BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

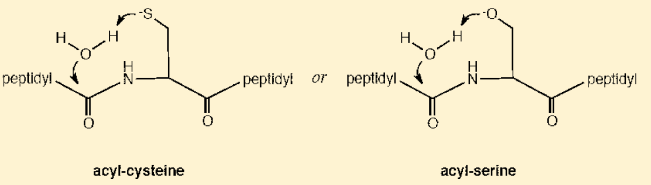
# Amide Bonds to the Nitrogen Atoms of Cysteine and Serine as "Weak Points" in the Backbones of Proteins

Charles A. Lewis, Jr. and Richard Wolfenden*

Department of Biochemistry and Biophysics, University of North Carolina, Chapel Hill, North Carolina 27599, United States

Supporting Information

**ABSTRACT:** During the initial event in protein self-splicing, a peptide bond to the nitrogen atom of an internal cysteine or serine residue is usually cleaved by the side chain −SH or −OH group to yield a thioester or oxyester intermediate that undergoes further reactions. Self-splicing reactions also accompany the maturation of hedgehog signaling proteins, plant-type asparaginases, and pyruvoyl enzymes. It would be of interest to know whether peptide bonds that involve the nitrogen atoms of cysteine or serine are more susceptible to cleavage than peptide bonds to amino acids that lack reactive side chains. Extrapolations of the results of model reactions conducted at elevated temperatures indicate that the −SH group of *N*-acetylcysteine enhances the rate of its hydrolysis by a factor of 70, while the OH group of *N*-acetylserine enhances the rate of its hydrolysis 12-fold, compared with the rate of hydrolysis of *N*-acetylalanine in neutral solution at 25 °C. Several lines of evidence suggest that the rate-enhancing effects of these −SH and −OH side chains arise from their ability to act as intramolecular general acid−base catalysts for hydrolysis, rather than as nucleophilic catalysts. The protein environment within self-splicing proteins appears to redirect the actions of these side chains to nucleophilic attack, generating rate enhancements that approach the rate enhancements generated by conventional enzymes.




The resistance of peptide bonds to hydrolysis accounts for the stability of proteins at ordinary temperatures.[1−3] But some proteins undergo self-splicing reactions during which a segment of polypeptide (the intein) is excised from the interior of the original sequence, and the two end-sections (exteins) become joined in the final sequence of the mature protein. The first step in protein splicing usually occurs at a peptide bond involving the nitrogen atom of cysteine or occasionally serine.[4] Other proteins undergo maturation in a single cleavage event that resembles the first step in protein splicing. During autoprocessing of the hedgehog signaling protein, the N-terminal section of the precursor protein becomes esterified to cholesterol, while the C-terminal section appears to be nonfunctional.[5] In the precursors of plant-type asparaginases, peptide cleavage generates a C-terminal section, the active enzyme, equipped with a catalytic nucleophile at its newly exposed N-terminus.[6] Likewise, PLP-independent amino acid decarboxylases, equipped with catalytic pyruvoyl residues at their N-termini, are generated by cleavage between adjacent serine residues of a precursor protein.[7] In each of these cases, peptide cleavage occurs at the nitrogen atom of a cysteine, serine, or threonine residue of the precursor protein through formation of a cyclic hydroxythiazolidine or hydroxyoxazolidine intermediate.[8]

Because of their speed and specificity, Paulus has compared the activities of self-splicing proteins to those of primitive catalysts that might have served as enzyme precursors.[9] To appreciate the extent to which catalysis is required for these cleavage reactions, it would be of interest to know whether peptide bonds to the nitrogen atoms of cysteine and serine are inherently susceptible to decomposition under ordinary conditions and, if so, by what mechanism decomposition occurs. Here, we show that amide bonds to the nitrogen atoms of amino acids whose side chains are equipped with 3-OH, 3-SH, and 3-$NH_2$ substituents are much less stable to hydrolysis than amide bonds to the nitrogen atom of alanine. However, intramolecular general base catalysis by the 3-substituent, rather than nucleophilic attack of the kind that occurs during protein splicing, appears to be responsible for that instability. In self-splicing proteins, the protein environment appears to redirect the mechanism to nucleophilic attack, producing a rate enhancement approaching the rate enhancements that are achieved by conventional enzymes.

## MATERIALS AND METHODS

*N*-Acetylated amino acids, *N*-acetylamines, *O*-acetylserine, and other compounds were purchased from Sigma-Aldrich Corp. and used without purification. To monitor hydrolysis, solutions of acetylated amines and amino acids (0.03 M) in potassium phosphate buffer (0.1 M, pH 7.0), sealed in quartz tubes under argon, were heated in Thermolyne 47900 ovens for times sufficient to achieve 10−90% reaction. For analysis by proton NMR, samples were then diluted with 99% $D_2O$ containing pyrazine (4H, $\delta$ 8.60 ppm) added as an internal integration standard. Proton NMR spectra were acquired using a Varian Unity 500 MHz instrument equipped with a cryo-probe, controlled by Solaris 9 software, with suppression of the water

**Received:** May 26, 2011
**Revised:** July 13, 2011
**Published:** July 14, 2011

signal. At these concentrations, 4−9 pulses were normally required to obtain integrated intensities of reactants and products with a signal-to-noise ratio of >10:1. The progress of each reaction was followed by measuring the integrated intensities of protons whose signals were well-removed (≥1 ppm) from the residual water signal that remained after water suppression. The chemical shifts of the starting materials and major reaction products are shown in Table S1 of the Supporting Information.

In each case, the disappearance of the starting material followed first-order kinetics under all conditions examined, and plots of the logarithm of the reaction rate as a function of the reciprocal of absolute temperature yielded linear Arrhenius plots. In some cases, more than one set of products was observed (see below), and the rate constant for deacetylation was estimated by multiplying the overall rate constant for disappearance of the starting materials by the fraction attributable to acetate release. In the cases of *N*-acetylcysteine (AcCys), *N*-acetylasparagine (AcAsn), and *N*-acetylglutamine (AcGln), disulfide formation and deamidation occurred to an appreciable extent, and at temperatures above 150 °C, *N*-acetylserine (AcSer) also yielded a complex mixture of products. The concentrations of the immediate reaction products were identified by analysis of the products of their breakdown and used to estimate the concentration of amino acid or amine that had been released initially by amide hydrolysis. In every case, reaction rates at 25 °C were obtained by extrapolation.

## RESULTS

**Decomposition of Acetylated Amino Acids.** An initial survey of the rates of decomposition and of acetate release from ten *N*-acetylated amino acids was conducted in potassium phosphate buffer (0.1 M, pH 7.0) at 150 °C. The results, summarized in Table S2 and Figure S1, indicate that AcSer and AcCys are deacetylated most rapidly, at rates that are similar in magnitude to the rates of hydrolysis of the amide side chains of AcAsn and AcGln.

Next, we determined the rates of decomposition of AcAla, AcSer, and AcCys at 100 °C in potassium phosphate buffer (0.1 M, pH 7.0). Under those conditions, AcAla and AcSer yielded only the expected hydrolysis products. AcCys underwent simultaneous oxidation to *N,N*′-diacetylcystine and hydrolysis to acetate and cysteine (with some appearance of cystine), in a ratio of ∼2:1. Schemes for the decomposition of AcSer and AcCys at 100 °C are presented in the Supporting Information. Rate constants for the decomposition of the three acetylated amino acids and for the appearance of acetate at 100 °C are shown in Table 1. At 100 °C, AcSer was hydrolyzed 18-fold more rapidly, while AcCys is hydrolyzed 114-fold more rapidly than AcAla.

**Table 1. Rate Constants ($s^{-1}$) for the Disappearance of *N*-Acetylated Amino Acids (AcAA) and for the Appearance of the Amino Acid (AA) and Acetate (Ac) in Potassium Phosphate Buffer (0.1 M, pH 7.0) at 100 °C[a]**

| substrate | −d(AcAA)/d*t* | | +d(AA)/d*t* | | +d(Ac)/d*t* | |
|---|---|---|---|---|---|---|
| AcAla | $2.2 \times 10^{-8}$ | | $2.1 \times 10^{-8}$ | | $2.4 \times 10^{-8}$ | |
| AcSer | $4.0 \times 10^{-7}$ | 18× | $3.2 \times 10^{-7}$ | 15× | $4.0 \times 10^{-7}$ | 16× |
| AcCys | $2.5 \times 10^{-6}$ | 114× | $6.9 \times 10^{-7}$ | 33× | $7.2 \times 10^{-7}$ | 30× |

[a]Numbers between columns are the ratio of the rate for AcSer or AcCys to the rate for AcAla.

The same reactions were monitored over the temperature range between 80 and 230 °C in potassium phosphate buffer (0.1 M, pH 7.0). Rate constants for substrate decomposition and for acetate formation yielded linear Arrhenius plots (see Figure 1), from which values at 25 °C were obtained by extrapolation (Table 2). In terms of their rate constants for appearance of the hydrolysis products extrapolated to 25 °C, AcSer and AcCys released acetate 12- and 68-fold more rapidly than did AcAla, respectively.

Over the temperature range between 80 and 200 °C, AcAla yielded only alanine and acetate, but small quantities of acetaldehyde, ethylamine, methylamine, and methanol began to appear at temperatures above 200 °C. AcSer produced only acetate and serine at temperatures below 100 °C, but at higher temperatures, glycine, pyruvate, ethanolamine, glycolate, acetaldehyde, and alanine began to appear. In separate experiments under similar conditions, these byproducts were also found to be generated from serine itself in similar proportions. At 100 °C, AcCys yielded a product mixture that was dominated by *N,N*-diacetylcystine (67% at 100 °C, decreasing to 15% at 150 °C) with significant concentrations of cystine, pyruvate, glycolate, acetaldehyde, and acetate accumulating at higher temperatures.

**Decomposition of Acetylated Amines at 150 °C.** To test these reactions in a simpler system, we examined the behavior of the corresponding *N*-acetylamines at temperatures between 100 to 210 °C, with the results shown in Table 2. Only *N*-acetylcysteamine yielded products other than acetate and the amine. Cysteamine was not observed in the product mixture (<2%). Instead, the relatively stable disulfide *N,N*-diacetylcystamine was the dominant product along with smaller amounts of glycolate and acetaldehyde that had presumably arisen from the decomposition of cysteamine.

To compare the effects of a 3-amino substituent with the effects of a 3-hydroxy or a 3-mercapto substitutent discussed above, we also examined the behavior of *N*-acetylethylenediamine—the decarboxylated version of 3-aminoalanine. The results, included in Table 2, show that a 3-amino group and a 3-hydroxyl group are similar in the rate enhancement that they produce.

**Equilibrium Constants for N-to-S Isomerization of Acetylcysteine and N-to-O Isomerization of Acetylserine in 0.1 M HCl.** To evaluate the likelihood of mechanisms involving participation of the 3-SH group of cysteine and the 3-OH group of serine as nucleophiles, rather than as general acid−base catalysts (see Discussion), it was necessary to evaluate the equilibrium constants for N-to-S isomerization of AcCys and N-to-O isomerization of AcSer (see Discussion). When the *N*-acetylated compounds (0.03 M) were incubated in 0.1 M HCl at elevated temperatures, equilibrium was was established within 25 h at 50 °C and within 5 h at 80 °C. At these low temperatures, no hydrolysis was found to have occurred. Positions of equilibrium were estimated from the $^1$H NMR resonances of the acetyl groups of the *N*-acetyl isomer of AcSer ($\delta$ 1.98 ppm), the *O*-acetyl isomer of AcSer ($\delta$ 2.04 ppm), the *N*-acetyl isomer of AcCys ($\delta$ 1.99 ppm), and the *S*-acetyl isomer of AcCys ($\delta$ 2.35 ppm). van't Hoff plots of the logarithm of $K_{eq}$ for N-to-S isomerization of acetylcysteine and N-to-O isomerization of acetylserine as a function of the reciprocal of absolute temperature were linear and yielded equilibrium constants of 0.08 ± 0.01 ($\Delta H$ = −0.8 kcal/mol) for conversion of the

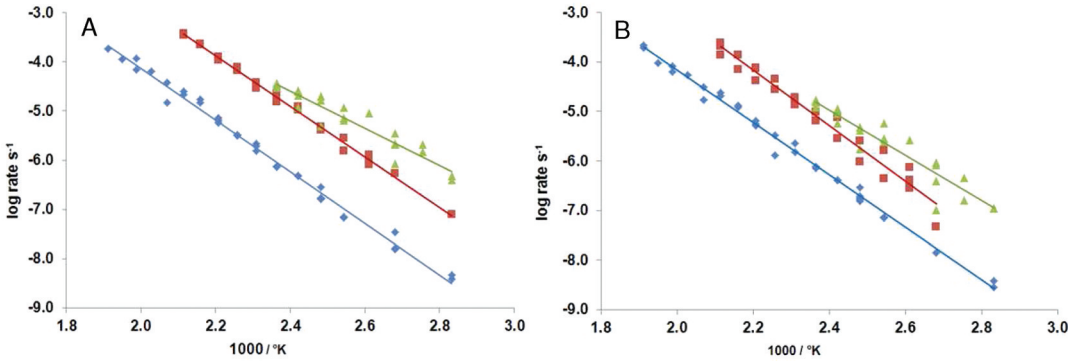



**Figure 1.** Arrhenius plots for decomposition and the release of acetate for acetylated alanine (◆), serine (■), and cysteine (▲). Panel A shows the decompositions of the acetylated amino acids; equations for trendlines: Ala, $y = -5.253x + 6.375$; Ser, $y = -5.132x + 7.413$; Cys, $y = -3.767x + 4.449$. Panel B shows the release of acetate; equations for trendlines: Ala, $y = -5.275x + 6.375$; Ser, $y = -5.606x + 8.162$; Cys, $y = -4.596x + 6.063$. Duplicate or more data points were collected at each temeperature. Alanine and serine reactions gave very good fits to the lines with $r^2$s of ≥0.99, while the cysteine data gave slightly poorer fits with $r^2$s of ~0.90.

**Table 2. Rate Constants at 25 °C and Thermodynamic Activation Parameters for the Decomposition of Acetylamino Acids and Acetylamines and for Acetate Release (pH 7.0)[a]**

| | $k$ ($s^{-1}$) at 25 °C | $\Delta H^{\ddagger}$ | $\Delta G^{\ddagger}$ | $T\Delta S^{\ddagger}$ | $t_{1/2}$ (years) |
|---|---|---|---|---|---|
| *alanine* | | | | | |
| *N*-acetylalanine | | | | | |
| −d(AcAla)/d*t* | $5.3 \times 10^{-12}$ | 23.6 | 32.7 | −9.2 | 4200 |
| +d(Ac)/d*t* | $6.5 \times 10^{-12}$ | 23.1 | 3.26 | −9.5 | 3400 |
| *N*-acetylethylamine | | | | | |
| −d(AcEt)/d*t* | $3.9 \times 10^{-12}$ | 25.7 | 32.0 | −7.5 | 5500 |
| +d(Ac)/d*t* | $3.9 \times 10^{-12}$ | 25.7 | 32.0 | −7.3 | 5500 |
| *serine* | | | | | |
| *N*-acetylserine[b] | | | | | |
| −d(AcSer)/d*t* | $2.0 \times 10^{-10}$ | 22.4 | 30.6 | −8.2 | 100 |
| +d(Ac)/d*t* | $7.6 \times 10^{-11}$ | 23.4 | 31.2 | −7.8 | 300 |
| *N*-acetylethanolamine[e] | | | | | |
| −d(AcEtO)/d*t* | $2.1 \times 10^{-10}$ | 21.7 | 30.6 | −9.4 | 100 |
| +d(Ac)/d*t* | $2.1 \times 10^{-10}$ | 21.7 | 30.6 | −9.4 | 100 |
| *cysteine* | | | | | |
| AcCys[c] | | | | | |
| −d(AcCys)/d*t* | $6.5 \times 10^{-9}$ | 16.7 | 28.5 | −11.9 | 3 |
| +d(Ac)/d*t* | $4.4 \times 10^{-10}$ | 20.5 | 30.1 | −9.7 | 50 |
| Ac-cysteamine[d] | | | | | |
| −d(AcCa)/d*t* | $6.6 \times 10^{-10}$ | 19.9 | 29.9 | −10.0 | 33 |
| +d(Ac)/d*t* | $1.8 \times 10^{-10}$ | 20.4 | 30.6 | −10.2 | 120 |
| *3-aminoalanine* | | | | | |
| Ac-3-aminoalanine | | | | | |
| Ac-ethylenediamine[e] | nd | nd | nd | nd | nd |
| −d(AcEda)/d*t* | $1.2 \times 10^{-10}$ | 22.6 | 30.9 | −8.3 | 180 |
| +d(Ac)/d*t* | $1.2 \times 10^{-10}$ | 22.6 | 30.9 | −8.3 | 180 |

[a]Approximate errors in the thermodynamic parameters are $\Delta H^{\ddagger} \pm 1$ kcal/mol, $\Delta G^{\ddagger} \pm 0.2$ kcal/mol, and $T\Delta S^{\ddagger} \pm 1.2$ kcal/mol. [b]Ac-serine decomposition was greater than acetate release because serine decomposed into other species at a rate which is about half the decomposition rate. [c]Ac-cysteine decomposition was greater than acetate release because cysteine decomposed into other species at a rate similar to the decomposition rate. [d]Ac-cysteamine decomposition was greater than acetate release because cysteamine readily decomposes into other species. [e]Ac-ethylamine and Ac-ethanolamine decompositions are similar to acetate release since ethylamine and ethanolamine are relatively stable over the temperature range of the studies.

*N*-acetyl isomer of acetylcysteine to the *S*-acetyl isomer and 0.02 ± 0.01 ($\Delta H$ = 5.8 kcal/mol) for conversion of the *N*-acetyl isomer of acetylserine to the *O*-acetyl isomer in 0.1 M HCl at 25 °C.

**Rate Constant of O-to-N Isomerization of *O*-Acetylserine in Neutral Solution.** To evaluate the likelihood of mechanisms involving participation of the 3-SH group of cysteine and the 3-OH group of serine as nucleophiles, rather than as general acid−base catalysts (see Discussion and Supporting Information), it was also necessary to evaluate the rate constant for O-to-N isomerization of *O*-acetylserine at pH 7. In his early experiments, Josefsson observed the isomerization of *O*-(1-$^{14}$C-*O*-acetyl)serine to *N*-(1-$^{14}$C-acetyl)serine in acid solution (0.1 M acetic acid) at room temperature, with


dx.doi.org/10.1021/bi200813s | *Biochemistry* 2011, 50, 7259–7264

no release of $^{14}C$,[10] but did not report the rate of reaction. In the present experiments, rate constants for isomerization of *O*-acetylserine to *N*-acetylserine were determined by incubating samples (0.03 M) at room temperature in potassium phosphate buffer (0.1 M, pH 7.0), removing samples at intervals to determine the concentration of the *O*-acetyl isomer ($\delta$ 2.08) and the *N*-acetyl isomer ($\delta$ 2.02) by $^1H$ NMR. Conversion of *O*-acetylserine to AcSer followed first order kinetics to completion with a rate constant of $7.4 \times 10^{-6}$ $s^{-1}$ at 25 °C ($t_{1/2}$ = 26 h) (Figure 2).

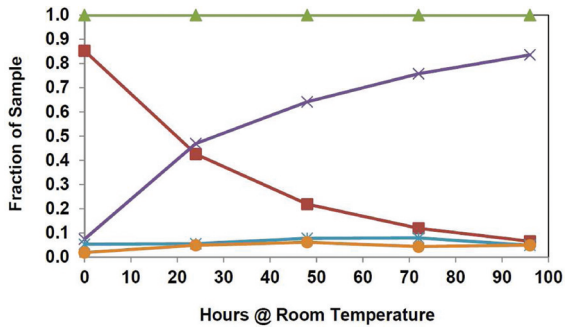



**Figure 2.** Conversion of *O*-AcSer to *N*-AcSer at room temperature in 0.1 M potassium phosphate buffer, pH 7. The decomposition of *O*-AcSer (red data) occurred with concurrent formation of *N*-AcSer (purple data). Note that only small amounts of serine (aqua data) and acetate (orange data) were initially released and did not increase. Shown for comparison is the decomposition of *N*-AcSer (green data).

Comparable rates of O-to-N isomerization have been reported for a series of benzyloxycarbonylglycine esters of thiazolidines designed to mimic intermediates in the first step of splicing ($t_{1/2}$ ~ 9 h)[11] and for a series of peptides containing an intein plus a fragment of the *C*-extein ($t_{1/2}$ = 9–110 h).[12] Shao and Paulus[13] used depsipseptides with ester or thioester (rather than amide) linkages between glutamine and either serine or cysteine as a model for the O-to-N and S-to-N transfer which occurs in the final step of the intein reaction. In both cases, the $t_{1/2}$ values obtained were ~1 m at neutral pH. The rapidity of these reactions reflects the presence of the glutamine residue, which may participate in the reaction in a manner similar to asparagine through the formation of a cyclic intermediate. That reaction differs from the O-to-N transfer reactions that is are subject of the present experiments.

## DISCUSSION

The rate of hydrolysis of AcAla, extrapolated to 25 °C at pH 7, is enhanced 12- and 68-fold by a neighboring hydroxyl and thiol group, respectively, while the rate of hydrolysis of *N*-acetylethylamine is enhanced 70-fold and 100-fold by the same substitutions. These results imply that peptide bonds to the nitrogen atoms of cysteine and serine are appreciably less stable to hydrolysis than amide bonds to the nitrogen atom of alanine and raise the question of whether their instability arises from intramolecular nucleophilic or general acid–base catalysis. These effects are neither large enough to exclude the possibility of intramolecular general acid–base catalysis nor small enough to exclude the possibility of intramolecular nucleophilic catalysis, since they lie near to the breakpoint suggested by Kirby's analysis of effective molarities in intramolecular reactions.[14] On the basis of the magnitudes of these effects, either (or both) mechanisms might be responsible for the observed rate enhancements.

The following features of these model reactions seem to favor the view that the rate-enhancing effects of a nearby thiol or hydroxyl group arise from general acid–base catalysis of hydrolysis, rather than from nucleophilic catalysis.

First, the observed values of $T\Delta S^{\ddagger}$ are similar in magnitude for both series of reactants (the *N*-acetylamino acids and the *N*-acetylamines (Table 2)), and the value of $T\Delta S^{\ddagger}$ is marginally less favorable for the thiol-catalyzed reaction than for the uncatalyzed reaction. If the rate-limiting step for the thiol- and hydroxyl-catalyzed reactions involved intramolecular cyclization but the uncatalyzed reaction did not, then the intramolecular cyclization reaction might have been expected to exhibit a lower entropy of activation.[a]

Second, nucleophilic catalysis can be ruled out as the source of the rate enhancement observed in *N*- acetylethylenediamine because intramolecular amine attack would result in regeneration of the starting material. Although that argument does not rule out side-chain attack in the cases of *N*-acetylcysteamine and *N*-acetylethanolamine, it is worth noting that the free amino group of *N*-acetylethylenediamine exerts a catalytic effect that is comparable in magnitude with those observed for the thiol and hydroxyl groups of *N*-acetylcysteamine and *N*-acetylethanolamine (Table 2). Thus, it seems reasonable to conjecture that general acid–base catalysis is at work in those cases as well.

Finally, any hydrolytic reaction proceeding through a thioester or oxyester intermediate could not proceed more rapidly than the equilibrium constant for formation of that intermediate, multiplied by the rate constant for its hydrolysis. The equilibrium constant for conversion of AcCys to S-acetylcysteine can be estimated as follows. At pH 1, where *N*-acetylcysteine is uncharged but the isomerization product *O*-acetylcysteine is almost fully protonated at its newly exposed amino group, the equilibrium constant observed here for N-to-S transfer is 0.08. The $pK_a$ value of the $\alpha$-ammonium group of cysteine (released by acetyl migration) is ~10.8. It follows that the equilibrium constant for N-to-S isomerization would be $1.3 \times 10^{-11}$ if only the uncharged reactant and product were present,[16] and that at pH 7, the observed equilibrium constant for N-to-S isomerization of acetylcysteine would be $9 \times 10^{-8}$. Multiplying that value by the rate constant ($2.5 \times 10^{-7}$ $s^{-1}$) observed for hydrolysis of aliphatic thioesters at pH 7[17] yields $2.2 \times 10^{-14}$ $s^{-1}$ as the maximum possible rate constant for decomposition of *N*-acetylcysteine via *S*-acetylcysteine. That rate constant is $10^4$-fold lower than the constant observed here for hydrolysis of *N*-acetylcysteamine ($1.8 \times 10^{-10}$ $s^{-1}$), from which it can be inferred that *N*-acetylcysteamine hydrolysis proceeds by direct water attack, not through a thioester intermediate.[b] Likewise, the equilibrium constant for conversion of *N*-acetylserine to *O*-acetylserine can be estimated as follows. At pH 1, where *N*-acetylserine is uncharged but the isomerization product *O*-acetylserine is almost fully protonated at its newly exposed amino group, the equilibrium constant observed here for N-to-O transfer is 0.02. The $pK_a$ value of the ammonium group of serine (released by acetyl migration) is ~10.8. It follows that the equilibrium constant for N-to-O isomerization would be $3.2 \times 10^{-12}$ if only the uncharged reactant and product were present[16] and that at pH 7 the observed equilibrium constant for N-to-S isomerization would be $2 \times 10^{-8}$. Multiplying that value by the rate constant ($3 \times 10^{-7}$ $s^{-1}$) observed for hydrolysis of aliphatic oxyesters at pH 7[16] yields $6 \times 10^{-15}$ $s^{-1}$ as the expected rate constant for


dx.doi.org/10.1021/bi200813s | *Biochemistry* 2011, 50, 7259–7264

decomposition of *N*-acetylserine via *S*-acetylserine. Because that rate constant is far less than the constant observed here for acetate release from *N*-acetylcysteamine ($7.5 \times 10^{-10}$ $s^{-1}$), this latter process presumably proceeds by direct water attack rather than through an oxyester intermediate.

Shao et al.[18] used several intein constructs to demonstrate the N to O and N to S acyl rearrangement reaction and estimated $t_{1/2}$s ranged from 5.8 h, without added amine nucleophiles, to 0.6 and 0.1 h when ethylenediamine or hydroxylamine respectively were present. The faster rates of reaction likely reflect the presence of the unique features of the intein protein structure and environment present to direct and accelerate the reaction.

Figure 3 shows the present rate constants and also a rate constant reported earlier[1] for diketopiperazine formation

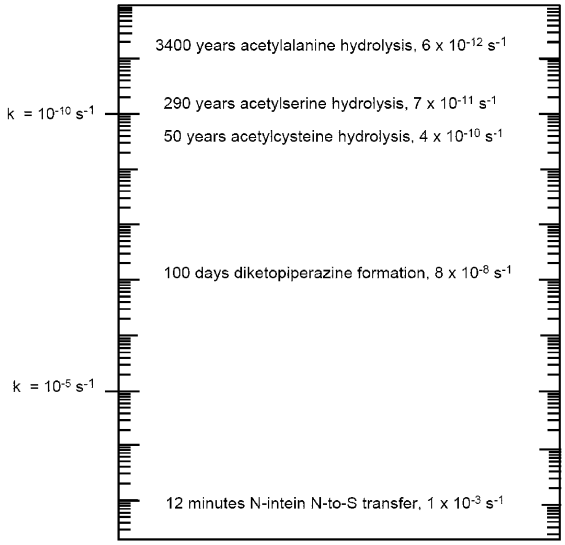



**Figure 3.** Half-lives and rate constants (25 °C, pH 7).

(a degradative reaction that occurs in proteins equipped with a free amino terminus). Comparison with a reported rate constant of $\sim 10^{-3}$ $s^{-1}$ for intein N-to-S transfer[19] indicates that the first step in protein splicing occurs at least $10^6$-fold more rapidly than either of the model hydrolytic reactions described here.

Since reactions in general tend to proceed by the mechanism that involves the lowest energy barrier, the rate constants observed for the model reactions described here set an upper limit on the rate constants that would be observed for any (hypothetical) model reaction that might bear a closer resemblance to the first step in intein splicing reaction in which an ester (or thioester) intermediate is generated. If the maximum possible rate constant estimated above for decomposition of *N*-acetylcysteine via *S*-acetylcysteine ($5.5 \times 10^{-15}$ $s^{-1}$) is compared with the rate constant ($\sim 10^{-3}$ $s^{-1}$) produced by a trans-splicing intein from *Synechocystis sp.*, the resulting rate enhancement is more than $10^{11}$-fold, comparable in magnitude with the values achieved by many enzymes.[20] Thus, intein-containing proteins are even better catalysts (in terms of the extent to which they enhance the rate of reaction by the double displacement mechanism that is actually followed) than simple rate comparisons suggest.[c] Similar considerations presumably apply to other processes mentioned in the Introduction, including the autoprocessing of the hedgehog signaling protein, the maturation of plant-type asparaginases, and the generation of amino acid decarboxylases equipped with catalytic pyruvoyl residues at their N-termini.

## ■ ASSOCIATED CONTENT

### Ⓢ Supporting Information

(1) NMR analysis, Table S1: Proton NMR data for alanine, serine, and cysteine derivatives used in these studies; (2) rate constants for decomposition of acetylated amino acids at 150 °C including Table S2, Figure S1, Schemes S1 and S2 for the decomposition of Ac-Gln and Ac-Glu, respectively, Table S3: first order rate constants for decomposition of serine and cysteine at 150 °C; (3) data for the decomposition of Ac-Ser and Ac-Cys at 100 °C. This material is available free of charge via the Internet at http://pubs.acs.org.

## ■ AUTHOR INFORMATION

### Corresponding Author

*E-mail: Richard_Wolfenden@med.unc.edu. Phone: (919) 966-1203. Fax: (919) 966-2852.

### Funding

This research was supported by the National Institutes of General Medical Sciences, Grant GM-18325.

## ■ ACKNOWLEDGMENTS

We thank David Edwards and Michael Caplow for helpful discussions during the preparation of this manuscript.

## ■ ADDITIONAL NOTES

[a]It should be noted that in the hydrolysis of a series of *N*-alkylmaleamic acids entropies of activation have been found to be less favorable for the more reactive maleamic acids,[15] implying the need for caution in efforts to predict the magnitude and sign of entropic effects in intramolecular reactions.

[b]In an effort to model the fourth step in protein splicing, in which the N-terminal extein is transferred from the SH group of cysteine to the $\alpha$-amino group of cysteine, Shao and Paulus[13] were able to follow the transfer of Fmoc-Gln from the −SH group to the N-terminus of Cys-Asp-Gly-Tyr-$NH_2$. That model reaction was found to proceed rapidly at 37 °C, with a $t_{1/2}$ of ~150 s corresponding to a rate constant of $\sim 5 \times 10^{-3}$ $s^{-1}$, at pH 4. If it is assumed that the rate of the Shao and Paulus model reaction is proportional to the fraction of the amino group that is uncharged, then the rate of S-to-N transfer in that model reaction would ~5 $s^{-1}$ at pH 7 at 37 °C or, because $\Delta H^{\ddagger}$ is only 4.2 kcal/mol, ~3 $s^{-1}$ at 25 °C. If $k_{S>N}$ is 3 $s^{-1}$ and $K_{eq(N>S)}$ is $9 \times 10^{-8}$ (see main text), then $k_{N>S}$ is $3 \times 10^{-7}$ $s^{-1}$, much smaller than the rate constant ($10^{-3}$ $s^{-1}$)[19] that is actually observed for N-to-S transfer in the first step in protein splicing. This comparison, of the rate constant for nucleophilic attack in the intein system, with the rate constant for nucleophilic attack in the Shao and Paulus model reaction, was suggested by a reviewer.

[c]A plausible mechanism of reaction, based on analysis of the results of $^1$H NMR experiments and QM/MM simulations, has been proposed for the initial acyl transfer event, in which an essential histidine residue ($pK_a$ = 7.3), that is distant in sequence from the scissile peptide bond but proximal in space, serves as a general base catalyst for nucleophilic attack by −SH and later acts as a general acid catalyst to assist elimination of the leaving $-NH_2$ group. Notably, the $pK_a$ value of that

histidine residue (H73 in an intein-containing protein from *M. tuberculosis*) is reduced by ≥3.8 units during the initial splicing event.[21]

## REFERENCES


(1) Radzicka, A., and Wolfenden, R. (1996) Rates of Uncatalyzed Peptide Bond Hydrolysis in Neutral Aqueous Solution and Transition State Affinities of Proteases. *J. Am. Chem. Soc. 118*, 6105−6109.

(2) Bryant, R. A. R., and Hansen, D. E. (1996) Direct Measurement of the Uncatalyzed Rate of Hydrolysis of a Peptide Bond. *J. Am. Chem. Soc. 118*, 5498−5499.

(3) Smith, R. M., and Hansen, D. E. (1998) The pH-Rate Profile for the Hydrolysis of the Peptide Bond. *J. Am. Chem. Soc. 120*, 8910−8913.

(4) Perler, F. B. (2002) InBase, the Intein Database. *Nucleic Acids Res. 30*, 383−384.

(5) Porter, J. A., Young, K. E., and Beachy, P. A. (1996) Cholesterol Modification of Hedgehog Signaling Proteins in Animal Development. *Science 274*, 255−259.

(6) Michalska, K., Hernandoez-Santoyo, A., and Jaskolski, M. (2008) The Mechanism of Autocatalytic Activation of Plant-type L-Asparaginases. *J. Biol. Chem. 283*, 13388−13397.

(7) van Poelje, P. D., and Snell, E. E. (1990) Pyruvoyl-Dependent Enzymes. *Annu. Rev. Biochem. 59*, 29−59.

(8) Perler, F. B., Xu, M.-Q., and Paulus, H. (1997) Protein Splicing and Autoproteolysis Mechanisms. *Curr. Opin. Chem. Biol. 1*, 292−299.

(9) Paulus, H. (2001) Inteins as Enzymes. *Bioorg. Chem. 29*, 119−129.

(10) Josefsson, L. (1963) Mechanism of O,N-Acyl Shift in O-acetyl-DL-serine. *Biochim. Biophys. Acta 74*, 774−775.

(11) Liu, C.-F., and Tam, J. P. (1994) Chemical Ligation Approach to Form a Peptide Bond Between Unprotected peptide Segments. Concept and Model Study. *J. Am. Chem. Soc. 116*, 4150−4153.

(12) Frutos, S., Goger, M., Giovanni, B., Cowburn, D., and Muir, T. (2010) Branched Intermediate Formation Stimulates Peptide Bond Cleavage in Protein Splicing. *Nat. Chem. Biol. 6*, 527−533.

(13) Shao, Y., and Paulus, H. (1997) Protein Spilcing: Estimation of the Rate of O-N and S-N acyl Rearrangements, the Last Step of the Splicing Process. *J. Peptide Res. 50*, 193−198.

(14) Kirby, A. J. (1980) Effective Molarities for Intramolecular Reactions. *Adv. Phys. Org. Chem. 17*, 183−277.

(15) Kirby, A. J., and Lancaster, P. W. (1972) Structure and Efficiency in Intramolecular and Enzymic Catalysis. Catalysis of Amide Hydrolysis by the Carboxy-group of Substituted Maleamic Acids. *J. Chem. Soc., Perkin Trans. 2*, 1206−1214.

(16) Carpenter, F. H. (1960) The free energy change in hydrolytic reactions: the non-ionized compound convention. *J. Am. Chem. Soc. 82*, 1111−1122.

(17) Schaefgen, J. R. (1948) Kinetics of the hydrolysis of ethyl thiolacetate in aqueous solution. *J. Am. Chem. Soc. 70*, 1308−1311.

(18) Shao, Y., Xu, M.-Q., and Paulus, H. (1996) Protein Splicing: Evidence for an N-O Acyl Rearrangement as the Initial Step in the Splicing Process. *Biochemistry 35*, 3810−3815.

(19) Martin, D. D., Xu, M.-Q., and Evans, T. C. Jr. (2001) Characterization of a Naturally Occurring Trans- Splicing Intein from *Synechocystis* sp. PCC6803. *Biochemistry 40*, 1393−1402.

(20) Wolfenden, R. (2006) Degrees of Difficulty of Water Consuming Reactions in The Absence of Enzymes. *Chem. Rev. 106*, 3379−3396.

(21) Du, Z., Shemella, P. T., Liu, Y., McCallum, S. A., Pereira, B., Nayak, S. K., Belfort, G., Belfort, M., and Wang, C. (2009) Highly Conserved Histidine Plays a Dual Catalytic Role in Protein Splicing: A $pK_a$ Shift Mechanism. *J. Am. Chem. Soc. 131*, 11581−11589.