to CFI_y (*18*). Given that in mammals CPSF interacts with poly(A) polymerase, cleavage factors, and the pre-mRNA, it is possible that the CPSF-related subunits of PFI also participate in specific interactions with CFI_y as well as with the RNA substrate. Thus in yeast, specificity and processivity may be conferred upon poly(A) polymerase by the combined action of CFI_y and PFI. The reason why PFI is not required for the cleavage reaction, in spite of its homology with CPSF, might be because the poly(A) polymerase itself does not participate in cleavage (*5*). However, the fact that neither the signal sequences on the RNA substrates nor the protein factors are interchangeable between yeast and metazoans may reflect subtle but important differences in protein-protein and RNA-protein interactions between the polyadenylation systems of yeast and metazoans.

Do bacteria have CPSF or PFI? Three open reading frames of unknown function in the M. *jannaschii* (*13*) are significantly related to the 73-kD protein and to Ysh1/Brr5. In view of recent surprising findings in the field of bacterial polyadenylation (*26*), this raises the question whether these proteins are involved in pre-mRNA 3′-end formation. However, M. *jannaschii* does not encode homologs of the 160-, 100-, or 30-kD subunits of CPSF nor of Fip1. As far as sequence data are available, this bias also holds true for eubacteria (*10, 16*). It thus remains an open question as to which degree pre-mRNA 3′-end formation is conserved between the different domains of life. In any case, the high degree of sequence conservation suggests that Ysh1/Brr5 and its related proteins have important and essential functions.

REFERENCES AND NOTES

1. E. Wahle and W. Keller, *Trends Biochem. Sci.* **21**, 247 (1996); J. L. Manley, *Curr. Opin. Genet. Dev.* **5**, 222 (1995).
2. S. Bienroth, W. Keller, E. Wahle, *EMBO J.* **12**, 585 (1993); E. Wahle, *Cell* **66**, 759 (1991).
3. U. Rüegsegger, K. Beyer, W. Keller, *J. Biol. Chem.* **271**, 6107 (1996); Y. Takagaki, L. C. Ryner, J. L. Manley, *Genes Dev.* **3**, 1711 (1989).
4. Z. Guo and F. Sherman, *Mol. Cell. Biol.* **15**, 5983 (1995).
5. J. Chen and C. Moore, *ibid.* **12**, 3470 (1992); L. Minvielle-Sebastia, P. J. Preker, W. Keller, *Science* **266**, 1702 (1994).
6. Y. Takagaki and J. L. Manley, *Nature* **372**, 471 (1994).
7. S. Bienroth, E. Wahle, C. Suter-Crazzolara, W. Keller, *J. Biol. Chem.* **266**, 19768 (1991); A. Jenny, H.-P. Hauri, W. Keller, *Mol. Cell. Biol.* **14**, 8183 (1994); A. Jenny and W. Keller, *Nucleic Acids Res.* **23**, 2629 (1995).
8. K. G. K. Murthy and J. L. Manley, *J. Biol. Chem.* **267**, 14804 (1992); K. G. K. Murthy and J. L. Manley, *Genes Dev.* **9**, 2672 (1995).
9. A rabbit was injected three times with purified CPSF 73-kD subunit and once with complete CPSF. The serum was affinity-purified and used for screening a calf aorta endothelial cDNA expression library (Stratagene 936705) as described (*24*). Additional clones were obtained with a NH_2-terminal probe (*24*).
10. A. Jenny, L. Minvielle-Sebastia, P. J. Preker, W. Keller, unpublished data.
11. Tryptic peptides were obtained from the same preparative SDS-polyacrylamide gel as described (*7*).
12. A cDNA fragment starting at the Hind III site and containing the authentic stop codon was subcloned into pQE11 (Qiagen) and expressed in *Escherichia coli* (*7*). A rabbit was injected four times with 100 μg of protein (*25*). Monoclonal antibodies were obtained from the fusion described earlier (*7*).
13. C. J. Bult *et al.*, *Science* **273**, 1058 (1996); the *M. jannaschii* gene numbers are: MJ0047, MJ0162, and MJ1236.
14. M. Johnston, S. Andrews, R. Waterson, GenBank accession number U17245 (1994).
15. S. M. Noble and C. Guthrie, *Genetics* **143**, 67 (1996).
16. G. Chanfreau, S. M. Noble, C. Guthrie, personal communication.
17. *YSH1* was cloned by PCR and the screening of a genomic library (*10*). A truncated form of *YSH1* starting at the internal Bgl II site and containing the authentic stop codon was subcloned into pQE10 (Qiagen) and expressed in *E. coli* (*7*). A rabbit was injected four times with 100 μg of SDS-polyacrylamide gel eluted fusion protein (*25*).
18. P. J. Preker, J. Lingner, L. Minvielle-Sebastia, W. Keller, *Cell* **81**, 379 (1995).
19. P. J. Preker and W. Keller, in preparation.
20. E.-C. Park, D. Finley, J. W. Szostak, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 1249 (1992); A. Varshavsky, *Cold Spring Harbor Symp. Quant. Biol.* **60**, 461 (1996); plasmids were modified as follows: an Xho I–Bam HI cassette containing the *CYC1* promoter, the ubiquitin coding sequence, and either a Met or Arg codon, was cloned between the *GAL1* upstream-activating sequence, and a Not I cloning site (plasmids pGUM1 and pGUR1, respectively). These plasmids are centromeric and *ADE2*-marked (*10*). A Not I PCR (polymerase chain reaction) fragment carrying the complete open reading frame of *YSH1* was introduced, yielding pGUM-YSH1 and pGUR-YSH1, respectively. Both plasmids can rescue a strain disrupted for the chromosomal copy of *YSH1* on galactose medium. However, on 5% glucose medium, only the pGUM-YSH1-containing strain [LM112 (*MAT***a**, *ysh1::TRP1*; *ura3-1*; *trp1*Δ; *ade2-1*; *leu2-3,112*; *his3-11,15*; and pGUM-YSH1)] is viable (LM111 is isogenic with LM112, except that it contains pGUR-YSH1).
21. Extracts were prepared as described [A. Ansari and B. Schwer, *EMBO J.* **14**, 4001 (1995)]. The reactions were processed as described (*5*), except that in cleavage assays, cordycepine triphosphate (0.5 mM final concentration) replaced CTP, and magnesium acetate (1.8 mM) was used instead of EDTA.
22. J. Lingner, J. Kellermann, W. Keller, *Nature* **354**, 496 (1991).
23. S. M. L. Barabino and W. Keller, unpublished data.
24. J. B. Olmstedt, *J. Biol. Chem.* **256**, 11955 (1981); J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular Cloning* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989).
25. E. Harlow and D. Lane, *Antibodies: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1988).
26. S. N. Cohen, *Cell* **80**, 829 (1995).

27. We thank R. Pöhlmann for help in DNA sequencing, E.-C. Park and J. W. Szostak for plasmids, H. Pick and P. Philippsen for the yeast genomic library, G. Chanfreau, S. M. Noble, and C. Guthrie for communicating unpublished results, and U. Rüegsegger and E. Wahle for critically reading the manuscript. Supported by the Kantons of Basel, the Swiss National Science Foundation, a European Union long-term fellowship (Human Capital and Mobility Program, L.M.-S.), and a predoctoral fellowship from the Boehringer-Ingelheim Fonds (P.J.P.).

19 June 1996; accepted 11 October 1996

Dependence of Yeast Pre-mRNA 3′-End Processing on CFT1: A Sequence Homolog of the Mammalian AAUAAA Binding Factor

Gabi Stumpf and Horst Domdey*

3′-End formation of pre-mRNA in yeast and mammals follows a similar but distinct pathway. In *Saccharomyces cerevisiae*, the cleavage reaction can be reconstituted by two activities called cleavage factor I and II (CFI and CFII). A CFII component, designated CFT1 (cleavage factor two) was identified by its sequence similarity to the AAUAAA-binding subunit of the mammalian cleavage and polyadenylation specificity factor (CPSF), even though the AAUAAA signal sequence appears to play no role in yeast pre-mRNA 3′ processing. Depletion of a yeast whole-cell extract with antibodies to CFT1 protein abolished cleavage and polyadenylation of pre-mRNAs. Addition of CFII restored cleavage activity, but not polyadenylation. Polyadenylation required the further addition of poly(A) polymerase and polyadenylation factor I, suggesting a close but not necessarily direct association of these two factors with the CFT1 protein.

Most eukaryotic pre-mRNAs are processed at their 3′ ends by endonucleolytic cleavage and subsequent addition of a poly(A) tail at the upstream fragment (*1*). Some studies have provided evidence for a possible evolutionary link between yeast and mammalian trans-acting factors involved in 3′-end formation (*2–4*). The poly(A) polymerase (PAP) has so far revealed the highest sequence similarity among the factors identified for cleavage and polyadenylation in yeast and mammals (*5*). The two known subunits of CFI (*6*), the Rna14 and Rna15 proteins (*7*), have been proposed to be the yeast counterparts of the 77-kD and 64-kD subunits of the mammalian cleavage stimulatory factor (CstF) (*2, 3*). Nevertheless,

Institut für Biochemie, Genzentrum, der Ludwig-Maximilians-Universität München Feodor-Lynen-Str. 25, D-81377 München, Germany

*To whom correspondence should be addressed.

the functional significance of CFI is more related to the mammalian CPSF (*8*) in that both factors (CFI and CPSF) are required for cleavage and polyadenylation, whereas CstF is only essential for the cleavage step. Furthermore, a weak similarity between the yeast protein Fip1 (*9*), which binds to both Rna14 protein and PAP, and the human 160-kD subunit of CPSF, which directly interacts with the AAUAAA signal, has been reported (*3*). The yeast *S. cerevisiae*, however, lacks the hexanucleotide AAUAAA, which is highly conserved in mammals, and seems to use different signals to direct mRNA 3′-end formation (*1*).

To identify in *S. cerevisiae* components of the 3′-end processing machinery, which are conserved between yeast and mammals, we made a computer search of the non-redundant NCBI protein database with sequences of all reported mammalian proteins participating in 3′-end formation of pre-mRNA. The open reading frame D9740.6 [which consists of 1357 amino acids (aa), (about 153 kD) and has been assigned accession number S61187] on chromosome IV of *S. cerevisiae* revealed 24% identity and 51% sequence similarity throughout the complete amino acid sequence with the mammalian AAUAAA-binding subunit of CPSF that coordinates pre-mRNA 3′-end formation. So far, all the described genes (*5*, *7*, *9*) that are needed for the basic 3′-end processing reaction in yeast are essential for viability (*PAP1*, *RNA14*, *RNA15*, and *FIP1*). To establish the role of this putative 3′-end processing factor, termed CFT1, we disrupted the corresponding gene by replacing more than 97% of the coding region by a kanamycin gene insertion (*10*). In no case, did more than two of the four separated spores from dissected tetrads form colonies, indicating that *CFT1* was indispensable for cell growth. Polyclonal antibodies to CFT1 protein (*11*) recognized a single, approximately 150-kD protein in immunoblots of an ammonium sulfate–fractionated whole cell extract from *S. cerevisiae*, which contained all of the mRNA 3′-end processing activities (*12*, *13*). Treatment of a yeast whole-cell extract with purified antibodies to CFT1 protein (*14*) abolished cleavage and polyadenylation of *CYC1*- and *ADH1*-derived pre-mRNA precursors (*15*) in vitro (Fig. 1, A and B, lanes 5). The preimmune serum and an unrelated antibody had no effect on 3′-end processing, demonstrating the specificity of the depletion (Fig. 1, A and B, lanes 4, 6, and 7).

Reconstitution of cleavage and polyadenylation of a pre-mRNA in vitro was achieved by the combination of four activities, which could be separated on a Mono Q column (PAP, CFI, CFII, and polyadenylation factor I) (*6*). Cleavage and polyadenylation could be uncoupled in vitro, with only CFI and CFII being essential for the cleavage reaction (*6*). As depletion of the yeast extract with antibodies to CFT1 protein had a drastic effect on cleavage, we assumed that the CFT1 protein was a component of either cleavage factor, and separated these activities as described (*6*, *16*).

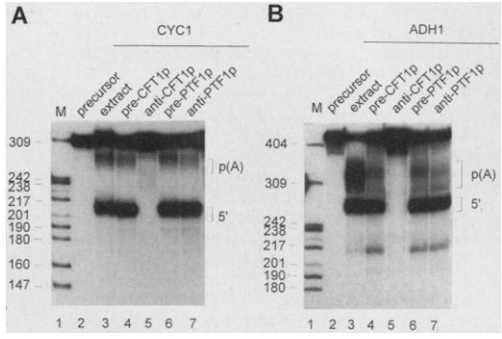


Fig. 1. Inactivation of 3′-end processing activity of a whole cell extract from *S. cerevisiae* by immunodepletion with antibodies to CFT1p. *CYC1* (**A**) and *ADH1* (**B**) derived pre-mRNAs (precursor, lanes 2) were assayed for cleavage and polyadenylation under standard reaction conditions (extract, lanes 3). The extract was immunodepleted with antibody to CFT1p (anti-CFT1p, lanes 5), to the unrelated protein PTF1 (*27*) (anti-PTF1p, lanes 7) or with the corresponding preimmune sera (pre-CFT1p, lanes 4 and pre-PTF1p, lanes 6). The upstream (5′) cleavage product and the polyadenylated 5′ cleavage product [p(A)] are indicated. M (lane 1), molecular size marker in nucleotides (pBR322 DNA digested with Hpa II).

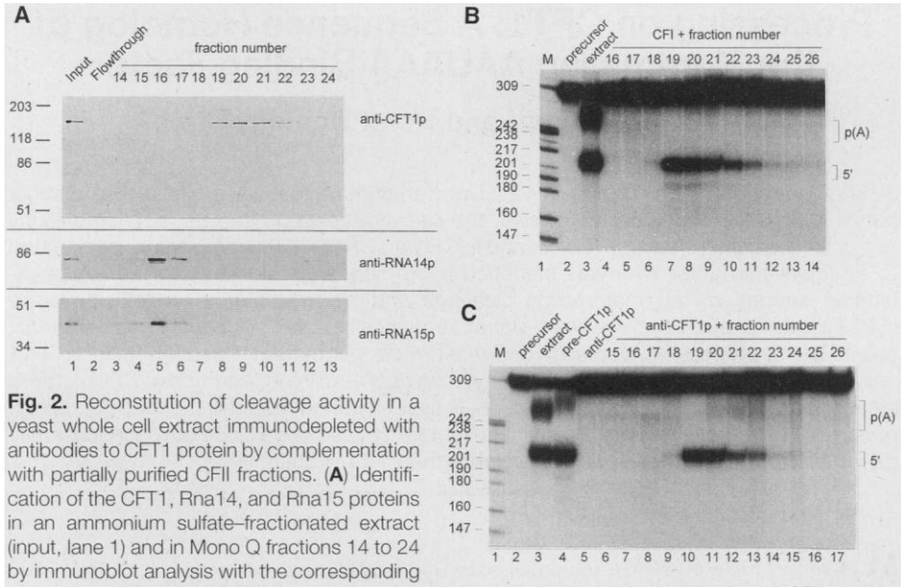


Fig. 2. Reconstitution of cleavage activity in a yeast whole cell extract immunodepleted with antibodies to CFT1 protein by complementation with partially purified CFII fractions. (**A**) Identification of the CFT1, Rna14, and Rna15 proteins in an ammonium sulfate–fractionated extract (input, lane 1) and in Mono Q fractions 14 to 24 by immunoblot analysis with the corresponding antibodies. The sizes of marker proteins are indicated on the left. (**B**) Identification of CFII activity-containing Mono Q fractions by combining 1 μl of fractions 16 to 26 (lanes 4 to 14) with 1 μl of CFI activity (fraction 16) and *CYC1*-derived pre-mRNA (precursor, lane 2) under standard reaction conditions (*15*). The same RNA was also assayed for cleavage and polyadenylation in a yeast whole cell extract (lane 3). (**C**) Cleavage activity of an extract immunodepleted with antibodies against the CFT1 protein (lane 5) after addition of Mono Q fractions (lanes 6 to 17). Reconstitution of 3′-end processing activity of immunodepleted extract was done by adding 1 μl of Mono Q fractions to 2 μl of immunodepleted extract under standard reaction conditions. As control, the *CYC1*-derived pre-mRNA (lane 2) was also incubated in an extract that had not been immunodepleted (lane 3) or treated with the preimmune sera (lane 4). The upstream (5′) cleavage product and the polyadenylated product [p(A)] are indicated. M (lane 1), molecular size marker in nucleotides (pBR322 DNA digested with Hpa II).

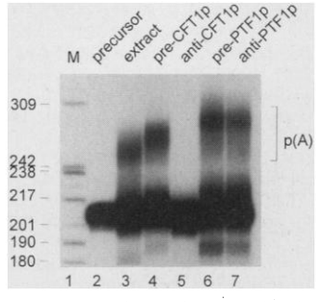


Fig. 3. Inhibition of specific polyadenylation of a precleaved *CYC1*-derived RNA in an extract immunodepleted with antibodies to the CFT1 protein. The precleaved RNA (lane 2) was assayed for specific polyadenylation under standard reaction conditions (lane 3). The extract was immunodepleted with antibodies to the CFT1 protein (lane 5), against the unrelated PTF1 protein (lane 7), or treated with the corresponding preimmune sera (lanes 4 and 6). The polyadenylated RNA [p(A)] is indicated.

Immunoblots with antibodies to CFT1 protein (*17*) detected the elution of a polypeptide of approximately 150 kD (Fig. 2A), correlating with the activity of CFII (Fig. 2B), thus providing evidence for CFT1 being a component of CFII. The final proof emerged from the complementation of the cleavage activity in the immunodepleted yeast whole cell extract with Mono Q fractions containing CFII activity (Fig. 2C, lanes 10 to 13) and coinciding with the elution of the CFT1 protein (Fig. 2A, lanes 8 to 11). The addition of CFII did not restore polyadenylation (Fig. 2C, lanes 10 to 13), a result that was unexpected, because CFII activity is essential only for the cleavage step but not for the specific polyadenylation of a pre-mRNA (*6*, *18*). We then tested precleaved *CYC1* RNA (*19*) in a polyadenylation assay with immunodepleted extract and found that the polyadenylation reaction was inhibited in an extract that had been treated with antibodies to CFT1 protein (Fig. 3, lane 5), whereas no inhibition occured in the control reaction with the corresponding preimmune sera (Fig. 3, lane 4) or when an unrelated antibody was used (Fig. 3, lanes 6 and 7).

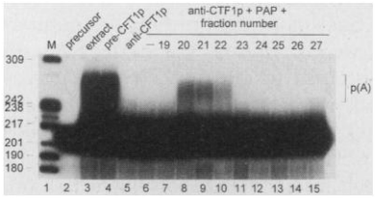


Fig. 4. Reconstitution of polyadenylation in a yeast whole cell extract immunodepleted with antibodies to the CFT1 protein by complementation with partially purified PAP (lanes 7 to 15) and PFI (lanes 8 to 10) fractions (*28*). Precleaved *CYC1*-derived RNA (lane 2) was assayed for specific polyadenylation under standard reaction conditions in an extract immunodepleted with antibodies to CFT1 protein (lanes 5 to 15), pretreated with the corresponding preimmune sera (lane 4), or in an untreated extract (lane 3). M (lane 1), molecular size marker in nucleotides (pBR322 DNA digested with Hpa II).

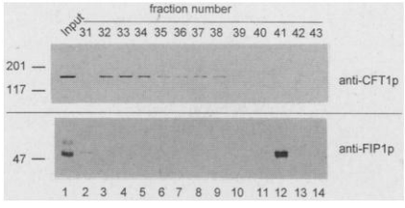


Fig. 5. Identification of the CFT1 and Fip1 proteins in an ammonium sulfate–fractionated extract (input, lane 1) and in Mono Q fractions 31 to 43 (lanes 2 to 14) by immunoblot analysis with the corresponding antibodies. A different Mono Q column profile (6) from the one shown in Figs. 3 and 4 was used. The sizes of marker proteins are indicated on the left.

These results further indicated that one or more of the components that are indispensable for the specific polyadenylation of precleaved RNA, coimmunoprecipitate with the CFT1 protein. Our previous experiments already excluded that CFT1 protein interacts with components of CFI, because reconstitution of the cleavage reaction with the immunodepleted extract was attained by adding only CFII-containing fractions (Fig. 2C, lanes 10 to 13). Coprecipitation of the polyadenylation factor I (PFI, contained in fractions 20 to 22) only can again be ruled out, because reconstitution of the polyadenylation reaction in the immunodepleted extract could not be achieved by the addition of Mono Q fractions containing both CFII and PFI (Fig. 2B, lanes 11 to 13). Therefore, we next analyzed polyadenylation of a cleaved *CYC1* RNA with the depleted extract in the presence of either PAP alone or PAP- and PFI-containing Mono Q fractions (Fig. 4). Polyadenylation could only be restored by adding PAP and PFI together (Fig. 4, lanes 8 to 10) but not with PAP alone (Fig. 4, lane 6). These results offer further evidence for an association of the CFT1 protein with PAP and a component of PFI, and also support the notion that the Fip1 protein, which is a component of the PFI activity, directly interacts with PAP (*9*). A further immunoblot analysis confirmed that the CFT1 protein and the Fip1 protein did not coelute on different Mono Q column with a better resolution for CFII and PFI (Fig. 5). This situation resembles the mammalian 3′-end processing machinery, where a specific interaction between PAP and CPSF was reported (*3*, *20*). The existence of a yeast protein with its sequence similarity to the 160-kD AAUAAA-binding subunit of the mammalian CPSF was unanticipated, because *S. cerevisiae* lacks highly conserved sequences that dictate the 3′-end processing reaction (*1*). The seeming disparity may be due to the broader binding specificity of CPSF that also binds to less defined upstream sequences found in nonviral and viral poly(A) sites (*21*, *22*). In the example HIV-1, the 160-kD subunit of CPSF enhances poly(A) addition by interacting directly with an element upstream of the core poly(A) site (*22*). Furthermore, CPSF contacts the upstream sequences of rare processing sites that have suboptimal variants of the AAUAAA hexamer, and the significance of such mammalian upstream sequences has been discussed (*22*, *23*). Even though the hexanucleotide AAUAAA does not play any role in yeast pre-mRNA 3′-end formation, we could show that the CFT1 protein is a key component of the *S. cerevisiae* 3′-end processing machinery, thus underlining the postulated evolutionary link between yeast and mammals.

REFERENCES AND NOTES

1. E. Wahle and W. Keller, *Annu. Rev. Biochem.* **61**, 419 (1992); W. Keller, *Cell* **81**, 829 (1995); E. Wahle, *Biochim. Biophys. Acta* **1261**, 183 (1995).
2. Y. Takagaki and J. L. Manley, *Nature* **372**, 471 (1994).
3. K. G. K Murthy, and J. L. Manley, *Genes Dev.* **9**, 2672 (1995).
4. Z. Guo and F. Sherman, *Mol. Cell. Biol.* **5**, 5983 (1995).
5. J. Lingner, J. Kellermann, W. Keller, *Nature* **354**, 496 (1991); A. M. Zhelkovsky, M. Kessler, C. L. Moore, *J. Biol. Chem.* **270**, 26715 (1995).
6. J. Chen and C. Moore, *Mol. Cell. Biol.* **12**, 3470 (1992).
7. L. Minvielle-Sebastia, B. Winsor, N. Bonneaud, F. Lacroute, *Mol. Cell. Biol.* **11**, 3075 (1991); L. Minvielle-Sebastia, P. J. Preker, W. Keller, *Science* **266**, 1702 (1994).
8. G. M. Gilmartin and J. R. Nevins, *Genes Dev.* **3**, 2180 (1989); S. Bienroth, E. Wahle, C. Suter-Crazzolara, W. Keller, *J. Biol. Chem.* **266**, 19768 (1991); K. G. K. Murthy and J. L. Manley, *ibid.* **267**, 14804 (1992); A. Jenny and W. Keller, *Nucleic Acids Res.* **23**, 2629 (1995).
9. P. J. Preker, J. Lingner, L. Minvielle-Sebastia, W. Keller, *Cell* **81**, 379 (1995).
10. A. Wach, A. Brachat, R. Pöhlmann, P. Philippsen, *Yeast* **10**, 1793 (1994).
11. For the production of antibodies, a DNA fragment representing the 889 COOH-terminal amino acids of the CFT1 protein was amplified by PCR (polymerase chain reaction) from genomic DNA with the upstream primer (5′-ACTCGAGCAT**ATGGACGATTTATACGCCGATG**-3′) and the downstream primer (5′-GGG**TCACTTACCCTGGCATAAAG**-3′). (Bold letters indicate the coding region of *CFT1*.) The PCR product was treated with Klenow fragment, cut with Nde I, inserted into plasmid pET15b (Novagen) behind an NH_2-terminal hexahistidine tag and expressed in *Escherichia coli* strain BL21(DE3) (*24*). The recombinant protein was purified on nickel-nitrilotriacetic acid-agarose (Qiagen). Antisera were prepared in rabbits with the use of Ribi Adjuvant System (Sigma).
12. The whole cell extract from *S. cerevisiae* was prepared as described (*13*).
13. J. S. Butler, P. P. Sadhale, T. Platt, *Mol. Cell. Biol.* **10**, 2599 (1990); J. S. Butler and T. Platt, *Science* **242**, 1270 (1988).
14. Immunodepletion of the ammonium sulfate fractionated extract was performed as follows. Fifty microliters of preimmune sera or antisera to CFT1 protein or to PTF1 protein were coupled to 40 μl of protein A–Sepharose (Sigma), equilibrated with buffer D (*6*) containing 0.01% NP-40 for 2 hours at 4°C. After three washings, the ammonium sulfate–fractionated extract (50 μl) was added to the resin and incubated for 4 hours at 4°C on a rotating wheel.
15. Standard in vitro processing reactions were performed in a volume of 10 μl as described (*6*, *13*, *25*). Reactions were incubated for 30 min at 30°C. Plasmids used for the in vitro transcription of the 3′-terminal region of *CYC1* (pSKCYC) and *ADH1* (pKSADHwt) were as described (*26*).
16. G. Stumpf, A. Goppelt, H. Domdey, *Mol. Cell. Biol.* **16**, 2204 (1996).
17. Ammonium sulfate–fractionated extract and partially purified Mono Q fractions 14 to 24 were separated on an 8% SDS-polyacrylamide gel and analyzed by immunoblotting. Proteins were detected with antisera and reagents in the enhanced chemilumenescence kit (Amersham).
18. M. Kessler, A. M. Zhelkovsky, A. Skvorak, C. L. Moore, *Biochemistry* **34**, 1750 (1995).
19. Precleaved *CYC1* RNA was obtained by gel purification of the upstream cleavage product of *CYC1* pre-mRNA incubated for 45 min in an *S. cerevisiae* whole cell extract under standard processing conditions. To prevent polyadenylation of the upstream fragment, we replaced ATP with CTP, omitted magnesium, and included EDTA at a final concentration of 1 mM (*6*). Specific polyadenylation of precleaved RNA was assayed under the same conditions as for the

standard cleavage and polyadenylation reaction (*15*) except that reactions were incubated for 60 min at 30°C.
20. S. Bienroth, W. Keller, E. Wahle, *EMBO J.* **12**, 585 (1993).
21. A. Moreira, M. Wollerton, J. Monks, N. J. Proudfoot, *EMBO J.* **14**, 3809 (1995).
22. G. M. Gilmartin, E. S. Fleming, J. Oetjen, B. R. Graveley, *Genes Dev.* **9**, 72 (1995).
23. A. Bilger, C. A. Fox, E. Wahle, M. Wickens, *ibid.* **8**, 1106 (1994).
24. F. W. Studier, A. H. Rosenberg, J. J. Dunn, J. W. Dubendorff, *Methods Enzymol.* **185**, 60 (1990).
25. S. Heidmann, B. Obermaier, K. Vogel, H. Domdey, *Mol. Cell. Biol.* **12**, 4215 (1992).
26. S. Heidmann, C. Schindewolf, G. Stumpf, H. Domdey, *ibid.* **14**, 4633 (1994).
27. J. Hani, G. Stumpf, H. Domdey, *FEBS Lett.* **365**, 198 (1995).
28. For the reconstitution of specific polyadenylation, immunodepleted extract (2 μl) was mixed with 1 μl of partially purified PAP (contained in fraction 11) and 1 μl of Mono Q fractions 19 to 27. The PAP-containing fraction was identified by an unspecific polyadenylation assay as described (*6*).
29. We thank J. Hani for advice with the computer search, U. Sterzer for technical assistance, H. Würl for materials, and A. Goppelt and K. Wiebauer for discussions and critical reading of the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft.

21 June 1996; accepted 21 October 1996

Parallel Synthesis and Screening of a Solid Phase Carbohydrate Library

Rui Liang, Lin Yan, Jennifer Loebach, Min Ge, Yasuhiro Uozumi, Klara Sekanina, Nina Horan, Jeff Gildersleeve, Chris Thompson, Andri Smith, Kaustav Biswas, W. Clark Still, Daniel Kahne*

A solid phase carbohydrate library was synthesized and screened against *Bauhinia purpurea* lectin. The library, which contains approximately 1300 di- and trisaccharides, was synthesized with chemical encoding on TentaGel resin so that each bead contained a single carbohydrate. Two ligands that bind more tightly to the lectin than Gal-β-1,3-GalNAc (the known ligand) have been identified. The strategy outlined can be used to identify carbohydrate-based ligands for any receptor; however, because the derivatized beads mimic the polyvalent presentation of cell surface carbohydrates, the screen may prove especially valuable for discovering new compounds that bind to proteins participating in cell adhesion.

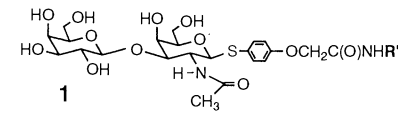
Cell surface carbohydrates play central roles in many normal and pathological biological recognition processes (*1*). For example, cell surface carbohydrates have been implicated in chronic inflammation, in viral and bacterial infection, and in tumorigenesis and metastasis (*2*). Considerable effort has been directed toward (i) understanding how carbohydrates function as recognition signals, and (ii) developing strategies to block undesirable interactions between cell surface carbohydrates and their protein targets. Ligands whose binding to the protein targets is better than that of the natural cell surface carbohydrates could provide an effective means of preventing or treating various diseases. But, progress in understanding how structure and function are related in biologically active carbohydrates has been slow because obtaining synthetic carbohydrate derivatives for biochemical studies is extremely difficult. In contrast, excellent chemical and biological methods to obtain large quantities of peptides and nucleic acids have been available for decades. As a consequence, much more is known about these other biopolymers.

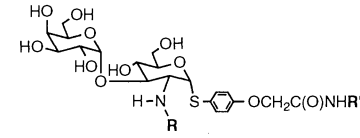
Screening combinatorial libraries of compounds is a very fast way to identify promising leads and elucidate structure-activity relationships (*3*). The first combinatorial libraries were built around peptides and nucleic acids because the chemistry to make them already existed (*4*). The successful use of peptide and nucleic acid libraries has stimulated efforts to develop combinatorial approaches to make other classes of molecules (*5*). One obvious area where a combinatorial approach could have tremendous impact is in carbohydrate chemistry and biochemistry.

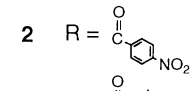
For the time being, the most satisfactory chemical approaches to the construction of compound libraries involve synthesis of the molecules on a solid support (*6*). Despite the efforts made, over more than 20 years, there has been little progress in the solid phase synthesis of carbohydrates. Carbohydrates present special difficulties for solid phase synthesis because the bonds between monomers must be formed stereospecifically and in high yield. Most glycosylation methods are extremely sensitive to structural variations in the glycosyl donor-acceptor pairs (*7*). Reaction conditions that provide excellent yields with one donor-acceptor pair may give virtually no product for another donor-acceptor pair. Furthermore, the stereochemical outcome is often difficult to predict. The unreliability of most glycosylation methods has precluded the construction of a solid phase carbohydrate library, which requires the ability to make a wide range of different glycosidic linkages both stereoselectively and in high yield.

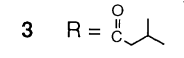
Several years ago we discovered a glycosylation method that makes use of anomeric sulfoxides as glycosyl donors. Anomeric sulfoxides can be activated almost instantaneously at low temperature regardless of the protecting groups on the sugar hydroxyl groups. The low temperature of reaction leads to excellent stereochemical control for a wide range of glycosyl donor-acceptor pairs while preventing side reactions so that it is possible to get nearly quantitative yields on the solid phase (*8*, *9*). We now report the synthesis of a solid phase carbohydrate library using the sulfoxide glycosylation reaction. The library was screened against a carbohydrate binding protein, and two ligands that bind more tightly than the natural ligand were identified. The strategy described can be used to identify carbohydrate-based ligands for any receptor; however, the on-bead screen may prove especially useful for discovering carbohydrate ligands that bind to their receptors in a polyvalent fashion.

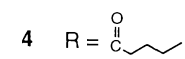
We designed the carbohydrate library to contain a diverse array of glycosidic linkages. The library consists of approximately

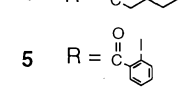












A series: R' = TentaGel
B series: R' = $CH_2CH_2OCH_3$

Fig. 1. Relevant structures contained in the library.

R. Liang, L. Yan, J. Loebach, M. Ge, N. Horan, J. Gildersleeve, C. Thompson, A. Smith K. Biswas, D. Kahne, Department of Chemistry, Princeton University, Princeton, NJ 08544, USA.
Y. Uozumi, K. Sekanina, W. C. Still, Department of Chemistry, Columbia University, New York, NY 10027, USA.
*To whom correspondence should be addressed.