

# Directed evolution of the surface chemistry of the reporter enzyme $\beta$ -glucuronidase

Ichiro Matsumura<sup>1</sup>, John B. Wallingford<sup>2,3</sup>, Neeraj K. Surana<sup>1,4</sup>, Peter D. Vize<sup>2</sup>, and Andrew D. Ellington<sup>1\*</sup>

<sup>1</sup>Institute of Cellular and Molecular Biology, ICMB A4800/MBB 3.424, 26th and Speedway, University of Texas, Austin, TX 78712. <sup>2</sup>Center for Developmental Biology, Department of Zoology, University of Texas, Austin, TX 78712. <sup>3</sup>Present address: Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA 94720. <sup>4</sup>Present address: Division of Biology and Biomedical Sciences, Washington University, St. Louis, MO 63108. \*Corresponding author (e-mail: [andy.ellington@mail.utexas.edu](mailto:andy.ellington@mail.utexas.edu)).

Received 29 December 1999; accepted 4 May 1999

The use of the *Escherichia coli* enzyme  $\beta$ -glucuronidase (GUS) as a reporter in gene expression studies is limited due to loss of activity during tissue fixation by glutaraldehyde or formaldehyde. We have directed the evolution of a GUS variant that is significantly more resistant to both glutaraldehyde and formaldehyde than the wild-type enzyme. A variant with eight amino acid changes was isolated after three rounds of mutation, DNA shuffling, and screening. Surprisingly, although glutaraldehyde is known to modify and cross-link free amines, only one lysine residue was mutated. Instead, amino acid changes generally occurred near conserved lysines, implying that the surface chemistry of the enzyme was selected to either accept or avoid glutaraldehyde modifications that would normally have inhibited function. We have shown that the GUS variant can be used to trace cell lineages in *Xenopus* embryos under standard fixation conditions, allowing double staining when used in conjunction with other reporters.

Keywords:  $\beta$ -glucuronidase, reporter gene, in vitro evolution, directed evolution, DNA shuffling, *Xenopus laevis*

Since plants express endogenous  $\beta$ -galactosidase activity, *lacZ* cannot be employed as a reporter gene<sup>1</sup>. Instead, the *Escherichia coli*  $\beta$ -glucuronidase gene (*gusA*, formerly *uidA*) has been developed as a reporter gene for plants, and has been widely used for over a decade<sup>2</sup>. Both chromogenic and fluorogenic GUS substrates have been synthesized<sup>3</sup>, allowing rapid nonradioactive assays. The GUS enzyme is stable and active under a variety of conditions<sup>1</sup>, even when fused to other sequences<sup>4</sup>.

The utility of GUS as a reporter, however, has been constrained in three ways. First, many animal systems, and some plants and plant-associated bacteria express endogenous glucuronidase activities<sup>2,3</sup>. Second, GUS activity is greatly reduced during tissue fixation by glutaraldehyde or formaldehyde, making it necessary to trade off retention of activity for preservation of tissue structure<sup>5</sup>. Third, both of these considerations drastically restrict the use of GUS as a reporter gene in vertebrate systems<sup>6</sup>.

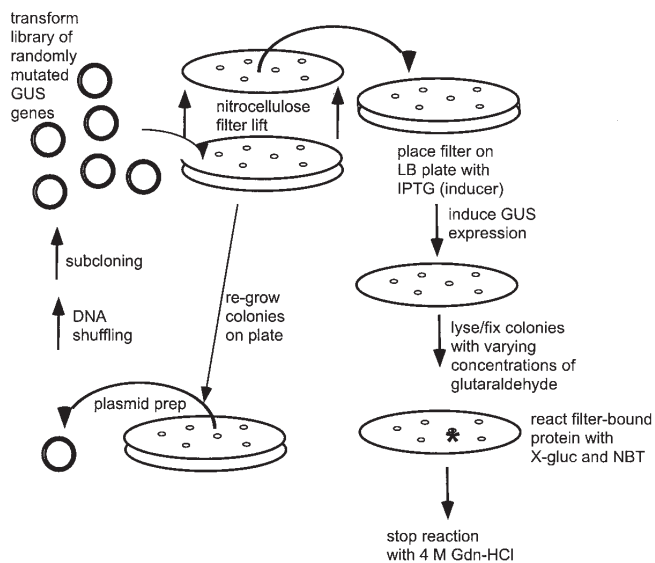
Enzymatic inactivation by aldehydes is largely due to the formation of Schiff bases with surface-accessible lysine residues<sup>7</sup>. While the removal of lysine residues by directed mutation might render an enzyme more resistant to fixatives, many surface lysines are critical for function and cannot be readily changed. The sequences of the *E. coli*<sup>8</sup>, human<sup>9</sup>, mouse<sup>10</sup>, rat<sup>11</sup>, and dog<sup>12</sup> homologs are known. Six of the 27 lysine residues in the *E. coli* protein are conserved in the other species and thus are likely essential. Moreover, to find what combination of the 27 lysine residues could be changed in order to increase resistance to fixatives without abrogating enzyme activity would require constructing and assaying a dauntingly large number of mutant enzymes. Therefore, in order to alter the surface chemistry of GUS, either to avoid or to accommodate aldehyde modifications without loss of enzyme activity, we employed a random mutation-

approach similar to those previously proven useful for altering enzyme substrate specificity<sup>13</sup> or thermostability<sup>14</sup>.

## Results

**Directed evolution of glutaraldehyde-resistant variants.** Random mutations were initially introduced into the *gusA* structural gene by mutagenic PCR<sup>15</sup>. Mutated PCR products were ligated into the expression vector *gusA*-pBS $\Delta$  and transformed into *E. coli*. When the library was induced on plates containing the chromogenic GUS substrate, 5-bromo-4-chloro-3-indolyl- $\beta$ -D-glucuronide (X-gluc), approximately 80% of the colonies were visibly less green than control colonies expressing only the chromosomal gene (see Experimental protocol).  $\beta$ -Glucuronidase functions as a tetramer<sup>16</sup>, so it was likely that many of the mutations in the highly expressed, plasmid-borne library had a dominant negative effect on the function of the chromosomal gene. This did not deter us from utilizing this library for screening experiments, since successive rounds of DNA shuffling should efficiently select against neutral or deleterious mutations<sup>13</sup>.

Nine thousand replica-plated colonies, each expressing a randomly mutated *gusA* gene, were exposed to buffer containing 0.2% glutaraldehyde for 20 min. The colony remnants were then incubated in buffer containing X-gluc and the histochemical indicator, nitroblue tetrazolium (NBT) (Fig. 1). The catalytic activity of the wild-type enzyme is greatly diminished under those conditions, indicating that the glutaraldehyde disrupts the cell membranes and covalently modifies many intracellular proteins, including GUS (cf. Figs 2A and B). Of all the variants examined, only 10 colonies reproducibly exhibited greater catalytic activity than control colonies expressing wild-type *gusA* (Fig. 2B). The corresponding colonies on master plates were isolated, and their expression vectors purified

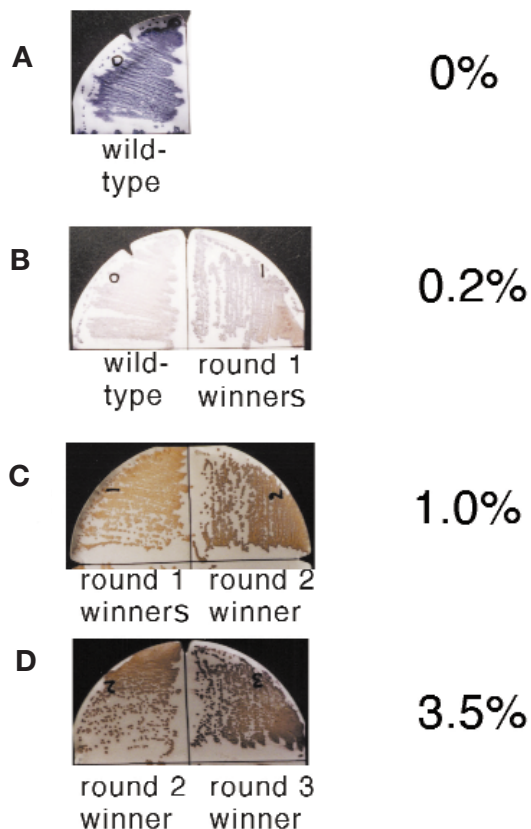


**Figure 1.** Screen for glutaraldehyde-resistant  $\beta$ -glucuronidase (GUS) function (sequence from top left). A library of randomly mutated  $\beta$ -glucuronidase genes (*gusA*) is subcloned into an inducible expression vector and transformed into *Escherichia coli*. The resulting colonies are transferred to a nitrocellulose filter, which is overlaid upon an agarose plate containing an inducer and incubated for 12–24 h at 37°C. The filter-bound colonies are incubated in buffer containing glutaraldehyde, then transferred to buffer containing the histochemical indicators of  $\beta$ -glucuronidase, X-gluc and NBT. The brief incubation in 4 M guanidine HCl (Gdn-HCl) arrests color development. Colonies that retain GUS activity are isolated from the original plate and randomly recombined by DNA shuffling for the next round of screening.

and pooled. The variant *gusA* genes were amplified using the PCR and randomly recombined by DNA shuffling<sup>17</sup>. We then screened 6,000 random recombinants in a second round for variants that retained catalytic activity after a 20 min incubation in 1.0% glutaraldehyde. Nine colonies contained variants that exhibited greater residual catalytic activity than the most resistant clone isolated in the first round of screening (Fig. 2C). These variants were again pooled, amplified, and randomly recombined. Then, 6,000 recombinants were screened in a third round for variants that retained catalytic activity after a 20 min incubation in 3.5% glutaraldehyde. Again, nine improved clones were isolated, one of which ( $GUS^{AR}$ ) reproducibly showed the greatest activity under the most stringent conditions (Fig. 2D).

**In vitro characterization of  $GUS^{AR}$ .** To determine whether the colony-lift assay significantly influenced the apparent fixative-resistant phenotype of  $GUS^{AR}$ , activity assays were also carried out in cell extracts. The *GUS*-deficient strain, pREP4/GMS407, was transformed with vectors that expressed either wild-type *gusA*, the  $GUS^{AR}$  variant, or the *lacZ*  $\alpha$ -fragment. Cell extracts from induced cultures were treated with glutaraldehyde or formaldehyde for 20 min at 23°C, and diluted 100-fold in buffer containing saturating concentrations of a *GUS* substrate, p-nitrophenyl  $\beta$ -D-glucuronide (PNPG). The extracts containing wild type or  $GUS^{AR}$  catalyzed the hydrolysis of PNPG; no hydrolysis was detected in the negative control extracts, in which only the *lacZ*  $\alpha$ -fragment was expressed (data not shown). Treatment of the extract containing wild-type *GUS* with only 0.04% glutaraldehyde for 20 min at 23°C reduced catalytic activity by  $99.6 \pm 0.24\%$ . In sharp contrast, the  $GUS^{AR}$  extract retained  $78.1 \pm 0.69\%$  of its activity after treatment with a fivefold higher (0.2%) concentration of glutaraldehyde (Fig. 3A).

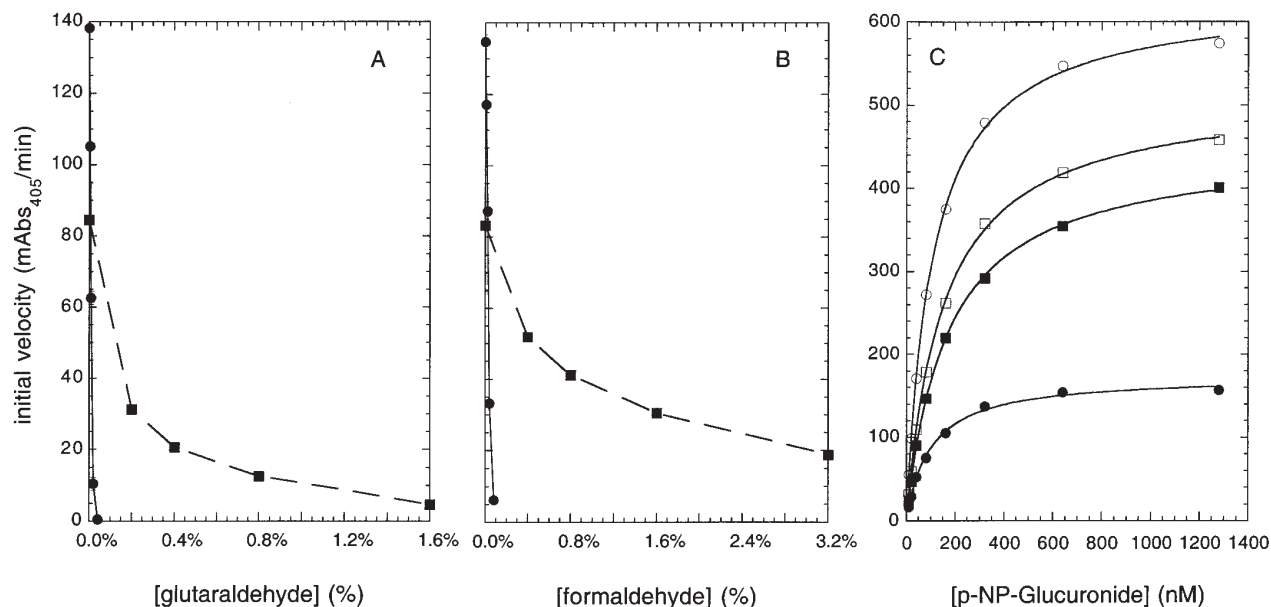
[Glutaraldehyde]



**Figure 2.** Detection of glutaraldehyde-resistant GUS activity. *Escherichia coli* cells transformed with vectors expressing the wild-type (A, B left), a pool of the ten glutaraldehyde-resistant variants from round 1 (B right, C left) or the most resistant variants from rounds 2 (C right, D left), or 3 (D right) were streaked onto noninducing plates. The colonies were propagated, induced, and treated for 20 min with the indicated concentrations of glutaraldehyde, then reacted with X-gluc and NBT, as described in the legend to Figure 1.

Extracts were also separately treated with formaldehyde to assess whether the fixative-resistant phenotype was specific to glutaraldehyde. Again, the  $GUS^{AR}$  variant exhibited much greater resistance to the fixative than did the wild-type *GUS*. The wild-type extract retained only  $4.6 \pm 0.05\%$  of its catalytic activity after treatment with 0.08% formaldehyde; the  $GUS^{AR}$  extract retained  $62.4 \pm 0.40\%$  activity after incubation with 0.4% formaldehyde (Fig. 3B). To determine how sequence and chemical modifications may have influenced *GUS* activity, we conducted kinetic studies of the mutant enzyme (Fig. 3C). The wild-type and evolved *gus* genes were subcloned, expressed as fusion proteins with N-terminal hexahistidine tags, and purified by immobilized metal ion adsorption chromatography. Purified enzymes were assayed with varying concentrations of PNPG (Fig. 3C). The kinetic parameters of the wild-type enzyme ( $K_M$  for the complex with PNPG =  $110 \pm 2.9 \mu\text{M}$ ;  $k_{cat}$  =  $920 \pm 7.3 \text{ s}^{-1}$ ) were very similar to those of the  $GUS^{AR}$  variant ( $K_M$  =  $150 \pm 3.9 \mu\text{M}$ ;  $k_{cat}$  =  $750 \pm 5.8 \text{ s}^{-1}$ ). The kinetic parameters of wild-type and evolved enzymes were also determined following reaction with a sublethal concentration (0.04%) of formaldehyde for 20 min at 23°C. Formaldehyde had a larger effect on the turnover number ( $250 \pm 17 \text{ s}^{-1}$  for the partially modified wild type,  $650 \pm 4.7 \text{ s}^{-1}$  for the modified  $GUS^{AR}$  variant), than on the Michaelis constants ( $99 \pm 6.7 \mu\text{M}$  and  $170 \pm 3.7 \mu\text{M}$  for the modified wild-type and  $GUS^{AR}$  enzymes,

## RESEARCH



**Figure 3.** GUS catalytic activity as a function of aldehyde concentration. *Escherichia coli* cell extracts containing the wild-type (circles) or evolved GUS<sup>AR</sup> (squares)  $\beta$ -glucuronidase were incubated for 20 min at 23°C in buffer containing the indicated concentrations of glutaraldehyde (A) or formaldehyde (B). The protein was then diluted 100-fold into buffer containing the chromogenic GUS substrate, p-nitrophenyl- $\beta$ -D-glucuronide (PNPG). Hydrolysis of the substrate was followed at 405 nm using a spectrophotometer (see Experimental Protocol). Each point represents the average of three initial velocity values; the points subsume the error bars. The control extract from an isogenic strain not expressing GUS does not have detectable activity (not shown). (C) Purified wild-type (circles) or evolved GUS<sup>AR</sup> (squares) enzymes were incubated for 20 min at 23°C in buffer containing 0% (empty symbols) or 0.04% (filled symbols) formaldehyde, then reacted with the indicated concentrations of PNPG. The initial velocity values were fitted to the Michaelis-Menten equation (lines); the derived kinetic parameters are presented in the text.

respectively). These results in conjunction with the cell extract data show that GUS<sup>AR</sup> is expressed at lower levels than the wild type, but is inherently more aldehyde resistant.

**Sequence and structural mapping of the evolved GUS<sup>AR</sup> variant.** Upon sequencing, the evolved *gusA* gene was found to contain the following amino acid substitutions: N66D, D151N, A219V, I396T, T480A, Q498R, D508E and K567R, as well as six silent mutations. Only one of these changes, the D508E mutation, results from a transversion, indicating a strong transition bias in our random mutagenesis methods. The amino acid sequences of the *E. coli* and human GUS proteins are 50% identical<sup>18</sup> and could be readily aligned by the algorithm devised by Needleman and Wunsch<sup>19</sup> using the program GAP 4.0 (Genetics Computing Group, Madison, WI). Both proteins are tetramers<sup>16,18</sup> and are virtually identical in substrate specificity<sup>1</sup>. The positions of the loci that were altered in the evolved *E. coli* enzyme could thus be tentatively mapped onto the crystal structure of the human GUS protein<sup>18</sup> (Fig. 4).

The 10 *gusA* mutants isolated in the first round of screening were sequenced; mutations were found at a frequency of three per 1.8 kb. Seven of the first-round variants contained amino acid substitutions (K567R, T480A, D508E or N66D/D151N) that were subsequently found in the most active third-round GUS<sup>AR</sup> variant. It is instructive that many of the single substitutions that confer modest resistance to aldehyde modification can interact additively or synergistically to confer robust resistance to aldehyde modification.

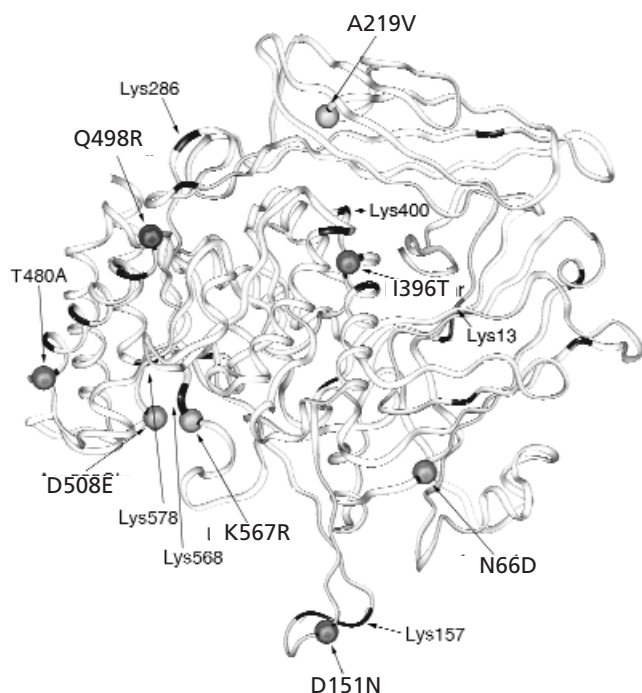
**GUS<sup>AR</sup> as a lineage tracer in *Xenopus* embryos.** The N358S mutant of GUS is a commercially available and commonly used reporter gene in plants<sup>4</sup>. The N358S mutation eliminates a cryptic glycosylation site, and should not affect its function in the cytoplasm; we chose this construct because it also contained the upstream sequences necessary for expression in eukaryotic cells<sup>20</sup>. In order to determine if the fixative-resistant GUS might also prove useful in other model organisms, transcripts encoding

GUS<sup>AR</sup> and N358S GUS were microinjected into 16- to 32-cell stage *Xenopus* embryos. Two days later, the embryos were fixed in 3.7% formaldehyde for 20 min and stained using a standard protocol for the detection of *lacZ* expression, except that X-gal was substituted for X-gal. The descendants of cells injected with the wild-type-like N358S GUS mRNA did not change color (Fig. 5A). Endogenous GUS activity was apparently also abrogated by the 20 min incubation in 3.7% formaldehyde. In contrast, the descendants of cells injected with the GUS<sup>AR</sup> mRNA turned bright blue-green (Fig. 5B).

In order to determine if multiple reporters might be used in tandem for lineage analysis, mRNAs encoding either N358S or the GUS<sup>AR</sup> were co-injected into embryonic cells along with mRNA encoding *lacZ*. When the embryos were first stained with X-gal, again only the cells that inherited the GUS<sup>AR</sup> mRNA turned blue-green (Fig. 6A and B). The embryos were subsequently stained with rose-gal, a  $\beta$ -galactosidase substrate that forms a red precipitate. In embryos that received either the N358S or GUS<sup>AR</sup>, some cells were colored red, indicating the inheritance of *lacZ*. However, in embryos that received GUS<sup>AR</sup> some cells or patches were also purple, indicating the co-inheritance of the GUS<sup>AR</sup> and *lacZ* (Fig. 6C and D). The *lacZ* mRNA in this experiment also served as an internal control that demonstrated that RNAs were entering cells and surviving until fixation.

## Discussion

**Mechanisms of aldehyde resistance.** The *E. coli* GUS protein contains 27 lysine residues, six of which are conserved among the sequenced GUS genes. Although the fraction of lysine residues that are modified by aldehydes is unknown, wild-type GUS activity is quite susceptible to even low levels of fixatives. For example, the catalytic activities of the wild-type (Fig. 3B) and N358S GUS (Figs 5A and 6A) are inactivated by <3.7% formaldehyde, while  $\beta$ -galactosidase, which contains 20 lysine residues, is not (Fig. 6). Taken together, these results suggest that one or more of the GUS



**Figure 4. Homology mapping of amino acid substitutions that confer aldehyde resistance.** The crystal structure of a subunit of the C $\alpha$  trace of human GUS<sup>18</sup> is shown. The amino acid sequences of the *E. coli* and human GUS proteins were aligned using the application GAP 4.0 (Genetics Computer Group) and were found to be 48.5% identical. The positions of lysine residues are darkened, and the conserved lysines are labeled. The positions of amino acid substitutions in the evolved GUS<sup>AR</sup> *E. coli* protein are shown as balls.

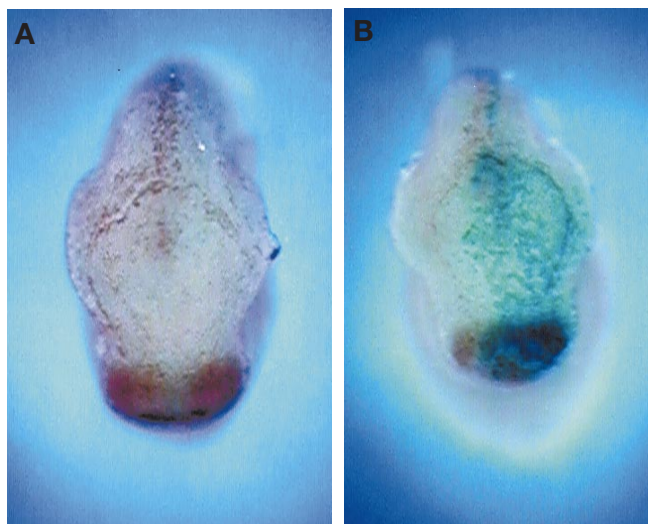
lysine residues is either itself critical for activity or presents a conjugation site that leads to functional disruption. However, identifying which of the many lysine residues in GUS were responsible for inhibition by fixatives would have been a daunting task. Instead, we relied on a random mutagenesis to identify GUS variants with catalytic activity resistant to aldehydes. Following three rounds of screening and amplification, we isolated an octuple-mutant GUS<sup>AR</sup> with catalytic activity resistant to roughly 80-fold higher levels of glutaraldehyde than the wild-type activity (Fig. 3A).

Surprisingly, only one of the amino acid substitutions, K567R, in the evolved GUS<sup>AR</sup> occurred at a lysine residue. Since AAA or AAG encodes lysine, the apparent transition bias in our random-mutagenesis method and the size of our initial library provided ample opportunities for each lysine to conservatively mutate into arginine (AGA or AGG). While it is possible that mutation of this single lysine was largely responsible for protection against aldehydes, this explanation is unlikely. Three of the 10 clones isolated in the first round of screening contain the K567R sequence substitution, but none of these first-round isolates are as resistant as any of the second-round isolates (Fig. 2C). The aldehyde resistance of GUS progressively increased over three rounds of screening and selection, and the final product had accumulated seven additional amino acid substitutions. The finding that amino acid substitutions that modulate protein function are dispersed in the primary and tertiary structure of GUS is congruent with previous attempts to evolutionarily engineer the physical and kinetic parameters of enzymes. Experiments that directed an increase in the catalytic activity of a *p*-nitrobenzyl esterase in organic enzymes yielded multiple sequence substitutions scattered throughout the tertiary structure<sup>21</sup>. Site-directed mutation studies of T4 lysozyme have shown that stabilizing amino acid changes, which occur in the core of that enzyme, are additive in effect<sup>22</sup>.

Interestingly, the seven non-lysine amino acid substitutions mapped onto the surface of the protein near lysine residues (Fig. 4). Protein structure is more highly conserved than protein sequence<sup>23</sup>, and since the primary sequences of the *E. coli* and human GUS enzymes are quite similar (48.5% identity<sup>8</sup>), it can be conservatively assumed that their tertiary structures also align well. Based on this assumption, we can advance hypotheses regarding the contributions of individual amino acid substitutions to aldehyde resistance. For example, Lys568 (*E. coli* numbering) is conserved among the sequenced GUS genes, and is in the active-site<sup>18</sup>. The C $\alpha$ -C $\alpha$  distance from Lys568 to the D508E substitution is 3.97 Å (Fig. 4). Since the lysine side chain is 7 Å in length, this adjacent sequence substitution might raise the pK<sub>a</sub> of the epsilon amino group of Lys568, thereby reducing its reactivity with aldehydes. Similarly, the K567R substitution already mentioned is within 3.81 Å of the active-site lysine, and mutation to arginine may prevent modification that could sterically interfere with substrate binding or catalysis.

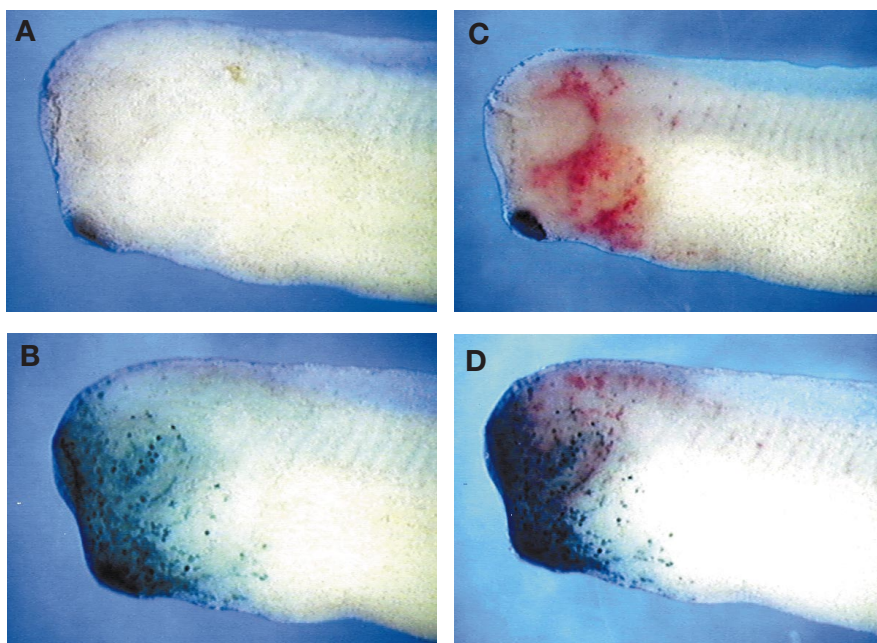
Similarly,  $\beta$ -glucuronidase is active only as a tetramer<sup>16,18</sup>, and lysines play a key role in its quaternary structure. To the extent that modifications of interfacial lysines disrupt quaternary structure and enzymatic function, adjacent amino acid substitutions could render these lysines less reactive. In this regard, the loop containing D151N and three lysines is <5 Å away from the  $\alpha$ -helix of the adjacent subunit containing T480A and three other lysines. These amino acid substitutions could also prevent structural and functional disruption by independently increasing the affinity between the subunits. For example, the A219V substitution also maps to the other interface, although it is not immediately adjacent to any lysines.

Overall, it appears the surface chemistry of the enzyme has coordinately evolved either to cause lysines to be less reactive or to functionally accommodate covalent modification of lysines. Our results suggest that there may be multiple possible routes by which proteins could be adapted to function in a wide variety of fixatives or solvent systems. More importantly, they suggest a way of augmenting protein chemistry by introducing amino acids with novel surface conjugates.



**Figure 5. Expression of GUS<sup>AR</sup> in *Xenopus* embryos.** Embryos at the 16- to 32-cell stage were injected with 1 ng of mRNA encoding N358S (A) or GUS<sup>AR</sup> (B) and fixed two days later in 3.7% formaldehyde for 20 min. GUS activity was detected using the chromogenic substrate X-gluc (light blue/green). The reddish-purple color of the cement gland of the embryo shown in (A) is from a natural pigment, and the blue color of the embryo shown in (B) is from GUS<sup>AR</sup> activity.

## RESEARCH



**Figure 6. Multiple marker staining of *Xenopus* embryos.** Embryos were co-injected with 0.5 ng of *lacZ* mRNA and 1 ng of either N358S GUS (A, C) or GUS<sup>AR</sup> (B, D), and subsequently fixed in 3.7% formaldehyde for 20 min. Following fixation, embryos were reacted with X-gluc (light blue/green in all frames). As in Figure 5, embryos injected with mRNA encoding GUS<sup>AR</sup> (B) stained much more intensely than those injected with N358S GUS mRNA (A). All embryos were subsequently rinsed free of X-gluc and stained with the chromogenic substrate for  $\beta$ -galactosidase, rose-gal (red in C and D).

**GUS<sup>AR</sup> as a universal reporter gene.** Since most naturally occurring  $\beta$ -glucuronidases are likely to be fixative-labile, the fixative-resistant GUS<sup>AR</sup> we have isolated should prove useful for expanding the range and power of GUS staining techniques. In addition, it should be possible to develop methods for following multiple genes or cell lineages in parallel. Such methods generally rely on protocols in which fixed tissues are reacted with antibodies conjugated to dyes or reporter enzymes (for example, see ref. 24).

Reporter genes are very commonly used in *Xenopus* as cell lineage tracers, and have proved important for gene expression studies in developing embryos<sup>25,26</sup>. Following mRNA microinjection, the fixative-resistant GUS could be specifically followed in *Xenopus* relative to both background activity and the wild-type reporter. Moreover, a lineage trace in tandem with  $\beta$ -galactosidase demonstrated the use of GUS in a multiple-reporter format. These experiments pave the way for the practical development of two-enzyme reporter systems, and could potentially be combined with a  $\beta$ -lactamase reporter system developed by Raz *et al.*<sup>27</sup> to create three-enzyme reporter systems.

The results in *Xenopus* embryos are notable in that that no special precautions were taken to enhance gene expression or enzymatic activity. In contrast to the transformation of reporter constructs, microinjected reporter mRNAs do not replicate and their dosage progressively decreases as messages are segregated or broken down. Further, no attempt was made to increase the signal intensity of GUS<sup>AR</sup> by fusing it to a nuclear localization signal, as was the case for the *lacZ* control. Nor were fluorescent or other highly sensitive commercially available GUS substrates<sup>3</sup> utilized. In short, the evolved enzyme is itself robust enough so that new staining techniques can easily be adapted from extant methods.

### Experimental protocol

**Materials.** DNA-modifying enzymes, including restriction enzymes and Vent polymerase, were purchased from New England Biolabs (Beverly, MA). Deoxyribonuclease I was from GIBCO-BRL (Gaithersburg, MD).

Taq polymerase was expressed and purified as described by Grimm and Arbutnot<sup>28</sup>. DNA sequencing kits were from Perkin-Elmer/Applied Biosystems (Foster City, CA). Cloning vector pGEM-5 was from Promega (Madison, WI), pBluescript II SK(+) was from Stratagene (La Jolla, CA), and the regulatory vector pREP4 was from Qiagen (Chatsworth, CA). pGUS N358S was from Clontech (Palo Alto, CA), and pET28a(+) from Novagen (Madison, WI). DNA purification columns were purchased from Qiagen (Chatsworth, CA). X-gluc was from Gold Biotechnology (St. Louis, MO) and Butterfly nitrocellulose membranes from Schleicher and Schuell (Keene, NH). The mMessage mMachine SP6 in vitro mRNA transcription kit was from Ambion (Austin, TX). MicroSpin G-25 Sephadex spin columns were from Pharmacia Biotech (Piscataway, NJ). *Escherichia coli* strain Inv $\alpha$ F' was from Invitrogen (Carlsbad, CA), W3110 (ATCC No. 27325) from the ATCC (Rockville, MD), GMS407 from the *E. coli* Genetic Stock Center (New Haven, CT), and BL21(DE3)pLysS from Novagen. Other chemicals, including glutaraldehyde, PNPG and NBT, were from Sigma Chemicals (St. Louis, MO).

**Cloning of *gusA*.** The *E. coli gusA* gene was amplified from W3110 cells using Vent polymerase and the primers 5'-CCGGATCCTCTAGAGATGT-TACGTCCTGTGAAAACC-3' and 5'-GCCGAATTCGACGTCATTGTTTGCCTCCCTGCT-3' (*Xba*I and *Eco*RI sites underlined). The PCR product was blunt-end ligated into the *Eco*RV site of pGEM-5 by standard methods<sup>29</sup>. *Escherichia coli* Inv $\alpha$ F' cells were transformed by the method described by Inoue *et al.*<sup>30</sup>. The *gusA* gene was subcloned into pBluescript II SK(+) using restriction endonucleases *Xba*I and *Eco*RI. The nucleotides encoding the *lacZ*  $\alpha$ -fragment that would normally have been located between the ribosome binding site and the *gusA* start codon were deleted by amplifying the remainder of the plasmid using primers 5'-CCGGATCCTC-TAGAGATGTTACGTCCTGTGAAAACC-3' and 5'-CGTCTAGAAGCT-GTTTCTCTGTGTGAAATTG-3', digesting with *Xba*I, ligating, and transforming pREP4/Inv $\alpha$ F' (see below). The resultant construct placed the *gusA* gene under direct control of the lac promoter. The GUS expression vector was named *gusA*-pBSA.

**Library construction and screening.** For the first round of screening, random mutations were introduced into the cDNA by mutagenic PCR<sup>15</sup> using primers 5'-CCAGTCACGACGTTGTAAAA CGACG-3' and 5'-ATGCTTC-CGGCTCGTATGTTGTGTGG-3', which anneal to the pBluescript II SK(+) vector outside of the boundaries of the *gusA* insert. The amplification reaction was carried out with 100 nM primers, 60 mM Tris-HCl pH 8.5, 15 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 3.2 mM MgCl<sub>2</sub>, 0.125 mM MnCl<sub>2</sub>, 0.2 mM dGTP, 0.2 mM dATP, 0.4 mM dTTP, 0.4 mM dCTP, for 35 cycles of 94°C × 30 s, 72°C × 2 min. The *gusA*-pBSA plasmid library was transformed into *E. coli* Inv $\alpha$ F' cells harboring the *lacI* expression vector pREP4. The plasmid was unstable when propagated in *E. coli* Inv $\alpha$ F' without pREP4, probably because the lac repressor is not present at high enough levels to limit expression of the *gusA*. For colony-lift assays, *gusA*-pBSA/pREP4/Inv $\alpha$ F' colonies were propagated on liquid Luria Broth supplemented with 25  $\mu$ g/ml kanamycin and 100  $\mu$ g/ml ampicillin (LB-kan/amp) + 0.4% glucose plates for 12 h at 37°C. The colonies adsorbed to a nitrocellulose filter and were transferred colony side up to LB-kan/amp plates containing 0.5 mM isopropyl  $\beta$ -D-thiogalactopyranoside (IPTG), and induced at 37°C for 12–24 h. The nitrocellulose-bound colonies were transferred to GUS buffer (50 mM sodium phosphate pH 7.0, 0.1% Triton X-100, 1 mM EDTA) containing 0.2% glutaraldehyde and incubated for 20 min at 23°C. The filters were then transferred to buffer containing 165  $\mu$ g/ml X-gluc and 330  $\mu$ g/ml NBT and incubated for 10–30 min. The filters were incubated briefly in 4 M guanidine hydrochloride to arrest color development. Those colonies on the master plate corresponding to the darkest colony remnants on the filter were isolated and amplified.

For subsequent rounds of screening, the alleles were randomly recombined and mutated by DNA shuffling as described by Stemmer<sup>17</sup>. In short, the *gusA* variants were PCR amplified using the same primers as in the

mutagenic PCR reactions already described, partially digested with DNase I, and reassembled in a PCR reaction without primers. The products were amplified in a PCR with primers, then subcloned back into *gusA*-pBSA for screening. The second and third rounds were carried out in the same way, except that 1.0% and 3.5% glutaraldehyde was used to fix the colonies before the incubation in X-gluc and NBT. The most resistant round 3 clone was isolated and sequenced at the University of Texas, Institute of Cellular and Molecular Biology Core Facility using the Applied Biosystems protocol, via the primers originally used for mutagenic PCR and two additional internal primers: 5'-CGCCGGGAATGGTGATTACC-3' and 5'-CTGATGGTATCGGTGTGAGCG-3'.

**In vitro characterization of enzyme activity.** For the preparation of lysates, *gusA*-pBSA/pREP4/GMS407 cells were propagated at 37°C in LB-kan/amp. The *gusA* gene was induced by the addition of 0.5 mM IPTG to mid-log (OD<sub>600</sub> = 0.3) cultures, and the induced cultures were grown overnight. Cells were centrifuged, resuspended in distilled water, centrifuged again, and resuspended in GUS buffer. Cells were lysed with the addition of 10 mM EDTA and 1 mg/ml chicken lysozyme. The insoluble fraction was centrifuged down, and the aldehyde resistance of the GUS in the supernatant was determined as follows. Glutaraldehyde or formaldehyde was added to an aliquot of supernatant and the mixture was incubated at 23°C for exactly 20 min. The mixture was then diluted 1/100 into GUS buffer containing 0.5 mM PNPG. The hydrolysis of the substrate was followed for 1 min at 23°C at 405 nm in a Shimadzu UV-1601 spectrophotometer. The absorption extinction coefficient of p-nitrophenol under these conditions was 11.50 mM<sup>-1</sup> cm<sup>-1</sup>. The initial rates of hydrolysis were linear (data not shown).

To generate purified GUS enzymes, the wild-type and evolved *gusA* genes were amplified by PCR with the primers: 5'-GCTCTAGAGCATATGT-TACGTCCTGTAGAAACC-3' and 5'-GCGAATTCGACGTCATTGTTTGC-CTCCCTGCT-3' and subcloned into the expression vector pET28a(+) using the restriction enzymes *NdeI* and *EcoRI* (sites underlined in primers). The resultant genes were sequenced as described already to confirm that no additional mutations had been introduced during amplification or cloning. The expression constructs were transformed into BL21(DE3)/pLysS. The transformed strains were propagated and induced, and the proteins purified by nickel chelate chromatography, as suggested by Novagen (Madison, WI). The protein preparations were judged to be >99% pure following SDS-PAGE and Coomassie Blue staining (data not shown). Purified protein concentrations were determined via Bradford protein assays (Bio-Rad, Hercules, CA).

A 10 pmol quantity of purified GUS protein was preincubated for 20 min at 23°C in 10 µl of GUS buffer (1 µM) containing 0% or 0.04% formaldehyde. Then, 5 µl of protein solution were added to 1 ml of buffer (5 nM) containing varying concentrations of PNPG, and the initial velocity of each reaction was determined as already described. The kinetic parameters of the wild-type and mutant enzymes were calculated by fitting the initial velocity values to the Michaelis-Menten equation using the application Kaleidagraph 3.0.5 (Adelbeck Software, Reading, PA).

**Expression of *gusA* in *Xenopus* embryos.** The GUS<sup>AR</sup> gene was subcloned into pGUS N358S; this placed the gene downstream of a Kozak sequence<sup>20</sup>, so that its transcript could be recognized by eukaryotic translation systems. N358S GUS and GUS<sup>AR</sup> were subcloned into the *Xenopus* expression vector p64TS. This plasmid provides in vitro-transcribed mRNAs with *Xenopus* globin 5'- and 3'-untranslated regions and greatly increases the amount of protein translated from the mRNA<sup>31</sup>. Capped mRNA was produced by *in vitro* transcription<sup>32</sup> of the clones already described using the Ambion mMessage mMaker SP6 protocol. In vitro transcriptions were also treated with DNase I, and the mRNA was purified using a Sephadex G-25 spin column to minimize nonspecific toxicity effects. Purified mRNAs were resuspended in sterile water for injections.

Female adult *Xenopus* were induced to ovulate with human chorionic gonadotropin, and eggs were fertilized in vitro. Embryos were dejellied in 3% cysteine solution and washed in 0.2× MMR<sup>33</sup>. Embryos were then reared at 13–18°C in 0.2× MMR. Microinjections were performed as described<sup>34</sup>. Embryos were fixed in MEMFA (0.1 M MOPS, pH 7.4 / 2 mM EGTA/1 mM MgSO<sub>4</sub> / 3.7% formaldehyde) for 20 min, and embryos were washed 5 × 5 min in 1× PBS. GUS activity was detected using 1 mg/ml X-gluc in a solution of 1× PBS / 20 mM potassium ferricyanide / 20 mM potassium ferrocyanide / 2 mM MgCl<sub>2</sub> / 0.02% NP-40 at 37°C for 2 h. β-Galactosidase was detected using the same buffer but substituting rose-gal for X-gluc. Injection experiments repeated on different days with different preparations of mRNA gave similar results (data not shown).

## Acknowledgments

We thank the Office of Naval Research for funding. I.M. was supported by a National Science Foundation/Alfred P. Sloan Postdoctoral Research Fellowship in Molecular Evolution (DBI-9750002). We thank Dr. Mary Berlyn of the E. coli Genetic Stock Center for sending us strain GMS407 and Ms. Sabine Bell for synthesizing oligonucleotides. Finally, we thank members of the Ellington group and Dr. William Wu of Ambion for helpful discussion.

- Jefferson, R.A., Burgess, S.M. & Hirsh, D. Beta-glucuronidase from *Escherichia coli* as a gene-fusion marker. *Proc. Natl. Acad. Sci. USA* **83**, 8447–8451 (1986).
- Martin, T., Woehner, R.-V., Hummel, S., Willmitzer, L. & Frommer, W.B. in *GUS Protocols: using the GUS gene as a reporter of gene expression*. (ed. Gallagher, S.R.), 23–43 Academic Press, New York; 1992).
- Naleway, J.J. in *GUS Protocols: using the GUS gene as a reporter of gene expression*. (ed. Gallagher, S.R.), 61–76 (Academic Press, New York; 1992).
- Farrell, L.B. & Beachy, R.N. Manipulation of beta-glucuronidase for use as a reporter in vacuolar targeting studies. *Plant Mol. Biol.* **15**, 821–825 (1990).
- Craig, S., in *GUS Protocols: using the GUS gene as a reporter of gene expression* (ed. Gallagher, S.R.), 115–124 (Academic Press, New York; 1992).
- Jefferson, R.A. The GUS reporter gene system. *Nature* **342**, 837–838 (1989).
- Habeeb, A.F.S.A. & Hiramoto, R. Reactions of proteins with glutaraldehyde. *Arch. Biochem. Biophys.* **126**:16–26 (1968).
- Schlamann, H.R., Risseuw, E., Franke-van Dijk, M.E. & Hooykaas, P.J. Nucleotide sequence corrections of the *uidA* open reading frame encoding beta-glucuronidase. *Gene* **138**, 259–260 (1994).
- Oshima, A. *et al.* Cloning, sequencing, and expression of cDNA for human beta-glucuronidase. *Proc. Natl. Acad. Sci. USA* **84**, 685–689 (1987).
- D'Amore, M.A., Gallagher, P.M., Korfhagen, T.R. & Ganschow, R.E. The complete sequence and organization of the murine beta-glucuronidase gene. *Biochemistry* **27**, 7131–7140 (1988).
- Nishimura, Y. *et al.* Nucleotide sequence of rat preputial gland beta-glucuronidase cDNA and *in vitro* insertion of its encoded polypeptide into microsomal membranes. *Proc. Natl. Acad. Sci. USA* **83**, 7292–7296 (1986).
- Ray, J. *et al.* Cloning of canine beta-glucuronidase cDNA, mutation identification in canine MPS VII, and retroviral vector-mediated correction of MPS VII cells. *Genomics* **48**, 248–253 (1988).
- Zhang, J.H., Dawes, G. & Stemmer, W.P. Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening. *Proc Natl Acad Sci USA* **94**, 4504–4509 (1997).
- Giver, L., Gershenson, A., Freskgard, P.O. & Arnold, F.H. Directed evolution of a thermostable esterase. *Proc. Natl. Acad. Sci. USA* **95**, 12809–12813 (1998).
- Cadwell, R.C. & Joyce, G.F. Randomization of genes by PCR mutagenesis. *PCR Methods Appl.* **2**, 28–33 (1992).
- Blanco, C. & Nemoz, G. One step purification of *Escherichia coli* beta-glucuronidase. *Biochimie* **69**, 157–161 (1987).
- Stemmer, W.P.C. DNA shuffling by random fragmentation and reassembly: *in vitro* recombination for molecular evolution. *Proc. Natl. Acad. Sci. USA* **91**:10747–10751 (1994).
- Jain, S., Drendel, W.B., Chen, Z.W., Mathews, F.S., Sly, W.S. & Grubb, J.H. Structure of human beta-glucuronidase reveals candidate lysosomal targeting and active-site motifs. *Nat. Struct. Biol.* **3**, 375–381 (1996).
- Needleman, S.B. & Wunsch, C.D. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* **48**, 443–453 (1970).
- Kozak, M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* **15**, 8125–8148 (1987).
- Moore, J.C. & Arnold, F.H. Directed evolution of a para-nitrobenzyl esterase for aqueous-organic solvents. *Nat. Biotechnol.* **14**, 458–467 (1996).
- Zhang, X.-J., Baase, W.A., Shoichet, B.K., Wilson, K.P. & Matthews, B.W. Enhancement of protein stability by the combination of point mutations in T4 lysozyme is additive. *Protein Eng.* **8**, 1017–1022 (1995).
- Holm, L. & Sander, C. Mapping the protein universe. *Science* **273**, 595–603 (1996).
- van der Loos, C.M., Becker, A.E. & van den Oord, J.J. Practical suggestions for successful immunoenzyme double-staining experiments. *Histochem. J.* **25**, 1–13 (1993).
- Wallingford, J., Seufert, D., Virta, V. & Vize, P. p53 Activity is essential for normal development in *Xenopus laevis*. *Curr. Biol.* **7**, 747–757 (1997).
- Wallingford, J., Carroll, T. & Vize, P. Precocious expression of the Wilms' tumor gene *xWT1* inhibits embryonic kidney development in *Xenopus laevis*. *Dev. Biol.* **202**, 103–112 (1998).
- Raz, E., Zlokarnik, G., Tsiens, R.Y. & Driever, W. Beta-lactamase as a marker for gene expression in live zebrafish embryos. *Dev. Biol.* **203**, 290–294 (1998).
- Grimm, E. & Arbutnot, P. Rapid purification of recombinant Taq DNA polymerase by freezing and high temperature thawing of bacterial expression cultures. *Nucleic Acids Res.* **23**, 4518–4519 (1995).
- King, P.V. & Blakesley, R.W. Optimizing DNA ligations for transformations. *FOCUS* **8**, 30–32 (1986).
- Inoue, H., Nojima, H. & Okayama, H. High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96**, 23–28 (1990).
- Krieg, P.A. & Melton, D.A. Functional messenger RNAs are produced by SP6 *in vitro* transcription of cloned cDNAs. *Nucleic Acids Res.* **12**, 7057–7070 (1984).
- Krieg, P. & Johnson, A. in *A laboratory guide to RNA: isolation, analysis, and synthesis*. (ed. Krieg, P.) 141–153 (Wiley-Liss, New York; 1996).
- Peng, H.B. Solutions and protocols. *Methods Cell Biology* **36**, 657–662 (1991).
- Vize, P., Melton, D., Hemmati-Brivanlou, A. & Harland, R. Assays for gene function in developing *Xenopus* embryos. **36**, 367–387 (1991).