Addition of HOBt Improves the Conversion of Thioester-Amine Chemical Ligation

Toni Todorovski,[1] David Suñol,[1] Antoni Riera,[1,2] Maria J. Macias[1,3]

[1]*Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, Baldiri Reixac, 10 08028 Barcelona, Spain*

[2]*Departament de Química Orgànica, University of Barcelona, Martí i Franquès, 1, Barcelona 08028, Spain*

[3]*Catalan Institution for Research and Advanced Studies (ICREA), Passeig Lluis Companys, Barcelona 23 08010, Spain*

Received 6 May 2015; revised 10 September 2015; accepted 18 September 2015
Published online 22 September 2015 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/bip.22745

ABSTRACT:

The syntheses of large peptides and of those containing non-natural amino acids can be facilitated by the application of convergent approaches, dissecting the native sequence into segments connected through a ligation reaction. We describe an improvement of the ligation protocol used to prepare peptides and proteins without cysteine residues at the ligation junction. We have found that the addition of HOBt to the ligation, improves the conversion of the ligation reaction without affecting the epimerization rate or chemoselectivity, and it can be efficiently used with peptides containing phosphorylated amino acids.

Keywords: ligation of peptide fragments through direct aminolysis; cysteine-free native chemical ligation; native chemical ligation with β-branched residues at ligation junction; formation of highly reactive HOBt esters

Additional Supporting Information may be found in the online version of this article.

Correspondence to: Maria J. Macias, Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, Baldiri Reixac, 10, 08028 Barcelona, Spain;
e-mail: maria.macias@irbbarcelona.org

Contract grant sponsor: European Union's Seventh Framework Programme for research, technological development and demonstration
Contract grant number: IRBPostPro 246557
Contract grant sponsor: Spanish National Research Program
Contract grant number: SAF2011-25119, BFU2014-53787-P, CTQ2014-56361, 2009SGR00901
Contract grant sponsor: Consolider
Contract grant number: CSD2009-00080

This article was originally published online as an accepted preprint. The "Published Online" date corresponds to the preprint version. You can request a copy of any preprints from the past two calendar years by emailing the Biopolymers editorial office at biopolymers@wiley.com.

INTRODUCTION

Solid-phase peptide synthesis (SPPS), developed in the 60s by B. Merrifield,[1] is nowadays the standard methodology for chemical preparation of peptides. Although peptides up to 50 residues can be obtained using this methodology [2,3] the linear synthesis of large or complex sequences frequently yield mixtures containing undesired products and is usually not practical. Improvement of these syntheses often require the application of a convergent approach, dissecting the native sequence in segments that are connected using a ligation reaction (Figure 1, Panel A).[4–6] These protocols also facilitate the introduction of modified amino acids in proteins (including phosphorylated or acetylated residues or non-natural amino acids) as well as to the addition of specific labels for antibody recognition, protein detection and cell imaging.[7]

The direct coupling of two peptides was reported for the first time by Kemp *et al.* in 1974.[8,9] This method allows a ligation reaction to occur independently of the amino acids pairs at the ligation junction. The reaction couples a peptide containing an activated carboxylic derivative at its C-terminus to a second peptide carrying an unprotected NH_2-group at its N-terminus. The success of the method depends on the aqueous solubility of the designed segments and suffers from the epimerization of the product at the ligation site. Improvements to the direct aminolysis method were described later by Blake

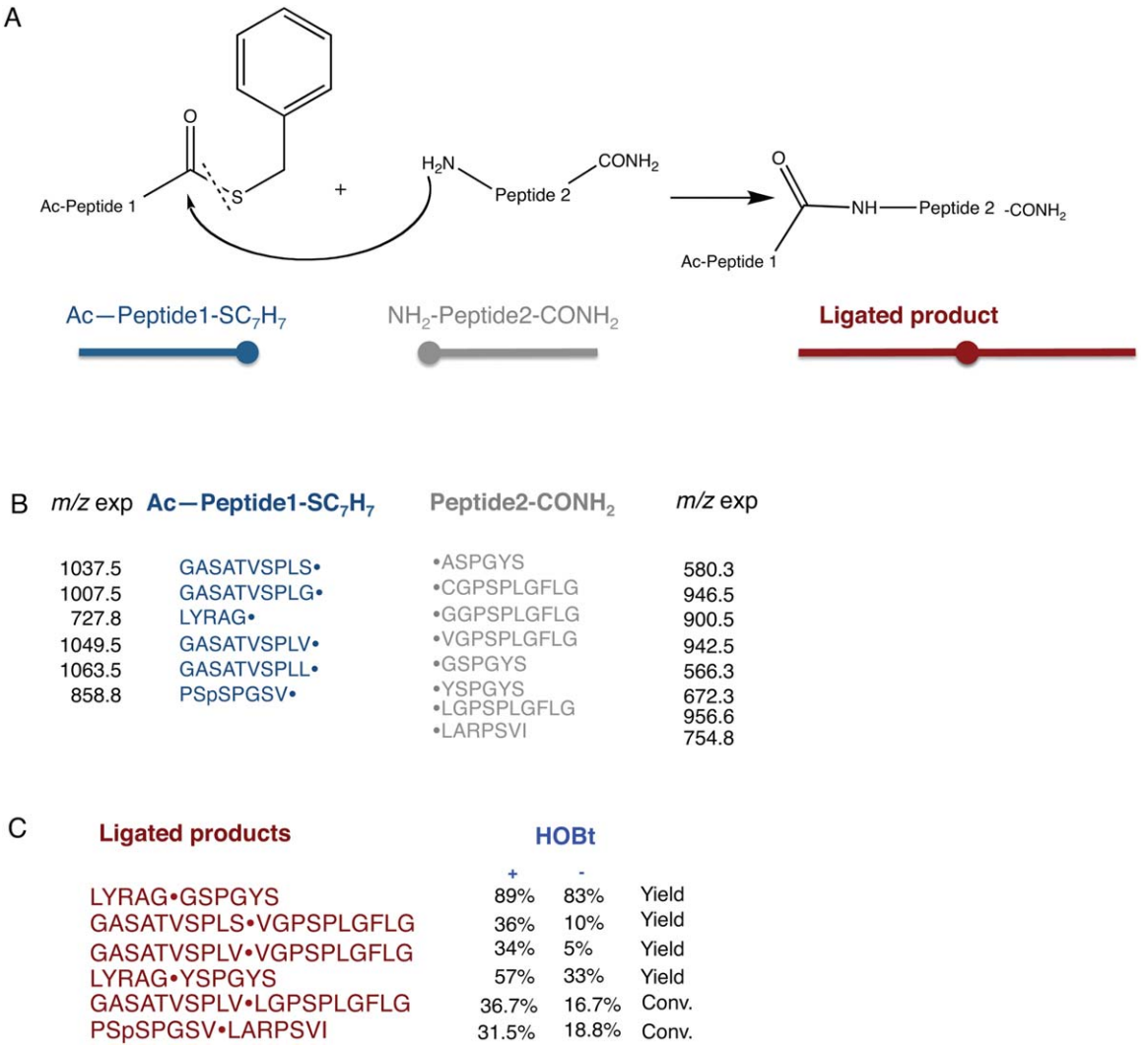

FIGURE 1 Schematic representation of the ligation reaction through direct aminolysis (Panel A) Peptides designed for the ligation reactions used in this study. The m/z experimental values (which are same as theoretically calculated ones) obtained for each peptide is shown next to the corresponding sequences prior the ligation (Panel B). Ligated peptides sequences (the ligation site is represented with a circle) are listed, with yields and conversion percentages obtained for ligation reactions carried in the presence or in the absence of HOBt (Panel C).

et al. and Aimoto *et al.*[10–12] However, in these protocols the reaction conditions cause epimerization of the C-terminal thioester residue, and therefore only non-epimerizable amino acids (glycine or proline) can be introduced at this position.

A breakthrough in the ligation strategy was the development of the native chemical ligation (NCL) approach.[4] In this reaction, carried out in aqueous buffers, the peptide containing a C-terminal thioester is attacked by the thiol group of a cysteine residue located at the N-terminus of the second peptide. The thioester intermediate forms a native amide bond after an intramolecular rearrangement. This ligation reaction can also be used with recombinant proteins to generate the segment carrying the thioester group at the C-terminus *via* intein-mediated cleavage.[13] A drawback in this process is that proteins synthesized by NCL are generated mainly from two segments and are often limited by the presence of cysteines in the sequence. If cysteines are not present in the native sequence, the ligated peptide would require a desulfurization step which is not always straightforward.[14,15]

Novel approaches have recently been described by Wong and coworkers, that partially overcome the above mentioned drawback.[16,17] Their works report the chemical ligation of peptides and proteins without cysteine residues at the ligation junction with high chemoselectivity and low levels of

epimerization. The main limitation of these strategies is the ligation rate, normally requiring long reaction times, ranging from a few hours to several days.

We considered that adding 1-hydroxybenzotriazole (HOBt), a common reagent used in SPPS, to the protocol described by Wong and coworkers,[16] could improve the ligation yield and the conversion of the ligation reaction without affecting the epimerization rate or chemoselectivity. To test this hypothesis, we prepared various peptides with different amino acids at the ligation junction including a WW domain sequence (of 35 aa) containing at non-natural residue. Some of the peptides included phosphorylated amino acids, which are often a challenge due to the steric hindrances that they introduce during couplings, as well as their tendency to undergo alpha, beta-elimination under strong reaction conditions. Our results indicate that, the presence of HOBt increases the yield and conversion of the ligation, thus reducing the possibility of side-products and degradation. These experimental conditions are especially suited for the preparation of intrinsically unfolded protein segments that are particularly prone to degradation.

MATERIALS AND METHODS

General Procedure for the Synthesis of Peptide Thioesters. Six peptide thioesters (Figure 1, Panel B) were synthesized using the SPPS approach and 4-sulfamylbutyryl resin (Supporting Information). The coupling of the first residue was performed as recommended by the manufacturer.[18] In brief: Fmoc-AA-OH (4 equiv.), DIC (4 equiv.) and 1-methylimidazole (4 equiv.) were dissolved in DCM (12 μL/μmol). The solution was added to 1 equiv. of the resin and the mixture was gently agitated for the next 18 h at 25°C. The resin was then washed 5 times with DMF and 5 times with DCM before being dried. Coupling efficiency was estimated through resin load determination by UV analysis of Fmoc-release. The remaining amino acids were incorporated manually using 5 equiv. of the corresponding amino acid derivatives activated with 4.9 equiv. of DIC in the presence of 4.9 equiv. of HOBt in DMF at 100 μmol scale using Fmoc/tBu chemistry. The efficiency of the couplings was verified with the Kaiser test.[19]

After completion of the sequences, all peptides were either acetylated or Boc-protected at the *N*-terminus prior to final peptide cleavage from the resin. The procedure for the peptide activation was the following: iodoacetonitrile (67 equiv.) and DIPEA (13 equiv.) in DMF (72 μL/μmol) were added to the resin and gently agitated for the next 18 h. The resin was then washed six times with DMF, six times with DCM and then dried. In the next step, the peptide was cleaved from the activated resin using benzyl mercaptan (50 equiv.) and DIPEA (13 equiv.) in DMF (72 μL/μmol) in an 18 h reaction with gentle agitation. The peptide solutions were filtered and collected in a round-bottom flask. The resin was washed twice with DMF (4 mL each time) and the combined filtrates were concentrated under vacuum in order to completely remove DMF. The crude product was dissolved in 6 mL of a TFA/triisopropylsilane/water mixture (95:2.5:2.5 by vol.) and stirred at room temperature for 3 h, yielding the peptide thioester without the side-chain protecting groups. The peptide thioesters were then precipitated in cold diethyl ether and centrifuged (3000*g*). The pellet was washed twice with cold ether, dried, and stored at −20°C.

The crude thioesters were analyzed by LC-MS and, depending on their purity, some of them were additionally purified by preparative RP-HPLC prior to their use in the ligation reactions.

General Procedure for Free N-Terminus Peptide Synthesis. Peptides containing a free *N*-terminus (Figure 1, Panel B) were synthesized manually using a Rink amide AM resin and Fmoc/tBu chemistry with 5 equiv. of the amino acids derivatives activated by 4.9 equiv. of DIC in the presence of 4.9 equiv. of HOBt in DMF at the 100-μmol scale. The efficiency of the manual coupling was verified by the Kaiser test.[19]

The resin-bound peptides were cleaved and deprotected with TFA containing a scavenger mixture of water, thioanisol, and ethanedithiol (90:5:2.5:2.5 by vol) at RT for 2 h. They were then precipitated in cold diethyl ether and centrifuged (3000 *g*). The pellet was washed twice with cold ether, dried and stored at −20°C.

Finally, the crude peptides were purified by preparative RP-HPLC. Pure fractions were collected, lyophilized, and stored at −20°C. They were then analyzed by MALDI-MS and LC-MS prior to their use in ligation reactions.

General Procedure for Peptide Ligation Reactions. The free *N*-terminus peptides (1.5 equiv. or 0.5 equiv.) were dissolved in 120 μL of ligation buffer *(Ligation buffer = NMP:6M GnHCL + 1M HEPES = 4:1 (v/v). A buffer containing 6 M GuHCl and 1M HEPES was prepared (50 mL) and adjusted to pH 8.5 using 25% NaOH solution and degassed. For each ligation trial 100 μL of the buffer was mixed with 400 μL of NMP*). The resulting solution was used to dissolve the thioester peptide (between 0.3 and 1.5 μmol) (*Depending of the peptide thioester amount the final volume of the ligation buffer was different (the volume of 80 μL and the described procedure is for peptide thioester amount of 1 μmol). However, the free NH_2 peptide was always 1.5 equiv. higher (*Table I*, entry 1, 2, 5, 6 and 7) or 0.5 equiv. lower (*Table I*, entry 3, 4 and 8) from the peptide thioester equivalents).* For each reaction, the solution was separated into two equal parts, and 15 μL of HOBt dissolved in the same ligation buffer (2 equiv. based on the amount of peptide thioester) were added to the ligation series with HOBt, while 15 μL of the ligation buffer was added to the other part, for comparison, as a control. For the ligations with HOAt and DIPEA, 2 equiv of HOAt, or, 4 equiv of DBU and DIPEA were added to the ligation mixture, respectively.

Finally, thiophenol (2% by volume, 1.5 μL) was added to the reactions in the presence or absence of activators, and the resulting mixture was incubated at 37°C with gentle agitation until the reaction was completed. At various time points the reaction was followed by LC-MS, and in some cases by MALDI-MS and RP-HPLC, depending on the peptide sequences and amino acids at the ligation junction (Table I, Supporting Information Table I). At each time point, 8 μL aliquots of the ligation mixture were taken and quenched by the addition of 0.1% TFA in water (32 μL) or *tris*(2-carboxyethyl)phosphine (TCEP) solution (32 μL, of a 10 mg/mL solution) when the products contained cysteine residues. The quenched mixtures were diluted up to 1.5 mL with % MeCN/0.1% FA in water (Supporting Information Table I) and stored at −20°C.

Table I Ligation Combinations Used in the Study With the Experimental Masses for the Product and Exchanged Thiophenol Thioester Used for Monitoring the Reaction Kinetics

	Ligation Combinations	Experimental *m/z* of the Product	Experimental *m/z* of Exchanged Thiophenol Thioester SC_7H_7» SC_6H_5	Yield (Y)/conv.(C) With HOBt addition (%)	Yield (Y)/conv.(C) Without HOBt Addition (%)
1	Ac-GASATVSPL***G***-SC_7H_7 + NH_2-***G***GPSPLGFLG-$CONH_2$	$[M+H]^+=1783.4$ $[M+2H]^{2+}=892.2$	$[M+H]^+=993.5$	n.d	n.d
2	Ac-GASATVSPL***S***-SC_7H_7 + NH_2-***G***GPSPLGFLG-$CONH_2$	$[M+H]^+=1813.2$ $[M+2H]^{2+}=907.1$	$[M+H]^+=1023.5$	n.d	n.d
3	Ac-LYRA***G***-SC_7H_7 + NH_2-***G***SPGYS-$CONH_2$	$[M+H]^+=1168.5$	$[M+H]^+=713.5$	89 (Y)	83 (Y)
4	Ac-LYRA***G***-SC_7H_7+ NH_2-***A***SPGYS-$CONH_2$	$[M+H]^+=1182.5$	$[M+H]^+=713.5$	n.d	n.d
5	Ac-GASATVSPL***S***-SC_7H_7+ NH_2-***V***GPSPLGFLG-$CONH_2$	$[M+H]^+=1855.4$ $[M+2H]^{2+}=928.2$	$[M+H]^+=1023.5$	36 (Y)	10 (Y)
6	Ac-GASATVSPL***V***-SC_7H_7+ NH_2-***V***GPSPLGFLG-$CONH_2$	$[M+H]^+=1867.2$ $[M+2H]^{2+}=934.1$	$[M+H]^+=1035.5$	34 (Y)	5.2 (Y)
7	Ac-GASATVSPL***V***-SC_7H_7+ NH_2-***L***GPSPLGFLG-$CONH_2$	$[M+H]^+=1881.4$ $[M+2H]^{2+}=941.2$	$[M+H]^+=1035.5$	36.7 (C)	16.7 (C)
8	Ac-LYRA***G***-SC_7H_7+ NH_2-***Y***SPGYS-$CONH_2$	$[M+H]^+=1274.6$	$[M+H]^+=713.5$	57 (Y)	33 (Y)
9	Ac-PSpSPGS***V***-SC_7H_7+ NH_2-***L***ARPSVI-$CONH_2$	$[M+H]^+=1488.6$ $[M+2H]^{2+}=744.8$	$[M+H]^+=844.8$	31.5 (C)	18.8 (C)
10	Ac-GASATVSPL***S***-SC_7H_7+ NH_2-***C***GPSPLGFLG-$CONH_2$	$[M+H]^+=1859$ $[M+2H]^{2+}=930$		n.d.	n.d

The ligation junctions are in bold and italic.

Synthesis of Modified Human Pin 1 Protein Using the Cysteine-Free Ligation Protocol With HOBt. Human Pin 1 protein was synthesized using the ligation protocol with HOBt, through ligation of segment 1 with segment 2 (Scheme 1, Supporting Information figures). Segment 1 was synthesized as thioester following the above described procedure and acetylated at the *N*-terminus. Segment 2, up to the Asn, was synthesized in a microwave-assisted peptide synthesizer (Liberty Blue, CEM Corporation). Starting from Phe, the rest of the sequence of segment 2 was coupled manually following the general procedure for peptide synthesis. The ligation was performed following the protocol described previously, using 3.2 μmol of segment 1 and 4.8 μmol of segment 2. The final volume of the ligation buffer was 700 μL due to the tendency of segment 2 to form a viscous gel (theoretically the final volume of the ligation buffer should be 240 μL). The reaction was followed by MALDI-MS and stopped after 78 h. The crude product was purified by RP-HPLC, and the fractions corresponding to the ligation product were collected, lyophilized, and stored at −20°C. In order to remove the Lys side-chain protecting group (ivdDe), the product was dissolved in DMF containing 3% of hydrazine. The solution was left at room temperature for 1h with a gentle mixing. Afterwards, DMF was removed under reduced pressure and the solid residue was dissolved in water solution containing 20% MeCN/0.1% FA. The solution was lyophilized and afterwards stored in the freezer at −20°C.

MALDI-MS. Mass spectra were acquired on a 4700 Proteomic analyzer or on a 4800 Plus MALDI TOF/TOF Analyzer (AB Sciex) calibrated with Calmix (Calmix 4700 Proteomics Analyzer Calibrating Mixture). The mass spectra were recorded in positive reflector TOF mode in the *m/z* range 500–2000 or 1500–4500 at a fixed laser intensity of 4800 using alpha-cyano-4-hydroxycinnamic acid as a matrix. Spectra were analyzed by Data Explorer software (Version 4.6, Applied Bisystems GmbH).

RP-HPLC. Crude peptides, peptide thioesters, and ligation mixtures were purified using an Aqua C_{18}-column (internal diameter 4.6 mm, length 150 mm, particle size 5 μm, pore size 12.5 nm, Phenomenx) with a linear gradient from 10% to 24% aqueous acetonitrile (0.1% TFA) over 1.2 min, followed by a gradient of 24% to 57% for the next 42 min with a flow rate of 1 mL/min using an ÄKTA purifier 10 HPLC System (GE Healthcare Life Sciences). Fractions were analyzed by MALDI-MS and those containing the products were collected, lyophilized, and stored at −20°C.

Ligation mixtures were analyzed on the microbore Aqua C_{18}-column using a linear gradient from 4.75% to 57% or from 9.5% to 57% aqueous acetonitrile (0.1% TFA) for 25 min at a flow rate of 1 mL/min.

LC-MS. The ligation mixtures (30 μL, at final concentration of approx. 60 μmol/L for each time point) were injected into the HPLC-MS system (Waters, model Alliance 2796 with a quaternary pump

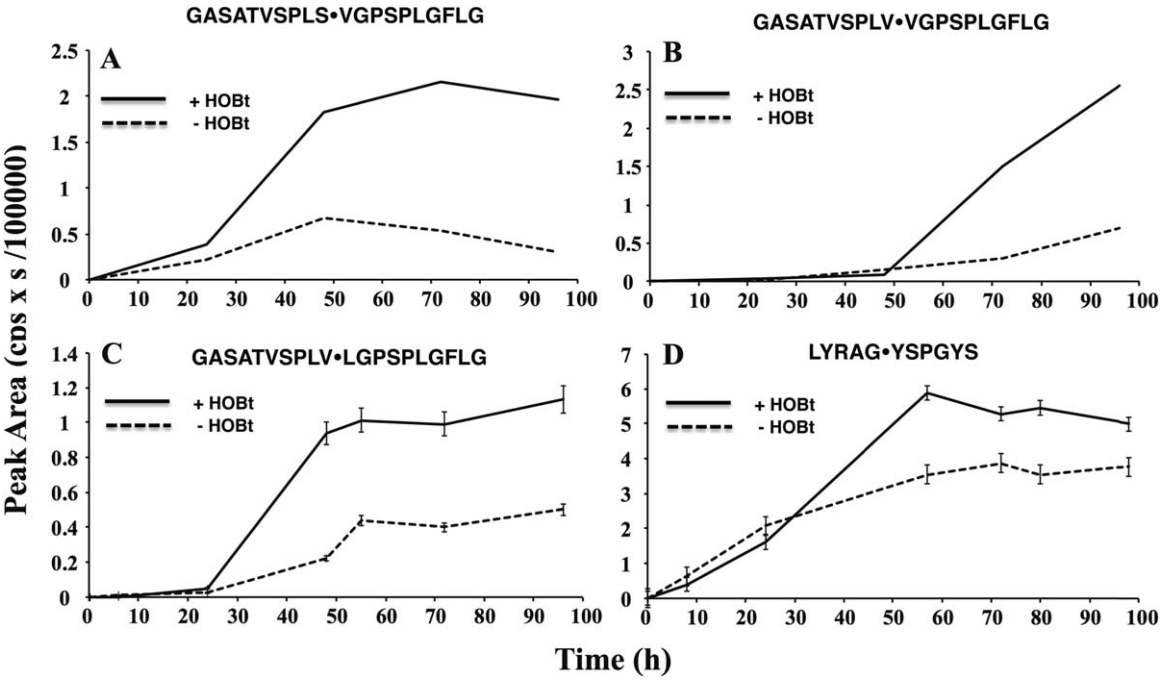

FIGURE 2 Kinetics of ligated product formation in ligation reactions with HOBt and without (solid or dashed lines respectively). The combinations shown are GASATVSPLS-VGPSPLGFLG (panel A), GASATVSPLVVGPSPLGFLG (panel B), GASATVSPLV-LGPSPLGFLG (panel C) and LYRAG-YSPGYS (panel D). The reaction was followed by LC-MS and through peak integration of the product ions at m/z 928.2 (doubly charged, panel A), m/z 934.1 (doubly charged, panel B), m/z 941.2 (doubly charged, panel C), m/z 1274.6 (singly charged, panel D).

and UV/Vis dual absorbance detector Waters 2487 connected with ESI-MS model Micromass ZQ). The separation was achieved on a Sunfire C_{18}-column (internal diameter 2.1 mm, particle size 3.5 μm, length 100 mm) using a linear gradient from 10% or 20% to 100% aqueous acetonitrile (0.1% FA) in 8 min at a flow rate of 0.3 mL/min. The mass spectra were acquired for a mass range from *m/z* 500 to 2000 in positive ion mode using five different cone voltages ranging from 5 to 70 V. The TIC spectra used for peak integration correspond to the cone voltage of 30V and were analyzed by Masslynx 4.0 software (Waters).

Kinetics of Cysteine-Free Ligation With and Without HOBt. As mentioned above, depending on the peptide sequence and amino acids at the ligation junction, the kinetics of the ligation reaction was followed using different time intervals (Supporting Information Table I). At each time point, aliquots of 8 μL of the ligation mixture were taken and quenched by the addition of 0.1% TFA in water (32 μL) or of a TCEP solution (32 μL, of a 10 mg/mL solution) if the products contained cysteine residues. The samples were diluted up to 1.5 mL with solution containing H_2O/MeCN/FA at different ratios by volume (Supporting Information, Table I, column 4) and analyzed by LC-MS, in some cases additionally by MALDI-MS or analytical RP-HPLC. The kinetics was followed through integrating the peak areas of the single or double charged ions of ligated product and exchanged thiophenol thioester (Table I). Yields were calculated after the RP-HPLC purification of the crude ligation mixture, while the given conversions where calculated from the integrated ions ratio of ligated product vs peptide thioester in the TIC spectra. In those cases, were the experiments were performed in triplicate the standard error was calculated (Reagents provided as supplementary material).

RESULTS AND DISCUSSION

To test if the addition of HOBt to the reaction protocol described by Wong and coworkers[16] can improve the rate and conversion of ligation we first selected a set of different native sequences, present in TGIF1 and FoxH1 proteins (Supporting Information Figure 1) and collected in Figure 1. The selected peptides contain β-branched (Table I, entries 5, 6 and 7) or aromatic residues (Table I, entry 8) at the ligation site, since these residues -and also phosphorylated amino acids- are often present in protein binding interfaces and are known to add complexity to the ligation reaction.[20,21] All sequences of the acetylated *C*-terminal peptide thioesters and free *N*-terminal amidated peptides used here are given in Figure 1. These peptides were subjected to the chemical ligation with and without HOBt as additive. We performed the ligation studies following the protocol described by Wong et al.,[16] in conjunction with a set of experiments where 2 equivalents of HOBt (1:2 molar ratio of peptide thioester:HOBt) were added to the ligation

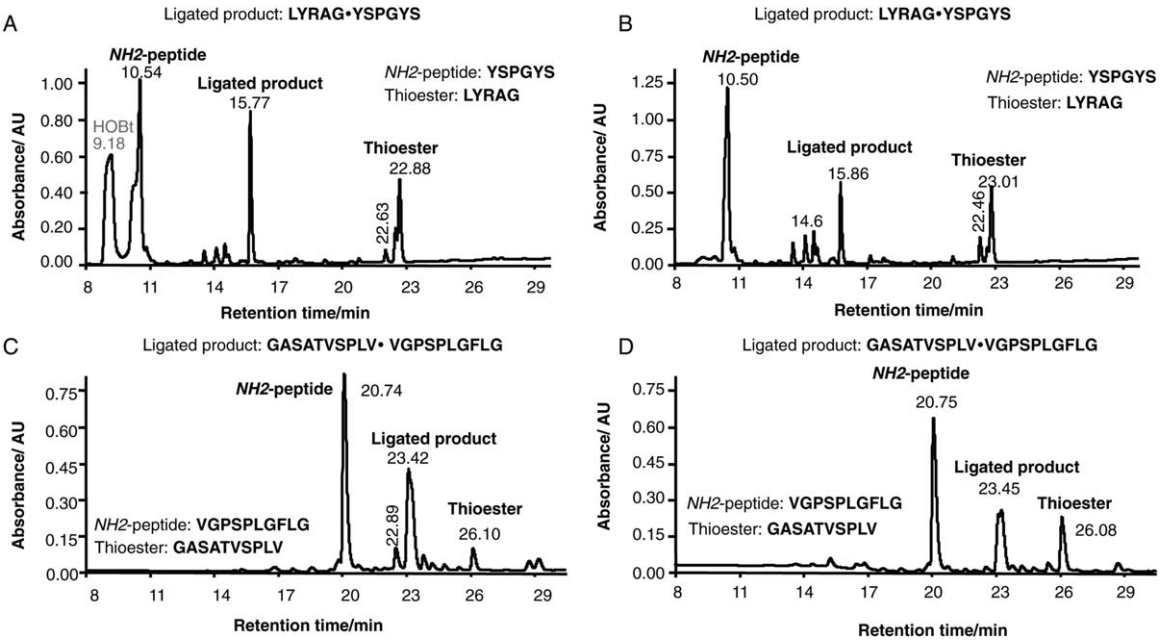

FIGURE 3 Reversed-phase HPLC chromatograms of the ligation reaction with HOBt (Panel A and C) and without HOBt (Panel B and D) using two peptides, LYRAG-YSPGYS (Panel A and B) and GASATVSPLVVGPSPLGFLG (Panel C and D).

mixture. The combinations are shown in Figure1, panel C and Table I.

In all cases, the final conversion, based on the peak integration of the product ions in the LC-MS experiments, and yield was higher when HOBt was present in the ligation mixture (Figures 1 panel C, Figure 2, Supporting Information Figure 2, Table I) independently of the peptide thioester/peptide molar ratio. Moreover, in the cases when aromatic or β-branched amino acids were present at the ligation junction, the influence of the added HOBt on the final conversion/yield was greater (Figure 1 panel C, Figure 2 bold lines) than when the peptide had less sterically hindered amino acids (Figure 1 panel C, Supporting Information Figure 2 bold lines). Interestingly, the HOBt addition specifically increased the final conversion/yield, while having very little or no influence on the rate of ligation (except in the case of GY ligation junction, Figure 2 panel D). The addition of 10 equiv. of HOBt did not improve the final conversion or the reaction rate further (Supporting Information Figure 3), but led to a reduction in the conversion rate (Supporting Information Figure 3 dashed lines). We attribute this observation to a decrease of the pH of the ligation mixture caused by the acidic nature of HOBt.[22] Furthermore, using the above-described conditions the product can be obtained with high purity since it elutes at a different retention time than the starting peptide reagents (thioester and the free N-terminus peptide), as displayed in Figure 3.

In another set of experiments performed as controls, we ligated the same set of NH_2-peptide sequences as before, in the

Table II NH_2-Free *C*-Terminal Peptide Thioesters and *N*-Terminal Peptide-Free Amines Used in the Ligation Studies

Peptide Thioesters	*m/z* theor	*m/z* exper	Peptides	*m/z* theor	*m/z* Exper	Lig. Com.
1. NH_2-GASATVSPLG-SC_7H_7	965.54	965.58	A. NH_2-GGPSPLGFLG-$CONH_2$	900.49	900.47	1-A 2-A 2-B 3-B 4-C
2. NH_2-GASATVSPLS-SC_7H_7	995.49	995.46	B. NH_2-VGPSPLGFLG-$CONH_2$	942.54	942.55	
3. NH_2-GASATVSPLV-SC_7H_7	1007.53	1007.55	C. NH_2-CGPSPLGFLG-$CONH_2$	946.48	946.46	
4. NH_2-GASATVSPLL-SC_7H_7	1021.54	1021.57				

Lig. Com.: ligation combinations.

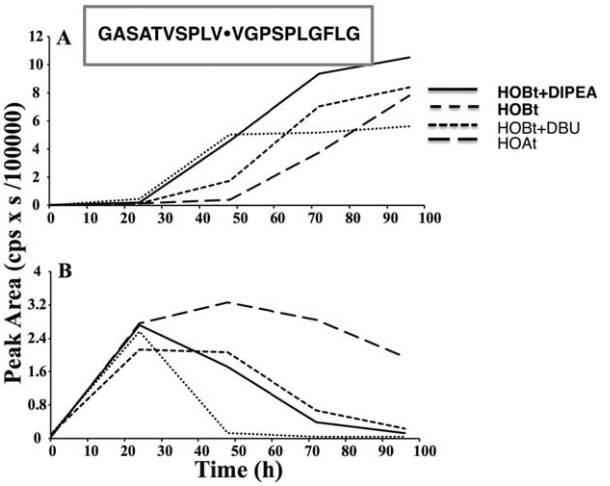

FIGURE 4 Kinetics of ligated product formation (panel A) and transthioesterification step (panel B) in the ligation reaction with HOBt+DIPEA (solid line), HOBt (dashed lines), HOBt+DBU (dotted lines) and HOAt (long-dashed lines). The ligation was performed using ligation combinations GASATVSPLV-VGPSPLGFLG. The reaction was followed by LC-MS and through peak integration of the product ion at m/z 934.1 (doubly charged, panel A), and through peak integration of the ion at m/z 1035.5 (singly charged, panel B).

presence of HOBt but using unprotected thioesters (at the N-terminus) (Table II). In all cases the formation of the expected ligation product was observed (Supporting Information Figure 4 solid lines); however, in the presence of cyclic side-products (Supporting Information Figure 4 dashed lines). According to the observed *m/z* values, the cyclization products are formed due to the intramolecular cyclization of the peptide thioesters, a competitive side reaction favored by the unprotected *N*-termini of the peptide thioester.

Finally, ligation reactions with or without *N*-terminal protected thioesters and peptides containing a N-term Cysteine were much faster than in all the other ligation reactions tested, (Supporting Information Figure 5).

The addition of HOBt suppresses the 5(4H)-oxazolone formation, preventing plausible racemization, and thus, improving the aminolysis rate through the formation of more effective acyl donors esters.[23] For this reason, we hypothesized that the presence of HOBt in the ligation mixture improves the ligation reaction yield and also contribute to increasing the conversion.

Effects of the Presence of DBU and DIPEA Activators

The addition of a base activator can improve the final conversion and speed up the reaction. To test this hypothesis, we performed ligation reactions with HOBt and DBU and DIPEA as activating bases, since they are strong bases but weak nucleophiles (Figure 4, panel A dotted line and solid line, correspondingly). To test the effect of the activators we chose Val-Val residues at the ligation site because HOBt-enhances the ligation reaction in cases where β-branched or aromatic amino acids were present at the junction (Figure 1 panel C, Figure 2 and Table I). In the case of DIPEA, we observed limited improvement in the final conversion (Figure 4, panel A solid line), while for DBU the final conversion was lower, despite the fact that up to 50 h the kinetics of the reaction was almost identical to that achieved using DIPEA (Figure 4, panel A dotted line). The observation is directly related to a transthioesterification step (which can be followed by a singly charged ion at *m/z* 1035.5), as after 50 h the number of ions corresponding to the thiophenol thioester exchange was close to zero in the case of DBU (Figure 4, panel B dotted line). Our results indicate that the presence of DBU do not contribute to increasing the ligation rate under the experimental conditions we have investigated, while in the case of DIPEA we could detect small improvements regarding the final conversion.

Potential Alternatives to HOBt

Regarding the choice of an alternative to HOBt, we used HOAt, a molecule similar to HOBt but with a lower pKa (pKa 3.28 and 4.60 respectively)[24,25] and therefore potentially a better leaving group. However, we observed that the final conversion and also the reaction rate were similar to those obtained with HOBt in the experimental conditions tested (Figure 4, panel A). This observation could be attributed to the fact that HOAt requires double the induction time relative to all other cases (Figure 4, panel A long-dashed line). This extended time can be explained by the fact that the transthioesterification is also slower for HOAt (Figure 4, panel B long-dashed line). These observations imply that if reactions were allowed to proceed for longer (more than 96 h), we would achieve a higher final conversion; however, this would also cause an increase in side products, thereby reducing the yield of the product of interest and adding complexity to the purification step. However, when the ligation reactions are more favorable, the use of HOAt could be beneficial with regards to the final conversion, as increases in the reaction time will not have a strong influence on the amount of side products.

Synthesis of a WW Domain Containing a Non-Natural Amino Acid

Finally, as an application of the ligation protocol here described we synthesized a modified WW domain using the peptidyl-prolyl isomerase I (Pin I, 34 AA, Supporting Information Scheme 1) sequence as the template. Pin1 is a modular protein that contains a WW domain, responsible for target recognition,

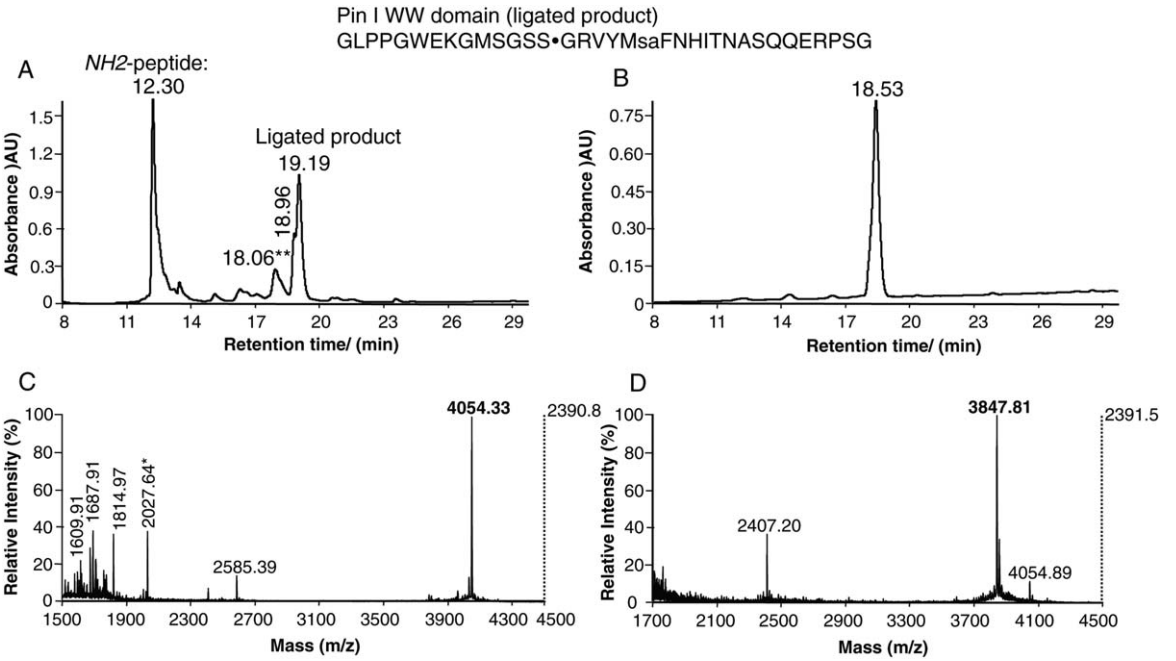

FIGURE 5 Reversed-phase chromatograms (Panel A and B) and MALDI-MS (Panel C and D) of modified human Pin I protein. After purification of the crude Pin I (Panel A) the pure sample (Panel B) was analyzed by MALDI-MS with ivdDe group present at the Lys residue (Panel C) and after removal of the ivdDe group (Panel D). The masses corresponding to the Pin 1 protein are labeled in bold, while the doubly charge ion of the Pin 1 protein is with asterisk. The peak in Panel A, labeled with double asterisk corresponds to the Pin I protein with a 16 Da higher mass.

and the catalytic domain involved in the cis-trans isomerization of phosphor-Ser/Thr-Pro bonds.[26,27] The folding pathway of the Pin1 WW domain has been the subject of many studies,[28] which have shown that the domain can be unfolded and refolded with high efficiency. In addition, many structures of this WW domain in complex with target peptides are described in the literature.[29] Since folding and binding studies often require the production of many WW proteins carrying mutations, we explored the possibility of using peptide synthesis and refolding as an efficient strategy to prepare these samples. If some of these mutations are designed to include non-natural amino acids, a ligation strategy might be the best choice to optimize the segment's design and purification. As an example, we prepared a WW sequence with a non-natural residue 2,4,6-trimethylphenyl Ala (Msa)[30–32] at position 19 and several point mutations: Lys 1 was replaced with Gly, whereas Arg 9 and Arg 12 were replaced by Ala. The replacement of both arginines prevented possible intramolecular cyclization with the guanidinium group of the Arg side-chain, while the Lys at position 8 was protected with ivdDe. The ligation junction was chosen at a Ser-Gly (position 14-15) due to the expected faster ligation rate when compared with the other options (Supporting Information Figure 2, panel B). The crude product was purified by RP-HPLC (Figure 5, panel A) and the modified Pin I was obtained with high purity (Figure 5, panel B and C). After successful removal of the ivdDe side chain group (Figure 5, panel D) the pure product was obtained at final yield of 15%. Despite the low yield obtained in this example, the synthesis of the modified Pin1 is a proof of principle of the applicability of the method, and we consider that the yield could be significantly improved with further optimizations.

Addition of HOBt to Peptides Containing Phosphorylated Residues

We also tested the HOBt-ligation protocol using peptide sequences that correspond to TGIF1 protein containing phosphorylated Ser (Table I entry 9) in the sequence of the acetylated *C*-terminal peptide thioester. The expression of proteins containing phosphorylated residues is complicated and expensive. In addition, the synthesis of phosphorylated polypeptides using microwave (MW)-assisted SPPS is inefficient due to the alpha,beta-elimination of the phosphate group under MW conditions catalyzed by piperidine.[33] Therefore, the cysteine-free ligation with HOBt could also be useful in the synthesis of phosphorylated polypeptides. The addition of HOBt significantly improved the final conversion (Figure 1 panel C, Figure 6, Table I entry 9 and Supporting Information). Furthermore,

the phosphorylated residue did not participate in any side reaction with the added HOBt, and the products that correspond to truncated peptide sequences or sequences with dephosphorylated Ser residue were not detected.

Mechanism of Reaction

Our experimental results suggest a possible general mechanism for the ligation reactions (Supporting Information Scheme 2). The initial conversion of the benzyl to the phenyl thioester is followed by the subsequent substitution by the more reactive HOBt ester. The lower pKa of HOBt (pKa 4.6) compared to that of thiophenol (pKa 6.6) would make the former a better leaving group.

Our results indicate that the limiting reaction step is the transthioesterification, as the ligation reaction does not occur in the absence of thiophenol. It is known that the thiophenol is able to undergo rapid and almost complete thiol-thioester exchange[34] and is an excellent leaving group. Therefore, in the absence of thiophenol, the initial benzyl thioesters won't be able to undergo aminolysis or to form active HOBt esters, since alkanethiols are poor leaving groups.[34] When thiophenol is replaced by 4-(carboxymethyl)thiophenol (MPAA), which has a pKa of 6.6, similar to that of thiophenol, the rate of transthioesterification is lower (Supporting Information Figure 7, panel A dashed line). We believe that the ligation buffer can affect this step since the presence of MPAA, which shows better solubility in aqueous buffers, induces a higher and faster transthioesterification than thiophenol.[34] Moreover, in our experiments, the concentration of the MPAA thioester formed after the transthioesterification remained constant during the experiment. Consequently, the ligation conversion is six times lower, as shown by peak integration in the LC-MS experiments (Supporting Information Figure 7, panel B solid and dashed line, respectively). In addition, when HOBt, was added to the ligation mixture in the absence of thiophenol, no ligation product was detected (data not shown). Again, as we mentioned before, we believe that this is due to the fact that alkanethiols are poor leaving groups and therefore the preformed benzyl thioesters are not able to react giving the active HOBt esters. In summary, the addition of HOBt facilitates the reaction of the exchanged phenyl thioester with the free *N*-terminus peptide, probably through formation of highly reactive HOBt-ester, However, it does not have a significant effect on the extent of transthioesterification. These observations are shown in Supporting Information Figure 8 (panel A-D), where the final conversion of the transthioesterification step (Supporting Information Figure 8, dotted and long-dashed lines) differs from the final conversion of the product (Supporting Information Figure 8, dashed and solid lines). In addition, the induction time of the reaction is related to the transthioesterification step, i.e., when the transthioesterification reaches the maximum and begins to drop, the formation of the product starts immediately.

CONCLUSIONS

We have found that the addition of HOBt to the ligation of *C*-terminal peptide thioesters with various free *N*-terminus peptides increases the ligation conversion, especially when β-branched or aromatic amino acids are present at the ligation junction. This increase is probably due to the *in-situ* formation of a highly reactive HOBt-ester. Furthermore, all reactions proceeded without epimerization and the method is chemoselective. Only when Lys residues are present in the sequence it is expected that the residue's side chain react with the *in-situ* formed HOBt-ester in a similar way as with a thioester previously reported by Wong et al.[16] Despite the long reaction times, the method shows potential for application in the synthesis of peptide sequences that cannot be generated by other ligation techniques, especially in the presence of β-branched or aromatic amino acids at the junction site. The improvement of the method is also applicable to the synthesis of phosphorylated peptides that, despite recent developments in protein chemistry, are not readily obtainable due to their tendency to suffer alpha,beta-elimination of the phosphate group.

T.T. is a recipient of a FP7 Marie Curie Actions (COFUND Grant program), and D.S. is a recipient of "la Caixa"-IRB Barcelona International PhD Fellowship. The authors thank IRB Barcelona for support, Dr. M. Royo for discussions regarding some peptide synthesis, Dr. M. Vilaseca and of Dr. I. Fernandez for MS analysis. They also thank Alvaro Rol Rua for the preparation of the Msa amino acid. M.J.M. is an ICREA Programme Investigator.

REFERENCES

1. Merrifield, R. B. J Am Chem Soc 1963, 85, 2149–2154.
2. Alewood, P.; Alewood, D.; Miranda, L.; Love, S.; Meutemans, W.; Wilson, D. Methods Enzym 1997, 289, 14–29.
3. Chandrudu, S.; Simerska, P.; Toth, I. Molecules 2013, 18, 4373–4388.
4. Dawson, P.; Muir, T.; Clark-Lewis, I.; Kent, S. Science 1994, 266, 776–779.
5. Dirksen, A.; Meijer, E. W.; Adriaens, W.; Hackeng, T. M. Chem Commun (Camb) 2006, 15, 1667–1669.
6. Dawson, P. E.; Kent, S. B. H. Annu Rev Biochem 2000, 69, 923–960.
7. Wombacher, R.; Cornish, V. W. J Biophotonics 2011, 4, 391–402.
8. Kemp, D. S.; Bernstein, Z. W.; McNeil, G. N. J Org Chem 1974, 69, 2831–2835.

9. Kemp, D. S.; Choong, S. L. H.; Pekaar, J. J Org Chem 1974, 39, 3841–3847.
10. Blake, J. Int J Pept Protein Res 1981, 17, 273–274.
11. Aimoto, S.; Mizoguchi, N.; Hojo, H.; Yoshimura, S. Bull Chem Soc Jpn 1989, 62, 524–531.
12. Aimoto, S. Biopolymers 1999, 51, 247–265.
13. Vila-Perelló, M.; Muir, T. W. Cell 2010, 143, 191–200.
14. Pentelute, B. L.; Kent, S. B. H. Org Lett 2007, 9, 687–690.
15. Crich, D.; Banerjee, A. J Am Chem Soc 2007, 129, 10064–10065.
16. Payne, R. J.; Ficht, S.; Greenberg, W. A.; Wong, C. H. Angew Chem Int Ed 2008, 47, 4411–4415.
17. Wang, P.; Danishefsky, S. J. J Am Chem Soc 2010, 132, 17045–17051.
18. Novabiochem, Novabiochem Innovations 4/99; 1999.
19. Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I. Anal Biochem 1970, 34, 595–598.
20. Hackeng, T. M.; Griffin, J. H.; Dawson, P. E. Proc Natl Acad Sci USA 1999, 96, 10068–10073.
21. Coltart, D. M. Tetrahedron 2000, 56, 3449–3491.
22. Saha, A.; Nadimpally, C.; Paul, A.; Kalita, S.; Mandal, B. Protein Pept Lett 2014, 21, 188–193.
23. König, W.; Geiger, R. Chem Ber 1970, 103, 788–798.
24. Novabiochem. Novabiochem Innovations: 2/09; 2009.
25. Carpino, L. A. J Am Chem Soc 1993, 115, 4397–4398.
26. Ping Lu, K.; Hanes, S. D.; Hunter, T. Nature 1996, 380, 544–547.
27. Ranganathan, R.; Lu, K. P.; Hunter, T.; Noel, J. P. Cell 1997, 89, 875–886.
28. Luo, Z.; Ding, J.; Zhou, Y. Biophys J 2007, 93, 2152–2161.
29. Aragón, E.; Goerner, N.; Zaromytidou, A. I.; Xi, Q.; Escobedo, A.; Massagué, J.; Macias, M. J. Genes Dev 2011, 25, 1275–1288.
30. Medina, E.; Moyano, A.; Pericas, M. A.; Riera, A. Helv Chim Acta 2000, 83, 972–988.
31. Martín-Gago, P.; Gomez-Caminals, M.; Ramón, R.; Verdaguer, X.; Martin-Malpartida, P.; Aragón, E.; Fernández-Carneado, J.; Ponsati, B.; López-Ruiz, P.; Cortes, M. A.; Colás, B.; MacIas, M. J.; Riera, A. Angew Chem Int Ed 2012, 51, 1820–1825.
32. Martín-Gago, P.; Aragón, E.; Gomez-Caminals, M.; Fernández-Carneado, J.; Ramón, R.; Martin-Malpartida, P.; Verdaguer, X.; López-Ruiz, P.; Colás, B.; Cortes, M. A.; Ponsati, B.; Macias, M. J.; Riera, A. Molecules 2013, 18, 14564–14584.
33. Attard, T. J.; O'Brien-Simpson, N. M.; Reynolds, E. C. Int J Pept Res Ther 2009, 15, 69–79.
34. Johnson, E. C. B.; Kent, S. B. H. J Am Chem Soc 2006, 128, 6640–6646.