Production, Purification, and Cleavage of Tandem Repeats of Recombinant Peptides

Athan Kuliopulos and Christopher T. Walsh*

Contribution from the Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115

*Received December 13, 1993**

Abstract: We describe a method to construct tandem repeats of coding sequences for polypeptides interspersed by single methionine residues and terminating in a His_6 tag that can be purified by Ni chelate chromatography and then cleaved by cyanogen bromide (CNBr) into homogeneous peptide units. Annealing and unidirectional ligation of complementary 42mer oligonucleotides encoding the 13 amino acid residue yeast α-mating factor (αF) followed by ligation of the oligomerized 42mers into a pET vector placed the α-factor tandem repeats downstream of the ketosteroid isomerase (KSI) gene and upstream of a His_6 cassette. A KSI-$(\alpha F)_5$-His_6 fusion was overproduced in *Escherichia coli*, purified by Ni chelate chromatography, and then cleaved with CNBr to release insoluble KSI, the His_6 tail, and the α-factor peptide units, each terminating with homoserine (HS) lactone. HPLC yielded pure peptide in a yield of 56 mg/L. The α-factor-HS(lactone) could be ammonolyzed or hydrolyzed to yield α-factor-HS-amide or α-factor-HS, respectively. The α-factor-HS peptide had similar biological potency as authentic α-factor in yeast cell arrest assays. The α-factor-HS(lactone) was also reacted with a number of other compounds including analogs of fluorescein, dansyl, and biotin to produce α-factor peptides derivatized exclusively at the C-terminus. To test the ability of the expression system to produce longer peptides, 60–67 amino acid residue peptides encoding the Gla domain of profactor IX (FIXQS, FIXQS-His_6) were also produced. Yields of 50–55 mg/L of pure FIXQS-His_6 and FIXQS-HS(lactone) were obtained.

Introduction

The expression of the synthetic somatostatin gene as a fusion protein in *Escherichia coli*[1] marked the advent of the production of a large number of recombinant heterologous proteins in bacteria.[2] The most stable and hence abundantly produced proteins tended to be bacterial in origin and protease resistant. Attempts to produce small polypeptides of less than 100 amino acids has proved difficult unless these peptides are part of a fusion protein. For example, the carrier portion of the fusion construct has ranged in size from the first 7 amino acid residues to 997 amino acid residues of β-galactosidase,[3,4] with both extremes resulting in stable gene expression of the fusion proteins. Aside from providing enhanced stability in *E. coli*, the appropriate carrier protein, such as the glutathione-*S*-transferase or maltose-binding proteins, could be used as an affinity tag to purify the fusion protein. Since many recombinant proteins are nonfunctional when tethered to a carrier protein, several cleavage methods have been developed in order to cleave the desired protein from the carrier protein.[5] Site-specific proteases such as thrombin or factor X_a which recognize short peptide sequences are widely used but often do not cleave quantitatively at the anticipated site. Chemical cleavage methods, such as CNBr, which cleaves quantitatively and specifically at methionine residues, require strongly acidic conditions, which are incompatible with the survival of native protein structure.[6]

One particular case where the requirement for preservation of native structure during workup is not essential is in the production of small polypeptides of less than 100 amino acid residues. Early work by Shen[7] demonstrated that the 80 amino acid residue insulin monomer could be stably produced in *E. coli* as tandem repeated units which could then be subsequently cleaved into monomers at junctional methionines using CNBr. The insulin monomer was unstable by itself ($t_{1/2}$ <30 s) but the larger tandem insulin polypeptides were produced as insoluble inclusion bodies which were resistant to degradation. This method of production of tandem repeated polypeptides was also used for the 28 amino acid residue human atrial natriuretic peptide[3] and the 10 amino acid residue substance P peptide.[4] Final yields were low in these early applications of the tandem gene method due in part to the lack of protease deficient strains and tightly controlled expression systems, but especially due to a lack of a general method of affinity purification under denaturing conditions. We have adapted the tandem gene method to produce large amounts of pure 14 amino acid residue yeast α-mating factor peptide and demonstrate the general utility of the method to produce and purify by Ni chelate chromatography[8,9] any small polypeptide lacking internal methionine residues. This method should prove to be particularly useful in those cases where traditional solid-phase peptide synthesis and purification fails such as in the production of larger peptides.

Experimental Section

Bacterial Strains and Restriction Endonucleases. *E. coli* strain DH5αF'IQ was from our laboratory; the BL21(DE3)pLysS strain was from Novagen. All restriction endonucleases were from New England Biolabs.

Construction of Steroid Isomerase-α-Factor$_5$-His_6 Genes. The procedure used to create the tandem repeat of the 14 codons of α-factor fused to the C-terminus of the 125 codons for the bacterial ketosteroid isomerase (KSI) gene[10] is shown in Figure 1A. The sense and antisense 42mer deoxyoligonucleotide strands encoding the yeast α-factor peptide were designed using the preferred codon usage of *E. coli*.[11] Each strand included a three-base overhang for unidirectional end-to-end self-ligation of the duplex 42mers. Each 42mer (200 μg) was purified on a 12.5% preparative denaturing polyacrylamide gel and desalted by passage

* Abstract published in *Advance ACS Abstracts*, May 1, 1994.

(1) Itakura, K.; Hirose, T.; Crea, R.; Riggs, A. D.; Heyneker, H. L.; Bolivar, F.; Boyer, H. W. *Science* **1977**, *198*, 1056–1063.
(2) Nilsson, B.; Forsberg, G.; Moks, T.; Hartmanis, M.; Uhlen, M. *Curr. Opin. Struct. Biol.* **1992**, *2*, 569–575.
(3) Lennick, M.; Haynes, J. R.; Shen, S.-H. *Gene* **1987**, *61*, 103–112.
(4) Kempe, T.; Kent, S. B. H.; Chow, F.; Peterson, S. M.; Sundquist, W. I.; L'Italien, J. J. *Gene* **1985**, *39*, 239–245.
(5) Uhlen, M.; Moks, T. *Methods Enzymol.* **1990**, *185*, 129–143.
(6) Fontana, A.; Gross, E. In *Practical Protein Chemistry*; Darbre, A., Ed.; J. Wiley & Sons: New York, 1986; pp 67–120.
(7) Shen, S.-H. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 4627–4631.
(8) Hochuli, E.; Dobeli, H.; Schacher, A. *J. Chromatogr.* **1987**, *411*, 177–184.
(9) Smith, M. C.; Furman, T. C.; Ingolia, T. D.; Pidgeon, C. *J. Biol. Chem.* **1988**, *263*, 7211–7215.
(10) Kuliopulos, A.; Shortle, D.; Talalay, P. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 8893–8897.
(11) Boer, H. A. d.; Kastelein, R. A. In *Maximizing Gene Expression*; Reznikoff, W., Gold, L., Eds.; Butterworths: Stoneham, MA, 1986; pp 225–285.

0002-7863/94/1516-4599$04.50/0 © 1994 American Chemical Society

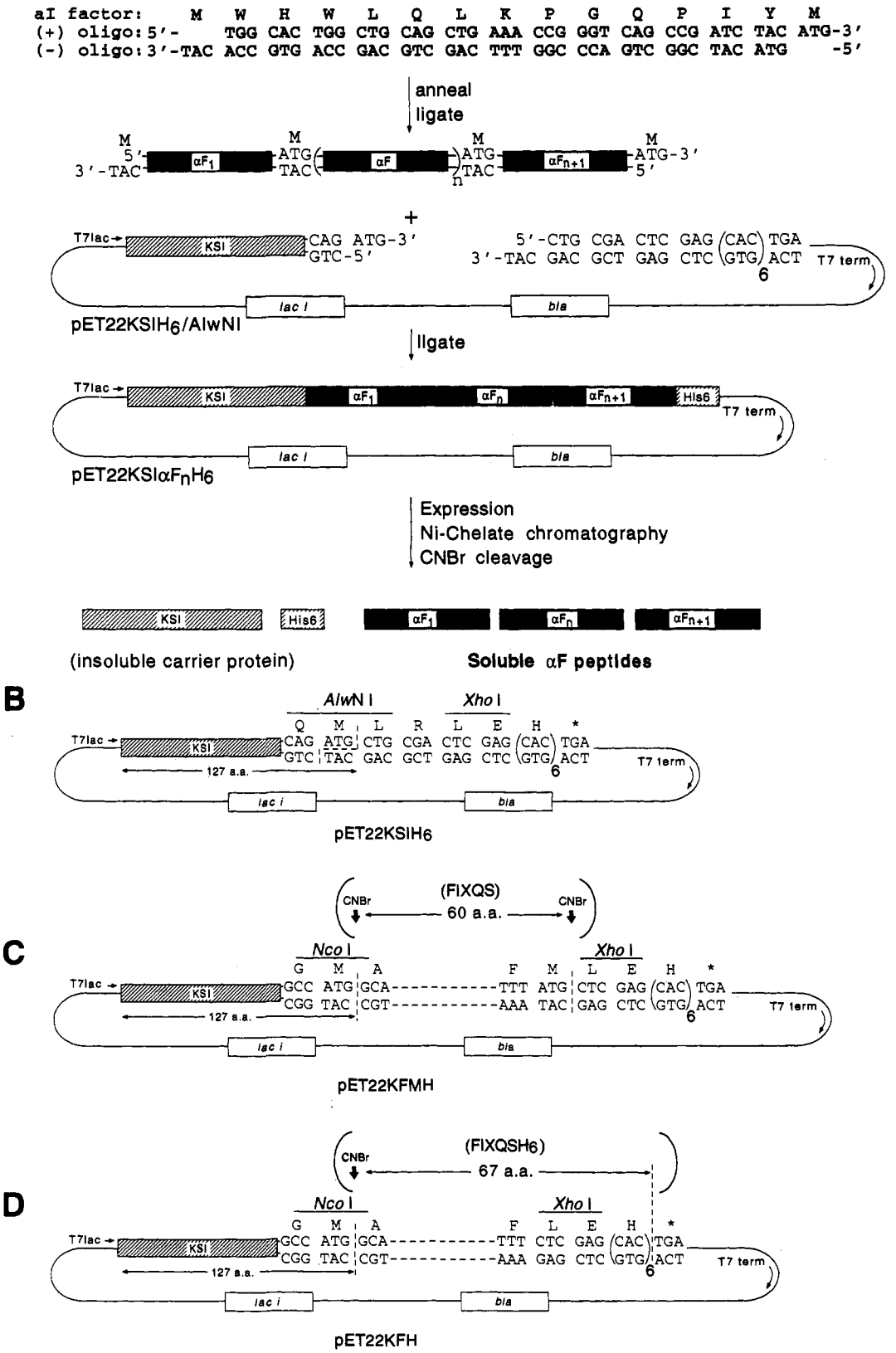


Figure 1. Strategy for the construction, production, Ni chelate purification, and CNBr cleavage of ketosteroid isomerase-tandem peptide-His_6 fusion proteins. (A) oligonucleotides encoding both strands of the 13 amino acid residue α-factor peptide were annealed, phosphorylated, and unidirectionally self-ligated to form an array of tandem repeats of α-factor-encoding units (αF_n) separated by ATG codons. These multimers were then ligated into the *Alw*NI site of pET22KSIH_6 (B) to create ketosteroid isomerase (KSI)αF_nH_6 fusion genes where n = 0–14. The genes were placed under the tight control of the T7*lac* promoter and expressed in *E. coli*. After expression levels were tested by SDS-PAGE, the highest yielding fusion proteins were purified by Ni chelate chromatography and cleaved by CNBr into individual soluble peptide units, insoluble KSI, and the His_6 tail. (B) The pET22KSIH_6 parent vector contains a unique *Alw*NI site suitable for unidirectional ligation of duplex DNA multimers terminating in ATG-3′ and 3′-TAC single-stranded DNA overhangs. (C) The pET22KFMH vector contains the KSI-FIXQSMH_6 fusion gene and was used to produce the FIXQS-homoserine-(lactone) 60 amino acid residue peptide. A methionine residue was placed upstream of the His_6 cassette to allow CNBr cleavage of the His_6 away from the 60 amino acid residue peptide after purification by Ni chelate chromatography. (D) The pET22KFH vector contains the KSIFIXQSH_6 fusion gene and was used to produce the FIXQSH_6 67 amino acid residue peptide. The C-terminal His_6 was designed to remain attached to the FIXQS peptide after CNBr cleavage.

through a C_{18}-SepPak (Millipore). Each complementary strand (25 μg) was annealed and phosphorylated by incubation of the two strands for 1 h at 37 °C in the presence of 125 units of T4 polynucleotide kinase (New England Biolabs), 1 mM ATP, in a total volume of 100 μL. After extraction with phenol and precipitation with ethanol, the 5′-phosphorylated duplex DNA was intermolecularly ligated at 8 °C for 16 h using 15 units of T4 DNA ligase (Boehringer) in a total volume of 22 μL. An additional 10 units of T4 DNA ligase and 1 mM ATP was added after 16 h and the reaction was allowed to proceed for another 24 h at 8 °C. The tandem DNA multimers were separated on a 3% low melting-point agarose gel, and purified individual bands were used for the construction of the array of tandem-α-factor genes in pET22KSIHis$_6$.

The KSI gene from pAK808KSIY14F was inserted into the *Hin*dIII/*Bam*HI sites of M13mp18, which served as the template by the mutagenesis procedure of Kunkel using the primer 5′-CTAGAAG-GAGATATACATATGAATACCCCAGAACACATGACC-3′ as previously described.[12] This introduced new *Nde*I and *Nsi*I sites, destroyed an old *Nde*I site, and created an additional mutation (N2H) at the 5′ end of the KSI gene. This new KSI gene (N2H/Y14F) had an additional *Xho*I site placed at the 3′ end of the gene by PCR mutagenesis using the (−) PCR primer 5′-GGAATTCCTCGAGCATAAATTCACTTGTC-TTTTC-3′. This KSI gene was inserted into the *Nde*I/*Xho*I sites of pET22b (Novagen), which was further digested with *Alw*NI, and the ends were made blunt by incubation with Klenow (Boehringer) and the 4 dNTPs. The resulting blunt ends were ligated, thus destroying the *Alw*NI site at position 3634 to create pET22KSIΔAlwNI. The methionine at codon 7 was changed to isoleucine by PCR using the (+) PCR primer 5′-ATACATATGCATACCCCAGAACACATCACCGCCGTGG-3′, and the (−) PCR primer 5′-GTGGTGCTCGAGTCGCAGCATCTG-GCCAGCGTGAAT-3′ was used to create an *Alw*NI site at the 3′ end of the KSI gene. The PCR-amplified product was inserted into the *Nsi*I/*Xho*I sites of pET22KSIΔAlwNI to create the universal tandem gene expression vector pET22KSIH$_6$ (Figure 1B).

This expression vector was digested with *Alw*NI, dephosphorylated with calf intestinal alkaline phosphatase (New England Biolabs), and ligated at 8 °C for 16 h with each different size tandem DNA multimer in separate reactions to create an array of pET22KSI(αF)$_n$H$_6$ constructs. The ligated DNA was transformed into DH5αF′IQ and the number (n) of α-factor encoding 42mer units of the resulting clones was ascertained by agarose gel electrophoresis. The appropriate constructs were then transformed into BL21(DE3)pLysS for protein overproduction. All fusion genes were sequenced to verify the number of tandem αF units and also to rule out adventitious mutations.

Construction of Steroid Isomerase-FIXQS-His$_6$ Genes. PCR mutagenesis was used to shuttle the previously constructed KSIFIX fusion gene from pAK808FIXQS[12] into the pET22b expression vector. The (+) PCR primer used in the construction of both pET22KFMH (encodes FIXQS, Figure 1C) and pET22KFH (encodes FIXQSH$_6$, Figure 1D) was used previously to make pAK808FIXQS. The (−) PCR primers 5′-GGAATTCCTCGAGCATAAATTCAGTTGTCTTTTC-3′ and 5′-GGAATTCCTCGAGAAATTCAGTTGTCTTTTC-3′ were used to make pET22KFMH and pET22KFH, respectively. The PCR amplified products were ligated into the *Nco*I/*Eco*RI sites of pAK808KSIFIX (N2H/Y14F). The KFMH and KFH genes were subsequently ligated into the *Nde*I/*Xho*I sites of pET22b to make the final expression vectors pET22KFMH and pET22KFH. These constructs were then transformed into BL21(DE3)pLysS for protein overproduction.

Expression of KSIαF$_n$H$_6$, KSIFIXQS, and KSIFIXQSH$_6$. The fusion proteins were produced in the pET22b vector under the control of the T7*lac* promoter, which is comprised of the T7 RNA polymerase promoter and the *lac* operator in tandem.[13] The protease deficient strain BL21-(DE3)pLysS was used for all protein production. Cultures of cells (25 mL) were grown overnight in LB media containing 50 mg of ampicillin/L at 37 °C. This starter culture was diluted 40-fold into 1 L of fresh LB-Amp, and the cells were induced with 1 mM IPTG when an absorbance of 0.3–0.5 at 595 nm was attained. The cells were harvested at OD values of ca. 2.0, which was generally reached 3–6 h after induction. The cell paste (2.5 g/L of *E. coli*) could be stored at −20 °C for at least a week without any loss in yield of fusion protein. The cell pellets were resuspended in 5 mM imidazole, 20 mM Tris-HCl, pH 7.9, and 500 mM NaCl (5 mM Im-BB), sonicated, and centrifuged at 12000*g* for 10 min at 4 °C. Levels of the total cellular protein, soluble protein, and insoluble protein were assessed by densitometry of SDS-polyacrylamide gels stained with Coomassie Blue R-250.

(12) Kuliopulos, A.; Cieurzo, C. E.; Furie, B.; Furie, B. C.; Walsh, C. T. *Biochemistry* **1992**, *31*, 9436–9444.

(13) Studier, F. W. *J. Mol. Biol.* **1991**, *219*, 37–44.

His$_6$-Ni Chelation Affinity Chromatography and CNBr Cleavage. The following three protocols are based on a 2-L culture of *E. coli*. All steps were performed at room temperature unless otherwise noted.

Protocol 1: Production and Purification of Recombinant Tandem 14 Amino Acid Residue α-Factor Peptides. The insoluble cellular protein pellet contains the majority of the fusion protein as dense inclusion bodies. This wet pellet (3.8 g) was dissolved in 45 mL of 5 mM Im-BB/6 M guanidine-HCl (5 mM Im-BB/6G). This mixture was recentrifuged at 12000*g*, 10 min, 4 °C, to remove particulate matter. The supernatant was loaded onto a 50-mL Ni-imino diacetic acid column (His-Bind resin/Novagen) which had been charged with $NiSO_4$ and equilibrated with 5 mM Im-BB/6G at a flow rate of 2.5 mL/min. The column was washed with 150 mL of 5 mM Im-BB/6G and then with 200 mL of 16 mM Im-BB/6G. The fusion protein was eluted with 150 mL of 300 mM Im-BB/6G. The protein in each fraction was assessed for purity on 10–15% SDS-PAGE PHAST gels (Pharmacia). The peak fractions were combined (96 mL), dialyzed overnight at 4 °C against 2 × 10 L of H_2O in 12–14 kDa molecular weight cutoff (MWCO) dialysis bags. The majority of the protein formed a white precipitate which could be pelleted by centrifugation at 2000*g* for 10 min at 4 °C.

The dense white pellet (5 g) was resolubilized in 60 mL of 80% formic acid and transferred to a 250-mL round-bottomed flask, and 2 g of CNBr was added. Nitrogen was bubbled in, and the flask was wrapped in aluminum foil and stirred for 18–22 h. This reaction mixture was rotoevaporated to dryness in 90 min at 28 °C. The resultant transparent proteinaceous gel that formed was resuspended in 20 mL of 40% CH_3CN/60% H_2O/0.1% TFA and stirred for 1 h. This suspension was centrifuged at 12,000*g* for 10 min at 4 °C. The supernatant was further clarified by passage through a 0.22-μm filter and the purity of the recombinant α-factor-HS(lactone) analyzed by reverse-phase HPLC and UV spectroscopy. The protein concentration was determined using a ϵ_{280} value of 12 300 M^{-1} cm^{-1}.

Protocol 2: Production and Purification of Recombinant FIXQS 60 Amino Acid Residue Peptides. The purification and CNBr cleavage of the FIXQS peptide from the KSIFIXQSMH$_6$ fusion protein was the same as in Protocol 1 until the CNBr cleavage step. Here the two protocols diverged due to the differing solubility between the very hydrophilic 60 amino acid residue peptide and the hydrophobic 14 amino acid residue α-factor peptide. After the CNBr-cleaved peptide mixture was dried by rotoevaporation, the gelatinous material was resuspended in 20 mM KPO_4 and 100 mM NaCl, and the pH was adjusted to 7.4 with 1 M $NaHCO_3$. This mixture was stirred overnight under N_2 and wrapped in aluminum foil. The suspension was centrifuged at 5000*g* for 15 min at 4 °C and the supernatant analyzed by UV spectroscopy. The 60 amino acid residue peptide in the supernatant was either concentrated by rotoevaporation or dried by lyophilization. The pellet contained ca. 50% of the FIXQS-HS(lactone) which could be extracted with 5 mM Im-BB/6G and dialyzed overnight against 500 mM NaCl in dialysis bags of 3350 MWCO (Spectrum) at 4 °C. This dialysate was centrifuged and the supernatant analyzed as before.

Protocol 3: Production and Purification of Recombinant FIXQSH$_6$ 67 Amino Acid Residue Peptides. The production of this recombinant peptide from the KSIFIXQSH$_6$ fusion protein was different than the other two peptides because the His$_6$ tag was designed to remain attached to the peptide after CNBr cleavage at the junctional methionine between the KSI carrier protein and the C-terminal FIXQSH$_6$ peptide. The inclusion body-containing insoluble cell pellets were directly dissolved in 80% formic acid, cleaved with CNBr, and dried by rotoevaporation as in Protocol 1. The resulting gelatinous material was resolubilized in a minimal volume of 6.5 M guanidine-HCl and transferred to 3350 MWCO dialysis bags. This material was dialyzed overnight at 4 °C against 20 mM Tris-HCl, pH 8.1, and 50 mM NaCl. The dialysate was centrifuged at 12000*g* for 15 min at 4 °C and the supernatant filtered through a 0.22- μm filter. The supernatant was loaded onto a Ni chelation affinity column equilibrated with 5 mM Im-BB. Note that unlike the other two protocols, no guanidine-HCl was used during this affinity purification. The column was sequentially washed with 5 mM Im-BB and 30 mM Im-BB and eluted with 300 mM Im-BB. The protein concentration of both FIXQSH$_6$ and FIXQS-HS(lactone) was determined by the method of Goodwin and Morton.[12]

Synthesis of Recombinant α-Factor Peptide Conjugates. Recombinant αI12–homoserine(lactone) (αI12–HSl) produced by Protocol 1 was purified to >99% purity by reverse-phase HPLC and lyophilized to a fluffy white powder. The peptide was stable at −20 °C for at least 6 months without degradation or loss of biological activity; however, the lactone moiety slowly hydrolyzed to the free acid form (<5% after 6 months). The lactone could be completely and immediately converted

to the free acid form of homoserine (αI12-HS) by addition of 0.2 N NaOH. For each of the following reactions, 0.3–0.8 mg (170–460 nmol) of αI12-HS(lactone) were placed in 1.7-mL polypropylene tubes, completely lactonized by addition of 20 μL of 100% TFA, and immediately dried in a SpeedVac (Savant). The dried pellets were dissolved in 50 μL of anhydrous DMF delivered by a gas-tight syringe, and 8 μL of Et_3N was added. All reaction tubes were equipped with a microstirbar. Synthesis of αI12-HSA: 50 μL of anhydrous DMF saturated with NH_3 gas was added directly to the dried lactonized pellet prior to the addition of Et_3N. This reaction was stirred at 44 °C, and amidation was complete in 30 min. Synthesis of αI12-HS-ED: 25 equiv of ethylenediamine was added to the lactonized peptide solution, and the reaction was complete after 68 min at 45 °C. Synthesis of αI12-HS-EO: 250 equiv of ethanolamine was added to the lactonized peptide solution, and the reaction was complete in 68 min at 45 °C. Synthesis of αI12-HS-biotin: 100 equiv of biotin ethylenediamine (Molecular Probes) was added to the lactonized peptide solution, and the reaction was allowed to proceed for 5 h 12 min at 48 °C. Synthesis of αI12-HS-fluorescein: 28 equiv of fluoresceinyl glycine amide (Molecular Probes) was dissolved in an additional 100 μL of DMF and added to the lactonized peptide solution, and the reaction was allowed to proceed for 7 h 45 min at 48 °C. Synthesis of αI12-HS-dansyl: 37 equiv of dansyl ethylenediamine (Molecular Probes) was added to the lactonized peptide solution, and the reaction was allowed to proceed for 7 h 10 min at 48 °C. Synthesis of αI12-HS-ANS: 100 equiv of 5((2-aminoethyl)amino)naphthalene-1-sulfonic acid (Molecular Probes) was dissolved in an additional 200 μL of DMF plus an additional 32 μL of Et_3N and added to the lactonized peptide solution, and the reaction was allowed to proceed for 5 h 50 min at 45 °C.

All syntheses of recombinant α-factor peptide conjugates were monitored by RP-HPLC. Peak fractions were collected and their compositions and concentrations determined by UV spectroscopy. The compositions of each recombinant peptide conjugate were also verified by both laser desorption mass spectrometry and FAB mass spectrometry, as already described.[12]

Yeast Cell Arrest Assays. HPLC-purified recombinant α-factor-HS, α-factor-HS-conjugates, and synthetic α-factors were dissolved in DMSO at concentrations of 1 mg/mL. A 10-fold diluted stock solution was also made for each peptide in 60% H_2O/40% CH_3CN/0.1% TFA. Wild-type α-factor was from Sigma and αI12 and αL12 were synthesized by standard Fmoc solid-phase chemistry. The EY957 yeast strain (MATa/sst1Δ) was kindly provided by E. Elion at the Harvard Medical School. The cell arrest assays were performed according to Elion et al.,[14] except that 100 μL of an overnight culture of EY957 was used instead of 50 μL.

Results

Strategy for the Construction and Production of Tandem Peptides as Fusion Proteins in *E. coli*. The stability of short polypeptides (<80 amino acids) produced in *E. coli* has been shown to be greatly enhanced if these peptides are either fused to a carrier protein[5] or linked together as a large tandem polymer of repeated peptide units.[7] To enhance the stability of the 13 residue yeast α-mating factor test peptide, we produced the peptide as tandem repeated units attached to the highly expressed 375 base pair (125 amino acid) bacterial steroid isomerase gene and upstream of a His_6 cassette. The resulting fusion proteins were then purified from inclusion bodies by Ni chelate chromatography under denaturing conditions and digested with CNBr at junctional methionines releasing free α-factor peptide units. The overall strategy for construction and production of the α-factor fusion proteins is shown in Figure 1A. Complementary oligonucleotides encoding the α-factor peptide were synthesized, annealed, and unidirectionally self-ligated using three-base -ATG-3′, 3′-TAC- overhangs to form a large array of tandem repeated α-factor units separated by single methionine codons. The expression vector pET22KSIH_6 (Figure 1B) was digested with *Alw*NI which recognizes CAGNNN↓CTG. We then unidirectionally ligated the different size tandem repeated α-factor units (αF_n) into the pET22KSIH_6 *Alw*NI site to produce the final pET22KSIαF_nH_6 constructs (n = 0–14), which could be tested for protein production levels. To test the ability of the expression system to produce longer peptides, 60–67 amino acid residue peptides encoding the highly soluble Gla domain of profactor IX (FIXQS, FIXQSH_6) were also constructed as KSI fusions (Figure 1C,D). Note that the FIXQS peptide is produced as a fusion protein, KSIFIXQSMH_6, which differs from the KSIFIXQSH_6 fusion by a single additional methionine residue placed just prior to the C-terminal His_6 tag. The His_6 remains attached to the FIXQSH_6 but not the FIXQSMH_6 peptide after CNBr cleavage.

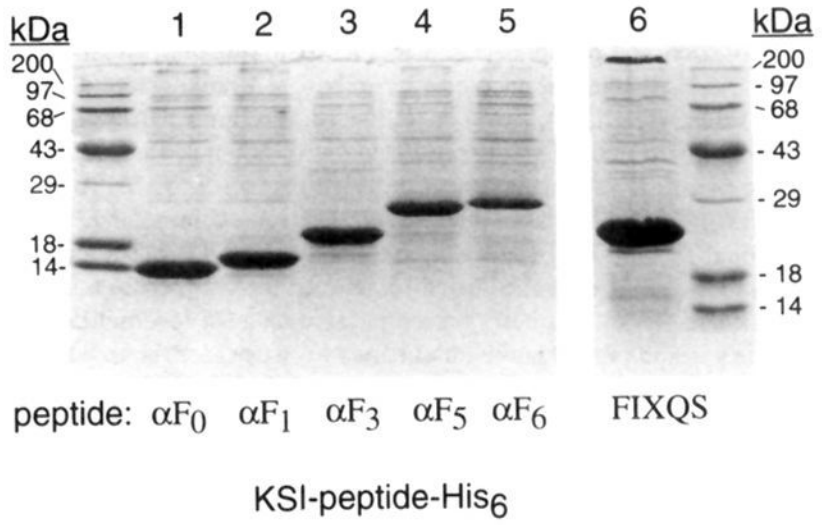


Figure 2. Production of KSIαF_nH_6 and KSIFIXQSMH_6 fusion proteins in *E. coli* as analyzed by 15% SDS-polyacrylamide gels stained with Coomassie Blue R-250. Gel 1: Lanes **1–5** are insoluble protein fractions from IPTG-induced *E. coli* BL21(DE3)pLysS harboring the pET22KSIαF_nH_6 fusion constructs where n = 0, 1, 3, 5, and 6. Molecular weights in kilodaltons are shown on the left for gel 1. Gel 2: Lane 6 is the insoluble protein fraction from the KSIFIXQSMH_6 fusion construct. Molecular weights in kilodaltons are shown on the right for gel 2.

Production and Purification of Recombinant α-Factor, FIXQS, and FIXQS-H_6 Peptides. Each of the different KSIαF_nH_6 and KSIFIXQS(M)H_6 constructs were expressed in the protease deficient strain BL21(DE3)pLysS under the tight control of the T7*lac* promoter in the pET22 vector. The production levels are shown in Figure 2, and the amounts of fusion protein as a percentage of total protein are given in Table 1. The KSIαF_nH_6 fusion proteins were processed into inclusion bodies and comprised the majority of the insoluble protein fraction as shown (Figure 2). The production levels of the KSIαF_nH_6 fusion proteins slightly increased and then gradually decreased as n increased from 0 (70% of total insoluble protein) to n = 6 (55% of total insoluble protein). When n was >6, the production levels dramatically decreased and became <1% of total cellular protein at $n \geq 11$ (data not shown), indicating instability of gene expression at high n values for this test peptide. Since, however, the peptide/KSI-peptide-H_6 mass ratio also increased from 0% (n = 0) to 42% (n = 6), the theoretical maximum yield of peptide occurred at n = 5 (Table 1). Thus, when n = 5 (5 × 14 = 70 a.a.), the α-factor peptide comprised 70 amino acids/206 amino acids or 37.3% (Da/Da) of the mass of the fusion polypeptide yielding the greatest amount of α-factor/L of *E. coli*.

The inclusion body protein pellets were solubilized in 6 M guanidine-HCl, and the KSI-peptide-H_6 fusion proteins were purified to homogeneity by Ni chelation affinity chromatography under denaturing conditions as determined by Ag-stained SDS-PAGE gels. Urea was avoided since it can carbamylate proteins.[15] Fusion proteins could also be purified from the soluble cellular fraction under native conditions, but these Ni chelate chromatography preparations were always contaminated with several minor proteins (ca. 5%) which coeluted from the Ni resin at 300 mM imidazole. The pure KSIαF_5H_6 protein purified in the presence of guanidine was then quantitatively cleaved by CNBr at the junctional methionines releasing KSI (1–112, 113–126), the His_6 tail, and the αF peptide units, each terminating in a HS(lactone) residue. The αF peptides could be easily extracted away from the extremely hydrophobic KSI peptides using 60% H_2O/40% CH_3CN/0.1% TFA. The HPLC chromatogram of

(14) Elion, E. A.; Grisafi, P. A.; Fink, G. R. *Cell* **1990**, *60*, 649–664.

(15) Marshall, R. C.; Inglis, A. S. In *Practical Protein Chemistry*; Darbre, A., Ed.; John Wiley & Sons: New York, 1986; pp 2–66.

Table 1. Production of Recombinant KSI-Peptide-His_6 Fusion Proteins[a]

KSI-Peptide-His_6	peptide monomer length (a.a.)	KSI-peptide length (a.a.)	total KSI-peptide mass (Da)	% peptide/ KSI-peptide (Da/Da)	% [KSI-peptide]/ [tot cell prot] ($\mu g/\mu g$)[b]	% [KSI-peptide]/ [tot insol prot] ($\mu g/\mu g$)[b]	max yield of peptide[c] (mg/L)	actual yield of pure peptide[d] (mg/L)
$KSI\alpha F_0H_6$	0	136	14 900	0	74	70	0	
$KSI\alpha F_1H_6$	14	150	16 678	10.7	76	72	24 ± 8	
$KSI\alpha F_3H_6$	14	178	20 234	26.3	65	74	57 ± 19	
$KSI\alpha F_5H_6$	14	206	23 790	37.3	68	65	72 ± 24	56
$KSI\alpha F_6H_6$	14	220	25 568	41.7	58	55	69 ± 23	
$KSIFIXQSMH_6$	60	194	21 614	32.6	67	65	48 ± 16	50
$KSIFIXQSH_6$	67	193	21 483	37.1	67	58	66 ± 22	55

[a] Fusion genes were expressed in the protease deficient strain BL21(DE3)pLysS. Overnight cultures were diluted 40-fold and induced with 1 mM IPTG at A_{595} = 0.3–0.5 OD. Cells were harvested 6 h after induction. Sonicated cells were centrifuged in order to pellet insoluble cellular proteins and inclusion bodies. Total cellular protein, insoluble protein, and soluble protein fractions were analyzed by SDS-PAGE and densitometry. [b] Determined by the Bradford method using IgG as a standard. [c] Obtained from (peptide/KSI-peptide)([KSI-peptide]/[tot insol prot])(300 mg/L). The total insoluble protein (300 ± 100 mg/L) was obtained from 14 separate measurements of sonicated insoluble material from cells harvested 6 h after induction with 1 mM IPTG. [d] KSI-peptide-His_6 fusion protein was purified from the inclusion body pellets, purified by Ni chelate chromatography, cleaved with CNBr, and purified by HPLC according to Protocols 1–3 as described in the Experimental Section.

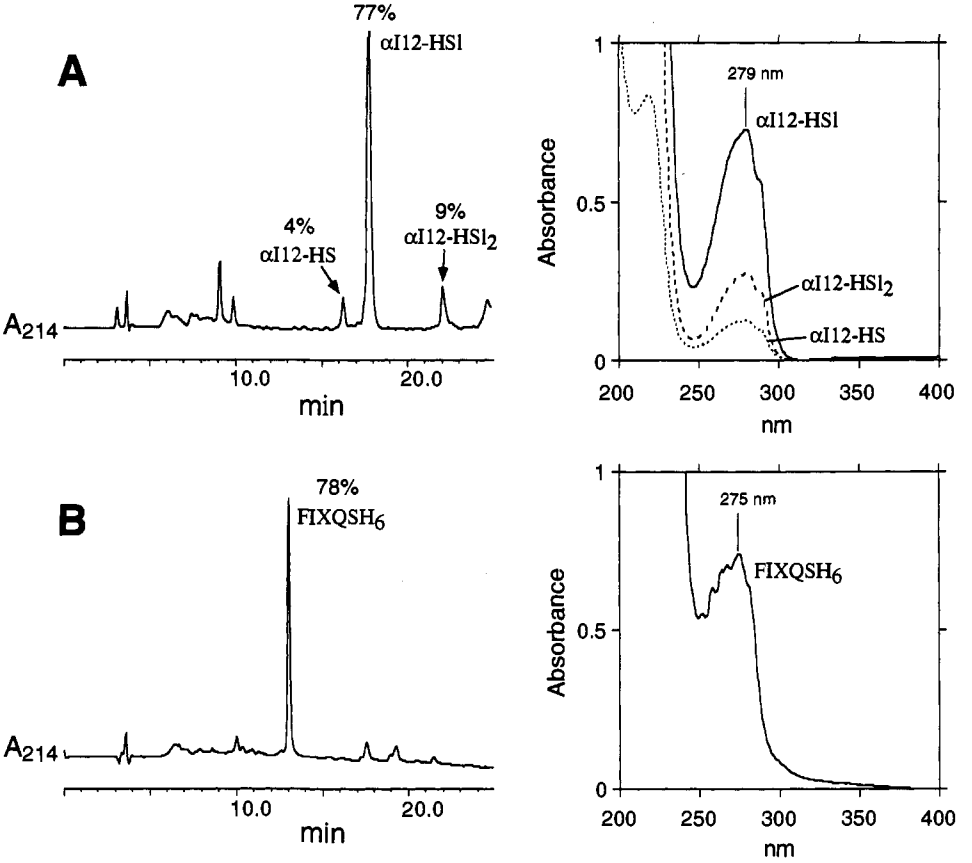


Figure 3. Determination of purity and composition of crude CNBr cleavage reaction mixtures of recombinant α-factor and $FIXQSH_6$ peptides. (A) Reverse-phase HPLC of α-factor CNBr cleavage products. The CNBr reaction mixture was extracted with 60% H_2O/40% CH_3CN/0.1% TFA (Protocol 1 in the Experimental Section) and injected onto a 4 × 250 mm Vydac RP-HPLC column, and a linear gradient (0% + 5%/min for 5 min and then 1%/min for 18 min) of acetonitrile/0.1% trifluoroacetic acid (TFA) in distilled–deionized water/0.1% TFA was employed. Peptide peaks were detected by UV absorbance at 214 nm. Fractions containing the desired peptides were combined and lyophilized. Reinjection of purified peptides showed them to be >99% homogeneous. The composition of each of the indicated peaks was verified by FAB-MS. The UV absorption spectra of each of the indicated HPLC peaks is shown to the right. (B) Same as in A except that the peptide was the recombinant 67 amino acid residue $FIXQSH_6$ peptide purified by Ni chelate affinity chromatography (Protocol 3 in the Experimental Section). For purposes of peak integration, a control HPLC chromatograph run before each of these injections was subtracted from chromatographs A and B. Trace contaminating peaks eluting prior to 9 min (<29% CH_3CN) and after 23 min (>50% CH_3CN) were also present in these control chromatographs.

this extracted CNBr cleavage mixture is shown in Figure 3A. The identity of each peptide peak was determined by mass spectrometry. The reaction mixture yielded 77% αI12-HSl and 4% αI12-HS. Thus, 81% of the peptide mixture remained in the correct form of the peptide; however, 9% dimerized to form $(\alpha I12\text{-}HSl)_2$. For comparison, the yield of the synthetic α-factor peptide produced by nonoptimized solid-phase peptide synthesis was 12–48%.[16] The exact linkage between the two αI12HS peptides in the dimeric species is not known, but the molecular mass was correct for αI12-HS-(αI12)-HSl. Incomplete CNBr cleavage was ruled out since that would have yielded peptide dimer that was 21 mass units heavier (HS versus Met).

Each α-factor peptide HPLC peak was collected and ultraviolet absorption spectra obtained (Figure 3) in order to confirm the aromatic amino acid composition and to provide a reliable method to determine purity and peptide concentration of all three species. Since each α-factor peptide contained 1 Tyr:2 Trp, the expected extinction coefficient was $\epsilon_{279.4}$ = 12 300 M^{-1} cm^{-1} on the basis

(16) Levin, Y.; Khare, R. K.; Abel, G.; Hill, D.; Eriotou-Bargiota, E.; Becker, J. M.; Naider, F. *Biochemistry* **1993**, *32*, 8199–8206.

of the additive spectral contributions of the three aromatic amino acids.[17] The final yield of pure recombinant αI12-HS 14 amino acid residue peptide was 56 mg/L of *E. coli*.

The KSIFIXQSMH$_6$ fusion protein, which eventually yielded the FIXQS-HS 60 amino acid residue peptide, was also overproduced. Again, a majority of the KSIFIXQSMH$_6$ fusion protein was processed into inclusion bodies (Figure 2); however, since the FIXQS sequence is quite hydrophilic,[12] a sizable portion (40%) of the KSIFIXQSMH$_6$ fusion protein remained with the soluble protein fraction. The KSIFIXQSMH$_6$ fusion was designed to have the C-terminal His$_6$ tag clipped off by CNBr from the rest of the FIXQS peptide after Ni chelation chromatography, much in the way of the KSIαF$_n$H$_6$ fusion (compare parts A and C of Figure 1). Thus, the FIXQS(HS) peptide was first purified to homogeneity by Ni chelate chromatography and then cleaved with CNBr to release the free FIXQS(HS) 60 amino acid residue peptide. The strategy for extraction of the highly water soluble FIXQS(HS) peptide from the CNBr cleavage mixture differed from the extraction of the more hydrophobic αI12HS 14 amino acid residue peptide. The FIXQS(HS) was either extracted with phosphate-buffered saline or by solubilization with 6 M guanidine-HCl and subsequently dialyzed against high salt, leaving behind the insoluble KSI peptides. The final yield of pure recombinant FIXQS-HS 60 amino acid residue peptide was 50 mg/L of *E coli*.

The longest recombinant peptide to be produced in this expression system was the FIXQSH$_6$ 67 amino acid residue peptide. As before, the KSIFIXQSH$_6$ fusion protein was overproduced and processed into inclusion bodies (Table 1). Since this peptide was designed to retain its C-terminal His$_6$ tag, no junctional Met residue was inserted before the His$_6$ cassette (Figure 1D). Unlike the first two purification protocols, the KSIFIXQSH$_6$ fusion protein was cleaved by CNBr prior to Ni chelate chromatography. The Ni chelate affinity purified peptide was shown to be 78% pure by HPLC, and ultraviolet absorption spectroscopy was performed as above (Figure 3B). The aromatic absorption spectrum was quite different from the α-factor peptide and was quite typical for a polypeptide containing 1 Tyr:5 Phe, giving a major peak at 275 nm from Tyr (ϵ_{275} 1400 M^{-1} cm^{-1}) and four minor absorption maxima at 251, 257, 263, and 267 nm corresponding to phenylalanine fine structure.[17] The final yield of pure recombinant FIXQSH$_6$ 67 amino acid residue peptide was 55 mg/L of *E. coli*. Since the final yields of pure peptide for both of these longer peptides were comparable to the yield of the expression-optimized (n = 5) tandem repeated α-factor peptide (Table I), we did not examine expression levels of fusion proteins containing tandem repeated FIXQS(H$_6$) peptides.

Synthesis of Peptide-Homoserine Lactone Conjugates. The cleavage of methionine-containing peptides with CNBr will generate a C-terminal homoserine lactone residue, which can be used to create a plethora of peptide-HS conjugates. Due to the unique reactivity of the HS lactone electrophile one should be able to couple a large number of primary-amine containing compounds to the C-terminus of any peptide-HS lactone. We have adopted the chemistries developed for C-terminal coupling of peptide-HS lactones to amino glass resins used in solid-phase peptide sequencing.[18,19] As shown in Figure 4, the αI12-HSl was hydrolyzed or ammonolyzed to form the αI12HS (B) and αI12-HS-amide (αI12-HSA, C), respectively. The HPLC chromatograms of these two reaction mixtures show the nearly complete conversion of the αI12-HSl to the correct products, as confirmed by mass spectrometry. More complex αI12-HS conjugates were also made using the DMF/Et$_3$N coupling system including the biotinylated (D) and fluoresceinated (E) peptides. A number of other recombinant peptide-HS conjugates were also synthesized, and their physical properties are shown in Table 2. The yield of each of the coupling reactions depended largely on the solubility and steric bulk of the attacking group; however, at least a 50% yield was obtained in each case. Hydrazide-containing compounds were unreactive as nucleophiles with the αI12-HSl under the conditions employed.

Biological Activity of Recombinant α-Factor Peptides. The mature wild-type α-factor peptide is a 13 amino acid residue pheromone secreted by type MATα yeast cells, which upon binding to the α-factor receptor transmits a signal to an internally located G protein. Once activated, this G protein triggers a series of intracellular events which lead to cellular differentiation.[14] In the presence of external α-factor, yeast cells in the early G$_1$ phase of growth will undergo cell cycle arrest. The bioactivity of each recombinant α-factor could easily be determined in halo assays as a region of low cell density surrounding a central disk spotted with α-factor (Figure 5). Prior to constructing the tandem recombinant α-factor gene which lacked the wild-type Met at position 12, we first synthesized and tested α-factor peptides which had this internal Met changed to both Ile and Leu. The cell arrest$_{50}$ values (CA$_{50}$) and maximal zones of inhibition (zoi$_{int}$) of both αI12 and αL12 were essentially identical with those of wild-type (αFWT) α-factor (Table 2). Thus, we proceeded to construct the tandem recombinant α-factor peptide with an Ile at position 12. The C-terminal HS-free acid form of the recombinant peptide, αI12HS, was shown to have the highest bioactivity. Its CA$_{50}$ value was 128% of the wild-type value, and its zoi$_{int}$ value was two-thirds of the of wild-type value. Interestingly, the amidation of the C-terminal HS lactone, αI12HSA, resulted in a >11-fold increase in CA$_{50}$, indicating that a negative charge at the C-terminus[20] is important for binding of these C-terminally extended α-factor analogs. An additional C-terminal extension of the amide to either the ethylenediamine (αI12HS-ED) or ethanolamine (αI12HS-EO) derivatives completely abolished binding to the α-factor receptor. The lactone form of the recombinant α-factor (αI12HSl), however, had a 2-fold higher CA$_{50}$ value relative to αFWT, and its zoi$_{int}$ was one-third that of αFWT, also consistent with the C-terminus requiring a free carboxylate to bind the receptor. It is possible that this partial activity could be due to hydrolysis of the lactone on the solid media during the bioassay. None of the larger αI12HS conjugates had any appreciable agonist activity (Table 2) nor did they antagonize αFWT in dual disk competition assays.[21] These data would agree with the evidence that the C-terminal portion of the α-factor is critical for receptor binding.[21]

Discussion

Two groups pioneered the use of the tandem gene approach to produce limited amounts of recombinant peptides in *E. coli*.[4,7] Shen was the first to describe the general idea of tandem production of peptides and their cleavage with CNBr. He made tandem copies of the insulin gene either alone or attached to the first 80 codons (80 amino acids) of the β-galactosidase gene (*lacZ*). These fusion constructs were placed under the control of a *lac* or *tac* promoter and expressed in *E. coli* strain JM101. Inclusion bodies, which were highly enriched with the insulin-containing fusion protein, were treated with CNBr and the products analyzed by polyacrylamide gel analysis. This process had two major limitations. First, the linker region between the tandem insulin monomers must encode four extra amino acids plus the HS residue. Thus, each peptide, except the last one, must terminate in an extra pentapeptide: Arg-Arg-Asn-Ser-HS. The other major limitation is the lack of an easy method of purification of the fusion construct prior to CNBr cleavage.

In a follow-up to this study, Lennick et al.[3] introduced a slight variation on Shen's original process. In this paper they described the production of 8-tandem copies of the human atrial natriuretic

(17) Kuliopulos, A.; Mildvan, A. S.; Shortle, D.; Talalay, P. *Biochemistry* **1989**, *28*, 149–159.
(18) Horn, M. J.; Laursen, R. A. *FEBS Letts.* **1973**, *36*, 285–288.
(19) Horn, M. J. *Anal. Biochem.* **1975**, *69*, 583–589.
(20) Naider, F.; Becker, J. M. *CRC Crit. Revs. Biochem.* **1984**, *21*, 225–248.
(21) Eriotou-Bargiota, E.; Xue, C.-B.; Naider, F.; Becker, J. M. *Biochemistry* **1992**, *31*, 551–557.

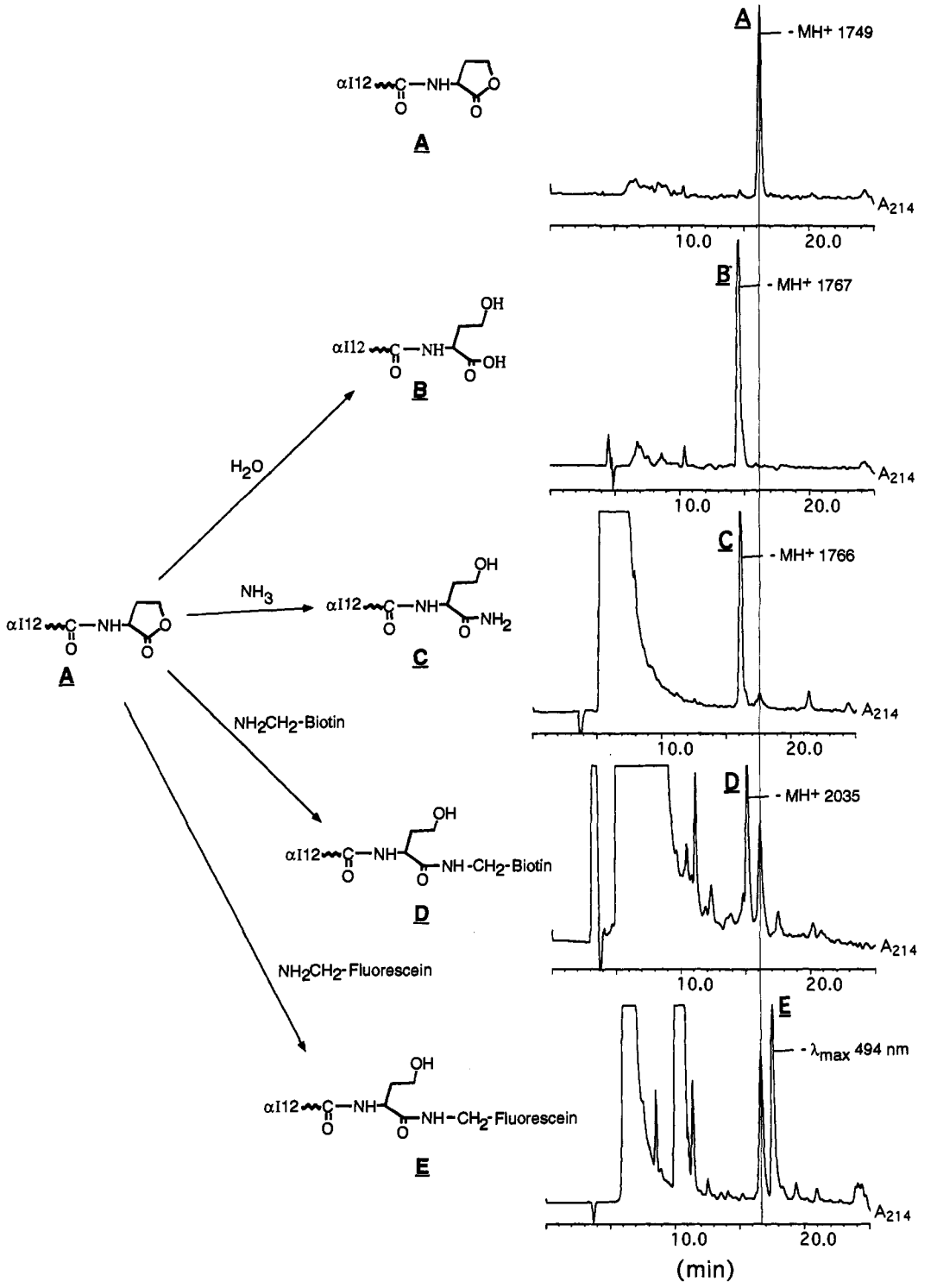
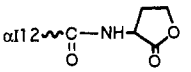
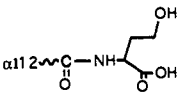
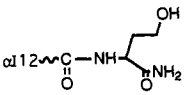
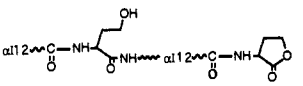
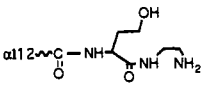
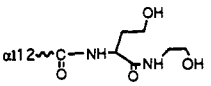
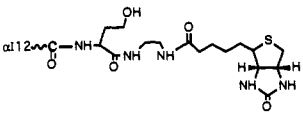
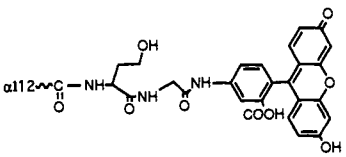
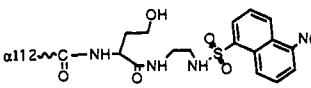
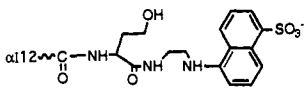


Figure 4. Synthetic scheme showing general routes to formation of recombinant α-factor peptide conjugates. The αI12-homoserine lactone (αI12HSl) was dissolved in anhydrous DMF and reacted with the indicated reagents in the presence of Et_3N. Incubation temperatures and times for each reaction are listed in the Experimental Section. HPLC chromatograms of each of the reaction mixtures are shown on the right. HPLC conditions are the same as described in the Figure 3 caption. The large peak eluting at 5–6 min in the bottom three traces is DMF. The masses (MH^+) of the peptide conjugates were determined by FAB-MS except for **E** where the fluorescein conjugate was characterized by its visible spectrum. **A**, αI12-HS(lactone); **B**, αI12-HS; **C**, αI12-HS-amide; **D**, αI12-HS-biotin; **E**, αI12-HS-fluorescein, the absorption maxima is indicated. The peaks eluting prior to 13 min in Figure 4D and prior to 15 min in Figure 4E were also present in the control injections of biotin or fluorescein in DMF alone (data not shown).

28 amino acid residue peptide (hANSP) placed behind the 7 N-terminal amino acids of β-galactosidase. Lysine-containing tetrapeptide linkers were necessary for *Bgl*II/*Bam*HI unidirectional ligation cycles. Lennick et al. extracted the fusion protein from the inclusion bodies with 10 M urea and did a further gel filtration purification step. This partially purified protein was then cleaved into individual 29 amino acid residue peptides with the endoprotease Lys-C at interspersed Lys residues rather than at Met residues with CNBr. Lennick et al. then removed the C-terminal Lys residue with carboxypeptidase B. Although the data were not shown, the final recombinant 28 amino acid residue peptide was reported to have biological activity similar to the chemically synthesized peptide.

A second group[4] expressed tandem copies of a 10 amino acid residue peptide (substance P) as a fusion construct consisting of cro repressor–peptide units–lac I–lac Z. Thus, their carrier protein was quite large: cro-21 a.a. + lac I-38 a.a. + lac Z-997 a.a. = 1056 a.a. totaling 120 kDa. Because this carrier protein was so

Table 2. Physical Properties and Biological Activity of Recombinant Yeast α-Factor Peptide Conjugates

peptide	structure	MH+[a] (m/z)	yield[b] (%)	z.o.i.$_{int}$[c] (mm)	CA$_{50}$[d] (pmoles)
Synthetic Peptides					
αFWT	WHWLQLKPGQP**M**Y	1687[e]		29 ± 3	74 ± 5
αI12	WHWLQLKPGQP**I**Y	1666[e]		27 ± 1	84 ± 33
αL12	WHWLQLKPGQP**L**Y	1666[e]		26 ± 2	91 ± 40
Recombinant Peptides					
αI12-HSl		1749[f]		11 ± 3	170 ± 27
αI12-HS		1767[f]	97	18 ± 3	95 ± 37
αI12-HSA		1766[f]	> 88		> 830
αI12-HSl$_2$[g]		3500[e]			
αI12-HSED		1813[e]	> 90	no effect at 5,500 pmoles	
αI12-HSEO		1814[e]	> 90	no effect at 5,500 pmoles	
αI12-HS-Biotin		2035[f]	58		> 950
αI12-HS-Fluorescein[h]			56	no effect at 6,700 pmoles	
αI12-HS-Dansyl		2042[f]	50	no effect at 4,900 pmoles	
αI12-HS-ANS		2017[e]	51	no effect at 740 pmoles	

[a] The molecular ion (MH+) of each HPLC-purified peak was determined by FAB-MS or by laser desorption-MS, as indicated. In each case, the observed MH+ mass agreed to within 1 Da of the calculated mass. [b] Determined by integration of HPLC traces at 214 nm of reaction mixtures described in the Experimental Section. [c] The zone of inhibition in millimeters (zoi) was measured for each series of different concentration α-factor and plotted against amount (pmol) of α-factor spotted per disk. The sigmoidal curves that were generated had three regions including a horizontal line drawn through 6.5 mm (the diameter of the disk), the sigmoidal portion of the curve, and an upper asymptope which always had a positive slope. The *y*-intercept of this asymptope (zoi$_{int}$) obtained for each α-factor peptide served as a reliable measure of the *relative* maximal effects of each α-factor peptide in the cell arrest assay. [d] The cell arrest$_{50}$ (CA$_{50}$) values in picomoles were calculated by drawing a tangent through the inflection point of the sigmoidal portion of the curve and obtaining the point of intersection with the upper asymptope. This value was divided by 2 and 3.25 nm was added to obtain the zoi value at which 50% of the maximal effect is seen (CA$_{50}$) or the amount of peptide required to cause 50% of cell arrest under the conditions employed. [e] Determined by laser desorption mass spectrometry. [f] Determined by FAB mass spectrometry. [g] Formed spontaneously during the CNBr cleavage reactions. Although the most likely structure is drawn, alternative amide bond linkages (i.e. via the Lys-ε-NH$_2$) could form with the C-terminal lactone and still have the molecular mass as shown. [h] The FAB-MS for this compound gave no peak, and the laser desorption-MS analysis was not reliable; however, the HPLC-purified compound, **E**, shown in Figure 4 had the expected UV–vis absorption characteristics of a fluorescein-peptide derivative with a maximal absorption occurring at 494.3 nm in 60% H_2O/40% CH_3CN/0.5% Et_3N, pH 10. The authentic fluorescein-glycine-amide starting material had a nearly identical absorption spectra and maximally absorbed at 497.6 nm under the same conditions.

large, they had to express up to 64 tandem copies of their peptide to obtain measurable amounts of peptide. In addition to the huge excess of carrier protein, the intrapeptide regions contained *Hin*dIII/*Bgl*II/*Bam*HI linkers which encoded an additional pentapeptide: Asp-Leu-Ser-Phe-Met. This linker was excised before the first Asp by CNBr cleavage after a preceding Met residue leaving the C-terminal HS residue common to all three methods. Thus, both Shen and Kempe et al. incorporated a penta- or tetrapeptide-encoding linker region which contained the restriction sites and base overhangs necessary to achieve unidirectional ligation. Recently, another group[22] has significantly advanced this technique to make ((AG)$_3$PEG)$_n$ peptides (n = 1–27) for studies of the solid-state properties of these polymers but have pared the linker region down to a hexanucleotide (*Ban*I site) encoding a Gly-Ala dipeptide.

In our method we have improved on these earlier ligation strategies by requiring only a single codon, ATG (methionine), for unidirectional ligation to occur with no restrictions on adjacent

(22) McGrath, K. P.; Fournier, M. J.; Mason, T. L.; Tirrell, D. A. *J. Am. Chem. Soc.* **1992**, *114*, 727–733.

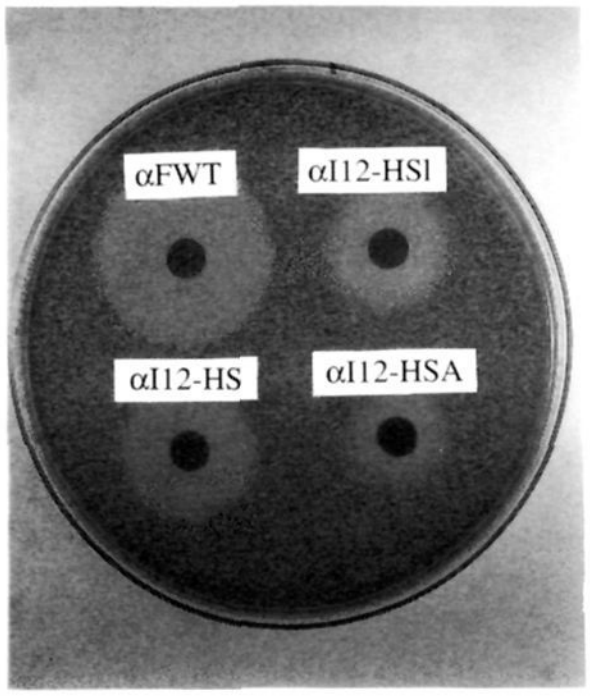


Figure 5. Bioactivity of recombinant α-mating factor peptides as determined by yeast cell arrest halo assays. Lawns of yeast MATa/sst1Δ strain EY957 were grown on YPD plates and were tested for α-factor sensitivity, as previously described (Elion et al., 1990). α-Factor peptides were dissolved in DMSO and spotted on sterile 6.5-mm paper disks. Peptide amounts were as follows: wild-type α-factor (αFWT), 595 pmol; αI12-homoserine(lactone) (αI12-HSl), 6850 pmol; αI12-homoserine (αI12-HS), 1130 pmol; αI12-homoserine-amide (αI12-HSA), 2820 pmol. The zone of inhibition (zoi) for each peptide is the diameter of the low cell density region.

coding sequence. This was made possible by the use of the restriction endonuclease, *Alw*NI, which recognizes CAGNNN↓CTG (Figure 1B). Since NNN (N = C, A, T, or G) can encode any amino acid including a relatively rare and chemically reactive amino acid such as Met or Trp, the tandem gene linker sequences could be trimmed down to a single nonpalindromic codon. Thus, we can cleave our expression vector pET22KSIH_6 with *Alw*NI, leaving a single -ATG-3′ three-base overhang suitable for accepting any duplex DNA with a 3′-TAC- overhang and a 5′-ATG- overhang, thereby preventing any potential frameshifts by using 1-, 2-, 4-, or 5-base overhangs generated by most other commercially available restriction endonucleases. Note that blunt-end digestion (0-base overhang) will not allow for unidirectional ligation. In our process a single ligation step will form all the necessary size DNA multimers (Figure 1A) for subsequent creation of a large array of different size fusion constructs which can be empirically tested for gene expression levels (Figure 2).

A second advantage to our method is that our protein production levels are consistently 2- to 5-fold higher than that demonstrated by the other methods. Our high gene expression levels result from the combination of a well-produced carrier protein, KSI, expression in the protease deficient *E. coli* strain BL21/DE3/pLysS and the use of the very tightly controlled pET22 vector system.[13] Thus, we were able to produce and purify three different test peptides by Ni chelate chromatography under denaturing conditions at very high yields of 50–56 mg/L of *E. coli* (Table 1). The KSI-peptide-H_6 fusion proteins eluted from the Ni affinity column as a single pure band, and the level of purity was independent of target peptide sequence. After CNBr cleavage, the extremely hydrophobic KSI carrier protein is essentially insoluble even in up to 40% acetonitrile/60% water (Figure 1A). The purification protocols have been tailored for each different recombinant peptide depending on their solubility in water (see Experimental Section). Thus our "crude" α-factor peptide is at least 80% pure (Figure 3) prior to further HPLC purification. The remaining impurities are primarily αF dimers formed via the lactone group.

Furthermore, this method has a tremendous advantage over solid phase peptide synthesis in terms of production and purification of peptides in the range 20 to ~100 amino acids. Nonoptimized solid-phase peptide synthesis often results in the cosynthesis of contaminating peptides, which can pose difficulties in the purification of peptides greater than 20 amino acids in length. In fact, a 60–67 amino acid residue peptide presents quite a challenging HPLC purification since the starting crude synthetic peptide mixture (only ca. 30% pure) has thousands of closely migrating contaminants such as peptides with deletions of single amino acids, which are sometimes impossible to purify away from the desired peptide using available chromatographic methods.

An important limitation to our method is that the tandemly produced target peptide must lack internal methionine residues if CNBr cleavage is to be used to release the target peptide from the fusion protein. One could circumvent this potential problem by mutating the internal methionine to a conserved residue such as isoleucine, as we have done for both the α-factor peptide and the FIXQS(H_6) peptides with no effect on bioactivity. Alternatively, one could easily incorporate unique Trp or Tyr residues at the junctional cleavage sites in place of Met and quantitatively cleave with oxidants such as *o*-iodosobenzoic acid or *N*-bromosuccinimide,[6] agents with which we have had preliminary success.[23] Since CNBr cleavage reactions can lead to some side reactions including nonspecific cleavage at Tyr or Trp by oxidative halogenation and oxidation of Cys residues,[6] it is essential to use fresh reagents and to carry out the reactions in the absence of oxygen. Another limitation of this method is that the C-terminal homoserine lactone residue generated by CNBr cleavage may have unforeseen effects on bioactivity, though these adverse effects were not seen with the peptides tested here.

Lastly, because of the fundamental difficulty of making peptide conjugates coupled uniquely at the C-terminus, we decided to develop a generally applicable method whereby compounds containing a primary amino group could be reacted with the C-terminal lactone group generated by CNBr cleavage. Previous systems employing 30% NH_4OH/MeOH4 as an amidating agent did not work since like most peptides the α-factor peptide was insoluble in MeOH and the NH_4OH generated too much hydrolysis side-product. A similar HS lactone coupling chemistry using dimethy lsulfoxide/triethylamine as the solvent system with ethylenediamine or 30% NH_4OH as the amine nucleophile was recently reported.[24] We instead employed anhydrous DMF with triethylamine as a base to achieve a simple, reliable, and widely applicable coupling system. In total, seven different α-factor peptide-HS-NH-R compounds were synthesized, purified, characterized, and tested for biological activity in an *in vitro* yeast cell cycle arrest assay (Table 2, Figure 5). Two of the recombinant α-factor peptides had biological activity similar to wild-type α-factor peptide. Together, the coupling of the α-factor peptide to a number of primary-amine-containing compounds has demonstrated the biological utility of our HS lactone coupling chemistry. Finally, we would like to emphasize that this relatively inexpensive method for production and purification of polypeptides in high yield and purity will allow researchers lacking solid-phase peptide synthesis capabilities ready access to the use of peptides in their research.

Acknowledgment. We are particularly indebted to Noele Nelson for her help in the construction and expression of the KSI fusion genes. We also thank Stan Tabor, Yian Shi, and the members of the Elion laboratory for stimulating discussions and suggestions and Lesley Stolz for critically reading the manuscript. Mass spectral data were provided by Ioannis Papayannopoulos who carried out the FAB mass spectral analysis at the MIT Mass Spectrometry Facility (supported by NIH Grant RR00317 to K. Biemann) and by Jim Lee who did the laser desorption mass spectrometry. This work was supported by a Judith Graham Pool Fellowship from the National Hemophilia Foundation (to A.K.) and a National Institutes of Health Grant HL42443-02.

(23) Kuliopulos, A. Unpublished data.

(24) Callaway, J. E.; Lai, J.; Haselbeck, B.; Baltaian, M.; Bonnesen, S. P.; Weickmann, J.; Wilcox, G.; Lei, S.-P. *Antimicrob. Agents Chemother.* **1993**, *37*, 1614–1619.