Annu. Rev. Biophys. Biomol. Struct. 1995. 24:435–62

SITE-DIRECTED MUTAGENESIS WITH AN EXPANDED GENETIC CODE

David Mendel

Lilly Research Laboratories, Eli Lilly and Company,
Lilly Corporate Center, Indianapolis, Indiana 46285-0540

Virginia W. Cornish and Peter G. Schultz

Department of Chemistry, University of California, Berkeley,
California 94720

KEY WORDS: unnatural amino acids, site-directed mutagenesis, in vitro suppression, protein structure-function

---

CONTENTS

ABSTRACT

A biosynthetic method has been developed that makes possible the site-specific incorporation of a large number of amino acids and analogues within proteins. In this approach, an amber suppressor tRNA chemically aminoacylated with the desired amino acid incorporates this amino acid site specifically into a protein in response to an amber codon introduced at the corresponding position in the protein's DNA sequence. Using this method, precise changes within a protein can be

1056-8700/95/0610-0435$05.00

made to address detailed structure-function questions. A series of fluorinated tyrosine analogues and linear, branched, and cyclic hydrophobic amino acids have been used to determine the impact of hydrogen bonding and hydrophobic packing, respectively, on protein stability. Glutamate analogues and conformationally restricted amino acids have been used to probe the mechanisms of staphylococcal nuclease and ras. In addition, this technique has been used to construct photocaged proteins and proteins containing photoaffinity labels, spin labels, and isotopic labels at specific positions in the protein sequence suitable for biophysical studies.

INTRODUCTION

Although proteins are involved in virtually every biological process, relatively little is understood about the detailed mechanisms by which these biopolymers, composed of 20 simple building blocks, carry out their remarkable functions. One important tool for probing the forces that govern protein structure and folding, biomolecular recognition, and catalysis is site-directed mutagenesis, in which a specific amino acid in a protein can be replaced with any of the other 19 common amino acids (57, 112). In contrast to small molecule synthesis where myriad changes in structure can be made, changes in protein structure are limited to the 20 natural amino acids. For example, few isostructural (e.g. Thr → Val) or isofunctional (e.g. Glu → Asp) substitutions can be made, making it difficult to dissect out the effects of a given mutation in terms of the specific steric or electronic features of a given amino acid. Ideally, the investigator would tailor replacements to address the particular structure-function question at hand. Such replacements would modify the size, acidity, nucleophilicity, or hydrogen-bonding or hydrophobic properties of an amino acid side chain or would modify the protein backbone itself. Alternatively, analogues could be introduced that have altogether new properties, beyond those specified by the genetic code. These novel molecules include spin labels, affinity labels, and redox-active or metal-chelating amino acids. The ability to substitute such unnatural amino acids into proteins would greatly expand the scope of physical organic studies on proteins.

Several methods can be used to incorporate unnatural amino acids into proteins. Solid-phase synthesis, particularly recent advances in the segment synthesis-condensation approach, has allowed for the synthesis of small proteins containing novel amino acids in milligram quantities (24, 44, 48, 76, 99). Protein semisynthesis, in which a synthetic peptide is ligated to a protein fragment to produce a full-length protein,

has also been used to incorporate unnatural amino acids into proteins (12, 82). In addition, chemical modification has been used to introduce a variety of unnatural side chains, including cofactors, spin labels, and oligonucleotides, into proteins (19, 49, 50, 77, 88, 89). Alternatively, biosynthetic methods that employ chemically modified aminoacyl-tRNAs have been used to incorporate several biophysical probes into proteins synthesized in vitro (11, 59). In general, these approaches are limited by difficulties in achieving site-specific incorporation of the amino acids, by the requirement that the amino acids be simple derivatives of the common 20 amino acids, or by problems inherent in the synthesis of large proteins or peptide fragments. However, recent developments in the stepwise enzymatic condensation of peptides to generate modified proteins show considerable promise (45).

Recently, a biosynthetic approach was developed that allows for the relatively facile site-specific incorporation of unnatural amino acids into proteins (Figure 1) (3, 27, 79). This method takes advantage of the fact that the genetic code contains three stop codons (23). Because only one stop codon is needed for translation termination, the other two

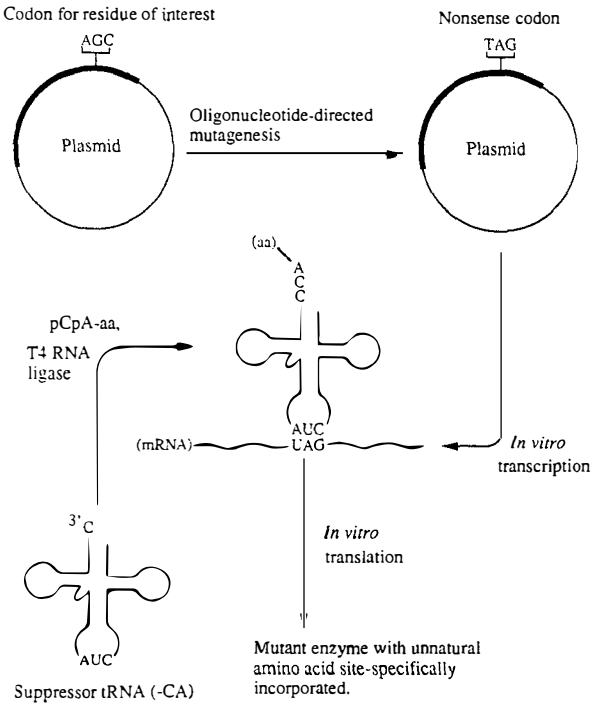


Figure 1 Method for the site-specific incorporation of unnatural amino acids.

can in principle be used to encode nonproteiogenic amino acids. A suppressor tRNA is prepared that recognizes the stop codon UAG (UAG is suppressed in vivo with relatively high efficiency) and that is chemically aminoacylated with the desired unnatural amino acid. Conventional site-directed mutagenesis (96) is used to introduce the stop codon UAG at the site of interest in the protein gene. When the acylated suppressor tRNA and the mutant gene are combined in an in vitro transcription-translation system, the unnatural amino acid is incorporated in response to the UAG codon to give a protein containing that amino acid at the specified position. Early experiments using [^{3}H]-Phe and later experiments with α-hydroxy acids demonstrated that only the desired amino acid is incorporated at the position specified by the UAG codon and that this amino acid is not incorporated at any other site in the protein (28, 79). More recent work has shown that the *Escherichia coli* protein biosynthetic machinery (the in vitro transcription-translation system) can tolerate a broad range of amino acids (27). α,α-Disubstituted amino acids, amino acids in which the side chain is a large reporter group, and even analogues in which the backbone is modified, such as lactic acid, have all been incorporated using this methodology. In vitro protein synthesis is inefficient relative to in vivo synthesis; however, this method, though labor intensive, can yield hundreds-of-microgram to single-milligram quantities of purified protein containing an unnatural amino acid. We have produced several such proteins in quantities sufficient for their characterization using circular dichroism (CD), nuclear magnetic resonance (NMR) spectrometry, and three-dimensional X-ray crystallography (28, 30, 47, 69).

METHODOLOGY

The biosynthetic site-specific incorporation of an unnatural amino acid or amino acid analogue into a protein requires generation of the requisite suppressor tRNA, a general method for acylating the tRNA with unnatural amino acids or amino acid analogues, and the demonstration that sufficient protein can be synthesized in vitro to carry out meaningful mechanistic or structural studies. The suppressor tRNA must be recognized by the protein biosynthetic machinery sufficiently well to suppress a stop codon with high efficiency and at the same time cannot be a substrate for any of the aminoacyl-tRNA synthetases present in the in vitro protein-synthesis extract. If the suppressor tRNA were recognized by any of the aminoacyl-tRNA synthetases, the tRNA could be subject to proofreading (deacylation of the noncognate amino acid) and/or aminoacylation with one of the common 20 amino acids. Such

recognition would result in either low suppression efficiencies or in the incorporation of a common amino acid, as well as the desired unnatural amino acid, in response to the stop codon.

Using the above criteria, we constructed an amber suppressor tRNA derived from yeast phenylalanine-tRNA ($tRNA_{CUA}$), in which residues 34–37 in the anticodon loop were replaced by 5′-CUAA-3′, for use in an *E. coli* in vitro protein transcription-translation system (2, 9, 10, 60, 74). Alternatively, a suppressor tRNA derived from *E. coli* $tRNA^{Gly}$ (92) has been used to suppress UAG codons in short peptides in a rabbit reticulocyte translation system (3). Runoff transcription can readily generate both suppressor tRNAs in relatively large quantities (75, 80). Even though these tRNAs are produced in vitro and may not be modified biochemically (40), both incorporate unnatural amino acids efficiently and exhibit little read-through (incorporation of any amino acid other than the desired unnatural one in response to the UAG codon) at optimal Mg^{2+} concentrations.

Given the high specificity of the aminoacyl-tRNA synthetases for their cognate amino acids (35, 97), an alternative method for acylation of the tRNA had to be developed to ensure that a wide range of unnatural amino acids could be incorporated using this methodology (Figure 2). A modification of the two-step method originally developed by Hecht (42) was used in which the dinucleotide pCpA is first chemically

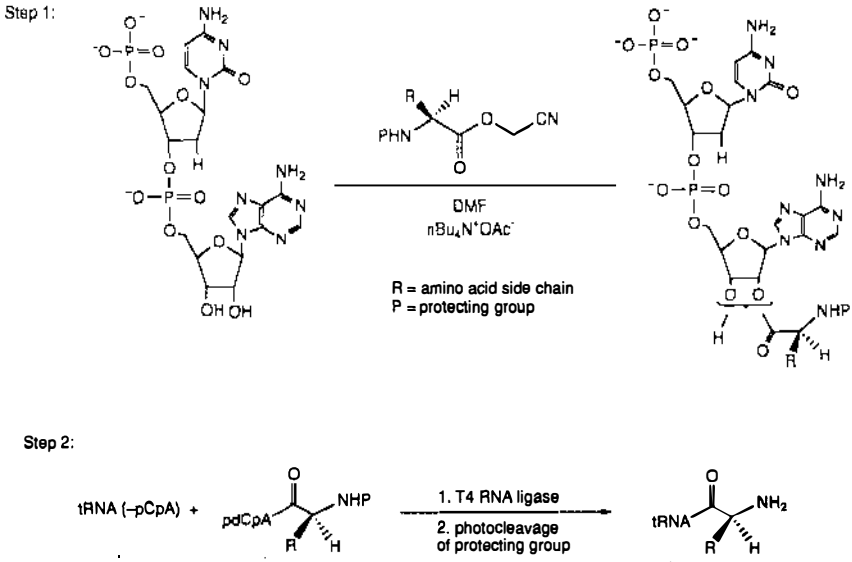


Figure 2 Strategy for the chemical aminoacylation of $tRNA^{Phe}_{CUA}$.

acylated with an Nα-protected amino acid and then enzymatically ligated to a truncated tRNA ($tRNA^{-CA}$) missing the terminal dinucleotide pCpA at the 3′-acceptor stem (94). We found that the cyanomethyl ester of an Nα-protected amino acid reacts selectively at the 2′,3′ hydroxyl functions of the ribose to give the desired monoacylated product, obviating the need to protect and deprotect the dinucleotide (Figure 2) (93). In addition, cytidine was replaced with deoxycytidine in the dinucleotide; this replacement significantly simplified the synthesis and eliminated another reactive 2′-OH group without affecting biological activity. Prior to acylation of the dinucleotide, the α-amino group of the amino acid and any reactive side chain functions are protected as their nitroveratryloxy (NVOC) carbamate, ester, or ether derivatives (1, 85). These protecting groups can be removed photochemically from the intact aminoacyl-tRNA (after ligation) in high yield under mildly acidic conditions; acid pH prevents deacylation of the aminoacyl-tRNA because aminoacyl-tRNAs hydrolyze rapidly at neutral pH (93). For light-sensitive amino acids, the biphenylisopropyloxycarbonyl (BPOC) protecting group can be used instead of the NVOC group. The BPOC group is cleaved with mild acid from the aminoacyl-pdCpA prior to ligation to $tRNA^{-CA}$ (3, 93). This aminoacylation protocol is relatively straightforward, proceeds in high yield (both the aminoacylation and ligation reactions), and applies to amino acids with various side chains (27, 37). Nevertheless, the generation of a nonspecific aminoacyl tRNA synthetase—either a lipase, mutagenized synthetase, or a catalytic antibody—could significantly simply this step (13, 53, 54).

Mutagenesis with the chemically acylated suppressor tRNA is currently carried out in vitro because no general methodology exists for introducing large quantities of the aminoacylated suppressor tRNA into intact, dividing cells (38). An *E. coli* in vitro transcription-translation protein synthesis system is being used in our laboratory, consisting of an *E. coli* S-30 extract containing all of the proteins and RNAs required for transcription and translation, nucleotide triphosphates, phosphoenolpyruvate, pyrophosphatase, a variety of salts (Mg^{2+} and Ca^{2+} concentration can significantly affect protein yields) and cofactors, and the DNA encoding the gene of interest. Typically, the gene of interest is placed under a strong T7 or bacterial RNA polymerase promoter. The in vitro reaction can be carried out on a 5-ml scale. The requisite S-30 can be prepared fairly easily on a 2-liter scale; the suppressor tRNA, on a 30-mg scale; the aminoacyl-pdCpA, on a 0.2-mmol scale; and the plasmid, on a 12-mg scale (27). In vitro protein synthesis remains limited by the relatively small quantities of protein that can be obtained. In general, protein yields greater than 100 $\mu g\ ml^{-1}$ have not been

achieved with either plasmid DNA using an *E. coli* transcription-translation system or with mRNA using a rabbit reticulocyte translation system (27). Unfortunately, in vitro protein yields do not necessarily correlate with those obtained in vivo, and we have found no simple and general rules for ensuring high levels of protein expression in vitro.

In order for this methodology to be useful, the protein biosynthetic machinery must be capable of incorporating a wide variety of amino acids into proteins (14, 41). In fact, approximately 70 different amino acids and analogues have been tested thus far with the *E. coli* system, and a large majority (>50) of these are accepted with reasonable efficiency by that system (Figure 3) (27). The structural and electronic requirements of the rabbit reticulocyte in vitro system are less well defined than the *E. coli* protein-synthesis apparatus (just over 10 nonnative amino acids and analogues have been tested) (5), but the available data suggest rather broad similarities. Conformationally restricted amino acids (e.g. methanoproline, cyclopropylglycine, and α-methyl-L-leucine), amino acids with spin labels and photoaffinity labels as side chains, amino acids with altered pK_a values and hydrogen-bonding properties, photocaged amino acids, α-hydroxy acids, and amino acids with unusual steric properties (e.g. *t*-butylglycine) have all been incorporated into proteins (Figure 3). Suppression efficiencies can vary widely depending on the nature of the amino acid, ranging from 100% for substitution of L-norleucine for Gln28 in 434 repressor (104) to 14% for incorporation of α-methyl-L-leucine in T4 lysozyme (68). Although the suppression efficiency of a given amino acid cannot be predicted a priori, several trends emerge when the identities of the amino acids tested are correlated with their suppression efficiencies (20, 68). In general, large hydrophobic amino acids such as *p*-benzoyl-L-phenylalanine are inserted with higher efficiency than are small amino acids such as alanine and glycine, or charged amino acids such as homoglutamate or ornithine. This correlation is in agreement with reports that large, hydrophobic amino acids such as Trp and Phe bind more tightly to *E. coli* elongation factor-Tu (EF-Tu) than do other natural amino acids such as Ala, Glu, and Lys (64).

The stereochemistry of an amino acid also affects suppression efficiency. L-amino acids and some α,α-disubstituted amino acids such as α-aminoisobutyric acid (AIB) and cyclopropyl through cyclohexylglycine can be incorporated into proteins (68). Neither the *E. coli* nor the rabbit reticulocyte in vitro protein-synthesis system seems able to accommodate D-amino acids or α-methyl-D-amino acids (5, 68). These stereochemical requirements are in agreement, for the most part, with previous studies on the affinity of a variety of unnatural aminoacyl-

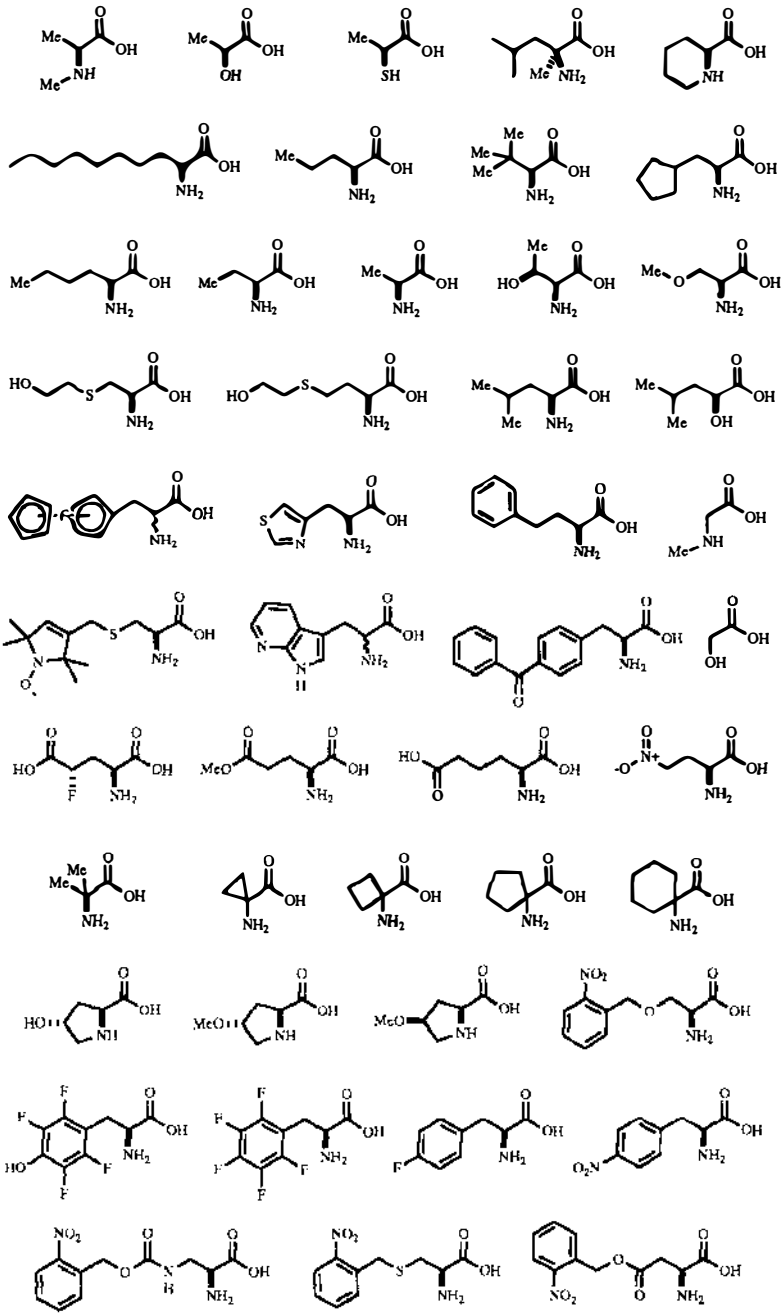


Figure 3 Examples of natural and unnatural amino acids successfully incorporated into proteins via in vitro suppression.

tRNAs for *E. coli* EF-Tu and the ribosome (14). Interestingly, previous experiments that examined the ability of an aminoacyl-tRNA or aminoacyl-CA to bind to the *E. coli* ribosomal A-site and participate in peptide elongation suggested that achiral amino acids such as AIB could not be used as substrates by the *E. coli* protein-biosynthetic machinery (91).

Both the *E. coli* and the rabbit reticulocyte protein biosynthetic machinery also incorporate α-hydroxy acids such as lactic, glycolic, and phenyllactic acid as well as mercapto acids (14, 22, 31). The resulting ester linkages, while similar in structure, are much easier to hydrolyze than amide linkages. A novel feature of mutants containing α-hydroxy acids is their clean and rapid hydrolysis under the SDS-PAGE conditions (17, 28). For example, T4 lysozyme (T4L) containing an Ala82 $\rightarrow$ Lac replacement retained wild-type enzymatic activity, and a band at the expected 18.7 kDa was observed when SDS-PAGE was carried out in a neutral pH buffer system. When the standard pH 8.8–9.0 buffer was used, a band at ~8 kDa was observed, which is consistent with hydrolysis at position 82. If any of the 20 natural amino acids had been inserted in response to the UAG signal at position 82 in this mutant, an amide linkage would have been formed that is stable to the alkaline conditions of conventional SDS-PAGE. The fact that no 18.7-kDa band was observed even when the gel was heavily overloaded reinforces the fact that heterogenous incorporation of natural amino acids does not occur.

The translational machinery accommodates various other backbone mutations. *N*-methyl-L-phenylalanine (rabbit reticulocyte), *N*-methylglycine (*E. coli*), and *N*-methyl-L-alanine (*E. coli*) are site-specifically incorporated into proteins, whereas *N*-ethyl-L-alanine is not (28). The *E. coli* system accepts pipecolinic acid, the six-membered version of proline, but will not incorporate the ring-contracted proline analogues aziridine-2-carboxylic acid and azetidine-2-carboxylic acid (this exclusion may stem from chemical instability rather than from discrimination by the biosynthesis machinery). However, other investigators (36; DA Tirrell & TJ Deming, personal communication) have reported the successful biosynthetic incorporation of L-azetidine-2-carboxylic acid into polypeptides. Other types of backbone replacements for α-amino acids that have been examined include β- and γ-amino acids such as β-alanine, γ-aminobutyric acid, and longer variants (D Mendel, JA Ellman & PG Schultz, unpublished results) as well as dipeptides and dipeptide mimetics (5). To date, attempts to incorporate these amino acids using either *E. coli* or rabbit reticulocyte lysate systems have proven unsuccessful.

The precise structure of an amino acid also influences its suppression efficiency dramatically. For example, even though the spin labels in Figure 10 (analogues 24 and 25, see Figure 10, below) both contain the 1-oxyl-2,2,5,5-tetramethylpyrroline *N*-oxide group, only analogue 24 could be incorporated in vitro (Figure 10, below) (20). In addition, suppression efficiencies may vary with the nature of the in vitro protein-synthesis system (prokaryotic vs eukaryotic). Both systems prohibit D-amino acid integration into peptide chains, and both systems readily accept α-hydroxy acids and *N*-methylamino acids. However, significant differences exist between the two systems for specific amino acids: L-2-amino-3,3-dimethylbutanoic acid is efficiently incorporated in the *E. coli* system but functions poorly in a rabbit reticulocyte system (5, 68). The situation is reversed for 2-amino-4-phosphonobutyric acid, a nonhydrolyzable phosphoserine analogue that is well tolerated in the rabbit reticulocyte system but that failed completely when tested with the *E. coli* extract (5; JA Ellman & PG Schultz, unpublished results). Finally, in contrast to in vivo results (9, 74), there appears to be little correlation between the codons adjacent to the UAG codon and suppression efficiencies (20). As the range of unnatural amino acids that are substituted into proteins expands, this methodology should provide additional insights into the specificity of the translational machinery and the factors governing in vitro suppression efficiency.

Future work will focus on broadening the scope of the methodology by increasing protein yields, by minimizing the effect of the amino acid identity on suppression efficiency, and by developing ways to introduce more than one unnatural amino acid into a protein at a time. Efforts to improve protein yields will require improvements in in vitro protein synthesis. Higher protein yields may be achieved using in vitro systems derived from other sources (38) or alternative approaches, such as continuous flow in vitro systems (103). Alternatively, because the *E. coli* transcription-translation system is active at 37°C for only ~1 hour, protein expression levels may be increased by using other lysates, such as lysed spheroplasts, capable of supporting protein synthesis for longer periods (61, 78).

Suppression efficiencies may be improved by (*a*) altering nucleotides in the tRNA molecule involved in tRNA recognition by the elongation factors, the ribosome, or the aminoacyl-tRNA synthetases; (*b*) utilizing suppressor tRNAs from different species; (*c*) deleting release factor-1 (*E. coli*), which competes with the suppressor tRNA for recognition of the UAG codon, from the in vitro extract (55); or (*d*) modifying the specificity of EF-Tu. The production of proteins containing several unnatural amino acids is another goal. A single unnatural amino acid

can be incorporated at more than one position in a protein simply by introducing several UAG codons in the coding sequence (VW Cornish & PG Schultz, unpublished results). Preliminary efforts to insert different unnatural amino acids into the same protein have focused on using the other nonsense codons (D Mendel & PG Schultz, unpublished results), infrequently occurring codons (in the latter approach the tRNA must be engineered to eliminate recognition by the corresponding aminoacyl-tRNA synthetase) (18), or a sixty-fifth codon composed of nonstandard nucleotides (4).

Finally, in an exciting recent development in the field, the stop codon–suppression methodology has been extended to the incorporation of unnatural amino acids into intact eukaryotic cells (81). A *Xenopus laevis* oocyte was coinjected with two mutant RNA species: (*a*) mRNA that was synthesized in vitro from a mutated cDNA clone and that contained the stop codon UAG at the residue of interest and (*b*) a suppressor tRNA that had been synthetically acylated with a series of unnatural amino acids. During translation by the oocyte's protein synthetic machinery, the unnatural amino acids were specifically incorporated at position 198 in the binding site of the nicotinic acetylcholine receptor α-subunit. Subtle changes in the structures of the amino acids resulted in readily detectable changes in the function of this ligand-gated ion channel as determined by electrophysiological analysis. This work suggests that unnatural amino acid mutagenesis may be applied to many questions concerning the structural and functional aspects of ion channels, receptors, and transporters.

APPLICATIONS

This section discusses several applications of the methodology in order to illustrate its scope as well as questions that can be addressed when unnatural amino acids are incorporated into proteins. These examples include studies of protein stability, enzyme mechanism, and signal transduction; the construction of photocaged proteins; and the site-specific incorporation of biophysical probes into proteins.

Protein Stability

Site-directed mutagenesis has contributed significantly to our understanding of the factors that determine the stability of proteins. However, because few structurally and electronically conservative changes can be made using the 20 natural amino acids, it is difficult to break down the energetic effects of a given mutation in terms of the contributions from van der Waals interactions, conformational entropy, hydro-

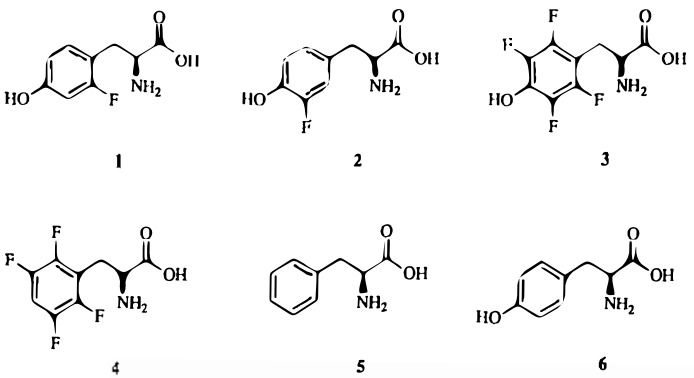
gen bonds, and other forces. The use of unnatural amino acids should allow us to make more defined mutations in proteins, thereby gaining a more precise picture of the forces that govern protein stability.

An outstanding issue in the protein stability and folding literature is the role of hydrogen bonding. Although hydrogen bonds clearly play an important role in determining the secondary and tertiary structure of proteins (86, 87), the magnitude of their contribution to protein stability has been difficult to assess (46, 95). Values of hydrogen-bond stabilization determined from mutational studies, in which one member of a hydrogen-bonded pair is deleted, vary considerably depending upon the nature and local environment of the hydrogen bond (25, 32). For example, two nondisruptive hydrogen-bond deletions (Tyr78 → Phe and Ser91 → Ala) in barnase each account for 1.4–1.9 kcal mol^{-1} in protein stabilization (33). On the other hand, deletion of a hydrogen bond in the active site of barnase had no effect on protein stability (67). As pointed out above, a concern in all such mutational studies is the possibility of introducing additional destabilizing interactions by leaving unfilled hydrogen-bond donors or acceptors and/or altering local solvation and packing interactions.

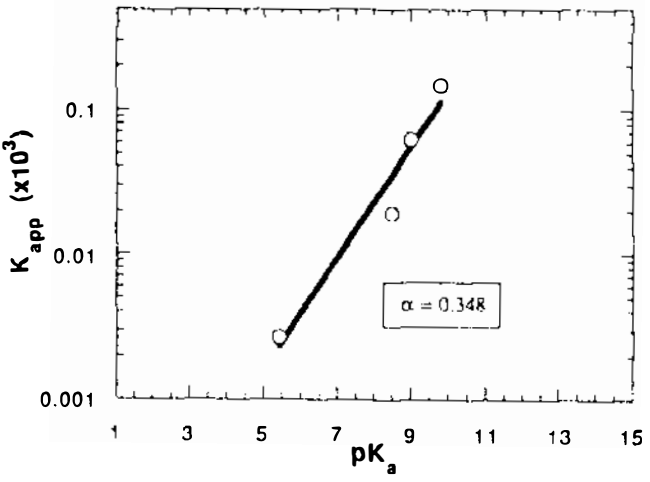
In an effort to determine the degree to which side-chain hydrogen bonding stabilizes the folded state of proteins while minimizing the steric and electronic perturbations associated with deleting one member of a hydrogen-bonded pair, Tyr27 in staphylococcal nuclease was substituted with each of several isosteric, fluorinated tyrosine analogues, including 2-fluorotyrosine (analogue 1), 3-fluorotyrosine (analogue 2), and tetrafluorotyrosine (analogue 3), as well as 2,3,5,6-tetrafluorophenylalanine (analogue 4) and phenylalanine (analogue 5) (Figure 4*a*) (105). These mutants were designed in order to examine the effect of increasing the strength of the Tyr27-Glu10 hydrogen bond on protein stability. Denaturation studies of the corresponding mutants revealed a unique free-energy correlation between $\Delta\Delta G_{H_2O}$ ($\alpha = 0.35$) and the pK_a of the Tyr27 hydroxyl group (Figure 4*b*). Indeed, substitution of Tyr ($pK_a \approx 10$) with tetrafluoro-L-tyrosine ($pK_a \approx 5$) increased the stability of staphylococcal nuclease (SNase) by 2.3 kcal mol^{-1}. This experiment provides strong evidence that side-chain intramolecular hydrogen bonds can preferentially stabilize the folded state of a protein relative to the unfolded state in water.

The role of backbone hydrogen bonds in determining the stability of the folded state of a protein was also examined (25). The surface amino acid Ala82 in T4L was replaced with the isostere lactic acid (Figure 4*c*) (28). Ala82 resides at a break between two helices where its NH group and side chain are exposed to water. The Ala82 → lactate muta-

a)



b)



c)

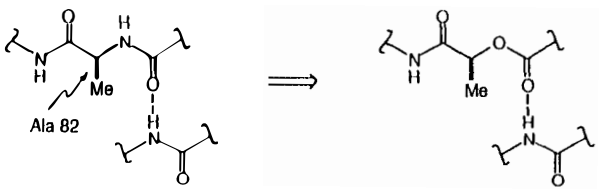


Figure 4 (*a*) Tyrosine and phenylalanine analogues incorporated at site Tyr27 in SNase. The numbered molecules in all figures are referred to as "analogue *n*" in the text. (*b*) Free energy relationship between the pK_a of the site 27 side chain and the stability of SNase. (*c*) Substitution of Ala82 in T4 lysozyme (T4L) with lactic acid.

tion effectively replaces a good hydrogen-bond acceptor, the amide carbonyl group, with a considerably weaker hydrogen-bond acceptor, the ester carbonyl group (51). Because both esters and amides usually adopt the *trans* conformation (110), the 3.7°C (1.0 kcal mol^{-1}) decrease in melting temperature (stability) caused by this substitution again likely reflects the stabilization energy associated with variations in hydrogen-bond strength.

Site-directed mutagenesis studies have also contributed much to our understanding of the role the hydrophobic effect plays in protein stabilization (25, 66). However, the limited substitutions that can be made with the common amino acids quite often lead to changes in factors other than hydrophobicity. For example, in T4 lysozyme (T4L), a relatively large cavity extends beyond the side chains of Leu133 and Ala129 into the core of the enzyme (52). In an attempt to increase hydrophobic-core packing density and, as a consequence, thermal stability, proteins containing the mutations Leu133 → Phe and Ala129 → Val were constructed (52). In both cases, the mutant enzymes were less thermally stable than the wild-type enzyme because of unfavorable packing interactions that increased local strain energy. In contrast, mutagenesis with nonproteiogenic amino acids successfully extended the aliphatic side chain of Leu133 into the T4 lysozyme core without increasing strain energy, thereby affording mutants significantly more stable than wild-type (69). Specifically, when Leu133 was replaced with S,S-2-amino-4-methylhexanoic acid (analogue 7) and S-2-amino-3-cyclopentylpropanoic acid (analogue 8) (Figure 5), the mutant T4Ls melted at temperatures 1.9°C (0.6 kcal mol^{-1} more stable) and 4.3°C (1.2 kcal mol^{-1} more stable) higher, respectively, than wild-type. Interestingly, the surface areas of these two amino acids differ by only ~1 $Å^2$, demonstrating that burial of hydrophobic surface area was not the sole factor contrib-

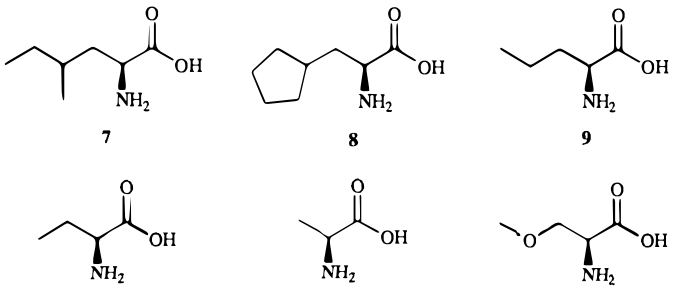


Figure 5 Structure of amino acids substituted for Leu133 in T4L.

uting to increased stability. Relative to the unfolded state, the cyclic variant loses less conformational entropy during folding than does the open chain variant, thus leading to a more stably folded protein.

A complement to experiments aimed at increasing packing density involved systematic pruning of the Leu133 side chain to determine the effect of enlarging the cavity within T4L. Substitution of S-2-aminopentanoic acid (analogue 9), S-2-aminobutyric acid (analogue 10), or alanine (analogue 11) for Leu133 (Figure 5) resulted in a nonlinear relationship between hydrophobic side-chain length and enzyme stability, in which the enzyme became increasingly unstable with removal of each methylene unit from its core (69). Semiquantitative estimates of $\Delta\Delta G$ values agreed well with the experimental values and suggest that stabilization results not only from hydrophobic effects, but also from entropic-, packing-, and cavity-energy terms.

Mutagenesis with nonproteiogenic amino acids has also been used to examine the effects of side-chain solvation on protein stability. We found that T4L containing S-2-aminopentanoic acid-133 (analogue 9) was more stable than the *O*-methylserine-133 (analogue 12) mutant by 1.7 kcal mol^{-1} (Figure 5) (69). This stabilization can be ascribed almost totally to the less polar nature of norvaline compared with *O*-methylserine because such isostructural replacements are unlikely to have significantly different packing interactions or side-chain conformational entropies. Similarly, the difference in the octanol-water partitioning ratios obtained with the *N*-acetyl amide derivatives of S-2-aminopentanoic acid and *O*-methylserine favored S-2-aminopentanoic acid in octanol by 1.8 kcal mol^{-1} and provided further evidence that octanol-water partitioning values accurately reflect solvation effects on protein stability (69).

The effect of β-branched amino acids on protein stability was examined by comparing mutant proteins containing alanine (analogue 11), L-2-aminobutanoic acid (analogue 10), L-2-aminopentanoic acid (analogue 9), valine (analogue 13), L-2-aminohexanoic acid (analogue 14), and L-2-amino-3,3-dimethylbutanoic acid (analogue 15) at two surface-exposed sites in the middle of two α-helices in T4L (Figure 6) (21). The natural β-branched amino acids valine, isoleucine, and threonine are thought to destabilize α-helices. These amino acids occur infrequently relative to amino acids such as leucine in α-helices in known protein structures, and they destabilize α-helical peptides and protein α-helices (7, 100). For example, when a series of linear and branched amino acids were incorporated into an α-helical peptide, the β-branched amino acids were destabilizing relative to the linear and γ-branched amino acids (65). Substitution of L-2-amino-3,3-dimethylbutanoic acid for L-2-aminohex-

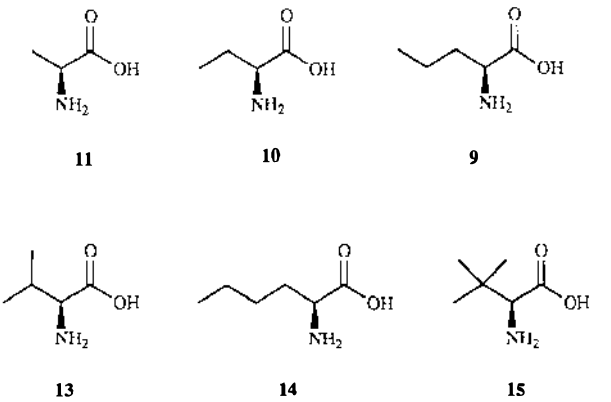


Figure 6 Examples of molecules designed to probe the effects of β-substitution on helix stability in a protein.

anoic acid or alanine destabilizes the peptide by 0.9 kcal mol^{-1}. The same substitution, however, either destabilizes T4L by 0.69 kcal mol^{-1} (2.5°C) at site Ser44 or stabilizes the protein by 0.27 kcal mol^{-1} (1.0°C) at site Asn68. This difference illustrates the difficulty of delineating simple rules as to which factors stabilize or destabilize proteins given the influence of context effects. In addition, the stabilities and simulated structures of the L-2-amino-3,3-dimethylbutanoic acid mutants have provided insight into the effects of β-branched side chains on α-helix stability in proteins (69). These results underscore the difficulty in interpreting mutagenesis data when more than one property of an amino acid is being altered at once. The ability to precisely alter the structure of an amino acid should allow us to better dissect the individual contributions of hydrophobicity, packing, entropy, and cavity formation on protein stability. We are currently extending these studies to hydrogen-bonding, π-π, and π-ion interactions in protein interiors.

Enzyme Mechanism

Unnatural amino acid mutagenesis also is being used to probe the catalytic mechanisms of several enzymes including staphylococcal nuclease, aspartate aminotransferase, methionine aminopeptidase, and ribonucleotide reductase. The enzyme staphylococcal nuclease (SNase) accelerates the hydrolysis of phosphodiester bonds in nucleic acids some 10^{16}-fold over the uncatalyzed rate. This enzyme has been the subject of many structural, mechanistic, and mutagenesis studies aimed at understanding how enzymes can achieve such extraordinary rate enhancements (63, 108, 109). Based on these studies, it has been

suggested that general base catalysis contributes significantly to the catalytic efficiency of this enzyme. Specifically, Glu43 in SNase is thought to act as a general base to activate a water molecule for attack on the phosphodiester backbone of DNA. Glu43 is known to be important for catalysis because replacement by the natural amino acids Asp and Gln results in a significant loss in activity (43).

Surprisingly, substitution of Glu43 with either homoglutamate (analogue 18) or the nitro analogue S-4-nitro-2-aminobutyric acid (analogue 17), which is both isoelectronic and isosteric to glutamate (analogue 16) but a much poorer base, yielded mutant enzymes with kinetic constants markedly similar to those of wild-type SNase under normal assay conditions (Figure 7) (47). The catalytic efficiencies of these mutants, coupled with their pH behavior and the crystal structure of the HGlu43 mutant, suggest that Glu43 may not be acting as a base but rather may play a structural role, serving as a bidentate hydrogen-bond acceptor to fix the conformation of the neighboring loop. Independent studies by Gerlt, in which the loop adjacent to Glu43 was deleted from SNase, have led to the same conclusion (39). Homoglutamate replacements would be of interest in other proteins as well. Conventional mutagenesis on triose phosphate isomerase (TIM) revealed that the Glu165 → Asp mutation dramatically decreases the catalytic efficiency of TIM (56).

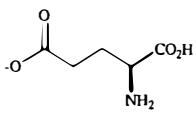
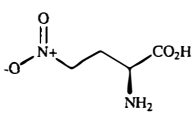
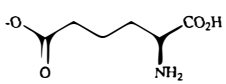
Amino Acid	Vmax (A_{260} min^{-1} μg^{-1})	K_M^{DNA} (μg/mL)	K_M^{Ca2+} (μM)
 **16**	6.7 ± 0.7	8 ± 3	320 ± 30
 **17**	3.0 ± 0.5	26 ± 8	470 ± 40
 **18**	5.2 ± 0.2	10 ± 2	290 ± 30

Figure 7 Catalytic efficiencies of wild-type SNase and the mutant enzymes Glu43 → S-2-amino-4-nitrobutanoic acid and Glu43 → homoglutamate.

It would be of interest to ask how the Glu165 → HGlu replacement would affect the catalytic power of TIM.

As with staphylococcal nuclease, we also asked how varying the pK_a of active site residues affected T4L activity. γ-Fluoroglutamic (FGlu) acid has a pK_a about 2–3 units lower than glutamic acid. Not only did the Glu11 → FGlu T4L retain a high degree of enzyme activity, but also the pH optimum for cell wall hydrolysis was not significantly shifted (D Mendel, JA Ellman & PG Schultz, unpublished results). These results suggest that Glu11 is active in the fully deprotonated state in wild-type T4L, possibly providing electrostatic stabilization to the transition state. The glutamate esters δ-methylglutamate and δ-methyl-γ-fluoroglutamate were also substituted for Glu11 in T4L, but both yielded inactive enzyme (D Mendel, JA Ellman & PG Schultz, unpublished results).

Signal Transduction

Unnatural amino acid mutagenesis has been used to probe the role of ras p21 in cellular signal-transduction pathways. Mammalian proteins encoded by the *ras* genes are thought to function as regulators of various signal-transduction processes involved in cell growth and differentiation (6, 8, 102). The chemical basis for signal regulation involves cycling of the protein between the inactive guanosine diphosphate (GDP)–bound state and the active guanosine triphosphate (GTP)–bound state. Point mutations that result in a decrease in the intrinsic GTPase activity of ras or of the GTPase-activating protein (GAP)–stimulated GTPase activity are associated with approximately 30% of human cancers (8, 107). In order to better understand the molecular basis by which mutations in ras lead to switch inactivation, we have substituted a series of unnatural amino acids (Figure 8) for residues in

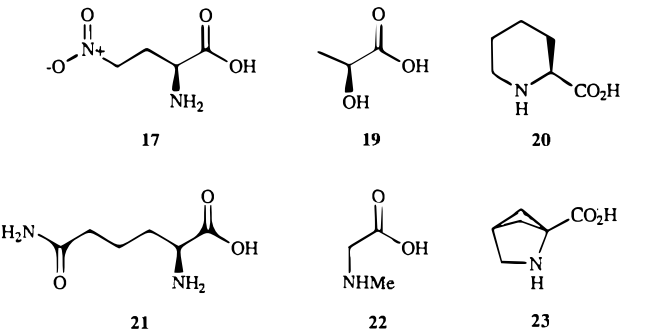


Figure 8 Structures of some unnatural amino acids used to probe the switch function of ras.

loop L4 (the switch II region), loop L2 (the switch I region), and loop L1 (the phosphate-binding loop) (15–17).

Mutations at Gln61 commonly lead to impaired intrinsic GTPase activity. This residue lies in loop L4 of ras, a region that undergoes conformational change upon GTP-GDP exchange. Biochemical studies and the elucidation of the three-dimensional X-ray crystal structures of the GTP-, Gpp(CH_2)p-, and Gpp(NH)p-bound forms of ras have led to the proposal that Gln61 is critical for γ-phosphate binding and catalysis (58, 73, 83, 90). Specifically, it has been proposed that the γ-carboxamide of Gln61 either polarizes water molecule 175 for attack on the γ-phosphate or stabilizes the incipient pentacoordinate transition state. To test these mechanistic hypotheses, we replaced Gln61 with the isoelectronic, isosteric nitro analogue (analogue 17) and with homoglutamine (analogue 21) (Figure 8) (17). Whereas 17 natural mutants at position 61 have reduced GTPase activity and are not activated by GAP, both Glu → analogue 17 and Glu → analogue 21 mutants had GTPase activity similar to that of the wild-type protein and were activated by GAP. The fact that these mutants retained activity suggests that residue 61 plays some other role in catalysis, perhaps serving to distort the bound GTP toward the transition-state geometry.

Mutations at Gly12 of ras also result in impaired intrinsic GTPase activity and are commonly associated with oncogenic activation (8, 101). Gly12 occurs in a highly conserved type II β-turn, a phosphate-binding loop found in many nucleotide-binding proteins (34, 111). Mutation of Gly12 to any common amino acid other than proline results in diminished GTPase activity (101). To more fully understand the role of Gly12 in switch function, several residues with backbone modifications were inserted at this site, including lactic acid (Lac) (analogue 19), pipecolic acid (Pip) (analogue 20), and *N*-methylglycine (MeGly) (analogue 22) (Figure 8) (17). The pipecolic acid mutant, which was expected to have a much more negative ϕ value than Gly12, retained GTPase activity similar to that of wild-type ras. The *N*-methylglycine mutant also had wild-type GTPase activity (Nα methyl substitution also decreases ϕ values). Whereas the Ala12 mutant had reduced GTPase activity, the isosteric lactic acid mutant, in which the backbone amide linkage is replaced by an ester linkage, had normal GTPase activity. The fact that only mutants that can adopt unusual backbone conformations are active suggests that a specific conformation of the loop L1 backbone may be required to avoid unfavorable side-chain interactions in the transition state for GTP hydrolysis. However, even though the Pro12, Pip12, Lac12, and MeGly12 mutants had intrinsic GTPase activity similar to that of wild-type ras, they were not activated by GAP.

Moreover, the Gly13 → Thr, Gly13 → *allo* Thr, and Gly13 → Ser mutants, which have intrinsic GTPase activity two to three times that of wild-type ras (attributable to the β-hydroxyl group), were not activated by GAP and were not transforming in a germinal vesicle breakdown assay with *X. laevis* oocytes. One interpretation of these results is that loop L1 can adopt two or more conformations in solution and that its positioning by one or more effector proteins modulates GTPase activity and oncogenic activation. The unusual backbone structures of the Gly12 and Gly13 mutants may prevent switch function.

The role of Pro34 in loop L2 function was also probed by substituting amino acids containing novel backbone structures for Pro34. The most significant difference between the GTP- and GDP-bound forms of ras is in the region encompassing residues 32–40 (58, 73, 84, 90). It has been proposed that Pro34, which is conserved in ras and is close to the active site, may play a role in controlling the conformation of loop L2, perhaps via a *cis-trans* isomerization of the Pro34 amide bond (84). To examine the structural and mechanistic role of this residue more precisely, Pro34 was replaced with 2,4-methanoproline (analogue 23) (Figure 8), which is strongly biased toward the *trans* configuration by virtue of the Cα substitution (15). The fact that this mutant has wild-type intrinsic and GAP-activated GTPase activity strongly suggests that a *cis-trans* isomerization of Pro34 does not play a key role in signal transduction. Similar substitutions may be useful in probing the role of backbone isomerization in protein-folding pathways.

Novel Functional Substitutions

Unnatural amino acid mutagenesis also allows the substitution of amino acids possessing novel spectroscopic properties including fluorescent, photoactivatable, or EPR- or NMR-active amino acid replacements.

PHOTOACTIVE RESIDUES Photoactivatable, or caged molecules, possess inactivating protecting groups that can be rapidly removed (<1 ms) to generate the active form of an agonist, substrate, or inhibitor. Caged molecules provide a noninvasive method for controlling a variety of biochemical and cellular processes, including intracellular second-messenger concentrations (85, 98). In the past, such studies have usually used low-molecular-weight caged compounds. Unnatural amino acid mutagenesis simplifies the construction of caged macromolecules bearing one or more blocked side-chain groups (71). For example, active-site residue Asp20, which is essential for catalytic activity in T4 lysozyme, when masked as the β-nitrobenzyl ester (NBAsp) in T4L yielded an enzyme that could be photoactivated. Although the nitrobenzyl ester

was incorporated with much better efficiency than native aspartic acid, enzymes containing the ester were completely inactive. Photolysis at 350 nm rapidly converted the ester to its parent acid, resulting in full enzymatic activity. In a second example, a serine residue thought to be involved in the self-splicing reaction of the protein vent DNA polymerase was replaced by *O*-nitrobenzylserine (Figure 9) (18a). The resulting caged polymerase did not undergo the self-splicing reaction, allowing the full-length protein to be isolated. Photolysis of the unspliced protein resulted in loss of the nitrobenzyl protecting group and subsequent protein splicing. This result demonstrates unequivocally

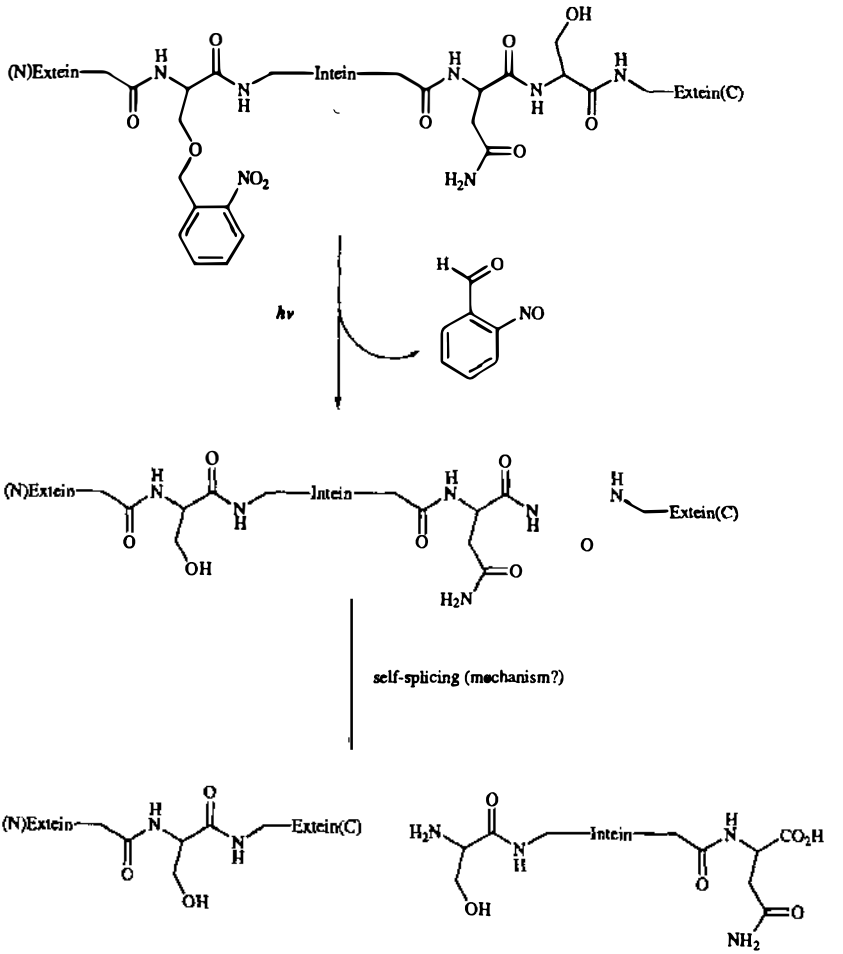


Figure 9 Photactivated self-splicing of vent polymerase using a caged serine analogue.

that protein splicing occurs at the posttranslational level. The ability to construct caged proteins of this sort should make possible a broad range of time-resolved experiments relevant to catalytic mechanism, biomolecular recognition, and protein folding.

Beyond simply masking native side chains, photoactivatable side chains can be used to cross-link proteins with unknown target molecules. Thus far, amino acids that form nitrene, carbene, and ketyl radical active species have been successfully incorporated into p21ras and T4L at specific positions (20; JA Ellman, DR Benson & PG Schultz, unpublished results). Photoaffinity probes can in principle be used to identify key components involved in complex biochemical processes, e.g. signal transduction or gene activation.

REPORTER GROUPS Biophysical probes have been incorporated into proteins using unnatural amino acid mutagenesis. For example, spin labels and fluorescent amino acids have been site-specifically incorporated into T4L (Figure 10) (20). Taking advantage of recent loop-gap resonator technology, we (20) measured the ESR spectrum of approximately 10 μl of a 1 μM solution of purified T4L containing a spin-labeled amino acid (analogue 24) at position 44. Linewidth analysis of the ESR spectrum revealed that the label was immobilized, consistent with it being protein-bound. Replacement of one of the tryptophans in T4L by 7-azatryptophan (analogue 26) resulted in a red shift in T4L's fluorescent emission maxima. The red shift, however, was not as large as seen in model systems in which water was the solvent, not a protein.

Unnatural amino acid mutagenesis has been employed to site-specifically insert isotopically enriched amino acids into proteins for NMR

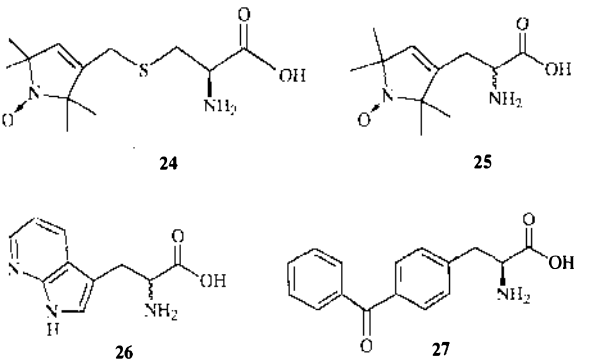


Figure 10 Structure of some biophysical probes tested for incorporation in T4L.

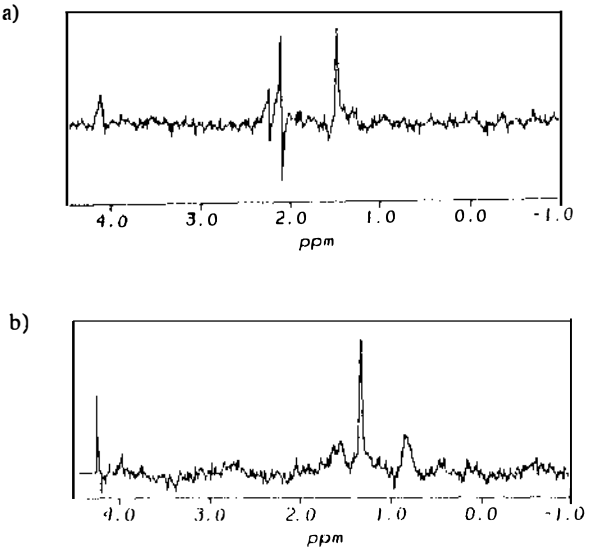


Figure 11 (*a*) ^{13}C-filtered ^{1}H NMR spectrum of the T4L mutant Ala82 → [^{13}C]Ala. (*b*) ^{13}C-filtered ^{1}H NMR spectrum of the denatured form of the same T4L mutant.

studies. A sufficient quantity of purified T4L specifically labeled at site Ala82 with [^{13}C]Ala was produced to observe the proton resonances of the selectively labeled protein, in both the native and denatured states, using ^{13}C-filtered NMR (Figure 11) (30). The ability to site-specifically incorporate probes capable of detecting local structure and dynamics into proteins should allow for more precise studies of protein folding and stability; conformational changes in proteins; and the interaction of proteins with other proteins, small molecules, and membranes. For example, it should be possible to carry out spectroscopy on both the folded and unfolded states of proteins, compare denaturation curves for various regions of a protein, or probe local structural changes that occur during protein translocation.

CONCLUSION

The ability to incorporate unnatural amino acids into proteins site-specifically makes it possible to carry out detailed physical organic studies on this important class of molecules. Mutations can be made to probe the precise nature of an individual amino acid's contribution to protein structure or function, often without dramatic losses in activity. This methodology should provide new insights into and perhaps challenge

assumptions about the nature of protein structure and function. In addition, it may be possible to generate mutant proteins with novel functions. What remains is to increase the protein quantities obtainable and to find ways to bypass the specificity of the protein-biosynthetic machinery so that more diverse amino acid structures can be incorporated to add to the power of this approach.

ACKNOWLEDGMENTS

We are grateful for financial support for this work from the National Institutes of Health (Grant No. R01 GM49220) and the Director, Office of Energy Research, Division of Material Sciences and Division of Energy Biosciences, Office of General Life Sciences, Structural Biology Division, of the US Department of Energy under Contract No. DE-AC03-76SF00098. VWC thanks the National Science Foundation for a predoctoral fellowship.

Literature Cited

1. Adams SR, Kao JPY, Tsien RY. 1989. Biologically useful chelators that take up Ca^{2+} upon illumination. *J. Am. Chem. Soc.* 111:7957–68
2. Ayer D, Yarus M. 1986. The context effect does not require a fourth base pair. *Science* 231:393–95
3. Bain JD, Glabe CG, Dix TA, Chamberlin AR, Diala ES. 1989. Biosynthetic site-specific incorporation of a non-natural amino acid into a polypeptide. *J. Am. Chem. Soc.* 111: 8013–14
4. Bain JD, Switzer C, Chamberlin A, Benner S. 1992. Ribosome-mediated incorporation of non-standard amino acids into a peptide through expansion of the genetic code. *Nature* 356: 537–39
5. Bain JD, Wacker DA, Kuo EE, Chamberlin AR. 1991. Site-specific incorporation of non-natural residues into peptides: effect of residue structure on suppression and translational efficiencies. *Tetrahedron* 47: 2389–2400
6. Barbacid M. 1987. *ras* genes. *Annu. Rev. Biochem.* 56:779–827
7. Blaber M, Zhang XJ, Lindstrom JD, Deprot SD, Baase WA, Matthews BW. 1994. Determination of α-helix propensity within the context of a folded protein. *J. Mol. Biol.* 235: 600–24
8. Bos JL. 1989. *ras* oncogenes in human cancer: a review. *Cancer Res.* 49:4682–89
9. Bossi L. 1983. Translation of UAG codon by suppressor tRNA is affected by the sequence following UAG in the message. *J. Mol. Biol.* 164:73–87
10. Bruce A, Atkins J, Wills N, Uhlenbeck O, Gesteland R. 1982. Replacement of anticodon loop nucleotides to produce functional tRNAs: amber suppressors derived from yeast $tRNA^{Phe}$. *Proc. Natl. Acad. Sci. USA* 79:7127–31
11. Brunner J. 1993. New photolabeling and crosslinking methods. *Annu. Rev. Biochem.* 62:483–514
12. Chaiken IM. 1981. Semisynthetic peptides and proteins. *CRC Crit. Rev. Biochem.* 11:255–301
13. Chinsky N, Margolin AL, Klibanov AM. 1989. Chemoselective enzymatic monoacylation of bifunctional compounds. *J. Am. Chem. Soc.* 111: 386–88
14. Chladek S, Sprinzl M. 1985. The 3′-end of tRNA and its role in protein

biosynthesis. *Angew. Chem. Int. Ed. Engl.* 24:371–91
15. Chung HH, Benson DR, Cornish VW, Schultz PG. 1993. Probing the role of loop 2 in ras function with unnatural amino acids. *Proc. Natl. Acad. Sci. USA* 90:10145–49
16. Chung HH, Benson DR, Schultz PG. 1993. Probing the role of lysine 16 in ras p^{21} protein with unnatural amino acids. *J. Am. Chem. Soc.* 115: 6414–15
17. Chung HH, Benson DR, Schultz PG. 1993. Probing the structure and mechanism of ras protein with an expanded genetic code. *Science* 259: 806–9
18. Chunhua M, Kudlicki W, Odom OW, Kramer G, Hardesty B. 1993. In vitro protein engineering using synthetic $tRNA^{Ala}$ with different anticodons. *Biochemistry* 32:7937–45
18a. Cook SN, Jack WE, Xiong X, Danley LE, Ellman JA, et al. 1995. Photochemically-initiated protein splicing. *Cell.* In press
19. Corey DR, Schultz PG. 1987. Generation of a hybrid sequence-specific single-stranded deoxyribonuclease. *Science* 233:1401–3
20. Cornish VW, Benson DR, Altenbach CA, Hideg K, Hubbell WL, Schultz PG. 1994. Site-specific incorporation of biophysical probes into proteins. *Proc. Natl. Acad. Sci. USA* 91: 2910–14
21. Cornish VW, Kaplan MI, Veenstra D, Kollman PA, Schultz PG. 1994. Destabilizing and stabilizing effects of placing β-branched amino acids in protein α-helices. *Biochemistry* 33: 120221–31
22. Cornish VW, Mendel D, Schultz PG. 1995. Probing protein structure and function with an expanded genetic code. *Angew. Chem. Int. Ed. Engl.* In press
23. Crick FHC, Barrett L, Brenner S, Watts-Tobin R. 1961. General nature of the genetic code for proteins. *Nature* 192:1227–32
24. Dawson PE, Muir TW, Clark-Lewis I, Kent SBH. 1994. Synthesis of proteins by native chemical ligation. *Science* 266:776–79
25. Dill KA. 1990. Dominant forces in protein folding. *Biochemistry* 29: 7133–55
26. Deleted in proof
27. Ellman JA, Mendel D, Anthony-Cahill S, Noren CJ, Schultz PG. 1992. Biosynthetic method for introducing unnatural amino acids site-specifically into proteins. *Methods Enzymol.* 202:301–36
28. Ellman JA, Mendel D, Schultz PG. 1992. Site-specific incorporation of novel backbone structures into proteins. *Science* 255:197–200
29. Deleted in proof
30. Ellman JA, Volkman BF, Mendel D, Schultz PG, Wemmer DE. 1992. Site-specific isotopic labeling of proteins for NMR studies. *J. Am. Chem. Soc.* 114:7959–61
31. Fahnestock S, Neumann H, Shashoua V, Rich A. 1970. Ribosome-catalyzed ester formation. *Biochemistry* 9:2477–83
32. Fersht AR. 1987. The hydrogen bond in molecular recognition. *Trends Biochem. Sci.* 12:301–4
33. Fersht AR, Shi J-P, Knill-Jones J, Lowe D, Wilkinson A, et al. 1985. Hydrogen bonding and biological specificity analysed by protein engineering. *Nature* 314:235–38
34. Gay NJ, Walker JE. 1983. Homology between human bladder carcinoma oncogene product and mitochondrial ATP-synthase. *Nature* 301:262–64
35. Giege R, Puglisi JD, Florentz C. 1993. tRNA structure and aminoacylation efficiency. *Progr. Nucleic Acids Res. Mol. Biol.* 45:129–206
36. Grant MM, Brown AS, Corwin LM, Troxler RF, Franzblau C. 1975. Effect of L-azetidine 2-carboxylic acid on growth and proline metabolism in *Escherichia coli. Biochim. Biophys. Acta* 404:180–87
37. Griffith MC. 1990. *Chemical aminoacylation of tRNA's with unnatural amino acids.* PhD thesis. Univ. Calif., Berkeley
38. Haines BD, Higgins SJ. 1984. *Transcription and Translation: A Practical Approach.* Oxford: IRL
39. Hale SP, Poole LB, Gerlt JA. 1993. Mechanism of the reaction catalyzed by staphylococcal nuclease: identification of the rate-determining step. *Biochemistry* 32:7479–87
40. Harrington KM, Nazarenko IA, Dix DB, Thompson RC, Uhlenbeck OC. 1993. In vitro analysis of translational rate and accuracy with an unmodified tRNA. *Biochemistry* 32:7617–22
41. Hecht SM. 1992. Probing the synthetic capabilities of a center of biochemical catalysis. *Acc. Chem. Res.* 25:545–52
42. Heckler TG, Chang LH, Zama Y, Naka T, Chorghade MS, Hecht SM. 1984. T4 RNA ligase mediated preparation of novel "chemically misacy-

lated" tRNA^{Phe}s. *Biochemistry* 23: 1468–73
43. Hibler DW, Stolowich NJ, Reynolds MA, Gerlt JA, Wilde JA, Bolton PH. 1987. Site-directed mutants of staphylococcal nuclease. Detection and localization by ^{1}H NMR spectroscopy of conformational changes accompanying substitutions for glutamic acid-43. *Biochemistry* 26:6278–86
44. Hofmann K, Bohn H. 1966. Studies on polypeptides. XXXVI. The effect of pyrazole-imidazole replacements on the S-protein activating potency of an S-peptide fragment. *J. Am. Chem. Soc.* 88:5914–19
45. Jackson DY, Burnier J, Quan C, Stanley M, Tom J, Wells JA. 1994. A designed peptide ligase for total synthesis of Ribonuclease A with unnatural catalytic residues. *Science* 266: 243–47
46. Jeffrey GA, Saenger W. 1991. *Hydrogen Bonding in Biological Molecules.* New York: Springer-Verlag
47. Judice JK, Gamble TR, Murphy EC, deVos AM, Schultz PG. 1993. Probing the mechanism of staphylococcal nuclease with unnatural amino acids: kinetic and structural studies. *Science* 261:1578–81
48. Kaiser ET. 1989. Synthetic approaches to biologically active peptides and proteins including enzymes. *Acc. Chem. Res.* 22:47–54
49. Kaiser ET, Lawrence DS. 1984. Chemical mutation of enzyme active sites. *Science* 226:505–11
50. Kaiser ET, Lawrence DS, Rokita SE. 1985. The chemical modification of enzymatic specificity. *Annu. Rev. Biochem.* 54:565–95
51. Kamlet MJ, Abboud JLM, Abraham MH, Taft RW. 1983. Linear solvation energy relationships. 23. A comprehensive collection of the solvatochromic parameters, π^*, α, and β, and some methods for simplifying the generalized solvatochromic equations. *J. Org. Chem.* 48:2877–87
52. Karpusas M, Baase WA, Matsumura M, Matthews BW. 1989. Hydrophobic packing in T4 lysozyme probed by cavity-filling mutants. *Proc. Natl. Acad. Sci. USA* 86:8237–41
53. Kast P, Hennecke H. 1991. Amino acid substrate specificity of *Escherichia coli* phenylalanyl-tRNA synthetase altered by distinct mutations. *J. Mol. Biol.* 222:99–124
54. Kern D, Giege R, Ebel JP. 1972. Incorrect aminoacylations catalysed by the phenylalanyl- and valyl-tRNA synthetases from yeast. *Eur. J. Biochem.* 31:148–55
55. Kleina LG, Masson JM, Normanly J, Abelson J, Miller JH. 1990. Construction of *E. coli* amber suppressor tRNA genes. II. Synthesis of additional tRNA genes and improvement of suppressor efficiency. *J. Mol. Biol.* 213:705–18
56. Knowles J. 1991. Enzyme catalysis: not different, just better. *Nature* 350: 121–24
57. Knowles JR. 1987. Tinkering with enzymes: what are we learning? *Science* 236:1252–58
58. Krengel U, Schlichting I, Scherer A, Schumann R, Frech M, et al. 1990. Three-dimensional structures of H-*ras* p21 mutants: molecular basis for their inability to function as signal switch molecules. *Cell* 62:539–48
59. Krieg UC, Walter P, Johnson AE. 1986. Photocrosslinking of the signal sequence of nascent preprolactin to the 54-kilodalton polypeptide of the signal recognition particle. *Proc. Natl. Acad. Sci. USA* 83:8604–8
60. Kwok Y, Wong JT. 1980. Evolutionary relationship between *Halobacterium cutirubrum* and eukaryotes determined by use of aminoacyl-tRNA synthetases as phylogenetic probes. *Can. J. Biochem.* 58:213–18
61. Landman O, Spiegelman S. 1955. Enzyme formation in protoplasts of *Bacillus megaterium. Proc. Natl. Acad. Sci. USA* 41:698–704
62. Deleted in proof
63. Loll PJ, Lattman EE. 1989. The crystal structure of the ternary complex of staphylococcal nuclease, Ca^{2+}, and the inhibitor pdTp, refined at 1.65 Å. *Proteins* 5:183–201
64. Louie A, Jurnak F. 1984. Kinetic studies of *E. coli* elongation factor Tu–guanosine 5′-triphosphate-amino acyl-tRNA complexes. *Biochemistry* 24:6433–39
65. Lyu PC, Sherman JC, Chen A, Kallenbach NR. 1991. α-Helix stabilization by natural and unnatural amino acids with alkyl side chains. *Proc. Natl. Acad. Sci. USA* 88:5317–20
66. Matthews BW. 1993. Structural and genetic analysis of protein stability. *Annu. Rev. Biochem.* 62:139–60
67. Meiering EM, Serrano L, Fersht AR. 1992. Effect of active site residues in barnase on activity and stability. *J. Mol. Biol.* 225:585–89
68. Mendel D, Ellman J, Schultz PG. 1993. Protein biosynthesis with conformationally restricted amino acids. *J. Am. Chem. Soc.* 115:4359–60

69. Mendel D, Ellman JA, Chang Z, Veenstra DL, Kollman PA, Schultz PG. 1992. Probing protein stability with unnatural amino acids. *Science* 256:1798–1802
70. Deleted in proof
71. Mendel D, Ellman JA, Schultz PG. 1991. Construction of a light-activated protein by unnatural amino acid mutagenesis. *J. Am. Chem. Soc.* 113:2758–60
72. Deleted in proof
73. Milburn MV, Tong L, de Vos AM, Brunger A, Yamaizumi Z, et al. 1990. Molecular switch for signal transduction: structural differences between active and inactive forms of protooncogenic *ras* proteins. *Science* 247: 939–45
74. Miller A, Albertini A. 1983. Effects of surrounding sequence on the suppression of nonsense codons. *J. Mol. Biol.* 164:59–71
75. Milligan JF, Groebe DR, Witherell GW, Uhlenbeck OC. 1987. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.* 15: 8783–98
76. Nakatsuka T, Sasaki T, Kaiser ET. 1987. Peptide segment coupling catalyzed by the semisynthetic enzyme thiosubtilisin. *J. Am. Chem. Soc.* 109:3808–10
77. Neet KE, Nanci A, Koshland DE. 1968. Properties of thiol-subtilisin. *J. Biol. Chem.* 243:6392–6401
78. Nomura M, Hosoda J, Nishimura S. 1958. Enzyme formation in lysozyme lysate of *Bacillus subtilis*. *Biochim. Biophys. Acta* 29:161–67
79. Noren CJ, Anthony-Cahill SJ, Griffith MC, Schultz PG. 1989. A general method for site-specific incorporation of unnatural amino acids into proteins. *Science* 244:182–88
80. Noren CJ, Anthony-Cahill SJ, Suich DJ, Noren KA, Griffith MC, Schultz PG. 1990. *In vitro* suppression of an amber mutation by chemically aminoacylated transfer RNA prepared by runoff transcription. *Nucleic Acids Res.* 18:83–88
81. Nowak MW, Kearny PC, Sampson JR, Saks ME, Lavarca GC. 1995. Side chain contributions at the nicotinic receptor binding site probed with unnatural amino-acid incorporation in intact cells. *Science*. Submitted
82. Offord RE. 1987. Protein engineering by chemical means? *Protein Eng.* 1: 151–57
83. Pai EF, Kabsch W, Krengel U, Homes KC, John J, Wittinghofer A. 1989. Structure of the guanine-nucleotide-binding domain of the H*a*-*ras* oncogene product p21 in the triphosphate conformation. *Nature* 341: 209–14
84. Pai EF, Krengel U, Petsko GA, Goody RS, Kabsch W, Wittinghofer A. 1990. Refined crystal structure of the triphosphate conformation of H-*ras* p21 at 1.35 Å resolution: implications for the mechanism of GTP hydrolysis. *EMBO J.* 9:2351–59
85. Patchornik A, Amit B, Woodward RB. 1970. Photosensitive protecting groups. *J. Am. Chem. Soc.* 92: 6333–35
86. Pauling L, Corey RB. 1951. Configurations of polypeptide chains with favored orientations around single bonds: two new pleated sheets. *Proc. Natl. Acad. Sci. USA* 37:729–40
87. Pauling L, Corey RB, Branson HR. 1951. The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl. Acad. Sci. USA* 37:205–11
88. Polgar L, Bender ML. 1966. A new enzyme containing a synthetically formed active site, thiol-subtilisin. *J. Am. Chem. Soc.* 88:3153–54
89. Pollack SJ, Nakayama G, Schultz PG. 1988. Introduction of nucleophiles and spectroscopic probes into antibody combining sites. *Science* 242:1038–40
90. Prive GG, Milburn MV, Tong L, de Vos AM, Yamaizumi Z, et al. 1992. X-ray crystal structures of transforming p21 ras mutants suggest a transition-state stabilization mechanism for GTP hydrolysis. *Proc. Natl. Acad. Sci. USA* 89:3649–53
91. Quiggle K, Kumar G, Oh TW, Ryu EK, Chladek S. 1981. Donor site of ribosomal peptidyltransferase: investigation of substrate specificity using 2′(3′)-*O*-(*N*-acylaminoacyl)-dinucleoside phosphates as models of the 3′ terminus of *N*-acylaminoacyl transfer ribonucleic acid. *Biochemistry* 20: 3480–85
92. Roberts JW, Carbon J. 1975. Nucleotide sequence studies of normal and genetically altered glycine transfer ribonucleic acids from *Escherichia coli*. *J. Biol. Chem.* 250:5530–41
93. Robertson SA, Ellman JA, Schultz PG. 1991. A general and efficient route for chemical aminoacylation of transfer RNAs. *J. Am. Chem. Soc.* 113:2722–29
94. Robertson SA, Noren CJ, Anthony-Cahill SJ, Griffith MC, Schultz PG.

1989. The use of 5′-phospho-2-deoxyribocytidylylriboadenosine as a facile route to chemical aminoacylation of tRNA. *Nucleic Acids Res.* 17: 9649–60
95. Rose GD, Wolfenden R. 1993. Hydrogen bonding, hydrophobicity, packing, and protein folding. *Annu. Rev. Biophys. Biomol. Struct.* 22: 381–415
96. Sayers JR, Schmidt W, Eckstein F. 1988. 5′-3′ Exonuclease in phosphorothioate-based oligonucleotide-directed mutagenesis. *Nucleic Acids Res.* 16:791–802
97. Schimmel P. 1987. Aminoacyl tRNA synthetases: general scheme of structure-function relationships in the polypeptides and recognition of transfer RNAs. *Annu. Rev. Biochem.* 56:125–58
98. Schlichting I, Almo SC, Rapp G, Wilson K, Petratos A, et al. 1990. Time-resolved x-ray crystallographic study of the conformational change in Ha-*ras* p21 protein on GTP hydrolysis. *Nature* 345:309–15
99. Schnolzer M, Kent SBH. 1992. Constructing proteins by dovetailing unprotected synthetic peptides: backbone-engineered HIV protease. *Science* 256:221–25
100. Scholtz JM, Baldwin RL. 1992. The mechanism of α-helix formation by peptides. *Annu. Rev. Biophys. Biomol. Struct.* 21:95–118
101. Seeburg PH, Colby WW, Capon DJ, Goeddel DV, Levinson AD. 1984. Biological properties of human c-Ha-*ras*1 genes mutated at codon 12. *Nature* 312:71–75
102. Spandidos D. 1989. *Ras Oncogenes.* New York: Plenum
103. Spirin AS, Baranov VI, Ryabova LA, Ovodov SY, Alahhov YB. 1988. A continuous cell-free translation system capable of producing polypeptides in high yield. *Science* 242: 1162–64
104. Suich D. 1993. *Mutagenesis of the recognition helix of 434 repressor with nonproteinogenic amino acids.* PhD thesis. Univ. Calif., Berkeley
105. Thorson JS, Judice JK, Chapman E, Murphy EC, Schultz PG. 1994. A linear free energy analysis of hydrogen bonding in proteins. *J. Am. Chem. Soc.* Submitted
106. Deleted in proof
107. Trahey M, McCormick F. 1987. A cytoplasmic protein stimulates normal N-*ras* p21 GTPase, but does not affect oncogenic mutants. *Science* 238: 542–45
108. Tucker PW, Hazen EE Jr, Cotton FA. 1979. Staphylococcal nuclease reviewed: a prototypic study in contemporary enzymology. II. Solution studies of the nucleotide binding site and the effects of nucleotide binding. *Mol. Cell. Biochem.* 23:3–16
109. Weber DJ, Serpersu EH, Shortle D, Mildvan AS. 1990. Diverse interactions between the individual mutations in a double mutant at the active site of staphylococcal nuclease. *Biochemistry* 29:8632–42
110. Wiberg KB, Laidig KE. 1987. Barriers to rotation adjacent to double bonds. 3. The C-O barrier in formic acid, methyl formate, acetic acid, and methyl acetate. The origin of ester and amide "resonances." *J. Am. Chem. Soc.* 109:5935–43
111. Wierenga RK, Terpstora P, Hol WG. 1986. Prediction of the occurrence of the ADP-binding ⟨bab⟩-fold in proteins, using an amino acid sequence fingerprint. *J. Mol. Biol.* 187:101–7
112. Zoller MJ, Smith M. 1983. Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors. *Methods Enzymol.* 100:468–500

Annual Review of Biophysics and Biomolecular Structure
Volume 24, 1995

CONTENTS