Chapter 1

Peptide Fragmentation/Deletion Side Reactions

Due to the inherent attributes of certain peptide individuals they could undergo a variety of fragmentation processes during synthesis, purification or even storage. Fragmentation could selectively address peptides with characteristic sequences like *N*-terminal *N*-Ac-*N*-alkyl moiety, *N*-acyl-*N*-alkyl-Aib-Xaa- bond, -Asp-Pro-, *N*-terminal His-Pro-Xaa- moiety, *C*-terminal *N*-Me-Xaa, *N*-terminal FITC, thioamide bond and guanidinyl group on Arg side chain, etc. Moreover, utilization of isodipeptide Boc-Ser/Thr(Fmoc-Xaa)-OH as the building block for peptide synthesis could result in the formation of des-Ser/Thr impurity. On top of these specific cases DKP formation could also affect general peptide assembly that leads to the deletion of affected dipeptide moiety from the parental peptide sequence. The occurrence of these fragmentation/deletion side reactions on peptide materials could decrease the manufacturing yield, cause challenges for the down-stream peptide purification, and affect peptide stability upon processing and/or storage. Phenomenon and mechanism of common fragmentation/deletion in peptide synthesis are described in this chapter. Corresponding solutions to minimize these side reactions are proposed.

1.1 ACIDOLYSIS OF PEPTIDES CONTAINING *N*-Ac-*N*-alkyl-Xaa MOTIF

Peptides with a motif of *N*-Ac-*N*-alkyl-Xaa sequence at the *N*-terminus have the distinctively high propensity to suffer from an acidolysis side reaction at the step of acid-mediated peptide cleavage from resin and side chain global deprotection. The *N*-terminal *N*-Ac-*N*-alkyl-Xaa unit might be split from the parental peptide as a 5-member ring derivative, leading to the formation of des-*N*-Ac-*N*-alkyl-Xaa truncated side product.

This kind of side reaction has been detected in the process of the preparation of a series of Arodyn peptides ((acetylated Dyn A) Arodyn 1, 2, 3, 4).[1] It was reasoned that the synthesis of Arodyn 2 resulted in the acidolytic cleavage of *N*-terminal motif *N*-Ac-*N*-Me-Phe during the TFA-mediated global deprotection step (Table 1.1).

The proposed mechanism of the subjected acidolysis side reaction is indicated in Fig. 1.1. It is reasoned in the corresponding investigation that the

TABLE 1.1 Sequences of Arodyn 1, Arodyn 2, Arodyn 3, Arodyn 4 Peptides

Dyn A(1-11)	H-Tyr-Gly-Gly-Phe-Leu-Arg-Arg-Ile-Arg-Pro-Lys-NH_2
Arodyn 1	Ac-Phe-Phe-Phe-Arg-Leu-Arg-Arg-D-Ala-Arg-Pro-Lys-NH_2
Arodyn 2	Ac-*N*-Me-Phe-Phe-Trp-Arg-Leu-Arg-Arg-D-Ala-Arg-Pro-Lys-NH_2
Arodyn 3	CH_3OCO-*N*-Me-Phe-Phe-Trp-Arg-Leu-Arg-Arg-D-Ala-Arg-Pro-Lys-NH_2
Arodyn 4	*N*-Me-Phe-Phe-Trp-Arg-Leu-Arg-Arg-D-Ala-Arg-Pro-Lys-NH_2

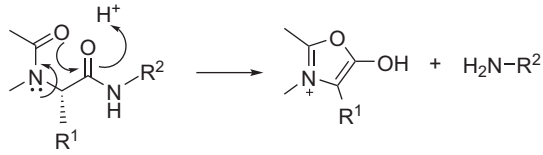

FIGURE 1.1 Proposed mechanism of the acidolytic cleavage of *N*-Ac-*N*-alkyl-Xaa from parental peptide.

occurrence of this side reaction is subject to the actual conditions under which the peptide global deprotection is conducted. It is verified that if the referred reaction is processed at 4°C in the absence of any scavengers the acidolysis of *N*-terminal *N*-Ac-*N*-alkyl-Xaa could be significantly suppressed. No similar impurities with deletion sequences have been detected in the process of Dyn A(1-11) or Arodyn 1 synthesis. The preparation of Arodyn 4 that is devoid of acetyl moiety on its *N*-terminus does not suffer from the concerned acidolysis side reaction upon TFA treatment, accounting for the involvement of the acetyl functional group in the process of *N*-Ac-*N*-alkyl-Xaa acidolysis. Significant *N*-terminus acidolysis side reaction has been invoked in the synthesis of Arodyn 2 in which Ac-*N*-Me-Phe is located on the *N*-terminus compared with the Ac-Phe motif from Arodyn1. This phenomenon is attributed to the presence of *N*-alkyl amino acid residue that favors the advantageous peptide secondary structure facilitating the acidolytic fragmentation of the *N*-terminal residue. In case the *N*-terminal acetyl is replaced by more electron-withdrawing group methyl carbamate, as is the case for Arodyn 3, the subjected acidolysis side reaction on the peptide *N*-terminus would be basically circumvented due to the decrease of the nucleophilicity of the carbonyl oxygen from the methyl carbamate that initiates the ring closure in the acidolytic fragmentation process. It could therefore be deduced from the aforementioned phenomenon that the acidolysis of peptide *N*-terminal *N*-Ac-*N*-alkyl-Xaa motif is induced by the acetyl oxygen nucleophilic attack on the amide bond between the subjected *N*-Ac-*N*-alkyl-Xaa and the neighboring amino acid at its *C*-terminus, facilitated by the advantageous local structure in that the ratio of *cis*-amide bond is significantly increased by the presence of an *N*-alkyl-amino acid residue. Under such

circumstances the *N*-acetyl group serves as a nucleophile that initiates the ring closure, and subsequent acidolytic fragmentation of the *N*-Ac-*N*-alkyl-Xaa unit.

1.2 Des-Ser/Thr IMPURITIES INDUCED BY *O*-acyl ISODIPEPTIDE Boc-Ser/Thr(Fmoc-Xaa)-OH AS BUILDING BLOCK FOR PEPTIDE SYNTHESIS

O-acyl isodipeptide derivatives have already found widespread application as effective building blocks in peptide synthesis, particularly for difficult peptide assemblies that are hardly quantitatively realized by the conventional stepwise coupling methods. This methodology takes advantage of the inherent feature of the base-induced reversible intramolecular acyl *O*→*N* shift that involves the ester bond from the Ser/Thr side chain and the α-amino group on the peptide backbone (Fig. 1.2).

The incorporation of the isodipeptide unit into the peptide sequence is intended to disrupt the adverse secondary structure of the subjected peptide that impedes the smooth coupling of the forthcoming amino acid to the elongating peptide chains, particularly for the "difficult couplings." Peptide secondary structures are basically induced and reinforced by diverse molecular interactions such as hydrogen bond, Van der Waals force, hydrophobic interaction, ionic bond, and so forth. The establishment of peptide secondary structure might considerably reduce the flexibility of the affected peptide chains that consequently adversely interferes with the subsequent amino acid couplings during peptide synthesis. This phenomenon is basically regarded as one of the major causes for the nonquantitative amino acid couplings occurred in peptide synthesis that accounts for the generation of peptide impurities with deletion sequences.

O-acyl isodipeptide building blocks[2–4] are utilized in an effort to address this inherent problem in peptide synthesis. The existence of -Xaa-Ser- or -Xaa-Thr- unit in the target peptide sequence is the prerequisite for the employment of

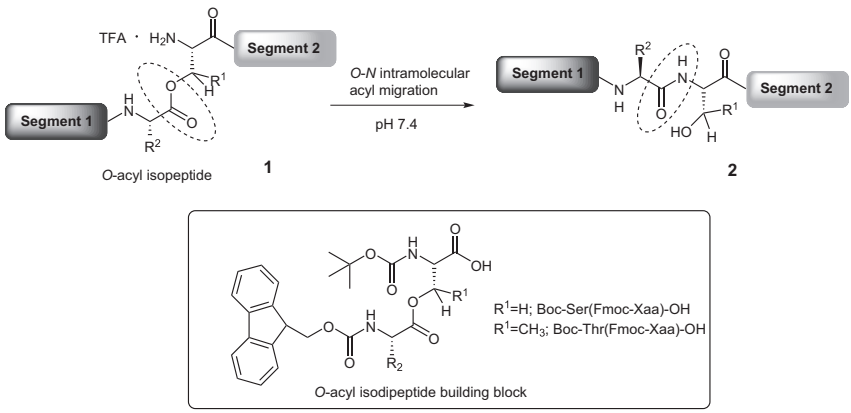

FIGURE 1.2 Peptide preparation via *O*-acyl isopeptide strategy.

O-acyl isodipeptide strategy. The subjected isodipeptide unit is incorporated in the manner of Boc-Ser/Thr(Fmoc-Xaa)-OH building block into the target peptide chains, functioning as the synthon for the natural -Xaa-Ser/Thr- counterpart. The intermediary product containing *O*-acyl isodipeptide structure is depicted as compound **1** in Fig. 1.2. The backbone carboxyl group of the -Xaa- unit is chemically linked with the hydroxyl side chain from Ser/Thr by means of an ester bond (highlighted in a dotted circle). The introduction of *O*-acyl isodipeptide moiety could manifestly disrupt the local peptide secondary structure. The solubility and liquid chromatographic properties of the peptide precursor **1** containing *O*-acyl motif are normally superior to those of its interchangeable *N*-acyl counterpart **2**. These outstanding features of *O*-acyl isopeptide could tremendously facilitate the otherwise challenging chromatographic purification. The purified *O*-acyl isopeptide **1** will be subsequently addressed to the base-catalyzed acyl *O*→*N* shift process that regenerates the natural form of the peptide amide bond via a five-member ring intermediate. The disadvantageous peptide secondary structure that impedes the smooth amino acid coupling is circumvented by this means, significantly facilitating the effective chemical preparation of the target peptide product.

In spite of the successful utility of *O*-acyl isodipeptide strategy manifested in the challenging peptide preparation such as β-amyloid 1-42,[5] it has been detected that this methodology could potentially induce side reactions such as β-elimination which leads to the formation of des-Ser/Thr impurities. The possible mechanism of this side reaction is originated from the formation of active ester Boc-Ser/Thr(Fmoc-Xaa)-OBt **4** derived from the carboxylate activation of its precursor *O*-acyl isodipeptide Boc-Ser/Thr(Fmoc-Xaa)-OH **3** (as depicted in Fig. 1.3). The lifespan of the activated derivative **4** in the reaction system is directly correlated to the kinetics of the subjected acylation reaction. If the referred reaction is proceeding

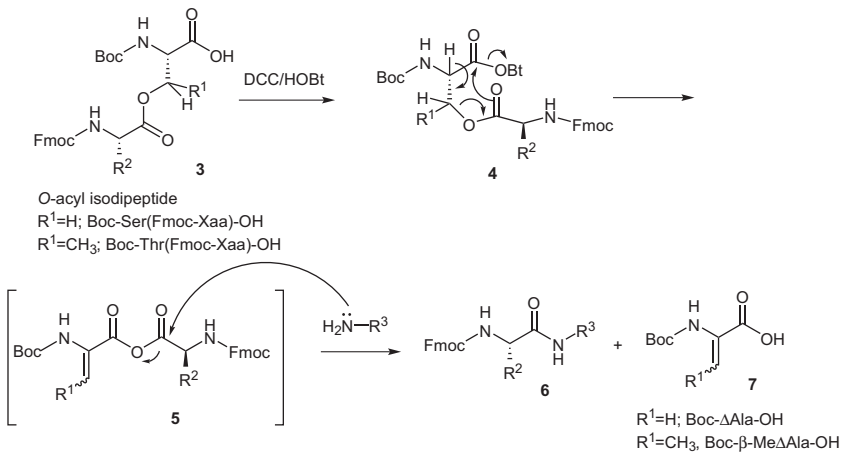

FIGURE 1.3 Proposed mechanism of *O*-acyl isodipeptide induced Ser/Thr elimination side reaction.

sluggishly, Boc-Ser/Thr(Fmoc-Xaa)-OBt **4** will be afforded with sufficient time to deviate from the target intermolecular condensation reaction and undergo intramolecular rearrangement by means of β-elimination, giving rise to the formation of the mixed anhydride 5 from Fmoc-Xaa-OH and Boc-(β-Me)ΔAla-OH, as indicated in Fig. 1.3. As a consequence, the unacylated peptide chain could possibly function with **5** at its two reactive sites, but the anhydride carbonyl at Fmoc-Xaa side is preferred due to the fact that the unsaturated (β-Me)ΔAla side chain unavoidably attenuates the electrophilicity of anhydride carbonyl on the Boc-(β-Me)ΔAla side. The unit of (β-Me)ΔAla is, therefore, excluded from the product structure as Boc-(β-Me)ΔAla-OH **7** upon the nucleophilic attack of the peptide N^{α} on the mixed anhydride **5**, giving rise to the formation of des-Ser/Thr impurity **6**.

In order to verify the proposed mechanism of *O*-acyl isodipeptide-induced deletion side reaction, Boc-Ser(Fmoc-Gly)-OH isodipeptide was incubated in NMP in the presence of DCC (2 equiv.)/HOBt (2 equiv.) for 2 h before 2.2 equiv. benzylamine was charged into the reaction system.[5] The obtained product was analyzed by MS and analytical RP-HPLC, and no Boc-Ser(Fmoc-Gly)-NHBzl was detected while large amount of Fmoc-Gly-NHBzl as well as Boc-ΔAla-NHBzl were located instead. As a matter of fact, the abundance of Fmoc-Gly-NHBzl side-product in the crude material is as high as 80%. In another experiment *O*-acyl isodipeptide Boc-Ser(Fmoc-Gly)-OH was subject to the activation process by 2 equiv. DIC/2 equiv. HOBt in DMF-d_7 for 2 h, ^{1}H-NMR analysis of the obtained product detected 2 types of olefin hydrogen signal which were assigned to E/Z isomers. This result combined with the corresponding MS and RP-HPLC analysis explicitly indicates that Boc-Ser(Fmoc-Gly)-OH has almost been quantitatively converted to the mixed anhydride composed of Fmoc-Gly-OH and Boc-ΔAla-OH within 2 h upon activation by DIC/HOBt.

Moreover, Boc-Ser(Fmoc-Ile)-OH, Boc-Thr(Fmoc-Gly)-OH and Boc-Thr(Fmoc-Ile)-OH were subject to DIC (2 equiv.)/HOBt (2 equiv.) activation in DMF for 2 h, respectively, before 2 equiv. benzylamine was charged into the reaction system to entrap the activated species. Abundant Fmoc-Gly-NHBzl, Fmoc-Ile-NHBzl, Boc-ΔAla/β-MeΔAla-NHBzl were detected as a consequence in the corresponding crude products.[5] All these results have unequivocally verified the susceptibility of *O*-acyl isodipeptide Boc-Ser/Thr(Fmoc-Xaa)-OH to suffer from the undesired rearrangement/deletion side reaction upon activation, while the inclination of this process is seemingly independent on the steric effect of the concerned amino acid.

Another indicative finding towards this side reaction is that when Boc-Ser(Fmoc-Gly)-OH was incubated in $CDCl_3$ in the presence of DIC/HOBt, ^{1}H-NMR of the obtained crude product did not indicate the existence of olefin signals from Boc-ΔAla-OH.[5] This result implies that the inclination of this side reaction of *O*-acyl isodipeptide is considerably influenced by the properties of the solvent. Polar solvents such as DMF or NMP would facilitate this process while unpolar solvents like DCM or $CHCl_3$ could minimize its occurrence. In light of this finding, it is advisable to utilize the unpolar solvents for the

activation and coupling of *O*-acyl isodipeptide in order to suppress the deletion side reaction in this process. Moreover, it has been figured out that the types of the coupling reagent additives, such as HOBt, HOAt and HOOBt, would not exert significant impacts on the propensity of this side reaction.

1.3 ACIDOLYSIS OF -*N*-acyl-*N*-alkyl-Aib-Xaa- BOND

Peptide *N*-terminal *N*-Ac-*N*-alkyl-Xaa moiety can not only be addressed to the aforementioned acidolysis process, but is also subjected to the *endo*-peptide bond scission side reaction taken place at the site of -*N*-acyl-*N*-alkyl-Aib-Xaa- sequence upon acid treatment.

The undesired acidolytic fragmentation process on -*N*-acyl-*N*-alkyl-Aib-Xaa- sequence was detected in the preparation of head-to-tail cyclic peptide *cyclo*-[Phe-D-Trp-Lys-Thr-Phe-*N*-Me-Aib].[6] The side chain protected precursor peptide *cyclo*-[Phe-D-Trp-Lys(Boc)-Thr(tBu)-Phe-*N*-Me-Aib] **8** was subjected to TFA/EDT-mediated global deprotection treatment. It is highlighted in Fig. 1.4 that acidolysis at the site of -*N*-Me-Aib-Phe- gives rise to ring disclosure and formation of linear peptide H-Phe-D-Trp-Lys-Thr-Phe-*N*-Me-Aib-OH **9** as well as its thioester counterpart H-Phe-D-Trp-Lys-Thr-Phe-*N*-Me-Aib-SCH_2CH_2SH **10**.

A plethora of peptides containing *N*-Me-Aib residue has been produced, and their crystal structures have been intensively investigated in order to study the mechanism of the acidolysis side reactions occurred at the site of -*N*-Me-Aib-Xaa. The X-ray crystallography analysis of these peptides combining with the kinetics of the acidolysis implies the origins of -*N*-acyl-*N*-alkyl-Aib-Xaa- acidolysis from the aspects of steric effects. It has been discovered in a dedicated investigation[6] that *N*-Me-Aib-containing cyclic peptide *cyclo*-[Phe-Ser(Bzl)-Ser(Bzl)-Phe-*N*-Me-Aib] **11** suffered from acidolytic fragmentation at the site of -*N*-Me-Aib-Phe- upon pure TFA treatment, generating the corresponding

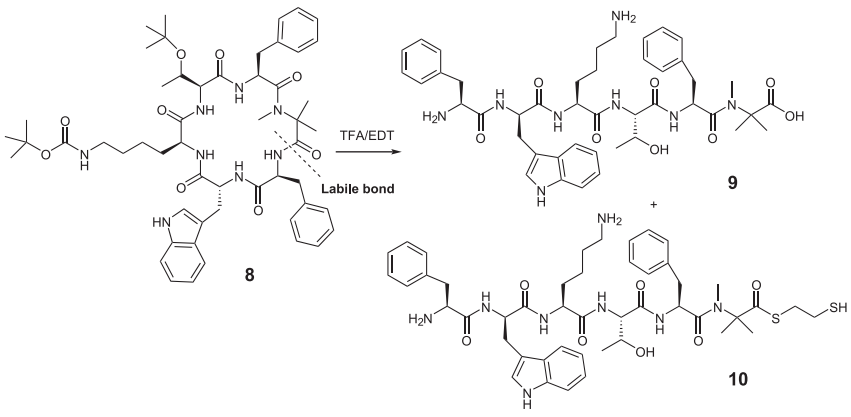

FIGURE 1.4 Acidolysis and ring disclosure of *cyclo*-[Phe-D-Trp-Lys(Boc)-Thr(tBu)-Phe-*N*-Me-Aib].

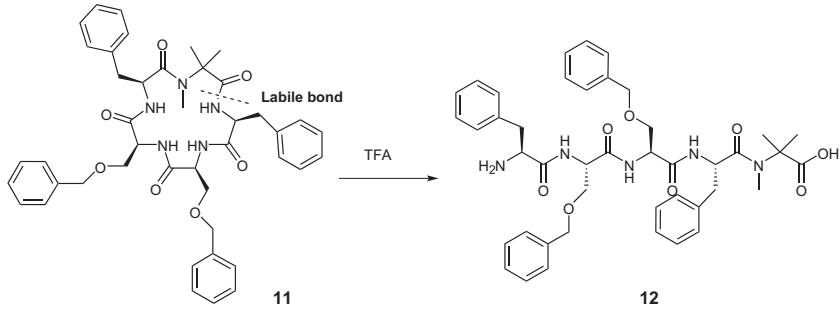

FIGURE 1.5 Acidolysis of peptide *cyclo*-[Phe-Ser(Bn)-Ser(Bn)-Phe-*N*-Me-Aib].

degraded linear peptide H-Phe-Ser(Bzl)-Ser(Bzl)-Phe-*N*-Me-Aib-OH **12** (Fig. 1.5). This ring disclosure process is identified as a pseudo first-order reaction according to its kinetics. The rate of acidolysis is reduced upon the addition of water into the reaction system, while the solvent polarity is significantly decisive for this process in that cyclic peptide **11** underwent a considerably faster acidolysis in TFA/CH_3CN (1:1) ($t_{1/2}$ = 1.1 h) than in TFA/DCM (1:1) ($t_{1/2}$ = 4.1 h). This feature is attributed to the formation of oxazolinium intermediate during acidolysis of the referred -*N*-acyl-*N*-alkyl-Aib-Xaa- peptide bond. Increase of the CH_3CN content will accelerate the kinetics of the ring disclosure.

It was illustrated from an X-ray crystallography study of cyclic peptide **11** that all amide bonds possess ordinary lengths and angles. On the other hand, it was detected that C^{α} of Aib and the carbonyl oxygen from the amino acid preceding Aib residue is spatially in proximity. This would imply that the subjected oxygen atom might be involved in a nucleophilic attack at the -*N*-Me-Aib-Xaa- bond that finally resulted in fragmentation at this site.

The proposed mechanism of the acidolysis of -*N*-acyl-*N*-alkyl-Aib-Xaa- is illustrated in Fig. 1.6 based on the above investigations. Peptide **13** containing

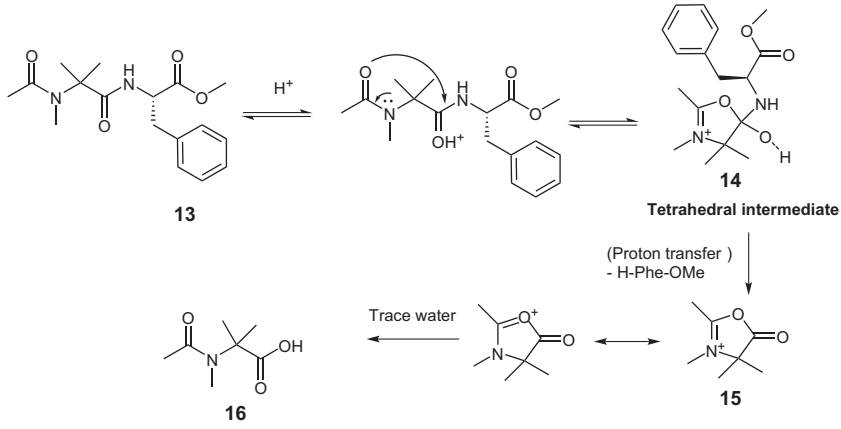

FIGURE 1.6 Proposed mechanism of *N*-acyl-*N*-alkyl-Aib-Xaa- acidolysis.

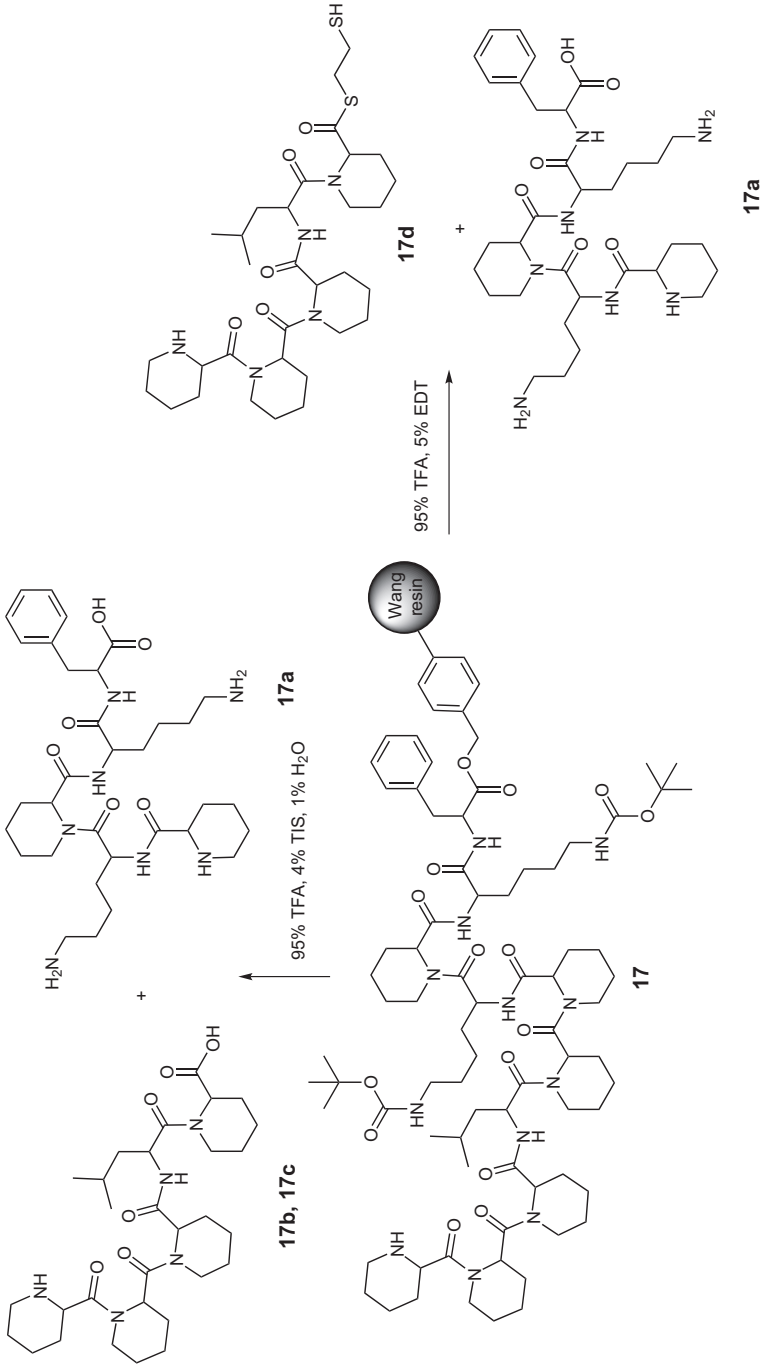

FIGURE 1.7 Acidolysis of Pip-abundant peptide.

N-Ac-*N*-Me-Aib-Xaa- moiety serves as the substrate in this connection. It is readily transformed into a tetrahedral intermediate **14** in acidic milieu. The nitrogen atom from Phe residue in compound **14** does not participate in the conjugation system with *N*-Me-Aib unit, rendering it into a proton acceptor in the acidic condition, dispelling H-Phe-OMe moiety off the intermediate **14** complex, and giving rise to the formation of oxo-oxazolinium derivative **15**. The latter is rapidly hydrolyzed into *N*-Ac-*N*-Me-Aib-OH **16**, finalizing the acidolysis process.

Similar side reactions have also been identified in the preparation of various Pip-abundant peptide derivatives.[7] Treatment of peptidyl resin **17** by 95% TFA/4% TIS/1% H_2O led to the formation of peptide fragments: **17a**, **17b**, and **17c** (Fig. 1.7). It was verified by MS and RP-HPLC that the amide bond between Pip^5 and Pip^6 was subjected to the fragmentation in this process, resulting in the formation of degraded peptide fragments: **17a**, **17b**, and **17c** – the two latter derivatives are diastereomers since the concerned acidolysis at the amide bond between Pip^5 and Pip^6 simultaneously induces configuration conversion on Pip^5-C^{α}. Meanwhile, treatment of **17** by 95% TFA/5% EDT released fragments **17a** and **17d**, the latter is the corresponding thioester of **17b/c** derivatives.

1.4 ACIDOLYSIS OF -Asp-Pro- BOND

It is known that the -Asp-Pro- peptide bond is labile under acidic conditions, such as in TFA,[8] HF,[9] formic acid,[10] and acetic acid.[11] Acidolysis of -Asp-Pro- peptide bond may not only take place in HF-mediated peptide side chain global deprotection reaction but also in weak acidic milieu (pH = 4).[12] The mechanism of this process (see also Fig. 1.8) is basically akin to that of aspartimide formation in that the amide nitrogen atom from Pro backbone attacks the carboxyl side chain of the preceding Asp residue, forming an instable cationic imide intermediate **18**[8] readily hydrolyzed into peptidyl fragments **19** and **20** whose *C*- and *N*-terminus are occupied by the subjected Asp and Pro, respectively.

In another separate investigation[13] protein E298D eNOS has been identified to suffer from acidolytic fission at -Asp^{298}-Pro^{299}- sequence, giving rise to 100 and 35 kDa fragments, while its native protein counterpart eNOS (Glu^{298})

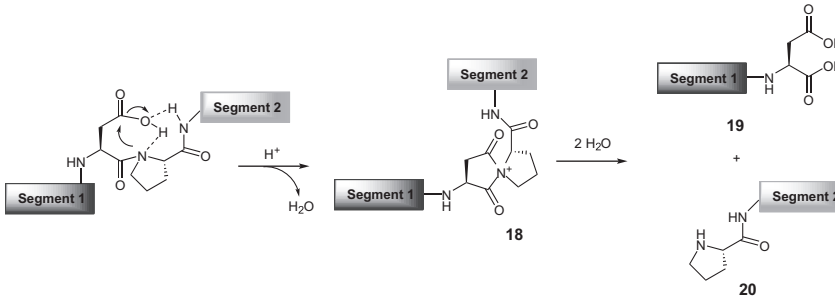

FIGURE 1.8 Proposed mechanism of -Asp-Pro- acidolysis.

is exempted from acidolysis under the same conditions. This distinctive contrast accounts for the notorious susceptibility of -Asp-Pro- to undergo acidolytic fragmentation, highly probable via the imide intermediate formation step.

The local peptide/protein secondary structure around -Asp-Pro- sequence plays a critically important role in dictating the readiness of the acidolysis side reaction. It has been discovered[12] that cellulosomal scaffoldin protein unit cohesin2-CBD undergoes fragmentation in a buffer solution at pH 4 and the acidolysis site is exactly -Asp-Pro- sequence. While this protein contains three -Asp-Pro- moieties located at -Asp^{40}-Pro^{41}-, -Asp^{50}-Pro^{51}-, -Asp^{57}-Pro^{58}-, respectively, only the -Asp^{57}-Pro^{58}- unit suffers from the acidolysis, and the other 2 remain intact upon the treatment. It is subsequently disclosed that the labile -Asp^{57}-Pro^{58}- sequence is located at a relatively rigid turn structure motif synergically stabilized by multiple hydrogen bonds.[14,15] The crystal structure of the parental protein[12] indicates that the carboxyl side chain of Asp^{50} does not lie in close proximity to Pro,[51] whereas Asp^{40} and Asp^{57} side chains are spatially closer to Pro^{41} and Pro,[58] respectively. Moreover, the oxygen atoms from Asp^{40} and Asp^{57} carboxyl side chains are noncovalently paired with reciprocal nitrogen atoms from Asn^{42} and Asn^{59} amide side chains, respectively, by means of hydrogen bond. This spatial alignment brings the carboxyl side chain from Asp, and the backbone amide on the neighboring Pro into proximity, and locks the local moiety into an advantageous conformation that promotes both the imide intermediate generation and the subsequent hydrolysis.

Some peptides such as Herpes simplex virion-originated peptide might suffer from -Asp-Pro- cleavage during FAB-MS analysis.[16] Meanwhile, when the labile -Asp-Pro- unit in the referred peptide was replaced by -Asn-Pro- the stability of the modified peptide could be considerably enhanced under FAB-MS analysis conditions, and no -Asn-Pro- fragmentation was detected. Laser irradiation might induce -Asp-Pro- fission as well.[17]

1.5 AUTODEGRADATION OF PEPTIDE *N*-TERMINAL H-His-Pro-Xaa- MOIETY

It is known that imidazolyl side chain from His endows many functional proteins with a wide variety of catalytic effects, whereas this functional group could also initiate various autocatalysis processes especially when the concerned His is located on the *N*-terminus of the subjected peptide/protein chain, and neighbored by a Pro residue. The amide bond between the pertinent Pro and the amino acid on its *C*-terminal side in peptide sequence could be suffered from fragmentation process catalyzed by the imidazole group on the *N*-terminal His.[18] Apparently, the presence of Pro residue in the referred peptide sequence facilitates the adoption of the *cis*-configuration of His-Pro amide bond which favors as a consequence the nucleophilic attack of the His-N^{α} on -Pro-Xaa- backbone amide, whereas the nucleophilicity of His-N^{α} is strengthened through the effect