Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 4151–4155, May 1993
Immunology

Recombination sequence-binding protein in thymocytes undergoing T-cell receptor gene rearrangement

KATHRIN MUEGGE, MICHAEL WEST, AND SCOTT K. DURUM

Biological Carcinogenesis and Development Program, Program Resources, Inc./Dyncorp, National Cancer Institute-Frederick Cancer Research and Development Center, Frederick, MD 21702-1201 and Laboratory of Molecular Immunoregulation, Biological Response Modifiers Program, National Cancer Institute, Frederick, MD 21702-1201

Communicated by Martin Gellert, February 17, 1993

ABSTRACT **Rearrangement of T-cell antigen receptor and immunoglobulin genes occurs in immature lymphoid cells by an unknown mechanism. To identify components of the rearrangement machinery, we isolated a population of murine thymocytes enriched for rearranging pre-T cells. In the nuclear fraction of these cells, we detected a protein that specifically bound the recombination sequences that flank T-cell receptor and immunoglobulin genes and are required for their rearrangement. This protein recognized both heptamer and nonamer motifs of the recombination sequence, separated by either 12 or 23 bp. The protein complexed with the recombination sequence oligonucleotide had an apparent molecular mass of 30 kDa. The binding characteristics of the protein and its presence in rearranging thymocytes and cell lines suggest that it could serve as the recognition unit of a recombinase complex.**

Gene rearrangement is a unique ability of lymphoid cells to enlarge their gene repertoire, creating the diversity of their immune recognition molecules which is necessary to confront the variability in the environment. T and B cells perform the rearrangement of their T-cell antigen receptor (TCR) and immunoglobulin genes early in their development, T cells primarily in thymus, B cells in bone marrow (reviewed in ref. 1). The actual process of gene rearrangement is unknown, the hypotheses about it being drawn from the structure of genes before and after rearrangement. A short motif, sometimes termed the "recombination sequence," flanks the rearranging genes and has been shown to be necessary and sufficient for rearrangement of exogenous plasmid substrates transfected into cell lines. The recombination sequence consists of highly conserved heptamer and nonamer motifs that are separated by 12 or 23 bp of spacer DNA (H-12-N or H-23-N). The recombination sequence could serve as a binding site for recombinase machinery, since cutting of DNA during rearrangement occurs precisely adjacent to the heptamer motif, and point mutations in the heptamer motif abrogate rearrangement (2, 3). In the following study the recombination sequence was used to identify a nuclear factor in rearranging thymocytes.

MATERIALS AND METHODS

Thymocytes from Radiation Chimeras. To enrich for a population of immature thymocytes undergoing rearrangement of TCR genes, a procedure using radiation chimeras was followed. SJL/J recipient mice, which lack $V_{\beta 8}$ variable-region genes were irradiated (950 rads; 1 rad = 0.01 Gy). Mice were reconstituted with liver cells (5×10^6) from C57BL/6 embryos at day 14 of gestation. C57BL/6 cells, which bear the $V_{\beta 8}$ genes, were analyzed in the recipient thymus at various times by using the polymerase chain reaction (PCR); thymus was also used to prepare DNA-binding proteins.

Detection of TCR β-Chain Gene Rearrangement. We used a PCR to determine the time point at which, following reconstitution, thymocytes were undergoing rearrangement of the TCR β-chain gene. Thymus DNA was extracted as described (4). Rearrangement of the β-chain variable (V_β) gene segment to the diversity (D_β) segment was detected by using a primer for the 5′ end of $V_{\beta 8}$ (5′-GAGGCTGCAGTCACCCAAAGTCCAA-3′) and a $J_{\beta 2.1}$ joining-region primer (5′-TGAGTCGTGTTCCTGGTCCGAAGAA-3′). After 30 PCR cycles, the product was electrophoresed through a 1.2% agarose gel, blotted to a Nytran membrane (Schleicher & Schuell), and hybridized with a 90-bp probe to the midregion of $V_{\beta 8}$. A control pair of primers was used which generated a PCR product from the $V_{\beta 8}$ gene in either germ-line or rearranged form and consisted of the same $V_{\beta 8}$ 5′ primer as above, together with a $V_{\beta 8}$ 3′ primer (5′-ACAGAAATATACAGCTGTCTGAGAA-3′).

Analysis of DNA-Binding Proteins. Nuclear proteins were extracted from thymuses and other tissues and cell lines, and DNA-binding proteins were then analyzed by a mobility-shift assay (5). Briefly, nuclear extracts were prepared as described (6) at a concentration of 4×10^8 cell equivalents per ml. The binding reaction was conducted by mixing 2 μl of nuclear extract with radiolabeled double-stranded oligonucleotides (0.2 ng) in 25 μl of buffer {25 mM Hepes, pH 7.5/10% (vol/vol) glycerol/2 μg of poly[d(I-C)]/50 mM NaCl/0.05% Nonidet P-40/1 mM dithiothreitol}. The mixture was incubated for 25 min at room temperature, electrophoresed in a 6% polyacrylamide gel, and visualized by autoradiography. For competition experiments, various amounts of unlabeled oligonucleotide were preincubated for 20 min at room temperature prior to addition of labeled oligonucleotide. For the experiments shown in Figs. 3–5, a different reconstituted thymus pool was used in each panel. The oligonucleotides, consisting of various motifs from recombinase sites in TCR β genes, various mutants, and unrelated controls, are shown in Table 1.

To estimate the molecular weight of DNA-binding proteins (4), they were crosslinked to the labeled oligonucleotide by irradiation with ultraviolet light (250 nm) for 15 min on ice. Complexes were separated by SDS/10% PAGE and visualized by autoradiography.

RESULTS AND DISCUSSION

The transformed pre-B- and pre-T-cell lines which have been reported to rearrange their endogenous immunoglobulin and TCR genes do so at a very low frequency. If we could capture precursor T cells in transition from germ-line to rearranged TCR gene, we would have a 100% recombination-active population. Such cells in transition are rare even among

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

Abbreviation: TCR, T-cell antigen receptor.

Table 1. Sequences of double-stranded oligodeoxynucleotides used in mobility-shift assays

Designation	Source	Sequence
N-12-H	$D_{\beta 2}$	AAGAAACTTTTTTGTATCACGATGTAACATTGTGGGGACT
N-12′-H	$D_{\beta 1}$	AGGGTCCTTTTTTGTATAAAGCTGTAACATTGTGGGGACA
N-12-H′	D_{β}/Igκ	AAGAAACTTTTTTGTATCACGATGTAACACTGTGGGGACT
N′-12-H	D_{β}/Igκ	AAGAAAGGTTTTTGTATCACGATGTAACATTGTGGGGACT
N″-12″-H′	Igκ	TTACAGTGTTTTTGTTCCAGTCTGTAGCACTGTGTGAATG
N-12-H^{mut}		AAGAAACTTTTTTGTATCACGATGTAACAATCTGGGGACT
N-23-H		AAGAAACTTTTTTGTATCACGATGTAAATCACGATGTACATTGTGGGGACT
N′-23-H′	Igκ	GATCCGGTTTTTGTACAGCCAGACAGTGGAGTACTACCACTGTGG
N-12-( )		AAGAAACTTTTTTGTATCACGATGTAA
( )-12-H		ATCACGATGTAACATTGTGGGGACT
N-10-H		AAGAAACTTTTTTGTATCGATGTAACATTGTGGGGACT
AP-1		TAAAGCATGAGTCAGACACCTCTG
cAMP		CTTGGCTGACGTCAGAGAG
NFκB		CAGAGGGGACTTTCCGAGAGGC
NF-IL6		ATGCTAAAGGACGTCACATTGCACAATCT

Nonamer and heptamer motifs are underlined.

immature ($CD3^-$, $CD4^-$, $CD8^-$) thymocytes, most having already rearranged their TCR β genes (K.M. and S.K.D., unpublished work). Embryonic thymocytes, which begin to rearrange their TCR β genes at day 15 of gestation in the mouse, provide very little material for study.

In an effort to obtain useful quantities of T cells rearranging their TCR genes, we produced synchronously developing thymocytes in radiation chimeras after reconstitution with stem cells and determined when the first rearrangement of the TCR β gene occurred. The irradiated SJL/J host strain lacked the $V_{\beta 8}$ gene, thus allowing the detection of $V_{\beta 8}$ rearrangement of donor-derived stem cells by a PCR technique. As shown in Fig. 1, V–D–J rearrangement of the TCR β locus occurred on days 10–12 after reconstitution. This represents an ≈5000-fold increase in rearranged DNA (based on titration compared with positive and negative controls; data not shown), which is a far greater increase than cell division could achieve in 48 hr. The percentage of cells having undergone rearrangement during this period may represent about 20%; this is based on titrating DNA and finding 20% of rearranged product on day 12 compared with whole adult thymus (data not shown). This population therefore represents pre-T cells captured in the transition from germ line to rearranged, and several million thymocytes per mouse could be obtained at day 12 for study.

Nuclear extracts were prepared from these pre-T cells at day 12 and analyzed for factors that could participate in the rearrangement of TCR genes by binding to the recombination sequence from $D_{\beta 2}$ (N-12-H). As shown in Fig. 2, a unique complex (termed RP, for recognition protein) was found in five nuclear extracts produced from independent batches each containing thymuses from 20 radiation chimeras. This assay is quite sensitive, and the actual number of RP molecules detected per cell is about 100 on the basis of densitometer scanning. Note that a larger complex ("ubiquitous") is not unique to the rearranging cells and is seen in all cell lines and tissues shown. By comparison to reconstituted thymus, other unfractionated thymus preparations showed much less RP, with a small amount being seen in the preparation of adult thymus, and none in newborn thymus. Both of the latter consist primarily of thymocytes which have already undergone TCR β rearrangement; thus RP is most plentiful in reconstituted thymus, which is enriched for immature T cells rearranging the TCR β-chain gene.

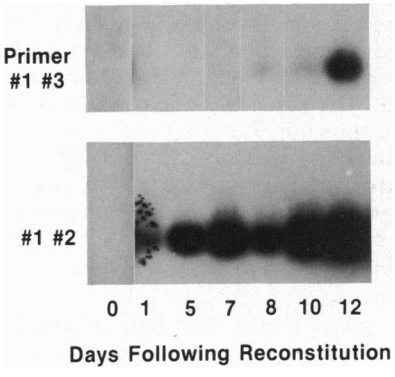

FIG. 1. Rearrangement of the TCR β-chain gene during thymic reconstitution. To produce a population of synchronized rearranging T cells, radiation chimeras were produced. SJL/J mice, which lack $V_{\beta 8}$, were irradiated and then received an intravenous injection of fetal liver cells from C57BL/6 embryos. Donor cells, which bear the $V_{\beta 8}$ genes, were detected in the recipient thymus by a PCR detecting either unrearranged $V_{\beta 8}$ (primers 1 and 2) or rearranged $V_{\beta 8}$–$D_{\beta 2}$ (primers 1 and 3).

We also detected RP in five cell lines (M1, PD31, 38B9, and 18-81, Fig. 2; NFS-70, data not shown) reported to undergo rearrangement of transfected genes (7), but not in the pre-B-cell line 70Z3, which was reported not to rearrange immunoglobulin genes *in vitro* (7). RP was not detectable in the other lines and normal tissue analyzed in Fig. 2, nor in the Ana-1 mouse macrophage line and HepG2 human hepatoma line (data not shown). Thus, the presence of RP corresponded to the ability of a cell to rearrange immune genes.

The binding specificity of RP was analyzed by using various oligonucleotides for competition or direct binding. Specificity for N-12-H was verified by competition with the same unlabeled sequence: 25 ng (≈100:1) gave complete inhibition (Fig. 3*A*). Unrelated motifs, such as the AP-1 or NF-κB recognition site, did not compete for RP binding (Fig. 3*B*).

Mutation of the heptamer motif inhibited binding of RP, thus indicating the specificity of RP binding to its sequence (Fig. 4*A*). This mutation of the heptamer has been shown to inhibit rearrangement of transfected plasmid substrates (2, 3). Also, a reduction in the spacer length to 10 bp has been reported to inhibit rearrangement efficiency (3) and was found to reduce binding of RP to the recombination sequence (Fig. 4*A*). The ubiquitous band, it should be noted, has a binding pattern distinct from that of RP, in that the former binds the point mutant N-12-H^{mut} and the construct with the 10-bp spacer (this makes it less likely that it is, for example, a precursor of RP). Thus, the binding characteristics of RP strongly correlated with known recombination rules.

RP recognized both heptamer and nonamer motifs simultaneously, since either motif competed with the whole N-12-H sequence (Fig. 3 *C* and *D*) (the full sequence competed better than the heptamer, which competed better than

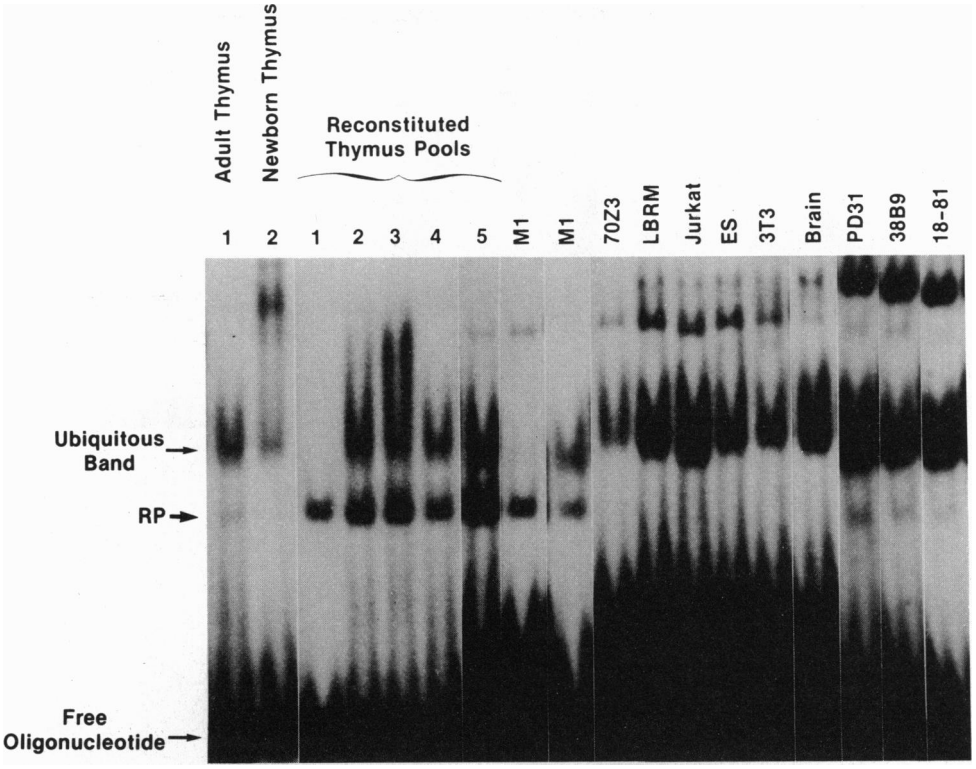

FIG. 2. Detection of a recognition protein (RP) that binds the heptamer–nonamer motif. Rearranging thymocytes and other cells were analyzed for nuclear factors that bound to the N-12-H sequence of the $D_{\beta 2}$ gene. Normal thymus preparations were from 8-week-old or newborn C57BL/6 mice. Five reconstituted thymus pools were prepared from chimeras as in Fig. 1, 12 days after reconstitution. Preparations are shown from various cell lines (M1 is a rearranging myeloblastoid leukemia, 70Z3 is a nonrearranging pre-B leukemia, LBRM and Jurkat are T lymphomas, ES is the embryonic stem cell line D3, 3T3 is a mouse fibroblast line, and PD31, 38B9, and 18-81 are rearranging pre-B leukemias) and normal brain tissue. Nuclear proteins were extracted and analyzed for binding to the radiolabeled H-12-N oligonucleotide.

the nonamer). However, the affinity of RP for heptamer or nonamer alone was not high enough to observe direct binding: both heptamer and nonamer had to be present (Fig. 4*B*) (although we have visualized direct binding to nonamer when the interaction was stabilized by ultraviolet crosslinking; data not shown). Another nuclear protein, SAP1 in yeast, shows a precedent for binding two sites separated by, in this case, a 13-bp spacer; neither of the two sites alone binds strongly to SAP1, which has the remarkable property of inducing a double-stranded DNA break involved in mating-type switching (8). The ability of RP to simultaneously recognize heptamer and nonamer motifs again conforms to the recombination rule that both heptamer and nonamer are necessary for rearrangement.

RP also bound slightly different recombination motifs. Fig. 3*C* shows Competition with N-12′-H from TCR $D_{\beta 1}$ and N′-23-H′ sequences from the immunoglobulin κ light-chain gene. Fig. 4*C* shows RP bound directly to three different sequences from TCR β: a 23-bp spacer (N-23-H) and a different consensus heptamer and consensus nonamer (N-12-H′, N′-12-H). RP also bound to sequences from Igκ with either a 23-bp spacer (N′-23-H′, shown) or a 12-bp spacer (N″-12″-H′; data not shown).

Finally, the size of RP was estimated by ultraviolet crosslinking. The RP–nucleotide complex migrated at 30 kDa (Fig. 5). The small size suggests that RP might serve as the recognition unit of a larger recombinase complex. Not shown is the sensitivity of RP to proteinase K, indicating the peptide nature of RP.

The recombinase machinery is undefined as yet; several candidates have been proposed from earlier studies, although none appear to correspond to RP. Mice with severe combined immunodeficiency (SCID) have a defect in a gene (*scid*) controlling repair of rearranging genes (9) as well as other, nonrearranging genes (10); thus the *scid* gene would not be expected to encode a protein such as RP, which recognizes recombination sequences, and in fact we observed RP in thymocytes from SCID mice (data not shown).

RAG-1 and RAG-2 were identified as genes which, upon transfection, induced rearrangement of cotransfected constructs (reviewed in ref. 11). Both genes have been shown to be necessary for rearrangement and normal lymphogenesis, by gene targeting in mice (12, 13). However, there are no data suggesting a direct participation in the recombinase machinery (binding to the recombination sequence, endonuclease activity, etc.). Also, transfection of fibroblasts with both genes results in only modest recombination activity (<1% in stable transfectants). This suggests that the RAG genes, although necessary (based on gene targeting), may require other lymphoid components to achieve optimal rearrangement. Perhaps the RAG proteins regulate the activity of RP or other members of the recombinase complex.

RAG-1 expression persists through the "double-positive" ($CD4^+CD8^+$) stage of thymopoiesis (14), whereas RP is much more prominent in immature thymocytes. The double-negative ($CD4^-CD8^-$) stage is the time when rearrangements of TCR γ, δ, β, and α genes all occur, and for the first three genes it is complete then. The α-chain gene, however, has been suggested to undergo successive rearrangements (of unknown frequency) during the double-positive stage (15), which parallels RAG-1 expression, but this observation suggests that little RP would be required.

Recently a protein (RBP-J_κ) was cloned from a pre-B cell line and shown to recognize the H-23-N form of the recognition sequence (6, 16). No functional evidence that this protein participates in rearrangement has appeared yet. A very curious finding is that a gene similar to RBP-J_κ is present in *Drosophila melanogaster*, a species lacking immunoglob-

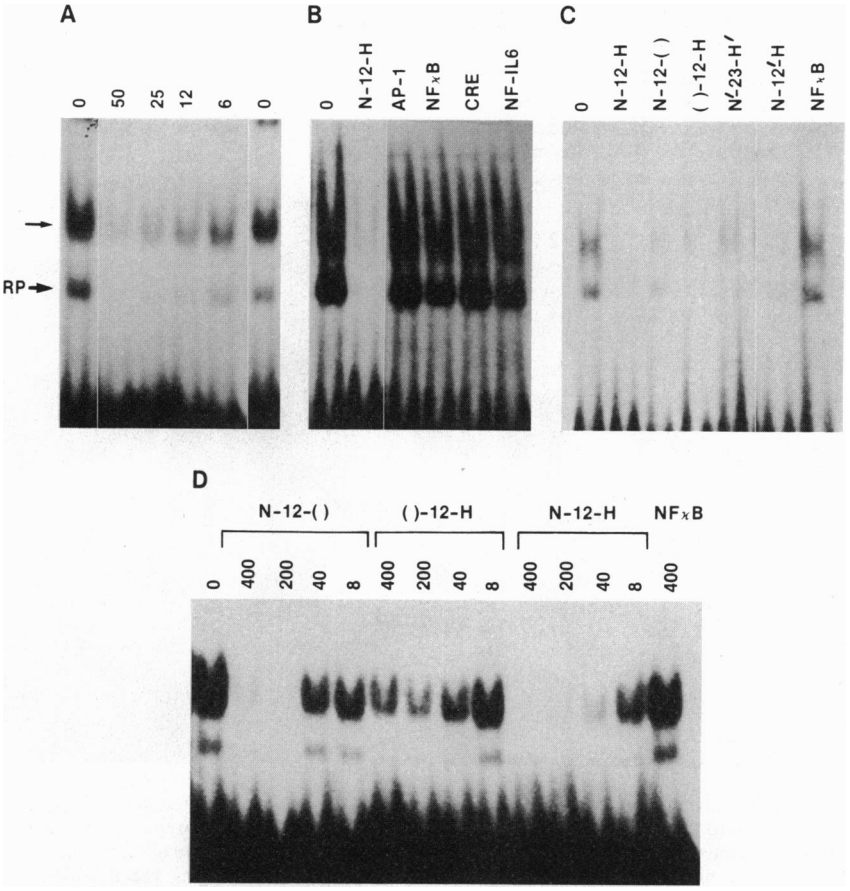

FIG. 3. Binding specificity of RP. The specificity of RP binding to the N-12-H sequence of $D_{\beta 2}$ was tested by comparing various sequences for competition with N-12-H. Approximately 0.25 ng of labeled probe was used per lane. The following unlabeled oligonucleotides were incubated with the nuclear proteins before addition of probe: N-12-H in the indicated fold excess (*A*); unrelated motifs (200-fold excess) (*B*); mutants (200-fold excess) (*C*); deletion mutants in the indicated fold excess (*D*).

ulin and TCR genes (17). The fly protein has the ability to bind murine recombination motifs (17); this suggests either that it may drive gene rearrangements in *Drosophila* or that it is not really involved in rearrangements but has some other function. Our RP protein appears to differ from RBP-J_κ by several criteria: size (<30 kDa vs. 60 kDa), binding specificity (binds H-12-N or H-23-N, whereas RBP-J_κ binds only H-23-N), and pattern of expression (unlike RP, RBP-J_κ is expressed in mature T-cell lines such as Jurkat).

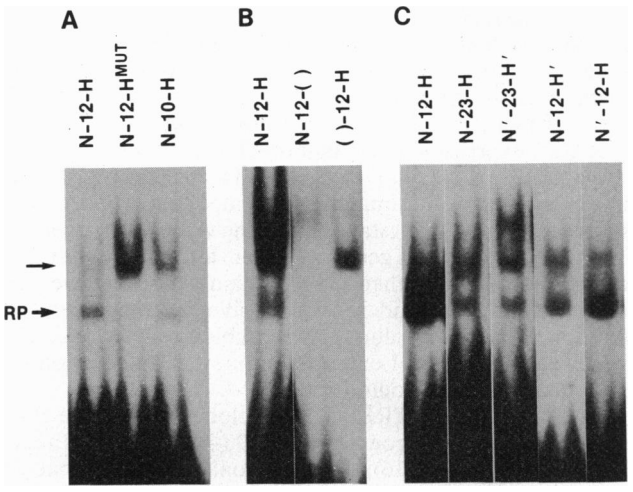

FIG. 4. Direct binding of RP to labeled oligonucleotides. Wild type N-12-H was compared with the indicated sequences.

Another protein, "nonamer-binding protein," NBP, has been isolated from calf thymus and proposed to participate in rearrangement (18, 19). NBP recognized only the nonamer

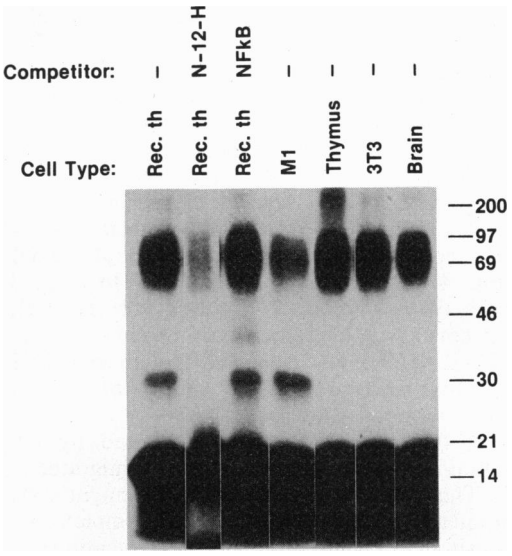

FIG. 5. Molecular size of RP estimated by ultraviolet crosslinking and SDS/PAGE. Nuclear extracts were mixed with labeled N-12-H in the absence or presence of the indicated competitor (200-fold excess). After ultraviolet crosslinking and SDS/PAGE; the RP–oligonucleotide complex migrated at 30 kDa. Rec.th, reconstituted thymus.

and not the heptamer motif, which would distinguish it from our RP, which binds both. NBP was also determined to be considerably larger than our RP (63 kDa vs. <30 kDa). NBP was isolated from whole thymus, mostly consisting of cells that have already rearranged, whereas our RP is from an enriched population of rearranging cells and is not abundant in whole thymus. Another protein (115 kDa) that bound a recombination sequence was extracted from a pre-B-cell line (20); it was visualized by Southwestern blotting and was rather abundant, but its binding specificity not determined.

Thus RP is a good candidate for a component of the recombination machinery, which, once it is defined, should have recognition and enzymatic activities that are interesting and useful. Moreover, knowledge of the recombination machinery might improve our understanding of leukemogenesis, in which translocations due to aberrant recombinase function may be a primary cause (reviewed in ref. 21).

We thank Rodney Wiles and Marta Vila for technical assistance; Drs. Kenneth Marcu, Ulrike Csaikl, and Steven Baur for providing cell lines; Dr. Leo Lee for preparation of oligonucleotides, Dr. Peter Johnson for helpful discussions, and Drs. Howard Young, Louis Matis, Joost Oppenheim, and Dan Longo for comments on the manuscript. This project was funded in part with funds from the Department of Health and Human Services under Contract N01-CO-74102 with Program Resources, Inc.

1. Lewis, S. & Gellert, M. (1989) *Cell* **59,** 585–588.
2. Akira, S., Okazaki, K. & Sakano, H. (1987) *Science* **238,** 1134–1138.
3. Hesse, J. E., Lieber, M. R., Mizuuchi, K. & Gellert, M. (1989) *Genes Dev.* **3,** 1053–1061.
4. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K., eds. (1987) in *Current Protocols in Molecular Biology* (Greene/Wiley-Interscience, New York), Sec. 10-2, 12-5.
5. Muegge, K., Williams, T. M., Kant, J., Karin, M., Chiu, R., Schmidt, A., Siebenlist, U., Young, H. A. & Durum, S. K. (1989) *Science* **246,** 249–251.
6. Parker, C. S. & Topol, J. (1984) *Cell* **36,** 357–369.
7. Lieber, M. R., Hesse, J. E., Mizuuchi, K. & Gellert, M. (1987) *Genes Dev.* **1,** 751–761.
8. Arcangioli, B. & Klar, A. J. S. (1991) *EMBO J.* **10,** 3025–3032.
9. Bosma, M. H. & Carroll, A. M. (1991) *Annu. Rev. Immunol.* **9,** 323–350.
10. Fulop, G. M. & Phillips, R. A. (1990) *Nature (London)* **347,** 479–482.
11. Oettinger, M. A., Schatz, D. G., Gorka, C. & Baltimore, D. (1990) *Science* **248,** 1517–1523.
12. Mombaerts, P., Iacomini, J., Johnson, R. S., Herrup, K., Tonegawa, S. & Papaioannou, V. E. (1992) *Cell* **68,** 869–877.
13. Shinkai, Y., Rathbun, G., Lam, K., Oltz, E., Stewart, V., Mendelsohn, M., Charron, J., Datta, M., Young, F., Stall, A. M. & Alt, F. W. (1992) *Cell* **68,** 855–867.
14. Guy-Grand, D., Vanden Broecke, C., Briottet, C., Malassis-Seris, M., Selz, F. & Vassalli, P. (1992) *Eur. J. Immunol.* **22,** 505–510.
15. Borgulya, P., Kishi, H., Uematsu, Y. & von Boehmer, H. (1992) *Cell* **69,** 529–537.
16. Matsunami, N., Hamaguchi, Y., Yamamoto, Y., Kuze, K., Kangawa, K., Matsuo, H., Kawaichi, M. & Honjo, T. (1989) *Nature (London)* **342,** 934–937.
17. Furukawa, T., Kawaichi, M., Matsunami, N., Ryo, H., Nishida, Y. & Honjo, T. (1991) *J. Biol. Chem.* **266,** 23334–23340.
18. Halligan, B. D. & Desiderio, S. B. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 7019–7023.
19. Li, M., Morzycka-Wroblewska, E. & Desiderio, S. V. (1989) *Genes Dev.* **3,** 1801–1813.
20. Miyake, S., Sugiyama, H., Tani, Y., Fukuda, T. & Kishimoto, S. (1990) *J. Immunogenet.* **17,** 67–75.
21. Tycko, B. & Sklar, J. (1990) *Cancer Cells* **2,** 1–8.