

16 Protease Footprinting

Rod Hori¹ and Noel Baichoo²

¹*Department of Molecular Sciences, College of Medicine, University of Tennessee, Memphis, Tennessee 38163;* ²*Department of Microbiology, Cornell University, Ithaca, New York 14853*

INTRODUCTION, 285

Advantages and limitations, 287

Types of proteolysis, 288

Alternative methods, 289

OUTLINE OF PROCEDURE, 290

End labeling, 290

Assembly of complexes, 293

Proteolysis, 294

Identification of cleavage sites, 296

Quantitative analysis, 297

PROTOCOLS, 299

End-labeling a protein using protein kinase A, 299

Fe-EDTA protease footprinting reactions, 301

Proteolysis using enzymatic proteases, 303

Tricine SDS-polyacrylamide protein gels, 305

Calibration ladder generated by cleavage of proteins by CNBr, 306

Calibration ladder generated by enzymatic cleavage, 308

REFERENCES, 310

INTRODUCTION

In this chapter, we focus on the use of chemical and enzymatic proteases to determine the regions of a protein that become buried or exposed upon assembly into a macromolecular complex *in vitro* (for review, see Hori and Carey 1998; Loizos and Darst 1998; Datwyler and Mearns 2000; Heyduk et al. 2001). The primary method described is protease (or protein) footprinting. Protease footprinting is analogous to the commonly used DNase I footprinting technique except that, rather than using DNase I cleavage of end-labeled DNA to identify where proteins bind, pro-

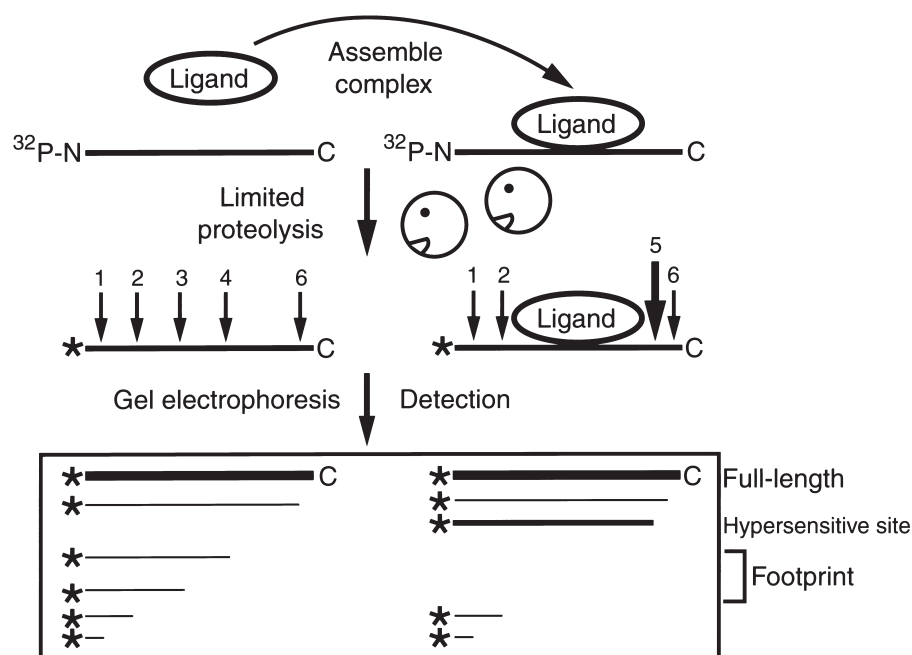


FIGURE 1. Principle of protease footprinting. A protein end-labeled with ^{32}P on its amino terminus (N) is assembled into complexes with its ligand. In parallel reactions, the free and complexed proteins are then immediately subjected to limited proteolysis. Cleavage sites are represented by arrows and are numbered. At this point, the end-labeled amino terminus is represented by an asterisk. The cleavage products are separated on a denaturing protein gel. Upon detection, a protected region (sites 3 and 4 are occluded) and a hypersensitive site (site 5 is cleaved only upon binding of the ligand) are identified.

teases are employed to identify the sites of contact and/or induced conformational changes made by an end-labeled protein upon complex formation (Fig. 1). One major advance in the development of protease footprinting has been the ability to generate an end label easily by engineering the recognition site for protein kinase A (PKA) at either the amino or carboxyl terminus of a protein of interest, allowing isotopic labeling with $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ or $[\gamma\text{-}^{33}\text{P}]\text{ATP}$ (Li et al. 1989; Heyduk et al. 1996). The end-labeled protein, both in its free state and within a macromolecular complex, can then be subjected to partial cleavage, generating a nested set of digestion products that initiate at the end label and terminate at the point of scission (see Fig. 1 in Chapter 11). Both enzymatic and chemical methods have been used to perform the proteolytic reaction and are described elsewhere in this chapter. The cleavage products are then resolved on a denaturing gel adjacent to the appropriate markers and detected using the end label. When the digestion pattern of the complexed protein is compared to that of the free protein, some sites show decreased proteolysis (protections), whereas others show increased proteolysis (hypersensitivity).

The protections are the result of either direct occlusion in the complex or a conformational change that inhibits the ability of the protease to access those regions of the protein. Hypersensitive sites are often caused by a conformational change that allows more efficient cleavage at those residues. The positions of the cleavage sites are assigned by creating “calibration ladders” (described in Identification of Cleavage Sites, p. 296) that serve as mobility standards to interpret amino acid position from migration. A relatively precise assignment of the affected cleavage sites (usually within 5 residues) can usually be made using these ladders. Proteolysis has historically been used for examining protein structure and mapping domains. By introducing an end label and the principles behind DNase footprinting, protease footprinting increases the resolution of the mapping and the amount of information that can be obtained using proteolysis, thus making it an attractive alternative (or supplement) to other methods.

Advantages and Limitations

Although there are many techniques for identifying interactions between a protein and another molecule, only a few can be used to specify the precise regions involved. Mutational analysis, chemical cross-linking, and structure determination by X-ray crystallography or nuclear magnetic resonance (NMR) are the most common methods for examining the structure/function relationship of a protein. Protease footprinting provides a direct biochemical method readily amenable to most laboratories for identifying the domains involved in complex assembly. Protease footprinting is an *in vitro* experiment executed using native proteins in solution and can be used to simultaneously examine multiple regions of a protein. In comparison, *in vivo* genetic analysis is often only inferential and requires the time-consuming construction of numerous mutations to study multiple domains. Structural determination under aqueous conditions of a single-chain protein by NMR has been limited to proteins that are 30 kD or less (Clare and Gronenborn 1998). Structural determination of proteins by X-ray crystallography requires non-physiological conditions to nucleate crystal formation and often requires complementary experiments such as a genetic analysis or protease footprinting for verification of their significance.

The limitations or difficulties of other protein-mapping techniques underscore the advantages of protease footprinting.

1. The only protein engineering required is the addition of the phosphorylation site to generate the end label. The termini of most proteins are exposed, thus an end label can usually be tolerated at those positions without disrupting the structure. If direct methods of end labeling perturb the structure, visualization by an indirect end-labeling procedure such as immunoblots using antibodies specific to one terminus can be used. In contrast, mutagenesis studies require amino acid changes throughout the coding region, which have an increased probability of causing structural alterations.
2. Protease footprinting can potentially identify sequential conformational changes that occur during the assembly of a complex. X-ray crystal structures can also be used to define these structural changes, but this requires solving and comparing the free (unbound) protein and each of the higher-order species. Mutational studies are usually not as straightforward in defining allosteric transitions, especially in the absence of structural data.
3. Protease footprinting experiments require only a small fraction of the amount of protein required for NMR or X-ray crystallography, and the necessary reagents and equipment are accessible to most laboratories.
4. The size of the protein makes little difference in executing a protease footprinting experiment. In contrast, large proteins are typically more difficult to characterize using NMR and X-ray crystallography because more data points are required to obtain the same resolution relative to a smaller protein. For protease footprinting experiments, additional footprinting reactions are typically not required for a larger protein, although multiple gels of varying polyacrylamide concentrations may be necessary to examine the entire length of the protein.

Protease footprinting has several limitations.

1. The affinity of the end-labeled molecule for its ligand(s) must be strong enough to assemble the complex quantitatively. Alternatively, there must be a method for purifying the bound protein from the free protein. One potential issue with purifying the bound complex is the presence of misfolded proteins among the free protein. If the initial solution of free protein contains a significant percentage of misfolded protein, it is possible to obtain a false footprint that reflects differences in the conformation between the active and inactive protein.
2. The protein should not contain any endogenous phosphorylation sites that preclude the introduction of a specific end label. However, the presence of an endogenous PKA consensus

sequence does not necessarily mean it will be accessible to phosphorylation in vitro; it is best to check this before proceeding to remove the site by mutagenesis.

- The protein should be purified homogeneously as full length. Degradation products (containing the phosphorylation site) will also be labeled and lead to background fragments that obscure the signal. Methods for eliminating this problem are described in the End-labeling section, p. 290.

Although there are some limitations, protease footprinting is a relatively simple method of directly mapping regions of proteins involved in interactions with other molecules. Table 1 includes a representative list of studies that have used protease footprinting and classifies them by the method of proteolysis employed and type of complex(es) characterized. A list of the studies that have comparisons to known three-dimensional structures and the types of detection methods used is summarized at the bottom of Table 1. This table helps to illustrate the flexibility this procedure has to examine nearly any type of macromolecular assembly.

Types of Proteolysis

The two most common reagents used for digestion of proteins are small chemicals (primarily metal–chelate complexes that produce hydroxyl radicals) and enzymes (e.g., proteinase K). Both methods have been used extensively (see Table 1). Chemicals such as cyanogen bromide and BNPS (3-bromo-3-methyl-2-(2 nitrophenylmercapto)-3H-indole skatole) cleave proteins at specific residues (methionine and tryptophan, respectively) and have been used for decades to map protein structures and interactions successfully. However, most of these reagents require extreme

TABLE 1. Examples of Complexes Characterized by Protease Footprinting

Type of proteolysis used	
enzymatic	chemical
Binary complexes:	
Protein: protein	
1. σ^{70} : phage T4 AsiA	4. α , β , and β' : other subunits of RNA polymerase
2. CRM1 (Exportin 1): Rev	5. gp55 (gene 55 protein: core RNA polymerase)
3. Transcription factor IIB: VP16 or TAF40	6. ADP ribosylation factor 1: Sec7 domain
Protein: nucleic acid	
7. TFIIIA: DNA or RNA	10. HMGI/C: DNA
8. cAMP receptor protein/cAMP:DNA	11. HMGI/Y: DNA
9. Rev: RNA	12. cAMP receptor protein/cAMP: DNA
Protein: small ligand	
13. DNA topoisomerase II: ATP	16. BirA: biotin
14. Myoglobin: heme	17. cAMP-dependent protein kinase: Mg-ATP
15. Iron regulatory protein-1: Fe-S cluster	18. I-PorI and I-DmoI: Fe ⁺⁺
Ternary complexes and higher	
19. TFIIA: TBP: DNA	22. B''': TATA-box binding protein: B-related factor: DNA
20. gp61 (gene 61 protein): gp41: DNA	23. GreB: RNA polymerase: DNA: RNA
21. σ^{70} : core RNA polymerase: DNA	24. α : β , β' , and σ (sequential assembly)

The first protein listed for each entry is the one that was directly or indirectly end-labeled.

(5, 6, 12, 14, 15, 16, 17, 19, 20, 21, 23) Results were mapped on a known or homologous three-dimensional structure.

(4, 7, 12, 17) Used an antibody against an endogenous terminus.

(13, 15) Used an antibody against an engineered epitope tag.

(8) His₆-CRP was bound to Ni-NTA affinity after digestion to purify the nested set initiating at the His₆ tag.

(21) Used biotinylated nitroloacetic acid to detect a hexahistidine tag as an end label.

(All others) Used either ³²P or ³³P.

conditions that make it difficult to use them with native proteins and generate sparse data because they cleave infrequently due to their specificity. In contrast, metal–chelate complexes, such as iron–EDTA and phenanthroline–copper (Baichoo and Heyduk 1999), generate a more comprehensive pattern of cleavage. Initially, these reagents were instrumental in producing DNA footprints with a uniform laddering pattern. Similar benefits are provided to those performing protease footprinting. Under mild conditions, iron–EDTA generates hydroxyl radicals that cleave proteins and nucleic acids nonspecifically. The fact that these radicals are comparable in size to a water molecule allows the entire solvent-exposed surface of a protein to be available for cleavage, resulting in more detailed and precise delineation of regions experiencing contact with other molecules, or conformational changes.

There are two major disadvantages to using Fe–EDTA chelates to digest proteins. First, their cleavage sites are not defined by sequence, so assigning cleavage positions requires detailed analysis of their relationship to the calibration markers. Second, extensive data analysis (described in *Quantitative Analysis*, p. 297) and averaging of several experiments is necessary to identify regions within proteins where only subtle changes in sensitivity to cleavage occur. A phosphor screen reader or exposure to film followed by densitometry or a fluorescence image reader, depending on the type of end label used, is necessary to collect the data required for this type of analysis.

Enzymatic proteases, on the other hand, are more familiar to most researchers and have several advantages. First, they are simpler to manipulate. Protease can easily be titrated over several orders of magnitude. Second, a relatively small number of digestion products are generated, usually making data simpler to interpret.

The main disadvantages of using enzymatic proteases are (1) the sequence specificity of the cleavage reaction and (2) their large size, which further limits their accessibility to potential cleavage sites. Thus, most proteins will have regions that contain only a few cleavage sites or none at all. Enzymatic proteases will be unable to characterize these regions fully.

Alternative Methods

Another way of employing chemical proteolysis to map macromolecular interactions is the use of tethered metal–chelate complexes. In this method, the tethered metal–chelate is covalently attached to either a cysteine or lysine residue on an unlabeled protein, which is allowed to form a macromolecular complex containing the end-labeled protein(s) of interest. Subsequently, cleavage is initiated by adding hydrogen peroxide and ascorbate to generate hydroxyl radicals (for review, see Datwyler and Meares 2000). If the cysteine or lysine residue is unique, this allows the specific cleavage and mapping of regions within other proteins that are close to that particular residue. Alternatively, the metal–chelate complex can be tethered to multiple lysine or cysteine residues to produce a more general characterization of interactions within the complex (Traviglia et al. 1999). When the cleavage products are resolved on a denaturing protein gel, residues within the end-labeled protein, juxtaposed within the complex near the site conjugated to the metal–chelate complex, are defined by identifying the positions of the cleavage sites (as described in *Quantitative Analysis*, p. 297).

One advantage of using a tethered protease is that it identifies only residues that are within its vicinity. Diffusible proteases do not allow the mapping of proximity to a specific residue. Another advantage is that results are easier to interpret than when diffusible metal–chelates are used, because relatively few fragments are produced. Additionally, this technique can be more sensitive because interactions generate the appearance of cleavage products against a null background rather than the decrease of a signal, and this is better for characterizing low-affinity associations. The disadvantages are the limitation of using only lysine or cysteine residues to attach the metal–chelate group, the identification of residues/regions within the distance of the linker arm used as a tether rather than the direct contact site, and the inability to define induced conformational changes.

Another protein footprinting method that can yield high-resolution data is based on chemical modification rather than direct cutting (Hanai and Wang 1994). In addition, this technique can be especially useful for examining proteins that denature upon even slight cleavage, but remain stable in response to the chemicals being used. In this method, free (unbound) protein and protein assembled into complexes are treated in parallel with a limiting amount of citraconic anhydride that reversibly modifies exposed lysine residues. Subsequently, the protein is denatured and all of the lysines that were not previously citraconylated are acetylated. Next, the citraconyl groups are removed by incubating the modified protein in acetic acid. The final product is recovered and the deblocked lysine residues are then digested to completion. Exposed lysine residues that become buried upon assembly of the complex exhibit a decreased rate of digestion (i.e., protection). The procedure has the advantages that a small chemical can be used that can access nearly all exposed regions, and the assembled complex is not subjected to proteolysis that could disrupt its integrity. The disadvantage is that only regions containing lysine residues can be characterized. An analogous method to examine domains containing cysteine residues has also been developed (Tu and Wang 1999) and can be used to gain additional information.

Protease footprinting is a method within the capabilities of the average molecular biology laboratory. A flowchart in Figure 2 outlines the basic steps that are common to the different versions of this technique. This method provides a powerful way to gain structural information about specific macromolecular interactions. With many genomes fully sequenced, the question remains—What is the function or role of these newly sequenced proteins? Direct structural studies of the macromolecular complexes that perform key enzymatic and regulatory functions in the cell are a critical adjunct to the genetic, computational, and other approaches of modern biology.

OUTLINE OF PROCEDURE

End Labeling

A specific end label is crucial to the success of protease footprinting, and protein kinase A provides a simple and reproducible method for achieving this goal (Pestka et al. 1999). Proteins containing a PKA recognition site (X-Arg-Arg-X-Ser-Y, where X and Y can be most amino acids, although arginine and hydrophobic residues, respectively, are optimal) engineered onto either the amino or carboxyl termini, respectively, have been used in protease footprinting experiments (Table 1). For greater detail regarding the recognition site, see Songyang et al. (1994). The most complete set of data is obtained when both amino- and carboxy-terminally labeled proteins are used for analysis. In the presence of either $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ or $[\gamma\text{-}^{33}\text{P}]\text{ATP}$ and the catalytic subunit of PKA, the serine residue becomes phosphorylated and creates an end label. Both ^{32}P and ^{33}P have been used for end-labeling proteins in protease footprinting experiments.

Several properties need to be examined prior to using this methodology for end-labeling a protein. First, the introduced PKA recognition site should be unique, or the other site(s) should be buried and inaccessible. This is important because internal or multiple sites will obscure the nested set of fragments containing the phosphorylation at the amino or carboxyl terminus. The sequence of the wild-type protein should be examined for the presence of endogenous PKA recognition sequences. If any are present, the phosphorylation efficiency of that wild-type protein can be compared in a trial reaction against a protein known to be phosphorylated by PKA. If the protein of interest is phosphorylated efficiently, these sites can be eliminated by site-directed mutagenesis, assuming the mutations do not affect the structure or function of the protein. (Alternatively, the end label can be generated using casein kinase II and a terminus engineered to contain its recognition site rather than that of PKA [Schwanbeck et al. 2000].) If the protein lacks endogenous PKA recognition sites, one is engineered onto a terminus of the protein. In addition, the presence of cryptic phosphorylation sites should be checked by incubating equal amounts of

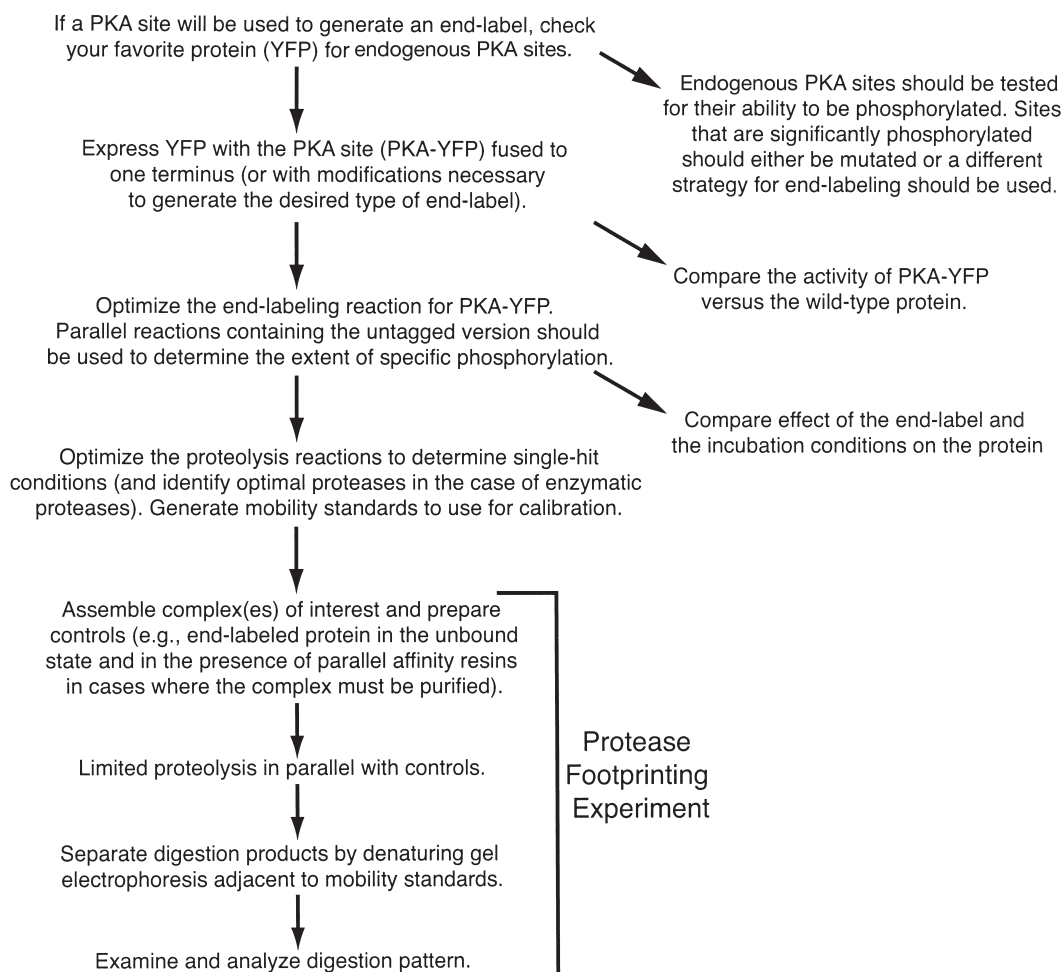


FIGURE 2. Flowchart of steps involved in protease footprinting.

the protein containing and lacking the engineered PKA-tag in a phosphorylation reaction and examining small aliquots of these reactions on an SDS-polyacrylamide protein gel. The percentage incorporation and the ratio of specific phosphorylation to fortuitous background phosphorylation should be determined using a phosphor screen reader or autoradiography followed by densitometry. For use in protease footprinting, this ratio should be at least 10:1. A time course of the reaction and a titration of enzyme should be performed to optimize the extent of labeling and the specific labeling.

The amount of protein in the end-labeling reaction depends on the efficiency of the phosphorylation reaction and complex assembly. For the highest specific-activity end-labeled protein, the optimal PKA site (Songyang et al. 1994) and a reaction containing 10–30 pmoles of protein and a three- to tenfold excess of $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ or $[\gamma\text{-}^{33}\text{P}]\text{ATP}$ should be used. The use of optimal labeling conditions is especially important when (1) only low incorporation rates can be obtained or (2) only a fraction of the free protein can be driven into complexes, due to low-affinity interactions, and purifying the bound protein results in significant loss of material (see Assembly of Complexes, p. 293). Alternatively, lower specific activity, achieved by increasing the protein or lowering the ATP concentration in the reaction, can be tolerated if the signal is sufficient to observe the digestion pattern and may provide an advantage if larger amounts of protein facilitate the reactions used to assemble complexes. Phosphorylation reactions involving human

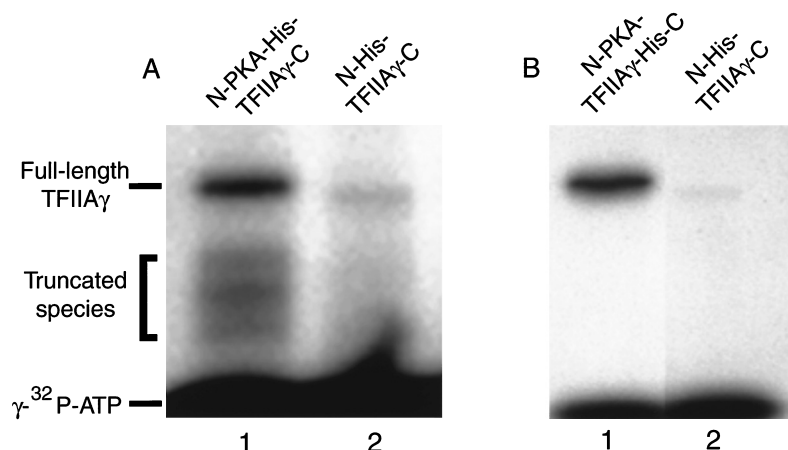


FIGURE 3. End-labeling of the γ subunit within human TFIIA. (A) End-labeling of TFIIA γ with both the PKA site and His₆ tag on the amino terminus (lane 1) versus one with only a His₆ tag on the amino terminus (lane 2). (B) End-labeling of TFIIA γ with the PKA site and the His₆ tag on opposite termini (lane 1) versus one with only a His₆ tag on the amino terminus (lane 2).

TFIIA, a general transcription factor, are illustrated in Figure 3. TFIIA can be reconstituted using two subunits named $\alpha\beta$ and γ . Reactions using molecules of TFIIA, in which the γ subunit either contains (lane 1) or lacks (lane 2) the PKA recognition site are shown in Figure 3B. The ratio of specific to nonspecific labeling is approximately 20:1.

Second, it is important to compare the specific activity of the phosphorylated protein to that of the wild-type version to ensure that the phosphorylation or the incubation conditions have not led to a significant loss in activity. As discussed above, the presence of inactive/misfolded protein can lead to a “false” footprint. A test phosphorylation reaction should be performed with and without [γ -³²P]ATP with a known amount of protein. If phosphorylation results in loss of activity, placing the PKA recognition site on the opposite terminus may be enough to eliminate this effect. If inactivity results from the incubation alone, the phosphorylation reaction can be performed at 4°C overnight, or alternative reaction conditions can be screened.

Third, the integrity of the protein to be labeled is also an important issue. Any degradation products that contain the PKA recognition site will be labeled and generate background bands upon detection. There are two methods for eliminating potential problems due to degradation. In the first, a sequence, such as a hexahistidine (His₆) tag, that can be used to purify the protein can be engineered on the terminus opposite the PKA site. Because the purification tag and the PKA site are on opposite termini, only full-length protein can be both purified and labeled. TFIIA, reconstituted using a γ subunit containing the His₆ tag and the PKA recognition site on the same terminus, yields background phosphorylated products due to degradation (see Fig. 3A, lane 1). The use of a His₆ tag on the opposite terminus to purify the same protein eliminated the background labeling (cf. Fig. 3A, lane 1, and 3B, lane 1). Alternatively, data analysis can be used to correct for the presence of degradation products. For a given preparation, there is a definable relationship between the full-length, end-labeled protein and the truncated species. These values can be used to normalize the pattern detected in each lane. A potential issue with this analysis is the increased possibility of misfolded forms among the truncated versions. In some cases, this may expose (or occlude) regions of the protein and generate (or diminish) those digestion fragments relative to the properly folded species.

Upon completing the phosphorylation reaction, the labeled protein can be purified from the unincorporated [γ -³²P]ATP using one of several methods, including a spin column or dialysis. Alternatively, an excess of nonradioactive ATP can be added to the reaction to prevent significant labeling of other proteins in subsequent experiments. This mixture can be used without further

purification as long as the components do not inhibit any subsequent steps. Because ATP is small, it migrates at a position that obscures fragments of less than 10 residues. For example, in Figure 5, the solution of end-labeled TFIIB also contains unincorporated [γ - ^{32}P]ATP, but it is not observed although fragments of ~5 residues generated by clostripain or trypsin (lanes 2–11) can be distinguished.

If phosphorylation is not an effective method for end-labeling a protein, there are other means available to achieve this goal. A direct end label can also be generated by specifically attaching a fluorescent tag at one terminus using either a lysine or cysteine residue. In one example, the Cy5 fluorescent dye was modified to allow it to be attached to a unique cysteine residue (Callaci et al. 1998). Alternatively, an indirect end label can be generated using immunoblots. The protein can be engineered to contain an epitope, such as the FLAG, myc, or His₆ tag, to which there are commercially available high-affinity antibodies (Lindsley and Wang 1993).

Some of these antibodies do not work when the epitope is at one of the termini, thus their ability to work should be carefully considered. As another option, instead of an antibody, biotinylated nitriloacetic acid can be used to probe a membrane containing a His₆-tagged protein (McMahan and Burgess 1999). If the protein does not tolerate the addition or alteration of residues on the termini, antibodies that recognize one of the endogenous termini can be used (Heyduk and Heyduk 1994; Cheng et al. 1998). In these cases, multiple proteins in a complex can be examined within the same experiment by using multiple fluorescent tags or antibodies (Greiner et al. 1996; Owens et al. 1998) to detect different proteins. Experiments to test the effect of any engineered residues on the activity of the protein and to measure the background generated from these methods must be performed.

Once a method for end-labeling the protein of interest has been established, the electrophoresis standards that will be used to determine which cleavage site generates each proteolytic fragment should be generated. These markers are described in more detail in Identification of Cleavage Sites (p. 296) and include those made by site-specific proteolysis of the end-labeled protein and engineered deletion mutants that contain the same end label, but are truncated from the opposite terminus.

Assembly of Complexes

In protease footprinting experiments, the end-labeled protein must be assembled quantitatively into the desired complexes because background signal that obscures the protections will be generated when the uncomplexed protein is digested at those sites. If the affinity is high, this can be accomplished by simply adding an excess of the ligand. A sufficient amount should be used so that the resulting digestion pattern can be observed, and that value depends on the specific activity of the end-labeled protein. A typical concentration to use in each digest is an amount of protein containing 10,000–30,000 Cerenkov counts of incorporated label. Higher concentrations of protein will be necessary to observe results when it has been less efficiently labeled. In other cases, the affinities are not sufficient and other measures must be employed to separate the complexed protein from the uncomplexed protein. Several of these methods are covered elsewhere in this manual and are only discussed briefly here. In the case of protein–protein complexes, the end-labeled protein can be bound on an affinity resin that displays the ligand protein. A simple procedure for generating such a resin is to express the ligand protein fused to the protein glutathione-S-transferase (GST) and to purify this fusion protein using glutathione–agarose. The fusion protein should be tested for activity as described in Chapter 4 and, if active, used subsequently to assemble complexes. Parallel affinity resins displaying the same fusion protein containing mutations that are known to abolish complex formation specifically, and/or GST alone, should be used as controls to establish conditions for specific complex assembly.

Protein–DNA and protein–RNA complexes can be purified by using biotinylated nucleic acids attached to a matrix containing streptavidin. Affinity resins displaying oligonucleotides contain-

ing crucial mutations, or the removal of a required protein from the assembly reaction, should be used as controls to establish conditions for specific complex formation and the subsequent proteolysis reactions. When the affinity is sufficiently low to require one of these (or some other) methods to purify the complex of interest, a significant fraction of the material is not recovered. This must be taken into account when designing the assembly. Typically, retaining >10,000 Cerenkov counts of end-labeled protein per digest (i.e., lane) is sufficient. The actual amount of end-labeled protein that must be incubated with the affinity matrix will depend on its specific activity and the efficiency of complex formation.

Proteolysis

As in the case of DNase footprinting, it is essential that each protein molecule be cut a maximum of once along its length. To achieve these “single-cut conditions,” the proteolysis reactions should be manipulated such that >50% of the protein is uncut (Brenowitz et al. 1986). This can be determined by comparing the intensity of the uncut band in a lane to the total intensity within the lane. The time of proteolysis necessary for single-cut conditions is determined empirically and should take into account the stability of the complexes being characterized. Stable complexes such as that of an antibody with myoglobin (Zhong et al. 1995) can survive longer incubations with lower concentrations of proteases. However, it is often preferable to minimize the time of digestion because the initial cleavage may lead to some protein unfolding. Increasing the length of the digestion reaction provides time for cleavage at secondary sites that have become exposed. When the complexes are less stable, shorter incubations with higher concentrations of proteases should be used. In either case, optimization using a range of concentrations and times is necessary to balance single-hit conditions with complex stability.

Parallel digestion reactions of the complex containing the end-labeled protein and, as a control, the free (unbound) end-labeled protein should always be performed. In addition, other controls may be necessary, depending on the method used to assemble (and purify) the complex being studied. If the complex has been assembled on an affinity matrix, the end-labeled protein should be digested in the presence of control affinity resin(s). For example, when a protein–protein complex is assembled and purified on a GST-based affinity matrix, the end-labeled protein should be examined after incubation with both GST–ligand and GST alone.

The GST resin will retain only a small fraction of end-labeled protein under conditions required for specific complex formation. In this case, it may be necessary to subject this assembly reaction to one less cycle of “washing” to retain sufficient end-labeled protein for detection. In another example, when ternary complexes are being examined, to determine the contribution of each component, the controls should include both the free (unbound) protein and binary species. In general, the controls required for protease footprinting will be similar to those required to establish the specificity of complex assembly.

Hydroxyl radical cleaving experiments require additional considerations. The time between preparation of reagents and their addition to proteins is crucially important. An iron–EDTA complex is generated by mixing solutions of ferrous ammonium sulfate and EDTA (Dixon et al. 1991; Heyduk and Heyduk 1994), or it may be prepared from preformed iron–EDTA (Greiner et al. 1996; Kumar et al. 1997). These solutions should be made just prior to use. A typical titration to determine the optimal concentrations uses 1 mM EDTA and 0.5 mM Fe as the minimal concentrations, with constant concentrations of ascorbate and hydrogen peroxide. Figure 4 illustrates an example of an experiment examining cAMP receptor protein (CRP)-cAMP bound to DNA with proteolysis by Fe–EDTA. The regions of CRP that become buried upon binding to DNA will produce a decrease in the signal intensity of fragments derived from that domain, whereas the ones that become more exposed will generate an increase in signal intensity. The sites that are significantly affected when CRP binds DNA are identified using a difference plot to determine the rela-

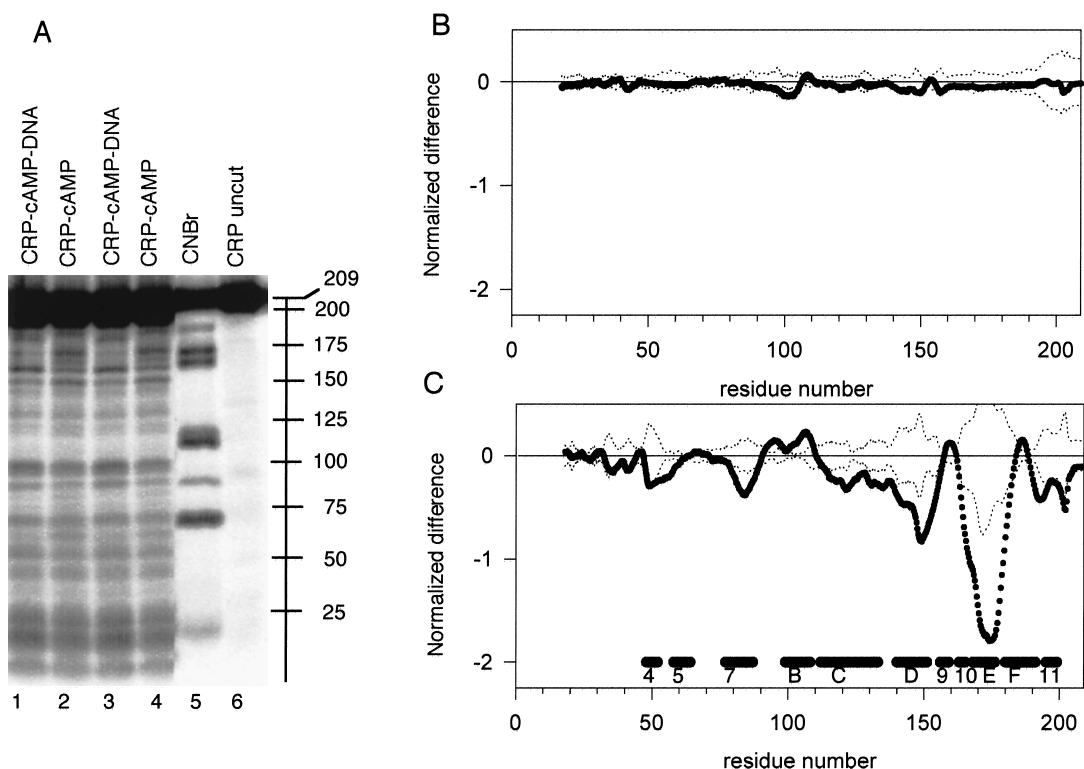


FIGURE 4. Fe-EDTA cleavage of CRP-cAMP and its complex with DNA. (A) Autoradiogram of a protease footprinting experiment for CRP-cAMP and its complex bound to DNA containing the consensus site. CRP indicates end-labeled CRP. Lanes 1 and 3 contain two parallel trials of CRP bound to cAMP and DNA containing the consensus binding site treated with Fe-EDTA. Lanes 2 and 4 contain two parallel trials of CRP bound to cAMP treated with Fe-EDTA. Lane 5 contains CRP digested with CNBr to produce a calibration ladder. Lane 6 contains CRP untreated by Fe-EDTA. (B) Averaged difference plot of CRP-cAMP versus CRP-cAMP complexes to determine the range of experimental variation. The dotted line indicates 2x standard error and the heavy line is the average normalized difference. The difference is calculated as (signal intensity of a given fragment within the bound protein – signal intensity of the same fragment within the free protein)/(signal intensity of that fragment within the bound protein). (C) Averaged difference plot for the CRP-cAMP-DNA ternary complex versus CRP-cAMP binary complex to determine the effect of binding to DNA. The dotted line indicates 2x standard error and the heavy line is the average normalized difference. Letters and numbers below the graph refer to the α helices and β strands, respectively.

tive change at any given site and are calculated as described in the figure legend. Negative values correspond to more buried (protected) regions, and positive values correspond to more exposed (hypersensitive) regions. Changes larger than two times the standard error (the thin lines) calculated from multiple trials (lanes 1 and 3) are considered to be significant. Note that the protection in residues 161–183 is evident from visual inspection of the gel, whereas the protections in residues 48–58 and in residues 79–88 require analysis to be discerned.

Because iron-chelate complexes must generate hydroxyl radicals to work in this procedure (Croft et al. 1992), it is necessary to perform reactions in buffers free from radical scavengers such as glycerol and thiols. If a protein is stored in buffers containing scavengers, dialysis or desalting columns can be used to exchange the buffer.

Many different enzymatic proteases have been used in protease footprinting experiments (see Table 1). They will cleave at a more limited set of positions than the chemical proteases, and the ones that generate the best signal for each situation must be determined empirically. When deter-

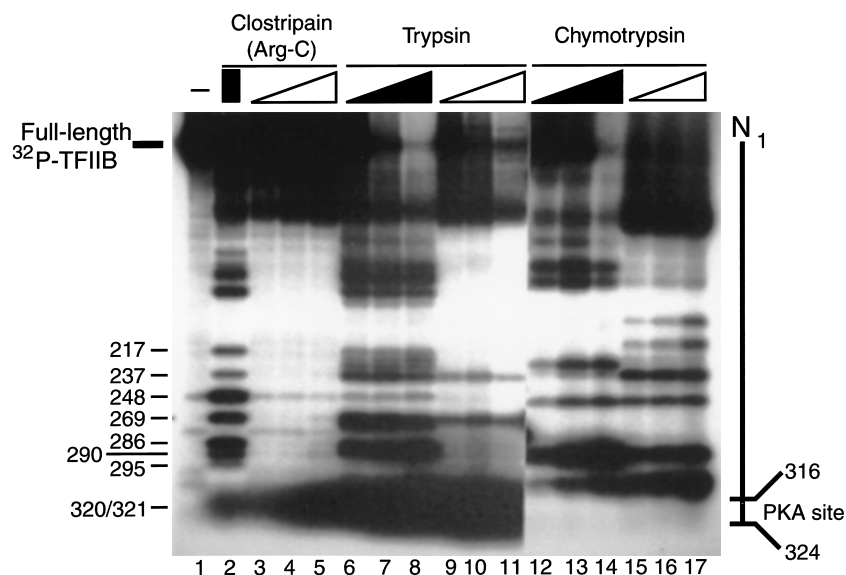


FIGURE 5. Test proteolysis reactions of TFIIB. Human TFIIB end-labeled on a carboxy-terminal PKA recognition site was digested for 2 minutes with increasing amounts (denoted by *triangles*) of clostripain (Endoproteinase Arg-C; lanes 2–5), trypsin (lanes 6–11), and chymotrypsin (lanes 12–17). Undigested end-labeled TFIIB is in lane 1. The open and closed triangles represent digestions using native and denatured TFIIB, respectively. The positions of the amino terminus and the PKA site on the 324-residue protein are shown on the right-hand side. The cleavage products corresponding to the nine arginine residues in the carboxy-terminal third of the protein identified when denatured TFIIB is digested with clostripain (lane 2) are identified on the left side.

mining the digestion conditions to use, the amount of protease used and the reaction time can be varied. The optimal protease concentrations can be defined by making serial dilutions and examining the resulting proteolysis products over a wide range of concentrations. A typical range of protease is 20–2000 ng. Figure 5 illustrates a set of reactions for establishing the conditions for digesting human TFIIB.

The optimal amount of protease is one in which the full-length protein is 30–50% digested. Under these conditions, the signal of the proteolytic fragments will be strongest while staying within the range that results in single-hit digestion. Different proteases digest different portions of the protein, and the ones that cleave in the regions of interest should be used in the experiments. Usually, a larger amount of protease is required to digest the assembled complex relative to the free protein. Even after an optimal concentration has been identified within one trial, it is often useful to use three concentrations centered around that amount in subsequent experiments to obtain the best results. To obtain the most consistent results, the proteases should be stored frozen in single-use aliquots at a convenient stock concentration, typically 1 mg/ml. Multiple use of aliquots results in increased variability due to either autoproteolysis and/or denaturation during the freeze/thaw cycle (R. Hori, unpubl.).

Identification of Cleavage Sites

The better the resolution of the system used for separating the digestion products, the more precise the assignment of cleavage sites and therefore the assignment of residues involved in protections and hypersensitivities. The use of tricine instead of glycine as a carrier ion allows higher resolution of smaller fragments (Schagger and von Jagow 1987) and, although it can be used to analyze frag-

ments as large as 100 kD, its biggest advantage is the increased resolution of fragments from 20 kD down to 1 kD. In addition, varying the percentage of acrylamide in the gel allows different ranges of cleavage products to be targeted and analyzed. Tricine- and glycine-based systems have been used in conjunction to analyze fragments from large proteins (Casaz and Buck 1999). Glycine gels are effective for resolving fragments from 15 to 100 kD. Larger resolving gels also increase resolution. Other potential candidates for separating fragments are capillary electrophoresis and mass spectrometry (Cohen et al. 1995; Kriwacki et al. 1996). The high sensitivity of mass spectrometry allows it to resolve fragments very close in size. However, results from cleavage by metal–chelate complexes can be difficult to analyze using this method. Hydroxyl radicals can attack amino acid side chains as well as peptide bonds; cleavage generates a number of different-sized fragments each derived from the same portion of the protein. Recently, *matrix assisted laser desorption ionization* (MALDI) was used to map modifications of side chains by hydroxyl radicals, a promising sign for further use of this method (Maleknia et al. 1999).

The method used for detecting the cleavage fragments depends on the type of end label and separation system used. In the case of a denaturing polyacrylamide gel, a phosphor screen reader or autoradiography can be used for a radioactive end label, and a fluorescence image reader for a fluorescent one. An immunoblot will be used in the case of using an antibody directed against the endogenous or engineered terminus. Regardless of how the cleavage products are detected, they need to be mapped to identify the sites of interaction and/or conformational changes. This is accomplished by comparing the mobility of the fragments generated proteolytically to that of known standards, including commercial prestained molecular-weight markers, deletion mutants, and site-specific cleavage products of the end-labeled protein. Because the molecular weights of discrete peptide fragments are not likely to retain a precise logarithmic relationship to their mobility, the prestained markers provide only a first approximation of the cleavage positions. The precise identification of sites (within about 5 residues) requires the use of the markers described next. To generate mobility standards by site-specific cleavage, the end-labeled protein is unfolded and then digested by either chemical or enzymatic proteases that target specific amino acid residues such as cyanogen bromide (which targets methionine residues) and endoproteinase Lys-C (which targets lysine residues), respectively. Cleavage sites are assigned by comparing fragments generated to the number and sizes of expected fragments predicted from the protein's primary sequence. In practice, however, some sites are cut more than others due to their environment or residual protein structure. To improve the confidence in assignment of fragments, one set of markers produced by a particular method can be used to assign another set produced by a different method. A set of deletion mutants can be engineered using standard molecular cloning techniques and must contain the same end label as the full-length molecule. These are particularly useful for delineating crucial regions that are affected within a complex but are not mapped by the site-specific cleavage markers. Using a combination of chemical and enzymatically produced markers along with deletion mutants, accurate assignment of cleavage sites is possible.

Quantitative Analysis

When using metal–chelate complexes, the large number of cleavage sites makes it difficult to see subtle, but significant, changes in cleavage susceptibility or to be certain that differences in the intensity of bands correspond to an effect due to complex formation. Therefore, it is necessary to use statistical or graphical methods to identify regions of protein involved in interactions. The first step for quantitative analysis is to obtain a digital image using either a densitometer, phosphor screen reader, or fluorescence image reader. Each lane is quantified by full-lane-width scanning along its length, producing a graph of intensity as a function of migration. This can be done using software associated with the image reader or by exporting the data points to a graphing or spread sheet program. Analogous bands in different lanes are then aligned to correct for minor

variations in migration. This can be done by manually adjusting the rows in a spread sheet or drawing program or using specialized software such as ALIGN (written in BASIC; available from T. Heyduk). Aligned scans are then corrected for differences in gel loading and cleavage efficiency, converted from mobility to amino acid cleavage site, and compared to each other as difference plots. In some cases, averaging of multiple experiments using statistical analysis will be necessary to observe differences in the cleavage patterns and can be accomplished using standard spread sheet software. Alternatively, specialized software such as DIFPLOT (written in SigmaPlot transform language; available from T. Heyduk) to consolidate normalization and averaging can be used (Baichoo and Heyduk 1999; Loizos and Darst 1999). These plots graph normalized differences between free and bound protein as a function of amino acid residue number. An example of this analysis can be observed in Figure 4 (described above). To eliminate artifacts and emphasize real differences, averaging of difference plots from multiple experiments is done and statistical methods such as standard error (the thin line in Figure 4) and Student's t-test can be applied. After this extensive analysis, regions of protein where changes occur can be assigned confidently.

Protocol 1

End-labeling a Protein Using Protein Kinase A

Protein kinase A provides a simple and reproducible method for end-labeling.

MATERIALS

CAUTION: See Appendix for appropriate handling of materials marked with <!.>.

Buffers and Solutions

5x Phosphorylation buffer

100 mM Tris (pH 7.6)

500 mM NaCl

60 mM magnesium chloride (MgCl₂) <!.>

If the protein solution contributes significant NaCl or MgCl₂, adjust the 5x phosphorylation buffer such that their concentrations in the reaction are 100 mM and 12 mM, respectively.

2x SDS gel-loading buffer

100 mM Tris-HCl (pH 6.8)

200 mM dithiothreitol (DTT) <!.>

4% SDS (electrophoresis grade) <!.>

0.2% bromophenol blue <!.>

20% glycerol

Your Favorite Protein (YFP) containing a terminal PKA recognition site

For high-specific-activity end-labeling, 10–30 pmoles is typically used per labeling reaction. For proteins between 10 and 100 kD, this is 0.1–3 μg.

Nontagged version of YFP (i.e., typically wild-type)

Enzymes and Buffers

Catalytic subunit (of PKA) from Bovine Heart (Sigma P 2465)

Resuspend the entire contents of the bottle in 6 mg/ml DTT, as specified by the manufacturer, at a concentration of 5 units/μl. Divide into single-use aliquots, and then freeze on dry ice and store at –70°C. No significant loss of enzyme activity has been observed with a single freeze/thaw cycle (R. Hori, unpubl.).

Gels

Gel filtration columns

SDS-polyacrylamide protein gel (4% stacking; 8–15% separating, depending on the size of the protein being labeled)

For further details on SDS-polyacrylamide gels, see Sambrook and Russell 2001, *Molecular Cloning*.

Additional Reagents

Bovine serum albumin
 Milli-Q (Millipore)
 Trichloroacetic acid <!\>

Labeled Compounds

[γ -³²P]ATP (>5000 Ci/mmmole) or [γ -³³P]ATP (1000–3000 Ci/mmmole) <!\>

METHOD

1. Combine the following:

5x phosphorylation buffer	20 μ l
catalytic subunit of PKA (30 units)	6 μ l
[γ - ³² P]ATP (>5000 Ci/mmmole; ~100 pmole)	500 μ Ci
<i>or</i>	
[γ - ³³ P]ATP (1000–3000 Ci/mmmole)	100–300 μ Ci
wild-type and PKA-tagged YFP, respectively	10–30 pmole

For high specific activity, use a three- to tenfold molar excess of labeled ATP. The optimal amount can be determined empirically.

Add Milli-Q H₂O to 100 μ l.

2. Incubate for 45 minutes at 25°C.
3. Add 1 μ l of 100 mM ATP to quench the reaction, and place on ice.
 Alternatively, the end-labeled protein can be purified from the unincorporated ATP using either a gel filtration column (Kumar et al. 1997) or dialysis.
4. Add 2 μ l of each reaction to 10 μ l of 2x SDS gel-loading dye and heat 3 minutes at 90°C.
5. Load these samples onto an SDS-polyacrylamide protein gel and electrophorese until the bromophenol blue dye has migrated approximately two-thirds of the way to the bottom. Transfer the gel to an old piece of film and wrap thoroughly with plastic wrap. Use autoradiography and densitometry (or a phosphor screen reader after drying the gel under a vacuum for 1 hour at 70°C) to determine the efficiency of radioactive incorporation and the ratio of specific to nonspecific labeling.

Alternatively, the reactions can be analyzed using gel filtration columns (Kumar et al. 1997) or by precipitation:

- a. Incubate 3 μ l of the labeling reaction, 0.5 ml of 0.1 mg/ml bovine serum albumin (as carrier), and 0.5 ml of ice-cold 20% trichloroacetic acid on ice for 30 minutes.
 - b. Centrifuge for 5 minutes at 10,000g, and then wash the pellet with 1 ml of ice-cold acetone.
 - c. Count the precipitated material and 3 μ l of the initial labeling reaction directly (Cerenkov) in a scintillation counter. Determine the incorporation rate by comparing the values.
6. Divide the protein into five 20- μ l aliquots to minimize freeze/thaw cycles (typically each one is sufficient for one experiment), freeze on dry ice, and store at –70°C.

Protocol 2

Fe–EDTA Protease Footprinting Reactions

This protocol was adapted from Heyduk and Heyduk (1994). In the following protocol, the end-labeled protein is assembled quantitatively into the desired complexes.

MATERIALS

CAUTION: See Appendix for appropriate handling of materials marked with <!.>.

Buffers and Solutions

Ammonium iron(II) sulfate (40 mM), hexahydrate <!.>
EDTA (80 mM) (diluted in reaction buffer from a 0.5 M EDTA, pH 8.0 stock solution)
Hydrogen peroxide (10 mM) <!.>
Reaction buffer
 10 mM MOPS (3-[*n*-morpholino]propane sulfonic acid) (pH 7.2) <!.>
 250 mM NaCl
 10 mM MgCl₂ <!.>
3X Sample buffer
 150 mM Tris-HCl (pH 6.8)
 36% glycerol
 12% SDS <!.>
 6% β-mercaptoethanol <!.>
 0.01% bromophenol blue <!.>
Sodium ascorbate (0.2 M) in reaction buffer <!.>

Gels

SDS-polyacrylamide protein gels (glycine or tricine-based)
For further details on SDS-polyacrylamide protein gels, see Sambrook and Russell 2001, *Molecular Cloning* (glycine-based) and this section (tricine-based).

Additional Reagents

Complexes from Protocol 1 and controls

METHOD

1. Prepare the Fe–EDTA complex just prior to use by combining 25 μl of 40 mM ammonium iron(II) sulfate with 25 μl of 80 mM EDTA and then adding 50 μl of reaction buffer. Alternatively, preformed Fe–EDTA can be used (Greiner et al. 1996; Kumar et al. 1997).

2. Assemble complexes for the proteolysis reactions. Controls for the digestion reaction should be prepared in parallel, and among others, typically include end-labeled protein alone (i.e., uncomplexed), in the presence of control affinity matrices that are unable to bind specifically, or in binary species (when a ternary complex is being characterized).
3. Add in the following order: Fe–EDTA, hydrogen peroxide, and ascorbate to assembled complexes and controls.

If there are numerous samples to process, it may simplify this step to add the reagents by placing the appropriate amounts of each reagent at different positions on the underside of the microcentrifuge tube cap, then closing it and centrifuging to simultaneously initiate the reaction.

For cAMP Receptor Protein (CRP), the following amounts were used:

- 1 μ l of Fe–EDTA complex
- 1 μ l of 0.2 M ascorbate
- 1 μ l of 10 mM hydrogen peroxide

Bring the reaction volume to 10 μ l with protein complex (6 μ M CRP) and reaction buffer.

Figure 4 is an example of a footprinting experiment comparing CRP in the presence and absence of its DNA-binding site.

4. Digest for the appropriate time. (Reaction times of 1–30 minutes have been used in various studies.)
5. Stop cleavage by adding 5 μ l of 3X sample buffer. Store the samples at -70°C until separated by electrophoresis.
6. Separate samples on a SDS-polyacrylamide protein gel (glycine or tricine, depending on the molecular weight of the end-labeled protein and the region[s] of interest). Controls that should be loaded onto the same gels include the undigested end-labeled protein, calibration ladders (ones made by site-specific cleavage and/or engineered truncations), and prestained molecular-weight markers. Heat each sample for 3 minutes at 90°C . Load 14 μ l of each sample.
7. Dry the gel under vacuum for 1.5 hours at 70°C . Expose the gel to either a phosphor screen reader or film.

Protocol 3

Proteolysis Using Enzymatic Proteases

In this protocol, each protein molecule is cut a maximum of once along its length.

MATERIALS

CAUTION: See Appendix for appropriate handling of materials marked with <!.>.

Buffers and solutions

Bovine serum albumin (10 mg/ml) (BSA)
Protease inhibitors (e.g., 100 mM phenylmethylsulfonyl fluoride [PMSF] <!.>, 1 mg/ml leupeptin)
2x SDS gel-loading buffer
100 mM Tris-HCl (pH 6.8)
200 mM dithiothreitol <!.>
4% SDS (electrophoresis grade) <!.>
0.2% bromophenol blue <!.>
20% glycerol
Complex assembly buffer (varies according to the particular system)
For the experiment in Figure 4, the buffer contained
10 mM MOPS (pH 7.2)
250 mM NaCl
10 mM MgCl₂

Enzymes and Buffers

Various proteases, e.g., alkaline protease (subtilisin), trypsin, chymotrypsin, typically resuspended at 1 mg/ml as specified by the manufacturer, stored at -70°C in single-use aliquots

Gels

SDS-polyacrylamide protein gels (glycine- or tricine-based)
For further details on SDS-PAGE, see Sambrook and Russell 2001, *Molecular Cloning* (glycine) and this section (tricine).

METHOD

Test Reaction to Optimize Proteolysis

1. On ice, make serial threefold dilutions, from 1:3 to 1:729 of each protease(s) in complex assembly buffer + 0.1 mg/ml BSA.

2. At room temperature, combine 1 pmole of end-labeled protein (~3% of the labeling reaction for high specific activity). Bring the reaction volume to 140 μl with complex assembly buffer. Divide this solution into 20- μl aliquots.
3. Add 2 μl of each protease dilution at 0, 12, 24, 36, 48, and 60 seconds.
4. Incubate each reaction for 2 minutes at room temperature. Add 2 μl of the appropriate protease inhibitor to each reaction and freeze immediately on dry ice.
5. Separate samples on SDS-polyacrylamide protein gel (glycine or tricine, depending on the molecular weight of the end-labeled protein and the region[s] of interest). Add 24 μl of 2 \times SDS gel-loading buffer. Heat each sample for 3 minutes at 90°C. Load 16 μl of each sample. Electrophorese until the bromophenol blue is ~75% of the distance to the bottom of the gel.
6. Dry the gel and expose to either film or a phosphor storage screen to detect the cleavage pattern.

The conditions established from this experiment can be used to perform the protease footprinting experiments. Substitute the following for step 2 above. The complexes are assembled along with the appropriate controls (e.g., uncomplexed protein or protein in the presence of control affinity matrices). After assembly of the complex, it is immediately divided into four aliquots, and three are subjected to limited proteolysis using varying amounts of protease and, as a control, one is not digested. The assembled complexes will generally require two- to five-fold more protease than the free protein.

Protocol 4

Tricine SDS-polyacrylamide Protein Gels

This protocol was adapted from Schagger and von Jagow (1987). This method for detecting cleavage fragments depends on the type of end label and separation system used. It is for a single 16 × 14 × 0.75-cm 16.5% gel. For increased resolution, the length of these gels can be increased.

MATERIALS

CAUTION: See Appendix for appropriate handling of materials marked with <!.>.

Buffers and Solutions

Acrylamide (50%) (48.5% w/v acrylamide, 1.5% w/v bisacrylamide) <!.>

Ammonium persulfate (10%) <!.>

Anode buffer (used in the bottom buffer tank)

0.2 M Tris-HCl (pH 8.9)

Cathode buffer (used in the top buffer tank)

0.1 M Tris base

0.1 M tricine

0.1% SDS <!.>

The pH of this solution is correct without adjusting it.

Gel buffer

3 M Tris-HCl (pH 8.45)

0.3% SDS <!.>

Glycerol (40%) (v/v)

H₂O (deionized)

TEMED (*N,N,N,N*-tetramethylethylenediamine) <!.>

METHOD

1. For the separating gel, combine 7 ml each of 50% acrylamide, gel buffer, and 40% glycerol. Filter, then add 70 µl of 10% ammonium persulfate and 7 µl of TEMED and pour into a vertical gel mold. Overlay with deionized H₂O (or 20% ethanol) until polymerized.
In some cases, eliminating or reducing the glycerol in the gel has no detrimental effects (R. Hori, unpubl.).
2. Pour off H₂O from the separating gel.
3. Overlay with a stacking gel and then insert a gel comb. For the stacking gel, combine 1 ml of 50% acrylamide, 3.1 ml of gel buffer, 12.5 ml of deionized H₂O; filter, then add 100 µl of 10% ammonium persulfate and 10 µl of TEMED.
4. Inject the samples into the wells at 14 mA and run the gels overnight at 24 mA.
5. Dry the gels under a vacuum for 1.5 hours at 70°C and expose to a phosphor storage screen or film.

If desired, the gels can be fixed using a solution of 45% methanol, 10% acetic acid. One problem associated with fixing the gels is an increased tendency to crack during drying.

Protocol 5

Calibration Ladder Generated by Cleavage of Proteins by CNBr

This method, adapted from Gross (1967), targets methionine residues to generate mobility standards.

MATERIALS

CAUTION: See Appendix for appropriate handling of materials marked with <!.>.

Buffers and Solutions

Acetonitrile <!.>
Crystalline cyanogen bromide (CNBr) (stored in a can at 4°C) (very toxic) <!.>
H₂O (deionized)
HCl (1 N) <!.>
3x Sample buffer (see Protocol 2)
SDS (10%) <!.>

Gels

SDS-polyacrylamide protein gels (glycine- or tricine-based)
For further details on SDS-PAGE, see Sambrook and Russell 2001, *Molecular Cloning*, and this section.

Special Equipment

Chemical fume hood
Phosphor storage screen
SpeedVac
Ultrafreezer

METHOD

1. Prefreeze a plastic rack in a -70°C ultrafreezer.
2. Allow CNBr (within its container) to warm up under a chemical fume hood to room temperature.
3. Tare a balance with an empty microfuge tube.
4. Under the fume hood, add a small amount (a few crystals) of CNBr to microfuge tube (as quickly as possible to reduce inhalation).
5. Weigh CNBr in the closed microfuge tube.

6. Under the fume hood, add appropriate amount of acetonitrile to obtain a 106 mg/ml stock of CNBr.
7. Assemble the following reaction mix:

10% SDS	0.75 μ l
1 N HCl	0.4 μ l
CNBr solution	10 μ l

Add CNBr last to initiate the reaction. Bring the final volume to 25 μ l with end-labeled protein and deionized H₂O.
8. Incubate at room temperature. Take the reactions to the -70°C ultrafreezer. Remove the samples (5 μ l) at different time points (e.g., typically every 5 or 10 minutes) and place in the rack in the -70°C ultrafreezer.
9. At the conclusion of the reactions, dry the frozen reactions in a SpeedVac.
10. Resuspend the dried pellet in 3 \times sample buffer (50 μ l). Run the samples (typically, 5 μ l, but it depends on the specific activity of the end-labeled protein) from different time points on a denaturing SDS-polyacrylamide gel to determine optimal digestion time.
11. Dry the gel under vacuum for 1.5 hours at 70°C and expose to film or a phosphor storage screen to detect the cleavage pattern. This pattern should be compared with the position of methionine residues in the protein's primary sequence.
12. Add H₂O to stock CNBr for disposal. Soak the spatula that was used to weigh the CNBr crystals overnight in H₂O.

Protocol 6

Calibration Ladder Generated by Enzymatic Cleavage

Here, mobility standards are generated by targeting arginines (Clostripain) or lysines (Endoproteinase Lys-C).

MATERIALS

CAUTION: See Appendix for appropriate handling of materials marked with <!.>.

Buffers and Solutions

Bovine serum albumin (10 mg/ml) (BSA)

Denaturation buffer

20 mM HEPES (pH 7.9)

100 mM KCl <!.>

2 M urea <!.>

0.2% SDS <!.>

0.5 M EDTA

EDTA (0.5 M)

2x SDS gel-loading buffer

100 mM Tris-HCl (pH 6.8)

200 mM dithiothreitol <!.>

4% SDS (electrophoresis grade) <!.>

0.2% bromophenol blue <!.>

20% glycerol

Enzymes and Buffers

Site-specific proteases (sequencing grade or the best available), e.g., Clostripain (which cleaves at arginines) or Endoproteinase Lys-C (which cleaves at lysines). Resuspend as specified by the manufacturer.

Additional Reagents

End-labeled protein of interest

Special Equipment

Autoradiography or phosphor image reader

METHOD

1. Add 1 pmole (~3% of a labeling reaction for high specific activity) of the end-labeled protein to 50 μl of denaturation buffer.
2. Heat for 30 minutes at 65°C.
3. Transfer the reaction mixture to 37°C.
4. Add 2 μg of Clostripain (or other site-specific protease such as Endoproteinase Lys-C).
5. Remove 10 μl at 2, 6, 18, and 54 minutes. Terminate the reaction by adding 1 μl of 0.5 M EDTA and 1 μl of 10 mg/ml BSA and freezing on dry ice. Store at -70°C.
6. Separate the samples on a tricine-based denaturing SDS-polyacrylamide gel to examine the proteolysis pattern. A standard SDS-polyacrylamide gel can be used if the end-labeled protein is very large or high resolution of small fragments is not necessary. Add 12 μl of 2 \times SDS gel-loading buffer. Heat each sample for 3 minutes at 90°C. Load 6 μl of each sample. Electrophorese until the bromophenol blue is ~75% of the distance to the bottom of the gel.
7. Dry the gel under vacuum for 1.5 hours at 70°C.
8. Visualize the cleavage pattern by autoradiography or a phosphor image reader.

REFERENCES

- Baichoo N. and Heyduk T. 1999. DNA-induced conformational changes in cyclic AMP receptor protein: Detection and mapping by a protein footprinting technique using multiple chemical proteases. *J. Mol. Biol.* **290**: 37–48.
- Brenowitz M., Senear D.F., Shea M.A., and Ackers G.K. 1986. Quantitative DNase footprint titration: A method for studying protein-DNA interactions. *Methods Enzymol.* **130**: 132–181.
- Callaci S., Heyduk E., and Heyduk T. 1998. Conformational changes of *Escherichia coli* RNA polymerase σ^{70} factor induced by binding to the core enzyme. *J. Biol. Chem.* **273**: 32995–33001.
- Casaz P. and Buck M. 1999. Region I modifies DNA-binding domain conformation of sigma 54 within the holoenzyme. *J. Mol. Biol.* **285**: 507–514.
- Cheng X., Shaltiel S., and Taylor S.S. 1998. Mapping substrate-induced conformational changes in cAMP-dependent protein kinase by protein footprinting. *Biochemistry* **37**: 14005–14013.
- Clore G.M. and Gronenborn A.M. 1998. NMR structure determination of proteins and protein complexes larger than 20 kDa. *Curr. Opin. Chem. Biol.* **2**: 564–570.
- Cohen S.L., Ferre-D'Amare A.R., Burley S.K., and Chait B.T. 1995. Probing the solution structure of the DNA-binding protein Max by a combination of proteolysis and mass spectrometry. *Protein Sci.* **4**: 1088–1099.
- Croft S., Gilbert B.C., Smith J.R., and Whitwood A.C. 1992. An E.S.R. investigation of the reactive intermediate generated in the reaction between FeII and H₂O₂ in aqueous solution. Direct evidence for the formation of the hydroxyl radical. *Free Radic. Res. Commun.* **17**: 21–39.
- Datwyler S.A. and Meares C.F. 2000. Protein-protein interactions mapped by artificial proteases: Where sigma factors bind to RNA polymerase. *Trends Biochem. Sci.* **25**: 408–414.
- Dixon W.J., Hayes J.J., Levin J.R., Weidner M.F., Dombroski B.A., and Tullius T.D. 1991. Hydroxyl radical footprinting. *Methods Enzymol.* **208**: 380–413.
- Greiner D.P., Hughes K.A., Gunasekera A.H., and Meares C.F. 1996. Binding of the sigma 70 protein to the core subunits of *Escherichia coli* RNA polymerase, studied by iron-EDTA protein footprinting. *Proc. Natl. Acad. Sci.* **93**: 71–75.
- Gross E. 1967. The cyanogen bromide reaction. *Methods Enzymol.* **11**: 238–255.
- Hanai R. and Wang J.C. 1994. Protein footprinting by the combined use of reversible and irreversible lysine modifications. *Proc. Natl. Acad. Sci.* **91**: 11904–11908.
- Heyduk E. and Heyduk T. 1994. Mapping protein domains involved in macromolecular interactions: A novel protein footprinting approach [published erratum appears in *Biochemistry* 1995 Nov 21; **34**(46): 15388]. *Biochemistry* **33**: 9643–9650.
- Heyduk T., Baichoo N., and Heyduk E. 2001. Probing of proteins by metal ions and their low molecular weight complexes. In *Metal ions in biological systems* (ed. A. Sigel and H. Sigel), vol. 38, pp. 255–287. M. Dekker, Inc., New York.
- Heyduk T., Heyduk E., Severinov K., Tang H., and Ebricht R.H. 1996. Determinants of RNA polymerase alpha subunit for interaction with beta, beta', and sigma subunits: Hydroxyl-radical protein footprinting. *Proc. Natl. Acad. Sci.* **93**: 10162–10166.
- Hori R. and Carey M. 1998. Protease footprinting analysis of protein-protein and protein-DNA interactions. In *Human genome methods* (ed. K. Adolph), pp. 73–90. CRC Press, Boca Raton.
- Kriwacki R.W., Hengst L., Tennant L., Reed S.I., and Wright P.E. 1996. Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: Conformational disorder mediates binding diversity. *Proc. Natl. Acad. Sci.* **93**: 11504–11509.
- Kumar A., Kassavetis G.A., Geiduschek E.P., Hambalko M., and Brent C.J. 1997. Functional dissection of the B'' component of RNA polymerase III transcription factor IIIB: A scaffolding protein with multiple roles in assembly and initiation of transcription. *Mol. Cell. Biol.* **17**: 1868–1880.
- Li B.L., Langer J.A., Schwartz B., and Pestka S. 1989. Creation of phosphorylation sites in proteins: Construction of a phosphorylatable human interferon alpha. *Proc. Natl. Acad. Sci.* **86**: 558–562.
- Lindsley J.E. and Wang J.C. 1993. Study of allosteric communication between promoters by immunotagging. *Nature* **361**: 749–750.
- Loizos N. and Darst S.A. 1998. Mapping protein-ligand interactions by footprinting, a radical idea. *Structure* **6**: 691–695.
- . 1999. Mapping interactions of *Escherichia coli* GreB with RNA polymerase and ternary elongation complexes. *J. Biol. Chem.* **274**: 23378–23386.
- Maleknia S.D., Brenowitz M., and Chance M.R. 1999. Millisecond radiolytic modification of peptides by synchrotron X-rays identified by mass spectrometry. *Anal. Chem.* **71**: 3965–3973.
- McMahan S.A. and Burgess R.R. 1999. Mapping protease susceptibility sites on the *Escherichia coli* transcription factor sigma70. *Biochemistry* **38**: 12424–12431.

- Owens J.T., Chmura A.J., Murakami K., Fujita N., Ishihama A., and Meares C.F. 1998. Mapping the promoter DNA sites proximal to conserved regions of sigma 70 in an *Escherichia coli* RNA polymerase-lacUV5 open promoter complex. *Biochemistry* **37**: 7670–7675.
- Pestka S., Lin L., Wu W., and Izotova L. 1999. Introduction of protein kinase recognition sites into proteins: A review of their preparation, advantages, and applications. *Protein Exp. Purif.* **17**: 203–214.
- Schagger H. and von Jagow G. 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166**: 368–379.
- Schwanbeck R., Manfioletti G., and Wisniewski J.R. 2000. Architecture of high mobility group protein I-C-DNA complex and its perturbation upon phosphorylation by Cdc2 kinase. *J. Biol. Chem.* **275**: 1793–1801.
- Songyang Z., Blechner S., Hoagland N., Hoekstra M.F., Piwnica-Worms H., and Cantley L.C. 1994. Use of an oriented peptide library to determine the optimal substrates of protein kinases. *Curr. Biol.* **4**: 973–982.
- Travaglia S.L., Datwyler S.A., and Meares C.F. 1999. Mapping protein-protein interactions with a library of tethered cutting reagents: The binding site of sigma 70 on *Escherichia coli* RNA polymerase. *Biochemistry* **38**: 4259–4265.
- Tu B.P. and Wang J.C. 1999. Protein footprinting at cysteines: Probing ATP-modulated contacts in cysteine-substitution mutants of yeast DNA topoisomerase II. *Proc. Natl. Acad. Sci.* **96**: 4862–4867.
- Zhong M., Lin L., and Kallenbach N.R. 1995. A method for probing the topography and interactions of proteins: Footprinting of myoglobin. *Proc. Natl. Acad. Sci.* **92**: 2111–2115.

