Use of Site-Specific Protein–DNA Photocrosslinking to Analyze the Molecular Organization of the RNA Polymerase II Initiation Complex

François Robert and Benoit Coulombe

1. Introduction

Site-specific protein–DNA photocrosslinking has proved to be the method of choice for analysis of the formation of nucleoprotein complexes such as those involved in transcription by mammalian RNA polymerase II (RNA Pol II). The method has two principal advantages. First, it yields structural information on large, multisubunit complexes that in general cannot be analyzed using standard high-resolution techniques such as X-ray crystallography or nuclear magnetic resonance (NMR). For example, site-specific protein–DNA photocrosslinking, in conjunction with complementary methods such as protein-affinity chromatography and electron microscopy, has produced information on both the molecular organization and the composition of the RNA Pol II pre-initiation complex on promoter DNA (1–4). This complex contains RNA Pol II and the general transcription factors TBP, TFIIA, TFIIB, TFIIE, TFIIIF (RAP74 and RAP30), and TFIIH, and is composed of more than 25 polypeptides ranging in M_r from 10 to 220 kDa (5). Neither X-ray crystallography nor NMR, which can only resolve the structure of complexes containing short protein fragments bound to small pieces of promoter DNA, could provide any detailed structural information on this complex. Second, the method has sufficient technical flexibility so as to allow the rapid analysis of complexes assembled under various conditions. Over the past few years, we have analyzed a large collection of complexes assembled in the presence of various combinations of the general transcription factors (wild-type or different deletion mutants) and RNA Pol II (1–4). These experiments have enabled us to draw conclusions on the dynamics of RNA Pol II pre-initiation complex assembly and has led to the notion that isomerization of the RNA Pol II pre-initiation complex proceeds through wrapping of the promoter DNA around the enzyme (4).

Site-specific protein–DNA photocrosslinking is a method composed of two successive steps. First, a number of photoprobes that place one (or a few) photoreactive nucleotide(s) into juxtaposition with one (or a few) radiolabeled nucleotide(s) at various specific positions along the promoter DNA are prepared. Second, transcription complexes are assembled onto the various photoprobes, irradiated with ultraviolet (UV) light so as to induce protein–DNA crosslinking, and the processed in order to identify the crosslinked polypeptides. Because the crosslinking of protein to DNA is site-specific, the use of a series of photoprobes that

place the photonucleotide derivative at various positions along the promoter DNA provides information on the relative position of the various factors within the complex.

2. Materials

1. Buffer A (10X): 300 mM Tris-HCl, pH 8.0, 500 mM KCl, and 70 mM MgCl₂, freshly prepared.
2. Bovine serum albumin (BSA) solution: Prepare a 25-mg/mL solution of BSA in deionized distilled water. Store in aliquots at -20°C.
3. Dilute with water to 5 mg/mL prior to use.
4. dNTP mix: 20 mM each of dATP, dCTP, dGTP, and dTTP in buffer A (1X), freshly prepared.
5. ND buffer: 20 mM HEPES, pH 7.9, 100 mM KCl, 20% glycerol, 0.2 mM EDTA, 0.2 mM EGTA, and 10 mM of β-mercaptoethanol. Store in aliquots at -20°C.
6. TBE buffer (10X): Prepare 1 L by mixing 108 g Tris base, 55 g boric acid, and 40 mL EDTA (0.5 M, pH 8.0).
7. Gel loading solution (10X): For polyacrylamide native gels, use a solution containing 0.25% bromophenol blue, 0.25% xylene cyanol, and 25% Ficoll (type 400) in deionized distilled water.
8. MBS (5X): 40 mM MgCl₂, 100 mM HEPES, pH 7.9, 100 µg/mL BSA, and 5 mM ATP.
9. Store in aliquots at -20°C.
10. Complex mix: 50 µL MBS (5X) and 2 µL (NH₄)₂SO₄ (2 M), freshly prepared.
11. Poly(dI.dC-dI.dC) stock: Prepare a 25-mg/mL solution of poly(dI.dC-dI.dC) in deionized distilled water. Store in aliquots at -20°C.
12. DNase mix: A solution containing 200 units/mL DNase I and 32 mM CaCl₂, freshly prepared.
13. Acid mix: Mix equal volumes of 5% acetic acid and 30 mM ZnCl₂, freshly prepared.
14. S1 mix: 30,000 U/mL S1 nuclease in deionized distilled water, freshly prepared.
15. Dilution mix: 50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 10 mM EGTA, 500 mM NaCl, 5 mM NaF, 1 mM phenylmethylsulfonyl fluoride (PMSF), and 1% (v/v) Triton X-100, freshly prepared.
16. Protein A mix: 100 mg/mL Protein A in the dilution mix.
17. Wash 1 solution: 10 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1% Triton X-100, 1% NP-40, 0.1% sodium dodecyl sulfate (SDS), 1 mg/mL BSA, 0.5% NaN₃, and 0.5% Na-deoxychlorate, freshly prepared.
18. Wash 2 solution: 10 mM Tris-HCl, pH 8.0, and 1.5 M NaCl.

19. Wash 3 solution: 10 mM Tris-HCl, pH 8.0, and 150 mM NaCl.

3. Methods

3.1. Day 1: Synthesis and Purification of the Photoprobes

The first step of the procedure is the synthesis of the photoprobes. This part is illustrated in Fig. 1 which shows, as an example, a scheme for the synthesis of a photoprobe designed to place the photoreactive nucleotide at position –2 relative to the transcriptional initiation site of the adenovirus major late promoter. The site-specific incorporation of the photoreactive nucleotide (*see* ^{Note 1}) and the radiolabeled nucleotide is directed through the annealing of a primer, referred to as the specific primer, with a single-stranded DNA template containing the promoter DNA. The promoter is flanked by two restriction sites (in this example, *DraI* and *SacI*). A second primer, referred to as the upstream primer, is annealed a few base pairs upstream of the *DraI* site. After annealing, the photoreactive and radiolabeled nucleotides are incorporated by primer extension using T4 DNA polymerase with limiting amounts of dNTPs (*see* ^{Note 2}). After the labeling step, the extension reaction is completed by the addition of an excess of cold dNTPs (*see* ^{Note 3}). The photoprobe is generated by digestion with the restriction enzymes and gel purified (*see* Fig. 2 for an example of a gel on which the products of a photoprobe synthesis reaction have been separated).

- 1 Mix 700 ng of single-stranded (ss) DNA (approx 0.5 pmol) with 40 ng (approx 5 pmol) of both the specific and upstream primers. Add 1 µL of Buffer A (10X) and complete to 10 µL with deionized distilled water.
- 2 Mix well and incubate for 3 min at 90°C.
- 3 Incubate for 30 min at room temperature.
- 4 From this point on, all manipulations must be carried out under reduced light conditions (*see* ^{Note 4}). Add 1 µL BSA (5 mg/mL), 1 µL N₃R–dUMP (*see* ^{Note 1}), 25 µCi of the appropriate [α -³²P]–dNTP ([α -³²P]–dCTP for the example shown in Fig. 1), 5–10 U T4 DNA polymerase, and 1 µL buffer A (10X). Complete to a final volume of 20 µL with deionized distilled water.
- 5 Incubate for 30 min at room temperature.
- 6 Add 5 µL of dNTP mix.
- 7 Incubate for 5 min at room temperature.

¹The nucleotide derivative we use, namely 5-[*N*-(azidobenzoyl)-3-aminoallyl]–dUMP (N₃R–dUMP or AB–dUMP) (*see* Chapter 25 for further details on the structure and the chemical synthesis of N₃R–dUMP), possesses a side chain that places a reactive nitrene 10 Å away from the DNA backbone in the major groove of the double helix. For this reason, the crosslinking of a polypeptide to the photoprobe does not require a direct interaction of the polypeptide with the DNA helix. The amount of N₃R–dUMP to be added to the reaction is determined empirically for each preparation of the photoreactive nucleotide and is generally between 0.5 and 2 µL (often 1 µL).

²The specific primer must be designed in such a manner that T4 DNA polymerase only adds a few nucleotides. In the example shown in Fig. 1, the incorporation is restricted to positions –4 to –1 by omitting dATP from the reaction (*see* **refs.** 1–4 for additional examples). The success of this step can be monitored by analysis of the reaction products on a sequencing gel.

³The addition of dNTPs in large excess is crucial because it is necessary to limit primer extension to the incorporation of standard dNTPs, preventing the addition of any more radiolabeled and photoreactive nucleotides.

⁴The use of a standard dark room is not necessary at this stage. As a rule, we find that conditions that provide just enough light to be able to work are acceptable.

- 8 Incubate for 20 min at 37°C.
- 9 Add 10–20 U of each restriction enzyme (*SacI* and *DraI* in the example shown in Fig. 1).
- 10 Incubate for 90 min at the temperature recommended by the supplier of restriction enzymes (37°C for *SacI* and *DraI*).
- 11 Add 3.5 µL of gel loading solution (10X).
- 12 Load on a native 8% polyacrylamide gel in TBE buffer (1X).
- 13 Run at 150 V for about 1 h in TBE buffer (1X).
- 14 Remove the glass plates containing the gel from the gel box.
- 15 Separate the glass plates and leave the gel on one of them.
- 16 Wrap the gel/glass plate in plastic wrap.
- 17 Wrap the entire package in aluminum foil.
- 18 Move to a dark room (*see* ^{Note 5}).
- 19 Place a Kodak X-OMAT AR film on a clean bench.
- 20 Remove the foil and place the gel on the film with the glass plate facing up (e.g., gel side down).
- 21 Expose 5 min.
- 22 During the exposition time, mark the film using a sharp tool by tracing the contour of the glass plate (this will be helpful later for the localization of the photoprobe in the gel).
- 23 Remove the gel and rewrap it with the foil.
- 24 Develop the film (*see* ^{Note 6}).
- 25 Using a scalpel, cut the film so that the square piece containing the band corresponding to the photoprobe is removed. This operation leaves the film with a window at the position of the photoprobe.
- 26 Superimpose the film on the gel by taking advantage of the marks made in **step 22**, and mark the square corresponding to the photoprobe on the saran wrap using a pen.
- 27 Cut out the gel slice containing the photoprobe using a clean scalpel.
- 28 Cut the gel slice in small pieces (six to eight fragments).
- 29 Place the gel fragments in an Eppendorf tube and add water in order to completely submerge the gel (usually 100–150 µL of water).

⁵A conventional red light can be used.

⁶An example of gel photoprobe purification is shown in Fig. 2. The position of the band corresponding to the photoprobe can be easily identified because the size of the DNA fragment generated by digestion with restriction enzymes is known.

- 30 Incubate overnight at room temperature.
- 31 Collect the liquid containing the probe.
- 32 Purify the probe on a Micro-Spin S-200 HR column (Pharmacia Biotech) to remove any salts and other putative contaminants.
- 33 Count 1 μL of the photoprobe solution by liquid scintillation, and dilute the probe to 1250 cpm/ μL with deionized distilled water.
- 34 The probe is ready for use and can be stored in the dark at 4°C for 1–2 wk (*see* Note 7).

3.2. Day 2: Protein–DNA Photocrosslinking

The gel-purified photoprobe is used for pre-initiation complex assembly with the purified transcription factors (TBP, TFIIA, TFIIB, TFIIE, TFIIIF, and TFIID) and RNA Pol II. Reactions are irradiated with ultraviolet (UV) light to induce protein–DNA crosslinking and treated with DNase I and S1 nuclease in order to liberate polypeptides that are covalently attached to a very short piece of DNA carrying one to four radiolabeled nucleotides. After sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) separation of the photocrosslinked polypeptides, the gel is dried and exposed to X-ray film. Examination of the ensuing autoradiogram permits identification of the protein(s) that interact with a particular site. For the RNA Pol II initiation complex, the specificity of the photocrosslinking signals can be assessed by comparing reactions performed with photoprobes containing either a wild-type or a mutated TATA element and/or by comparing reactions performed in either the presence or the absence of TBP (*see* Note 8). The photocrosslinked polypeptides can be identified according to their M_r (Fig. 3A,B) and to their immunoreactivity with specific antibodies (Fig. 3C). In the latter case, crosslinking reactions are submitted to immunoprecipitation with antibodies directed against specific factors prior to SDS-PAGE (*see* Note 9).

⁷Fresh probes (less than a week old) give the best results.

⁸Because some of the general transcription factors and RNA Pol II bind non-specifically to DNA, it is important to discriminate between specific and nonspecific crosslinking signals. For the adenovirus major late promoter, it is well-documented that mutations of two bases in the TATA box (TATAAA to TAGAGA) completely abolish pre-initiation complex formation. The comparison of crosslinking signals obtained with probes containing either a wild-type or a mutated TATA box permits differentiation between specific and nonspecific signals. However, this strategy is not simple, as it doubles the number of probes to be synthesized. Because the basic function of the TATA box is to bind TBP, the specificity can be assessed by comparing crosslinking reactions performed in the presence or the absence of TBP. The absence of TBP was found to give the same result as the use of a probe with a mutated TATA box (1–4). An example is shown in Fig. 3A. The crosslinking signal at the top of the gel (RPB1) is considered as nonspecific because its intensity is not affected by either a mutation in the TATA box or the absence of TBP. The two additional crosslinking signals (RPB2 and RAP74) are specific because their intensities are significantly affected by both mutation of the TATA box and the omission of TBP in the crosslinking reaction.

⁹Identification of the photocrosslinked polypeptides is a central issue in the method. The main difficulty comes from the fact that several factors have a M_r between 30 and 40 kDa. At least three different means can be used to identify the crosslinking signals. First, SDS-PAGE analysis provides direct information on the size of photocrosslinked polypeptides (*see* Fig. 3A). Second, the use of truncated forms of a factor can be useful. An example is shown in Fig. 3B in which all factors included in the crosslinking reactions are the same, except for RAP74, which is either full-length (lanes 1 and 2) or truncated in its C-terminus (lanes 3 and 4). The different fragments of RAP74 migrate with different mobilities, allowing the identification of the photocrosslinked polypeptide. Third, the photocrosslinked polypeptides can be identified after immunoprecipitation with a specific antibody. Following nuclease treatment, the crosslinking products are immunoprecipitated and then submitted to SDS-PAGE analysis. An example is shown in Fig. 3C, where the photocrosslinked polypeptides shown in lane 1 have been immunoprecipitated using an antibody raised against RAP74 (lane 3) or a control antibody (lane 4). This shows that the photocrosslinked polypeptide is RAP74.

- 35 Mix the proteins (50–200 ng each) and complete the volume to 16 μ L with ND buffer (*see* ^{Note 10}). Add 5.2 μ L of the complex mix, 1 μ L of diluted poly(dI.dC–dI.dC) (*see* ^{Note 11}) and 5000 cpm of the photoprobe (4 μ L of 1250 cpm/ μ L). The final volume is 26.2 μ L.
- 36 Mix well and incubate for 30 min at 30°C.
- 37 Open the lids of the tubes and irradiate 10 min with UV light (*see* ^{Notes 12 and 13}).
- 38 Add 5 μ L of DNase mix.
- 39 Incubate for 20 min at 37°C.
- 40 Add 1.5 μ L of 10% SDS.
- 41 Incubate for 3 min at 90°C.
- 42 Spin for 10 s in a microfuge (16,000*g*)
- 43 Add 2 μ L of Acid mix.
- 44 Add 1 μ L of S1 mix.
- 45 Incubate for 20 min at 37°C.
- 46 For regular reactions, go directly to **step 62**.
- 47 For immunoprecipitation reactions, add 315 μ L of the dilution mix.
- 48 Add 5 μ L of antibody (*see* ^{Note 14}).
- 49 Incubate for 60 min at 4°C.
- 50 Add 80 μ L of Protein A mix.
- 51 Agitate 60 min at 4°C using a rocker.
- 52 Spin down the Protein A beads (1 min in a microfuge at 16,000*g*).
- 53 Remove the supernatant by pipetting.
- 54 Wash with 400 μ L of wash 1 solution.
- 55 Repeat **steps 52–54** twice.
- 56 Wash with 400 μ L of wash 2 solution.
- 57 Spin down the Protein A beads (1 min in a microfuge at 16,000*g*).

¹⁰In the crosslinking reactions, we routinely use 200 ng of each recombinant human (rh) TFIIB, rhRAP30, rhRAP74, rhTFIIE34, and rhTFIIE56, 100 ng of calf thymus RNA Pol II, 50 ng of natural human TFIIF, 50–200 ng of rhTFIIA, and 200 ng of recombinant yeast TBP. The amounts of the different protein factors should be optimized for each different combination of proteins and for each protein preparation.

¹¹The poly(dI.dC–dI.dC) stock should be diluted just prior to use. The exact dilution should be determined experimentally in order to favor specific versus non-specific signals without adversely affecting the intensity of the specific signals.

¹²Irradiation time with UV light should be optimized by performing a time-course with the particular system to be used. We use a UV Stratalinker 2400 (Stratagene) with 254-nm bulbs.

¹³From this point on, normal light conditions can be used.

¹⁴The amounts of antibody to be used should be optimized. Purified antibodies provide the best results.

- 58 Remove the supernatant by pipetting.
- 59 Wash with 400 μ L of wash 3 solution.
- 60 Spin down the Protein A beads (1 min in a microfuge at 16,000*g*).
- 61 Remove the supernatant by pipetting.
- 62 Add SDS gel loading solution and boil 5 min (*see* ^{Note 15}).
- 63 Resolve the photocrosslinked polypeptides by SDS-PAGE (run at 30 mA in the stacking gel and 50 mA in the separating gel) (*see* ^{Note 15}).
- 64 Transfer the gel to Whatman paper and dry.
- 65 Expose the dried gel to X-ray film using an intensifying screen (*see* ^{Note 16}).

Acknowledgments

We thank members of our laboratory for valuable discussions and Will Home for critical reading of the manuscript. B.C. is the recipient of funding from the Medical Research Council of Canada, the Cancer Research Society and the Fonds de la Recherche en Santé du Québec. F.R. holds a studentship from the FCAR.

References

1. Coulombe B, Li J, Greenblatt J. Topological localization of human transcription factors IIA, IIB, TATA box-binding protein and RNA polymerase II-associated protein 30 on a class II promoter. *J Biol Chem.* 1994; 269:19,962–19,967.
2. Robert F, Forget D, Li J, Greenblatt J, Coulombe B. Topological localization of transcription factors IIE and IIF immediately upstream the transcription start site of a class II promoter. *J Biol Chem.* 1996; 271:8517–8520. [PubMed: 8621472]
3. Forget D, Robert F, Grondin G, Burton ZF, Greenblatt J, Coulombe B. RAP74 induces promoter contacts of RNA polymerase II upstream and downstream of a DNA bend at the TATA box. *Proc Natl Acad Sci USA.* 1997; 94:7150–7155. [PubMed: 9207059]
4. Robert F, Douziech M, Forget D, Egly JM, Greenblatt J, Burton ZF, et al. Wrapping of promoter DNA around the RNA polymerase II initiation complex induced by TFIIF. *Mol Cell.* 1998; 2:341–351. [PubMed: 9774972]
5. Hampsey M. Molecular genetics of the RNA polymerase II general transcription machinery. *Microbiol Mol Biol Rev.* 1998; 62:465–503. [PubMed: 9618449]
6. Sambrook, J., Fritsch, EF., Maniatis, T. *Molecular Cloning: A Laboratory Manual.* 2. Cold Spring Harbor Laboratory; Cold Spring Harbor, NY: 1989.

¹⁵Detailed procedures for SDS-PAGE electrophoresis have been described (e.g., *see* **ref.** 6).

¹⁶The use of BioMax (Kodak) screens and films is recommended.

Fig. 1. Synthesis and purification of the photoprobes. Schematic representation of the synthesis of a photoprobe that places a photoreactive nucleotide at position -2 and three radiolabeled nucleotides at positions -4, -3, and -1 of the adenovirus major late promoter. The sequence of the adenovirus major late promoter flanked by plasmid DNA is shown. The primers used to direct the incorporation of N_3R -dUMP (U) and radiolabeled nucleotides (*) are shown in boxes. The TATA element is in bold type and the transcriptional initiation site is indicated by an arrow. *Dra*I and *Sac*I restriction sites used for the excision of the photoprobe are indicated. In the first step, the primers are annealed to single-stranded DNA containing the promoter sequence (Annealing). In the second step, the photoreactive and the radiolabeled nucleotides are incorporated using T_4 DNA polymerase (Labeling). In the third step, the extension reaction is completed by a chase with excess of cold dNTPs (Chase). In the fourth

step, the site-specifically labeled promoter fragment is excised using restriction enzymes and gel purified (Digestion and purification).

Fig. 2.

Example of the autoradiogram of a gel after electrophoresis of the products of a synthesis reaction. The positions of the photoprobe, the free nucleotides, and the plasmid fragment are shown. The photoprobe is then purified from the gel.

Fig. 3.

Typical SDS-PAGE gels of photocrosslinked proteins. **(A)** Crosslinking reactions assembled either in the absence of TBP or using a probe with a mutated TATA box give identical results. Crosslinking reactions were performed with TFIIB, RAP30, RAP74, TFIIE34, TFIIE56, and RNA Pol II in the presence (+) or in the absence (–) of TBP using either a wild-type (TATAAA) or a mutated (TAGAGA) TATA element. In the example shown here (the photonucleotide derivative is placed at position –15 of the adenovirus major late promoter), the crosslinking of both the second largest subunit of RNA Pol II (RPB2) and RAP74 are considered to be specific (dark arrowhead). The crosslinking of RPB1 is considered as nonspecific (open arrowhead) because it is not affected by the omission of TBP or the use of a probe with a mutation in the TATA box. The positions of the molecular weight markers (MW) are indicated. **(B)** The use of truncated polypeptides as a tool to identify the photo-crosslinked polypeptides. Crosslinking reactions were performed using TFIIB, RAP30, TFIIE34, TFIIE56, and RNA Pol II in the presence (+) or in the absence (–) of TBP, using full-length (1–517) or truncated (1–409) RAP74. The different mobilities of RAP74 fragments are diagnostic for RAP74 contact at this promoter position. **(C)** Immunoprecipitation of photocrosslinked polypeptides. Crosslinking reactions were performed using TFIIB, RAP30, RAP74, TFIIE34, TFIIE56, and RNA Pol II in the presence (+) or in the absence (–) of TBP. The photocrosslinked polypeptides were either processed normally (first two lanes) or immunoprecipitated using an antibody directed against RAP74 or a control antibody.