SUPPLEMENTARY MATERIALS

SUPPLEMENTARY EXPERIMENTAL PROCEDURES

Cloning of truncated and mutant forms of HMGB1- A 648 bp *Xho*I-*Eco*R1 fragment from pET11d-hHMGB1 encoding human HMGB1 (1) was subcloned into pBluescript II KS+ (3.0 kb, Stratagene, La Jolla, CA), creating pBSHMGB1. Plasmid DNA was purified using a QIAprep Spin Miniprep Kit (QIAGEN, Valencia, CA). The pBSHMGB1 construct was digested with either *Xho*I and *Bam*HI or *Bam*HI and *Eco*RI to remove fragments encoding box A or B (and the acidic tail), respectively, and replaced with DNA linkers assembled from the following oligonucleotides: 5'-GATCCCTAGTAG-3' and 5'-AATTCTACTAGG-3' (inserted into *Bam*H1-*Eco*RI sites) or 5'-TCGAGTTCAAG-3' and 5'-GATCCTTGAAC-3 (inserted into *Xho*I-*Bam*HI sites), creating pBS-Box A and pBS-Box B. The *Xho*I-*Eco*R1 cassette from pET11d-HMGB1 was then replaced by those obtained by *Xho*I-*Eco*RI digestion of pBS-Box A and pBS-BoxB, yielding pET11d-BoxA and pET11d-BoxB, respectively.

To disrupt the “hydrophobic wedge” present in each HMG-box domain of HMGB1, we introduced a large, contiguous block of alanine substitutions beginning at position 18 of box A or position 102 of box B using a strategy we have termed "ligation-assisted recombination polymerase chain reaction" (LAR-PCR). In this procedure, which is generally based on a technique called recombination PCR (2), two separate PCR amplifications are performed to generate two DNA fragments that, when assembled together, reconstitute a plasmid (Supp. Fig. 1A). Primers are designed such that after PCR, the two amplicons each contain the same unique restriction site on one end and homologous ends within an antibiotic resistance gene on the other (Supp. Figure 1B). After digesting one set of DNA ends with an appropriate restriction enzyme, the ends are ligated together and the product is introduced into *E. coli*, where it subsequently undergoes homologous recombination at the other pair of DNA ends *in vivo* to yield a recombinant plasmid that confers antibiotic resistance to the transformed bacteria. A similar approach that does not rely on recombination to regenerate the antibiotic resistance gene has recently been described (3).

For PCR, pBSHMGB1 was linearized by digestion with either *Xho*I or *Eco*RI (for amplification of the "right" or "left" half of the plasmid, respectively) in the presence of calf intestinal phosphatase (all restriction and modifying enzymes were purchased from New England Biolabs, Beverly, MA, USA). Digested DNA was purified using the QIAquick PCR Purification Kit (QIAGEN, Valencia, CA) and analyzed by agarose gel electrophoresis to verify DNA recovery and integrity. Reactions assembled for PCR (100 µL) contained linearized pBSHMGB1 (~30 ng), both a mutagenic and nonmutagenic primer (500 pmol each), and recombinant *Taq* polymerase (2.5 units; Invitrogen, Carlsbad, CA) in PCR buffer (20 mM Tris-HCl [pH 8.4], 50 mM KCl, 200 µM each deoxynucleotide triphosphate (dNTP), 1.5 mM $MgCl_2$). Samples were overlaid with mineral oil and PCR amplification was performed using a PTC-100 Thermal Cycler (MJ Research, Waltham, MA). Reactions were subjected to initial denaturation (94 °C, 3 min), 25 cycles of amplification (94 °C x 45 s, 50-55 °C x 30 s, and 72 °C x 90 s) and a final extension (72 °C x 5 min). In these experiments, mutagenic primers were synthesized that contain an *Sfi*I site at the 5' end, which overlaps a stretch of alanine codons (both the open reading frame of HMGB1 and pBluescript lack an *Sfi*I site), followed by 18-21 gene-specific nucleotides. The following mutagenic primers were used to introduce alanine substitutions into box A and box B: MutBoxA-Forward (5'-GCTGCTGCGGCCGCTGCGGCCGCGGCGGCTAAGAAGAAGCACCCAGATGC-3'), MutBoxA-Reverse (5'-GCCGCCGCGGCCGCAGCGGCCGCAGCAGCTGCATATGATGACATTTTGCC-3'), MutBoxB-Forward (5'-GCCGCCGCGGCCGCGGCGGCCGCTGCCGCAAAAATCAAAGGAGAACATCCT-3), and MutBoxB-Reverse (5'-GCGGCAGCGGCCGCCGCGGCCGCGGCGGCGGCCGAAGGAGGCCTCTTGG-3'). Nonmutagenic primers specific for the ampicillin resistance gene in pBluescript II KS+ were also prepared (sense primer, 232 bp downstream of start codon: 5'-GTATTATCCCGTATTGACGCCGGG-3'; antisense primer, 129 bp upstream of stop codon: 5'-AATGATACCGCGAGACCCACGCTC-3').

A portion of each PCR reaction (10 µL) was analyzed on a 1% agarose gel to verify the presence of an amplicon of the expected size. The remaining PCR product was purified using a QIAquick PCR Purification Kit, eluting in buffer EB (10 mM Tris-HCl [pH 8.0]; 30 µL). Each PCR product (~1 µg) was digested with *Sfi*I and purified using the QIAquick PCR Purification Kit as described above. Paired *Sfi*I-digested amplicons (~75 ng each) were ligated in reactions (20 µl) containing 3 Weiss units T4 DNA ligase in ligation buffer (50 mM Tris-HCl [pH 7.5], 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP and 25 µg/mL BSA) and incubated at 15 °C for 16 h. An aliquot of the ligation reaction (2 µL) was used to transform Max Efficiency competent DH5 *E. coli* according to the manufacturer's instructions (50 µL, Invitrogen), spreading $1/8^{th}$ of the cells onto LB agar plates containing carbenicillin (50 µg/mL). For each mutation, ten colonies were grown overnight in LB broth supplemented with carbenicillin to isolate plasmid DNA for screening with *Sfi*I (designated pBSmtAHMGB1 and pBSmtBHMGB1, respectively). Using the two different sets of mutagenic primers targeting the A and B box domains of HMGB1, we found in both cases that >90% of colonies screened using this protocol possessed plasmid DNA linearized by *Sfi*I. Sequencing of several clones revealed no errors in the region being mutated or in flanking regions up to restriction sites we utilized for subcloning. To introduce mutations into both HMG boxes, the *Bam*HI-*Eco*RI fragment from pBSmtAHMGB1 was replaced by the same cassette removed from pBSmtBHMGB1, generating pBSmtABHMGB1. The *Xho*I-*Eco*RI cassette from pET11d-HMGB1 was then replaced by those obtained by *Xho*I-*Eco*RI digestion of pBSmtAHMGB1, pBSmtBHMGB1, and pBSmtABHMGB1, yielding pET11d-mtAHMGB1, pET11d-mtBHMGB1, and pET11d-mtABHMGB1, respectively.

During the course of these studies, we became aware that the T7 termination signal was absent in pET11d-hHMGB1, and thus wished to move our coding sequences back into the native pET11d vector. Moreover, to ensure proper termination of translation and to introduce a convenient restriction site, 3 stop codons and an *Xma*I site were appended to the C-terminus of the human HMGB1 coding sequence by PCR. The reaction included pET11d-hHMG1 as the template, Platinum Pfx DNA polymerase (Invitrogen), the T7promoter primer (5'-TAATACGACTCACTATAG-3'), and the RevStop primer (5'-GTGCTGAGCTCAGCCCGGGCTACTATTATTCATCATCATCATCTTC-3'). Reactions were subjected to initial denaturation (94 °C, 2 min), 30 cycles of amplification (94 °C x 15 s, 49 °C x 30 s, and 68 °C x 1 min). The resulting product was digested with *Nco*I and *Blp*I, purified by agarose gel electrophoresis, recovered using the QIAquick Gel Extraction Kit (Qiagen), inserted into *Nco*I-*Blp*I-digested pET-11d to produce pET11d-HMGB1-T7, and the resulting construct verified by DNA sequencing. The constructs pET11d-boxB-T7, pET11d-mtAHMGB1-T7, and pET11d-mtABHMGB1-T7 were created by inserting a *Pst*I fragment from pET11d-HMGB1-T7 into *Pst*I-digested, alkaline-phosphatase-treated pET11d-boxB, pET11d-mtAHMGB1, and pET11d-mtABHMGB1, respectively. Furthermore, pET11d-mtBHMGB1-T7 was created by inserting a *Bam*HI-*Xma*I fragment from pET11d-mtABHMGB1 into *Bam*HI/*Xma*I-digested pET11d-HMGB1-T7.

The construction of pET11d-basic-T7 utilized PCR with the T7promoter and BASIC (5'-GCGGCGCCCGGGCTACTATTACTTCTTTTTCTTGCTTTTTTCAGCC-3') primers (the PCR profile was the same as that described for construction of pET11d-HMGB1-T7). The PCR product was digested with *Xho*I and *Xma*I, and inserted into *Xho*I-*Xma*I-digested pET11d- HMGB1-T7. pET11d-tailless-T7 was similarly constructed using the T7promoter and 3HMGStop (5'-GTGCTGACCCGGGCTACTATTACTTTCCTTTAGCTCGATATG-3') primers. A *Pst*I fragment from pET11d-tailless-T7 was inserted into *Pst*I-digested, alkaline-phosphatase-treated pET11d-boxB-T7, pET11d-mtAHMGB1-T7, and pET11d-mtBHMGB1-T7, yielding pET11d-taillessboxB-T7, pET11d-taillessmtA-T7, and pET11d-taillessmtB-T7, respectively. To make pET11d-shuffled-T7, box A was amplified by PCR using primers 5HMGB1A (5'-AGCTAAAGGAAGGAGACAAAAAAGAAGGGCAAAGGAGATCCTAAG-3') and 3HMGB1A (5'-GTGCTGACCCGGGCTACTATTACCCTTTGGGAGGGATATAG-3'), while primers 5HMGB1B (5'-GCCATATGCTCGAGTTCAAG-3') and 3HMGB1B (5'-TTCTTTTTTGTCTCCTTTCCTTTAGCTCGATATG-3') were used for box B. The PCR profile was the same as that described for construction of pET11d-HMGB1-T7. The box A product was digested with

*Bsm*AI and *Xma*I, and domain B with *Xho*I and *Bsm*AI. The two fragments were gel purified and ligated simultaneously into *Xho*I-*Xma*I-digested pET11d-HMGB1-T7. All PCR-generated HMGB1 plasmid constructs were verified by DNA sequencing.

HMGB1 expression and purification- The *E. coli* strain BL21(DE3)pLysS was transformed with the various pET-11d HMGB1-T7 constructs. Protein expression was induced by adding isopropyl-ß-D-thiogalactopyranoside (IPTG; 1 mM final concentration) to a log-phase culture and incubating at 30°C for 4 h. Cell pellets were resuspended in binding buffer (40 mM Tris [pH 8.0], 0.5 M KCl, 0.25% (vol/vol) Tween-20, 50 mM imidazole, 1 mM phenylmethylsulphonyl fluoride (PMSF), 10 µM leupeptin, 1 µM pepstatin A) and sonicated on ice (10 cycles of 10 sec pulse, 10 sec pause at 20% power). Lysates were clarified by centrifugation at 46,000 x *g* (Beckman SW55Ti rotor, 22,000 rpm) for 40 min at 4°C, followed by filtration of the supernatant through 0.45 µm syringe filters. Each protein was initially purified by immobilized metal affinity chromatography (IMAC) using 1 ml Chelating Sepharose Fast Flow (FF) columns (Amersham BioSciences) charged with Ni^{2+}, and eluting with elution buffer (40 mM Tris [pH 8.0], 0.5 M KCl, 500 mM imidazole, 1 mM PMSF, 10 µM leupeptin, 1 µM pepstatin A). Protein samples analyzed by SDS-PAGE followed by staining with SYPRO Orange revealed the presence of two major protein bands that were also immunoreactive toward rabbit polyclonal anti-HMGB1 antibodies raised against residues 161-181. In addition, two minor high molecular weight bacterial proteins (~83 kDa) were also detected (data not shown). Therefore we pursued additional purification using ion exchange chromatography (IEC). All purification steps involving IEC were performed using Econo-Pac ion exchange cartridges (BioRAD) with a BioLogic LP chromatography system connected to a Model 2110 Fraction collector (BioRAD). All washes and gradients were prepared with buffer A (40 mM Tris [pH 8.0], 1 mM DTT). The flow rate for all loading and washing steps was 1.5 ml/min, which was reduced to 1 ml/min for the elution step.

We found that minor bacterial contaminants were separated from full-length HMGB1 by IEC using a strongly acidic cation exchange support (High S) which preferentially retains HMGB1 over the contaminating bacterial proteins. However, the various HMGB1 protein species could not be separated from one another by IEC using High S (not shown). We speculated that the negative charge associated with the acidic C-terminal tail of HMGB1 might be exploited to separate degraded forms of HMGB1 lacking part or all of the acidic tail from full-length HMGB1 by IEC using a strongly basic anion exchange support (High Q). This approach proved successful. Specifically, to purify full-length WT, mtA, mtB, and mtAB HMGB1, the eluate obtained from the Ni^{2+}-FF chromatography step was diluted to 65 mM KCl with buffer A and then loaded onto a 1 ml High S cartridge. Bound proteins were washed with 80 mM KCl (9 ml), and eluted using a linear KCl gradient of 80 mM to 0.65 M developed over 25 min. The target protein eluted between about 300 mM and 365 mM KCl (depending on the form of HMGB1). Appropriate fractions were pooled, diluted 2-fold with buffer A, and loaded onto a 1 ml High Q cartridge. Bound proteins were washed with 220 mM KCl (10.5 ml), and eluted using a linear KCl gradient of 0.22 M to 0.7 M developed over 40 min. Desired proteins eluted between about 360 mM and 425 mM KCl (depending on the protein). Appropriate fractions were pooled, dialyzed against dialysis buffer (25 mM Tris-HCl [pH 8.0], 150 mM KCl, 2 mM DTT, 10% glycerol), snap frozen in liquid nitrogen, and stored at –80 °C until use.

A slightly different strategy was required to purify forms of HMGB1 lacking the acidic tail (including the individual HMG-box domains), as these proteins are not retained on High Q. To prepare box A, basic, tailless, tailless mtA, tailless mtB, shuffled, and box B', the Ni^{2+}-FF eluate was diluted to 125 mM KCl with buffer A and subjected to negative purification through a 1 ml High Q cartridge to remove high molecular weight contaminants. With the exception of basic HMGB1, the flow-through was then diluted 2-fold with buffer A, loaded onto a 1 ml High S cartridge, washed with 80 mM KCl, and eluted with a KCl gradient of 80 mM to 0.65 M. Box B' eluted around 150 mM, whereas the other proteins eluted between 300 mM and 365 mM. For basic HMGB1, IEC using High S failed to provide sufficient separation from degradation products, but a weakly acidic cation exchange support (CM) proved adequate for this purpose. Therefore, the flow-through was diluted 4-fold and loaded onto a 1 ml CM cartridge and

washed with 30 mM KCl. The protein was eluted with a KCl gradient of 30 mM to 0.75 M, with the desired protein eluting around 425 mM KCl. Appropriate fractions were pooled and dialyzed as above.

To purify box B, the eluate obtained from the Ni^{2+}-FF chromatography step was passed over an uncharged FF column to remove residual Ni^{2+} leached from the first purification step. The flow-through was diluted 4-fold with buffer A and loaded onto a 1 ml High Q cartridge. Bound proteins were washed with 220 mM KCl, and eluted with a 0.22 M to 0.7 M KCl gradient as described above. The box B protein eluted at about 425 mM KCl. Desired fractions were pooled and dialyzed as described above.

SUPPLEMENTARY DISCUSSION

For introducing 1-3 point mutations (affecting ≤ 3 codons), standard R-PCR works well. However, if R-PCR were used to generate mutations in 10 consecutive codons, the length of the mutagenic primer would be at least 49 nucleotides (30 mutagenic nt, preceded by 3 nonmutagenic nt, and followed by a 3' stretch of 16 nonmutagenic nt that anneals to the PCR template, as recommended (4)). Such long primers generally require synthesis on at least a 50 nmol scale, and HPLC or PAGE purification is generally suggested, both of which increase the overall cost of the procedure. In LAR-PCR, mutations in up to 10 codons can be introduced using mutagenic primers less than 45 nt in length, which are typically prepared on a 25-35 nmol scale and desalted. In this case, the extended recognition site of *Sfi*I required synthesis of longer mutagenic primers than might otherwise be necessary. In principle, an approach analogous to enzymatic inverse PCR (EIPCR) could be used to amplify an entire plasmid sequence using only the mutagenic primers (5). After PCR, the linear DNA is digested with a selected restriction enzyme and the ends ligated together. Positive clones are then identified based on their ability to be linearized by the chosen restriction enzyme. Although an EIPCR-like approach was successfully used previously for alanine scanning mutagenesis of *RAG-1* (6), it has several drawbacks compared to LAR-PCR, including (i) higher frequency of nonmutated clones due to contaminating plasmid template (40-60%), even when digested, linear DNA is gel isolated before ligation; (ii) longer PCR cycling times; (iii) lower yields of full-length amplicon (especially for large plasmids); and (iv) greater abundance of undesired PCR products.

Many restriction enzymes yielding cohesive ends are suitable for alanine scanning mutagenesis using LAR-PCR, including *Not*I, *Sfi*I, *Pst*I, *Bgl*I, and *Eag*I (Supp. Fig. 1B). In this study, *Sfi*I was used because its recognition site, containing a rare 8 bp sequence, was absent in the HMGB1 ORF and because pBSHMG1 retained *Not*I and *Pst*I sites in the polylinker of pBluescript. While LAR-PCR was developed primarily for alanine scanning mutagenesis, LAR-PCR in principle could be applied to any site-directed mutagenesis effort where large blocks of amino acids are substituted with other residues. The only limitation of this strategy is the availability of restriction enzymes whose recognition sites contain codons appropriate for the residues being introduced.

SUPPLEMENTARY REFERENCES

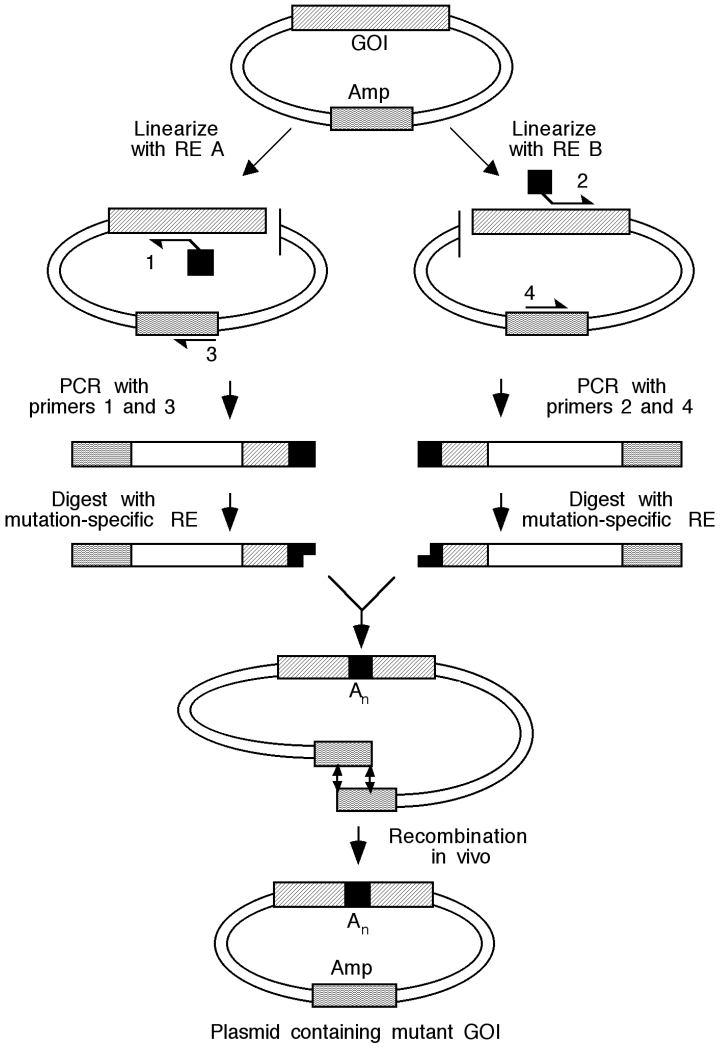
1. Ge, H., and Roeder, R. G. (1994) *J Biol Chem* **269,** 17136-17140.
2. Jones, D. H., and Winistorfer, S. C. (1992) *Biotechniques* **12,** 528-530, 532, 534-525
3. Allemandou, F., Nussberger, J., Brunner, H. R., and Brakch, N. (2003) *J Biomed Biotechnol* **2003,** 202-207.
4. Howorka, S., and Bayley, H. (2002) *Methods Mol Biol* **182,** 139-147.
5. Stemmer, W. P., and Morris, S. K. (1992) *Biotechniques* **13,** 214-220.
6. Swanson, P. C., and Desiderio, S. (1998) *Immunity* **9,** 115-125

SUPPLEMENTARY FIGURE LEGENDS

Supp. Fig. 1. (A) LAR-PCR strategy. Restriction enzymes (RE) A and B cut on either side of the gene of interest (GOI). Hemiarrows indicate primers; shaded blocks in primers 1 and 2 contain mutant sequences and mutation-specific restriction site. Double arrows indicate region of homology in ampicillin resistance gene (Amp) that undergoes recombination *in vivo*. (B) Primer design and generation of mutant plasmid constructs. (B) Representative examples of primers containing an *Sfi* I site (1; this work) or a *Pst* I site (2) (6) aligned to a 30 nt sequence in a GOI targeted for mutagenesis to create consecutive alanine codons. Flanking nucleotides (N) are added to facilitate RE cleavage close to the DNA end. Additional REs whose recognition sites are compatible for use in alanine scanning mutagenesis are also shown.

Bergeron et al. Supp. Figure 1.

A.



B.

Reading frame of gene of interest

5'-XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX-3'

1. NNN NNN NNG GCC GCX GCG GCC GCX GCX GCX-18-21nt (gene-specific)

2. NNN GCT GCA GCX GCX GCX GCX-18-21nt (gene-specific)

G GCC NNN NNG GCC	*Sfi*I
GCG GCC GC	*Not*I
CT GCA G	*Pst*I
GCC NNN NNG GC	*Bgl*I
CG GCC G	*Eag*I