23. The human fetal brain plasmid cDNA library (Invitrogen) was amplified in toto, and cesium-purified plasmid DNA was linearized at the Xba I site. In vitro transcription was performed from a T7 promoter as described (*18, 19*).
24. D. Eilat, M. Hochberg, R. Fischel, R. Laskov, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 3818 (1982).
25. The specific activity of [^{35}S]methionine used was 1275 Ci/mmol. Given the molecular weight (69 kD) and that nine methionines existed in FMRP (*7*), the expected number of disintegrations per minute per mole of protein was calculated and adjusted according to the predetermined efficiency of counting (69%) for the isotope. The number of moles of bound protein was determined by dividing the observed counts per minute of bound protein at saturation by the expected disintegrations per minute per mole; the concentration value given was determined by dividing by the reaction volume.
26. The standard binding equation used was $b = b_m([L]/K_d + [L])$, where b is the amount of protein bound, b_m is the maximum amount bound, [L] is the concentration of ligand, and K_d is the dissociation constant. Maximum amounts of bound protein were determined from direct scintillation counting of binding reactions performed in triplicate at constant amounts of ligand. Amounts of bound protein were determined either through direct scintillation counting or densitometric analysis of fluorograms. Data points were fit to this equation with the nonlinear least squares method furnished in the plotting program Delta Graph (Deltapoint, Inc., Monterey, CA), and equations were solved for the apparent K_d. A variation of this equation, $b = b_m[L]/K_d(1 + [I]/K_i) + [L]$ (where [I] = concentration of inhibitor), was used to fit the RNA competition curve and to solve for K_i.
27. M. Gribskov, R. Luethy, D. Eisenberg, *Methods Enzymol.* **183**, 146 (1989).
28. Biotinylated RNA binding assays were performed as described by Boelens *et al.* (*17*), except that magnetic beads with conjugated strepavidin (Dynal) were used. Captured material was resuspended in 20 μl of 1× SDS sample buffer, and bound protein was eluted by boiling for 10 min and resolved by SDS-PAGE. The 12% SDS-polyacrylamide gels were soaked for 30 min in destain solution (7.5% methanol and 10% acetic acid), followed by 1 hour in 1 M salicylate solution (1 M salicylate, 30% methanol, and 3.0% glycerol). Gels were then dried and exposed at −80°C with an intensifying screen for 24 to 72 hours.
29. Biotinylated DNA was prepared with a nick translation kit (Amersham) in the presence of biotinylated deoxythymidine triphosphate (dTTP) (BRL) at an equal molar ratio with dTTP and passed through a G-50 Sephadex spin column (Boehringer Mannheim) followed by precipitation. Single-strand DNA was obtained by heat denaturation of the product above.
30. Immunoprecipitations were carried out in the presence of 2 μl of in vitro–translated FMRP, D44 (20 ng/μl), and 80 ng of nonbiotinylated *FMR1* RNA by means of the same protocol as used in the biotinylated RNA binding assays (*17*). RNA molecules bound to antibody and FMRP were captured with protein A that had been linked to agarose beads (BRL; an amount equivalent to 20 μl). After centrifugation, pellets were washed as described (*17*), resuspended in 20 μl of 1× SDS buffer, and resolved by SDS-PAGE.
31. Supported by NIH grant HD20521 (S.T.W.). C.T.A. is a predoctoral fellow of the March of Dimes Birth Defects Foundation, and S.T.W. is an Investigator of the Howard Hughes Medical Institute.

21 July 1993; accepted 30 August 1993

A Yeast Protein Similar to Bacterial Two-Component Regulators

Irene M. Ota and Alexander Varshavsky*

Many bacterial signaling pathways involve a two-component design. In these pathways, a sensor kinase, when activated by a signal, phosphorylates its own histidine, which then serves as a phosphoryl donor to an aspartate in a response regulator protein. The Sln1 protein of the yeast *Saccharomyces cerevisiae* has sequence similarities to both the histidine kinase and the response regulator proteins of bacteria. A missense mutation in *SLN1* is lethal in the absence but not in the presence of the N-end rule pathway, a ubiquitin-dependent proteolytic system. The finding of *SLN1* demonstrates that a mode of signal transduction similar to the bacterial two-component design operates in eukaryotes as well.

In bacteria, a broad spectrum of responses to an often rapidly changing environment is mediated by mechanistically similar pathways known as two-component systems. The functions of two-component pathways include chemotaxis, sporulation, osmoregulation, transformation competence, virulence, and responses to changes in the sources of carbon, nitrogen, oxygen, and phosphorus (*1–3*). The sensor component of these systems is often an integral membrane protein containing a cytosolic transmitter domain that acts as a histidine-phosphorylating autokinase, when activated by a specific signal. The signal is sensed by a distinct input domain, often located in the periplasmic space. The histidine-linked phosphoryl group of an activated transmitter is transferred to an aspartate in the receiver domain of a response regulator protein, the second component of the pathway. Receiver phosphorylation regulates the activity of its output module, which is often a DNA-binding domain (*1–3*).

Although two-component pathways are common in bacteria, evidence for their existence in eukaryotes has been scarce. A protein kinase from rat mitochondria has sequence similarities to bacterial histidine kinases; however, in vitro it phosphorylates a Ser residue (*4*). Another eukaryotic candidate is phytochrome, a plant regulatory protein that has weak but potentially significant similarities to the sequences of bacterial histidine kinases (*5*). Histidine kinase activity has been detected in extracts from *S. cerevisiae* and other eukaryotic cells (*6*).

This report describes the *S. cerevisiae* *SLN1* gene which encodes a 134-kD product with strong sequence similarities to both the transmitter and receiver domains of the bacterial two-component regulators. We found *SLN1* while studying the N-end rule, a relation between the in vivo half-life of a protein and the identity of its N-terminal residue (*7, 8*). The N-end rule is a consequence of a set of ubiquitin-dependent degradation signal called N-degrons (*9*). Ubiquitin is a protein whose covalent conjugation to other proteins plays a role in a number of cellular processes, primarily through routes that involve protein degradation (*8*). In *S. cerevisiae*, the recognition component of the N-end rule pathway is encoded by the *UBR1* gene (*10*). A *ubr1Δ* mutant is viable, grows at nearly wild-type rates, but is unable to degrade N-end rule substrates, which are short-lived in wild-type (*UBR1*) cells (*8, 10*).

In a search for the functions of the N-end rule, we carried out a "synthetic lethal" screen to identify mutants whose viability requires the presence of the *UBR1* gene (*11*). To isolate such mutants, termed *sln* (*s*ynthetic *l*ethal of *N*-end rule), we used a screen based on 5-fluoroorotic acid (FOA) (*12*). In this method, yeast cells lacking chromosomal copies of both *URA3* and a (nonessential) gene of interest are transformed with a plasmid that expresses both of these genes. The cells are mutagenized and examined for growth on plates containing FOA and uracil. Because FOA selects against *URA3*-expressing cells (*12*), mutants that grow in the absence but not in the presence of FOA should include those whose viability requires the plasmid carrying the gene of interest linked to *URA3*.

A synthetic lethal screen with *UBR1* yielded a recessive mutant, *sln1-1*, which was viable in the *UBR1* but not in the *ubr1Δ* background (*11*). Our earlier attempts to clone *SLN1* yielded *PTP2*, which encodes a putative phosphotyrosine phosphatase, and is an extragenic multicopy suppressor of *sln1-1* (*11*). The *SLN1* gene was isolated as described (*13*). That the subcloned DNA fragment containing a single open reading frame (ORF) was indeed *SLN1* (rather than an extragenic suppressor of *sln1-1*) was confirmed by linkage analysis (*13*). The 3.66-kb ORF of *SLN1* encodes a

Division of Biology, California Institute of Technology, Pasadena, CA 91125.

*To whom correspondence should be addressed.

1220-residue protein (Fig. 1) with a molecular mass of 134 kD and a calculated *p*I (isoelectric point) of 7.5 (*14*). The codon adaptation index (*15*) of *SLN1* is 0.134, characteristic of weakly expressed yeast genes. The 3′ end of the *SLN1* ORF is 268 bp from the 3′ end of the convergently oriented *UBI1* gene, which encodes a fusion of ubiquitin to a ribosomal protein (*16*).

Comparisons of Sln1 to sequences in databases revealed extensive similarities to both the sensor and response regulator proteins of bacterial two-component systems (Fig. 1). These systems are organized in a variety of ways, including one in which the input and transmitter modules of the sensor, and the receiver module of the response regulator, are domains of a single polypeptide (*1*). Sln1 has this "input-transmitter-receiver" organization; moreover, the transmitter (histidine kinase) domain of Sln1

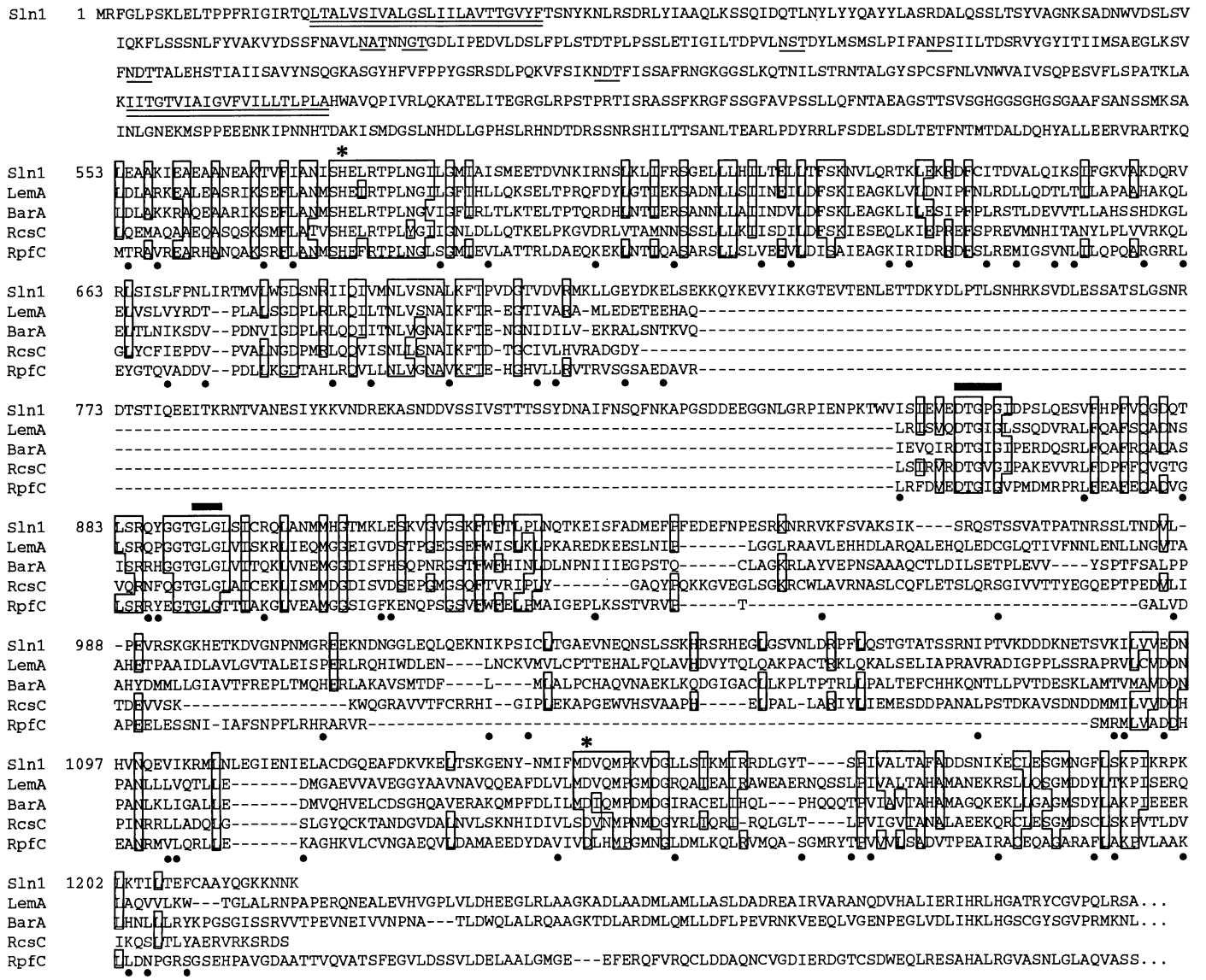

Fig. 1. Comparison of the yeast Sln1 and bacterial two-component regulators. The deduced sequence of the 1220-residue Sln1 is compared to portions of the deduced sequences of four bacterial proteins (LemA, BarA, RcsC, RpfC) in the PIR Protein Sequence database (*26*) that had the highest Sln1 similarity scores as defined by the BLAST program (National Center for Biotechnology Information, Bethesda, Maryland). These sequences were aligned to each other using the PileUp program (GCG package, version 7.2, Genetics Computer Group, Madison, Wisconsin). The presence of at least two identities to Sln1 at a given position is indicated by boxes. A conservative replacement (defined as follows: Arg ≈ Lys; Leu ≈ Ile ≈ Val; Tyr ≈ Phe; Asp ≈ Glu; Asn ≈ Gln; Ser ≈ Thr) in at least two of the sequences compared to Sln1 is indicated by a filled circle. Gaps were used to maximize alignment. The regions of Sln1 homologous to those of bacterial proteins include the putative histidine kinase domain (between residues ~546 and ~912) and the putative response regulator domain (between residues ~1085 and ~1220). Asterisks at His^{576} and Asp^{1144} of Sln1 indicate its putative phosphoryl group acceptors. Two potential membrane-spanning regions flanking the putative signal-sensing (input) domain of Sln1 are doubly underlined. Potential N-linked glycosylation sites between the putative membrane-spanning regions of Sln1 are singly underlined. The motif characteristic of ATP-binding proteins is indicated by two horizontal bars. The *barA* gene, which has homologies to sensor-regulator proteins, encodes a suppressor of an *envZ* deletion in *E. coli*; EnvZ is required for the adaptive response to changes in media osmolarity (*27*). The LemA protein is a two-component regulator in *Pseudomonas syringae*, a bacterial pathogen of beans (*28*). RcsC is a two-component regulator of capsule production in *E. coli* (*3*). RpfC is a two-component regulator in *Xanthomonas campestris*, a pathogen of cruciferous plants; RpfC regulates the production of extracellular enzymes such as the Prt protease (*29*). Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

(between residues ~546 and ~912) is most similar to those histidine kinase modules that reside in the similarly organized bacterial regulators (Fig. 1). This domain of Sln1 contains all of the residues conserved among bacterial histidine kinases, including the putative phosphoryl group acceptor His^{576} and the conserved Asn^{695}; the ~100-residue spacing between these residues is also characteristic of bacterial histidine kinases (*1*). Other conserved motifs in this domain include the sequences DTGPG (residues 859–863) and GLG (residues 891–893; Fig. 1), which fit the consensus DXGXG-[X_{15-50}]-GXG for glycine-rich loops characteristic of adenosine triphosphate (ATP)–binding proteins (*1, 2*).

The residues Glu^{1094}, Asp^{1095}, Asp^{1144}, and Lys^{1195} of Sln1 align with the conserved residues of bacterial response regulators (Fig. 1). In CheY, a response regulator of the bacterial chemotaxis pathway (*1*), the side chains of Asp^{12}, Asp^{13} and Asp^{57} form an active site pocket, with the ϵ-amino group of Lys^{109} positioned close to the carboxyl group of Asp^{57}, which is the acceptor of the phosphoryl group from the phosphorylated His residue of CheA, a histidine kinase (*1*). Inferring from sequence similarities, Asp^{1144} of Sln1 is the predicted site of phosphorylation, and the (phosphorylated) His^{576} of the Sln1 histidine kinase domain (Fig. 1) is the predicted donor of the phosphoryl group.

Another similarity between Sln1 and bacterial sensor proteins is the presence of two hydrophobic (potential) transmembrane segments in the N-terminal region of Sln1 (residues 23–46 and 334–354; Fig. 1). By analogy to the bacterial sensor proteins, the N-terminal region of Sln1 is likely to contain a membrane-bound signal-sensing domain which, if Sln1 is a plasma membrane protein, would be exposed on the surface of yeast cells. The Sln1 protein is also similar to bacterial proteins of the "input-transmitter-receiver" subclass in that Sln1 apparently lacks an output module; the receiver domain of Sln1 is close to its C-terminus (Fig. 1).

The method of gap repair (*17*) was used to locate the mutation in *sln1-1*. A transition G to A was detected at nucleotide position 2671 of the *SLN1* ORF that resulted in replacement of Gly with Asp at residue 891 (*18*), within the putative ATP-binding site of Sln1 (Fig. 1). Thus, *sln1-1* is a missense allele that encodes a protein with a presumably perturbed ATP-binding site. A null allele of *SLN1* (*sln1-Δ1::HIS3*) was constructed in vitro and used to replace the wild-type *SLN1* in a diploid [*his3/his3*] strain (*19*). Tetrad analysis of the resulting heterozygous [*sln1-Δ1::HIS3/SLN1*] His^+ diploid on rich (YPD) medium yielded two viable and two inviable segregants per tetrad, with the viable segregants always being His^-. The *sln1-Δ1::HIS3* spores germinated but formed microcolonies which did not progress beyond ~100 cells. However, when the sporulated [*sln1-Δ1::HIS3/SLN1*] diploid was dissected on poorer, synthetic media (SD), most tetrads yielded two normal (*SLN1*, His^-) and two slowly growing but viable (*sln1-Δ1::HIS3*, His^+) colonies per tetrad.

Similar results were obtained in the absence of sporulation and spore germination. *SLN1* was subcloned into a *URA3*-based vector (*20*), yielding pSLN1(URA3), which was transformed into the [*sln1-Δ1::HIS3/SLN1*] diploid. All of the His^+ segregants of this diploid (they contained the *sln1-Δ1::HIS3* allele) were found to be Ura^+ [that is, containing pSLN1(URA3)] when germinated on a rich (YPD) medium, indicating that *sln1Δ* segregants required the presence of the plasmid-borne *SLN1* for viability. However, [*sln1-Δ1::HIS3*, pSLN1(URA3)] cells grew (albeit very slowly) on FOA-containing SD plates, indicating that the loss of pSLN1(URA3) is compatible with cell viability on a poorer (SD) medium. The FOA-resistant *sln1-Δ1* colonies [which were Ura^-, indicating that they had lost pSLN1(URA3)] were retested on various media; they grew on SD but not on YPD plates, the latter result indicating that inviability of *sln1-Δ1* spores on YPD medium could not be caused exclusively by defects in germination. We conclude that *SLN1* is required for wild-type growth rates but not for cell viability on relatively poor (SD) media, and that it is essential for cell viability on a richer (YPD) medium.

The function of *SLN1* and the signal it senses are unknown. The apparent functional connections between Sln1, Ubr1, and Ptp2 remain to be understood as well. As *sln1-1* cells grow more slowly than congenic *SLN1* cells (*11*), and as *sln1Δ* cells are viable in poor but not in rich medium, Sln1 may sense a nutrient or another aspect of extracellular environment, and relay information about the medium to circuits that control the rate of cell growth and division. Given their impaired growth and hypersensitivity to rich medium, *sln1Δ* cells might be defective in adjusting the activity of a pathway that is normally regulated by Sln1 to the activities of other growth-related pathways whose inputs are also media-derived but are not regulated by Sln1. The resulting imbalance may be more severe in rich than in poor medium, consistent with the observed viability pattern of *sln1Δ* cells.

The identification of eukaryotic counterparts of bacterial two-component signal transduction systems in fungi (Fig. 1) and plants (*21*) suggests that mechanisms of this type may prove to be as important among nucleated cells as they are known to be in prokaryotes.

REFERENCES AND NOTES

1. R. B. Bourret, K. A. Borkovich, M. I. Simon, *Annu. Rev. Biochem.* **60**, 401 (1991); J. B. Stock, G. S. Lukat, A. M. Stock, *Annu. Rev. Biophys. Biophys. Chem.* **20**, 109 (1991); L. M. Albright, E. Huala, F. M. Ausubel, *Annu. Rev. Genet.* **23**, 311 (1989); J. S. Parkinson and E. C. Kofoid, *ibid.* **26**, 71 (1992); G. L. Hazelbauer, H. C. Berg, P. Matsumura, *Cell* **73**, 15 (1993).
2. D. E. Koshland, Jr., *Biochemistry* **27**, 5829 (1988); B. Magasanik, *Trends Biochem. Sci.* **13**, 475 (1988); E. C. C. Lin and S. Iuchi, *Annu. Rev. Genet.* **25**, 361 (1991); B. L. Wanner, *J. Bacteriol.* **174**, 2053 (1992); R. Gross, B. Aricò, R. Rappuoli, *Mol. Microbiol.* **3**, 1661 (1989).
3. S. Gottesman and V. Stout, *Mol. Microbiol.* **5**, 1599 (1991).
4. K. M. Popov, Y. Zhao, Y. Shimomura, M. J. Kuntz, R. A. Harris, *J. Biol. Chem.* **267**, 13127 (1992).
5. H. A. W. Schneider-Poetsch, *Photochem. Photobiol.* **56**, 839 (1992).
6. J. Huang, Y. Wei, Y. Kim, L. Osterberg, H. R. Matthews, *J. Biol. Chem.* **266**, 9023 (1991).
7. A. Bachmair, D. Finley, A. Varshavsky, *Science* **234**, 179 (1986); T. Shrader, J. W. Tobias, A. Varshavsky, *J. Bacteriol.* **175**, 4364 (1993).
8. A. Varshavsky, *Cell* **69**, 725 (1992).
9. A. Bachmair and A. Varshavsky, *ibid.* **56**, 1019 (1989); E. S. Johnson, D. K. Gonda, A. Varshavsky, *Nature* **346**, 287 (1990).
10. B. Bartel, I. Wünning, A. Varshavsky, *EMBO J.* **9**, 3179 (1990); K. Madura, R. J. Dohmen, A. Varshavsky, *J. Biol. Chem.* **268**, 12046 (1993).
11. I. M. Ota and A. Varshavsky, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2355 (1992).
12. M. E. Basson, R. L. Moore, J. O'Rear, J. Rine, *Genetics* **117**, 645 (1987); J. D. Boeke, F. LaCroute, G. R. Fink, *Mol. Gen. Genet.* **197**, 345 (1984).
13. The *S. cerevisiae* genomic DNA library that yielded *SLN1* was carried in the *LEU2*, *CEN4*-containing vector p366 (*22*). Strain IOY7 [*sln1-1 ubr1Δ::HIS3*, pYULU] was transformed with the above library; Leu^+ transformants were selected and tested for FOA resistance as described (*11*). IOY7 differed from the original *sln1-1* strain IOY1 [*sln1-1 ubr1Δ::LEU2*, pUBR1] (*11*) in that it contained the *HIS3*-based null allele of *UBR1* and pYULU instead of pUBR1. pYULU is a derivative of YCplac33 (*20*) that bears *UBR1* and *LYS2*. All ~20 plasmids that conferred FOA resistance contained the same complementing DNA fragment (*SLN1*) in identical or overlapping inserts. The ~5.4-kb Eco RI–Mlu I fragment from one of the complementing plasmids (pG16) was subcloned into YCplac111 (*20*). This fragment (whose Eco RI–produced end was located in the vector, ~100 bp from the 5′ end of the putative *SLN1* ORF) also complemented *sln1-1*, and in addition lacked the adjacent *UBI1* gene (*16*). That the putative *SLN1* is not an extragenic suppressor of *sln1-1* was confirmed by demonstrating its tight linkage to the *sln1-1* locus. A ~3.3-kb Xba I fragment of p2a1 (another *sln1-1*–complementing plasmid) containing a 3′ proximal region of the putative *SLN1*, and *SLN1-UBI1* intergenic region, and a 5′ proximal region of *UBI1* (*16*) was subcloned into Xba I–cut YCplac111. A Sal I–Bam HI fragment of the resulting plasmid that contained the initial Xba I fragment was then subcloned into Sal I–Bam HI–cut pRS304, a *TRP1*-based yeast integration vector (*23*), yielding pSLNK. The latter was cut at the unique Pst I site within the putative *SLN1* ORF, and was then introduced into BBY48 [*MATα SLN1 UBR1 trp1-1*] (*10*). Trp^+ transformants were selected, and the integration of Pst I–cut pSLNK at the putative *SLN1* locus was confirmed by Southern hybridization. The resulting strain, IOY301, containing the putative *SLN1* tightly linked to *TRP1*, was crossed to IOY3 [*MATa sln1-1 UBR1 trp1 ura3 his3 leu2 lys2*; produced as described (*11*)], yielding IOY305 [*MATa/MATα sln1-1/SLN1::pSLNK(TRP1)/trp1 UBR1/UBR1*], which was subjected to tetrad analysis. A haploid [*sln1-1 UBR1*] strain can be identified by its small-colony phenotype (*11*). If the putative (cloned) *SLN1* gene maps to the locus of the

sln1-1 mutation, the (slowly growing) *sln1-1* segregants of IOY305 must almost always be Trp$^-$, whereas the (normally growing) *SLN1* segregants must almost always be Trp$^+$ (bearing the *TRP1* gene tightly linked to *SLN1*). Nineteen tetrads were dissected, and all of the large-colony (*SLN1*) segregants were found to be Trp$^+$, whereas all of the small-colony (*sln1-1*) segregants were Trp$^-$, indicating that the cloned ORF was indeed the *SLN1* gene.
14. Nucleotide sequence of the *SLN1* locus determined in this work encompassed 428 bp upstream of the (inferred) start codon of *SLN1*, the *SLN1* ORF, and 320 bp downstream of the *SLN1* stop codon. This sequence, determined on both strands with the use of the chain termination method and standard strategies (*24*), has been submitted to GenBank; its accession number is U01835. Hybridization of the ^{32}P-labeled ~4-kb Xba I fragment that contained the 5′ flanking and 5′ proximal regions of *SLN1* to the blot of *S. cerevisiae* chromosomes fractionated by pulse-field electrophoresis (Clontech) localized *SLN1* to Chromosome IX. More detailed mapping, with the use of an ordered set of lambda phages containing *S. cerevisiae* genomic DNA inserts [Prime Clone Blots; a gift from L. Riles and M. Olson] positioned *SLN1* at the left arm of Chromosome IX, between *SUC2* and *HIS5*.
15. P. M. Sharp and W. H. Li, *Nucleic Acids Res.* **15**, 1281 (1987).
16. E. Özkaynak, D. Finley, M. J. Solomon, A. Varshavsky, *EMBO J.* **6**, 1429 (1987); D. Finley, B. Bartel, A. Varshavsky, *Nature* **338**, 394 (1989).
17. R. Rothstein, *Methods Enzymol.* **194**, 281 (1991).
18. Yeast DNA inserts of the library-derived, *CEN4*, *LEU2*-based, *SLN1*-containing plasmids that complemented *sln1-1* were cut at different restriction sites within and beyond the *SLN1* ORF, and the resulting linear ("gapped") (*17*) fragments were introduced into IOY7 [*sln1-1 ubr1Δ::HIS3 leu2*, pYULU]. Leu$^+$ transformants were selected and tested for their sensitivity to FOA (that is, for their inability to lose the *URA3*, *UBR1*-expressing pYULU plasmid without losing viability). FOA-sensitive Leu$^+$ transformants would be expected to have been transformed with plasmids whose "gap" encompassed the mutation-containing region of *sln1-1*. The gap-repaired plasmids (*17*) were isolated and sequenced at the gap-encompassing regions. This analysis detected a G-to-A transition at nucleotide position 2671 that resulted in a Gly-to-Asp alteration at residue 891 of Sln1.
19. The *sln1-Δ1::HIS3* deletion allele was constructed by subcloning into pUC19 (*24*) the 655-bp Eco RI–Bam HI fragment containing the 5′ flanking sequence of *SLN1* and the first 527 bp of the *SLN1* ORF, as well as the 452-bp Bam HI–Sal I fragment containing the last 40 bp of the *SLN1* ORF, the *SLN1-UBI1* intergenic region, and part of the *UBI1* ORF [this fragment was produced with the use of the polymerase chain reaction (PCR) and standard protocols (*24*)], yielding pSLN-21. The ~1.8-kb *HIS3*-containing Bam HI fragment of YEp6 (*25*) was subcloned into the Bam HI site of the pSLN-21 insert, yielding pSLN-22. The ~2.9-kb Eco RI–Sal I fragment of pSLN-22 was introduced into the diploid DF5 [*his3/his3*] (*16*), and His$^+$ transformants were selected; Southern hybridization was used to confirm the replacement (*17*) of one of two copies of *SLN1* with *sln1-Δ1::HIS3*.
20. R. D. Gietz and A. Sugino, *Gene* 74, 527 (1988).
21. C. Chang, S. F. Kwok, A. B. Bleeker, E. M. Meyerowitz, *Science* **262**, 539 (1993); L. Alex and M. Simon, personal communication.
22. The library was constructed by F. Spencer and P. Hieter; it was purchased from the American Type Culture Collection, Rockville, MD.
23. R. S. Sikorski and P. Hieter, *Genetics* **122**, 19 (1989).
24. F. M. Ausubel *et al.*, *Current Protocols in Molecular Biology* (Wiley-Interscience, New York, 1988).
25. K. Struhl, D. T. Stinchcomb, S. Scherer, R. W. Davis, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 1035 (1979).
26. W. C. Barker, D. G. George, H. W. Mewes, A. Tsugita, *Nucleic Acids Res.* **20**, 2023 (1992).
27. S. Nagasawa, H. Tokishita, H. Aiba, T. Mizuno, *Mol. Microbiol.* **6**, 799 (1992).
28. E. M. Hrabak and D. K. Willis, *J. Bacteriol.* **174**, 3011 (1992).
29. J. L. Tang *et al.*, *Mol. Gen. Genet.* **226**, 409 (1991).
30. We thank E. Meyerowitz, C. Chang, M. Simon, and L. Alex for sharing unpublished results; L. Riles for Prime Clone Blots; members of the laboratory for helpful discussions and advice; and C. Byrd, T. Clandinin, E. Johnson, N. Johnsson, F. Lévy, and K. Madura for comments on the manuscript. Supported by grants to A.V. from the National Institutes of Health.

20 July 1993; accepted 22 September 1993

The Conserved Pre-mRNA Splicing Factor U2AF from *Drosophila*: Requirement for Viability

Roland Kanaar, Siobhan E. Roche, Eileen L. Beall, Michael R. Green, Donald C. Rio*

The large subunit of the human pre–messenger RNA splicing factor U2 small nuclear ribonucleoprotein auxiliary factor (hU2AF^{65}) is required for spliceosome assembly in vitro. A complementary DNA clone encoding the large subunit of *Drosophila* U2AF (dU2AF^{50}) has been isolated. The dU2AF^{50} protein is closely related to its mammalian counterpart and contains three carboxyl-terminal ribonucleoprotein consensus sequence RNA binding domains and an amino-terminal arginine- and serine-rich (R/S) domain. Recombinant dU2AF^{50} protein complements mammalian splicing extracts depleted of U2AF activity. Germline transformation of *Drosophila* with the dU2AF^{50} complementary DNA rescues a lethal mutation, establishing that the dU2AF^{50} gene is essential for viability. R/S domains have been found in numerous metazoan splicing factors, but their function is unknown. The mutation in *Drosophila* U2AF will allow in vivo analysis of a conserved R/S domain–containing general splicing factor.

Generation of functional mRNAs in eukaryotes requires removal of noncoding regions (introns) from pre-mRNAs by a process termed RNA splicing (*1*). Pre-mRNA splicing takes place in the spliceosome, a dynamic RNA-protein complex that assembles in a stepwise adenosine triphosphate (ATP)–dependent manner on the pre-mRNA (*1*). The spliceosome is composed of small nuclear ribonucleoprotein (snRNP) particles and extrinsic RNA binding proteins. Studies with human cell (HeLa) nuclear splicing extracts have shown that the targeting of U2 snRNP to the branch site on the pre-mRNA requires a protein factor called U2 auxiliary factor (U2AF) (*1*). U2AF binds site-specifically to the polypyrimidine tract located between the branch site and the 3′ splice site of the pre-mRNA (*2*). Human U2AF consists of two associated polypeptides with apparent molecular sizes of 65 and 35 kD, termed hU2AF^{65} and hU2AF^{35}, respectively (*3*). Splicing activity of HeLa cell nuclear extracts depleted of U2AF activity can be restored by the addition of hU2AF^{65} (*3*). Sequence analysis of the complementary DNA (cDNA) encoding hU2AF^{65}, in conjunction with biochemical experiments, have revealed that hU2AF^{65} contains three COOH-terminal ribonucleoprotein consensus sequence (RNP-CS) domains that mediate RNA binding and an NH_2-terminal arginine- and serine-rich (R/S) domain that is essential for splicing activity in vitro (*2*). Biochemical and immunochemical assays have shown that U2AF activity and antigenically related proteins of 50 and 38 kD, corresponding to hU2AF^{65} and hU2AF^{35}, respectively, are present in the fruitfly *Drosophila melanogaster* (*3*).

We isolated a cDNA encoding the *Drosophila* homolog of hU2AF^{65} by screening a 4- to 8-hour *Drosophila* embryonic cDNA library under low-stringency conditions with a radiolabeled hU2AF^{65} cDNA fragment (*4*). The dU2AF^{50} cDNA contains a 1248–base pair open reading frame predicted to encode a 416–amino acid protein with a calculated molecular size of 46,652 daltons. The predicted amino acid sequence of dU2AF^{50} indicates that it contains functional domains similar to those found in hU2AF^{65} (Fig. 1). Amino acids 93 to 408 contain three RNP-CS domains and amino acids 6 to 41 contain an R/S domain. Comparison of the *Drosophila* and mammalian proteins shows that the amino acid identity is greatest within the RNP-CS RNA binding domains (80% identity and 90% similarity). Optimal alignment of the amino acid sequences requires that gaps and insertions be added in the regions connecting the RNP-CS domains, but not within

R. Kanaar, S. E. Roche, E. L. Beall, D. C. Rio, Department of Molecular and Cell Biology, Division of Genetics and Division of Biochemistry and Molecular Biology, University of California, Berkeley, CA 94720.
M. R. Green, Program in Molecular Medicine, University of Massachusetts Medical Center, Worcester, MA 01605.

*To whom correspondence should be addressed.