifactual results due to perturbation of the structure of the complex. In addition, treatment of very large complexes with these nonselective reagents can lead to many different products, sometimes rendering analysis of the results difficult.

Because of these deficiencies, several investigators have explored linking a probe of some sort (crosslinking reagent, cleavage reagent or a transferrable label) to a particular residue of a single protein. For example, Ebright and coworkers, attached a radiolabeled nitrene precursor to a unique surface-exposed cysteine of the *Eschericia coli* catabolite gene activator protein (CAP) by disulfide bond formation. After forming a CAP-RNA polymerase complex and subjecting it to photolysis, the disulfide bond was reduced, transferring the radioactive tag to the polymerase and identifying the subunit of RNA polymerase contacted by CAP.[5]

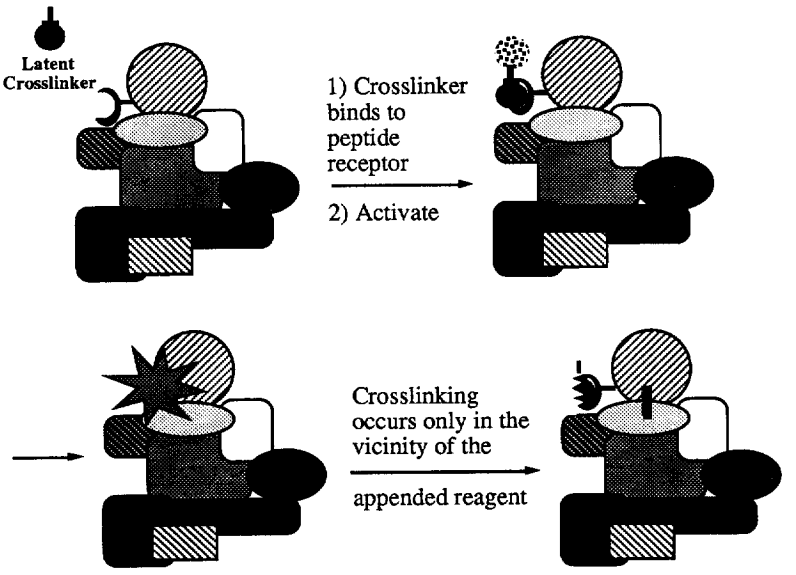

Figure 1. Schematic diagram of an affinity crosslinking strategy in which the activatable crosslinking reagent is delivered to a specific protein in a large complex via noncovalent interactions.

While powerful, the site-specific modification of proteins with probe molecules is of limited utility for the study of protein machines due to the fact that attachment of the probe requires covalent modification of a purified protein. Often it is unfeasible to remove a particular factor from a complex and then reconstitute it with the modified protein. We have therefore begun to explore strategies to deliver an activatable probe molecule to a particular protein in a noncovalent fashion (Fig. 1). The strategy is to fuse to the protein of interest at the genetic level a peptide that binds the probe molecule with high affinity and specificity. The modified protein could be expressed and allowed to assemble into its native complex. The probe molecule could then simply be introduced into the cell or added to a crude lysate at which point it would bind selectively to the peptide "receptor" and hopefully nowhere else in the complex. Activation of the probe molecule would then lead to a highly localized crosslinking, cleavage, or labeling reaction, without wholesale modification of other factors in the complex.

To engineer such a system, we took advantage of the invention of Brown, et al. of an oxidative protein crosslinking reaction mediated by a nickel-peptide complex.[6] In this reaction, a preformed complex of nickel acetate and the tripeptide NH_2-GGH is activated by a peracid, presumably producing some type of Ni(III) species, which crosslinks proteins stably associated with one another in solution. There is some evidence that crosslinking in at least some cases involves oxidative coupling of aromatic residues (K. Brown, personal communication, D. Fancy, unpublished results). From our perspective, an attractive feature of this system is that crosslinking is almost completely dependent on the presence of the peptide ligand. Free nickel ions are not

effective catalysts. This suggested that by fusing a nickel-binding peptide to the protein of interest, a potent crosslinking reagent could be delivered specifically to that fusion product in the context of an intact protein machine. The six histidine (His_6) tag is used widely to purify recombinant proteins by metal chelation chromatography and is known to have a high affinity for Ni^{+2}. Therefore, His_6-tagged proteins were employed to evaluate the feasibility of an affinity crosslinking scheme based on the noncovalent formation of a His_6-Ni^{+2} complex followed by activation with a peracid.

A study of this possibility revealed that the His_6-Ni complex could indeed mediate oxidative crosslinking.[7] An amino terminal His_6-tagged Glutathione S-transferase (His_6-GST) which is a dimer in solution, was used as a test case. The reaction was found to be dependent on nickel, the His_6 tag, and oxidant. It was also found to be highly localized, since any protein without a His_6 tag, or not intimately associated with a His_6-tagged protein, did not crosslink. Finally, it was shown that crosslinked products of His_6-tagged proteins could be highly enriched by chromatography over nickel-saturated NTA agarose beads. This technology is currently being employed to elucidate the organization of several multiprotein complexes.

In this paper, we report two modifications of the original protocol that increase the utility of the system. The first is the introduction of second peptide tag that allows the His_6-tagged protein to be radiolabeled with ^{32}P. The second is the development of a protocol that allows the crosslinking reaction to be carried out while the protein complex is bound to nickel-saturated NTA resin under native conditions. The radioactive label allows the results of a crosslinking reaction to be followed in a straightforward fashion by autoradiography. The bead-bound crosslinking protocol allows the reaction to be carried out in an environment highly enriched for the complex in question, providing higher yields of product.

RESULTS

^{32}P-labeled , His $_6$ -tagged Gal4 activation domain crosslinks to the Gal80 repressor. For the studies reported here, we employed two polypeptides known to associate specifically, the activation domain (AD) of the yeast Gal4 protein[8], and the Gal80 transcriptional repressor[9,10]. Gal4 protein is a transcriptional activator that stimulates the expression of several yeast genes involved in galactose metabolism. Gal80 protein is a repressor of Gal4p that blocks activation at low galactose concentrations. While the detailed mechanism of action of Gal80p is unknown, it is clear that binding of the repressor to the C-terminal activation domain of Gal4p is crucial. A polypeptide corresponding to Gal4p residues 841-874 binds Gal80p about as well as the intact protein and this small fragment is often employed in studies of Gal4p-Gal80p interactions.[11-13]

A construct was made that allowed high-level expression of a fusion protein in *E. coli* comprised of the following fragments: a His_6 tag (immediately after the N-terminal methionine), two pentameric bovine heart kinase recognition sites[14] separated by a glycine, a TEV protease cleavage site[15], and finally, Gal4p 841-874 (see Fig. 2). The fusion protein was expressed from a T7 promoter in BL21DE3 cells, purified to homogeneity and labeled to a high specific activity with ^{32}P-γ-ATP and bovine heart kinase. The labeled AD was incubated with purified Gal80 protein in the presence of nickel acetate and monomagnesium perphthalic acid (MMPP) was

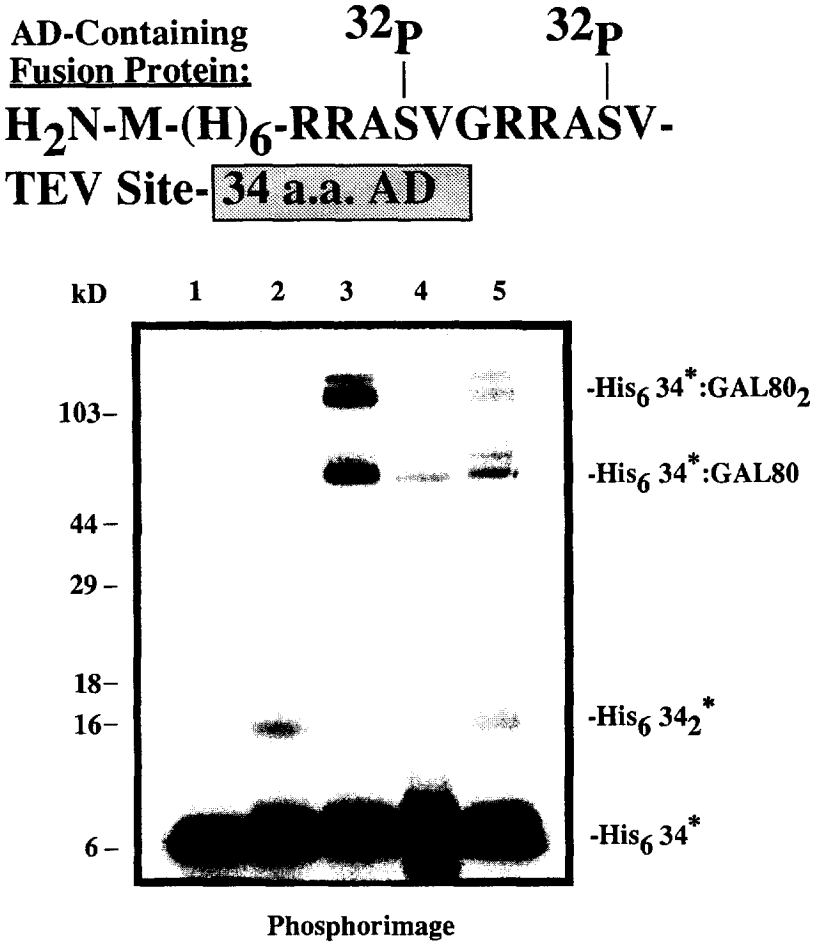

Figure 2. His_6-Ni-mediated oxidative crosslinking of the Gal4p AD and Gal80p. Top: schematic representation of the labeled fusion protein. Bottom: phosphorimage of a denaturing polyacrylamide gel, revealing only the radioactive AD-containing fusion protein and crosslinked products. Lane 1: Purified ^{32}P-labeled AD + purified Gal80p, no oxidant added. Lane 2: ^{32}P-labeled AD treated with nickel acetate and MMPP. Lane 3: ^{32}P-labeled AD + Gal80p, treated with nickel acetate and MMPP. Lane 4: Same as lane 3 except an excess of yeast extract was included. Lane 5: Same as lane 4 except that the AD-Gal80p complex was concentrated on nickel-saturated NTA-agarose beads prior to addition of MMPP.

then added to initiate crosslinking. The results were assessed by denaturing gel electrophoresis and phosphorimagery (Fig. 2). Besides the band corresponding to the AD, two new radioactive bands appeared. One has the apparent molecular mass expected for a Gal80p-Gal4 AD heterodimer and the other the expected apparent mass of a $(Gal80p)_2$-AD trimer. In addition, bands running just above these products were observed, which were attributed to the further crosslinking of another AD molecule. These products are expected from a $(Gal4p\ AD\text{-}Gal80)_2$ complex. Gal80p-Gal80p crosslinking also occurred, as shown by a Coomassie Blue-stained gel (data not shown). Gal80p is known to be a native dimer[12] and crosslinking in this case probably results from reaction of the oxidized His_6-Ni complex with Gal80p residues located close in space. A similar result was observed when related Gal4p AD fragments were crosslinked to the yeast TATA-binding protein, which is also a native dimer.[7]

Control reactions demonstrated that the production of these high molecular mass bands required the Gal80 protein and was not the result of AD multimerization. Lane 1 shows a control in which no oxidant was added to the AD-Gal80p complex. The only band is due to the labeled AD. Lane 2 shows an experiment in which Gal80p was omitted. In this case addition of nickel and MMPP yielded a small amount of AD homodimer, but no lower mobility products.

The Gal4 AD and Gal80p crosslink in a yeast lysate, but less efficiently. Lane 4 of Figure 2 shows an experiment in which a preformed Ni^{+2}-His_6Gal4 AD/Gal80p complex was incubated with an approximately 15-fold excess of crude yeast extract prior to addition of the peracid MMPP. The extract was prepared from yeast cells deleted for *GAL4* and GAL80. A radiolabeled band corresponding to the AD-Gal80p heterodimer was produced, but in lower yield than was the case in the absence of the extract. The utility of the ^{32}P label is readily apparent from this experiment. The product band would not have been discernible in a silver or Coomassie stained gel. Very small amounts of the $AD\text{-}(Gal80p)_2$ product were observed. In general, the crosslinking reaction was less efficient in the presence of the large excess of yeast proteins.

Since the Gal4 AD-Gal80p complex is long-lived and was preformed prior to addition to the extract, the reduced yield of crosslinked products was presumably due to interference by the lysate-derived proteins with crosslinking rather than disruption of the complex. The most likely explanation for the reduced crosslinking efficiency is that the peracid will oxidize exposed cysteine residues of proteins directly. This is the major deviation of the present system from the ideal depicted in Figure 1, in which reactions occur only at the site of the peptide-complexed crosslinker. Thus, the presence of excess proteins is equivalent to the addition of inhibitors of the oxidation process that leads to crosslinking. Although more MMPP was employed than was the case when the purified complex was crosslinked, this inhibition cannot be overcome completely. It should be noted that although a large excess of other proteins were present, no evidence for spurious crosslinking of the labeled Gal4 AD to random proteins in solution was observed.

Crosslinking a Ni-NTA agarose-bound complex restores the efficiency of the reaction. Given the above hypothesis for the reduced crosslinking yield in the presence of yeast extract, a reasonable response was to partially purify the complex prior to crosslinking, but to do so in a way that would not perturb the complex and would be convenient as a general protocol. To do so, the same solution employed in the Lane 4 reaction was

incubated with Ni-saturated NTA agarose beads, which were then pelleted, washed and resuspended. MMPP was then added but nickel acetate was not, the hope being that the nickel ions on the column would support the crosslinking reaction. After a brief incubation, beads were boiled in denaturing loading buffer containing imidazole and the results assessed by gel electrophoresis and phosphorimagery. As can be seen in Fig. 2, lane 5, the column-bound crosslinking protocol worked nicely. All of the crosslinked products observed in the reaction lacking competitor proteins (see lane 3) were observed, indicating an increase in crosslinking efficiency relative to the lane 4 experiment.

Discussion

In this report, two important technical improvements to the His_6-Ni/peracid affinity crosslinking system[7] are described. One is the inclusion of a radiolabel on the His_6-tagged protein[16], which allows the detection of even low-level products. This is extremely useful when evaluating reactions conducted in complex solutions such as cell extracts, where crosslinking yields may be low, possibly due to competing reactions with cysteines in unlabeled proteins. The sensitivity of the assay could be increased even further by purifying and concentrating the His_6-tagged crosslinked products by binding to nickel-saturated NTA agarose beads under denaturing conditions prior to electrophoresis.

The decreased efficiency of the reaction in cell extracts is a deficiency of this first generation affinity crosslinking reaction, since it presumably means that undesired covalent modifications of other proteins occur at some level. This was expected since peracids and thiols are known to react directly with one another. However, there is no evidence that these undesired side reactions lead to spurious crosslinking products. Future work will focus on the noncovalent delivery of other types of triggerable crosslinking reagents, such as photoactivated compounds.

Despite this drawback, the present system can be employed in crude solutions to monitor protein-protein interactions. Furthermore, the efficiency of the reaction can be increased significantly by first binding the His_6-tagged protein and any associated factors to nickel-saturated NTA agarose beads and washing away proteins not in the complex prior to introduction of the peracid. The bead-bound nickel atoms support the oxidative crosslinking reaction. This simple protocol is currently the method of choice for analyzing interactions of His_6-tagged proteins in crude cellular lysates, particularly when the tagged protein is present at low levels.

EXPERIMENTAL

Proteins and reagents. The expression vector for His_6-tagged Gal80p was kindly donated by Prof. Stephen Johnston (University of Texas Southwestern Medical Center). Recombinant His_6Gal80p was purified by a procedure that will be published elsewhere. The His_6-tag was removed from the Gal80 protein with TEV protease according to a published protocol.[15] All proteins were dialyzed from their storage buffers into 15 mM sodium phosphate (pH 7.6) and 150 mM NaCl. Protein concentrations were determined using the Bio-Rad reagent with bovine serum albumin (BSA) as the standard. All other concentrations were based on dry weight. Restriction enzymes were purchased from New England BioLabs. BSA was purchased from United States

Biochemical. Bovine kinase was purchased from Sigma. NTA agarose was purchased from Qiagen. Magnesium monoperoxyphthalic acid (MMPP) and nickel acetate were purchased from Aldrich.

Construction of the pETHKT-34 E.Coli Expression Plasmid. The Parent Vector pETHisK,[16] which contains sequences encoding a His_6 and two bovine kinase sites (HK) was kindly donated by Dr. Tsonwin Hai (Ohio State University. pGEX-34 which contains Gal4p residues 841-874 fused to Glutathione S-transferase was kindly donated by Prof. Stephen Johnston. The DNA fragment containing the sequences which encode the 34 residue AD fragment of Gal4p was isolated by restriction digest with XhoI and SalI, followed by agarose gel purification. This fragment, encoding the TEV site, the AD peptide and a stop codon, was ligated into XhoI/EcoRV-digested pETHisK. The coding region of the final construct was sequenced to confirm its identity.

Expression and Purification of His_6 -HK $_2$ -TEV-Gal4 AD peptide . A freshly transformed colony (BL21DE3 cells) was used to inoculate 5.0 mL of LB + 50 μg/mL Ampicillin. This culture was shaken for 8 hours at 37°C until turbid and then used to inoculate 4 L of identical liquid media. This solution was shaken at 37°C until the optical density at 595 nm was 0.6, and then made 1.0 mM with IPTG to induce protein production. The culture was allowed to shake an additional 4 hours at 37 °C and then the cells were harvested by centrifugation in a JA10 Beckman rotor at 8k rpm. The cells were then resuspended in 4.5 volumes of binding buffer (20% Glycerol, 20 mM Tris Cl (pH 8.0), 1 mM $MgCl_2$, 20 mM imidazole, 200 mM KCl and 0.1 mM phenylmethylsulfonylfluoride), lysed with a French Press at 1260 psi and insoluble material was separated by centrifugation in a Beckman ultracentrifuge Ti45 rotor at 35k rpm. A fraction of the soluble and insoluble material was dissolved in 4x sample loading buffer (0.2 M Tris, 8% SDS, 2.88M β-mercaptoethanol, 40% glycerol, 0.4% bromophenol blue and 0.4% xylene cyanol) and peptide content was examined by tricine gel electrophoresis followed by staining with Coomassie Brilliant Blue. The entire portion of His_6-tagged Gal4 AD peptide was found to be insoluble. The insoluble pellet were resuspended in denaturing binding buffer (6M guanadinium hydrochloride, 20 mM Tris Cl (pH 8.0), 20 mM imidazole), centrifuged at 35k rpm and again analyzed by gel electrophoresis. The soluble fraction, which contained all of the Gal4 AD, was mixed with previously equilibrated NTA agarose in the same denaturing buffer. After mixing for 2.5 hours at 4 °C, the resin was washed sequentially with 2x 5 volumes of denaturing binding buffer, non denaturing binding buffer containing 500 mM KCl, and finally bound protein was eluted with 3x 2 volumes of nondenaturing binding buffer containing 250 mM imidazole. Eluted protein was checked by gel electrophoresis and found to be >98% pure. 11.0 mg of the Gal4 AD peptide was purified from 22 grams of cells.

Radiolabeling of His_6 -34 0.3 nmoles of His_6Gal4 AD peptide was diluted to 50 μL in labeling buffer (20 mM Tris Cl (pH 7.6), 12 mM $MgCl_2$, 100 mM NaCl and 1 mM β-mercaptoethanol), 5 μCi of γATP (NEN). Labeling was initiated with 1.0 μL of bovine kinase (Sigma). After a 30 minute incubation on ice, 20 μL of a 1:1 slurry of NTA agarose (Quiagen) was added and the solution was vortexed 3 times over a 10 minute period. Unreacted nucleotide was washed from the resin with 4 consecutive 300 μL washes with PBS (15 mM potassium phosphate (pH 8.0), 150 mM NaCl). Bound protein was eluted with 50 μL of PBS containing 500 mM imidazole. The eluant was used directly for crosslinking reactions.

Crosslinking reactions Unless otherwise noted, crosslinking reactions were carried out as follows in a total

volume of 30 μL. Final concentrations were 15 mM phosphate (pH 8.0), 150 mM NaCl, and 10 μM Gal80p and Gal4 AD. To form protein complexes, the Gal4 AD peptide was incubated with Gal80p for 30 minutes on ice. Nickel acetate was then added to a final concentration of 200 μM followed by a further five minute incubation. Reactions were initiated by the addition of MMPP to a concentration of 100 μM and were incubated at room temperature for 1 minute. The reaction was then quenched by the addition of 8 μL of 4x loading buffer. Samples were then heated to 100°C for 5 minutes and then separated by gel electrophoresis through a 10% tricine-SDS polyacrylamide gel. Proteins were visualized by fixing the gel in 7% acetic acid then drying on Whatmann 3 mm paper. The products were visualized by phosphorimagery.

Crosslinking in a yeast lysate The same conditions as those described above were used, except that yeast lysate (prepared as described from a *GAL4,GAL80* deletion strain, YJ0) was present in a final concentration of 5 mg/mL, and ten times the normal amount of MMPP and nickel acetate were employed.

Crosslinking on NTA-Agarose resin A protein complex was formed between the His_6-tagged Gal4 AD and Gal80p in 30 μL of PBS on ice as above, except that no nickel acetate was added. To this solution was added 20 μL of a 1:1 slurry of NTA agarose and buffer. The suspension was incubated for 20 minutes on ice with periodic mixing. Unbound protein was removed by washing the resin with 40 volumes of PBS. 30 μL of PBS was then added back to the resin, and crosslinking was initiated by the addition of MMPP.

Acknowledgments: This work was supported by grants from the Council for Tobacco Research and the Welch Foundation. We would like to thank Dr. Sang-Hwa Yang for providing purified Gal80 protein.

References

1. Koleske, A. J.; Young, R. A. *Nature* **1994**, *368*, 466-469.
2. Alberts, B. M. *Cold Spring Harbor Symp. Quant. Biol.* **1984**, *49*, 1-12.
3. Fields, S.; Song, O.-k. *Nature* **1989**, *340*, 245-246.
4. Phizicky, E. M.; Fields, S. *Microb. Rev.* **1995**, *59*, 94-123.
5. Chen, Y.; Ebright, Y. W.; Ebright, R. H. *Science* **1994**, *265*, 90-92.
6. Brown, K. C.; Yang, S.-H.; Kodadek, T. *Biochemistry* **1995**, *34*, 4733-4739.
7. Fancy, D.; Melcher, K.; Johnston, S. A.; Kodadek, T. *Chem. & Biol.* **1996**, *3*, 551-559.
8. Johnston, M. *Microbiol. Rev.* **1987**, *51*, 458-476.
9. Johnston, S. A.; Salmeron, J. M., Jr.; Dincher, S. S. *Cell* **1987**, *50*, 143-146.
10. Ma, J.; Ptashne, M. *Cell* **1987**, *50*, 137-142.
11. Melcher, K.; Johnston, S. *Mol. Cell. Biol.* **1995**, *15*, 2839-2848.
12. Van Hoy, M.; Leuther, K. K.; Kodadek, T.; Johnston, S. A. *Cell* **1993**, *72*, 587-594.
13. Leuther, K. K.; Salmeron, J. M.; Johnston, S. A. *Cell* **1993**, *72*, 575-585.
14. Blanar, M. A.; Rutter, W. J. *Science* **1992**, *256*, 1014-1018.
15. Parks, T. D.; Leuther, K. K.; Howard, E. D.; Johnston, S. A.; Dougherty, W. G. *Anal. Biochem.* **1994**, *216*, 413-417.
16. Chen, B. P. C.; Hai, T. *Gene* **1994**, *139*, 73-75.

(*Received* 23 *December* 1996; *accepted* 1 *March* 1997)