thermore, we recently isolated a less stable product with the axial nitro group in the reaction of **2** with acetylacetone (unpublished data).

(18) Although the reason why our conditions, different from the conditions employed by Paulsen et al., gave no manno isomer is not clear, it was noteworthy that the reaction of **2** (the α anomer of **1**) as well as the analogous α-acetate with the excess of hydrazoic acid in chloroform solutions afforded exceptionally large amounts of cis adduct (as a major product) with respect to the aglycone; these conditions used excess of reagent in nonpolar solvent to be close to those employed by Paulsen et al. rather than those described here.

(19) Nitronate ion derived from the gluco isomer might exist in the chair conformation. Equatorial attack of proton on this ion may occur, at least to a considerable extent, because the $A^{(1,3)}$ strain would force both the cyano group and oxygen at C-4 up and block approach of proton from the axial side.[5,21]

(20) Nitrite ion is not a poor leaving group and basicity of this reaction is not high; therefore, the possibility of Elcb reaction may be neglected. The anti elimination of the allo isomer should predominate over syn elimination of the gluco isomer owing to the torsional strain and the principle of least motion; J. E. Bunnett. *Surv. Prog. Chem.*, **5**, 53 (1969).

(21) Bordwell and co-worker, *J. Am. Chem. Soc.*, **92**, 5933 (1970), found that *cis*-2-phenyl-1-nitrocyclohexane was deprotonated 350-fold more rapidly than its trans isomer. Angyal and co-worker, *Aust. J. Chem.*, **23**, 1485 (1970), showed that *cis*-2-nitrocyclohexane-*cis*-1,3-diol rapidly epimerized in solution to the more stable trans,trans isomer. These facts suggest that abstraction of H-3 of the allo isomer should be much more rapid than that of the gluco isomer. As already described, abstraction of H-3 of the gluco isomer was much easier than that of H-2. Therefore it is likely that before a base works on H-2 of the allo isomer deprotonation of H-3 should occur to give a nitronate ion, from which the more stable gluco isomer is formed.

(22) P. M. Collins, D. Gardiner, S. Kumar, and W. G. Overend, *J. Chem. Soc., Perkin Trans. 1*, 2596 (1972).

(23) T. Sakakibara and R. Sudoh, *J. Chem. Soc., Chem. Commun.*, in press.

(24) Combined steric hindrance due to H-4 (1,3-nonbonded interaction) and H-5 (1,4-nonbonded interaction), the former retarding axial attack and the latter equatorial attack, appears to be less important than that due to the glycosidic methoxy group.

(25) H. H. Baer and T. Neilson, *Can. J. Chem.*, **43**, 840 (1965).

(26) H. H. Baer, F. Kienzle, and T. Neilson, *Can. J. Chem.*, **43**, 1829 (1965).

(27) T. Sakakibara, R. Sudoh, and T. Nakagawa, *Bull. Chem. Soc. Jpn.*, submitted for publication.

The Chemistry of a Method for the Determination of Carboxyl-Terminal Residues in Peptides

Marvin J. Miller, F. Eugene DeBons, and G. Marc Loudon*

Spencer Olin Laboratory, Department of Chemistry, Cornell University, Ithaca, New York 14853

Received October 26, 1976

We here describe in detail our method whereby the carboxyl-terminal residue of a peptide may be determined. The procedure involves, first, formation of an O-substituted hydroxamic acid by reaction of the peptide C-terminal carboxyl group with a water-soluble carbodiimide and a nucleophilic O-substituted hydroxylamine. Considerations leading to the choice of *O*-pivaloylhydroxylamine (**10**, OPHA) as the nucleophile are described, and optimization of the conditions for this reaction are discussed. The O-substituted hydroxamic acid then undergoes, at higher pH, a Lossen rearrangement, leading to degradation of the carboxyl-terminal residue and observed loss of that residue upon subsequent amino acid analysis. The fate of the rearranged residue has been thoroughly explored, and the rates of the Lossen rearrangement have been characterized for representative amino acids. Potential interferences from aromatic amino acids, the amino-terminal residue, carboxyl-terminal dicarboxylic amino acids, and residues with nucleophilic side chains have been characterized. In some cases, these do not occur; in cases in which such interferences do occur, these have been characterized, and most do not affect the utility of the method. Only carboxyl-terminal Asp and Glu do not degrade satisfactorily, and the reasons for this observation have been investigated.

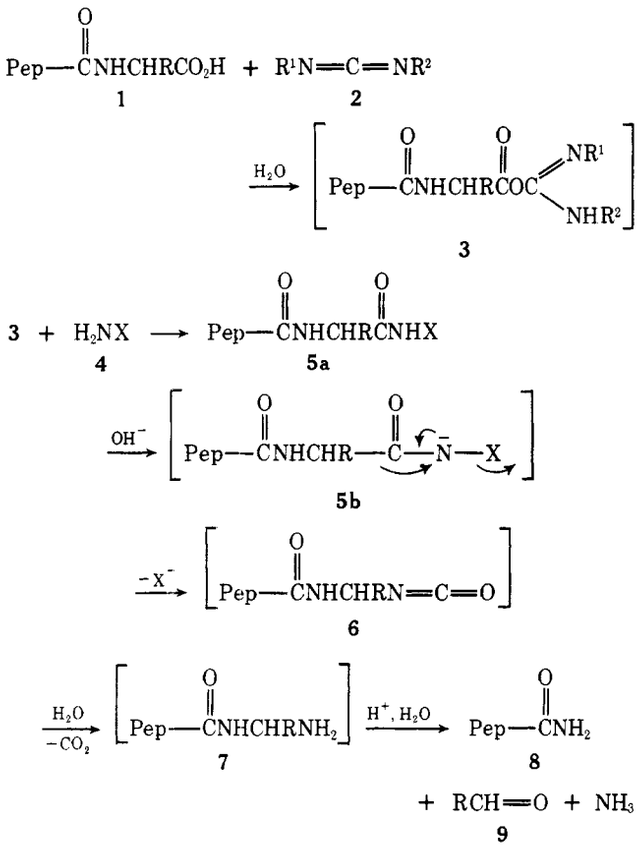
Rapid experimental development has provided the protein chemist with a variety of useful analytical tools. A number of reagents are now available for the specific modification and cleavage of peptide chains,[2,3] and highly efficient amino terminal[3,4] and sequential[5] methods have been described. However, "no entirely satisfactory chemical method of carboxyl-terminal analysis exists".[3] Of the several carboxyl-terminal methods that have been proposed,[4,6] all suffer limitations, and few have been useful in actual practice. Vigorous conditions, solvent restrictions, and failure at certain amino acid residues have all contributed to their lack of general utility. It becomes important to search for new methods of carboxyl-terminal analysis which complement other methods now in use. We have developed a method of carboxyl-terminal residue analysis of peptides which not only circumvents some of the usual limitations, but is also mild, efficient, and easily used in aqueous solution, mixed solvents, or 8 M urea. We here wish to describe this method and the detailed chemistry of its development.

The concept of this method is shown in Scheme I, and involves the activation of the C-terminal carboxyl group followed by reaction with a nucleophile NH_2X. When X is a good leaving group, the resulting amide **5** is susceptible to rearrangement leading to eventual loss of the C-terminal residue as an aldehyde. Although the Lossen and related rearrangements have had substantial appeal for such a degradation,[7,8] until now such a degradation has been impractical because of the large number of reactions hitherto required to generate the intermediate **5** which is capable of rearrangement. Once the rearrangement occurs, the carboxyl terminal residue will be absent in the amino acid analysis, and subsequent reactions determine its ultimate fate, but not the fact of its loss. Critical to the fulfillment of Scheme I, then, is the quantitative conversion of peptide carboxyl-terminal carboxyl groups to a derivative such as **5**. This objective required careful selection of the nucleophile, NH_2X.

Choice of the Nucleophile. In order to be useful, the nucleophile NH_2X is required to be stable, reasonably water soluble, nucleophilic enough to compete with hydrolysis of a carbodiimide-derived activated intermediate, and, most importantly, it must incorporate a critical substituent, X, a potential leaving group which would be compatible with a reasonably facile rearrangement. Rearrangements of the type illustrated in Scheme I include the Hofmann,[9] Curtius,[10] Wawzonek,[11] and Lossen[12,13] reactions. Consideration of these types of reactions led to the synthesis and investigation of a variety of potential nucleophiles shown in Table I.[14] Of all these, only *O*-pivaloylhydroxylamine (OPHA, **10**) satisfactorily fulfilled the requisite criteria.

O-Pivaloylhydroxylamine was first prepared by Marmer and Maerker so that they could study its potential as an aminating agent.[15] These workers showed that its hydrochloride

Scheme I[a]



[a] Pep = N-terminal portion of a peptide.

salt is stable and can be conveniently purified by sublimation. In the manner of other *O*-acylhydroxylamines, the neat free amine does isomerize to pivaloylhydroxamic acid within a few hours at room temperature, but solutions in chloroform are stable for longer than 1 month. The only demonstration of its nucleophilicity, however, was the formation of the substituted cyclohexanone oxime derivative in low yield. Therefore, as with the other derivatives studied in Table I, the first test of its usefulness was to determine its reactivity toward carbodiimide-activated acids.

Coupling Reactions with Model Compounds. Koshland[16] and his coworkers have demonstrated that water-soluble carbodiimides (WSC) can be used in the presence of nucleophiles to quantitatively modify carboxyl groups in proteins. We used two such carbodiimides with about equal success: 1-cyclohexyl-3-(2-morpholinoethyl)carbodiimide metho-*p*-toluenesulfonate (WSC-I) and 1-ethyl-3-dimethylaminopropylcarbodiimide (WSC-II). The latter was generally preferred only for reasons of economy. Using benzoic acid and hippuric acid as models, it was found that, in the presence of excess WSC and OPHA at pH 4.25, the desired *O*-pivaloylhydroxamates **11** were formed in 88 and 100% yields, respectively (eq 1). Although the conditions employed here were similar to those used by Koshland et al.,[16] much less carbodiimide and nucleophile were required, presumably because OPHA exists as the free amine under these conditions. Thus, despite its low pK_a (ca. 2),[17] OPHA, an α-effect nucleophile, was sufficiently nucleophilic to compete with hydrolysis for the carbodiimide-activated carboxylic acid. The fact that more than 1 equiv of carbodiimide is required indicates that hydrolysis of the active anhydride-like intermediate does occur. This observation is of no consequence, since the carboxylic acid is regenerated and can be reactivated by the excess WSC. Hydrolysis also avoids the slow accumulation of inactive *N*-acylureas **12** from O → N acyl migration in the

Table I. Rearrangements and Nucleophiles Considered for Carboxyl-Terminal Degradation

Rearrangement	Nucleophile	
1. Hofmann	NH_2Cl[a,b]	
2. Curtius	N_3^-[c]	
3. Wawzonek	$NH_2-\overset{+}{N}(CH_3)_3$[d,e]	$NH_2-\overset{+}{N}$[d,e]
	$NH_2-\overset{+}{N}(CH_3)_2C_6H_5$[d,e]	
4. Lossen	NH_2OSO_3H[a,b,d]	$NH_2OSO_2CH_3$[a,b]
	$NH_2-OSO_2-C_6H_4-CH_3$[a,f]	$NH_2-OSO_2-C_6H_2(CH_3)_3$[f]
	$NH_2-O-C_6H_3(NO_2)_2$[d,f]	$NH_2-O-C_6H_3(NO_2)-\overset{+}{N}(CH_3)_3$[g]
	$NH_2-S-C_6H_3(NO_2)$[d,e,f]	$NH_2-O-C(=O)-CH_3$[a]
	$NH_2-O-C(=O)-C_6H_5$[a,f]	$NH_2-O-C(=O)-C(CH_3)_3$ **10**

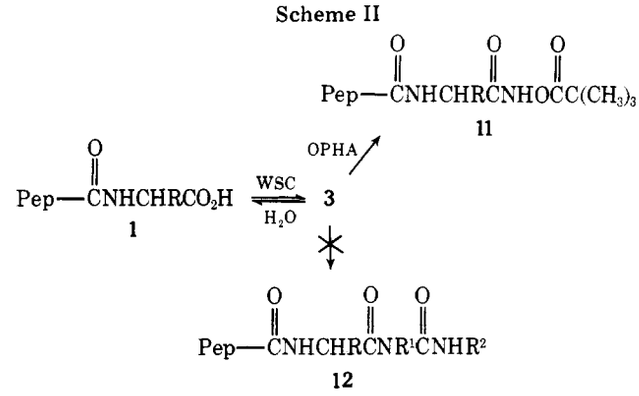
[a] Unstable. [b] Undesired reactivity (e.g., disproportionates in water to interfering products). [c] Acyl azides susceptible to hydrolysis. [d] Insufficiently nucleophilic toward carbodiimide-derived active esters. [e] Excessively vigorous rearrangement conditions required. [f] Water insoluble. [g] Could not be prepared. Although the immediate precursors were readily prepared, acid decomposition, even with nonnucleophilic acids, gave only the corresponding phenol.

$$RCO_2H + H_2NO_2CC(CH_3)_3 \xrightarrow[\text{pH 4.25}]{\text{WSC}} RCONHO_2CC(CH_3)_3 \quad (1)$$

10 **11**

activated intermediate. These considerations are shown in Scheme II.[18]

Scheme II



Lossen Rearrangement of Model Compounds. Titration of several *O*-pivaloylhydroxamate derivatives (Table II) revealed that the anion **5b** required to initiate the Lossen rearrangement could easily be formed at mild pH values. This result implied that no benefit can be obtained by performing the rearrangement at a pH higher than about 8.5. These titrations were also reversible, an observation which indicated that no rapid rearrangement or detrimental reaction with adjacent amide bonds occurs.

Hauser[19] has shown that the rate of rearrangement of *O*-acylbenzohydroxamic acids is related to the pK_a of the conjugate acid of the carboxylate leaving group. Pivalic acid has a relatively high pK_a (5.01).[20] Extrapolation of Hauser's data led to the prediction that the first-order rate constant, k_r, for rearrangement of *O*-pivaloyl benzohydroxamate (**13**) should be 1.8×10^{-4} min^{-1} at 30 °C. Using the data of Berndt and

Table II. Properties of *O*-Pivaloylhydroxamates $RCONHOCOC(CH_3)_3$

Registry no.	No.	R[a]	pK_a	Lossen rearrangement[b]	
				k, min^{-1}	Product
61650-22-4	13	Ph	7.4[c]	3.97×10^{-3}	$PhNH_2 + (PhNH)_2CO$ (trace)
61689-15-4	14	$PhCH_2$	7.25[c]	*e*	$(PhCH_2NH)_2CO$
61650-23-5	15	$PhCONHCH_2$	6.75[c]	1.41×10^{-3}	$(PhCONHCH_2NH)_2CO$
61650-24-6	16	AcNHCH $(CH_2)_2CONH_2$	6.0[d]	5.40×10^{-3}	*f*

[a] Abbreviations: Ph = phenyl, Ac = acetyl. [b] At pH 8.5 and 50 °C, monitored by standard base consumption using the pH stat; see the Experimental Section for other conditions. The rate constant k is the observed first-order rate constant for the rearrangement under these conditions. [c] Determined in 30 vol % aqueous ethanol by titration. [d] Determined in water by titration. [e] Not determined. [f] See text for discussion of products.

Schechter,[21] we estimated that the rate of rearrangement of **13** at 50 °C should be about ten tmes greater than the rate at 30 °C. Subjection of **13** to Lossen rearrangement conditions (pH 8.5, 50 °C) yielded aniline (eq 2) with $k_r = 3.97 \times 10^{-3}$

$$C_6H_5CO\bar{N}OCC(CH_3)_3 \longrightarrow C_6H_5N{=}C{=}O + {}^{-}O_2CC(CH_3)_3 \quad (2a)$$

$$C_6H_5N{=}C{=}O \xrightarrow{H_2O} C_6H_5NH_2 + CO_2 \quad (2b)$$

min^{-1} (Table II; half-life 174 min at 50 °C), a factor of 2 faster than predicted. These observations implied that under these conditions the rearrangement was 99% complete in 18 h. Although slower than one might consider ideal, this result indicated that the pivaloyl leaving group was usable, since the rearrangement can be monitored with a pH stat and left unattended.

The Lossen rearrangement is also accelerated by electron-donating substituents in the migrating group.[19,21,22] On this basis, it was expected that the rearrangement of the *O*-pivaloyl hydroxamates, **11**, would be faster than that observed for **13**, since the migrating group, which has a substituent selectivity similar to carbonium ion **18a**, would not only be secondary in all cases except glycine, but also would be stabilized as indicated in structure **18b** by the α-amido nitrogen.

$$\left[Pep{-}\overset{O}{\overset{\|}{C}}\overset{+}{N}HCHR\right] \longleftrightarrow \left[Pep{-}\overset{O}{\overset{\|}{C}}\overset{+}{N}H{=}CHR\right]$$

18a **18b**

To test this idea, the *O*-pivaloyl derivative of N^{α}-acetylglutaminohydroxamic acid, **16**, was subjected to the same rearrangement conditions. Indeed, its rearrangement was slightly faster than that observed for the benzo analogue **13** (Table II). Subsequent acid hydrolysis followed by analysis showed no glutamic acid (≤0.5%) remaining, but that conversion to the corresponding aldehyde **9** (succinic semialdehyde in this case) had occurred.

O-Pivaloylhippurohydroxamic acid (**15**) underwent relatively slow rearrangement (Table II) presumably because the migrating group is primary. However, because of the anticipated ease of product identification, this compound was chosen as a model peptide in order to study the rearrangement and its products in more detail.

The Fate of the C-Terminal Residue after Rearrangement. Several intermediates are produced during the Lossen rearrangement. The first, the isocyanate **6**, is not stable in aqueous solution but reacts with water to form a carbamate which in turn decarboxylates[47] to form the *N*-aminomethylamide **7**. A number of subsequent reaction paths can be envisioned for **7** in aqueous solution. The most desirable pathway would have been one which generates a new carboxyl terminus, since a sequential degradation procedure would then be possible. However, the product obtained upon rearrangement of **15** at an initial concentration of $>10^{-2}$ M (considerably higher concentration than that used in subsequent analytical procedures) was the symmetrical urea **17**. Its formation could easily be decreased to about 30% by subjecting **15** to rearrangement under very dilute conditions (ca. 10^{-4} M). In no case was benzoic acid detected after rearrangement, and controls established that 1% could easily have been detected. The formation of the symmetrical urea requires that the *N*-aminomethyl amide **7a** not only be reasonably stable under the rearrangement conditions, but also be nucleophilic enough to compete with a large molar excess of water for the isocyanate **6a** as it is generated (eq 3). *N*-Aminomethylamides such

$$\underset{\textbf{7a}\ (=\textbf{7},\ \text{with Pep} = C_6H_5)}{C_6H_5\overset{O}{\overset{\|}{C}}NHCH_2NH_2} \longrightarrow C_6H_5\overset{O}{\overset{\|}{C}}NH_2 + NH_3 + CH_2O \quad (3a)$$

$$\textbf{7a} + \underset{\textbf{6a}}{C_6H_5\overset{O}{\overset{\|}{C}}NHCH_2N{=}C{=}O} \longrightarrow \underset{\textbf{17}}{\left(C_6H_5\overset{O}{\overset{\|}{C}}NHCH_2NH\right)_2C{=}O} \quad (3b)$$

as **7**, in fact, are stable enough to be isolated under certain conditions, and have been shown to decompose to the corresponding amide and aldehyde only under more vigorous hydrolytic conditions.[8,23] Whereas aniline generated from the rearrangement of **13** has a pK_a of 4.6,[20] the pK_a of the *N*-aminomethylamide produced from **15** is about 7.5.[23] The difference in nucleophilicity of amines as reflected by their pK_a difference accounts for the products obtained under similar rearrangement conditions.[48] Wieland[7] also obtained symmetrical ureas during the Lossen rearrangement of *O*-benzoylpeptidohydroxamates; he, too, found that hydrolysis of these ureas generates an amide and an aldehyde (Wieland conditions 0.02 N HCl, 100 °C, 5–10 min).

The isocyanate could be trapped as it was formed by various added amine nucleophiles provided that the pK_a of the amine conjugate acid was greater than about 7. The hydrolysis of the ureas obtained yielded the same products as those obtained from the symmetrical urea **17**, except that the trapping amine was also obtained in these cases. The hydrolysis of these urea trapping products required equivalent hydrolytic conditions to those required for the hydrolysis of **17**. Significantly less than quantitative yields of these ureas were often obtained when excess nucleophile was added because of attack of the basic added amine at the pivaloyl carbonyl group resulting in formation of an unproductive, unsubstituted hydroxamic acid.

Table III. Effect of pH of the Coupling Step on the Overall Carboxyl-Terminal Degradation Yield with Peptides

Registry no.	Peptide (Ac = acetyl)	pH	Degradation yield, %
13716-72-8	Ac-Gly-Phe	4.0	83
		3.5	87
29852-55-9	Ac-Gly-Leu	4.5	83
		3.5	90
57282-76-5	Ac-Ala-Leu-Gly	4.75	51
		3.5	77

Thus, attempts to intercept the isocyanates with other nucleophiles were abandoned. In general, we found that these ureas could be hydrolyzed by Wieland's conditions (see above) or by reaction at pH 1–2 and 50 °C for several hours. In one case, stirring with an acidic resin (Amberlite IR-120) at room temperature also liberated the aldehyde without destruction of the remaining peptide bonds.[24]

Optimization of the Experimental Conditions with Peptides. After confirming the nature of the intermediates, investigations centered on maximizing the overall yield of the degradation.To avoid mechanical losses, the entire reaction sequence (activation, coupling, and rearrangement) was carried out with several peptides without isolation of intermediates. Subsequent direct hydrolysis of the reaction mixture and subtractive amino acid analysis indicated the extent of the loss of the carboxyl-terminal residue relative to others (degradation yield). In all cases, an aliquot of the hydrolysate was tested for the presence of aldehyde.[25] Using this general approach, conditions during both the coupling (i.e., *O*-pivaloyl hydroxamate formation) and rearrangement stages were varied to find the optimum.

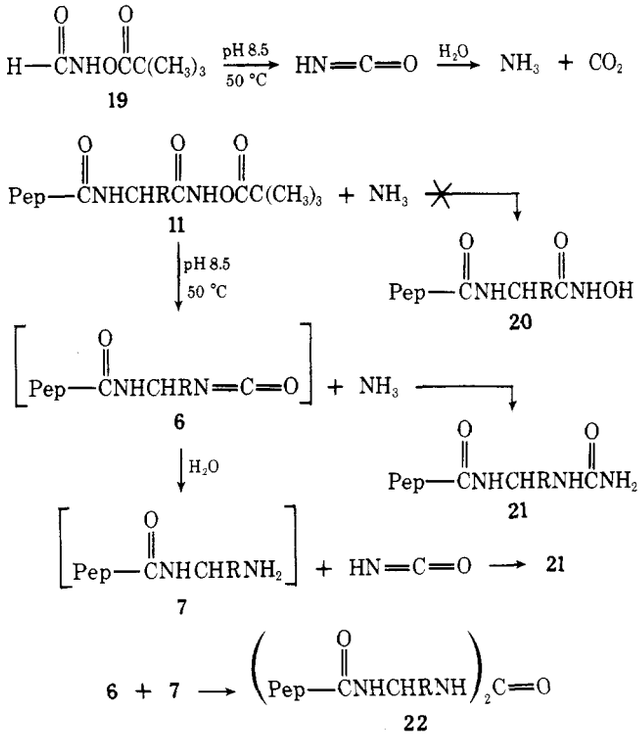
An estimate of the time and amount of carbodiimide required for the quantitative formation of the *O*-pivaloyl peptidohydroxamates was obtained by adding aliquots of a carbodiimide solution to a solution of peptide and excess OPHA at pH 4.5 while maintaining the pH with a pH stat. Once the free carboxyl groups are consumed by formation of the O-substituted hydroxamic acid, the addition of WSC has no effect, and no titrant is then consumed. Using these results, 1 h was considered ample time for the activation and coupling process.

As shown in Table III, variation of the pH during the coupling step also affected the overall degradation yield to a small extent. Based on the data shown in this table, pH 3.5 was chosen as the optimum pH for the coupling step. Hydrolysis at the pivaloyl carbonyl group of the *O*-pivaloyl hydroxamates was initially hypothesized to be the reason for this effect, but in no case was a hydroxamic acid detected with a $FeCl_3$ test. The ability of the anion of the O-substituted hydroxamic acid nitrogen to act as a nucleophile (see below) may be responsible for this effect, but the origin of the effect has not been definitively characterized.

To avoid spurious side reactions the excess WSC and OPHA were quenched by addition of excess formate buffer (pH 3.5) before initiation of the Lossen rearrangement. This addition results in the formation of a large amount of *O*-pivaloylformohydroxamic acid (**19**) (Scheme III). Thus, the rearrangement of **19** dominates the uptake of titrant on the pH stat; this titrant uptake then serves as an indicator for the extent of reacton. Based on this work and the previously described model experiments, rearrangement conditions of 20 h, 50 °C, and pH 8.5 were considered to be reasonable.

There was some concern that the ammonia generated from the rearrangement of **19** would attack the pivaloyl carbonyl group and regenerate the free hydroxamic acid **20** in a manner

Scheme III



similar to that observed for the added nucleophiles in the model studies (Scheme III). Under the conditions adopted, the ammonia concentration never approaches 0.1 M. Hauser[19] and Berndt[21] used 0.1 M ammonia solutions in their studies of the Lossen rearrangement and showed that no aminolysis at the acyl carbon of their leaving group occurred; the pivaloyl group should be even less susceptible to such aminolysis than the groups used in the studies just cited. Thus, no problems of this nature were encountered. One result of the fomate quenching step is presumably the generation of a mixture of ureas (Scheme III), but this is not detrimental since they are all susceptible to similar hydrolysis conditions. In fact, the only consequence of this procedure is a large ammonia peak obtained during amino acid analysis of the final reaction mixture.

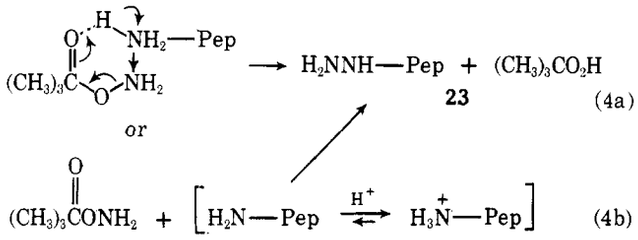
Results of the Degradative Procedure. The final procedure used in the degradation is presented in the Experimental Section. The results of a series of degradations of various peptides are presented in Table IV. In this table is given only the degradation yield of the carboxyl-terminal amino acid; this is the loss of the carboxyl-terminal amino acid relative to an internal residue known to be unaffected by degradation conditions. *Except as noted below,* other residues were unaffected and were recovered unchanged. Complete analytical data for the A and B chains of insulin are presented in Tables V and VI, respectively.

Interferences, Real and Potential. A. The Amino Terminus. Despite the putative quenching of OPHA after the coupling step, acetylation of the amino terminus of dipeptides prior to the degradation procedure was often found to increase the apparent degradation yield. Two factors were considered potentially responsible. Polymerization of the peptide in the presence of WSC during the coupling step would reduce the number of carboxyl groups available for degradation. The high concentration of OPHA relative to peptide, and the isolation of quantitative yields of coupled derivatives prior to Lossen rearrangement, showed that this problem was not significant under the conditions used. Since OPHA is known to be a poor electrophilic aminating agent,[15] it was considered possible that hydrazine derivatives **23** were being formed from the amino-

Table IV. Results of the Carboxyl-Terminal Peptide Degradation

Peptide[a]	% degradn	Registy no.	Peptide	% degradn	Registry no.
Ac-Gly-Asn	71	57282-73-3	Ac-Ala-Leu-Gly	77	
Ala-Ser	72	3303-41-1	Gly-Trp	86[c]	2390-74-1
Ac-Ala-Glu	41 (60)[b]	57282-74-3	Ac-Gly-Leu-Tyr	92	57282-77-6
Ac-Ala-Asp	42 (49)[b]	57282-75-4	Ac-Pro-Phe-Gly-Lys(Ac)	94	57282-78-7
Pro-Gly	77	2578-57-6	Ac-Met-Arg-Phe-Ala	92	57282-79-8
Gly-Met	81	554-94-9	Ac-Phe-Asp-Ala-Ser-Val	93	57282-80-1
Ac-Ala-Pro	82	41036-60-6	Ac-Leu-Trp-Met-Arg-Phe	94	57282-81-2
Gly-Phe	87	3321-03-7	Ac-γ-Glu-Cys(SO_3H)-Gly (oxidized acetylglutathione)	95 (Glu) 75 (Gly)	57601-76-0
Ac-Gly-Leu	90				
Bz-Gly-Arg	93	744-46-7	Insulin A-chain (C-terminal Asn)	75	
			Insulin B-chain (C-terminal Ala)	99	

[a] Abbreviations: Ac = acetyl, Bz = benzoyl. Yields are degradation yields. For detailed yields of each amino acid, see the data in Tables V and VI for the insulin chains. In the cases in this table, only the C-terminal amino acid was lost, *with the exceptions*, however, noted in the text. [b] The higher yields were obtained by performing the coupling step at a lower pH (0.75–1.5). [c] Determined using chloride-free conditions and the hydrolysis procedure of Liu and Chang;[30] see Experimental Section.



terminal group before the OPHA was quenched (eq 4). Such an eventuality would lead to a lower ratio of N-terminal to C-terminal amino acids in dipeptides, and therefore a lower apparent degradation yield. Studies with larger peptides, in which unaffected internal residues can serve as indicators, appeared to verify the existence of this possibility.

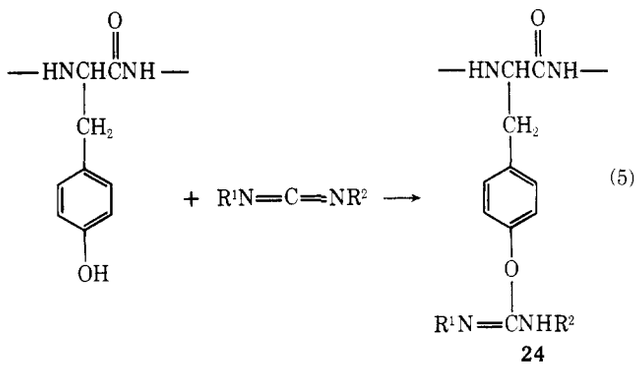
In a control, the tripeptide Gly-Leu-Tyr was subjected to the conditions of the normal rearrangement, except that WSC was omitted. Amino acid analysis after the hydrolysis revealed a 25% loss of N-terminal glycine. This loss was not observed when the peptide was acetylated prior to degradation.

In order to further clarify the nature of this reaction, phenylalanine was treated with OPHA for 21.5 h at pH 8.5, and α-hydrazino-β-phenylpropionic acid (identified by comparison with authentic material[26]) was detected in the resulting product mixture. Furthermore, the peptide Phe-Gln-Asn was treated with OPHA at pH 8.5 for 21 h and the resulting product was hydrolyzed in 6 N HCl. Thin layer chromatography of the hydrolysate showed clearly the presence of Asp, Glu, Phe, and α-hydrazino-β-phenylpropionic acid; amino acid analysis showed the loss of 37% Phe. (It should be observed that this loss is higher than that observed under analytical conditions because the formate quench was not used.)

Although the extent of N-terminal loss is small compared to the amount of C-terminal loss under degradation conditions, the N-terminal loss can be completely avoided by prior acetylation of the peptide. One consequence of acetylation, however, is the formation of *O*-acetylserine and *O*-acetylthreonine derivatives which undergo β-elimination (ca. 55% under the conditions of our degradation). The combination of degradations of acetylated and unacetylated peptides has, however, proven sufficient to eliminate ambiguities caused by these side reactions.

These results suggested that other residues within peptides might also be susceptible to loss by reaction with the electrophilic OPHA. Especially suspect were the basic residues histidine, lysine, and arginine. However, in only one case, lysine of the insulin B chain, was any significant effect noticed (44% loss); acetylation prevented this loss.

B. Reactive Aromatic Residues. Reactive aromatic residues also appearto be susceptible to destruction under the degradation conditions. Approximately 30% of tyrosine was consistently lost when this residue was internal (see, for example, Tables V and VI). Further studies showed the reasons for this loss to be complex. The peptide Gly-Leu-Tyr was treated with WSC-II at pH 3.5 for 1 h in the absence of OPHA. Amino acid analysis revealed 33% loss of tyrosine (peptide–peptide coupling, if it occurred, would have little effect on reactions at the side chain). The reaction with WSC-II is hardly surprising in view of the results of Carraway and Koshland,[27] who demonstrated the reaction of WSC and the tyrosine phenol (eq 5). The product **24** was shown[27] to be acid



stable under normal hydrolysis conditions preparatory to amino acid analysis. We were able to regenerate tyrosine from WSC-treated Gly-Tyr-NH_2 with hydroxylamine, albeit under more vigorous conditions than those used by Carraway and Koshland.[27]

Treatment of Gly-Leu-Tyr with OPHA at pH 8.5 for 20 h at 50 °C followed by amino acid analysis of the reaction mixture gave 12–15% loss of tyrosine. This loss is due to a reaction of unquenched OPHA with tyrosine either during the pH 8.5 incubation or during the 6 N HCl hydrolysis preparatory to amino acid analysis (OPHA is not removed prior to this hydrolysis). That the latter can contribute to tyrosine destruction was shown by combination of Gly-Leu-Tyr and 0.16 M OPHA in a hydrolysis tube in 6 N HCl followed by the usual 22-h hydrolysis at 110 °C; 23% tyrosine was lost. The presence of 20 mg of sodium dithionite in the hydrolysis tube of an identical duplicate sample prevented tyrosine loss, evidently by destruction of OPHA. Thus, addition of dithionite prior

Table V. Carboxyl-Terminal Degradation of the Oxidized A Chain of Insulin[a]

Amino acid	Before degradn		After degradn
	Theor	Obsd[b]	
Asx[c]	2	1.99	1.24
DAPA[d]	0	0	0
Asx + DAPA	2	1.99	1.24
Ser	2	1.93	1.89
Glx	4	4.04	3.56
DABA[d]	0	0	0.19
Glx + DABA	4	4.04	3.75
Gly[e]	1	1.01	0.67
Ala	1	1.00	1.00
Cys-SO_3H	4	4.07	3.91
Val	2	1.41[f]	1.46[f]
Ile	1	0.47[f]	0.49[f]
Leu	2	1.92	2.07
Tyr	2	1.80	1.62

[a] Degradation carried out in 8 M urea. [b] Average of two analyses; values calculated relative to alanine. [c] Asn is the C-terminal residue. [d] DAPA = diaminoproprionic acid; DABA = diaminobutyric acids; these arise from the Lossen rearrangement of the side chain carboxylic acid groups of aspartic and glutamic acids, respectively. [e] The N-terminal amino acid. [f] These values are low because the Ile-Val sequence is incompletely hydrolyzed in 22 h at 110 °C.[49]

to 6 N HCl hydrolysis eliminates any destruction of tyrosine by OPHA during hydrolysis. When OPHA and Gly-Leu-Tyr were incubated at 50 °C, pH 8.5, for 20 h with addition of dithionite preceding amino acid analysis, 8% tyrosine was still lost. This is thus the amount of tyrosine loss which can be attributed to reaction with OPHA over the long reaction period at pH 8.5.

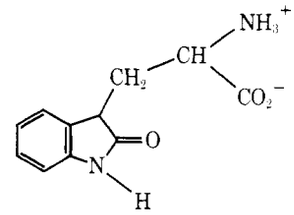
The reaction(s) of tyrosine with OPHA under the various conditions are not all clear. Several small new peaks in the amino acid analysis appear following these treatments (these are not found when OPHA alone is hydrolyzed), and an undefined electronic absorption out to 500 nm is observed (the solutions are yellow after the pH 8.5 treatment). *All* destruction of tyrosine from every source, including eq 5, is eliminated by O-methylation of tyrosine in Gly-Leu-Tyr. This observation is completely consistent with destruction of this residue by eq 5 and with oxidative processes on tyrosine by OPHA involving the phenolic hydroxyl, some occurring during pH 8.5 treatment, and others occurring during 6 N HCl hydrolysis. It is worth noting that O-methylation protects tyrosine from oxidation by *N*-bromosuccinimide,[28,29] and oxidation of tyrosine in proteins and model systems by Fremy's salt (potassium nitrosodisulfonate) is accmpanied by the production of several new species and long wavelength absorption.[31,32] It is clear, however, that the major loss of tyrosine is due to eq 5, which cannot be avoided if carbodiimides are to be used. Although prolonged incubation of tyrosine-containing peptides with hydroxylamine before analysis, removal of hydroxylamine, and amino acid analysis in the presence of dithionite would undoubtedly eliminate most of the tyrosine loss, under the analysis conditions the loss of tyrosine which is observed does not in practice interfere with the identification of the carboxyl-terminal residue, so these precautions were not routinely taken.

Using the procedure of Liu and Chang[30] and avoiding chloride ion to obviate destruction of tryptophan during acid hydrolysis preparatory to amino acid analysis, it was found that up to 50% of internal tryptophan residues are also destroyed by the degradation conditions. The nature of this reaction was determined to be an OPHA-induced oxidation

Table VI. Carboxyl-Terminal Degradation of the Oxidized B Chain of Insulin[a]

Amino Acid	Before degradn		After degradn
	Theor	Obsd[b]	
Asx	1	1.09	1.04
DAPA[c]	0	0	0
Thr	1	0.91	0.91
Ser	1	1.06	1.04
Glx	3	3.26	2.41
DABA[c]	0	0	0.41
Glx + DABA	3	3.26	2.82
Pro	1	1.01	1.01
Gly	3	3.00	3.00
Ala[d]	2	*1.98*	*1.01*
Cys-SO_3H	2	2.01	2.05
Val	3	2.93	2.87
Leu	4	3.80	3.68
Tyr	2	1.53	0.93
Phe[e]	3	2.72	2.69
His	2	1.62	1.52
Lys	1	0.86	0.47
Arg	1	0.84	0.86

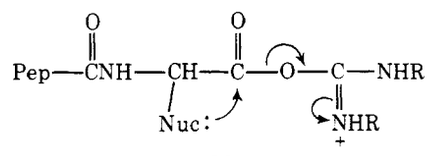
[a] Degradation carried out in 8 M urea. [b] Average of two analyses; values calculated relative to glycine. [c] DAPA = diaminopropionic acid; DABA = diaminobutyric acid. [d] The C-terminal residue. [e] The N-terminal residue.



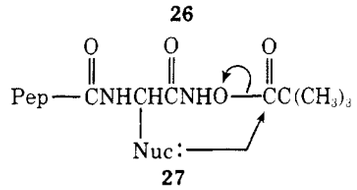
25

of tryptophan to oxindole **25.** This derivative has been synethsized and shown[33] to be clearly identifiable on amino acid analysis as a peak occurring just prior to tryptophan on the short column of the analyzer; a similar peak was noted by Moore and Stein[34] for hydrolysis of tryptophan-containing peptides in 6 N HCl. This peak was clearly identifiable in the amino acid analyses of reactions of tryptophan-containing peptides with OPHA, and was identical with the peak observed when **25** was generated as described by Ohta.[33] Although the 50% loss of tryptophan is serious, internal tryptophan is still distinguishable (albeit with less certainty) from carboxyl-terminal tryptophan, which is degraded virtually quantitatively.

C. Carboxyl-Terminal Residues with Functionalized Side Chains. Potential interferences by the reactions of proximal nucleophiles with the intermediate *O*-acylisourea to give a cyclic derivative **26** during the activation step were also considered. The reaction of serine, threonine, and cysteine



26



27

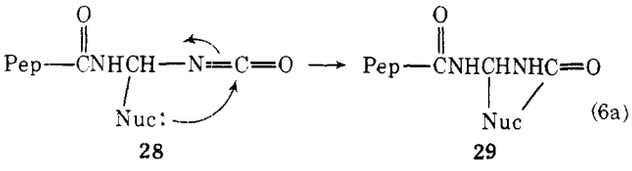
Table VII. Degradation of Carboxyl-Terminal Dicarboxylic Acids (Asp and Glu) in Acetylalanylaspartic Acid and Acetylalanylglutamic Acid

Coupling pH	Rearrangement (pH 8.5, 50°C, 20 h)	
	% Asp (Glu) lost[a]	% DAPA (DABA)[b] found[a]
3.5	42 (41)	0 (15)
1.5	44.5 (60)	0 (9)
0.75	49 (c)	0 (c)

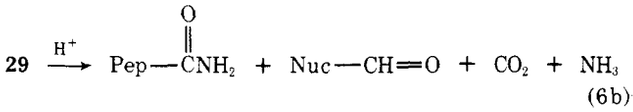
[a] Determined by difference amino acid analysis relative to alanine. [b] Abbreviations: DAPA = 2,3-diaminopropionic acid, **44a**; DABA = 2,4-diaminobutyric acid, **44b**. [c] Not determined.

in this manner would give a four-membered ring, and lysine would yield a seven-membered cycle; these rings were not expected to form readily. Histidine would yield an acylimidazole which would be susceptible to reopening. Aspartic and glutamic acids would yield anhydrides, but these would also be expected to be susceptible to reopening by nucleophilic attack of OPHA (but see below). Attack of nucleophilic side chains at the pivaloyl carbonyl, as in **27**, was not expected, since large rings would be formed in all cases.

Nucleophilic side chains might also react with the isocyanate **6** generated during the Lossen rearrangement. However, hydrolysis of the resulting products prior to amino acid analysis would still result in the formation of the aldehyde with net loss of the carboxyl-terminal residue as shown in eq 6. The high degradation yields obtained with various peptides[1,36] indicate that none of the side reactions indicated above were significantly detrimental if they occurred at all.



(6a)

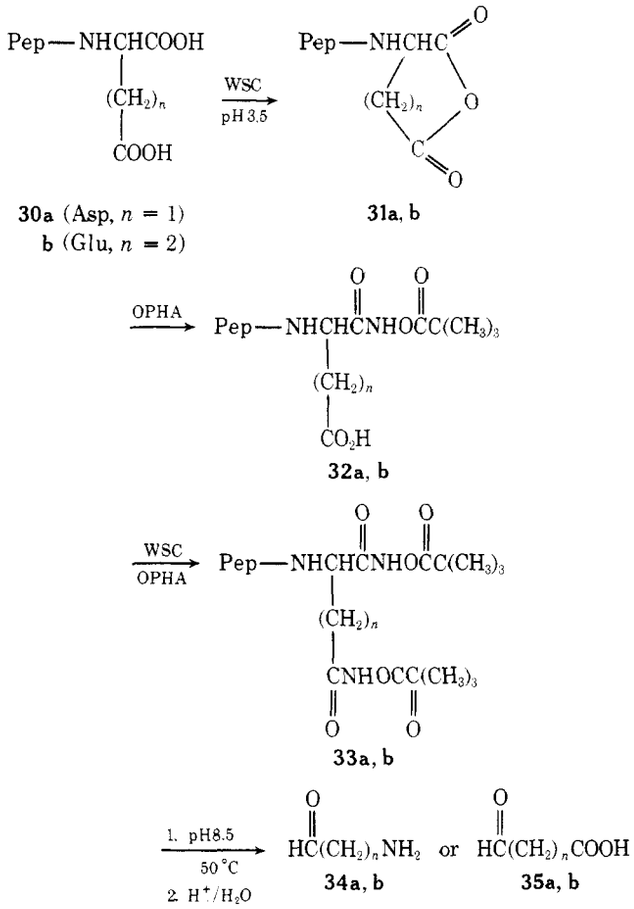


(6b)

D. Carboxyl-Terminal Aspartic and Glutamic Acids. As suggested previously, treatment of the carboxyl-terminal dicarboxylic acid residues with WSC under the coupling conditions was expected to generate cyclic anhydrides; reaction of these with OPHA and reaction of the remaining carboxyl group with more WSC and OPHA was then expected to give the dihydroxamic acid derivative **33** according to Scheme IV.[35] Rearrangement of the α-carbonyl group of **33** followed by hydrolysis would eventually give an aldehyde **34** or **35**, or the cyclic internal Schiff base of **34**. Rearrangement at the ω position was expected to be considerably slower because of migratory aptitude considerations (see below); if it did occur, it would generate an amine **34**; if not, subsequent hydrolysis would yield the semialdehyde **35**. In either case, loss of the C-terminal aspartic or glutamic acid was expected. In actual practice, however, only 40% of the C-terminal aspartic and glutamic acid residues could be removed under normal conditions (Table VII). The reason for these observations was revealed by the following model studies.

Subjection of *N*-acetylaspartic acid and *N*-acetylglutamic acid to the coupling conditions with WSC yielded only the cyclic imides **39** (Scheme V).[35] Hydrolysis of **39** in 6 N HCl, of course, merely regenerates the starting acids. Under the rearrangement conditions, however, hydrolysis was expected at positions of **39** corresponding to pathways a, b, and c of

Scheme IV [35]

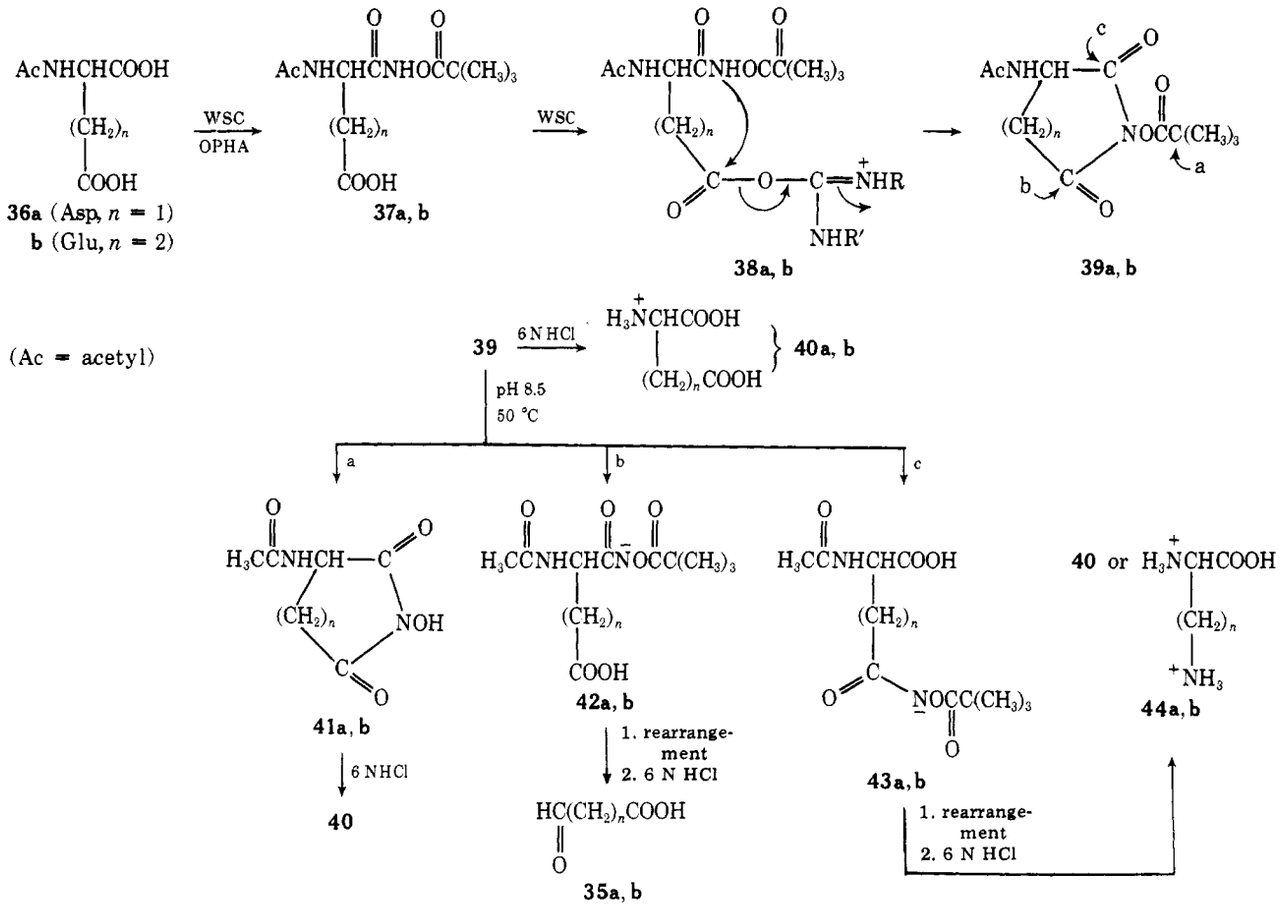


Scheme V. By analogy to the use of *N*-hydroxysuccinimide for active ester formation in peptide synthesis,[50] hydrolysis at position a was expected to prevail, except for the steric considerations introduced by the adjacent *tert*-butyl group. Such hydrolysis produces the *N*-hydroxyimide **41** which, upon subsequent acid hydrolysis prior to amino acid analysis, regenerates the starting amino acid with no net loss. Hydrolyses at carbonyl groups similar to b and c have also been reported.[51] Hydrolysis at b generates **42** which can undergo the normal Lossen rearrangement with eventual net loss of the carboxyl-terminal dicarboxylic amino acid. Hydrolysis at c gives **43** which again, because of migratory aptitude considerations, would be expected to undergo little or no rearrangement under the usual conditions. That rearrangement which does occur would produce the diaminocarboxylic acid **44** with a net loss of the carboxyl-terminal residue. If the rearrangement did not occur, subsequent analysis would regenerate the carboxyl-terminal residue **40**.

To test these possibilities the cyclic glutarimide derivative **39b** was subjected to the rearrangement conditions. Hydrolysis followed by amino acid analysis revealed both glutamic acid and 2,4-diaminobutyric acid (DABA), **44b**, with the ratio Glu/DABA = 3.73. Since this molecule contains no internal standard for use in amino acid analysis, the absolute amount of glutamic acid was not determined. The results indicate that path c of Scheme V does occur, but not its extent.

To get an accurate estimate of the loss of the C-terminal dicarboxylic acids, and the extent to which each pathway in Scheme V, a, b, or c, is significant, Ac-Ala-Glu and Ac-Ala-Asp were subjected to the same coupling and rearrangement conditions. The first entry of Table VII gives the results. Of the 41% of Glu lost, 15% appeared as DABA. Thus, a maximum of 59% of the reaction followed path a, a minimum of 15% followed path c, and 26% followed path b. In no case was di-

Scheme V

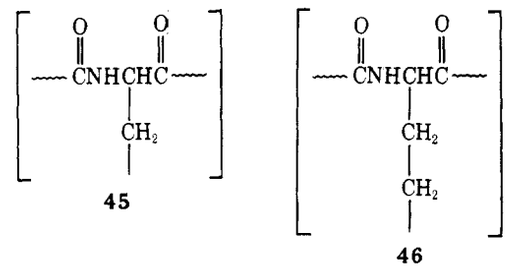


aminopropionic acid (DAPA, **44a**) detected from the same type of experiment with Ac-Ala-Asp. Assuming that all the hydrolysis pathways a, b, and c occur as indicated by the Ac-Ala-Glu experiment, with the exception that **43a** does not rearrange under these conditions, then 58% of the hydrolysis of **39a** must occur via paths a and c, whereas 42% goes by path b.

Lowering the pH of the coupling step was expected to decrease the detrimental cyclization to **39** by decreasing the already small amount of the anionic form of the amide nitrogen in **40**, thus promoting the sequence shown in Scheme IV. Table VII illustrates partial fulfillment of these expectations.

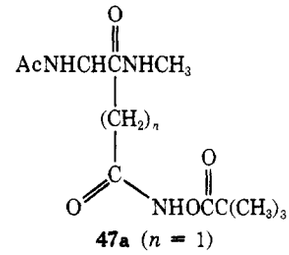
The low degradation yields of C-terminal Asp and Glu due to the above reactions represent one serious difficulty with the method; even so, indications that these residues are carboxyl-terminal are provided by their partial degradation; such results in unknown peptides would best be accompanied by other evidence. It should be noted that Asn and Gln undergo normal degradation; the reasons for the differences between amidated and nonamidated dicarboxylic amino acids are readily understood from the above chemistry.

E. Internal Aspartic and Glutamic Acid Residues. The ω-carboxyl groups of aspartic and glutamic acid residues within the peptide chain are, of course, also converted to the corresponding *O*-pivaloyl hydroxamates during the coupling step. After subjection to rearrangement conditions, however, in no case was a loss of aspartic acid or its conversion to DAPA, **44a**, observed; however, up to 20% of the internal glutamic acid residues are converted to DABA, **44b**. Again, the results appeared to be readily explained by the differences in migratory aptitude of the groups involved. The ω-migrating groups **45** and **46** are not only primary, but they have no possible stabilizing contributions such as those shown in structure **18**. In



fact, in these cases the strong electron-withdrawing effect of the nearby amide nitrogen results in further destabilization of the "carbonium ion-like" migrating groups. Such an inductive effect should reduce the migratory aptitude of **45** more than that of **46**.

Whereas *O*-pivaloyl N^{α}-acetylglutaminohydroxamate (**16**) and glutathione served as models for the ω-linked (free α-carboxyl) dicarboxylic amino acid residues, the only studies with α-linked (free ω-carboxyl) dicarboxylic amino acid residues were those involving complex peptides. To substantiate the results found with larger peptides, model compounds **47a,b** were synthesized from the corresponding N^{α}-acetyl-ω-benzyl esters of the corresponding amino acids by a



47a (n = 1)
47b (n = 2)

straightforward route detailed in the Experimental Section. Compounds **47a** and **47b** were, in respective separate experiments, subjected to the rearrangement conditions at pH 8.5 and 50 °C, hydrolyzed, and subjected to amino acid analysis. Under these conditions, **47a** showed no loss (<2%) of aspartic acid relative to methylamine, and glutamic acid was lost from **47b** to the extent of 20%, thus verifying the observations with peptides. These observations have been found to be the basis of a method for the determination of β-aspartyl and γ-glutamyl linkages in peptides.[36]

Results with Peptides. Summary. The method developed here successfully removes all residues encountered under relatively mild conditions, with the exception of carboxyl-terminal aspartic and glutamic acids. Peptides as large as the A and B chains of insulin were successfully degraded, and it was found that 8 M urea could be used as solvent. The interferences are the low (40–60%) degradation yields of C-terminal Asp and Glu, the partial loss of internal tyrosine (ca. 30%) and tryptophan (ca. 50%), and the amination of the amino-terminal residue (25%), with the latter avoided completely by acetylation of the peptide. If, after the normal degradation procedure, one suspects the possibility of C-terminal tryptophan, repetition of the procedure with the slight alterations described in the Experimental Section may be carried out.

The detailed analysis of the interferences provided here should not obscure the fact that the degradation is a simple, one-vessel procedure of great simplicity, and that most of the interferences do not contravene the method's great utility. The success of the procedure with larger peptides, the use of aqueous solvents, the high yields, the successful degradation of Asn, Pro, and Gln, and the overall simplicity of the method make it practical for general use. The method relies on difference amino acid analysis, and is therefore restricted to peptides of a size for which this technique is applicable. Many proteolytic and degradation peptides derived from sequencing studies as well as peptide hormones are of a size such that this method can be applied. The method can be used on a scale as small as one's analysis facilities will permit.

Experimental Section

Amino acid analyses were performed on a Beckman Model 116 or 120C automatic amino acid analyzer. Melting points were determined on a Büchi melting point apparatus, and are uncorrected. ^{1}H NMR spectra were taken at 60 MHz with a Varian A-60A instrument. Mass spectra were taken with an AEI MS 902 instrument, and elemental analyses were performed by Galbraith Laboratories, Inc. Titrations and other operations involving a pH stat were performed with a Radiometer TTT-1c instrument.

The carbodiimides used in this work (WSC-I and WSC-II) were obtained from Sigma and used without further purification. All peptides were commercially available. Insulin (Sigma) was oxidized according to the method of Sanger[38] and the chains separated on a 1 × 10 cm A-25 DEAE-Sephadex column by the method of Fittkau.[39]

***O*-Pivaloylhydroxylamine** (OPHA, **10**) was prepared as the hydrochloride salt by the method of Marmer and Maerker[15] using either *tert*-butyl *N*-hydroxycarbamate or ethyl *N*-hydroxyacetimidate as the precursor; the latter was preferred. After sublimation (60–70 °C, 15–20 Torr) the product, mp 120–122 °C, could be stored for months in the freezer (−26 °C).

General Procedure for the C-Terminal Degradation of Peptides. In a typical experiment 154 mg (1 mmol) of OPHA·HCl (**10** HCl) was dissolved in 4 mL of water, and the solution was adjusted to pH 3.5 with about 1 mL of 1 N NaOH. A solution of the peptide, previously amino acid analyzed, in 2–5 mL of H_2O was added. A 0.4-mL aliquot of a 0.5 m aqueous solution of WSC-II, 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride, was added immediately. The pH was maintained at 3.5 with 0.1 N HCl using the pH stat. Three more identical aliquots of the carbodiimide solution were added at 15-min intervals. The addition of 0.5 mL of a 5.0 N pH 3.5 formate buffer and stirring for 20 min quenched the carbodiimide and OPHA by formation of *O*-pivaloylformohydroxamic acid and hydrolysis of the excess carbodiimide. The Lossen rearrangement was initiated by adjustment of the pH to 8.5 and the temperature to 50 °C; the pH was then maintained with the pH stat using 1.0 N NaOH. Within 20 h the base consumption was complete. A portion of the reaction mixture was then made 6 N in HCl, hydrolyzed directly, and amino acid analyzed.

Tryptophan analysis by the method of Liu and Chang[30] requires the absence of chloride ion. Consequently, an aqueous solution of free OPHA was obtained by addition of 154 mg (1 mmol) of OPHA·HCl to a solution of 2.5% $NaHCO_3$, extraction with ether, drying of the combined ether layers, concentration to about 5 mL, addition of 5 mL of water, and evaporation of the remaining ether. The solution was then used in the previously described procedure after changing the titrant to 0.1 N H_2SO_4 and the carbodiimide to 1-cyclohexyl-3-(2-morpholinoethyl)carbodiimide metho-*p*-toluenesulfonate (WSC-I). Hydrolysis prior to amino acid analysis was performed with 4 N methanesulfonic acid containing 2% tryptamine,[46] or with 3 N mercaptoethanesulfonic acid (Pierce).

***O*-Pivaloylbenzohydroxamic acid** (**13**) was prepared by two independent routes.

A. Benzoic acid (2 mmol) was dissolved in 35 mL of water. OPHA·HCl (0.34 g, 2.2 mmol) was added and the pH was adjusted to 4.5 with 1.0 N NaOH solution. A 0.33 M aqueous solution of 1-cyclohexyl-3-(2-morpholinoethyl)carbodiimide metho-*p*-toluenesulfonate (WSC-I, 6 mL) was added and the pH maintained at 4.5 with the pH stat using 1.0 N HCl. Further aliquots were added as required until no more acid consumption was observed upon addition; the total amount of WSC-I used was 4.6 mmol. The resulting white precipitate was filtered, washed with water, and air dried to afford the desired product in 88% yield: mp 108–110.5 °C; NMR δ_{CDCl_3} (downfield from Me_4Si) 1.33 (s, 9 H), 7.2–7.95 (m, 5 H), 9.66 (s, 1 H, broad); pK_a (30% ethanol) = 7.4.

Anal. Calcd for $C_{12}H_{15}NO_3$: C, 65.14; H, 6.83; N, 6.33. Found: C, 65.10; H, 6.76; N, 6.29.

B. Benzohydroxamic acid (20 mmol)[40] was dissolved in 150 mL of acetonitrile. One equivalent of triethylamine was added and the solution stirred at room temperature while 20 mmol of pivaloyl chloride was added dropwise in 25 mL of dry acetonitrile. After addition, the mixture was stirred for an additional 1 h, then filtered to remove the triethylamine hydrochloride. Evaporation of the filtrate yielded a white solid which, after recrystallization from ethyl acetate–hexanes, gave a material with analytical and spectral properties identical with those obtained under A, in 90% yield.

***O*-Pivaloylphenylacethydroxamic acid** (**14**) was prepared in 80% yield by method B: mp 98–99.5 °C; NMR δ_{CDCl_3} 1.28 (s, 9 H), 3.11 (s, 2 H), 7.32 (s, 5 H), 8.95 (s, broad, –NH–); pK_a (30% ethanol) = 7.25.

***O*-Pivaloylhippurohydroxamic acid** (**15**) was prepared by both method A in quantitative yield and a modification of method B using water instead of acetonitrile and sodium carbonate rather than triethylamine. Careful acidification to pH 3 after the reaction precipitated the product, which was filtered and recrystallized from ethyl acetate–hexanes (64%): mp 119.5–121.5 °C; NMR δ_{CDCl_3} 1.28 (s, 9 H), 4.25 (d, 2 H), 7.2–7.5 (m, 6 H), 10.0 (s, broad, –NH–); pK_a (30% ethanol) = 6.7; mass spectrum (chemical ionization with CH_4, giving P + 1) *m/e* 279.

Anal. Calcd for $C_{14}H_{18}N_2O_4$: C, 60.42; H, 6.52; N, 10.07. Found: C, 60.31; H, 6.43; N, 9.89.

***O*-Pivaloyl-N^{α}-acetylglutaminohydroxamic acid** (**16**) was prepared in 86% yield by method A: mp 153.5–155 °C; NMR $\delta_{(CD_3)_2S=O}$ 1.24 (s, 9 H), 1.86 (s, 3 H), 1.7–2.2 (m, 4 H), 4.15–4.39 (m, 1 H), 6.70–7.4 (broad, NH_2), 8.1 (broad, –NH–), 9.4 (broad, –NH–); pK_a = 6.0.

Anal. Calcd for $C_{12}H_{21}N_3O_5$: C, 50.17; H, 7.37; N, 14.63. Found: C, 50.18; H, 7.49; N, 14.47.

Lossen Rearrangements. *O*-Pivaloyl Benzohydroxamate (**13**). Compound **13** (0.1 mmol) was suspended in 150 mL of water maintained with an oil bath at 50 °C. The pH was adjusted to 8.5 and maintained with the pH stat using 0.1 N NaOH. First-order consumption of approximately 1 equiv of base was observed (Table I). After 18 h, 10 μL of benzylamine was added as an internal standard. The pH was adjusted to 11.5 and the solution was immediately extracted with three 25-mL portions of ether. The combined ethereal extracts were dried ($MgSO_4$), filtered, and evaporated to 0.5 mL. Gas chromatographic analysis (0.125 in. × 8 ft 20% Carbowax) at 170 °C indicated a quantitative yield of aniline within analytical error (±5%). However, complete evaporation of the ethereal portion left a visible film which was identified by comparative thin layer chromatography as diphenylurea (silica gel, methanol–ethyl acetate, 1:9, R_f 0.67).

***O*-Pivaloyl phenylacethydroxamate** (**14**) was subjected to the same rearrangement conditions, except that the solvent was 33 vol

% acetonitrile in water, and the reaction was allowed to continue for about 20 h. Workup afforded a 92% yield of dibenzylurea, mp 170–171.5 °C (lit.[41] 171 °C). A trace of benzylamine (<1% by GC) was also obtained.

***O*-Pivaloyl hippurohydroxamate (15)** was subjected to the same conditions used for the rearrangement of **13** for 36 h. First-order consumption of base was observed (Table I), and a precipitate formed which was filtered to afford *N,N'*-bis(benzoylaminomethyl)urea in 63% yield, mp 247–249 °C (lit.[7] 247 °C). One equivalent of phenylacetic acid was added as an internal standard. The pH was adjusted to 3 and the solution extracted with ether. The ethereal portion was dried and treated with diazomethane generated from *N*-nitrosomethylurea and KOH. Concentration of the ethereal solution followed by gas chromatography gave methyl phenylacetate and methyl pivalate, but no methyl benzoate. Performing the same reaction on a larger scale at $>10^{-2}$ M for 36 h gave the corresponding symmetrical urea in 92% yield. At 10^{-4} M, a 30% yield of this compound was obtained.

***O*-Pivaloyl N^{α}-Acetylglutaminohydroxamate (16).** In 15 mL of water at 50 °C was dissolved 0.1 mmol of **16**. The pH was adjusted to 8.5, and maintained with the pH stat using 0.1 N NaOH. A first-order base consumption was observed (Table I) which ceased within 12–15 h. An aliquot of the reaction mixture corresponding to 5 μmol of **16** was removed, added to a standard glycine solution, made 6 N in HCl, hydrolyzed, and subjected to amino acid analysis in the customary manner. The analysis indicated about 0.6% glutamic acid remaining. The remaining aqueous solution was made 6 N in HCl and heated to reflux. After 1 h, a 0.5-mL aliquot was removed, made basic by the addition of 25% NaOH, and tested for aldehyde by the method of Dickinson.[25] A positive test was obtained. After refluxing for an additional 2 h, 2.75 mL of a 0.8% solution of 2,4-dinitrophenylhydrazine in 1 N HCl (warmed to make it homogeneous) was added to the remaining aqueous solution and stirred at room temperature for 3.5 h. The solution was evaporated to dryness to give an orange solid. A solution of 2.5% $NaHCO_3$ (25 mL) and ethyl acetate (15 mL) were added, and the layers separated. After extraction of the aqueous layer with more ethyl acetate, the aqueous solution was then acidified to pH 1–2 and extracted with ethyl acetate. The combined ethyl acetate layers from the latter extraction were dried, filtered, and evaporated to yield the 2,4-dinitrophenylhydrazone of succinic semialdehyde in 77% yield, mp 198–200 °C (lit.[42] 202–203 °C); a mixture melting point with 2,4-dinitrophenylhydrazine was depressed (167–174 °C).

Unsymmetrical Urea Formation. Isocyanate trapping experiments were performed by adding an amine nucleophile (CH_3NH_2, $C_6H_5CH_2NH_2$, $(CH_3)_2NNH_2$, glycine, and others, 1–30 equiv) during the Lossen rearrangement of **15** or other peptides. The unsymmetrical ureas were obtained by filtration or ethyl acetate extraction of the aqueous solution and were characterized by NMR. In those cases in which the pK_a of the conjugate acid of the added amine was low (2-aminopyridine, $pK_a = 6.7$, or 2-aminothiazole, $pK_a = 5.4$) the symmetrical urea **17** formed exclusively. The yield of the symmetrical urea was inversely related to the concentration of the added nucleophile, a fact indicating inhibition of the rearrangement by attack of the excess nucleophile at the pivaloyl carbonyl group. Positive tests for free hydroxamic acid were obtained in these cases by adding an acidified aliquot of the reaction mixture to an aqueous solution of 3% $FeCl_3$.

Urea Hydrolysis. In general the ureas **17, 21, 22,** and others could be hydrolyzed using Wieland's conditions (0.02 N HCl, 100 °C 5–10 min)[7] or by reaction at pH 1–2 and 50 °C for several hours. In all cases amide **8** (determined by comparative thin layer chromatography) and the aldehyde **9** (determined by paper chromatography of the 2,4-dinitrophenylhydrazone[37] or by the general test of Dickinson[25]) were obtained. Stirring the urea with the acidic resin Amberlite IR-120 at room temperature for 2 h also gave the same products.

Controls. A. For Unprotected Amino-Terminal Residue and Tyrosine Loss. The tripeptide Gly-Leu-Tyr (4 mg) was subjected to the normal degradation procedure, but without addition of the carbodiimide. Hydrolysis followed by amino acid analysis gave a 26% loss of glycine and a 32% loss of tyrosine. Aliquots taken from the degradation at various stages for amino acid analysis were hydrolyzed with and without ca. 20 mg of sodium dithionite in the hydrolysis tubes as described in the text.

B. Prevention of N-Terminal Residue and Lysine Loss by Acetylation. Peptides were acetylated in a manner similar to that used by Stark.[6] The peptide was dissolved in 5 mL of acetic acid or 50% aqueous acetic acid and 5 mL of distilled acetic anhydride added (or the anhydride was added until the solution just became turbid). The solution was heated at 50 °C for 4 h, diluted with one volume of water, and allowed to stand for 1 h to quench the acetic anhydride as well as any mixed anhydrides formed from peptide carboxyl groups. The resulting solution was evaporated, redissolved in water, and freeze dried. The effect of acetylation is clearly demonstrated by the following example. Prior to acetylation, subjection of the tetrapeptides Pro-Phe-Gly-Lys to the normal degradation procedure gave a 16% loss of N-terminal proline in addition to an expected 93% loss of lysine. Following acetylation and repetition of the degradation of acetylated material, 94% loss of lysine was observed, but no loss of proline took place. In a control experiment acetylated peptide was subjected to the conditions of the degradation but in the absence of OPHA; no proline or lysine was lost.

C. Effect of Acetylation on Serine Residues. This effect was investigated using the carboxyl-terminal pentapeptide of ribonuclease, Phe-Asp-Ala-Ser-Val, as a model. Under the normal conditions valine was removed in 89% yield, and no serine was lost. After acetylation 93% of the valine was lost, but 54% of the serine was also lost. Subjection of the acetylated pentapeptide to the rearrangement conditions alone (pH 8.5, 50 °C, 20 h) followed by amino acid analysis gave a 56% loss of serine without affecting any other residues. A corresponding amount of ammonia was obtained (Ser + NH_3 = 1.04 residues), an observation which is consistent with elimination to dehydroalanine, which is, in turn, hydrolyzed to pyruvic acid and ammonia.

Reaction of OPHA with Free Amino Groups. L-**α-Bromo-β-phenylpropionic acid** was prepared by the method of Maimind et al.[43] and was converted to **D-α-hydrazino-β-phenylpropionic acid (48)** by te method of Niedrich and Grupe,[26] mp 181–185 °C (lit.[26] 185–190 °C); recrystallization from water gave material with mp 208–212 °C $pK_a = 7.7$. Direct elution on the long column of the amino acid analyzer showed the presence of no α-amino acid.

In 4 mL of water in a jacketed pH stat titration vessel was dissolved 0.164 g (0.99 mmol) of phenylalanine. A solution of 0.156 g (1.01 mmol) of OPHA·HCl in 4 mL of water was added and the pH and temperature were adjusted, respectively, to 8.5 and 50 °C. The pH was held constant for 21.5 h by titration with 1 N NaOH. The solution was neutralized and lyophilized for analysis. Thin layer chromatography (TLC) was carried out on analytical silica gel plates by spotting the reaction mixture, phenylalanine, and **48**, spraying the spots with Ehrlich's reagent, and elution with 1-butanol–pyridine–water (6:4:3). The hydrazone of **48** (R_f 0.83) and phenylalanine (R_f 0.50) were clearly separated by this method. The prior spraying with Ehrlich's reagent was found to yield a much superior separation without streaking, in marked contrast to a number of separations attempted without this modification.

Reaction of Phe-Gln-Asn with OPHA. In 4 mL of 1:1 (v/v) acetonitrile–water was dissolved 81 mg (0.2 mmol) of Phe-Gln-Asn, and the pH and temperature were adjusted, respectively, to 8.5 and 50 °C. A solution of 154.8 mg (1 mmol) of OPHA·HCl in 4 mL of water was added, and the pH and temperature were maintained for 21 h. The solution was placed in a hydrolysis tube, mixed with concentrated HCl to a final HCl concentration of 6 N, and hydrolyzed for 21 h at 110 °C. Small samples of the peptide were analyzed before and after reaction with OPHA after hydrolysis.

Anal. (before reaction). Asp (0.96), Glu (1.0), Phe (0.99). Anal. (after reaction). Asp (1.0), Glu (0.98), Phe (0.63).

The solution after reaction with OPHA both before and after hydrolysis gave a positive test for hydrazine with Ehrlich's reagent; OPHA alone gave no test with this reagent.

The hydrolysis solution was evaporated to dryness and subjected to TLC using the 6:4:3 1-butanol–pyridine–water system after spraying the spotted sample with Ehrlich's reagent. Clearly identified were α-hydrazino-β-phenylpropionic acid (R_f 0.82), phenylalanine (R_f 0.57), glutamic acid (R_f 0.21), aspartic acid (R_f 0.16), and a faint, uncharacterized spot at R_f 0.67.

Methylation of Gly-Leu-Tyr with Dimethyl Sulfate. To a 25-mL, three-necked, round-bottomed flask equipped with a magnetic stirrer, condenser, serum cap, and Ar bubbler was added 10.54 mg (93.9 μmol) of potassium *tert*-butoxide followed by 0.5 mL of dimethylformamide (DMF). To 2 mL of DMF in a small beaker was added 15 mg (42.7 μmol) of Gly-Leu-Tyr. Heating this mixture resulted in a turbid suspension which was transferred to the round-bottomed flask using an additional 1 mL of DMF. The turbid suspension was stirred for 1 h to effect solution, and 8.9 μL (93.9 μmol) of dimethyl sulfate was added. The solution was stirred at room temperature for 3 h, the solvent was removed in vacuo at 50 °C, and the residue was lyophilized. The lyophilization product was dissolved in 1 mL of water, applied to a 2 × 50 cm Sephadex G-10 column, and eluted with water. Fractions of 3 mL were collected and assayed by absorbance at 230 and 275 nm. Two peaks were observed, a minor one peaking at fraction 24 and a major one at fraction 37. Fractions 29–50

were pooled and lyophilized to give 25.9 mg of a residue which was hydrolyzed in 3 N mercaptoethanesulfonic acid for amino acid analysis.

Anal. Gly (0.74), Leu (1.0), Tyr (0.0), O-Methyl-Tyr (0.97).

Cyclization of Glutamic Acid Residues.[44] N^{α}-Acetylglutamic acid (1 mmol) and OPHA·HCl (3.29 mmol) were dissolved in 25 mL of water, and the pH was adjusted to 3.5 with 1.0 M NaOH. A 2-mL aliquot of a 1.5 M aqueous solution of WSC-II was added, and the pH was maintained with the pH stat using 1.0 N HCl. A voluminous precipitate began forming within a few minutes. Three more aliquots of the carbodiimide solution were added at 15-min intervals. The solution was then saturated with NaCl, filtered, and air dried to give N-pivaloyloxy-N^{α}-acetylglutarimide (**39b**) in 88% yield, mp 197–199 °C after recrystallization from ethyl acetate–hexanes, mass spectrum (chemical ionization with methane, giving P + 1) *m/e* 271.

Anal. Calcd for $C_{12}H_{18}N_2O_5$: C, 51.56; H, 6.29; N, 10.93. Found: C, 51.63; H, 6.13; N, 10.85.

Hydrolysis of 39b. N-Pivaloyloxy-N^{α}-acetylglutarimide (**39b,** 0.1 mmol) was suspended in 15 mL of water at pH 8.5, 50 °C. These conditions were maintained using a pH stat and 0.1 N NaOH. An initial uptake of approximately 1 equiv of base occurred as the compound slowly dissolved, and slow base consumption was then observed. The reaction was allowed to continue for 46 h (base uptake was still slightly observable). An aliquot was removed and tested for hydroxamic acids with $FeCl_3$, and a positive test for the corresponding N-hydroxyglutarimide derivative **41b** was obtained. Another aliquot was removed and prepared for amino acid analysis in the usual manner to yield glutamic acid and 2,4-diaminobutyric acid (DABA, **44b**) in the ratio 3.73:1.

***N^{α}*-Acetyl-β-benzylaspartic Acid *N*-Methylamide (49a).** N^{α}-Acetyl-L-aspartic acid β benzyl ester (2.1 g, 7.9 mmol) was suspended in 100 mL of acetonitrile. A 40% methylamine solution (0.7 mL, ca. 1.1 equiv) was added, whereupon the solution became homogeneous. A solution of 1.65 g (8 mmol) of dicyclohexylcarbodiimide in 50 mL of acetonitrile was added and the solution stirred at room temperature for 14 h. The resulting white precipitate of dicyclohexylurea was filtered off, and the filtrae was evaporated to dryness. The residue was taken up in warm ethyl acetate, allowed to cool to 30–35 °C, and filtered to remove a trace more dicyclohexylurea. The filtrate was concentrated to 15–20 mL, heated to boiling, and diluted with hexanes to a slight turbidity. The hot solution was just clarified with ethyl acetate and cooled to give small white crystals. After a second recrystallization, the desired material was obtained in 74% yield: mp 132–134 °C; NMR δ_{CDCl_3} 2.0 (s, 3 H), 2.76 (d, 3 H), 2.6–2.9 (m, 2 H), 4.6–4.9 (m, 1 H), 5.15 (s, 2 H), 6.6–6.8 (broad, –NH–), 6.8–7.1 (broad, –NH–), 7.34 (s, 5 H); mass spectrum (chemical ionization with methane, giving P + 1) *m/e* 279.

The same procedure was used to prepare the glutamyl analogue, **49b, *N^{α}*-acetyl-γ-benzylglutamic acid *N*-methylamide,** in 63% yield, mp 144–146 °C. Mass spectrum (chemical ionization with methane, giving P + 1) *m/e* 293.

***N^{α}*-Acetylaspartic Acid *N*-Methylamide (50a).** Compound **49a** (1 g, 3.6 mmol) was dissolved in 75 mL of absolute methanol in a Parr hydrogenation bottle. The catalyst, 5% Pd/C (0.2 g), was added, and the bottle attached to a Parr hydrogenator, flushed with H_2, pressurized to 15 psi, and shaken for 1 h. The reaction mixture was filtered and the filtrate evaporated to give a white solid (96%), mp 163–167 °C. Recrystallization from methanol–ethyl acetate–hexanes gave an analytical sample, mp 166–168 °C.

Anal. Calcd for $C_7H_{12}N_2O_4$: C, 44.68; H, 6.43; N, 14.89. Found: C, 44.56; H, 6.37; N, 14.75.

The same procedure was used to prepare the glutamyl analogue **50b, *N^{α}*-acetylglutamic acid *N*-methylamide,** in 82% yield, mp 127–130 °C (also reported as 130 °C).[45]

***N^{α}*-Acetylaspartic Acid α-*N*-Methylamide β-*O*-Pivaloylhydroxamic Acid (47a).** Compound **50a** (1 mmol) was dissolved in a solution of OPHA (2.6–3 mmol) in 10 mL of water at pH 3.5. A 1.0-mL aliquot of a solution of 2 M WSC-II was added, and the pH maintained at 3.5 using 1.0 N HCl. Four more aliquots of the carbodiimide were added at 15-min intervals. The pH was then adjusted to 8.5 and the solution immediately extracted with three 15-mL portions of $CHCl_3$ to remove the excess OPHA. The aqueous solution was then reacidified to pH 3 and extracted with $CHCl_3$, and the resulting $CHCl_3$ extract was worked up in the usual manner to yield, after concentration, a white solid which was recrystallized from ethyl acetate–hexanes to give the desired compound in 45% yield, mp 163–164 °C, pK_a = 6.4–6.5. Amino acid analysis: Asp/CH_3NH_2 1.0.

Anal. Calcd for $C_{12}H_{21}N_3O_5$: C, 50.17; H, 7.37; N, 14.63. Found: C, 50.02; H, 7.50; N, 14.49.

***N^{α}*-Acetylglutamic acid α-*N*-methylamide γ-*O*-pivaloylhydroxamic acid (47b)** was prepared in a manner analogous to that used for **47b** except that the final extraction was performed with ethyl acetate. Concentration of the washed and dried ethyl acetate extracts gave the crude product in 91% yield, and recrystallization from ethyl acetate–hexanes gave the analytical sample, mp 173–175 °C, pK_a = 6.6–6.7. Amino acid analysis: Glu/CH_3NH_2 1.0.

Anal. Calcd for $C_{13}H_{23}N_3O_5$: C, 51.82; H, 7.69; N, 13.94. Found: C, 51.92; H, 7.70; N, 13.92.

Acknowledgments. We gratefully acknowledge support of this work by a grant from the National Institute of General Medical Sciences, and one of us (M.J.M.) wishes to acknowledge Training Grant support from the same institute. We thank Professor Vincent du Vigneaud for the use of amino acid analysis facilities, and Mrs. Patricia Miller for assistance with some of the analyses. We also express our appreciation to Dr. Francis Cardinaux for helpful and stimulating discussion.

Registry No.—**10** HCl, 35657-35-3; **7,** 61650-25-7; **39b,** 61650-26-8; **41b,** 61650-27-9; **47a,** 61650-28-0; **47b,** 61650-29-1; **49a,** 61650-30-4; **49b,** 56663-25-3; **50a,** 33067-37-7; benzoic acid, 65-85-0; benzohydroxamic acid, 495-18-1; diphenylurea, 102-07-8; dibenzylurea, 1466-67-7; Gly-Leu-Tyr, 4306-24-5; Phe-Asp-Ala-Ser-Val, 28817-13-2; Phe-Gln-Asn, 51442-24-1; dimethyl sulfate, 77-78-1; Gly-Leu-Tyr-(O-methyl-Tyr), 61650-31-5; N^{α}-acetylglutamic acid, 1188-37-0; N^{α}-acetyl-L-aspartic acid β-benzyl ester, 10144-33-9; methylamine, 74-89-5; N^{α}-acetylglutamic acid N-methylamide, 33878-44-3.

References and Notes

(1) This work has been reported in preliminary form: M. J. Miller and G. M. Loudon, *J. Am. Chem. Soc.*, **97,** 5297 (1975).
(2) Review: T. F. Spande, B. Witkop, Y. Degani, and A. Patchornik, *Adv. Protein Chem.*, **24,** 97 (1970).
(3) G. R. Stark, *Adv. Protein Chem.*, **24,** 261 (1970).
(4) L. Bailey, "Techniques in Protein Chemistry", 2nd ed, American Elsevier, New York, N.Y., 1967.
(5) P. Edman and G. Begg, *Eur. J. Biochem.*, **1,** 80 (1967).
(6) G. R. Stark, *Methods Enzymol.*, **25,** 369 (1972).
(7) T. Wieland and H. Fritz, *Chem. Ber.*, **86,** 1186 (1953).
(8) M. Bergmann and L. Zervas, *J. Biol. Chem.*, **113,** 341 (1936).
(9) E. S. Wallis and J. F. Lane, *Org. React.*, **3,** 267 (1946).
(10) P. A. S. Smith, *Org. React.*, **3,** 337 (1946).
(11) S. Wawzonek and R. C. Gueldner, *J. Org. Chem.*, **30,** 3031 (1965).
(12) H. L. Yale, *Chem. Rev.*, **33,** 209 (1943).
(13) L. Bauer and O. Exner, *Angew. Chem., Int. Ed. Engl.*, **13,** 376 (1974).
(14) For details concerning the preparation of novel O-substituted hydroxylamines, see L. A. Carpino, *Acc. Chem. Res.*, **6,** 191 (1973), and references cited therein.
(15) W. N. Marmer and G. Maerker, *J. Org. Chem.*, **37,** 3520 (1972).
(16) D. G. Hoare and D. E. Koshland, Jr., *J. Am. Chem. Soc.*, **88,** 2057 (1966); *J. Biol. Chem.*, **242,** 2447 (1967).
(17) A 0.1 M solution of OPHA·HCl has a pH of 1.0–1.5 and requires 1 equiv of base to adjust the pH to 3.5–4.0.
(18) This is a common problem in carbodiimide-promoted coupling reactions using poor nucleophiles in organic solvents. For example, see Y. Klausner and M. Bodansky, *Synthesis*, 455 (1972). An attempt has also been made to use the O → N acyl migration for carboxyl-terminal analysis: H. G. Khorana, *J. Chem. Soc.*, 2081 (1952).
(19) C. R. Hauser and W. B. Renfrow, *J. Am. Chem. Soc.*, **59,** 2308 (1937); C. R. Hauser and R. D. Bright, *ibid.*, **61,** 618 (1939).
(20) H. A. Sober, Ed., "Handbook of Biochemistry", Chemical Rubber Publishing Co., Cleveland, Ohio, 1975, p J-187.
(21) D. C. Berndt and H. Schechter, *J. Org. Chem.*, **29,** 916 (1964).
(22) C. Walling and A. N. Naglieri, *J. Am. Chem. Soc.*, **82,** 1820 (1960).
(23) G. M. Loudon, M. E. Parham, and F. E. DeBons, unpublished observations.
(24) For a similar specific peptide cleavage without affecting peptide bonds, see S. Yamashita, *Biochem. Biophys. Acta*, **229,** 301 (1971).
(25) In several cases the aldehyde was identified as its 2,4-dinitrophenylhydrazone. In all cases, the presence of aldehyde was confirmed by the method of R. G. Dickinson and H. W. Jacobsen, *Chem. Commun.*, 1719 (1970).
(26) H. Niedrich and R. Grupe, *J. Prakt. Chem.*, **27,** 108 (1965).
(27) K. L. Carraway and D. E. Koshland, Jr., *Biochem. Biophys. Acta*, **160,** 272 (1968).
(28) S. Shaltiel and A. Patchornik, *J. Am. Chem. Soc.*, **85,** 2799 (1963).
(29) S. Shaltiel and A. Patchornik, *Isr. J. Chem.*, **1,** 187 (1963).
(30) T. Y. Liu and Y. H. Chang, *J. Biol. Chem.*, **246,** 2842 (1971).
(31) R. B. Horner, S. R. Allsop, and J. E. Arrieta, *FEBS Lett.*, **59,** 173, (1975).
(32) S. Dukler, M. Wilchek, and D. Lavie, *Tetrahedron*, **27,** 607 (1971).
(33) T. Nakai and T. Ohta, *Biochem. Biophys. Acta*, **420,** 258 (1976).
(34) S. Moore and W. H. Stein, *Methods Enzymol.*, **6,** 819 (1963).
(35) In these schemes, the α-carboxyl, which is more sensitive to nucleophilic attack, is shown reacting first. Obviously, initial attack at the ω position would give the same result.
(36) M. J. Miller, Ph.D. Dissertation, Cornell University, 1976, p 115 ff.

(37) D. F. Meigh, *Nature (London)*, 579 (1952).
(38) F. Sanger, *Biochem. J.*, **44**, 126 (1949).
(39) S. Fittkau, *Naturwissenschaften*, **15**, 522 (1963).
(40) C. R. Hauser and W. B. Renfrow, Jr., "Organic Syntheses", Collect. Vol. II, Wiley, New York, N.Y., 1943, p 67.
(41) J. Chapman, *J. Chem. Soc.*, 507, 511 (1931).
(42) H. Woelsch, *Methods Enzymol.*, **3**, 570 (1957).
(43) W. J. Maimind, K. M. Ermoljajew, and M. M. Shemjakin, *Zh. Obshch. Khim.*, **26**, 2313 (1956).
(44) Aspartic acid residues were cyclized in a similar manner.
(45) Private communication from Dr. Francis Cardinaux.
(46) Although our analytical facilities are limited to a few tenths of a micromole or above, our procedures seem to work best at the lowest peptide levels attempted. There seems to be no reason why nanomole-level degradations using the procedure described here and elsewhere[1] should not be successful.
(47) M. Caplow, *J. Am. Chem. Soc.*, **90**, 6795 (1968).
(48) A. Williams and W. P. Jencks, *J. Chem. Soc., Perkin Trans. 2*, 1753 (1974).
(49) G. R. Stark, *Biochemistry*, **7**, 1796 (1968).
(50) (a) F. Weygand, D. Hffmann, and E. Wünsch, *Z. Naturforsch.*, **216**, 426 (1966); (b) E. Wünsch and F. Drees, *Chem. Ber.*, **99**, 110 (1966).
(51) H. T. Clark and L. D. Behr, "Organic Syntheses", Collect. Vol. II, Wiley, New York, N.Y., 1943, p 19.

A Computerized Infrared Spectral Interpreter as a Tool in Structure Elucidation of Natural Products[1]

Hugh B. Woodruff and Morton E. Munk*

Department of Chemistry, Arizona State University, Tempe, Arizona 85281

Received September 29, 1976

The process of structure elucidation as practiced by the natural products chemist requires the determination of structural constraints (the polyatomic fragments known to be present in the molecule and others known to be absent). The structural constraints may be chemist derived or computer derived. This paper decscribes an interactive computer program that interprets infrared spectra in order to help the chemist determine structural constraints. The program attempts to parallel the reasoning a chemist uses in interpreting a spectrum. If the program is to be of value to the chemist, it must be able to interpret the spectra of relatively complex molecules and make decisions concerning the presence or absence of a large number of functional groups. This program makes decisions concerning 159 different classes of compounds and has been tested on a wide variety of sample spectra.

Chemists in general, and natural products chemists in particular, are frequently faced with the task of piecing together results from chemical experimentation and spectroscopic work to deduce the molecular structure of an unknown compound. The chemist interprets these data, expresses the result in terms of a partial structure (structural fragments plus unaccounted for atoms), and intuitively attempts to reduce the partial structure to all molecular structures consistent with the available evidence. While intuition is a valuable asset to the chemist, an asset that cannot be adequately programmed into a computer model of the structure elucidation process, the chemist may frequently overlook a valid combination pathway, especially when dealing with the relatively complex molecules of nature. To preclude such an occurrence, and to relieve the chemist of the tedious task of manually assembling molecular structures, several computerized structure generators have been developed that ensure that *all* chemically feasible molecules are considered.[2-9]

For all but the simplest molecular formula, the number of valid structures generated without structural constraints (the polyatomic fragments known to be present and others known to be absent) is unmanageably large. As structural constraints are imposed on the process, the number of structures generated is diminished, ultimately to one. The structural constraints may be chemist derived or computer derived. Clearly, an essential component of any computerized structure elucidation package is an effective spectral interpretation procedure.

This paper describes an artificial intelligence program that aids the chemist in the interpretation of infrared spectra. The program is designed to present the chemist wit the most logical interpretation of a spectrum, not to interpret it definitively and pass the results directly to the structure generator, bypassing the chemist entirely. The chemist is an integral part of the decision making process and, for that reason, the program is interactive in nature.

When the task is to identify an unknown compound from its spectral data, and it is suspected that the unknown might be included in an accessible large library of spectra, a search and compare scheme is probably the best approach.[10,11] Pattern recognition procedures have enjoyed some success in predicting the functional groups that are present using infrared spectra.[12-14] A purely empirical approach for interpreting spectra has recently been described by Gray.[15] Like the method of Gray, the program described in this paper is completely empirical. It attempts to parallel the chemist's reasoning in interpreting an infrared spectrum as much as possible. The advantage in using an interpretation program such as this one in a structure elucidation package is evident when one considers that it is imperative that no information passed on to the structure generator can be incorrect. When an error is found, this program can be altered to correct the error, a capability that does not exist with pattern recognition programs.

General Approach

The chemist uses an empirical approach to interpret infrared spectra. After observing a sufficient number of spectra, or alternatively reading textbooks and learning from the observations of others, the chemist develops the ability to associate certain absorption peaks with their corresponding functional groups. In other words, a set of "rules" for interpreting infrared spectra is learned. For example, the "rules" for identifying a carboxylic acid might be to look for a broad, medium to strong peak centered around 3000 cm^{-1}, a strong peak near 1720 cm^{-1}, and a broad, medium intensity band around 920 cm^{-1}. It is this type of reasoning that must be programmed in order to develop a successful spectral interpreter. One of the first problems encountered is the need to digitize the spectra. It must be decided how much peak shape and intensity information should be retained. Many search systems and some pattern recognition procedures have been