Mapping the Active Site of Yeast RNA Polymerase B (II)*

Michel Riva‡‡, Christophe Carles§, and André Sentenac‡

From the Département de Biologie, Service de Biochimie et Laboratoire d’Ingénierie des Protéines, Centre d’Études Nucléaires de Saclay, 91191 Gif-sur-Yvette Cedex, France and the Station de Recherches Laitières, Groupe de Biochimie des Protéines, Centre de Recherche Jouy Institut National de la Recherche Agronomique, 78350 Jouy-en-Josas, France

Mikhail A. Grachev, Arkady A. Mustaev, and Evgeny F. Zaychikov

From the Limnological Institute of the Siberian Branch of the Academy of Sciences, Irkutsk, Union of Soviet Socialist Republics

Affinity Reagents—The structures of the affinity reagents used in this work are indicated in Scheme 1. The affinity reagents I–IV and affinity labeling technique to investigate the active site of prokaryotic and eukaryotic RNA polymerases. The substrate-binding site was found to reside in the E. coli β subunit (23) and in the second largest subunit of wheat germ enzyme B (II) (24), as well as of yeast RNA polymerases A (I), B (II), and C (III) (25), confirming the functional homology of these polypeptides. Mapping of the affinity labeled site in E. coli and an archaebacterial RNA polymerase identified a few candidate tryptophan residues in two conserved domains of the C-terminal part of the β or B subunit (26, 27).

The availability of the amino acid sequence for the R150 subunit of yeast RNA polymerase B (II) (13) prompted us to localize the site of affinity labeling using microsequence analysis in combination with a mapping technique (26), based on extensive or single-hit chemical cleavage of the labeled subunit. We present here the results of the localization of the nucleotide-binding site and show that it is restricted to the C terminus of the B150 subunit between Asn546 and Met699. This domain is part of one of the nine conserved regions that have been previously defined.

MATERIALS AND METHODS

Affinity Reagents—The structures of the affinity reagents used in this work are indicated in Scheme 1. The affinity reagents I–IV and VII–XIII were prepared as previously described (23, 26–28). The synthesis of derivatives V and VI was performed as for derivatives II and III using p-hydroxybenzaldehyde instead of o-hydroxybenzaldehyde. The availability of the amino acid sequence of the R150 subunit of yeast RNA polymerase B (II) (13) prompted us to localize the site of affinity labeling using microsequence analysis in combination with a mapping technique (26), based on extensive or single-hit chemical cleavage of the labeled subunit. We present here the results of the localization of the nucleotide-binding site and show that it is restricted to the C terminus of the B150 subunit between Asn546 and Met699. This domain is part of one of the nine conserved regions that have been previously defined.

MATERIALS AND METHODS

Affinity Reagents—The structures of the affinity reagents used in this work are indicated in Scheme 1. The affinity reagents I–IV and VII–XIII were prepared as previously described (23, 26–28). The synthesis of derivatives V and VI was performed as for derivatives II and III using p-hydroxybenzaldehyde instead of o-hydroxybenzaldehyde.

Affinity Labeling of Yeast RNA Polymerase B (II)—Yeast RNA polymerase B (II), prepared according to Sawadogo et al. (29) showed the subunit content previously described (30), with a B220 to B185 subunit ratio of about 50%. RNA polymerase B (II) was incubated for 30 min at 30 °C in 9 μl of a mixture containing 3 mM nucleotide derivative (2.5 mM in the case of the III, XI, and XIII analogues), 10 mM Hepes/NaOH, pH 7.9, 75 mM KC1, 1 mM dithio-
Hydroxylamine cleavage of proteins in gel pieces was done according to Saris et al. (32). Briefly, the labeled bands in the dried gel were excised after autoradiography. The gel pieces were swollen in 5% methanol, the paper backing was peeled off, and SDS was removed by washing 3 times for 10 min at room temperature in 5% methanol. Gel pieces were dried in a Speedvac concentrator (Savant) for 15 min at room temperature, submerged in 200 μl of either the reaction or the control mixture and incubated at 45°C for 3 h. The reaction mixture contained 2 M hydroxylamine hydrochloride (Fluka), 6 M guanidine hydrochloride (Pierce), 15 mM Tris (Sigma), 4.5 mM lithium hydroxide, pH 9.3. The control mixture contained no hydroxylamine. Gel pieces were then washed with three changes of 5% methanol, dried as before, and incubated in 50 μl of electrophoresis sample buffer at 100°C for 5 min. Gel pieces and buffer were applied to a polyacrylamide slab gel for analysis of protein fragments by electrophoresis.

**Sodium Dodecyl Sulfate Electrophoresis in Polyacrylamide Slab Gels**—Electrophoresis in polyacrylamide slab gels in the presence of sodium dodecyl sulfate were performed as described (33), or according to Fling and Gregerson (34) in the case of Fig. 5. Subunits of yeast RNA polymerases A (I) and B (II) (1) were used as molecular weight markers, as well as molecular weight standard kits (Pharmacia).

**RESULTS**

RNA polymerase B (II) was affinity labeled with 13 different nucleotide derivatives, as previously described (25). Briefly, the enzyme is first cross-linked to a nucleotide derivative and then labeled in a second step by allowing phosphodiester bond formation with [α-32P]UTP in the presence of DNA. This highly specific labeling procedure is based on the enzymatic competence of the enzyme. As described previously, using derivative IV (25), the B150 subunit was specifically labeled. As shown in Fig. 1, the efficiency of the labeling of B150 greatly depended on the nature of the nucleotide derivative, the highest yields being obtained with derivatives III, IV, VII, and VIII. Remarkably, in one case, with reagent IX, bands migrating at the level of the largest subunit B220 and its proteolyzed by-products were labeled as efficiently as B150. This labeling reaction was not further investigated. Derivatives III, IV, and VIII were selected for the subsequent mapping experiments.

Preliminary coarse mapping of the cross-linked product was performed by microsequencing a labeled subfragment of the B150 subunit. The most straightforward approach would be to microsequence the shortest possible labeled peptide after purification by high pressure liquid chromatography. However, this approach was excluded by the low efficiency of the labeling procedure (23), as only a small percentage of enzyme molecules catalyzed phosphodiester bond formation after modification with the nucleotide analogues. Indeed, the effi-

**Fig. 1. Affinity labeling of the B150 subunit of RNA polymerase B (II) with different derivatives.** Yeast RNA polymerase B (II) (2 μg) was affinity labeled with 13 distinct derivatives. The derivatives and the labeling procedure are described under “Materials and Methods” except that the RNase and DNase treatments were omitted. The labeled subunit was identified as B150 by autoradiography after separation of the enzyme subunits by electrophoresis in the presence of SDS through a 6–15% polyacrylamide gradient gel. Lanes 1–13 refer to derivatives I–XIII shown in Scheme 1. The subunit molecular weights are indicated (×10^3).
P-labeled fragments were revealed by autoradiography. Lane 1, B150 subunit was digested with S. aureus V8 protease as described under "Materials and Methods." Proteolytic fragments were separated by electrophoresis under denaturing conditions in an 11% polyacrylamide gel and the 58-, 43-, and 34-kDa bands are indicated. B, gel pieces corresponding to the 58-, 43-, and 34-kDa digestion products (lanes 1–3, respectively) were excised and further incubated with S. aureus V8 protease as above. Gel bands and buffer were directly applied on a second 10–20% polyacrylamide gel in the presence of SDS and the electrophoretic peptide patterns were revealed by autoradiography. Molecular weights of markers are indicated (×10^4).

The B150 subunit labeled with derivative IV was treated with limiting concentrations of different endoproteinases: Glu specific protease (S. aureus V8 protease), trypsin, or α-chymotrypsin (results not shown). A suitable peptide pattern was obtained with S. aureus V8 protease that gave reproducibly, and in good yield, a few labeled high molecular weight cleavage products in the presence of SDS. Four major products of 58, 43, 34, and 30 kD were detected (Fig. 2A). A similar pattern was obtained when RNA polymerase B (II) was labeled with derivatives III and VIII. Gel pieces corresponding to the 58-, 43-, and 34-kDa fragments labeled with derivative IV were excised and treated more extensively in situ with the same S. aureus V8 protease. The gel pieces were directly applied to a second polyacrylamide gel, and the protein fragments were separated by electrophoresis under denaturing conditions. As shown in Fig. 2B, the proteolytic patterns of the three fragments were similar; the 34-kDa fragment was generated by cleavage of both the 58- and 43-kDa fragments, and the 43-kDa fragment was a cleavage product of the 58-kDa one. These results suggest that the labeling sites are located in the 34-kDa fragment.

For microsequence analysis, the B150 subunit labeled with derivative IV was hydrolyzed with S. aureus V8 protease and the 43-kDa product was purified as above. At this point, it was verified that the radioactivity comigrated with the protein band on an SDS-polyacrylamide gel (result not shown). The N-terminal sequence of the 43-kDa fragment was HI-MMNQS. Analysis of the B150 subunit sequence (13, 35) indicated the presence of the sequence MI MNQS at position 839–844. Therefore, the 43-kDa polypeptide had been produced by cleavage at residue Glu840. The calculated molecular weight of the Asp837-Phe1224, 43,870, was in good agreement with the measured molecular weight. This result indicated that the affinity labeled site was located within the C-terminal part of the B150 subunit, somewhere between Asp837 and Phe1224 (see Fig. 6). Since the 43-kDa fragment is C-terminal, the 58- and 34-kDa polypeptides are probably generated by cleavage at Glu residues in the 710–720 region and at a cluster of 3 Glu residues at positions 922, 923, and 924, respectively.

The position of the label within the C-terminal region was further investigated by extensive CNBr cleavage of the whole B150 subunit, as well as of the 58- and 34-kDa fragments. Exhaustive CNBr cleavage of RNA polymerase B (II) affinity labeled with derivatives III or IV resulted in a major labeled peptide of 14 kDa and, to a minor extent, in a 15.6-kDa polypeptide, probably resulting from partial cleavage (Fig. 3A). When the RNA polymerase B (II) was labeled with derivative VIII, the 14-kDa peptide was present but was not the major product (Fig. 3A, lane 1). Lower molecular weight-labeled peptides were observed, which were also present in trace amounts in the case of derivative IV but not with derivative III. These results showed that the three derivatives affinity labeled the same site in the B150 subunit, but also suggested that a second region could be labeled with the longer derivatives IV and VIII. In the following experiments, we mapped the strong affinity labeled site observed with derivatives III and IV. A polypeptide of approximately 110–120 amino acids cross-linked to a dinucleotide could account for the observed molecular weight of 14,000. Within the C-terminal region, from Asp837 to Phe1224, only cleavage at Met885 and Met999 would generate a peptide of the correct size (114 amino acids); all the other possible CNBr peptides are smaller (M, < 6,000) (see Fig. 6). Extensive CNBr cleavage of the 58-kDa S. aureus V8 protease fragment labeled with derivative III generated the same major 14-kDa product (Fig. 3B, lane 1), while cleavage of the 34-kDa fragment yielded a fragment...
of smaller size, around 10-kDa (Fig. 3B, lane 2), suggesting that the label is located between Glu^923 and the nearest methionine residue, Met^996. The same results were found with the 58- and 34-kDa fragments labeled with derivative IV (results not shown). Furthermore, the labeled peptide pattern generated by limited CNBr cleavage of the 34-kDa fragment labeled with reagents III and IV was found to be very similar to that predicted for N-terminal single-hit peptides, indicating that the label was located between Glu^923-924 and Met^996 (results not shown). Remarkably, the Met^986-Met^990 fragment contains one of the four hydroxylamine cleavage sites present in the B150 subunit sequence, affinity labeled RNA polymerase B (II) with derivatives III or IV was subjected to a limited hydroxylamine treatment (Fig. 4). Size analysis of the single-hit cleavage products should provide the length of the N-terminal and C-terminal labeled peptides, and, by deduction, the position of the label relative to the Asn^946-Gly^947 cleavage site. Fig. 4 shows the pattern of the labeled bands obtained, in comparison with the three different patterns of bands that can be predicted depending on the position of the label. The doublet band of high molecular weight (≈120,000) was characteristic of the band patterns that would be predicted if the label were located either between Asn^946 and Asn^948 or between Asn^946 and Asn^1074. However, the presence of a labeled fragment of about 30 kDa unambiguously positioned the label in-between Asn^946 and Asn^1074. The observed pattern of bands exactly corresponded to the predicted one, except for two additional bands of 48.5 and 81 kDa (the upper band of the doublet), but these bands were also present in the control reaction (Fig. 4, lanes C). Note that the same band pattern was obtained when the B150 subunit was modified with derivatives III or IV (Fig. 4).

To obtain independent confirmation that the label was C-terminal relative to the Asn^946 hydroxylamine cleavage site, RNA polymerase B (II) was labeled with derivative IV and the three major products generated by S. aureus V8 protease (58-, 43-, and 34-kDa fragments) were subjected to partial hydroxylamine (2 M) cleavage for 3 h at 45°C. Cleavage products were separated by electrophoresis in the presence of SDS through a 10-25% polyacrylamide gradient gel. Control reactions were carried out under the same conditions without hydroxylamine. The autoradiogram of the gel is shown with the size of the main cleavage products. A schematic diagram of the 58-, 43-, and 34-kDa fragments showing the position of the hydroxylamine cleavage sites (NG) is presented with two possible locations of the label (between Glu^923-924 and Asn^946 or between Asn^946 and Asn^1074) indicated by different symbols. Depending on label location, a different pattern of radioactive peptides is predicted (left).
polymerase B (II). The B150 subunit was previously identified by affinity labeling as interacting with the initiator nucleotide (25). By the same approach and with a set of 13 distinct derivatives, we demonstrate that labeling of yeast B (II) enzyme is best achieved when using nucleotide analogues with an aldehyde reactive function. These derivatives are presumed to react essentially with lysyl residues (23) that are in close proximity to the active site. Furthermore, labeling efficiency was shown to depend critically on the position of the aldehyde function and on the length of the derivative. For example, derivatives of ATP with the aldehyde function in the ortho position on the benzene ring labeled the enzyme weakly as compared to derivatives with the aldehyde function in the para position (see Fig. 1, lanes 1 and 3). Furthermore, when comparing ortho (or para) derivatives, the labeling was dependent on the length of the nucleotide analogues. With para derivatives, the strongest labeling occurred with the ATP analogue, while the AMP analogue was the most efficient with ortho derivatives. A possible explanation is that ortho and para derivatives labeled different lysyl residues within the active site. Labeling of larger bands, most probably corresponding to the largest subunit, by the imidazolide derivative of ATP (IX) was more surprising. The possibility of enzyme labeling via the synthesis of a short RNA chain was not excluded in this case. It was shown previously that nascent RNA chains can be linked to the two large subunits (25). The labeling was specific to the ATP derivative since imidazolide analogues of ADP (X) or AMP (XI) only weakly labeled the B150 subunit (see Fig. 1).

Localization of the active site in the bacterial β subunit showed that the initiator nucleotide was cross-linked to two different regions, depending on the size of the affinity reagent (26). Therefore, three derivatives of various lengths that efficiently label the B150 subunit were selected for subsequent mapping of the labeled position. In this paper, we demonstrate that in each case the initiator nucleotide was cross-linked to the C-terminal part of the B150 subunit, between Asn962 and Met979. This region represents a high affinity labeled site with derivatives III and IV. Coarse mapping of the cross-link site was achieved by microsequencing a 43-kDa labeled subfragment of the B150 subunit generated by hydrolysis of the subunit with S. aureus V8 protease. The 43-kDa fragment was found to be C-terminal and to result from cleavage at Glu936. The position of the label was further restricted to the Met985-Met999 fragment by extensive CNBr cleavage of the B150 subunit. This localization was confirmed by hydroxylamine cleavage of the Met962-Met989 fragment, which contains one of the four hydroxylamine cleavage sites present within the B150 subunit. Finally, size analysis of single-hit hydroxylamine cleavage products of the labeled B150 subunit placed the label between Asn962 and Asn974. Altogether, these results mapped the position of the cross-linked nascent RNA chain to between Asn962 and Met989. Additional evidence obtained from CNBr and hydroxylamine cleavage of S. aureus V8 fragments confirmed this mapping.

Between Asn962 and Met989, 5 lysyl residues at positions 962, 965, 972, 979, and 987 could be the target of the affinity reagents (Fig. 7). Whether the same lysyl residue reacts with the three different nucleotide analogues is not known. In the vicinity of these residues, no specific chemical cleavage sites are available that could be used to further identify the modified lysine(s). Interestingly, the labeled region is included in one of the nine domains (H) conserved between the yeast B150 subunit and the bacterial β subunit (13). This region is also conserved in the second largest subunit of Drosophila RNA polymerase B (14), in the B or B' subunit of RNA polymerase from different archaeabacteria (16, 18), and in the coding region of the chloroplast RNA polymerase from tobacco (36) and of the Killer plasmid pGKL2 from Kluiveromyces lactis (37) (Fig. 7). Interestingly, when using nucleotide derivatives III and IV, the cognate RNA polymerase subunits from E. coli (26), Sulfolobus acidocaldarius (27), and Methan-
polymerase is involved in binding the nascent RNA chain and in the activation of the enzymatic activity of the RNA polymerase, which is implicated in the active site, found in the most conserved part of domain H (Fig. 7). Nevertheless, it is unlikely that lys^57^ is the targeting label, since the corresponding lysyl residue of E. coli β subunit (at position 1073) is not labeled (26). The lysyl residue at position 979 is therefore the most likely target of the affinity reagents. Nevertheless, the labeling of lys^6^, lys^5^, or lys^7^ cannot be excluded, especially the lys^57^ that is conserved in all but one sequence (the tobacco chloroplastic ORF 1070). The Asn^946^-Met^959^ sequence does not include a nucleotide-binding consensus sequence of the type G-X-X-X-G-K(T) (where X is any amino acid) common to purine nucleotide-binding or processing proteins (39). A related sequence, G EDVI CI GKT, was found in the B150 subunit at position 907–915 in homologous region G and was proposed as the nucleotide-binding site (13). Our results do not support this hypothesis, although the G^907^-K^915^ domain may be proximal to the Asn^946^-Met^959^ site in the tertiary structure and possibly contribute to substrate binding.

Previous observations with affinity labeled E. coli RNA polymerase indicated that, depending on the length of the nucleotide analogues, two different regions could be labeled. The second domain found in the bacterial subunits and labeled by longer derivatives was located in the last C-terminal conserved domain I. This domain is conserved in eukaryotic enzymes, especially with derivative VIII and, to a lower extent, with derivative IX. This difference between the three derivatives is more evident with derivative IV. This difference between the three derivatives could be correlated with their size: the greater the length of the derivative, the more was the second domain labeled. Whether this second region is homologous to the E. coli C-terminal labeled domain was not investigated.

The C-terminal part of the B150 subunit contains a sequence of the form C-X^2^-C-X^5^-C-X^2^-C that has been implicated in coordinating zinc in nucleic acid-binding proteins (40). Several lines of evidence suggest that, together with the other subunits, the second large subunit of eukaryotic RNA polymerase is involved in binding the nascent RNA chain (21) and DNA (22). The close proximity of the nucleotide-binding site with this potential zinc-binding domain may be functionally significant. Site-directed mutagenesis in the domain identified in this study will provide information on the amino acids implicated in the active site, taken in a broad sense.

Acknowledgments—We are grateful to Prof. G. Hartmann, A. Schaffner, and S. Glaser (Ludwig-Maximilians-Universität, München) for helpful discussions, for communicating results prior to publication, and for the gift of derivative IV, to P. Parouteau (Applied Biosystems, France) for high pressure liquid chromatography purification, and to J.-C. Huot (Institut National de la Recherche Agronomique, Versailles) for microsequencing facilities. We thank R. Swanson for a critical reading of the manuscript and the referee for helpful comments.

REFERENCES

6. Lys^5^, Lys^6^, or Lys^7^ cannot be excluded, especially the lys^57^ that is conserved in all but one sequence (the tobacco chloroplastic ORF 1070). The Asn^946^-Met^959^ sequence does not include a nucleotide-binding consensus sequence of the type G-X-X-X-G-K(T) (where X is any amino acid) common to purine nucleotide-binding or processing proteins (39). A related sequence, G EDVI CI GKT, was found in the B150 subunit at position 907–915 in homologous region G and was proposed as the nucleotide-binding site (13). Our results do not support this hypothesis, although the G^907^-K^915^ domain may be proximal to the Asn^946^-Met^959^ site in the tertiary structure and possibly contribute to substrate binding.

Previous observations with affinity labeled E. coli RNA polymerase indicated that, depending on the length of the nucleotide analogues, two different regions could be labeled. The second domain found in the bacterial β subunit and labeled by longer derivatives was located in the last C-terminal conserved domain I. This domain is conserved in eukaryotic enzymes, especially with derivative VIII and, to a lower extent, with derivative IX. This difference between the three derivatives could be correlated with their size: the greater the length of the derivative, the more was the second domain labeled. Whether this second region is homologous to the E. coli C-terminal labeled domain was not investigated.

The C-terminal part of the B150 subunit contains a sequence of the form C-X^2^-C-X^5^-C-X^2^-C that has been implicated in coordinating zinc in nucleic acid-binding proteins (40). Several lines of evidence suggest that, together with the other subunits, the second large subunit of eukaryotic RNA polymerase is involved in binding the nascent RNA chain (21) and DNA (22). The close proximity of the nucleotide-binding site with this potential zinc-binding domain may be functionally significant. Site-directed mutagenesis in the domain identified in this study will provide information on the amino acids implicated in the active site, taken in a broad sense.

Acknowledgments—We are grateful to Prof. G. Hartmann, A. Schaffner, and S. Glaser (Ludwig-Maximilians-Universität, München) for helpful discussions, for communicating results prior to publication, and for the gift of derivative IV, to P. Parouteau (Applied Biosystems, France) for high pressure liquid chromatography purification, and to J.-C. Huot (Institut National de la Recherche Agronomique, Versailles) for microsequencing facilities. We thank R. Swanson for a critical reading of the manuscript and the referee for helpful comments.

REFERENCES

Mapping the active site of yeast RNA polymerase B (II).
M Riva, C Carles, A Sentenac, M A Grachev, A A Mustaev and E F Zaychikov


Access the most updated version of this article at http://www.jbc.org/content/265/27/16498

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/265/27/16498.full.html#ref-list-1