Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 10979–10983, November 1992
Biochemistry

A conserved double-stranded RNA-binding domain

DANIEL ST JOHNSTON*, NICHOLAS H. BROWN*, JOSEPH G. GALL†‡, AND MICHAEL JANTSCH†

*Wellcome/Cancer Research Campaign Institute, Tennis Court Road, Cambridge CB2 1QR, United Kingdom; and †Department of Embryology, Carnegie Institution, Baltimore, MD 21210

Contributed by Joseph G. Gall, August 31, 1992

ABSTRACT **We have identified a double-stranded (ds)RNA-binding domain in each of two proteins: the product of the *Drosophila* gene staufen, which is required for the localization of maternal mRNAs, and a protein of unknown function, Xlrbpa, from *Xenopus*. The amino acid sequences of the binding domains are similar to each other and to additional domains in each protein. Database searches identified similar domains in several other proteins known or thought to bind dsRNA, including human dsRNA-activated inhibitor (DAI), human trans-activating region (TAR)-binding protein, and *Escherichia coli* RNase III. By analyzing in detail one domain in staufen and one in Xlrbpa, we delimited the minimal region that binds dsRNA. On the basis of the binding studies and computer analysis, we have derived a consensus sequence that defines a 65- to 68-amino acid dsRNA-binding domain.**

The identification of regions of amino acid sequence similarity between proteins has led to the characterization of a number of conserved motifs, which are often found in proteins that are otherwise unrelated. In many cases, these motifs define structural domains that confer particular biochemical properties on the proteins in which they occur (discussed in ref. 1). Only a few protein motifs that interact with RNA have been identified, most notably the zinc finger, first characterized in TFIIIA (2), and the RNA recognition motif or RRM (3–5). Here we identify by database searches a consensus sequence of 65–68 amino acids that is found in a diverse group of double-stranded (ds)RNA-binding proteins. Sequence similarities between some of these proteins have been noted before, and parts of the binding domain have also been identified by filter binding assays (6–9). We have used deletion analysis to show that in two cases the minimal dsRNA-binding domain corresponds essentially to the 65- to 68-amino acid consensus sequence.

MATERIALS AND METHODS

Clones. We used the cDNA clone E10 of the staufen gene from *Drosophila melanogaster* (ref. 10; accession no. M69111), subcloned between the *Hin*dIII and *Not* I sites of Bluescript KS (Stratagene). The cDNA clone Xlrbpa was isolated from a *Xenopus laevis* ovary library (11) by Northwestern screening with a mixture of ^{32}P-labeled U1 and U2 small nuclear RNAs (snRNAs). The screen was performed according to Vinson *et al.* (12), except that 8 M urea replaced 6 M guanidine hydrochloride as the chaotropic agent. The binding buffer contained 50 mM NaCl, 10 mM $MgCl_2$, 10 mM Hepes (pH 7.5), 0.1 mM EDTA, and 1 mM dithiothreitol. The probe was a mixture of ^{32}P-labeled transcripts produced by T7 polymerase from U1 and U2 snRNA gene clones (13). The Xlrbpa insert was sequenced after conversion of the λZAP bacteriophage into the corresponding Bluescript phagemid. The sequence is available from GenBank under accession no. M96370.

Fusion Proteins. Regions of the staufen and Xlrbpa cDNAs were amplified by the polymerase chain reaction (PCR) with primers that contained restriction sites for cloning the PCR products. The 5′ and 3′ primers used for the amplification of staufen fragments had *Bam*HI and *Eco*RI sites, respectively; those for Xlrbpa had *Bam*HI and *Kpn* I sites. The staufen fragments were cloned between the *Bam*HI and *Eco*RI sites of pGEX-2T (14) or pRSET (Invitrogen, San Diego). The fragments of Xlrbpa were cloned between the *Bam*HI and *Kpn* I sites of pGEX-2T that had been modified by introduction of a polylinker to provide a unique *Kpn* I site. Clones with correct inserts were picked from single colonies and were grown in 3-ml cultures for 4 hr. Fusion protein production was induced by addition of isopropyl β-D-thiogalactoside to a concentration of 0.5 mM, after which the cultures were grown for an additional 2 hr and harvested by centrifugation. Crude lysates were produced by sonication of the cell pellet in 500 μl of SDS sample buffer.

RNA-Binding Assays. Fifteen-microliter samples of crude cell extracts were separated by SDS/PAGE and blotted onto Immobilon P membranes (Millipore). The blotted proteins were denatured on the membranes in 8 M urea and slowly renatured by incubation in 10 stepwise dilutions of urea in Tris-buffered saline (TBS), 2:3 (vol/vol), for 10 min each. The membranes were then rinsed in TBS and blocked for 1 hr in 25 mM NaCl/10 mM $MgCl_2$/10 mM Hepes, pH 8/0.1 mM EDTA/1 mM dithiothreitol/5% (wt/vol) Carnation nonfat dry milk. RNA binding was carried out in 50 mM NaCl/10 mM $MgCl_2$/10 mM Hepes, pH 8/0.1 mM EDTA/1 mM dithiothreitol/2.5% milk plus ^{32}P-labeled T7 transcripts of U1 or U2 snRNA genes (13), adenovirus VA1 transcripts (15), or the 3′ untranslated region (UTR) of bicoid RNA (16) (2×10^5 cpm/ml). As controls we used a T7 transcript of the small nuclear ribonucleoprotein B gene or the dorsal gene of *Drosophila*, neither of which is predicted to form strong secondary structure. Poly(rI) and poly(rC) (Pharmacia) were partially hydrolyzed in Na_2CO_3 buffer (pH 10.2) at 70°C for 40 min, end-labeled using T4 polynucleotide kinase and [γ-^{32}P]ATP, and purified over a Sephadex G-25 column. Equal amounts of poly(rI) and poly(rC) were annealed to produce poly(rI)-poly(rC) by heating to 95°C for 5 min and cooling slowly. All single-stranded RNA probes were denatured prior to binding by boiling for 5 min and quick cooling on ice. After incubation, membranes were washed 3 × 10 min in binding buffer and exposed to x-ray film.

Search Programs and Databases. The similarities between Xlrbpa, staufen, and human TRBP were initially identified in searches of the GenBank (release 72) and European Molecular Biology Laboratory (EMBL) (release 31) databases, using the TFASTA program of Pearson and Lipman (17). The presence of multiple regions of similarity among these pro-

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.

Abbreviations: DAI, double-stranded RNA-activated inhibitor; ds, double-stranded; HIV, human immunodeficiency virus; TAR, trans-activating region; UTR, untranslated region; snRNA, small nuclear RNA.

‡To whom reprint requests should be addressed at: Department of Embryology, Carnegie Institution, 115 West University Parkway, Baltimore, MD 21210.

teins was subsequently detected using the DIAGON program of the Genetics Computer Group package (18), and these were further analyzed using the protein subsequence analysis functions of the MacVector program (IBI/Kodak). To identify other proteins that contain the dsRNA-binding consensus, we searched the Protein Identification Resource protein database (Version 32) on the Edinburgh Distributive Array Processor using the motif searching program PTNSEARCH (19) and a pam50 matrix. Since this program has a maximum motif length of 28 residues, the consensus was divided into two overlapping N-terminal regions, P(VIM)XXL(NQ)E(YL)X-Q(KR)XXX-XPX(YF)X(LVI)XXXSGPAH and (KR)-XXXXPX(YF)X(LVI)XXXSGPAHX(KR)XFTFX(VLIM-C)X(VLI), and a C-terminal consensus (GA)XGXSKKX-AKXXAAXXALXXL. Similar results were also obtained using the PROFILE search program (20) of the Genetics Computer Group package, although this program does not identify multiple regions of similarity within the same protein. Hssona, a human cDNA sequence of unknown function whose translation is not yet in the protein databases, was identified in searches of the EMBL database with the consensus sequence, using TFASTA. In calculating the similarity between domains, we considered amino acids in the following sets as equivalent to each other: V,L,I,M; K,R; S,T; E,D; Y,F; N,Q; A,G.

Sequences. The following nucleic acid sequences were used to obtain the protein sequences shown in Fig. 2: staufen, accession no. M6911 (10); human trans-activating region (TAR)-binding protein, M60801 (6); Xlrbpa, M96370; human dsRNA-activated inhibitor (DAI), M35663 (21); mouse TIK, M65029 (22); vaccinia (strain WR) E3L protein, M36339 (ORF1) translated from the minus-strand 4530–3961 (23); Hssona, X63753; *Escherichia coli* RNase III, X02673 (24); porcine rotavirus group C ns34 protein, M669115 (25); *Schizosaccharomyces pombe* pac1, X54998 (26).

RESULTS AND DISCUSSION

In screens of a *Xenopus* expression library made from immature ovary poly(A)$^+$ RNA (11), we recovered several clones that bound to ^{32}P-labeled U1 and U2 snRNAs. These cDNAs defined a dsRNA-binding protein, Xlrbpa, of unknown function. Initial binding experiments with a partial clone of Xlrbpa, which lacked N-terminal coding sequences, identified a 76-amino acid region (see xlrbpa-2 in Fig. 2*a*) that was sufficient to bind U1 and U2 snRNAs in a filter binding assay. Full-length Xlrbpa was later found to contain a second region (see xlrbpa-1 in Fig. 2*a*) that is similar to the identified RNA-binding domain. Computer searches showed that Xlrbpa shares extensive regions of similarity with human TRBP, a protein that binds *in vitro* to the TAR stem–loop of human immunodeficiency virus (HIV) RNA (6). In addition, the RNA-binding domain of Xlrbpa showed significant similarity to three regions in the staufen protein of *Drosophila*, the product of a maternal gene required for the correct localization of bicoid and oskar RNAs to the anterior and posterior poles of the egg (10, 27–29). Because Xlrbpa and human TRBP were known to bind double-stranded regions of RNA, it seemed probable that staufen might also function as a dsRNA-binding protein.

To investigate whether the regions of sequence similarity defined a conserved dsRNA-binding domain, we compared the RNA-binding properties of the 76-amino acid domain of Xlrbpa and the region of staufen that is most similar to it (see Dmstau-3 in Fig. 2*a*). These regions of the two proteins were expressed from the pGEX-2T vector (14) as fusion proteins with a portion of glutathione *S*-transferase or as 6xHis fusions from the pRSET vector (Invitrogen) and were analyzed on Northwestern blots with several RNA probes. The Xlrbpa and staufen constructs bound strongly to RNAs with extensive secondary structure, such as U1 and U2 snRNAs, VA1 RNA of adenovirus (15), and the 3′ UTR of bicoid RNA (16, 30), but failed to bind control RNAs that were not predicted to form double-stranded regions (data not shown). To confirm that the binding was specific for dsRNA, we compared the ability of the fusion proteins to bind poly(rI), poly(rC), or poly(rI)-poly(rC); the double-stranded product was formed by annealing the two single-stranded homopolymers. The fusion proteins did not bind either single-stranded polymer, whereas they strongly bound the double-stranded product (Fig. 1).

These results demonstrate that small regions in Xlrbpa and staufen, which contain the conserved motif, bind dsRNA. To determine if other proteins contained this motif, we performed several database searches using a consensus sequence derived from the seven copies of the motif in Xlrbpa, staufen, and human TRBP. The searches identified five other proteins that each contain one full-length copy of the domain; the alignment of all 12 domains is shown in Fig. 2*a*. In pairwise comparisons (excluding comparisons between proteins that may be homologous), the domains show an average of 29% amino acid identity in a region of 65–68 residues, or 42% similarity when conservative changes are taken into account. These are the only proteins that show strong similarity to the complete consensus. However, several proteins contain regions that match the C-terminal part of the consensus but fit the N-terminal part only poorly (Fig. 2*b*). Two arguments suggest that the partial similarity of these shorter regions is significant. (*i*) Less than one protein among the 40,298 entries in the Protein Identification Resource data base is expected to match the consensus as closely as the most divergent of the shorter domains (31). (*ii*) Six of the shorter domains occur in proteins that also contain one or more complete domains. Fig. 3 shows the arrangement of the 12 complete and 8 shorter domains in the 10 proteins in which they occur.

Although all of the similarities in Fig. 2 are significant on statistical grounds, the strongest argument that the complete consensus sequence defines a dsRNA-binding domain is provided by what is known about the proteins themselves. In addition to staufen, Xlrbpa, and TRBP, three of the other proteins with full-length domains are known to bind double-stranded regions of RNA. Perhaps the best characterized of

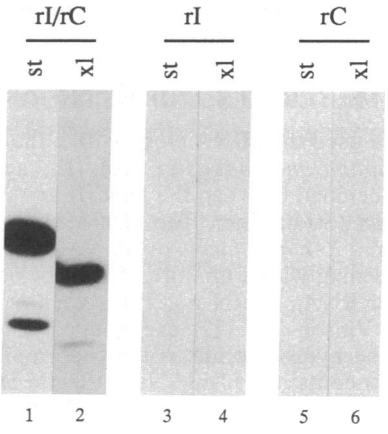

FIG. 1. RNA-binding domains of staufen and Xlrbpa specifically bind dsRNA. Fusion proteins that contained 101 residues of staufen (amino acids 569–669) or 76 residues of Xlrbpa (amino acids 102–177) were electrophoresed on a polyacrylamide gel, transferred to an Immobilon filter, and probed with double-stranded poly(rI)-poly(rC), single-stranded poly(rI), or single-stranded poly(rC) (Northwestern blots). Strong binding of dsRNA (lanes 1 and 2) contrasted with no binding of single-stranded RNA (lanes 3–6). st, Staufen; xl, Xlrbpa. The fainter bands in lanes 1 and 2 represent breakdown products of the fusion proteins.

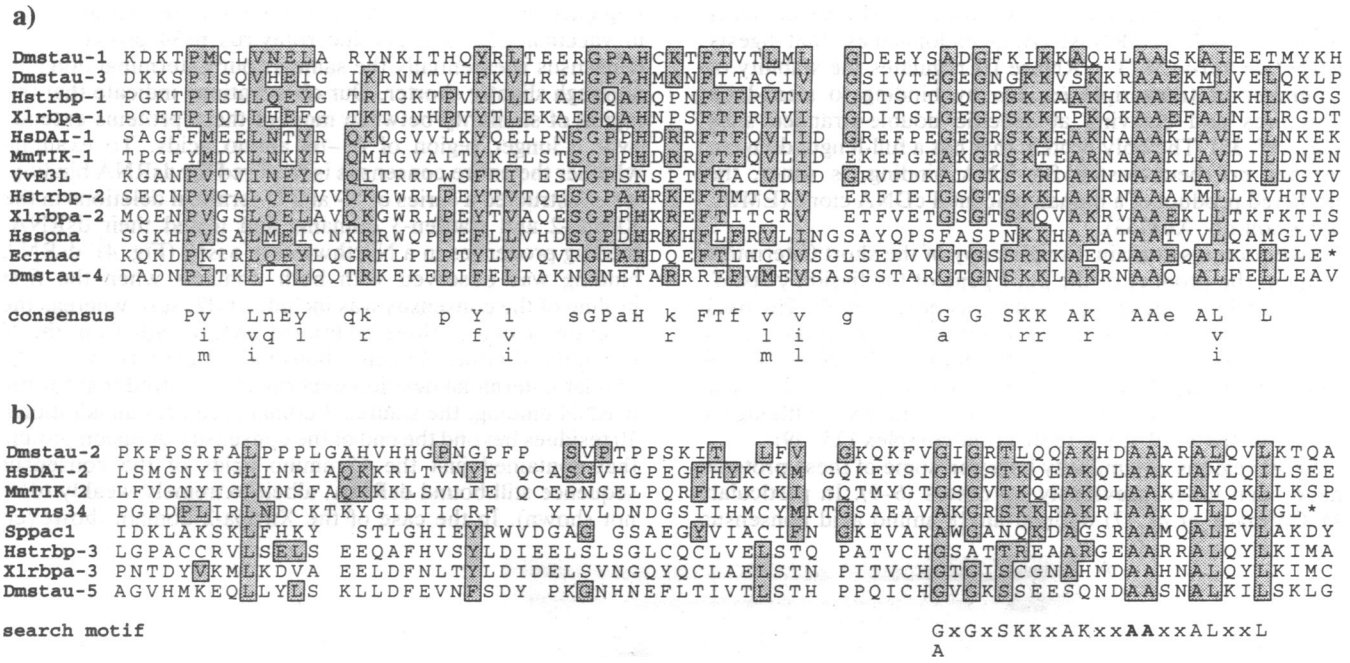

FIG. 2. Alignment of the protein sequences that contain the dsRNA-binding domain. (*a*) Twelve sequences that contain the full-length domain. Boxes indicate amino acids that are identical in at least five sequences or similar in at least eight. The consensus sequence is shown below. Uppercase letters indicate residues that occur in at least eight sequences; lowercase letters denote either residues shared by five to seven sequences or conservative substitutions (shown in their order of frequency). The sequences are as follows: Dmstau-1, staufen 308–380; Dmstau-3, 575–647; Hstrbp-1, human TAR-binding protein 6–78; Xlrbpa-1, *Xenopus* RNA-binding protein A 17–89; HsDAI-1, human dsRNA-dependent kinase 6–79; MmTIK-1, mouse TIK 5–78; VvE3L, vaccinia virus E3L 114–186; Hstrbp-2, 135–208; Xlrbpa-2, 109–182; Hssona, human son-a 1361–1436; Ecrnac, *E. coli* RNase III 152–227; Dmstau-4, 708–782. (*b*) Eight C-terminal short domains. Boxed amino acids correspond to consensus residues in the full-length domain. These sequences were found by a computer search using the PTNSEARCH motif searching program and the indicated search sequence. The sequences are as follows: Dmstau-2, 490–559; HsDAI-2, 96–169; MmTIK-2, 91–164; Prvns34, porcine group C rotavirus ns34 protein, 332–402; Sppac1, *S. pombe* pac1 protein, 287–358; Hstrbp-3, 269–342; Xlrbpa-3, 222–295; Dmstau-5, 948–1020.

these is the human dsRNA-dependent protein kinase DAI, an interferon-induced protein that acts in the cellular defense against viral infection (21, 32, 33). DAI kinase becomes activated when it binds to dsRNA; it then phosphorylates eukaryotic initiation factor eIF-2a, resulting in an inhibition of translation that prevents viral replication. McCormack *et al.* (7) used a Northwestern assay to show that the first 98 amino acids of human DAI are sufficient to bind several viral RNAs. This region contains one complete domain, extending from residues 10 to 75. Furthermore, DAI no longer binds to dsRNA when residues 39–50 or 58–69 are deleted (8). The mouse TIK gene, which has the same arrangement of one full-length and one short domain as DAI, shares extensive amino acid similarity with this protein and is probably its mouse homologue (8, 22, 34). A single copy of the domain occurs in a 25-kDa protein encoded by the vaccinia virus E3L gene, which inhibits the activation of DAI (9, 23, 35, 36). This protein binds poly(rI)-poly(rC) and may block DAI activity

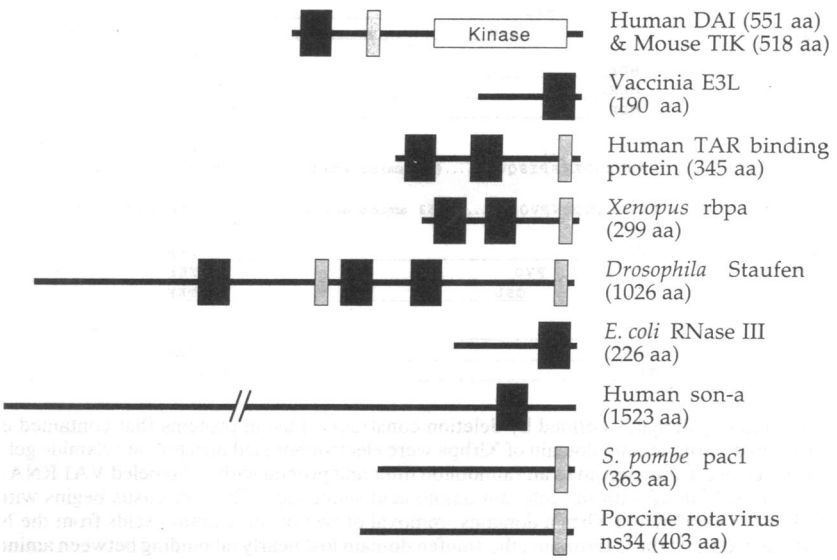

FIG. 3. Location of full-length and short domains in the 10 proteins in which they occur. Full-length domains are indicated by dark grey boxes and short domains are indicated by smaller light grey boxes.

by competing for dsRNA. A prokaryotic protein that contains the domain is *E. coli* RNase III, an endonuclease that digests dsRNA *in vitro* (24, 37). RNase III regulates the stability of a number of transcripts *in vivo* by binding to stem–loop structures and cleaving at specific double-stranded sites (e.g., ref. 38). The only protein that has a full-length domain, but has not been tested for RNA binding, is son-a, the predicted product of a human placental cDNA clone (EMBL accession no. X63753).

Two proteins contain a single copy of the short C-terminal domain: the product of the *pac1* gene of the fission yeast *S. pombe* and porcine group C rotavirus protein ns34. The pac1 protein, which was originally identified as a suppressor of meiosis, shares 25% amino acid identity with RNase III and can degrade dsRNA when expressed in *E. coli* (26). The rotavirus ns34 protein is not known to bind RNA, although it is part of the viral RNA replication complex (25, 39).

Short regions of similarity between some of these proteins have been noted previously by others (6–9). In particular, McCormack *et al.* (7) defined an 18-amino acid consensus sequence that occurs twice in DAI and mouse TIK and once in vaccinia E3L and porcine rotavirus ns34 protein. This consensus is essentially the same as our C-terminal domain, although slightly shorter. Our observations indicate that the region of similarity between most of these proteins extends over a longer region of 65–68 amino acids. To examine whether the entire consensus is required for dsRNA binding, we constructed a series of N- and C-terminal deletions in the Xlrbpa-2 and staufen-3 domains and tested their dsRNA-binding properties in a Northwestern assay (Fig. 4). dsRNA binding was observed with both domains when the first proline of the consensus was included (xl2, st3), whereas the deletion of two or three additional amino acids from the N terminus of either domain abolished all activity (xl3, st4). Similar C-terminal deletion experiments showed that for full dsRNA binding, the staufen-3 domain requires an additional 10 residues beyond the end of the consensus. A fusion protein that contained just the 65 amino acids of the consensus sequence still bound dsRNA, albeit extremely weakly (data not shown). In the case of the Xlrbpa-2 domain, however,

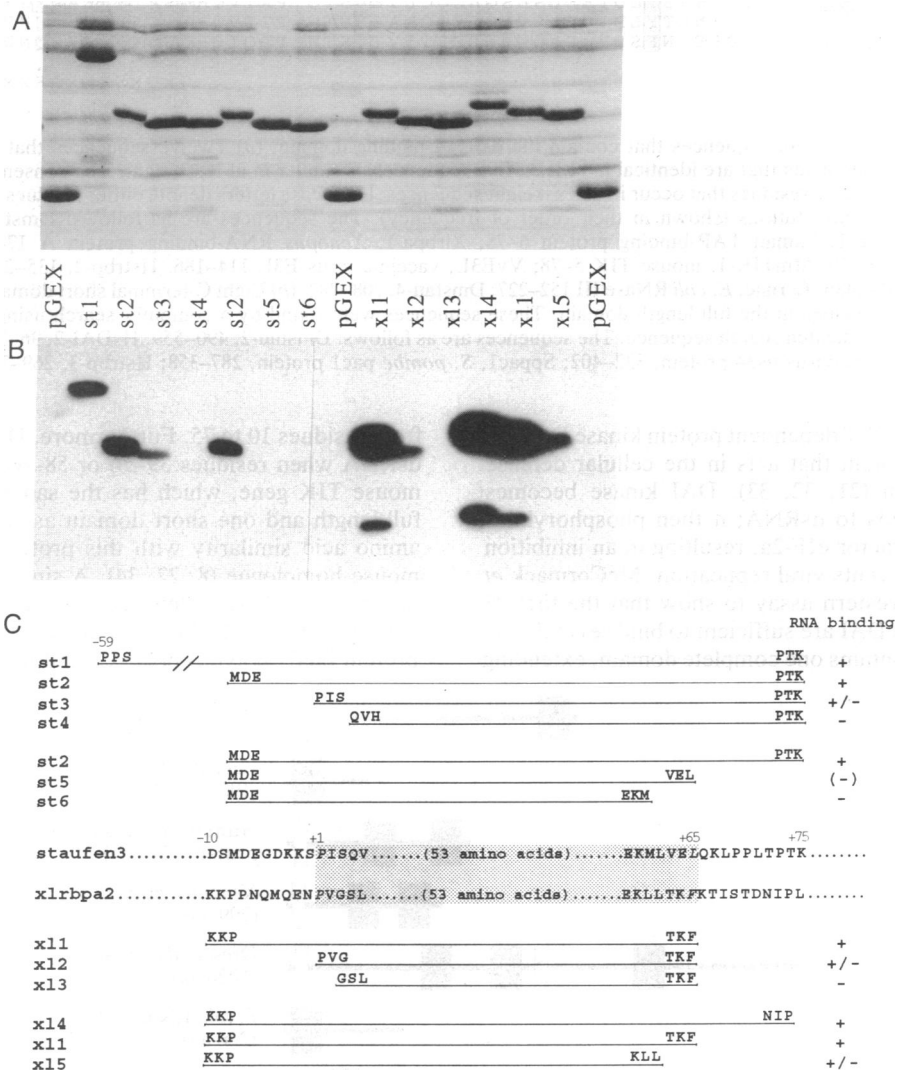

FIG. 4. Limits of the dsRNA-binding domains defined by deletion constructs. Fusion proteins that contained either a complete or deleted portions of the third domain of staufen or the second domain of Xlrbpa were electrophoresed on a polyacrylamide gel and stained with Coomassie blue (*A*). Proteins from a duplicate gel were transferred to an Immobilon filter and probed with ^{32}P-labeled VA1 RNA (*B*). The deletion constructs used in this experiment are shown in *C* along with the relevant amino acid sequences. The consensus begins with P at +1 and extends to L (staufen) or F (Xlrbp) at +65. For the staufen and Xlrbpa domains, removal of two or three amino acids from the N terminus of the consensus sequence eliminated binding (st4 and xl3). At the C terminus, the staufen domain lost nearly all binding between amino acids 75 and 65 (constructs st2 and st5), whereas the Xlrbpa domain retained binding at amino acid 65 (construct xl1) and had greatly diminished binding at amino acid 63 (construct xl5). The minimal binding domain thus extends for about 65 amino acids in each case and is essentially coextensive with the consensus sequence.

fusions that extended no further than the final phenylalanine of the consensus (xl1, xl2) still showed strong dsRNA binding. Deletion of the C-terminal three amino acids (xl5) reduced binding but did not abolish it entirely. Thus, the consensus sequence accurately defines the N-terminal extent of the dsRNA-binding domain, whereas the C-terminal requirement differs slightly between the two examples analyzed. Because the other domains in Fig. 2*a* likewise show similarity throughout a region of 65–68 amino acids, we suggest that features of this entire region are required for binding. We refer to the conserved amino acids within this domain as the dsRNA-binding consensus sequence.

Earlier studies have identified a single-stranded RNA-binding domain in various eukaryotic nuclear and cytoplasmic proteins, including several involved in mRNA processing (3–5). This RNA-binding domain is ≈90 amino acids long and contains two well-conserved sequences, RNP1 and RNP2, separated by about 30 amino acids. The dsRNA-binding domain that we describe has no obvious sequence similarity to the RNP1 and RNP2 consensus sequences. It is perhaps noteworthy that both RNA-binding domains are long, suggesting that structural features as well as specific residues are important. How the dsRNA-binding domain interacts with RNA and whether that interaction depends on specific RNA sequences are questions that must await further biochemical and structural analysis.

D.StJ. and M.J. noticed this domain independently and should be considered equal first authors. This work was supported by Wellcome Trust Senior Fellowships to D.StJ. and N.H.B., a fellowship from the Austrian Science Foundation to M.J., and Research Grant GM33397 from the National Institutes of Health. J.G.G. is American Cancer Society Professor of Developmental Genetics.

1. Branden, C. & Tooze, J. (1991) *Introduction to Protein Structure* (Garland, New York).
2. Klug, A. & Rhodes, D. (1987) *Trends Biochem. Sci.* **12**, 464–469.
3. Adam, S. A., Nakagawa, T., Swanson, M. S., Woodruff, T. K. & Dreyfuss, G. (1986) *Mol. Cell. Biol.* **6**, 2932–2943.
4. Bandziulis, R. J., Swanson, M. S. & Dreyfuss, G. (1989) *Genes Dev.* **3**, 431–437.
5. Kenan, D. J., Query, C. C. & Keene, J. D. (1991) *Trends Biochem. Sci.* **16**, 214–220.
6. Gatignol, A., Buckler, W. A., Berkhout, B. & Jeang, K. T. (1991) *Science* **251**, 1597–1600.
7. McCormack, S. J., Thomis, D. C. & Samuel, C. E. (1992) *Virology* **188**, 47–56.
8. Feng, G.-S., Chong, K., Kumar, A. & Williams, B. R. G. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 5447–5451.
9. Chang, H.-W., Watson, J. C. & Jacobs, B. L. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 4825–4829.
10. St Johnston, D., Beuchle, D. & Nüsslein-Volhard, C. (1991) *Cell* **66**, 51–63.
11. Tafuri, S. R. & Wolffe, A. P. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 9028–9032.
12. Vinson, C. R., LaMarco, K. L., Johnson, P. F., Landschulz, W. H. & McKnight, S. L. (1988) *Genes Dev.* **2**, 801–806.
13. Hamm, J., Dathan, N. A. & Mattaj, I. W. (1989) *Cell* **59**, 159–169.
14. Smith, D. B. & Johnson, K. S. (1988) *Gene* **67**, 31–40.
15. Akusjärvi, G., Mathews, M. B., Andersson, P., Vennström, B. & Pettersson, U. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 2424–2428.
16. Macdonald, P. M. & Struhl, G. (1988) *Nature (London)* **336**, 595–598.
17. Pearson, W. R. & Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2444–2448.
18. Devereux, J., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395.
19. Coulson, A. F. W., Collins, J. F. & Lyall, A. (1987) *Comput. J.* **30**, 420–424.
20. Gribskov, M., McLachlan, A. D. & Eisenberg, D. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 4355–4358.
21. Meurs, E., Chong, K., Galabru, J., Thomas, N. S., Kerr, I. M., Williams, B. R. & Hovanessian, A. G. (1990) *Cell* **62**, 379–390.
22. Icely, P. L., Gros, P., Bergeron, J. J., Devault, A., Afar, D. E. & Bell, J. C. (1991) *J. Biol. Chem.* **266**, 16073–16077.
23. Ahn, B. Y., Gershon, P. D., Jones, E. V. & Moss, B. (1990) *Mol. Cell. Biol.* **10**, 5433–5441.
24. March, P. E., Ahnn, J. & Inouye, M. (1985) *Nucleic Acids Res.* **13**, 4677–4685.
25. Qian, Y. A., Jiang, B. M., Saif, L. J., Kang, S. Y., Ojeh, C. K. & Green, K. Y. (1991) *Virology* **184**, 752–757.
26. Iino, Y., Sugimoto, A. & Yamamoto, M. (1991) *EMBO J.* **10**, 221–226.
27. St Johnston, D., Driever, W., Berleth, T., Richstein, S. & Nüsslein-Volhard, C. (1989) *Development* **107**, Suppl., 13–19.
28. Ephrussi, A., Dickinson, L. K. & Lehmann, R. (1991) *Cell* **66**, 37–50.
29. Kim-Ha, J., Smith, J. L. & Macdonald, P. M. (1991) *Cell* **66**, 23–35.
30. Macdonald, P. M. (1990) *Development* **110**, 161–171.
31. Collins, J. F., Coulson, A. F. W. & Lyall, A. (1988) *Comput. Appl. Biosci.* **4**, 67–71.
32. Levin, D. H., Pethryshyn, R. & London, I. M. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 832–836.
33. Hovanessian, A. G. (1989) *J. Interferon Res.* **9**, 641–647.
34. Thomis, D. C., Doohan, J. P. & Samuel, C. (1992) *Virology* **188**, 33–46.
35. Broyles, S. S. & Pennington, M. J. (1990) *J. Virol.* **64**, 5376–5382.
36. Watson, J. C., Chang, H. W. & Jacobs, B. L. (1991) *Virology* **185**, 206–216.
37. Robertson, H. D., Webster, R. E. & Zinder, N. D. (1968) *J. Biol. Chem.* **243**, 82–91.
38. Saito, H. & Richardson, C. C. (1981) *Cell* **27**, 533–542.
39. Gallegos, C. O. & Patton, J. T. (1989) *Virology* **172**, 616–627.