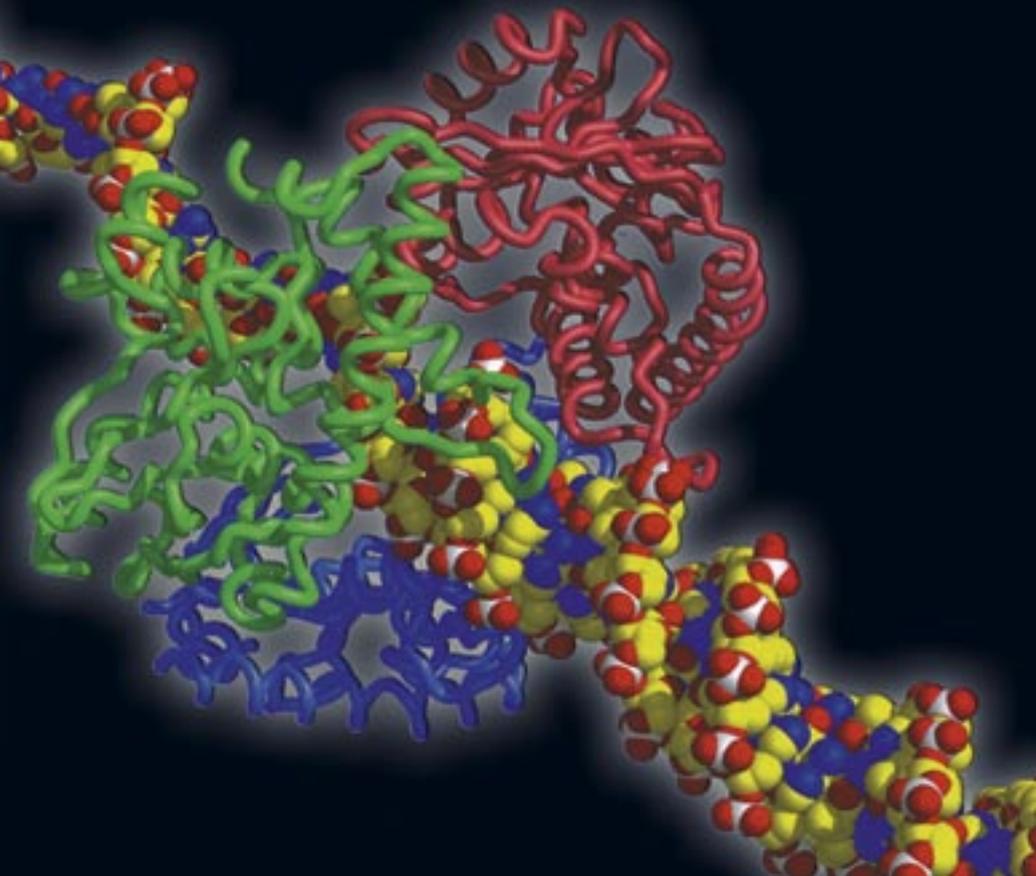


Mutagenesis Techniques collection



This booklet is brought to you by the AAAS/Science Business Office

Sponsored by



Over 20 Years of Innovation

As a global leader in the development of innovative technologies, our products simplify, accelerate, and improve drug discovery and drug development research. Since 1984, our reagents, kits, software, and instrumentation systems have been used in pharmaceutical and biotechnology laboratories worldwide to determine the molecular mechanisms of health and disease as well as to search for new drug therapies and diagnostic tests.

To order a copy of our current product catalog, visit www.stratagene.com/catalog.

AMPLIFICATION

CELL BIOLOGY

CLONING

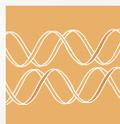
MICROARRAYS

NUCLEIC ACID ANALYSIS

PROTEIN FUNCTION & ANALYSIS

QUANTITATIVE PCR

SOFTWARE SOLUTIONS



www.stratagene.com

Stratagene USA and Canada

Order: (800) 424-5444 x3

Technical Services: (800) 894-1304 x2

Stratagene Europe

Order: 00800-7000-7000

Technical Services: 00800-7400-7400

Stratagene Japan K.K.

Order: 03-5159-2060

Technical Services: 03-5159-2070



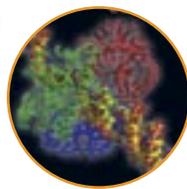
STRATAGENE

An Agilent Technologies Company

- 2 Introductions:
Bringing About Change**
Sean Sanders
- The Power of Mutagenesis**
Patricia Sardina
- 4 Bypassing a Kinase Activity with an ATP-Competitive Drug**
Feroz R. Papa, Chao Zhang, Kevan Shokat, Peter Walter
Science 28 November 2003 **302**: 1533-1537
- 12 Human Catechol-O-Methyltransferase Haplotypes Modulate Protein Expression by Altering mRNA Secondary Structure**
A. G. Nackley, S.A. Shabalina, I. E. Tchivileva, K. Satterfield, O. Korchynskyi, S. S. Makarov, W. Maixner, L. Diatchenko
Science 22 December 2006 **314**: 1930-1933
- 17 An mRNA Surveillance Mechanism That Eliminates Transcripts Lacking Termination Codons**
Pamela A. Frischmeyer, Ambro van Hoof, Kathryn O'Donnell, Anthony L. Guerriero, Roy Parker, Harry C. Dietz
Science 22 March 2002 **295**: 2258-2261
- 24 Direct Demonstration of an Adaptive Constraint**
Stephen P. Miller, Mark Lunzer, Antony M. Dean
Science 20 October 2006 **314**: 458-461
- 30 Technical Notes: Large Insertions: Two Simple Steps Using Quikchange® II Site-Directed Mutagenesis Kits**

About the Cover:

γ -Exonuclease binds to an end of double-stranded DNA and degrades one of the strands in a highly processive manner. The structure of the exonuclease consists of a toroidal trimer (~94 angstroms in diameter) and is presumed to enclose its substrate in the manner illustrated. [Image from the cover of *Science* **277** (19 September 1997)]



Design and Layout: Amy Hardcastle
Copy Editor: Robert Buck

Bringing About Change

This new method of mutagenesis has considerable potential in genetic studies.... [S]pecific mutation within protein structural genes will allow the production of proteins with specific amino acid changes. This will allow precise studies of structure-function relationships, for example the role of specific amino acids in enzyme active sites, and the more convenient production of proteins whose action is species specific, for example the conversion of a cloned gene coding for a lower vertebrate hormone to that for a human.

— closing paragraph of the seminal paper from Michael Smith's lab first describing site-directed mutagenesis¹

The art of life lies in a constant readjustment to our surroundings.

— Okakura Kakuzō, Japanese scholar

It is a truism that, to survive, we must be able to change. This, too, is the essence of evolution on the most fundamental biological level: without the innate ability for DNA to undergo mutation, it would not be possible for organisms to adapt and survive changes in their environment. Since humankind achieved an understanding of this mutagenesis process and its power, we have been trying to control it and bend it to our will.

First described in 1978 by Michael Smith at the University of British Columbia in Vancouver,¹ site-directed mutagenesis (SDM) has become an invaluable tool for molecular biologists. Originally named oligonucleotide-directed mutagenesis, the first experiments made use of short, 12-nucleotide strands of synthetic DNA containing a mismatch to an intrinsic reporter gene of the bacteriophage Φ X174. Using these oligonucleotides, together with DNA polymerase I from *E. coli* and T4 DNA ligase, Smith and his team demonstrated that it was possible to introduce a permanent mutation in the circular DNA of the phage, resulting in a phenotypic change. The prescient quote above shows the authors' understanding of the manifold applications for the new methodology.

Since its development, a number of modifications and improvements have been made to the SDM technique. The greatest advance came with the invention of the polymerase chain reaction (PCR) in 1983. Since then, SDM and PCR have been inextricably linked, a circumstance reflected in the 1993 Nobel in Chemistry being shared by Kary Mullis, the pioneer of PCR, and Michael Smith. This advance made SDM both quicker and more flexible and has enabled it to become a standard and necessary technique employed in labs worldwide.

In this booklet, we bring together a collection of influential papers that all depended on SDM for their success. As one of the most used techniques in molecular biology, it plays an often unacknowledged—but essential—role in many research projects. There is little doubt that future refinements and improvements will be made to the technique, allowing for an even broader range of applications. As these modifications are commercialized and made available to all scientists, new and innovative uses for them will be discovered, facilitating enrichment of our knowledge in many areas of research, from basic cellular processes to complex diseases such as neurodegenerative disorders and cancer.

Sean Sanders, Ph.D.
Commercial Editor, *Science*

¹ C. A. Hutchison III, *et al.* Mutagenesis at a specific position in a DNA sequence. *J. Biol. Chem.* **253**:6551-6560 (1978).

The Power of Mutagenesis

In vitro mutagenesis is a very powerful tool for studying protein structure-function relationships, altering protein activity, and for modifying vector sequences to incorporate affinity tags and correct frame shift errors. A variety of mutagenesis techniques have been developed over the past decade that allow researchers to generate point mutations, modify (insert, delete or replace) one or more codons, swap domains between related gene sequences, and create diverse collections of mutant clones.

Mutagenesis strategies can be divided into two main types, random or site-directed. With random mutagenesis, point mutations are introduced at random positions in a gene-of-interest, typically through PCR employing an error-prone DNA polymerase (error-prone PCR). Randomized sequences are then cloned into a suitable expression vector, and the resulting mutant libraries can be screened to identify mutants with altered or improved properties. By correlating changes in function to alterations in protein sequence, researchers can create a preliminary map of protein functional domains.

Site-directed mutagenesis is the method of choice for altering a gene or vector sequence at a selected location. Point mutations, insertions, or deletions are introduced by incorporating primers containing the desired modification(s) with a DNA polymerase in an amplification reaction. In more complex protein engineering experiments, researchers can design mutagenic primers to incorporate degenerate codons (site-saturation mutagenesis).

Stratagene, now an Agilent Technologies company, is a worldwide leader in developing innovative products and technologies for life science research. Our expertise in enzyme engineering has translated into the most efficient and easy-to-use mutagenesis kits available today. Our QuikChange® Site-Directed Mutagenesis products provide the most reliable and accurate method for performing targeted mutagenesis, saving time and valuable resources. These products have been cited in thousands of publications, including the papers featured in this booklet. In addition, our GeneMorph® Random Mutagenesis products offer an easy, robust method for creating diverse mutant collections.

Our mutagenesis product offering has been substantially enhanced by our newest product introduction, the QuikChange® Lightning Site-Directed Mutagenesis Kit. This kit offers significant time-savings by incorporating new, exclusive fast enzymes which allow researchers to shorten the duration of the thermal cycling and digestion steps while maintaining ultra-high fidelity and mutation efficiency.

The efficacy of mutagenesis as a tool for studying protein function and structure may lead to new experimental therapeutic approaches to diseases such as cancer. Precise mapping of protein-protein interactions may ultimately lead to a greater understanding of disease mechanics. This *Science* Mutagenesis Techniques Collection booklet illustrates the utility of mutagenesis in a wide variety of research studies. We are pleased to have the opportunity to sponsor this publication.

Patricia Sardina
Product Manager
Stratagene, An Agilent Technologies Company

Bypassing a Kinase Activity with an ATP-Competitive Drug

Feroz R. Papa,^{1,3*} Chao Zhang,² Kevan Shokat,² Peter Walter^{3,4}

Unfolded proteins in the endoplasmic reticulum cause trans-autophosphorylation of the bifunctional transmembrane kinase Ire1, which induces its endoribonuclease activity. The endoribonuclease initiates nonconventional splicing of *HAC1* messenger RNA to trigger the unfolded protein response (UPR). We explored the role of Ire1's kinase domain by sensitizing it through site-directed mutagenesis to the ATP-competitive inhibitor 1NM-PP1. Paradoxically, rather than being inhibited by 1NM-PP1, drug-sensitized Ire1 mutants required 1NM-PP1 as a cofactor for activation. In the presence of 1NM-PP1, drug-sensitized Ire1 bypassed mutations that inactivate its kinase activity and induced a full UPR. Thus, rather than through phosphorylation per se, a conformational change in the kinase domain triggered by occupancy of the active site with a ligand leads to activation of all known downstream functions.

Secretory and transmembrane proteins traversing the endoplasmic reticulum (ER) during their biogenesis fold to their native states in this compartment (1). This process is facilitated by a plethora of ER-resident activities (the protein folding machinery) (2). Insufficient protein folding capacity causes an accumulation of unfolded proteins in the ER lumen (a condition referred to as ER stress) and triggers a transcriptional program called the UPR. In yeast, UPR targets include genes encoding chaperones, oxido-reductases, phospholipid biosynthetic enzymes, ER-associated degradation components, and proteins functioning downstream in the secretory pathway (3). Together, UPR target activities afford proteins passing through the ER an extended opportunity to fold and assemble properly, dispose of unsalvageable unfolded polypeptides, and increase the capacity for ER export.

The yeast UPR is signaled through the ER stress sensor Ire1, a single-spanning ER transmembrane protein with three

functional domains (4). The most N-terminal domain, which resides in the ER lumen, senses elevated levels of unfolded ER proteins. Dissociation of ER chaperones from Ire1 as they become engaged with unfolded proteins is thought to trigger Ire1 activation (3).

Ire1's most C-terminal domain is a regulated endoribonuclease (RNase), which has a single known substrate in yeast: the *HAC1^u* mRNA (u stands for uninduced), which encodes the Hac1 transcriptional activator necessary for activation of UPR targets (5). *HAC1^u* mRNA is constitutively transcribed, but not translated, because it contains a nonconventional translation-inhibitory intron (6). Upon activation, Ire1's RNase cleaves *HAC1^u* mRNA at two specific sites, excising the intron (7). The 5' and 3' exons are rejoined by tRNA ligase (8), resulting in spliced *HAC1ⁱ* mRNA (i stands for induced), which is actively translated to produce the Hac1 transcriptional activator, in turn upregulating UPR target genes (5).

A functional kinase domain precedes the RNase domain on the cytosolic side of the ER membrane (9). Activation of Ire1 leads to its oligomerization in the ER membrane, followed by trans-autophosphorylation (9, 10). Mutations of catalytically essential kinase active-site

¹Department of Medicine, ²Department of Cellular and Molecular Pharmacology, ³Department of Biochemistry and Biophysics, ⁴Howard Hughes Medical Institute, University of California, San Francisco, CA 94143-2200, USA.

*To whom correspondence should be addressed. E-mail: frpapa@medicine.ucsf.edu

residues or mutations of residues known to become phosphorylated prevent *HAC1*^u mRNA splicing and abrogate UPR signaling, demonstrating that Ire1's kinase phosphotransfer function is essential for RNase activation (4, 9, 11). Thus, Ire1 communicates an unfolded-protein signal from the ER to the cytosol using the luminal domain as the sensor and the RNase domain as the effector. Although clearly important for the circuitry of this machine, it is unknown why and how Ire1's kinase is required for activation of its RNase. To dissect the role of Ire1's kinase function, we used a recently developed strategy that allows us to sensitize Ire1 to specific kinase inhibitors (12).

Unexpected behavior of Ire1 mutants sensitized to an ATP-competitive drug. The mutation of Leu⁷⁴⁵—situated at a conserved position in the adenosine 5'-triphosphate (ATP)-binding site—to Ala or Gly is predicted to sensitize Ire1 to the ATP-competitive drug 1-tert-butyl-3-naphthalen-1-ylmethyl-1*H*-pyrazolo[3,4-*d*]pyrimidin-4-ylamine (1NM-PP1) by creating an enlarged active-site pocket not found in any wild-type kinase (13). Most kinases are not affected by these mutations, but others are partially or severely impaired (14, 15). For Ire1, these substitutions markedly decreased UPR signaling (assayed through an *in vivo* reporter), which is indicative of decreased kinase activity. Ire1(L⁷⁴⁵→A⁷⁴⁵) [Ire1(L745A) (16)] showed a 40% decrease relative to the wild-type kinase, whereas Ire1(L745G) showed a >90% decrease, approaching that of the *Δire1* control (Fig. 1A)

Unexpectedly, the addition of 1NM-PP1 to cells expressing partially active Ire1(L745A) caused no inhibition, even at concentrations that completely inhibit other 1NM-PP1-sensitized kinases (15, 17). Instead, 1NM-PP1 increased reporter activity slightly (Fig. 1A, bar 4). To our further surprise, 1NM-PP1 provided significantly restored signaling to the severely crippled Ire1(L745G) (Fig. 1A, bar 6).

This result presents a paradox: An ATP-competitive compound that was expected to function as an inhibitor of the rationally engineered mutant enzyme instead permits activation.

UPR activation strictly required induction by dithiothreitol (DTT) (Fig. 1) or tunicamycin (18); 1NM-PP1 by itself had no effect, indicating that it does not induce the UPR by directly affecting ER protein folding. Moreover, the effects of 1NM-PP1 were specific: other structurally similar compounds, 2-naphthylmethyl PP1, which is a regio-isomer, and 1-naphthyl PP1, which lacks the methylene group between the heterocyclic ring and the naphthyl group, neither activated nor inhibited wild-type or mutant Ire1 (18).

Activation rather than inhibition by an ATP-competitive inhibitor raised the question of whether the kinase activity is necessary in the 1NM-PP1-sensitized mutants. To this end, we combined the L745A or L745G mutation with a "kinase-dead" variant D828A, which lacks an active-site Asp residue needed to chelate Mg²⁺ (11, 19). Whereas the Ire1(L745A,D828A) double mutant was detectable in cells at levels near those of wild-type cells, Ire1(L745G,D828A) was undetectable, probably because of instability, and could not be assayed. As expected, Ire1(L745A,D828A) was unable to induce the UPR reporter with DTT alone, but the addition of 1NM-PP1 allowed activation to levels approaching 80% of the wild-type activation level (Fig. 1B). This suggests that the binding of 1NM-PP1 to 1NM-PP1-sensitized Ire1 renders the kinase activity dispensable.

Trans-autophosphorylation by Ire1 of two activation-segment Ser residues (S⁸⁴⁰ and S⁸⁴¹) is necessary for signaling (9), prompting the question of whether this requirement is also bypassed in 1NM-PP1-sensitized mutants. As expected, the unphosphorylatable Ire1 (S840A,S841A) mutant was inactive in the absence and presence of 1NM-PP1 (Fig. 1C). In

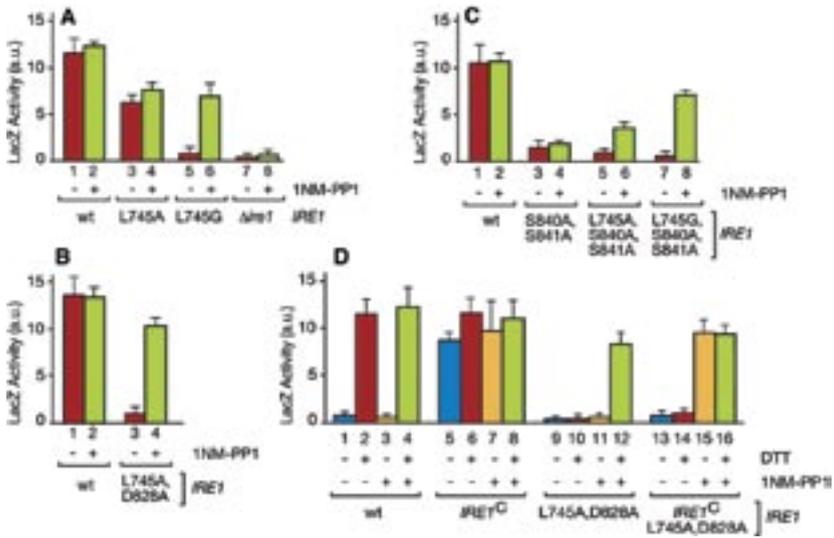


Fig. 1. (A to D) In vivo UPR assays using a UPR element-driven *lacZ* reporter. The relevant genotypes and the presence or absence of 1NM-PP1 and the UPR inducer DTT are indicated. For (A) to (C), DTT was used in all assays. wt, wild type; a.u., arbitrary units.

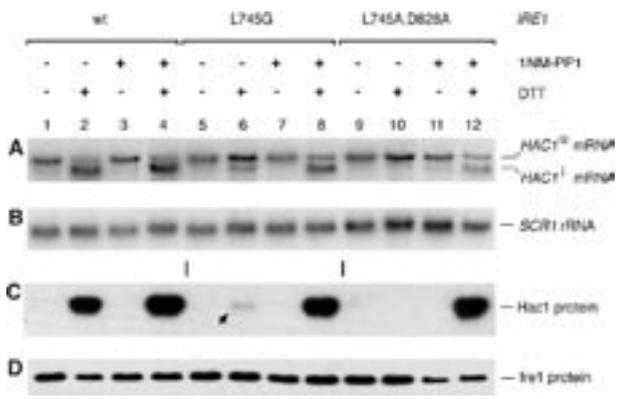


Fig. 2. In vivo *HAC1* mRNA splicing and Hac1 protein production. Northern blots showing (A) in vivo *HAC1* mRNA splicing and (B) *SCR1* ribosomal RNA (rRNA) as loading control. Western blots showing (C) Hac1 and (D) Ire1 protein levels. The arrow in (C), lane 6, calls attention to the small amount of Hac1 protein present in these cells.

contrast, the 1NM-PP1-sensitized Ire1 (L745A,S840A,S841A) and Ire1(L745G, S840A,S841A) mutants were significantly activated by 1NM-PP1, indicating that phosphorylation of S⁸⁴⁰ and S⁸⁴¹ can be bypassed. Two trends are apparent: With DTT alone, relative to the parent mutant Ire1(S840A,S841A), signaling decreased as the side chain of residue 745 was progressively reduced from Leu to Ala to Gly (Fig. 1C, bars 3, 5, and 7). This is consistent with the results of Fig. 1A, which

show an increasing sensitivity of the active-site to side-chain reduction at residue 745. Reciprocally, in the presence of 1NM-PP1, activation of 1NM-PP1-sensitized Ire1(S840A,S841A) mutants increased in the same order (Fig. 1C, bars 4, 6, and 8), perhaps because of progressively enhanced binding by 1NM-PP1 as the residue 745 side chain was trimmed back

1NM-PP1-sensitized Ire1 requires 1NM-PP1 as a cofactor for HAC1 mRNA splicing. We next confirmed that activa-

tion by 1NM-PP1-sensitized Ire1 mutants followed the expected pathway of activating *HAC1* mRNA splicing and Hac1 protein production. As expected, UPR induction of wild-type cells with DTT caused splicing of *HAC1* mRNA, as revealed by near-complete conversion of *HAC1*^u to *HAC1*^l mRNA (Fig. 2A, lanes 1 and 2) (5); 1NM-PP1 had no effect by itself or with DTT (Fig. 2A, lanes 3 and 4). In strict correlation with *HAC1* mRNA splicing, Hac1 protein accumulated significantly (Fig. 2C, lanes 2 and 4)

In contrast, cells expressing 1NM-PP1-sensitized Ire1(L745G) spliced *HAC1* mRNA weakly with DTT alone, but displayed a significant increase when 1NM-PP1 was also provided (Fig. 2A, lanes 5 to 8). 1NM-PP1 alone did not activate *HAC1* mRNA splicing. As expected, Hac1 protein levels correlated with splicing (e.g., Fig. 2C, lanes 6 and 8). Activation by 1NM-PP1 was even more pronounced in cells expressing the Ire1(L745A,D828A) double mutant, which neither spliced *HAC1* mRNA nor produced Hac1 protein with DTT alone, but exhibited robust splicing and Hac1 protein production when provided both 1NM-PP1 and DTT (Fig. 2, A and C, compare lanes 10 and 12).

Importantly, steady-state levels of Ire1 proteins were comparable in all experiments, excluding as a trivial explanation for the observed effects the possibility that 1NM-PP1-sensitized Ire1 mutants are only stably expressed in the presence of 1NM-PP1 (Fig. 2D). Taken together, our data show that 1NM-PP1 permits but does not instruct 1NM-PP1-sensitized Ire1 mutants to splice *HAC1* mRNA in response to ER stress. Thus, in the genetic background of 1NM-PP1-sensitized *IRE1* alleles, 1NM-PP1 is, in effect, a cofactor for signaling in the UPR

1NM-PP1 uncouples the kinase and RNase activities of 1NM-PP1-sensitized Ire1. To rule out indirect effects, we assayed the activities of the Ire1 mutants described above in vitro. The cytosolic

portion of Ire1 consisting of the kinase domain and RNase domain (Ire1*) was expressed and purified from *Escherichia coli* (7). Using γ -³²P]-labeled ATP, we assayed wild-type Ire1* and Ire1*(L745G) for autophosphorylation in the presence or absence of 1NM-PP1. Wild-type Ire1* exhibited strong autophosphorylation, which was not inhibited by 1NM-PP1 (Fig. 3A, lanes 1 and 2). In contrast, Ire1*(L745G) exhibited markedly diminished autophosphorylation, consistent with poor signaling by Ire1(L745G) in vivo (Figs. 1 and 2), which was completely extinguished by 1NM-PP1 (Fig. 3A, lane 3 and 4). This suggested that Ire1*(L745G) is diminished in its ability to use ATP as substrate and that 1NM-PP1 is, as predicted, an ATP competitor.

In contrast, when provided with [γ -³²P]-labeled *N*-6 benzyl ATP, an ATP analog with a bulky substituent (Fig. 3E), Ire1*(L745G) displayed enhanced autophosphorylation activity relative to wild-type Ire1* (Fig. 3B, lane 1 and 2). Therefore, the L745G mutation does not destroy Ire1's catalytic function, but rather alters its substrate specificity from ATP to ATP analogs having complementary features that permit binding in the expanded active site.

To ask how the RNase activity of Ire1 is affected by manipulation of its kinase domain, we incubated wild-type Ire1* and Ire1*(L745G) with a truncated in vitro transcribed *HAC1* RNA substrate (*HAC1*₆₀₀) and monitored the production of cleavage products (7). Wild-type Ire1* cleaved *HAC1*₆₀₀ RNA at both splice junctions, producing the expected fragments corresponding to the singly and doubly cut species (Fig. 3C, lane 2). The cleavage reaction was not affected by 1NM-PP1 (Fig. 3C, lane 3). In contrast, Ire1*(L745G) was completely inactive, whereas the addition of 1NM-PP1, at the same concentration that was shown in Fig. 3A (lane 4) to completely inhibit the kinase activity, restored RNase activity significantly (Fig. 3C, lanes 4 and 5), strongly supporting the no-

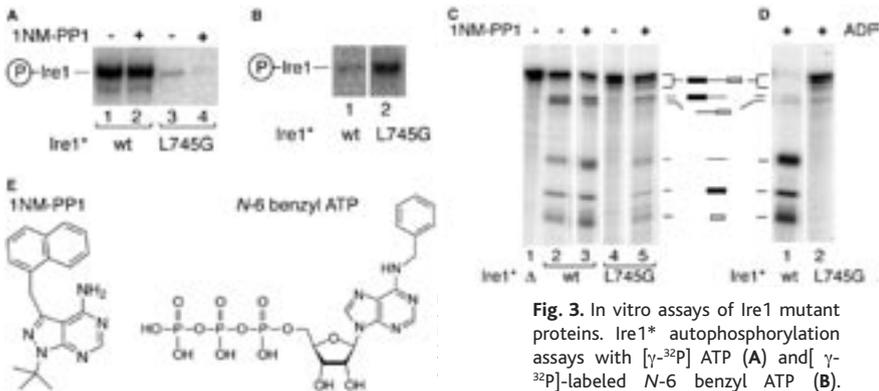


Fig. 3. In vitro assays of Ire1 mutant proteins. Ire1* autophosphorylation assays with $[\gamma\text{-}^{32}\text{P}]$ ATP (A) and $[\gamma\text{-}^{32}\text{P}]$ -labeled N-6 benzyl ATP (B). Phosphorylated Ire1* proteins (wild-type and L745G) are indicated. Ire1* RNase assays (against in vitro transcribed *HACT*₆₀₀ RNA) with stimulatory cofactors 1NM-PP1 (C) and ADP (D). Icons in (C) and (D) indicate *HACT*₆₀₀ RNA cleavage products, denoting all combinations of singly and doubly cut RNA. (E) The chemical structures of 1NM-PP1 and N-6 benzyl ATP are shown.



Fig. 4. Glycerol-density gradient velocity sedimentation of wild-type and mutant Ire1*. Ire1* proteins (A to C) and Ire1*(L745G) proteins (D to F) were incubated in the presence of ADP [(B) and (E)] or 1NM-PP1 [(C) and (F)] and subjected to velocity sedimentation analysis. Sedimentation is from left to right.

tion that 1NM-PP1 uncouples the kinase and RNase activities of 1NM-PP1-sensitized Ire1.

We have previously shown that cleavage of *HAC1* mRNA by wild-type Ire1* is stimulated by adenosine 5'-diphosphate (ADP) (7). This is consistent with the notion derived from in vivo studies that Ire1 requires an active kinase domain, because we know that recombinant wild-type Ire1* is already phosphorylated, thereby alleviating a necessity for de novo phosphorylation in vitro. ADP efficiently stimulated the RNase activity of Ire1* (Fig. 3D, lane 1) but not that of Ire1*(L745G) (Fig. 3D, lane 2).

For wild-type Ire1 and 1NM-PP1-sensitized Ire1(L745G), ADP and 1NM-PP1, respectively, function as stimulatory cofactors. Therefore, we asked whether a

conformational change occurs in response to cofactor binding by analyzing Ire1* and Ire1*(L745G) through glycerol-density gradient velocity sedimentation.

Ire1* sedimented as a broad peak in the gradient (Fig. 4). Upon the addition of ADP (Fig. 4B), but not 1NM-PP1 (Fig. 4C), the peak of Ire1* shifted as a result of faster sedimentation. Conversely, the peak of Ire1*(L745G) shifted by a corresponding amount upon the addition of 1NM-PP1 (Fig. 4F) but not upon the addition of ADP (Fig. 4E). This shift may be indicative of a conformational change and/or change in the oligomeric state of the enzymes in response to the binding of their respective cofactors.

1NM-PP1-sensitized, kinase-dead Ire1 signals a canonical UPR. To ask whether bypassing the Ire1 kinase activity with 1NM-PP1 allows the induction of a full UPR, we profiled mRNA expression on yeast genomic microarrays. To this end, mRNA from wildtype *IRE1*, *Δire1*, and *IRE1(L745A,D828A)* cells was isolated

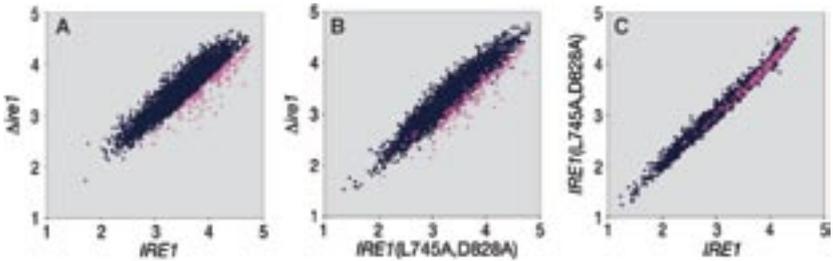


Fig. 5. (A to C) x-y scatter plot analysis of mRNA abundance. The plots (\log_{10}) compare yeast genomic microarray analyses between the genotypes indicated. In each case, mRNA was purified from cells that were provided both DTT and 1NM-PP1. Pink dots represent genes with expression more than two times as high in wild-type cells as in $\Delta ire1$ cells.

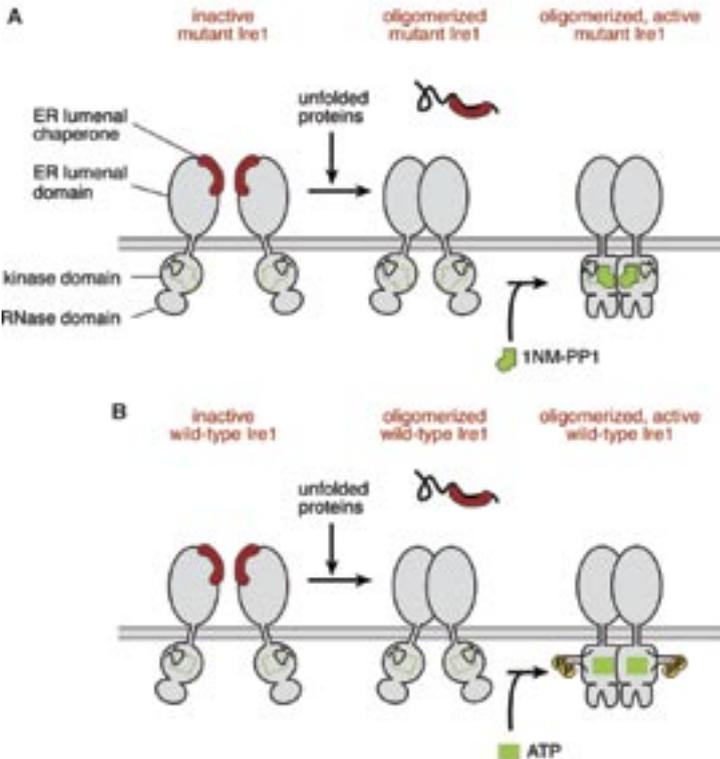


Fig. 6. Model of activation of 1NM-PP1-sensitized and wild-type Ire1. For both proteins, chaperone dissociation resulting from unfolded-protein accumulation causes oligomerization in the plane of the ER membrane, which is a precondition for further activation. For mutant Ire1 (A), 1NM-PP1 binding to the inactive kinase domain active site causes a conformational change, which activates the RNase. For wild-type Ire1 (B), trans-autophosphorylation, with ATP, of the activation segment fully opens the kinase domain for binding of ADP or ATP, which activates the RNase.

at different points in time after the addition of 1NM-PP1 and DTT. cDNAs derived from these mRNA populations were

labeled with different fluorescent dyes, mixed pairwise, and hybridized to microarrays (20). Scatter plots of pairwise

comparisons of mRNA expression levels are shown in Fig. 5; the *x-y* scatter plots represent relative expressions for every mRNA between the specified genotypes.

The comparison of wild-type *IRE1*-expressing cells with *Aire1* cells revealed a distinct set of up-regulated UPR target genes in wild-type *IRE1*-expressing cells (Fig. 5A; genes that were up-regulated more than twofold are shown in pink). These *IRE1*-dependent genes include previously described UPR transcriptional targets (21, 22). Notably, the same set of (pink-colored) UPR target genes was up-regulated to a similar magnitude in *IRE1(L745A,D828A)*-expressing cells relative to Δ *Aire1* cells (Fig. 5B), suggesting that a canonical UPR was induced in the 1NM-PP1-sensitized, kinase-dead mutant. This conclusion was confirmed by the tight superimposition of the scatter diagonals of both UPR targets and the rest of the transcriptome evident in the expression profiles of wild-type *IRE1*- compared with *IRE1(L745A,D828A)*-expressing cells (Fig. 5C).

Building an instructive UPR switch.

The 1NM-PP1-sensitized *IRE1* mutants described so far act permissively, because activation requires both 1NM-PP1 and protein-misfolding agents. This indicates that an unfolded-protein signal needs to be received by the ER luminal domain for activation. We asked if we could bypass this requirement using a constitutively on version of *IRE1* (called *IRE1^C*) isolated previously through random mutagenesis of the ER luminal domain (18, 21). *IRE1^C* significantly induced the UPR without DTT, resulting in about 75% of the maximal activity of wild-type *IRE1* with DTT (Fig. 1D, bars 2 and 5). We predicted that a chimera of the *IRE1^C* and 1NM-PP1-sensitized, kinase-dead *IRE1(L745A,D828A)* mutants would activate the UPR with 1NM-PP1 alone, even in the absence of DTT. The data in Fig. 1D (bars 13 to 16) show that this prediction holds true. Whereas wild-type *IRE1* re-

quires DTT for induction and is immune to 1NM-PP1 (Fig. 1D, bars 1 to 4), and *IRE1(L745A,D828A)* requires both DTT and 1NM-PP1 (Fig. 1D, bars 9 to 12), the chimeric *IRE1^C* (L745A,D828A) requires only 1NM-PP1 for activation, irrespective of the addition of DTT (Fig. 1D, bars 13 to 16). Thus, *IRE1^C* (L745A,D828A) is an instructive switch that turns on the UPR upon the addition of 1NM-PP1 alone (i.e., even in the absence of ER stress).

Summary and implications. Since our discovery that Ire1's RNase initiates splicing of *HAC1* mRNA, the function of its kinase domain has remained an enigma (7). Although mutational analyses have indicated a requirement for an enzymatically active kinase and have defined activation-segment phosphorylation sites (9), to date no targets besides Ire1 itself have been identified. Here, we report that both the kinase activity and activation-segment phosphorylation can be bypassed if the small ATP-mimic 1NM-PP1 is provided to a mutant enzyme to which it can bind. Instead of inhibiting the function served by ATP, 1NM-PP1 rectifies the signaling defect of 1NM-PP1-sensitized Ire1 mutants, leading to the induction of a canonical UPR (21). Thus, ligand binding to the kinase domain, rather than a phosphotransfer function mediated by the kinase activity, is required for RNase activation. The kinase module of Ire1, therefore, uses an unprecedented mechanism to propagate the UPR signal.

One model that could explain our data is proposed in Fig. 6B. According to this model, autophosphorylation of wild-type Ire1 occurs in trans between Ire1 molecules after they have been released from their chaperone anchors and have oligomerized by lateral diffusion in the plane of the ER membrane (9, 23). Trans-autophosphorylation of the activation segment locks the activation segment in the "swung-out" state (24, 25), and fully opens the active site, allowing ADP or ATP to bind efficiently, leading in turn to conformational

changes that activate the RNase domain.

In contrast, the binding of 1NM-PP1 to the active site of 1NM-PP1-sensitized Ire1 occurs even in the absence of phosphorylation of the activation segment. It is plausible that because of its small size, 1NMPP1 can bypass the "closed" activation segment, which is proposed to occlude access to the active site for the larger ADP and ATP molecules (Fig. 6A). Alternatively, 1NM-PP1 may enter the active site when the unphosphorylated activation segment opens transiently, which occurs with low frequency in other kinases (26). If the offrate of 1NM-PP1 binding to 1NM-PP1-sensitized Ire1 is slow compared to that of ADP and ATP, most mutant Ire1 molecules could be trapped in a drug-bound state. Based on either scenario, we propose that drug binding alone causes conformational rearrangements in the kinase that activate the RNase. The nucleotide-bound state of phosphorylated wild-type Ire1, therefore, mimics the 1NM-PP1-bound state of unphosphorylated 1NM-PP1-sensitized Ire1.

The model discussed so far does not explain why 1NM-PP1 should not bind to individual 1NM-PP1-sensitized Ire1 molecules and activate them, even without oligomerization induced by activation of the UPR through the ER luminal domain. Two possibilities could account for the observations: (i) The binding of 1NM-PP1 may be enhanced in oligomerized, chaperone-free mutants, or (ii) 1NM-PP1 may bind equally well to chaperone-anchored and chaperone-free (or constitutive-on) Ire1, but oligomerization may be required because interaction between neighboring molecules is required for the conformational change that activates the RNase function. Currently, our data do not allow us to distinguish between these possibilities.

Our findings provoke the question of what the "natural" stimulatory ligand of Ire1's kinase domain may be. One likely candidate is ADP, which is generated by Ire1 from ATP during the autophosphorylation reaction. In vitro, ADP is a better

activator of Ire1 than ATP (7). Ire1 would therefore be "self priming," generating an efficient activator already bound to its active site. In many professional secretory cells such as the β -cells of the endocrine pancreas (which contain high concentrations of Ire1), ADP levels rise (and ATP levels decline) temporarily in proportion to nutritional stress (27). ADP therefore is physiologically poised to serve as a cofactor that could signal a starvation state. It is known that protein folding becomes inefficient as the nutritional status of cells declines, triggering the UPR (28). Thus, in the face of ATP depletion, ADP-mediated conformational changes might increase the dwell time of activated Ire1. As such, Ire1 could have evolved this regulatory mechanism to monitor the energy balance of the cell and couple this information to activation of the UPR.

Although unprecedented in proteins containing active kinase domains, our findings are consistent with previous observations of RNase L, which is closely related to Ire1, bearing a kinase-like domain followed C-terminally by an RNase domain (29, 30). Similar to Ire1, RNase L is activated by dimerization (31). However, RNase Ls kinase domain is naturally inactive (29), whereas its RNase activity is stimulated by adenosine nucleotide binding to the kinase domain. Therefore, like Ire1, the ligand-occupied kinase domain of RNase L serves as a module that participates in activation and regulation of the RNase function. Insights gained from Ire1 and RNase L may extend to other proteins containing kinase or enzymatically inactive pseudokinase domains (32).

References and Notes

1. M. J. Gething, J. Sambrook, *Nature* **355**, 33 (1992).
2. F. J. Stevens, Y. Argon, *Semin. Cell Dev. Biol.* **10**, 443 (1999).
3. C. Patil, P. Walter, *Curr. Opin. Cell Biol.* **13**, 349 (2001).
4. J. S. Cox, C. E. Shamu, P. Walter, *Cell* **73**, 1197 (1993).
5. J. S. Cox, P. Walter, *Cell* **87**, 391 (1996).
6. U. Rueggsegger, J. H. Leber, P. Walter, *Cell* **107**, 103 (2001).
7. C. Sidrauski, P. Walter, *Cell* **90**, 1031 (1997).
8. C. Sidrauski, J. S. Cox, P. Walter, *Cell* **87**, 405 (1996).
9. C. E. Shamu, P. Walter, *EMBO J.* **15**, 3028 (1996).

10. A. Weiss, J. Schlessinger, *Cell* **94**, 277 (1998).
11. K. Mori, W. Ma, M. J. Gething, J. Sambrook, *Cell* **74**, 743 (1993).
12. K. Shah, Y. Liu, C. Deirmengian, K. M. Shokat, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 3565 (1997).
13. A. C. Bishop *et al.*, *Curr. Biol.* **8**, 257 (1998).
14. E. L. Weiss, A. C. Bishop, K. M. Shokat, D. G. Drubin, *Nature Cell Biol.* **2**, 677 (2000).
15. A. C. Bishop *et al.*, *Nature* **407**, 395 (2000).
16. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; D, Asp; G, Gly; L, Leu; and S, Ser.
17. A. S. Carroll, A. C. Bishop, J. L. DeRisi, K. M. Shokat, E. K. O'Shea, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 12578 (2001).
18. F. R. Papa, C. Zhang, K. Shokat, P. Walter, data not shown.
19. M. Huse, J. Kuriyan, *Cell* **109**, 275 (2002).
20. J. L. DeRisi, V. R. Iyer, P. O. Brown, *Science* **278**, 680 (1997).
21. K. J. Travers *et al.*, *Cell* **101**, 249 (2000).
22. Materials and methods are available as supporting material on Science Online.
23. F. Urano, A. Bertolotti, D. Ron, *J. Cell Sci.* **113**, 3697 (2000).
24. S. R. Hubbard, M. Mohammadi, J. Schlessinger, *J. Biol. Chem.* **273**, 11987 (1998).
25. T. Schindler *et al.*, *Mol. Cell* **3**, 639 (1999).
26. M. Porter, T. Schindler, J. Kuriyan, W. T. Miller, *J. Biol. Chem.* **275**, 2721 (2000).
27. F. Schuit, K. Moens, H. Heimberg, D. Pipeleers, *J. Biol. Chem.* **274**, 32803 (1999).
28. R. J. Kaufman *et al.*, *Nature Rev. Mol. Cell Biol.* **3**, 411 (2002).
29. B. Dong, M. Niwa, P. Walter, R. H. Silverman, *RNA* **7**, 361 (2001).
30. S. Naik, J. M. Paranjape, R. H. Silverman, *Nucleic Acids Res.* **26**, 1522 (1998).
31. B. Dong, R. H. Silverman, *Nucleic Acids Res.* **27**, 439 (1999).
32. M. Krohner, M. A. Miller, R. E. Steele, *Bioessays* **23**, 69 (2001).
33. We thank M. Stark for constructing the *IRE1^C* allele, S. Ullrich for his initial effort in constructing Ire1 mutants, and J. Kuriyan and members of the Walter and Shokat labs for valuable discussions. Supported by grants from the NIH to P.W. (GM32384) and K.S. (AI44009), and postdoctoral support from the UCSF Molecular Medicine Program (to F.P.). P.W. receives support as an investigator of the Howard Hughes Medical Institute.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1090031/DC1
 Materials and Methods
 Microarray Excel Spreadsheets S1 to S3
 References and Notes

4 August 2003; accepted 1 October 2003
 Published online 16 October 2003;
 10.1126/science.1090031
 Include this information when citing this paper.

Human Catechol-*O*-Methyltransferase Haplotypes Modulate Protein Expression by Altering mRNA Secondary Structure

A. G. Nackley,¹ S. A. Shabalina,² I. E. Tchivileva,¹ K. Satterfield,¹ O. Korchynskiy,³ S. S. Makarov,⁴ W. Maixner,¹ L. Diatchenko^{1*}

Catechol-*O*-methyltransferase (COMT) is a key regulator of pain perception, cognitive function, and affective mood. Three common haplotypes of the human *COMT* gene, divergent in two synonymous and one nonsynonymous position, code for differences in COMT enzymatic activity and are associated with pain sensitivity. Haplotypes divergent in synonymous changes exhibited the largest difference in COMT enzymatic activity, due to a reduced amount of translated protein. The major *COMT* haplotypes varied with respect to messenger RNA local stem-loop structures, such that the most stable structure was associated with the lowest protein levels and enzymatic activity. Site-directed mutagenesis that eliminated the stable structure restored the amount of translated protein. These data highlight the functional significance of synonymous variations and suggest the importance of haplotypes over single-nucleotide polymorphisms for analysis of genetic variations.

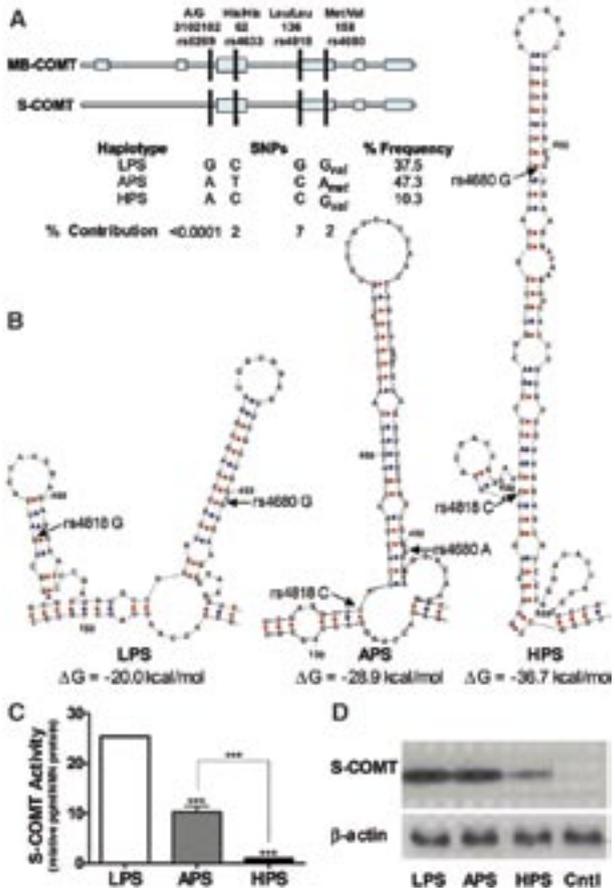
The ability to predict the downstream effects of genetic variation is critically important for understanding both the evolution of the genome and the molecular basis of human disease. The effects of nonsynonymous polymorphisms have been widely characterized; because these variations directly influence protein function, they are relatively easy to study

statistically and experimentally (*1*). However, characterizing polymorphisms locat-

¹Center for Neurosensory Disorders, University of North Carolina, Chapel Hill, NC 27599, USA. ²National Center for Biotechnology Information, National Institutes of Health, Bethesda, MD 20894, USA. ³Thurston Arthritis Center, University of North Carolina, Chapel Hill, NC 27599, USA. ⁴Attagene, Inc., 7030 Kit Creek Road, Research Triangle Park, NC 27560, USA.

*To whom correspondence should be addressed. E-mail: lbdiatch@email.unc.edu

Fig. 1. Common haplotypes of the human *COMT* gene differ with respect to mRNA secondary structure and enzymatic activity. **(A)** A schematic diagram illustrates *COMT* genomic organization and SNP composition for the three haplotypes. Percent frequency of each haplotype in a cohort of healthy Caucasian females, and percent independent SNP contribution to pain sensitivity, are indicated. **(B)** The local stem-loop structures associated with each of the three haplotypes are shown. Relative to the LPS and APS haplotypes, the HPS local stem-loop structure had a higher folding potential. **(C)** and **(D)** The LPS haplotype exhibited the highest, while the HPS haplotype exhibited the lowest enzymatic activity and protein levels in cells expressing *COMT*. *** $P < 0.001$, ≠ LPS. +++ $P < 0.001$, ≠ APS.

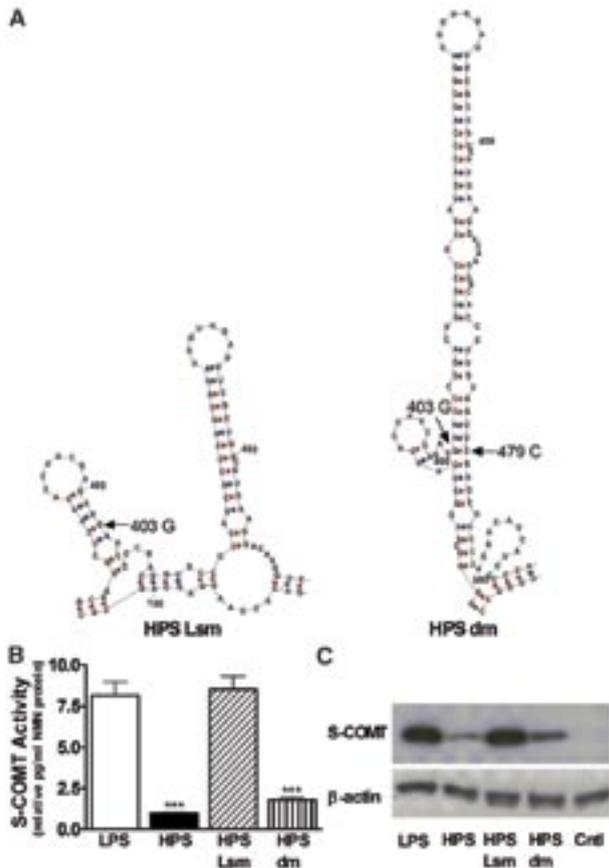


ed in regulatory regions, which are much more common, has proved to be problematic (2). Here, we focus on the mechanism whereby polymorphisms of the catechol-*O*-methyltransferase (*COMT*) gene regulate gene expression.

COMT is an enzyme responsible for degrading catecholamines and thus represents a critical component of homeostasis maintenance (3). The human *COMT* gene encodes two distinct proteins: soluble *COMT* (S-*COMT*) and membrane-bound *COMT* (MB-*COMT*) through the use of alternative translation initiation sites and promoters (3). Recently, *COMT* has been implicated in the modulation of persistent pain (4–7). Our

group demonstrated that three common haplotypes of the human *COMT* gene are associated with pain sensitivity and the likelihood of developing temporomandibular joint disorder (TMJD), a common chronic musculoskeletal pain condition (4). Three major haplotypes are formed by four single-nucleotide polymorphisms (SNPs): one located in the S-*COMT* promoter region (A/G; rs26269) and three in the S- and MB-*COMT* coding region at codons *his*⁶²*his* (C/T; rs4633), *leu*¹³⁶*leu* (C/G; rs4818), and *val*¹⁵⁸*met* (A/G; rs4680) (Fig. 1A). On the basis of subjects' pain responsiveness, haplotypes were designated as low (LPS; GCGG), average (APS; ATCA), or high (HPS; ACCG) pain sen-

Fig. 2. Site-directed mutagenesis that destroys the stable stem-loop structure corresponding to the HPS haplotype restores COMT enzymatic activity and protein expression. **(A)** The mRNA structure corresponding to the HPS haplotype was converted to an LPS haplotype-like structure (HPS Lsm) by single mutation of 403C to G. The original HPS haplotype structure (HPS dm) was restored by double mutation of interacting nucleotides 403C to G and 479G to C. **(B and C)** The HPS Lsm exhibited COMT enzymatic activity and protein levels equivalent to those of the LPS haplotype, whereas the HPS dm exhibited reduced enzymatic activity. $***P < 0.001$, \neq LPS.



sitive. Individuals carrying HPS/APS or APS/APS diploypes were nearly 2.5 times as likely to develop TMJD. Previous data further suggest that

COMT haplotypes code for differences in COMT enzymatic activity (4); however, the molecular mechanisms involved have remained unknown.

A common SNP in codon 158 (*val¹⁵⁸met*) has been associated with pain ratings and μ -opioid system responses (7) as well as addiction, cognition, and common affective disorders (3, 8–10). The substitution of valine (Val) by methionine (Met) results in reduced thermostability and activity of the enzyme (3). It is generally accepted that *val¹⁵⁸met* is the main source of individual variation in human COMT activity; numerous studies have identified associations between the low-activity *met* allele and several complex phenotypes (3, 8, 10).

However, observed associations between these conditions and the *met* allele are often modest and occasionally inconsistent (3). This suggests that additional SNPs in the COMT gene modulate COMT activity. Indeed, we found that haplotype rather than an individual SNP better accounts for variability in pain sensitivity (4). The HPS and LPS haplotypes that both code for the stable *val¹⁵⁸* variant were associated with the two extreme-pain phenotypes; thus, the effect of haplotype on pain sensitivity in our study cannot be explained by the sum of the effects of functional SNPs. Instead, the *val¹⁵⁸met* SNP interacts with other SNPs to determine phenotype. Because variation in the S-COMT promoter

region does not contribute to pain phenotype (Fig. 1A), we suggest that the rate of mRNA degradation or protein synthesis is affected by the structural properties of the haplotypes, such as haplotype-specific mRNA secondary structure.

Previous reports have shown that polymorphic alleles can markedly affect mRNA secondary structure (11, 12), which can then have functional consequences on the rate of mRNA degradation (11, 13). It is also plausible that polymorphic alleles directly modulate protein translation through alterations in mRNA secondary structure, because protein translation efficiency is affected by mRNA secondary structure (14–16). To test these possibilities, we evaluated the effect of LPS, APS, and HPS haplotypes on the stability of the corresponding mRNA secondary structures (17).

Secondary structures of the full-length LPS, APS, and HPS mRNA transcripts were predicted by means of the RNA Mfold (18, 19) and Afold (20) programs. The mRNA folding analyses demonstrated that the major *COMT* haplotypes differ with respect to mRNA secondary structure. The LPS haplotype codes for the shortest, least stable local stem-loop structure, and the HPS haplotype codes for the longest, most stable local stem-loop structure in the *val*¹⁵⁸ region for both *S-COMT* and *MB-COMT*. Gibbs free energy (ΔG) for the stem-loop structure associated with the HPS haplotype is ~17 kcal/mol less than that associated with the LPS haplotype for both *S-COMT* and *MB-COMT* (Fig. 1B and fig. S1A). Additional evidence supporting predicted RNA folding structures was obtained by generating consensus RNA secondary structures based on comparative analysis of *COMT* sequences from eight mammalian species (fig. S2). The consensus RNA folding structures were LPS-like and did not contain highly stable local stem-loop structures analogous to the human HPS-like form. Thus, substantial deviation

from consensus structure, as observed for the HPS haplotype, should have notable functional consequences. Additional studies were conducted to test this molecular modeling.

We constructed full-length S- and MB-COMT cDNA clones in mammalian expression vectors that differed only in three nucleotides corresponding to the LPS, APS, and HPS haplotypes (17, 21). Rat adrenal (PC-12) cells were transiently transfected with each of these six constructs. COMT enzymatic activity, protein expression, and mRNA abundance were measured. Relative to the LPS haplotype, the HPS haplotype showed a 25- and 18-fold reduction in enzymatic activity for S- and MB-COMT constructs, respectively (Fig. 1C and fig. S1B). The HPS haplotype also exhibited marked reductions in S- and MB-COMT protein expression (Fig. 1D and fig. S1C). The APS haplotype displayed a moderate 2.5- and 3-fold reduction in enzymatic activity for S- and MB-COMT constructs, respectively, while protein expression levels did not differ. The moderate reduction in enzymatic activity produced by the APS haplotype is most likely due to the previously reported decrease in protein thermostability coded by the *met*¹⁵⁸ allele (3). These data illustrate that the reduced enzymatic activity corresponding to the HPS haplotype is paralleled by reduced protein levels, an effect that could be mediated by local mRNA secondary structure at the level of protein synthesis and/or mRNA degradation. Because total RNA abundance and RNA degradation rates did not parallel COMT protein levels (fig. S2), differences in protein translation efficiency likely results from differences in the local secondary structure of corresponding mRNAs

To directly assess this hypothesis, we performed site-directed mutagenesis (17). The stable stem-loop structure of S- and MB-COMT mRNA corresponding to the HPS haplotype is supported by base pairs between several critical nucleotides, including 403C and 479G in *S-COMT* and

625C and 701G in *MB-COMT* (Fig. 2A and fig. S3A). Mutation of 403C to G in *S-COMT* or 625C to G in *MB-COMT* destroys the stable stem-loop structure and converts it into a LPS haplotype-like structure (HPS Lsm). Double mutation of mRNA in position 403C to G and 479G to C in *S-COMT* or 625C to G and 701G to C in *MB-COMT* reconstructs the original long stem-loop structure (HPS dm). The single- and double-nucleotide HPS mutants (HPS Lsm and HPS dm, respectively) were transiently transfected to PC-12 cells. As predicted by the mRNA secondary-structure folding analyses, the HPS Lsm exhibited increased COMT enzymatic activity and protein levels equivalent to those of the LPS haplotype, whereas the HPS dm exhibited reduced enzymatic activity and protein levels equivalent to those of the original HPS haplotype (Fig. 2, B and C; fig. S3, B and C). These data rule out the involvement of RNA sequence recognition motifs or codon usage in the regulation of translation. In contrast to the HPS haplotype, protein levels did not parallel COMT enzymatic activity for the APS haplotype and site-directed mutagenesis confirmed that the *met*¹⁵⁸ allele, not a more stable mRNA secondary structure, drives the reduced enzymatic activity observed for the APS haplotype (fig. S4). This difference is moderate relative to the mRNA structure-dependent difference coded by LPS and HPS haplotypes. These results were verified by an alternate approach of modifying mRNA secondary structure (fig. S5).

Our data have very broad evolutionary and medical implications for the analysis of variants common in the human population. The fact that alterations in mRNA secondary structure resulting from synonymous changes have such a pronounced effect on the level of protein expression emphasizes the critical role of synonymous nucleotide positions in maintaining mRNA secondary structure and suggests that the mRNA secondary structure, rather than indepen-

dent nucleotides in the synonymous positions, should undergo substantial selective pressure (22). Furthermore, our data stress the importance of synonymous SNPs as potential functional variants in the area of human medical genetics. Although non-synonymous SNPs are believed to have the strongest impact on variation in gene function, our data clearly demonstrate that haplotypic variants of common synonymous SNPs can have stronger effects on gene function than nonsynonymous variations and play an important role in disease onset and progression.

References and Notes

1. L. Y. Yampolsky, F. A. Kondrashov, A. S. Kondrashov, *Hum. Mol. Genet.* **14**, 3191 (2005).
2. J. C. Knight, *J. Mol. Med.* **83**, 97 (2005).
3. P. T. Mannisto, S. Kaakkola, *Pharmacol. Rev.* **51**, 593 (1999).
4. L. Diatchenko et al., *Hum. Mol. Genet.* **14**, 135 (2005).
5. J. J. Marbach, M. Levitt, *J. Dent. Res.* **55**, 711 (1976).
6. T. T. Ravvag et al., *Pain* **116**, 73 (2005).
7. J. K. Zubietta et al., *Science* **299**, 1240 (2003).
8. G. Winterer, D. Goldman, *Brain Res. Brain Res. Rev.* **43**, 134 (2003).
9. B. Funke et al., *Behav. Brain Funct.* **1**, 19 (2005).
10. G. Oroszi, D. Goldman, *Pharmacogenomics* **5**, 1037 (2004).
11. J. Duan et al., *Hum. Mol. Genet.* **12**, 205 (2003).
12. L. X. Shen, J. P. Basilion, V. P. Stanton Jr., *Proc. Natl. Acad. Sci. U.S.A.* **96**, 7871 (1999).
13. I. Puga et al., *Endocrinology* **146**, 2210 (2005).
14. K. Mita, S. Ichimura, M. Zama, T. C. James, *J. Mol. Biol.* **203**, 917 (1988).
15. T. D. Schmittgen, K. D. Danenberg, T. Horikoshi, H. J. Lenz, P. V. Danenberg, *J. Biol. Chem.* **269**, 16269 (1994).
16. A. Shalev et al., *Endocrinology* **143**, 2541 (2002).
17. Materials and methods are available as supporting material on *Science* Online.
18. D. H. Mathews, J. Sabina, M. Zuker, D. H. Turner, *J. Mol. Biol.* **288**, 911 (1999).
19. M. Zuker, *Nucleic Acids Res.* **31**, 3406 (2003).
20. A. Y. Ogurtsov, S. A. Shabalina, A. S. Kondrashov, M. A. Roytberg, *Bioinformatics* **22**, 1317 (2006).
21. Sequence accession numbers: S-COMT clones BG290167, CA489448, and BF037202 represent LPS, APS, and HPS haplotypes, respectively. Clones corresponding to all three haplotypes and including the transcriptional start site were constructed by use of the unique restriction enzyme Bsp MI. The gene coding region containing all three SNPs was cut and inserted into the plasmids through use of the S-COMT (BG290167) and MB-COMT (B1835796) clones containing the entire COMT 5' and 3' ends as a backbone vector.
22. S. A. Shabalina, A. Y. Ogurtsov, N. A. Spiridonov, *Nucleic Acids Res.* **34**, 2428 (2006).

23. We are grateful to the National Institute of Child Health and Human Development, National Institute of Neurological Disorders and Stroke, and National Institute of Dental and Craniofacial Research at the NIH for financial support of this work. This research was also supported by the Intramural Research Program of the NIH, National Center for Biotechnology Information.

Supporting Online Material

www.sciencemag.org/cgi/content/full/314/5807/1930/DC1
Materials and Methods

SOM Text

Figs. S1 to S6

References

13 June 2006; accepted 16 November 2006

10.1126/science.1131262

An mRNA Surveillance Mechanism That Eliminates Transcripts Lacking Termination Codons

Pamela A. Frischmeyer,¹ Ambro van Hoof,^{3,4} Kathryn O'Donnell,¹
Anthony L. Guerrero,² Roy Parker,^{3,4} Harry C. Dietz^{1,4*}

Translation is an important mechanism to monitor the quality of messenger RNAs (mRNAs), as exemplified by the translation-dependent recognition and degradation of transcripts harboring premature termination codons (PTCs) by the nonsense-mediated mRNA decay (NMD) pathway. We demonstrate in yeast that mRNAs lacking all termination codons are as labile as nonsense transcripts. Decay of "nonstop" transcripts in yeast requires translation but is mechanistically distinguished from NMD and the major mRNA turnover pathway that requires deadenylation, decapping, and 5'-to-3' exonucleolytic decay. These data suggest that nonstop decay is initiated when the ribosome reaches the 3' terminus of the message. We demonstrate multiple physiologic sources of nonstop transcripts and conservation of their accelerated decay in mammalian cells. This process regulates the stability and expression of mRNAs that fail to signal translational termination.

Eukaryotes have evolved surveillance mechanisms that are intimately linked to translation to eliminate errors in mRNA biogenesis. The decay of transcripts containing PTCs by the NMD pathway effectively prevents expression of deleterious truncated proteins. In prokaryotes, protein products encoded by transcripts lacking termination codons are marked for degradation by the addition of a COOH-terminal tag encoded by tmRNA (1, 2). Thus, both the presence and context of translational termination can regulate

gene expression. In order to determine whether the presence of translational termination influences mRNA stability, we assayed PGK1 transcripts in *Saccharomyces cerevisiae* derived from the following constructs (3): wild-type PGK1 (WT-PGK1), a nonsense form of PGK1 harboring a PTC at codon 22 [PTC(22)-PGK1], and nonstop-PGK1 that was created by removing the bona fide termination codon and all in-frame termination codons in the 3' UTR (untranslated region) from the WT-PGK1 transcript. Nonstop-PGK1 transcripts were as labile as their nonsense-containing counterparts (Fig. 1). At least three *trans*-acting factors (Upf1p, Upf2p, and Upf3p) are essential for NMD in *S. cerevisiae* (4, 5). Remarkably, nonstop transcripts were not stabilized in strains lacking Upf1p, distinguishing the pathway

¹Institute for Genetic Medicine, ²Department of Biophysics and Biophysical Chemistry, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. ³University of Arizona, Department of Molecular and Cellular Biology, Tucson, AZ 85721, USA. ⁴Howard Hughes Medical Institute.

*To whom correspondence should be addressed. E-mail: hdietz@jhmi.edu

tive determinants of message stability (13, 14). To determine whether this mechanism is relevant to nonstop decay, we examined the performance of transcript Ter-poly(A)-PGK1 which contains a termination codon inserted one codon upstream of the site of polyadenylate [poly(A)] addition (15) in transcript nonstop-PGK1. The addition of a termination codon at the 3' end of the 3' UTR of a nonstop transcript resulted in substantial (threefold) stabilization (Fig. 1). Unlike nonstop-PGK1 transcripts, the half-life of Ter-poly(A)-PGK1 was increased in the absence of Xrn1p (Fig. 1), indicating that this transcript is degraded by the pathway for normal mRNA and not by the nonstop pathway. These data suggest that the instability of nonstop-PGK1 transcripts cannot fully be explained by ribosomal displacement of factors bound to the 3' UTR.

There is an emerging view that the 5' and 3' ends of mRNAs interact to form a closed loop and that this conformation is required for efficient translation initiation and normal mRNA stability. Participants in the interaction include components of the translation initiation complex eIF4F as well as Pab1p. The absence of a termination codon in nonstop-PGK1 is predicted to allow the ribosome to continue translating through the 3' UTR and poly(A) tail, potentially resulting in displacement of Pab1p, disruption of normal mRNP structure, and consequently accelerated degradation of the transcript. However, while mRNAs in *pab1Δ* strains do undergo accelerated decay, the rapid turnover is a result of premature decapping and can be suppressed by mutations in *XRNI* (16), allowing distinction from nonstop decay. Lability of nonstop transcripts might also be a consequence of ribosomal stalling at the 3' end of the transcript. It is tempting to speculate that 3'-to-5' exonucleolytic activity underlies the decapping- and 5'-to-3' exonuclease-independent, translation-dependent accelerated decay of nonstop transcripts. An appealing candidate

is the exosome, a collection of proteins with 3'-to-5' exonuclease activity that functions in the processing of 5.8S RNA, rRNA, small nucleolar RNA (snoRNA), small nuclear RNA (snRNA), and other transcripts. Data presented in van Hoof *et al.* (17) validate this prediction.

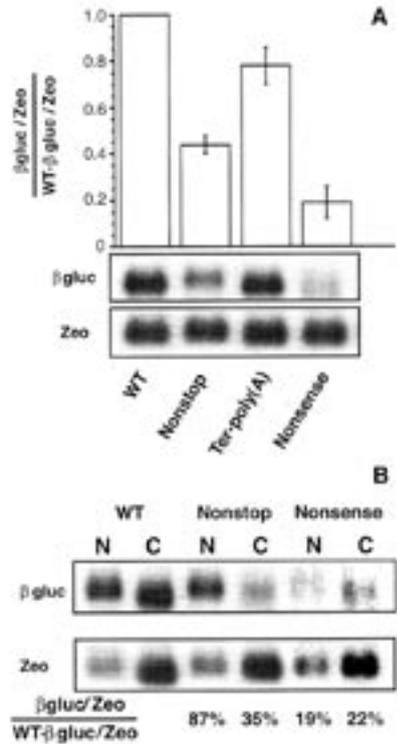


Fig. 2. Nonstop decay is conserved in mammalian cells. (A) Steady-state abundance of β -glucuronidase transcripts derived from wild type (WT), nonsense-, nonstop-, and Ter-poly(A)- β gluc minigene constructs after transient transfection into HeLa cells (28). Zeocin (zeo) transcripts served as a control for loading and transfection efficiency. For each construct, the β gluc/Zeo transcript ratio is expressed relative to that observed for WT- β gluc. (B) Steady-state abundance of WT- β gluc, nonstop- β gluc, and nonsense- β gluc transcripts were determined in the nuclear (N) and cytoplasmic (C) fractions of transfected cells (28). The ratio of normalized nonstop- or nonsense- β gluc transcript levels to WT- β gluc transcript levels are shown for each cellular compartment. Reported results are the average of three independent experiments.

In order to determine whether nonstop decay functions in mammals, we assessed the performance of transcripts derived from β -glucuronidase (β gluc) minigene constructs containing exons 1, 10, 11, and 12 separated by introns derived from the endogenous gene (18). The abundance of transcripts derived from a nonstop- β gluc version of this minigene was significantly reduced relative to WT- β gluc, suggesting that a termination codon is also essential for normal mRNA stability in mammalian cells (Fig. 2A). Addition of a stop codon one codon upstream from the site of poly(A) addition in nonstop- β gluc [Ter-poly(A)- β gluc (15)] increased the abundance of this transcript to near-wild-type levels (Fig. 2A). Thus, all functional characteristics of nonstop transcripts in yeast appear to be relevant to mammalian systems.

Both the abundance and stability of most nonsense transcripts, including nonsense- β gluc (Fig. 2B), are reduced in the nuclear fraction of mammalian cells (19, 20). This has been interpreted to suggest that translation and NMD initiate while the mRNA is still associated with (if not within) the nuclear compartment. If translation is initiated on nucleus-associated transcripts, so might nonstop decay. Subcellular fractionation studies localized mammalian nonstop decay to the cytoplasmic compartment, providing further distinction from NMD (Fig. 2B).

The conservation of nonstop decay in yeast and mammals suggests that the pathway serves an important biologic role. There are many potential physiologic sources of nonstop transcripts that warrant consideration. Mutations in bona fide termination codons would not routinely initiate nonstop decay due to the frequent occurrence of in-frame termination codons in the 3' UTR and could not plausibly provide the evolutionary pressure for maintenance of nonstop decay. In contrast, alternative use of 3'-end processing signals embedded in coding sequence has been documented

in many genes including CBP1 in yeast and the growth hormone receptor (GHR) gene in fowl (21–23). Moreover, a computer search of the human mRNA and *S. cerevisiae* open reading frame (ORF) databases revealed many additional genes that contain a strict consensus sequence for 3' end cleavage and polyadenylation within their coding region (Fig. 3A). Utilization of these premature signals would direct formation of truncated transcripts that might be substrates for the nonstop decay pathway. Indeed, analysis of 3425 random yeast cDNA clones sequenced from the 3' end (i.e., 3' ESTs) revealed that 40 showed apparent premature polyadenylation within the coding region (24).

The truncated GHR transcript in fowl is apparently translated, as evidenced by its association with polysomes (22). As predicted for a substrate for nonstop decay, the ratio of truncated-to-full-length GHR transcripts was dramatically increased after treatment of cultured chicken hepatocellular carcinoma cells with the translational inhibitor emetine (Fig. 3B). Other truncated transcripts, both bigger and smaller than the predicted 0.7-kb nonstop transcript, did not show a dramatic increase in steady-state abundance upon translational arrest suggesting that they are derived from other mRNA processing events, perhaps alternative splicing (Fig. 3B) (12). To directly test whether the nonstop mRNA pathway degrades prematurely polyadenylated mRNAs, we analyzed CBP1 transcripts in yeast. The CBP1 gene produces a 2.2-kb full-length mRNA and a 1.2-kb species with premature 3' end processing and polyadenylation within the coding region (25). As expected from the observation that nonstop decay requires the exosome in yeast (17), deletion of the gene encoding the exosome component Ski7p stabilized the prematurely polyadenylated mRNA but had no effect on the stability of the full-length mRNA (Fig. 3C). These data indicate that physiologic transcripts arising from premature poly-

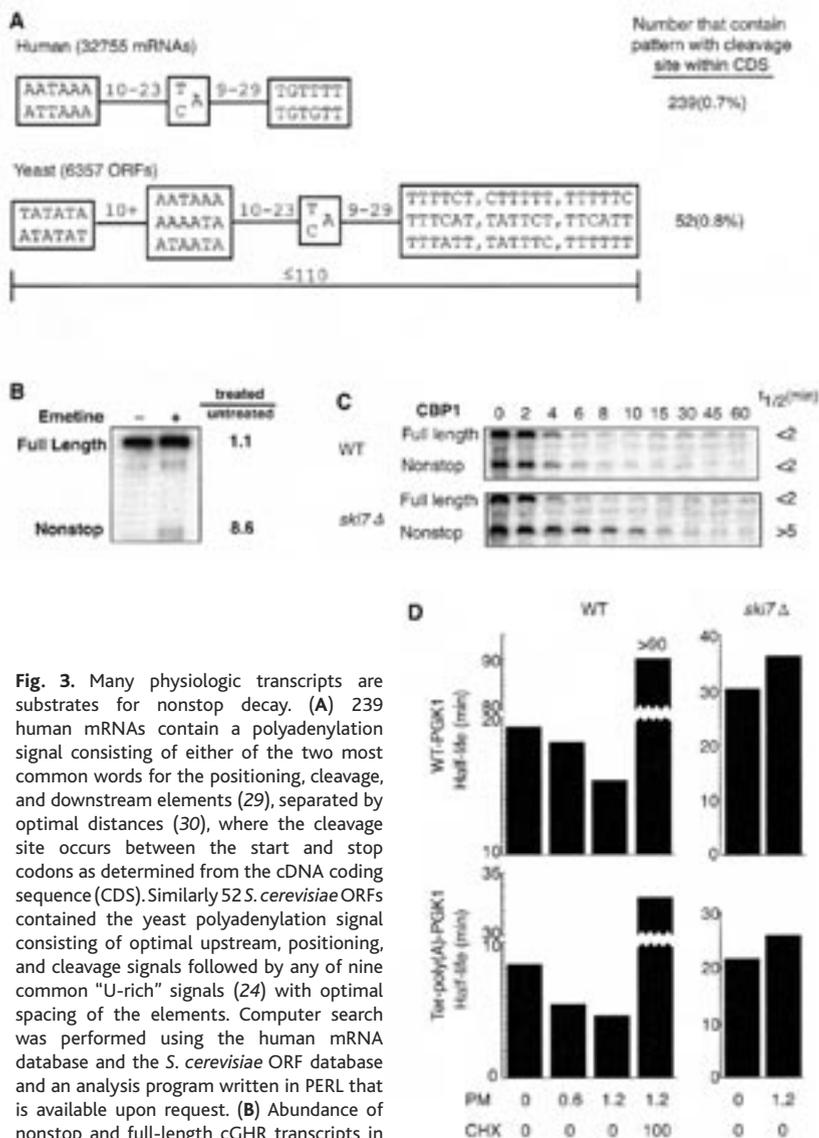


Fig. 3. Many physiologic transcripts are substrates for nonstop decay. **(A)** 239 human mRNAs contain a polyadenylation signal consisting of either of the two most common words for the positioning, cleavage, and downstream elements (29), separated by optimal distances (30), where the cleavage site occurs between the start and stop codons as determined from the cDNA coding sequence (CDS). Similarly 52 *S. cerevisiae* ORFs contained the yeast polyadenylation signal consisting of optimal upstream, positioning, and cleavage signals followed by any of nine common "U-rich" signals (24) with optimal spacing of the elements. Computer search was performed using the human mRNA database and the *S. cerevisiae* ORF database and an analysis program written in PERL that is available upon request. **(B)** Abundance of nonstop and full-length cGHR transcripts in chicken hepatocellular carcinoma cells (CRL-2117 purchased from the American Type Culture Collection, Manassas, VA) treated with 100 μ g/ml of emetine for 6 hours. Northern analysis was performed as described (22). The relative ratio of each transcript in untreated and treated cells is reported. **(C)** Half-lives ($t_{1/2}$) of the full-length and nonstop CBP1 mRNAs measured in wild-type-(WT) and *SKI7*-deleted (*ski7Δ*) yeast strains carrying plasmid pG::26 (25, 31). **(D)** Half-lives of WT-PGK1 and Ter-poly(A)-*pgk1* transcripts in wild-type-(WT) and *SKI7*-deleted (*ski7Δ*) yeast strains treated with the indicated dose of paromomycin (mg/ml) (PM) for 20 hours. For cycloheximide (CHX) experiments, yeast were treated identically except 100 μ g/ml of CHX was added during the last hour before half-life determination. Half-lives were determined as described (27).

adenylation are subject to nonstop mRNA decay. The cytoplasmic localization of ski7p is consistent with our observation that nonstop decay occurs within the cytoplasm (Fig. 2B).

Any event that diminishes translational fidelity and promotes readthrough of termination codons could plausibly result in the generation of substrates for nonstop decay. In view of recent attempts to treat genetic disorders resulting from PTCs with long-term and high-dose aminoglycoside regimens, this may achieve medical significance. As a proof-of-concept experiment, we examined the performance of transcripts with one [Ter-poly(A)-PGK1] or multiple (WT-PGK1) in-frame termination codons (including those in the 3' UTR) in yeast strains after treatment with paromomycin, which induces ribosomal readthrough. Remarkably, both transcripts showed a dose-dependent decrease in stability that could be reversed by inhibiting translational elongation with CHX or prevented by deletion of the gene encoding Ski7p (Fig. 3D). These data suggest that nonstop decay can limit the efficiency of therapeutic strategies aimed at enhancing nonsense suppression and might contribute to the toxicity associated with aminoglycoside therapy.

The degradation of nonstop transcripts may be regulated. The relative expression level of truncated GHR transcripts compared to full-length transcripts varies in a tissue-, gender-, and developmental stage-specific manner (22) and the relative abundance of truncated CBP1 transcripts varies with growth condition (21). Data presented here warrant the hypothesis that regulation may occur at the level of nonstop transcript stability rather than production. The conservation of the GHR coding sequence 3'-end processing signal throughout avian phylogeny and conservation of premature polyadenylation of the yeast RNA14 transcript and its fruitfly homolog su(f) support speculation that nonstop transcripts or derived protein products serve essential

developmental and/or homeostatic functions that are regulated by nonstop decay (21, 26).

Many processes contribute to the precise control of gene expression including transcriptional and translational control mechanisms. In recent years, mRNA stability has emerged as a major determinant of both the magnitude and fidelity of gene expression. Perhaps the most striking and comprehensively studied example is the accelerated decay of transcripts harboring PTCs by the NMD pathway. Nonstop decay now serves as an additional example of the critical role that translation plays in monitoring the fidelity of gene expression, the stability of aberrant or atypical transcripts, and hence the abundance of truncated proteins.

References and Notes

1. A. W. Karzai, E. D. Roche, R. T. Sauer, *Nature Struct. Biol.* **7**, 449 (2000).
2. K. C. Keiler, P. R. Waller, R. T. Sauer, *Science* **271**, 990 (1996).
3. The nonstop-PGK1 construct was generated from WT-PGK1 [pRP469 (8)] by creating three point mutations, using site-directed mutagenesis (Quik-Change Site-Directed Mutagenesis Kit, Stratagene) which eliminated the bona fide termination codon and all in-frame termination codons in the 3' UTR. The Ter-poly(A)-PGK1 construct was created from the nonstop-PGK1 construct, using site-directed mutagenesis (Quik-Change Site-Directed Mutagenesis Kit, Stratagene). All changes were confirmed by sequencing. Primer sequences are available upon request. The PTC(22)-PGK1 construct is described in (8) (pRP609).
4. P. Leeds, J. M. Wood, B. S. Lee, M. R. Culbertson, *Mol. Cell. Biol.* **12**, 2165 (1992).
5. Y. Cui, K. W. Hagan, S. Zhang, S. W. Peltz, *Genes Dev.* **9**, 423 (1995).
6. D. Muhrad, C. J. Decker, R. Parker, *Genes Dev.* **8**, 855 (1994).
7. C. A. Beelman *et al.*, *Nature* **382**, 642 (1996).
8. D. Muhrad, R. Parker, *Nature* **370**, 578 (1994).
9. G. Caponigro, R. Parker, *Microbiol. Rev.* **60**, 233 (1996).
10. M. Tucker *et al.*, *Cell* **104**, 377 (2001).
11. C. L. Wolfe, A. K. Hopper, N. C. Martin, *J. Biol. Chem.* **271**, 4679 (1996).
12. P. A. Frischmeyer, H. C. Dietz, data not shown.
13. I. M. Weiss, S. A. Liebhaber, *Mol. Cell. Biol.* **14**, 8123 (1994).
14. X. Wang, M. Kiledjian, I. M. Weiss, S. A. Liebhaber, *Mol. Cell. Biol.* **15**, 1769 (1995).
15. We performed 3' rapid amplification of cDNA ends (RACE) using the Marathon cDNA Amplification Kit (Clontech) and the Advantage 2 Polymerase Mix (Clontech) according to the manufacturer's instructions. Poly(A) RNA isolated from cells (Oligotex mRNA midi kit, Qiagen) transiently transfected with

- WT- β gluc or from a wild-type yeast strain carrying a plasmid expressing WT-PGK1 was used as template. A single RACE product was generated, cloned (TOPO TA 2.1 kit, Invitrogen), and sequenced.
16. G. Caponigro, R. Parker, *Genes Dev.* **9**, 2421 (1995).
 17. A. van Hoof, P. A. Frischmeyer, H. C. Dietz, R. Parker, *Science* **295**, 2262 (2002).
 18. To create the WT- β gluc minigene construct, mouse genomic DNA was used as template for amplifying three different portions of the β -glucuronidase gene: from nucleotide 2313 to 3030 encompassing exon 1 and part of intron 1 (including the splice donor sequence), from nucleotide 11285 to 13555 encompassing part of intron 9 (including the splice acceptor sequence), exon 10, and part of intron 10 (including the splice donor sequence), and from 13556 to 15813 which included the remainder of intron 10, exon 11, intron 11, and exon 12. All three fragments were TA cloned (TOPO TA 2.1 kit, Invitrogen), sequenced, and then cloned into pZeoSV2 (Invitrogen). A single base-pair deletion corresponding to the *gus*^{mps} allele was generated by site-directed mutagenesis (Quik-Change Site-Directed Mutagenesis Kit, Stratagene) of the WT- β gluc construct to create nonsense- β gluc. Nonstop- β gluc was generated from WT- β gluc by two rounds of site-directed mutagenesis (Quik-Change Site-Directed Mutagenesis Kit, Stratagene), which removed the bona fide termination codon and all in-frame termination codons in the 3' UTR. Ter-poly(A)- β gluc was generated from nonstop- β gluc via a single point mutation, which created a stop one codon upstream of the poly(A) tail. All changes were confirmed by sequencing. Primer sequences are available upon request.
 19. P. Belgrader, J. Cheng, X. Zhou, L. S. Stephenson, L. E. Maquat, *Mol. Cell. Biol.* **14**, 8219 (1994).
 20. O. Kessler, L. A. Chasin, *Mol. Cell. Biol.* **16**, 4426 (1996).
 21. K. A. Sparks, C. L. Dieckmann, *Nucleic Acids Res.* **26**, 4676 (1998).
 22. E. R. Oldham, B. Bingham, W. R. Baumbach, *Mol. Endocrinol.* **7**, 1379 (1993).
 23. C. Hilger, I. Velhagen, H. Zentgraf, C. H. Schroder, *J. Virol.* **65**, 4284 (1991).
 24. J. H. Graber, C. R. Cantor, S. C. Mohr, T. F. Smith, *Nucleic Acids Res.* **27**, 888 (1999).
 25. K. A. Sparks, S. A. Mayer, C. L. Dieckmann, *Mol. Cell. Biol.* **17**, 4199 (1997).
 26. A. Audibert, M. Simonelig, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 14302 (1998).
 27. All PGK1 minigene constructs are under the control of the GAL1 upstream activation sequence (UAS). Cultures were grown to mid-log phase in synthetic complete media-uracil (SC-ura) containing 2% galactose (transcription on). Cells were pelleted and resuspended in SC-ura media. Glucose was then added (transcription off) to a final concentration of 2% and aliquots were removed at the indicated time points. Approximately 10 μ g of total RNA isolated with the hot phenol method (4) was electrophoresed on a 1.2% agarose formaldehyde gel, transferred to a nylon membrane (Gene-Screen Plus, NEN), and hybridized with a 23 bp oligonucleotide end-labeled radioactive probe that specifically detects the polyG tract in the 3' UTR of the PGK1 transcripts (8). Half-lives were determined by plotting the percent of mRNA remaining versus time on a semi-log plot. All half-lives were determined at 25°C except for the *ccr4 Δ* (performed at 30°C) and *dcp1-2* [performed after shift to the nonpermissive time (37°C) for 1 hour] experiments.
 28. HeLa cells [maintained in DMEM with 10% fetal bovine serum (FBS)] were grown to ~80 to 90% confluency, trypsinized, and resuspended in 300 μ l of RPMI 1640 without FBS, 10 mM glucose and 0.1 mM dithiothreitol (DTT). Plasmid DNA (10 μ g) was added and cells were electroporated (Bio-Rad, Gene Pulser II Electroporation System) at a voltage of 0.300 kV and a capacitance of 500 μ F. Poly(A) RNA (Oligotex midi kit, Qiagen, 2 μ g), isolated 36 hours after transfection, was separated on a 1.6% agarose formaldehyde gel, transferred to a nylon membrane, and hybridized with a polymerase chain reaction (PCR) product comprising exons 10 through 12 of the β -gluc gene that had been radioactively labeled by nick translation (Random Primed DNA Labeling Kit, Boehringer Mannheim). The membrane was subsequently stripped with boiling 0.5% SDS and probed with radioactively labeled zeocin cDNA. Northern blot results were quantitated using an Instant Imager (Packard). For subcellular fractionation, transiently transfected HeLa cells were trypsinized, washed in cold 1 \times phosphate buffered saline, and pelleted by centrifuging at 2040g for 5 min at 4°C. Cells were then resuspended in 200 μ l of 140 mM NaCl, 1.5 mM MgCl₂, 10 mM Tris-HCl (pH 8.6), 0.5% NP-40, and 1 mM DTT, vortexed, and incubated on ice for 5 min. Nuclei were pelleted by centrifuging at 12,000g for 45 s at 4°C and subsequently separated from the aqueous (cytoplasmic) phase.
 29. J. H. Graber, C. R. Cantor, S. C. Mohr, T. F. Smith, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14055 (1999).
 30. F. Chen, C. C. MacDonald, J. Wilusz, *Nucleic Acids Res.* **23**, 2614 (1995).
 31. Plasmid pG-26 (25) containing the CBP1 gene under the control of the GAL promoter and the LEU2 selectable marker was introduced into wild-type and *SKI7*-deleted strains. Yeast were grown in media lacking leucine and containing 2% galactose at 30°C to an OD₆₀₀ of 0.5. Cells were pelleted and resuspended in media lacking a carbon source. Glucose was then added to a final concentration of 2% at time 0 and time points were taken. RNA was extracted and analyzed by northern blotting with a labeled oligo probe oRP1067 (5'-CTCGGTCCTG-TACCGAACGAGACGAGG-3').
 32. We thank N. Martin, A. Atkin, C. Dieckmann, and M. Sands for the provision of valuable reagents. This work was supported by a grant from NIH (GM55239) (H.C.D.), the Howard Hughes Medical Institute (H.C.D., R.P.), and the Medical Scientist Training Program (P.A.F.).

22 October 2001; accepted 31 January 2002

Direct Demonstration of an Adaptive Constraint

Stephen P. Miller,¹ Mark Lunzer,¹ Antony M. Dean^{1,2,*}

The role of constraint in adaptive evolution is an open question. Directed evolution of an engineered β -isopropylmalate dehydrogenase (IMDH), with coenzyme specificity switched from nicotinamide adenine dinucleotide (NAD) to nicotinamide adenine dinucleotide phosphate (NADP), always produces mutants with lower affinities for NADP. This result is the correlated response to selection for relief from inhibition by NADPH (the reduced form of NADP) expected of an adaptive landscape subject to three enzymatic constraints: an upper limit to the rate of maximum turnover (k_{cat}), a correlation in NADP and NADPH affinities, and a trade-off between NAD and NADP usage. Two additional constraints, high intracellular NADPH abundance and the cost of compensatory protein synthesis, have ensured the conserved use of NAD by IMDH throughout evolution. Our results show that selective mechanisms and evolutionary constraints are to be understood in terms of underlying adaptive landscapes.

The old notion of natural selection as an omnipotent force in biological evolution has given way to one where adaptive processes are constrained by physical, chemical, and biological exigencies (1–4). Whether constraint and/or stabilizing selection explain phenotypic stasis, in the fossil record and in phylogenies, remains an open question (5). Direct experimental tests of constraint are scarce (6–9). Even tight correlations among traits, at once suggestive of constraint, can be broken by artificial selection to produce new phenotypic combinations (8, 9). Despite all circumstantial evidence, results from direct experimental tests imply that selection is largely unconstrained.

The direct experimental test for constraint is conceptually simple. A phenotype is subjected to selection (natural or artificial) in an attempt to break the postulated constraint (6–9). A response to selection indicates a lack of constraint. No response to selection indicates the presence of a constraint. However, the cause of a constraint is rarely specified because the etiologies of most phenotypes are not well understood, their relationships to fitness

are usually opaque, and a lack of response to selection may reflect nothing more than a lack of heritable variation (4, 7). If the cause of a constraint is to be elucidated, it must be for a simple phenotype whose relationship to fitness is understood.

Coenzyme use by β -isopropylmalate dehydrogenase (IMDH) is a simple phenotype whose etiology and relationship to fitness are understood (10, 11). IMDHs catalyze the oxidative decarboxylation of β -isopropylmalate to α -ketoisocaproate during the biosynthesis of leucine, an essential amino acid. All IMDHs use nicotinamide adenine dinucleotide (NAD) as a coenzyme (cosubstrate). This invariance of function among IMDHs hints at the presence of ancient constraints, even though some related isocitrate dehydrogenases (IDHs) use NADP instead (12, 13).

Structural comparisons with related NADP-using IDHs identify amino acids controlling coenzyme use (14–16) (Fig. 1A). Introducing five replacements (Asp²³⁶ \rightarrow Arg, Asp²⁸⁹ \rightarrow Lys, Ile²⁹⁰ \rightarrow Tyr, Ala²⁹⁶ \rightarrow Val, and Gly³³⁷ \rightarrow Tyr) into the coenzyme-binding pocket of *Escherichia coli* *leuB*-encoded IMDH by site-directed mutagenesis causes a complete reversal in specificity (10, 11): NAD performance ($k_{\text{cat}}^{\text{NAD}}/K_m^{\text{NAD}}$, where K_m is the Michaelis constant) is reduced by a factor of 340, from $68 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$ to 0.2×10^3

¹BioTechnology Institute, ²Department of Ecology, Evolution and Behavior, University of Minnesota, St. Paul, MN 55108, USA.

*To whom correspondence should be addressed. E-mail: deanx024@umn.edu

$M^{-1} s^{-1}$, whereas NADP performance ($K_{cat}^{NADP}/K_m^{NADP}$) is increased by a factor of 70, from $0.49 \times 10^3 M^{-1} s^{-1}$ to $34 \times 10^3 M^{-1} s^{-1}$. The engineered LeuB[RKYVYR] mutant (the final R represents Arg³⁴¹, already present in wild-type *E. coli* IMDH) is as active and as specific toward NADP as the wild-type enzyme is toward NAD. Evidently, protein architecture has not constrained IMDH to use NAD since the last common ancestor.

Despite similar *in vitro* performances, the NADP-specific LeuB[RKYVYR] mu-

tant is less fit than the NAD-specific wild type (11). IMDHs, wild-type and mutant alike, display a factor of 30 higher affinity for the reduced form of NADP (NADPH) (coproduct) than for NADP (Fig. 1B). We suggested the LeuB[RKYVYR] mutant is subject to intense inhibition by intracellular NADPH, which is far more abundant *in vivo* than is NADP (11, 17). The inhibition slows leucine biosynthesis, reduces growth rate, and lowers Darwinian fitness (Fig. 1C). The wild type retains high fitness because its affinity for NADPH is

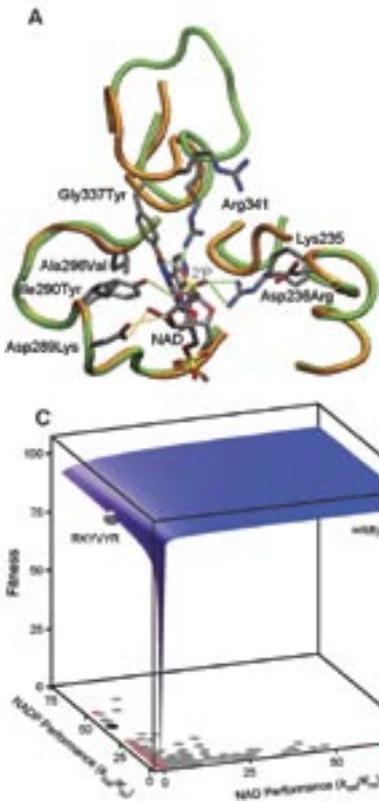


Fig. 1. The molecular anatomy of an adaptive constraint. **(A)** Structural alignment of the coenzyme binding pockets of *E. coli* IMDH (14) (brown main chain) with bound NAD modeled from *Thermus thermophilus* IMDH (15) and *E. coli* IDH (16) (green main chain) with bound NADP. Only key residues are shown (gray, carbon; red, oxygen; blue, nitrogen; yellow,

phosphorus) with labels designating the amino acid and site number in IMDH followed by the amino acid in IDH. Coenzyme use is determined by H bonds to NAD (brown lines) and to the 2'-phosphate (2'P) of NADP (green lines). **(B)** Mutations in the coenzyme binding pocket that stabilize NADP also stabilize NADPH. The ensuing correlation in affinities for NADP (K_m^{NADP}) and NADPH (K_i^{NADPH}) seen with engineered mutants (circles) (11) is retained in the screened mutants (dots). **(C)** The adaptive landscape for coenzyme use by IMDH (11) showing the phenotype-fitness map (blue surface) determined in chemostat competition and described by equation S5 (11, 18) and the predicted distribution of performances for 512 mutants in the coenzyme binding pocket (gray dots), with wild type (red dot), LeuB[RKYVYR] (black dot), and single amino acid replacements in LeuB[RKYVYR] coenzyme binding pocket (pink dots). The fitness of LeuB[RKYVYR] is hypothesized to be lower than in the wild type because of strong inhibition by abundant intracellular NADPH (17)

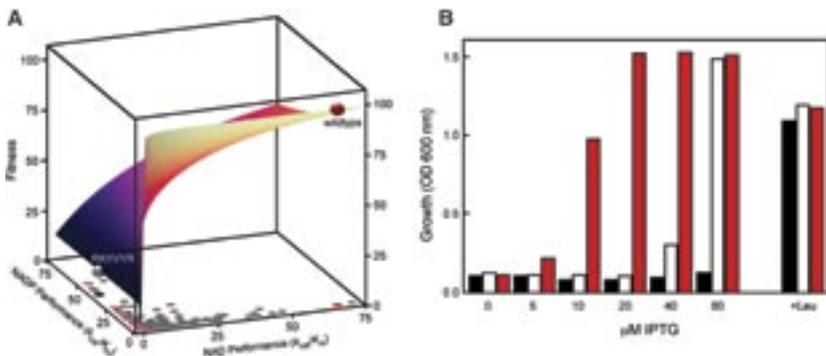


Fig. 2. The phenotypic basis of the genetic screen. Lowered expression in the presence of excess glucose brings IMDH to saturation with isopropylmalate, increasing coenzyme affinities. **(A)** Lowering IMDH expression in the chemostat-derived adaptive landscape [lower concentration of E in equation 55 (21, 18)] is predicted to reduce the fitness of LeuB[RKYVYR] (white sphere) far more than that of the wild type (red sphere). **(B)** IPTG-controlled expression of IMDHs ligated downstream of the T7 promoter in pETCoco in strain RFS[DE3] ($leuA^+B^{am}C^+$, with T7 RNA polymerase expressed from a chromosomal $lacUV5$ promoter) confirms that lower expression affects growth in minimal glucose medium of the LeuB[RKYVYR] (white) more than in the wild type (red). Plasmids lacking LeuB (black) are incapable of growth except in the presence of leucine.

low, whereas inhibition by NADH, which is far less abundant than NAD *in vivo*, is ineffective. Perhaps as a consequence of differences in Michaelis complex structure (18), IDH is not subject to such intense NADPH inhibition and hence could evolve NADP use.

We hypothesize that IMDH is constrained to use NAD because the strong inhibition associated with NADP use reduces fitness. However, identifying the mechanism of selection (NADPH inhibition) is not synonymous with identifying the causes of constraint. Mutations that increase k_{cat}^{NADP} (maximum rate of NADP turnover), that break the correlation in NADP and NADPH affinities (Fig. 1B), or that eliminate the trade-off in coenzyme performances (Fig. 1C) could each benefit the LeuB[RKYVYR] mutant without compromising its performance with NADP (18). We therefore hypothesize that IMDH is constrained to use NAD by three causes: an upper limit to k_{cat}^{NADP} , an unbreakable correlation in the affinities of NADP and NADPH, and an inescapable trade-off in coenzyme performance.

We used directed evolution (targeted random mutagenesis and selective screen-

ing) (19–21) to test whether NADP-specific IMDHs with higher fitness could be isolated. Random substitutions were introduced into $leuB$ [RKYVYR] by means of error-prone polymerase chain reaction (18). Mutated alleles were ligated downstream of the T7 promoter in pETCoco (a stable single-copy vector) and transformed into strain RFS[DE3] ($leuA^+B^{am}C^+$, with T7 RNA polymerase expressed from a chromosomal $lacUV5$ promoter). Sequencing unscreened plasmids revealed that, on average, each 1110–base pair $leuB$ [RKYVYR] received two nucleotide substitutions. From the pattern of base substitutions in these mutants and assuming Poisson statistics, we estimate that only 10.5 (0.4%) of the 2431 possible amino acid replacements were missing in our mutant library (18).

Our experimental design used decreased IMDH expression to provide a simple selective screen for mutations in $leuB$ [RKYVYR] that increase growth and/or fitness. As predicted from the phenotype-fitness map (Fig. 2A), selection against $leuB$ [RKYVYR] intensified as IMDH expression was lowered in the presence of excess glucose (Fig. 2B). Using 40 μ M isopropyl- β -D-thiogalacto-

Table 1. Kinetic effects of amino acid replacements in *LeuB*[*RKYVYR*] isolated at 48 hours growth. Standard errors are <13% of estimates. Single-letter abbreviations for amino acid residues: 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; Y, Tyr.

Enzyme	NADP			NAD			Preference (A/B)	K_i^{NADPH} (μM)
	Performance (A)			Performance (B)				
	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($\text{mM}^{-1} \text{s}^{-1}$)	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($\text{mM}^{-1} \text{s}^{-1}$)		
Site								
DDIAGR (wild type)	8400	4.10	0.4900	101	6.90	68.3200	0.007	254.30
RKYVYR	183	6.20	33.8800	4108	0.83	0.2000	169.400	9.90
			<i>Beneficial replacements</i>					
K235N	3919	5.95	1.5200	6356	0.50	0.0800	19.000	129.00
K235N,-3(M),A60V	1744	5.06	2.9000	5395	0.44	0.0800	35.700	56.00
K235R	554	6.39	11.5000	3292	0.36	0.1100	106.100	26.00
K235R,E63V	658	6.75	10.3000	3748	0.46	0.1200	84.100	28.00
K289E*	3449	2.71	0.7900	4897	1.98	0.4000	2.000	133.50
K289E*,D317E	3554	3.62	1.0200	5580	2.09	0.3700	2.800	112.00
K289E*,D87Y,S182F	2551	2.54	1.0000	5750	3.65	0.6300	1.600	105.20
K289E*,F102L	3891	3.66	0.9400	4823	1.10	0.2300	4.100	131.90
K289M	2216	4.13	1.8600	4034	0.48	0.1200	15.500	78.60
K289M,F170L	1945	2.50	1.2900	2947	0.48	0.1600	8.100	87.20
K289N*	1141	6.32	5.5400	4997	1.14	0.2300	24.300	44.00
K289N*,R187H	1146	5.83	5.0900	3069	0.79	0.2600	19.600	44.70
K289N*,H367Q	1248	5.27	4.2200	4555	1.05	0.2300	18.300	57.80
K289T	1596	6.18	3.8700	4056	0.78	0.1900	20.400	71.60
K289T,Q157H	1696	5.66	3.3400	5038	0.93	0.1800	18.600	72.60
Y290C	988	5.21	5.2700	3792	1.02	0.2700	19.600	49.00
Y290C,R152C	910	3.15	3.4600	3844	0.90	0.2300	14.800	43.00
Y290D	10213	3.13	0.3100	9040	0.82	0.0900	3.400	344.00
Y290F*	1658	5.59	3.3700	3110	0.65	0.2100	16.000	59.70
Y290F,P975,D314E	1097	4.64	4.2300	4158	0.71	0.1700	24.800	57.40
Y290N,N52T	2381	3.44	1.4400	10660	0.26	0.0200	59.500	90.00
			<i>Replacements with beneficial 5'-leader mutations</i>					
RKYVYR	183	6.20	33.8800	4108	0.83	0.2000	169.400	9.90
-3(M)†,E66D	208	2.97	14.2800	5002	0.30	0.0600	238.000	11.90
-3(M)†,E66D,E331D	267	7.99	29.9000	3403	0.34	0.1000	296.900	14.00
E82D	221	5.44	24.6200	4735	0.35	0.0700	351.700	14.10
K100R,A229T,L248M	137	5.58	40.3700	4079	0.45	0.1100	370.300	11.50
G156R,K289R,Y311H	292	5.40	18.4900	4464	0.54	0.1200	144.800	17.40
E173D	219	6.39	29.1800	5304	0.89	0.1700	171.600	6.60
Y337N,G131	171	5.71	33.3900	3367	0.24	0.0700	468.500	8.00
Y337H,P855,E181K	318	5.21	16.3800	2409	0.22	0.0900	183.600	12.80

*Switch from NADP to NAD requires two base substitutions in one codon. This is a possible transitional amino acid produced by a single base substitution (11). †The -3(M) designates a G to A substitution at base -3 that creates a new start codon.

pyranoside (IPTG) to induce a low level of IMDH expression, we found that cells harboring *leuB*[*WT*] formed large colonies at 24 hours, whereas cells harboring *leuB*[*RKYVYR*] barely formed pinprick colonies at 48 hours. We chose colony formation at 48 hours as the least stringent cri-

terion compatible with reliably identifying beneficial mutations in *leuB*[*RKYVYR*]. Longer periods of growth (or higher concentrations of IPTG) allowed unmutated *leuB*[*RKYVYR*] to form colonies, whereas shorter periods (or lower concentrations of IPTG) produced no colonies.

Of the 100,000 mutated plasmids screened, 134 (representing 107 distinct isolates) formed colonies within 48 hours (table S1). Each had either a substitution in the 5' leader sequence upstream of *leuB*[*RKYVYR*], an amino acid replacement in the coenzyme binding pocket, or both. Upstream substitutions occurred at three nucleotide positions: -3, -9, and -14 relative to the *leuB* AUG start codon (fig. S1). Ten amino acid replacements were found at three codons in the coenzyme binding pocket.

At first glance, the positive response to selection might suggest that coenzyme use by IMDH is unconstrained. Upstream substitutions in the Shine-Dalgarno sequence (positions -9 and -14), as well as a new AUG start codon (position -3) that replaces the less efficient GUG start codon, presumably derive their benefits through increases in expression because their kinetics are unchanged. Unexpectedly, however, beneficial amino acid replacements in the coenzyme binding pocket eliminate H bonds to the 2'-phosphate of NADP to cause striking reductions in NADP performance (Table 1). Although some mutants have improved NAD performance, others remain unchanged and several show reduced NAD performance. Isolated amino acid replacements outside the coenzyme binding pocket, which might have been expected to increase $k_{\text{cat}}^{\text{NADP}}$ or to break the correlation in affinities between NADP and NADPH (K_m^{NADP} and K_i^{NADPH}), have no detectable functional effects (table S2, those associated with beneficial 5'-leader mutations in Table 1). No doubt they, along with 126 silent substitutions (table S1), hitchhiked through the genetic screen with the beneficial mutations.

Our results suggest that increases in expression are beneficial, whereas increases in NADP performance are not. This seeming paradox is resolved if there are no mutations capable of breaking the upper limit to $k_{\text{cat}}^{\text{NADP}}$ the correlation in affinities for NADP and NADPH, or the trade-off in

coenzyme performance. With these constraints, and with reduced expression in the genetic screen, the phenotype-fitness map near *LeuB*[*RKYVYR*] remains flat with respect to increases in NADP performance (Fig. 2A). By contrast, severe losses of NADP performance are predicted to be beneficial as correlated reductions in the affinities for NADPH free up IMDH for use with abundant NAD. As predicted, all beneficial *LeuB*[*RKYVYR*] mutants have reduced affinities for both NADP and NADPH (Table 1). Unaffected by constraints, increases in expression are unconditionally beneficial (18). These results support the hypothesis that NADP-specific IMDHs function poorly in vivo because of strong inhibition by abundant NADPH. That reductions in NADP performance and increases in expression are both beneficial are the predicted consequences of a phenotype-fitness map constrained by an upper limit to $k_{\text{cat}}^{\text{NADP}}$, a correlation in affinities for NADP and NADPH, and a trade-off in coenzyme performance.

Breaking any one constraint would allow NADP-specific IMDHs to evolve. Yet no mutant increases $k_{\text{cat}}^{\text{NADP}}$, no mutant uncouples the affinities for NADP and NADPH, and no mutant breaks the trade-off in coenzyme performance. These conclusions are not the result of a selective screen that is too stringent. Of 35 colonies, representing 26 distinct mutants, that appeared after the 48-hour limit, 17 had no nucleotide substitutions in the 5' leader sequence or amino acid replacements in *LeuB*[*RKYVYR*] (table S3). Amino acid replacements in the other mutants neither improved enzyme performance nor decreased NADPH inhibition (table S4). That no additional beneficial mutations were recovered with relaxed criteria demonstrates that the selective screen was not overly stringent.

Nor are the results a consequence of inadequate sampling of protein sequence space. Natural adaptive evolution fixes advantageous mutations sequentially (22–

24). Indeed, experimental evolution demonstrates that advantageous double mutants in the evolved β -galactosidase of *E. coli* are not evolutionarily accessible and therefore must be accumulated as sequential advantageous mutations (25). Hence, screening all single substitutions is a sufficient sampling of protein sequence space for robust evolutionary conclusions. We estimate that only 0.04 advantageous amino acid replacements are missing from the mutant *leuB*[*RKYVYR*] library (18). We conclude that mutations capable of breaking the limit, the correlation, or the trade-off are unlikely to ever be fixed in populations because they are exceedingly rare (they may not exist), because they are minimally advantageous, or both.

The two remaining ways to evolve an NADP-specific IMDH are to reduce intracellular NADPH pool and, as our results show, to increase expression. Reducing intracellular NADPH relieves the inhibition but, as experiments deleting sources of NADPH show (13), the disruption to the rest of metabolism costs far more than the

benefit to be gained. The phenotype-fitness map (Fig. 1C) imposes a law of diminishing returns such that *LeuB*[*RKYVYR*] must be expressed above wild-type levels by a factor of 100 to overcome the inhibition by NADPH (18). Diverting resources away from other metabolic needs toward compensatory protein synthesis would impose a protein burden (26–29) sufficient to prevent the evolution of NADP-specific IMDHs.

The production of unnatural phenotypes, by artificial selection or molecular engineering, is not sufficient to conclude that evolutionary constraints are absent entirely. Rather, potential constraints underlying a conserved phenotype can be identified from the relationships among genotype, phenotype, and fitness that define an adaptive landscape. Experimental evolution can then be used to test their existence. Using this approach, we have shown how certain structure-function relationships in IMDH have constrained its coenzyme phenotype since the last common ancestor.

References and Notes

1. S. J. Gould, R. C. Lewontin, *Proc. R. Soc. London Ser. B* **205**, 581 (1979).
2. J. Antonovics, P. H. van Tienderen, *Trends Ecol. Evol.* **6**, 166 (1991).
3. M. R. Rose, G. V. Lauder, *Adaptation* (Academic Press, San Diego, CA, 1996).
4. M. Pigliucci, J. Kaplan, *Trends Ecol. Evol.* **15**, 66 (2000).
5. N. Eldredge *et al.*, *Paleobiology* **31**, 133 (2005).
6. M. Travisano, J. A. Mongold, A. F. Bennett, R. E. Lenski, *Science* **267**, 87 (1995).
7. H. Teotônio, M. R. Rose, *Nature* **408**, 463 (2000).
8. P. Beldade, K. Koops, P. M. Brakefield, *Nature* **416**, 844 (2002).
9. W. A. Frankino, B. J. Zwaan, D. L. Stern, P. M. Brakefield, *Science* **307**, 718 (2005).
10. R. Chen, A. Greer, A. M. Dean, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 12171 (1996).
11. M. Lunzer, S. P. Miller, R. Felsheim, A. M. Dean, *Science* **310**, 499 (2005).
12. A. M. Dean, G. B. Golding, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 3104 (1997).
13. G. Zhu, G. B. Golding, A. M. Dean, *Science* **307**, 1279 (2005); published online 13 January 2005 (10.1126/science.1106974).
14. G. Wallon *et al.*, *J. Mol. Biol.* **266**, 1016 (1997).
15. J. H. Hurley, A. M. Dean, *Structure* **2**, 1007 (1994).
16. J. H. Hurley, A. M. Dean, D. E. Koshland Jr., R. M. Stroud, *Biochemistry* **30**, 8671 (1991).
17. T. Penfound, J. W. Foster, in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, F. C. Neidhardt, Ed. (ASM Press, Washington, DC, ed. 2, 1996), pp. 721–730.
18. See supporting material on Science Online.
19. O. Kuchner, F. H. Arnold, *Trends Biotechnol.* **58**, 1 (1997).
20. F. H. Arnold, P. L. Wintrobe, K. Miyazaki, A. Gershenson, *Trends Biochem. Sci.* **26**, 100 (2001).
21. L. G. Otten, W. J. Quax, *Biomol. Eng.* **22**, 1 (2005).
22. S. Wright, in *Proceedings of the Sixth International Genetics Congress*, D. F. Jones, Ed. (Brooklyn Botanic Garden, Menasha, WI, 1932), pp. 356–366.
23. J. Maynard Smith, *Nature* **225**, 563 (1970).
24. D. M. Weinreich, N. F. Delaney, M. A. Depristo, D. L. Hartl, *Science* **312**, 111 (2005).
25. B. G. Hall, *Antimicrob. Agents Chemother.* **46**, 3035 (2002).
26. S. Zamenhof, H. H. Eichhorn, *Nature* **216**, 456 (1967).
27. K. J. Andrews, G. D. Hegeman, *J. Mol. Evol.* **8**, 317 (1976).
28. D. E. Dykhuizen, *Evolution* **32**, 125 (1978).
29. A. L. Koch, *J. Mol. Evol.* **19**, 455 (1983).
30. Supported by NIH grant GM060611 (A.M.D.). We thank three anonymous reviewers whose constructive criticisms we found invaluable.

Supporting Online Material

www.sciencemag.org/cgi/content/full/314/5798/458/DC1
Materials and Methods
SOM Text
Figs. S1 to S3
Tables S1 to S5
References

4 August 2006; accepted 20 September 2006
10.1126/science.1133479

Large Insertions: Two Simple Steps Using QuikChange® II Site-Directed Mutagenesis Kits

Abstract

This Technical Note describes a modified method to introduce large insertions in plasmids in two simple steps using the QuikChange® II Site-Directed Mutagenesis Kit. Dr. Wenge Wang in Dr. Wafik S. El-Deiry's lab at the University of Pennsylvania has used the QuikChange kit and a modification of the method described by Geiser et al. (2001)¹ to perform a large insertion of a PCR product into a target plasmid. He had inserted the enhanced green fluorescent protein (EGFP) open reading frame (760 bp) following the C-terminus of death receptor TRAIL receptor 1/death receptor 4 (TRAIL-R1/DR4) right after the signal sequence, using the QuikChange II site-directed mutagenesis kit with some modifications to the basic protocol. Dr. Wang transfected cells with the recombinant plasmid and it generated a fusion protein with the correct molecular size (Figure 1).

Background

Dr. Wafik S. El-Deiry's lab is focused on studying the mechanism of action of the tumor suppressor p53 and the contribution of its downstream target genes to control cell growth. Their lab has identified a number of genes that are directly regulated by p53 and which can inhibit cell cycle progression (p21WAF1), induce apoptosis (KILLER/DR5, Bid, caspase 6, Traf4 and others) or activate DNA repair (DDB2). The research has provided knowledge into the tissue specificity of the DNA damage response in vivo and into the mechanism by which wild-type p53 sensitizes cells to killing by anti-cancer drugs.

High-Efficiency Mutagenesis Method

All QuikChange® kits² offer a one-day method to introduce point mutations, amino acid substitutions, small insertions, and deletions in virtually any double-stranded

plasmid template at efficiencies greater than 80% (Figure 2). The QuikChange II and QuikChange II XL kits feature our highest fidelity DNA polymerase, our gold standard for accuracy, and a linear amplification strategy. Together, they reduce the mutation frequency due to incorporation errors typically caused by using an exponential PCR amplification approach. The result is high-efficiency mutagenesis without unwanted errors.

Ideal for Large Insertions and Deletions

We have used the QuikChange kit method for small insertions and deletions (~12 bp) and observed greater than 80% efficiency². Many of our customers have used the QuikChange method or have modified it to introduce deletions of 31 bp³, 87 bp⁴, and 3 kb⁵ or insertions of 31 bp³ and up to ~1 kb⁵. In order to achieve large insertions of up to 1 kb, megaprimers must be generated from an initial PCR reaction. To ensure the highest mutagenesis efficiency, we recommend gel purification with our StrataPrep® PCR Purification Kit. To create the megaprimers, a minimum of 20 bp upstream and downstream of the insertion sequence should be complementary to your template vector that will be used in the QuikChange kit reaction. Therefore, each oligo used in the initial PCR reaction will require 20 bp overhangs on the 5' ends.

Using the QuikChange kit for large deletions is even easier since oligos can be synthesized, rendering the initial PCR reaction and fragment purification unnecessary. Again, 20 to 30 bp may be required to bind to your template DNA both upstream and downstream of the region that you want to delete. However, the largest oligo required should not exceed 60 bp.

This Technical Note describes performing

a large insertion that can be completed in two simple steps. First, PCR-amplify the megaprimer. Second, add your gel purified megaprimer to our QuikChange site-directed mutagenesis kit (Figure 2). This strategy avoids tedious and time-consuming sub-cloning. With our QuikChange kits, you will have confidence in generating an error free clone while saving valuable time.

Materials & Methods

Part I. PCR reaction to generate the megaprimer

Reaction and cycling conditions.

PCR reagents: PfuUltra[®] II Fusion HS DNA Polymerase (Stratagene)

Thermal Cycler used: TECHNE: Touchgene[®] Gradient

- The forward primer: GACTCCGAATCC CGGGAGCGCAGCGGTGAGCAAGG GCGAGGAG, the first 25 bp overlaps the target vector, the remaining primer sequence is complementary to EGFP. The primers were resuspended in 50 mM NaCl.
- The reverse primer: CGCGGCTGCCTC TGCCACTCTTGACAGCTCGTCCA TGCC, the first 22 bp targets the vector and the remaining primer sequence is complementary to EGFP. The primers were resuspended in 50 mM NaCl.
- Template for PCR reaction: 200 ng of plasmid DNA (pEGFP-N1; Clontech)
- PfuUltra[®] II Fusion HS DNA Polymerase (Stratagene)
- PCR product was purified with QIAGEN'S QIAEX II Gel Extraction Kit
- The quantity of the PCR product was estimated on agarose gel with Ethidium Bromide

Note: The size of the PCR product is 760 bp, which includes the insert (EGFP) of 714 bp plus overlapping sequence of the target vector.

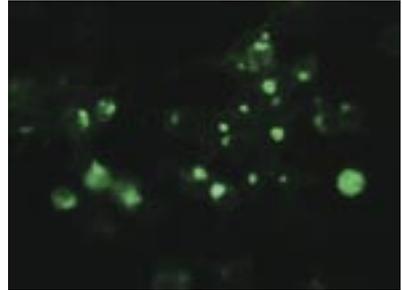


Fig. 1 Enhanced green fluorescent protein (EGFP) was inserted in the open reading frame (760 bp) following the C-terminus of death receptor TRAIL receptor 1/death receptor 4 (TRAIL-R1/DR4) right after the signal sequence. Cell staining shows the member protein fused with EGFP.

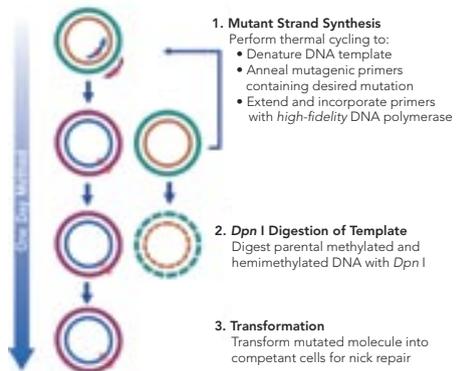


Fig. 2. The QuikChange[®] II One-Day Site-Directed Mutagenesis Method. 1. Mutant strand synthesis. 2. Dpn I Digestion of parental DNA template. 3. Transformation of the resulting annealed double-stranded nicked DNA molecules. After transformation, the XL-1 Blue *E. coli* cell repairs nicks in the plasmid.

Part II. The QuikChange II mutagenesis reaction can be set up using the following guidelines (Tables 1 and 2).

- Primer for the insertion reaction 300-500 ng of PCR product (760 bp EGFP) to serve as a megaprimer.

Note: Follow reaction with *Dpn* I digestion and transformation per the QuikChange® II manual. For PfuUltra® II Fusion HS DNA Polymerase information please refer to the manual.

Results ■

The results of this mutagenesis experiment demonstrate that large PCR fragments can be successfully inserted into the target plasmid DNA. The 760 bp PCR fragment had at each end only short regions of homology where the insertion occurred. There were approximately 100 colonies produced. About 20 colonies were selected for further analysis of mutagenesis efficiency, and they were 100% positive by sequencing.

Acknowledgments ■

This Technical Note was written based on data provided by Wenge Wang, M.D., Ph.D., in Dr. Wafik S. El-Deiry's laboratory, Division of Hematology and Oncology in the Department of Medicine at the University of Pennsylvania.

TABLE 1. Reaction Set-up

Component	Amount per reaction
Distilled water (dH ₂ O)	X µl
10x QuikChange reaction buffer	5.0 µl
dNTP mix	1 µl
DNA template (50 ng)	X µl
Megaprimer (300 - 500 ng)	X µl
High Fidelity DNA polymerase (2.5 U/µl)	1.0 µl (2.5U)
Total reaction volume	50 µl

TABLE 2. Cycling Method

Segment	Cycles	Temperature	Time
1	1	95°C	30 seconds
2	5	95°C	30 seconds
		52°C	1 minute
		68°C	1 min/kb of plasmid* (7 min)
3	13	95°C	30 seconds
		55°C	1 minute
		68°C	1 min/kb of plasmid* (7 min)

* For example, a 7 kb plasmid would have a 7 minute extension time

References

1. Geiser, M., et al. (2001) *BioTechniques* 31: 88-92.
2. Papworth, C., Braman, J., and Wright, D.A. (1996) *Strategies* 9: 3-4.
3. Wang, W. and Malcolm, B. (1999) *BioTechniques* 26: 680-682.
4. Burke, T., et al. (1998) *Oncogene* 16: 1031-1040.
5. Makarova, O., et al. (2000) *BioTechniques* 29: 970-972.

Legal

QuikChange® and StrataPrep® are registered

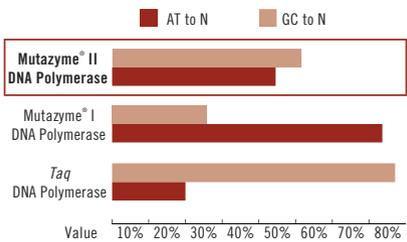
trademarks of Stratagene, an Agilent Technologies company. QIAGEN® is a registered trademark of QIAGEN. Touchgene® is the registered trademark of Techne Incorporated. Clontech® logo is the trademark of Clontech Laboratories, Inc. a. U.S. Patent Nos. 7,176,004; 7,132,265; 7,045,328; 6,734,293; 6,489,150; 6,444,428; 6,391,548; 6,183,997; 5,948,663; 5,932,419; 5,866,395; 5,789,166; 5,545,552, and patents pending.



The perfect assortment

With GeneMorph® II mutagenesis kits, a balanced spectrum of mutations is right at your fingertips

Our GeneMorph® II Kits include Mutazyme® II Polymerase, which delivers a balanced mutational spectrum.



GeneMorph® II random mutagenesis kits* feature our patented Mutazyme® II DNA polymerase, which delivers a balanced mutational spectrum with more robust yields than *Taq* polymerase under error-prone PCR conditions. This allows you to discover more key residues responsible for protein function easier and faster than before, thus enhancing the evolution of your protein.

- Simple protocol to control mutation frequency
- Efficient mutagenesis rates of 1 to 16 bases per kb
- Overcome poor PCR yield and mutational bias of *Taq* polymerase

u.s. and canada 1-800-424-5444 x3
europe 00800-7000-7000

www.stratagene.com

GeneMorph® II Mutagenesis Kits

GeneMorph® II Random Mutagenesis Kit	30 rxns	200550
GeneMorph® II EZClone Domain Mutagenesis Kit	10 rxns	200552

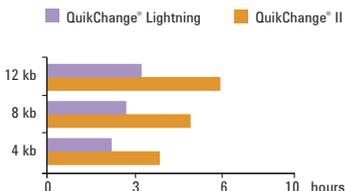
U.S. Patent No. 7,045,328; 6,863,216; 6,734,293; 6,480,150; 6,444,428; 6,183,997; 5,489,523 and patents pending. GeneMorph® and Mutazyme® are registered trademarks of Stratagene, an Agilent Technologies Company, in the United States.

Engineered for speed

2X Faster



QuikChange® Lightning Saves You Time



Time Comparison of QuikChange® Lightning to QuikChange® II

Our new QuikChange® Lightning Site-Directed Mutagenesis Kit* introduces point mutations, insertions, or deletions significantly faster than our previous kits. Using our most advanced high-fidelity enzyme technology, we accelerated the protocols while guaranteeing the highest accuracy for your site-directed mutagenesis.

- Greater than 80% mutagenesis efficiency
- High-fidelity linear amplification minimizes unwanted errors
- Entire procedure performed in under 3 hours

u.s. and canada 1-800-424-5444 x3
europe 00800-7000-7000

www.stratagene.com/QCLighting

Want twice the speed? Go to www.stratagene.com/QCLighting

QuikChange® Lightning Site-Directed Mutagenesis Kit 10 rxn 210518
30 rxn 210519

QuikChange® is a registered trademark of Stratagene, an Agilent Technologies Company, in the United States.
*U.S. Patent Nos 6,734,293; 6,489,150; 6,444,428; 6,391,548; 6,183,997; 5,948,663;
5,932,419; 5,866,395; 5,789,166; 5,545,552; 6,713,285 and patents pending.