ACCOUNTS of chemical research

Article

pubs.acs.org/accounts

Traceless Chemical Ligation from *S*-, *O*-, and *N*-Acyl Isopeptides

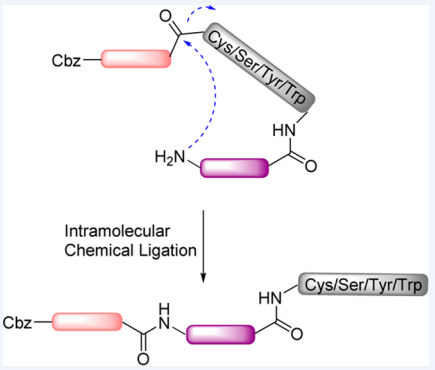
Siva S. Panda,[‡] C. Dennis Hall,[‡] Alexander A. Oliferenko,[‡] and Alan R. Katritzky*[,‡,§]

[‡]Center for Heterocyclic Compounds, Department of Chemistry, University of Florida, Gainesville, Florida 32611-7200, United States

[§]Chemistry Department, King Abdulaziz University, Jeddah, 21589 Saudi Arabia

ABSTRACT: Peptides are ubiquitous in nature where they play crucial roles as catalysts (enzymes), cell membrane ion transporters, and structural elements (proteins) within biological systems. In addition, both linear and cyclic peptides have found use as pharmaceuticals and components of various conjugate molecular systems. Small wonder then that chemists throughout the ages have sought to mimic nature by synthesis of the amide polymers known as peptides and proteins.

The fundamental reaction in the formation of a peptide bond is condensation of an amine of one amino acid with the activated carbonyl group of another. This "fragment condensation" has been achieved in many ways both in solution and by solid-phase peptide synthesis (SPSS) on resin. The most successful method for in-solution coupling is known as native chemical ligation (NCL), and the technique dates back to the pioneering work of Wieland (1953) and subsequently Kent (1994) among many others. This Account builds on the established principles of NCL as applied specifically to *S*-, *O*-, and *N*-isopeptides, molecules that are generally more soluble and less prone to aggregation than native peptides.



This Account also covers NCL of isopeptides containing terminal and nonterminal *S*-acylated cysteine units, reactions that enable the synthesis of native peptides from *S*-acyl peptides without the use of auxiliaries. With C-terminal *S*-acyl isopeptides, NCL was carried out under microwave irradiation in phosphate buffer (pH 7.3) at 50 °C. Intramolecular acyl migration was observed through 5−19-membered transition states with relative rates, as assessed by product analysis, in the order, 5 > 10 > 11 > 14, 16, or 17 > 12 > 13, 15, or 19 > 18 ≫ 9 > 8. The rate/pH profile for the 15-membered TS showed a maximum for ligated product versus transacylation at pH 7.0−7.3 presumably associated with the pK_a of the N-nucleophile in the hydrogen-bonded TS. Cysteine occurs at low abundance (1.7%) in natural peptides and is rarely available in a terminal position thus limiting the utility of the method. This Account reports, however, NCL at nonterminal acyl cysteine through 5-, 8-, 11-, and 14-membered TSs with relative rates of ligation in the order, 5 ≫ 14 > 11 ≫ 8, thus paralleling the results with acylated terminal cysteine residues.

In an obvious sequel to the work with acylated cysteine, we discuss intramolecular *O*- to *N*-acyl shift in *O*-acyl serine and *O*-acyl tyrosine isopeptides where the story becomes more complex in terms of viable conditions and optimum size of the cyclic TS. *N*- to *N*-acyl migration in acyl tryptophan isopeptides is described, and finally, chemical ligation is applied to the synthesis of cyclic peptides. Conformational analysis and quantum chemical calculations are used to rationalize ligation through a range of cyclic transition states.

This Account highlights the fact that NCL of acyl isopeptides is an extremely useful strategy for the synthesis of a wide variety of native peptides in good yields and under mild conditions. Mechanistic aspects of the ligations are not fully resolved, but theoretical studies indicate that hydrogen bonding within the various cyclic transition states plays a major role.

1. INTRODUCTION

Proteins are macromolecules that carry out most of the biochemical functions of a cell and also feature widely as structural units. The biological function of a particular protein derives from its unique folded structure, which in turn is defined by the amino acid sequence of its polypeptide chain. The importance of peptides and proteins in biology and medicine has inspired chemists for over a century to investigate methods that bring about their synthesis.[1]

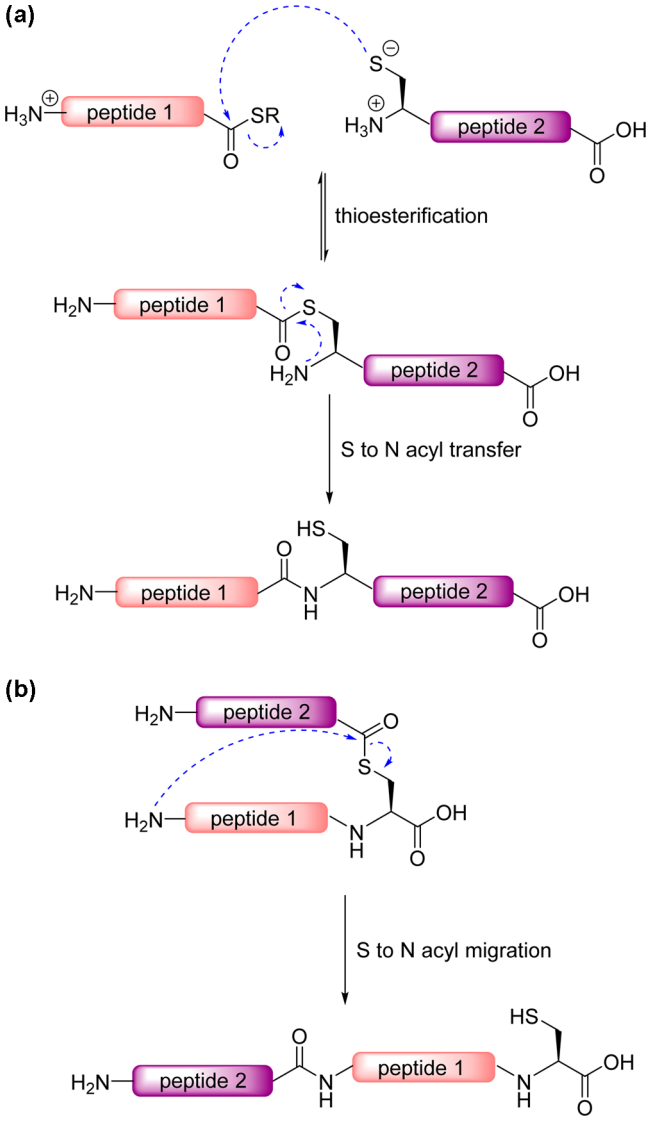
2. NATIVE CHEMICAL LIGATION

Methods for the synthesis of polypeptides have advanced greatly over several decades.[2] As a consequence of Merrifield's solid phase peptide synthesis (SPPS), the synthesis of peptides has become routine in most research laboratories.[1a] However, SPPS cannot meet the increasing need for the synthesis of large polypeptides or proteins, since it is best suited to peptide chains of less than 50 amino acids. Therefore, strategies involving solid phase peptide synthesis followed by in-solution fragment coupling have been introduced in order to facilitate the synthesis of large polypeptides.[3,4] The most successful method of in-solution fragment condensation for chemical synthesis of proteins is native chemical ligation (NCL). The concept of NCL dates back to the pioneering work of Wieland et al. (1953),[5] but it was not until 1994 that the method gained

Received: October 9, 2013
Published: March 12, 2014

   dx.doi.org/10.1021/ar400242q | *Acc. Chem. Res.* 2014, 47, 1076−1087

Scheme 1. (a) Intermolecular Chemical Ligation and (b) Intramolecular Ligation (Acyl Migration)



widespread attention when Kent and co-workers reported its application to the synthesis of interleukin-8, a cytokine responsible for the proliferation of B cells during immune response.[6] NCL involves the chemoselective coupling of two protein fragments, one containing a weakly activated C-terminal thioester and the other containing an unprotected N-terminal cysteine residue. The components combine to give a native peptide bond at the point of ligation and the fact that they do this in aqueous solution in the absence of protecting groups has placed this powerful technology at the forefront of protein synthesis. The driving force for the NCL reaction is the thermodynamic stability of an amide bond over a thioester, which occurs around neutral pH.[7] Recently Zheng et al.[8] reported new thioester equivalents for protein chemical synthesis, and Fang et al.[9] described protein chemical synthesis by ligation of peptide hydrazides. In general, NCL means intermolecular ligation with two molecules involved (Scheme 1a) whereas intramolecular ligation, often called *acyl migration*, occurs when an acyl group migrates from S → N within an isopeptide (Scheme 1b). Kemp and Galakatos reported peptide bond formation by intramolecular acyl transfer across a rigid template.[10]

3. *N*-, *O*-, AND *S*-ACYL ISOPEPTIDES

In peptide segment coupling, an N-terminal peptide fragment is coupled with a C-terminal fragment, and then the process is repeated to reach the target peptide through a convergent route. Some rare examples in which solubility problems of the fully deprotected native peptide or the parent fragments creates problems for the successful segment coupling led to the use of alternative organic solvents. A peptide segment extension from a lateral amino acid chain, rather than an extension following the α-amide backbone, defines an isopeptide bond (Figure 1).[11] The formation of a (thio)ester bond between a suitable β-hydroxyl or β-sulfhydryl at a serine, threonine, or cysteine residue leads to *O*- or *S*-acyl isopeptide bonds (Figure 1a). Similarly, amide and ester bonds are formed in tryptophan and tyrosine acyl isopeptides (Figures 1b,c). Other techniques have also been reported for the synthesis of difficult sequences, including the substitution of the amide proton with a removable protecting group or with protecting groups that induce solubilization.[4b]

An important improvement of the isopeptide method was disclosed in 2004 by Kiso and co-workers with the introduction of the *O*-acyl isodipeptide or depsidipeptide method. Facing substantial epimerization in the classical *O*-isopeptide approach, Kiso prepared a library of 40 *O*-acyl isodipeptide units with the general structure Boc-Ser/Thr(Fmoc-AA)-OH without epimerization.[12] Lewandowski et al. reported acyl transfer is a key feature of nonribosomal peptide synthesis.[13] We recently reported a simplified solution-phase entry to enantiopure *S*-, *O*-, and *N*-acyl isopeptides using acyl-benzotriazoles.[14]

The *S*-, *O*-, and *N*-acyl isopeptide methods have the major advantages of utilizing conventional amino acids and providing an isomer of the native peptide bond that can be isomerized under a variety of controlled conditions. In contrast to conventional NCL, in which two unprotected peptides are condensed by thioester-mediated amide bond formation, *N*-, *O*-, and *S*-isopeptide-based strategies require conventional chemical synthesis of an amide or (thio)ester linked peptide, which subsequently rearranges to a native peptide bond.[15] This

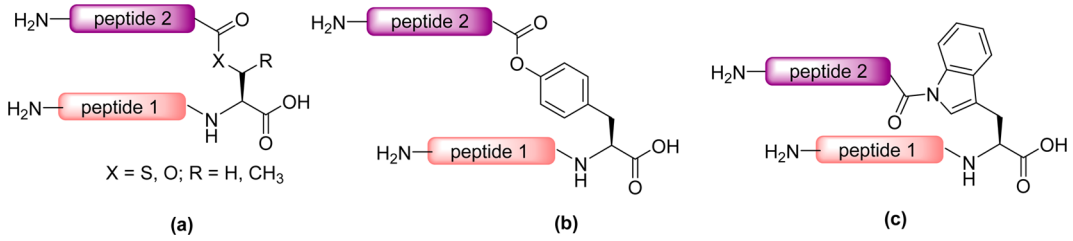


Figure 1. (a) Cysteine/serine/threonine isopeptide; (b) tyrosine isopeptide; (c) tryptophan isopeptide.

dx.doi.org/10.1021/ar400242q | *Acc. Chem. Res.* 2014, 47, 1076−1087

Table 1. Feasibility of Acyl Migration of Isopeptides via various Cyclic Transition States

sequence (isopeptide)	cyclic TS size (X → N)	condition[a]	relative area (%)[b] LP	BA
Cys(Fmoc-Phe)-OH	5 (S → N)	A	80	
Cys(Fmoc-Gly)-Gly-OMe	5 (S → N)	A	72	
Gly-Cys(Cbz-Ala)-OH	8 (S → N)	B	3	85
Gly-Cys(Cbz-Ala)-Gly-OMe	8 (S → N)	C		
β-Ala-Cys(Cbz-Ala)-OH	9 (S → N)	C	8	91
GABA-Cys(Cbz-Ala)-OH	10 (S → N)	C	80	17
Gly-Leu-Cys(Cbz-Ala)-Gly-OH	11 (S → N)	B	95	
Gly-Leu-Cys(Cbz-Ala)-Gly-OMe	11 (S → N)	C	70	30
Gly-Phe-Cys(Cbz-Ala)-Gly-OMe	11 (S → N)	C	33	67
Gly-β-Ala-Cys(Cbz-Ala)-OH	12 (S → N)	C	54	46
β-Ala-β-Ala-Cys(Cbz-Ala)-OH	13 (S → N)	C	82	17
Gly-Leu-Gly-Cys(Cbz-Ala)-Gly-OH	14 (S → N)	B	90	
Gly-Phe-Gly-Cys(Cbz-Ala)-Gly-OMe	14 (S → N)	C	57	43
β-Ala-Leu-Gly-Cys(Cbz-Ala)-Gly-OMe	15 (S → N)	C	35	65
GABA-Phe-Gly-Cys(Cbz-Ala)-Gly-OMe	16 (S → N)	C	57	43
β-Ala-Phe-β-Ala-Cys(Cbz-Ala)-Gly-OMe	16 (S → N)	C	64	36
Gly-Gly-Gly-Cys(Cbz-Ala)-OH	17 (S → N)	D	60	29
β-Ala-Gly-Gly-Cys(Cbz-Ala)-OH	18 (S → N)	D	13	82
GABA-Gly-Gly-Cys(Cbz-Ala)-OH	19 (S → N)	D	31	67
Phe-Ser-(Boc-Gly)-OH	8 (S → N)	E	57	
Phe-Ser-(Boc-Gly)-OH	8 (O → N)	E	27	
Gly-Phe-Ser-(Boc-Gly)-OH	11 (O → N)	E	99	
Gly-Phe-Ser-(Boc-Gly)-OH	11 (O → N)	E	18	31
Gly-Tyr(Cbz-Ala)-OH	12 (O → N)	E	95	
β-Ala-Tyr(Cbz-Ala)-OH	13 (O → N)	E	85	10
GABA-Tyr(Cbz-Ala)-OH	14 (O → N)	E	97	
Gly-Gly-Tyr(Cbz-Ala)-OH	15 (O → N)	D	95	5
β-Ala-Gly-Tyr(Cbz-Ala)-OH	16 (O → N)	D	87	9
β-Ala-β-Ala-Tyr(Cbz-Ala)-OH	17 (O → N)	D	97	3
β-Ala-GABA-Tyr(Cbz-Ala)-OH	18 (O → N)	D	100	
GABA−GABA-Tyr(Cbz-Ala)-OH	19 (O → N)	D	98	
Trp(Cbz-Ala)-OH	7 (N → N)	E	2	
Gly-Trp(Cbz-Ala)-OH	10 (N → N)	E	44	19
β-Ala-Trp(Cbz-Ala)-OH	11 (N → N)	E	71	
GABA-Trp(Cbz-Ala)-OH	12 (N → N)	E	99	

[a](A) CH_3CH/H_2O (3:1), 20 °C, 2 h; (B) phosphate buffer (pH = 7.8), MW (50 °C, 50 W), 1 h; (C) CH_3CN/H_2O (7:1), MW (70 °C, 50 W), 3−4 h; (D) phosphate buffer (pH = 7.3), MW (50 °C, 50 W), 3 h; (E) piperidine−DMF 20% v/v, 1−3 h. [b]Semiquantitative determination by HPLC-MS. The area of ion-peak resulting from the sum of the intensities of the $[M+H]^+$ and $[M+Na]^+$ ions for each compound was integrated (corrected for starting material). LP = ligated peptide; BA = bis-acylated product.

rearrangement may proceed through various transition states associated with X-to-*N*-acyl transfer (X = *S*, *O*, or *N*) (Table 1).

4. CHEMICAL LIGATIONS FROM CYSTEINE ISOPEPTIDES

4.1. Chemical Ligation of *S*-Acyl Isopeptides Containing C-Terminal Cysteine Residues

Although NCL has been studied extensively in peptidic compounds bearing a cysteine residue at the N-terminus,[16] an alternative approach for the synthesis of cysteine-containing peptides uses isopeptide ligation methodology. *S*-Acylated cysteine isopeptides form native peptides via intramolecular chemical ligation by an entropically favored mechanism. Yoshiya et al.[17] synthesized *S*-acyl peptides containing N-terminal cysteine residues; subsequent *S* → *N* intramolecular acyl migration furnished native peptide bonds. Our group studied chemical ligation of S-acylated cysteine isopeptides via various cyclic transition states. This method allows the synthesis of native peptides from *S*-acyl isopeptides with a C-terminal cysteine without utilizing auxiliaries (Scheme 2).[18−22]

The feasibility of intramolecular acyl migrations via 5−19-membered cyclic transition states was demonstrated. Long-range *S*- to *N*-acyl transfers were found to depend significantly on the size of the macrocyclic transition state (TS), with relative rates, as judged from yields, following the TS ring-size trend, 5 > 10 > 11 > 14, 16, or 17 > 12 > 13, 15, or 19 > 18 ≫ 9 > 8.[18−22]

Chemical ligation is pH sensitive: ligation usually proceeds rapidly around pH 7 but is less efficient at pH < 5.5.[23] In some cases, the yields for traceless Staudinger ligation in water increased to 80% at pH ≥ 8.5.[24] The ratio of the ligation product vs intermolecular transacylation for a 15-membered TS was studied under different pH conditions (Scheme 3). The relative rate of ligation versus transacylation (based on product ratio) varied with pH (Table 2) with optimum formation of ligated product at pH 7.0−7.3.[21]

Further evidence for intramolecular ligation as the origin of compound **8** (Scheme 4) was provided by a competitive experiment in which chemical ligation of *S*-(Pg-R-aminoacyl)-tripeptide **6** (via an 11-membered cyclic transition state) was carried out in the presence of 5 equiv of dipeptide **7** (H-Gly-Leu-OH). HPLC analysis of the isolated product showed 70% of ligated product **8** together with 10% of the bisacylated product **9**. Compound **10**, the N-acylated product of dipeptide **7**, was not detected.

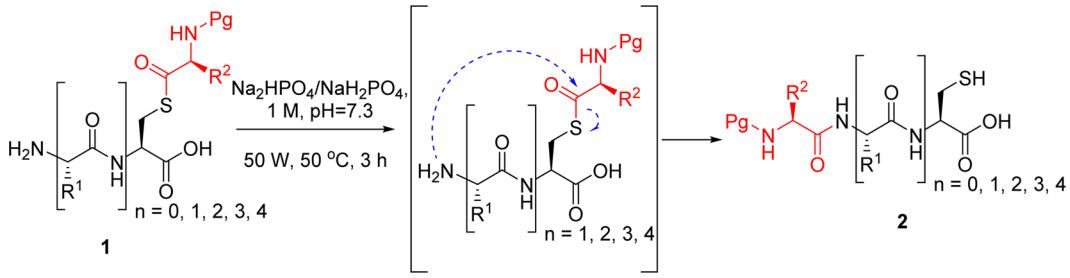
Thus, nucleophilic attack of the SH group of **8** on **6** is the most likely source of the bis-acylated product **9**.

4.2. Chemical Ligation of *S*-Acyl Isopeptides Containing Nonterminal Cysteine Residues

The development of a synthetic strategy for the preparation of a polypeptide by means of native chemical ligation relies on a suitable cysteine ligation site. Cysteine, however, is a relatively rare amino acid (1.7% average content)[3b] and is not always available in a terminal position. Haase and Seitz reported a chemical ligation approach utilizing internal cysteine residues to accelerate thioester-based peptide ligations. This method, however, requires relatively long reaction times (48−72 h), and the ligation products were not isolated.[25]

Our previously reported isopeptide methodology[18−22] was used to synthesize several novel *S*-acyl peptides containing internal cysteine residues. Using these isopeptides, we achieved the first isopeptide ligations to form native peptides from nonterminal cysteine residues. The chemical ligations via 5-, 8-, 11-, and 14-membered cyclic transition states showed that the 8-membered transition state was disfavored, whereas the

Scheme 2. Ligation Studies via *S*- to *N*-Acyl Migration



Scheme 3. Chemical Ligation Reaction of 3 via 15 TS

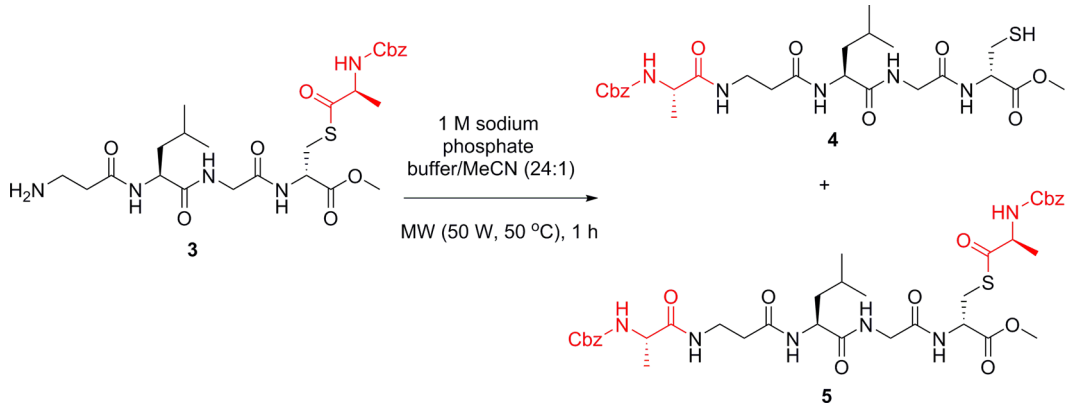


Table 2. Dependence on pH of the Product Ratio 4/5 in the Ligation Reaction of 3

		product ratio[a]	
entry	pH	ligated peptide[b] **4**	bis-acylated product **5**
1	6.2	20	80
2	7.0	45	55
3	7.3	43	57
4	7.6	36	64
5	8.2	15	85

[a]Determined by HPLC–MS semiquantitation. The area of the ion peak resulting from the sum of the intensities of the $[M + H]^+$ and $[M + Na]^+$ ions of each compound was integrated. [b]Analyzed as disulfide dimer.

remaining transition states gave yields of **13** in the order, 5 ≫ 14 > 11 ≫ 8 (Scheme 5).[26]

We also repeated the chemical ligation of *S*-acyl tetrapeptide **14** in the presence of 10 equiv of glycinamide **15** under similar reaction conditions. Glycinamide **15** was chosen as a model compound to mimic possible intermolecular reaction at the thioester moiety. Compound **18**, the potential product of N-acylation of **15**, was not detected by HPLC-MS (ESI), but the mixture contained the ligation product **16** and the intermolecular bis-acylated product **17** in 30:70 ratio (Scheme 6) thus providing further support for the intramolecular nature of the isopeptide ligation.

Homocysteine-containing isopeptide **19** was also prepared, and ligation via a 6-membered cyclic transition state was studied. The formation of ligated product with 98% abundance was observed by HPLC-MS analysis, showing that *S* → *N*-acyl group migration via a 6-membered transition state is preferred over intermolecular acylation (Scheme 7).[27]

5. CHEMICAL LIGATIONS AT NONCYSTEINE SITES

In attempts to overcome the low abundance of cysteine,[3b] considerable effort has been devoted to developing auxiliary thiol groups, but their use encountered difficulties due to steric hindrance and problems associated with extraneous groups in the ligated product.[28] Furthermore, the position of a cysteine residue in a peptide sequence is often not compatible with an efficient retrosynthetic disconnection. Synthetic approaches have therefore been developed to extend NCL to unprotected peptides at non-cysteine sites. Seminal work by Dawson et al. achieved ligation at alanine, combining cysteine-based NCL with desulfurization to alanine on the peptide construct.[29]

Raney Ni[29,30] and Pd/Al_2O_3 under hydrogen[29] were reported as metal-based desulfurization techniques. Wan and Danishefsky[31] introduced a mild and convenient metal-free protocol for desulfurization of thiols by organophosphites.[32] Hojo et al.[33] and Li et al.[34] developed thiol-containing auxiliaries for native chemical ligation.

6. CHEMICAL LIGATIONS FROM SERINE ISOPEPTIDES

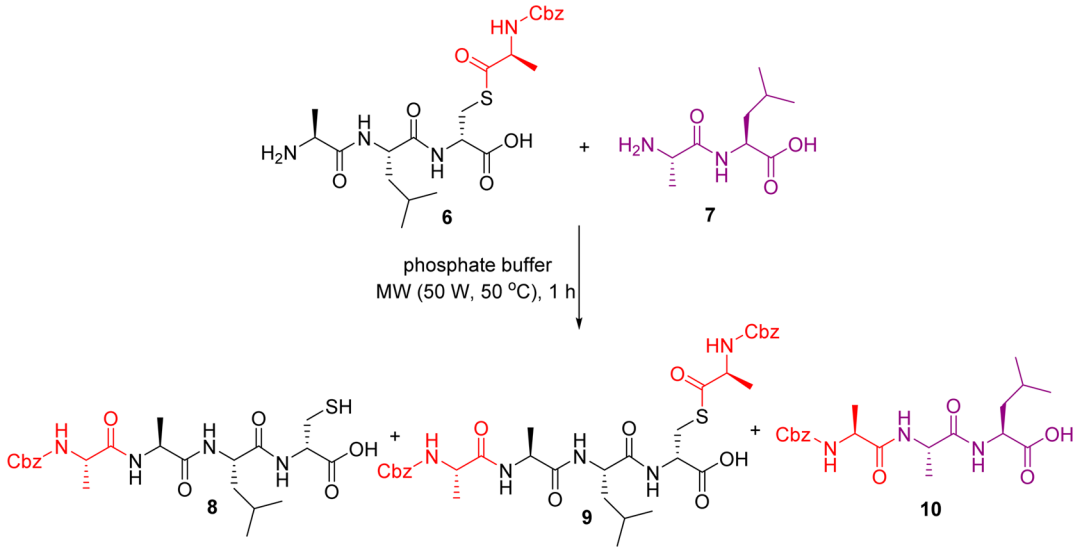
Kiso et al.[35] demonstrated that *O*-acyl residues within a backbone significantly altered the secondary structure of the native peptide. "*O*-Acyl isopeptides" are more hydrophilic and easier to purify by HPLC than the corresponding native peptides. N-Terminal serine isopeptides rapidly generate the corresponding native peptides by *O* → *N* intramolecular acyl migration via a 5-membered transition state.[36]

We also observed this classic *O* → *N*-acyl shift via 8- and 11-membered cyclic transition states in serine isopeptides.[37] Thus "traceless" chemical ligation involving *O*- to *N*-acyl shift without cysteine or an auxiliary group at the ligation site was achieved.[37]

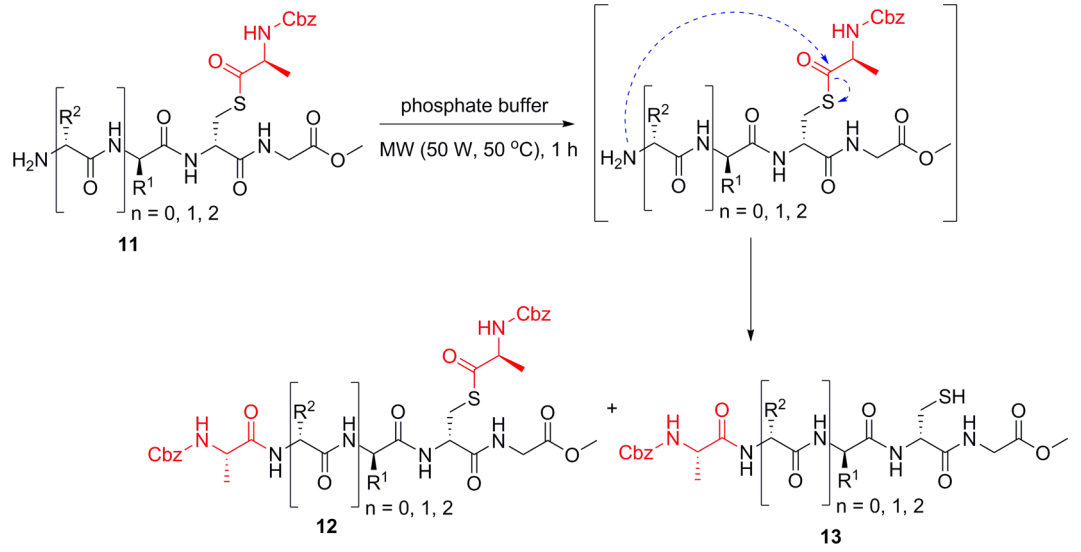
Deprotected *O*-acyl isodipeptides (**22** and **24**) were subjected to microwave irradiation in piperidine–DMF 20% v/v, 50 °C, 50 W, 1 h (Scheme 8) under anhydrous conditions

dx.doi.org/10.1021/ar400242q | Acc. Chem. Res. 2014, 47, 1076–1087

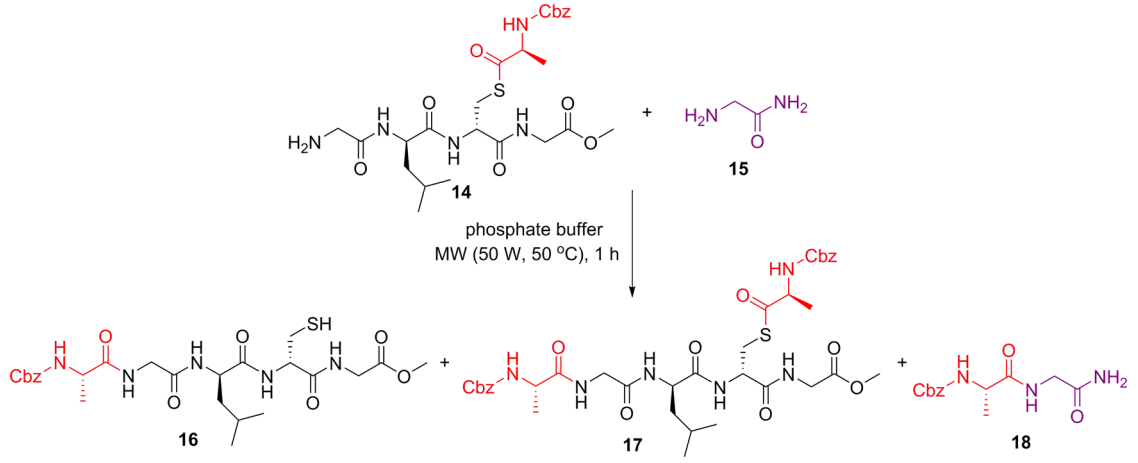
Scheme 4. Competitive Chemical Ligation of Cysteine Isopeptide 6



Scheme 5. Chemical Ligation of Nonterminal Cysteine Isopeptides 11



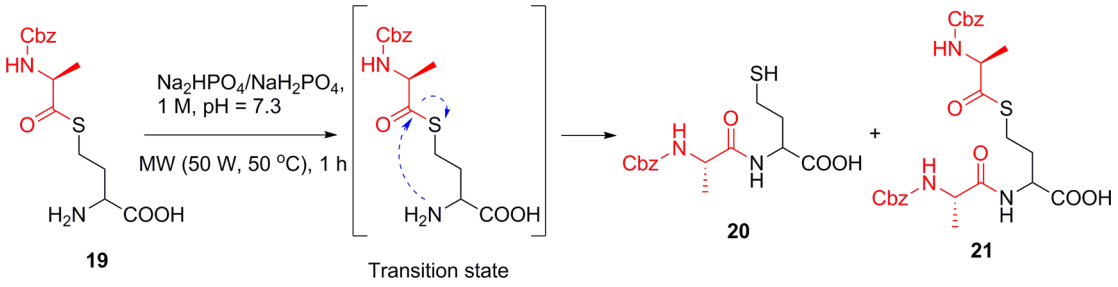
Scheme 6. Competitive Chemical Ligation of Cysteine Isopeptide 14



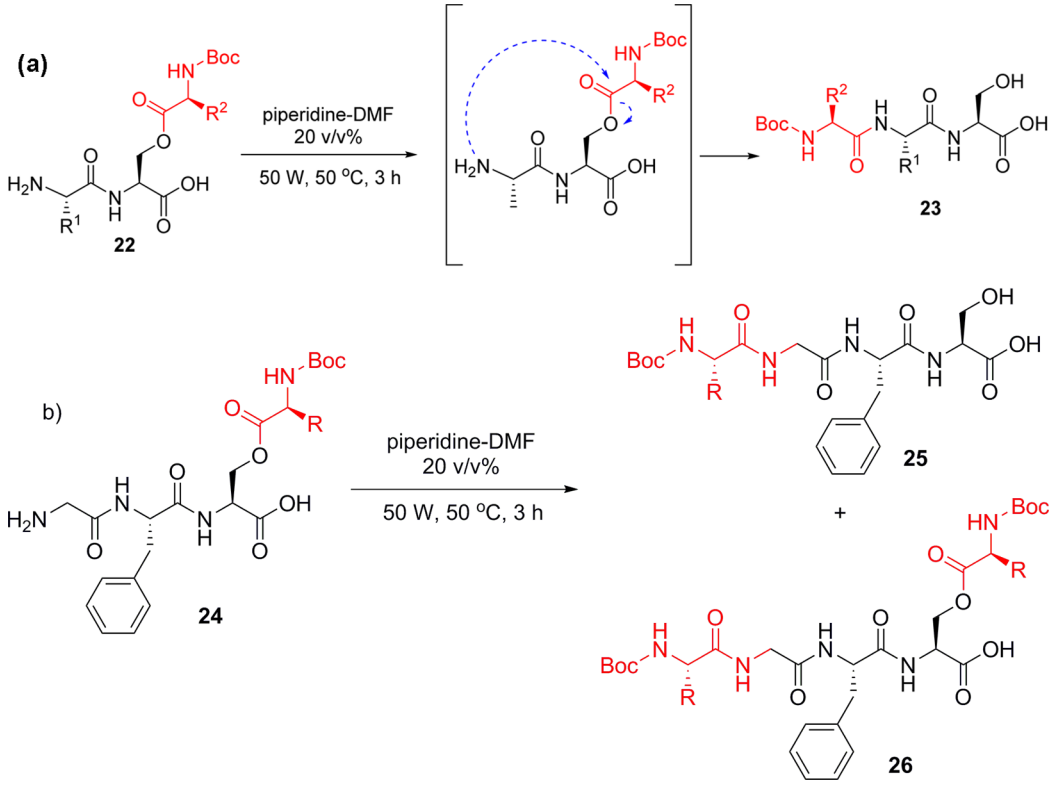
chosen to avoid ester hydrolysis. HPLC-MS indicated the formation of intramolecular ligated products (**23** and **25**, respectively), and HPLC-MS, via (−)ESI-MS/MS, confirmed that isopeptides and ligated products with the same molecular

dx.doi.org/10.1021/ar400242q | *Acc. Chem. Res.* 2014, 47, 1076−1087

Scheme 7. Chemical Ligation of Homocysteine Isopeptide 19



Scheme 8. Chemical Ligation of Serine Isopeptide: (a) 8-membered TS; (b) 11-membered TS



weight had different mass fragmentation patterns. The results revealed formation of ligated products through both 8- and 11-membered transition states, whereas in the cysteine case, an 8-membered transition state was disfavored even under basic conditions.

Ligation of **24** was also examined under aqueous conditions (pH 7.6, 1 M buffer strength, MW 50 °C, 50 W, 1 h), but HPLC-MS of the aqueous product revealed only a small amount of ligated product **25**, together with a major peak that corresponding to removal of the Boc-group from **24**.

7. CHEMICAL LIGATIONS FROM TYROSINE ISOPEPTIDES

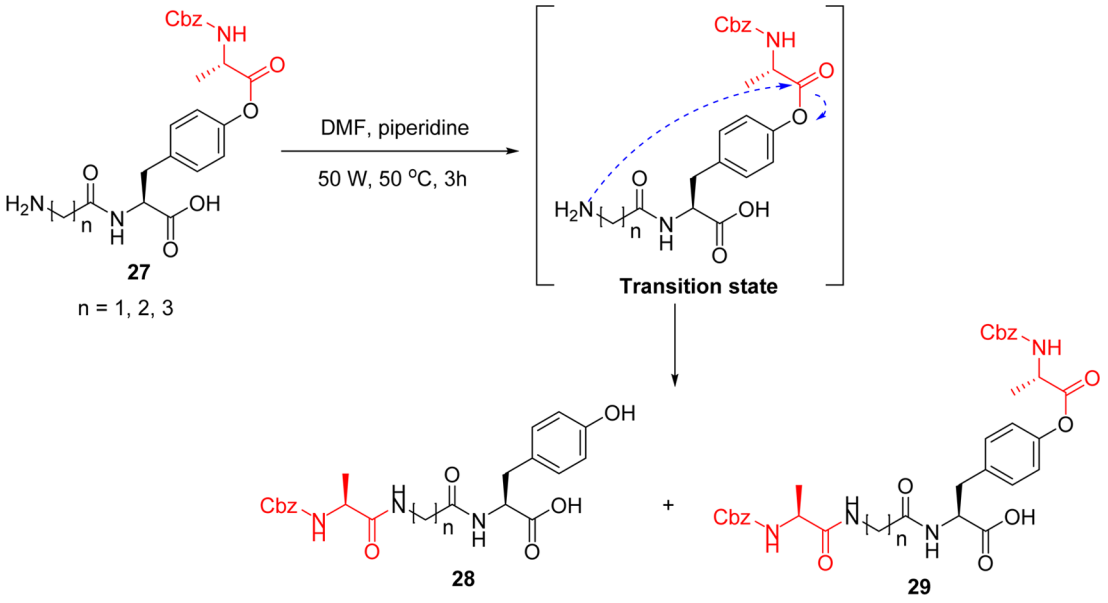
Tyrosine isotripeptides **27** were used for ligation studies via 12−14-membered cyclic transition states and classical coupling of tyrosine isotripeptides **27** with α-, β-, or γ-amino acids to provide the starting isotetrapeptides **30** for ligation studies via 15−19-membered cyclic transition states.[38] In order to enhance migration rates, glycine, β-alanine, and GABA units were used in the *O*-acylisopeptide intermediates.

Attempts to ligate **27** (Scheme 9) under aqueous conditions (pH 7.3, 1 M buffer concentration, MW 50 °C, 50 W, 3 h) failed to yield the ligated product via a 13-membered transition state. However when the reaction was carried out under microwave irradiation (Scheme 8), HPLC−MS indicated the formation of 85% of the intramolecular ligated product **28**, with 10% bis-acylated product **29** and 5% of the starting material. The retention times and fragmentation patterns of **27** and **28** were also studied in control experiments. HPLC−MS, via (−)ESI-MS/MS, confirmed that compounds **27** and **28**, each with MW 457, had different fragmentation patterns, proving intramolecular ligation to **28**. The results demonstrate that *O* → *N*-acyl group migrations via 12- and 14-membered cyclic transition states are even more highly preferred over intermolecular acylation than with the 13-membered cyclic transition state.

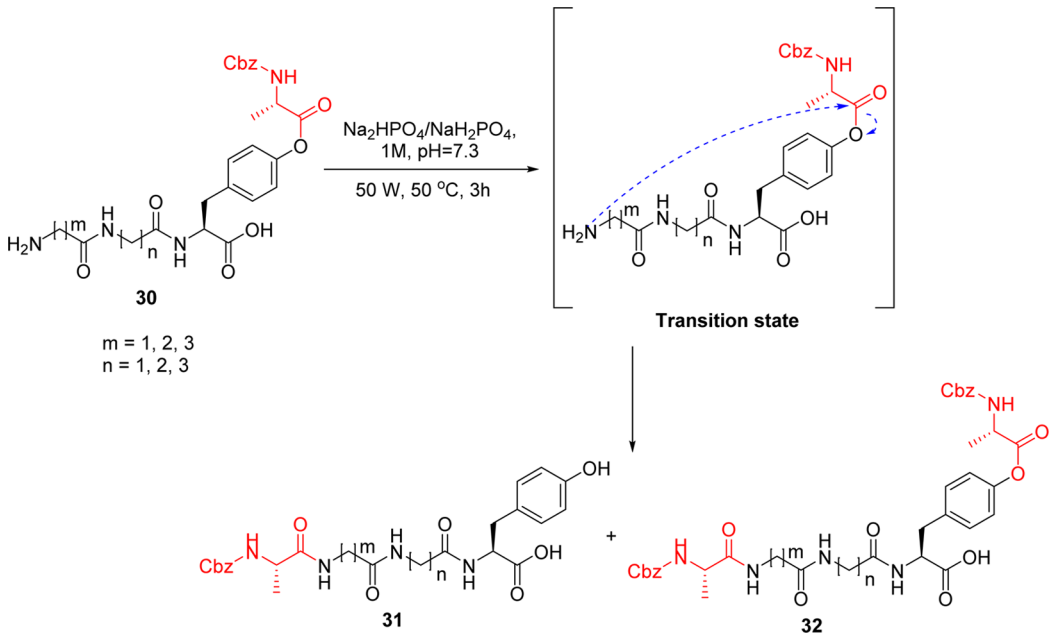
Under aqueous conditions (pH = 7.3), the 12-, 13-, and 14-membered ring transition state systems gave no ligation. To explain this, force field calculations (see section 10 for more details) were used to show that preorganization in aqueous media was poor in terms of both strain energy and preferred

dx.doi.org/10.1021/ar400242q | *Acc. Chem. Res.* 2014, 47, 1076−1087

Scheme 9. Chemical Ligation of Tyrosine Isopeptides 27



Scheme 10. Chemical Ligation of Tyrosine Isopeptides 30



conformation.[39] However in a DMF−piperidine mixture, the attacking nucleophile is expected to be totally deprotonated, and ligation takes place efficiently, with yields as high as 97%.

Tyrosine isotetrapeptides **30** under aqueous conditions produce the expected ligation products **31** (87−100%, Scheme 10). In the case of 15-, 16-, and 17-membered transition states, HPLC−MS (ESI) showed that intermolecular bis-acylated products were also formed together with small amounts of unreacted **30**.

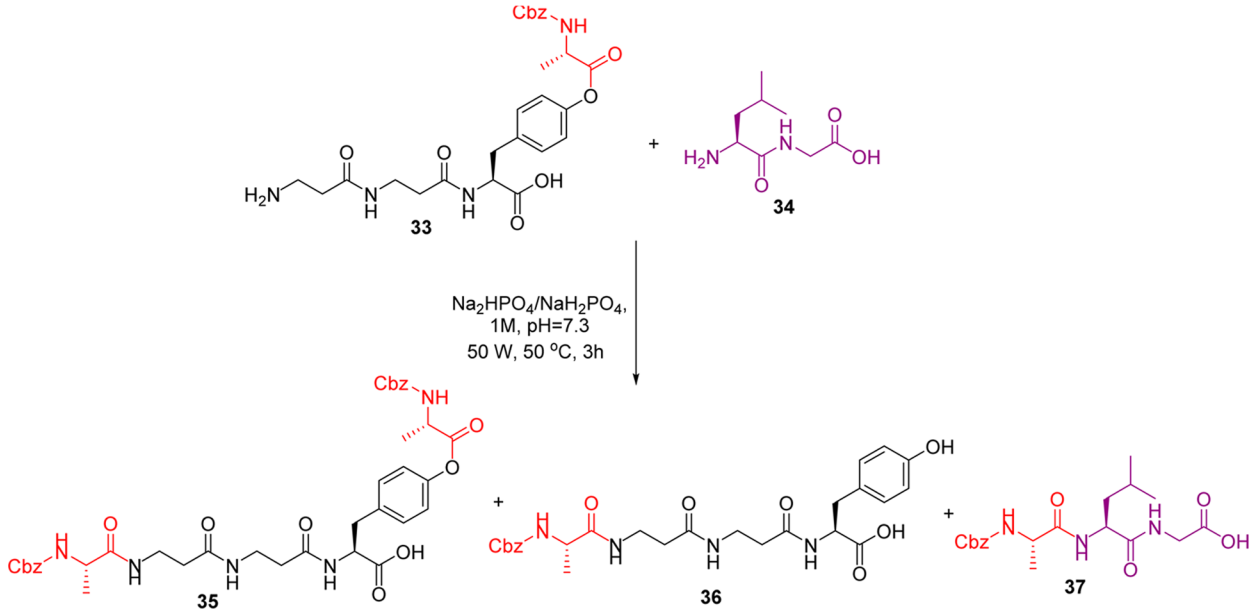
The intramolecular ligation of N-terminal unprotected *O*-isopeptides **27** and **30** to form native peptides **28** and **31**, respectively, through 12−19-membered cyclic transition states was supported by a competitive ligation of N-terminal unprotected *O*-isotetrapeptide **33** in the presence of 5 equiv of dipeptide **34** (H-Leu-Gly-OH) under aqueous conditions (Scheme 11). HPLC−MS analysis of the crude product confirmed the formation of 96% of ligation product **35** together with 3% of intermolecular bis-acylated product **36** and 1% of the starting material **33**. Cbz-protected tripeptide **37**, which is the N-acylated product of dipeptide **34**, was not detected by HPLC−MS.

8. CHEMICAL LIGATIONS OF TRYPTOPHAN ISOPEPTIDES

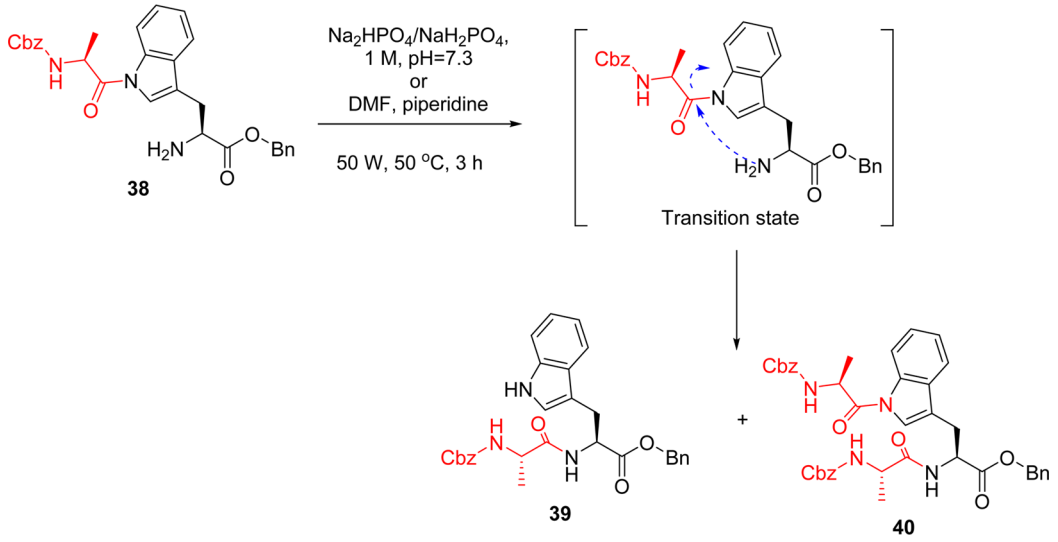
Tryptophan isodipeptide **38** was synthesized in order to study *N*-acyl migration from indole nitrogen to the N-terminal group of tryptophan via a 7-membered cyclic transition state. It also served as starting material for the synthesis of tryptophan isotripeptides, which were used to study the possibility of $N \rightarrow$ *N*-acyl migration via 10-, 11-, and 12-membered cyclic transition states.[40]

dx.doi.org/10.1021/ar400242q | *Acc. Chem. Res.* 2014, 47, 1076−1087

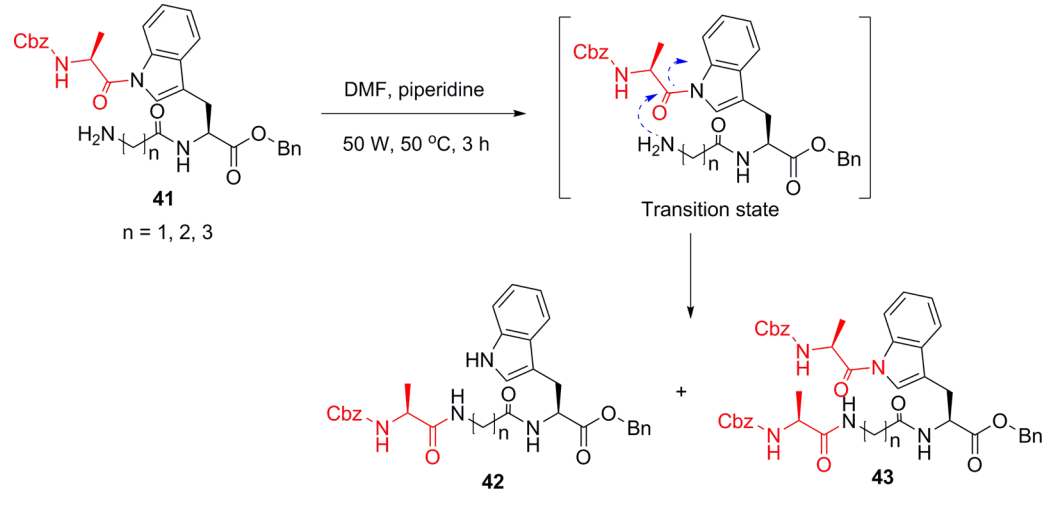
Scheme 11. Competitive Chemical Ligation of Tyrosine Isopeptide 33



Scheme 12. Chemical Ligation of Tryptophan Isopeptides 38



Scheme 13. Chemical Ligation of Tryptophan Isopeptides 41



dx.doi.org/10.1021/ar400242q | Acc. Chem. Res. 2014, 47, 1076−1087

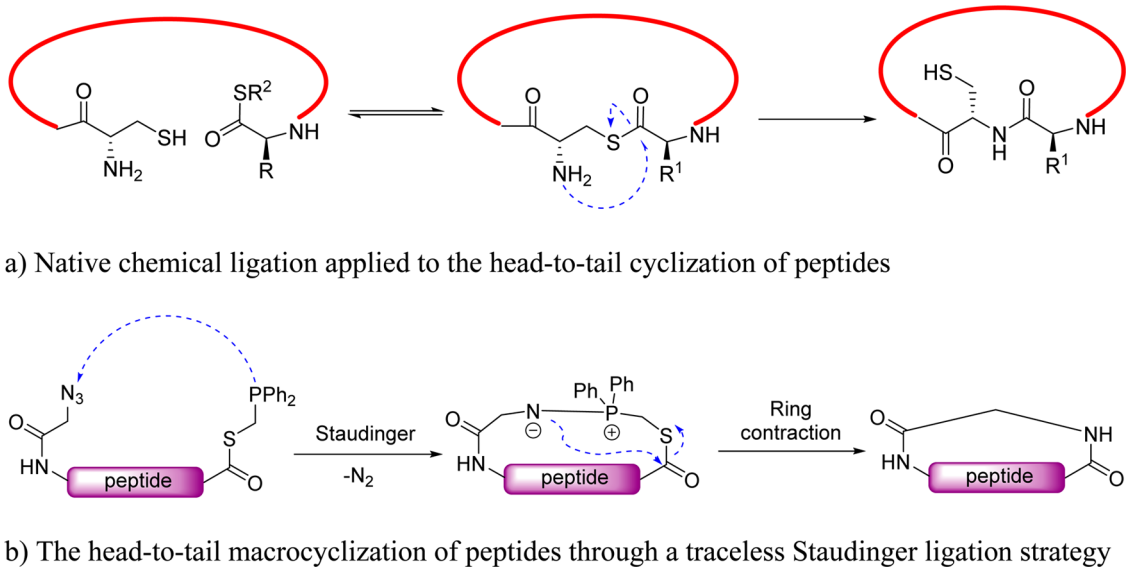
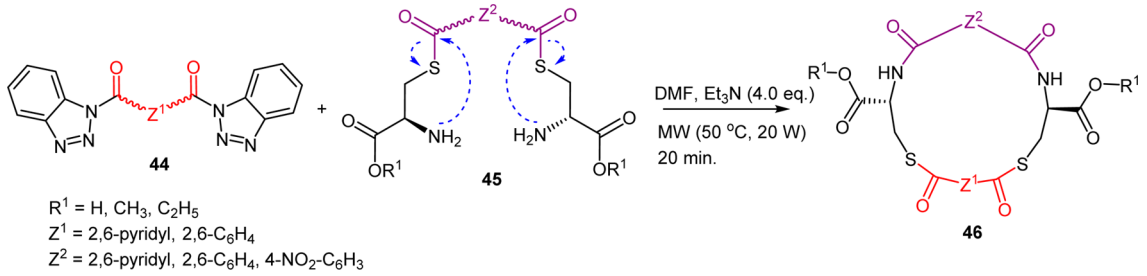


Figure 2. Peptide macrocyclizations.

Scheme 14. Preparation of Cyclic Peptidomimetics 46



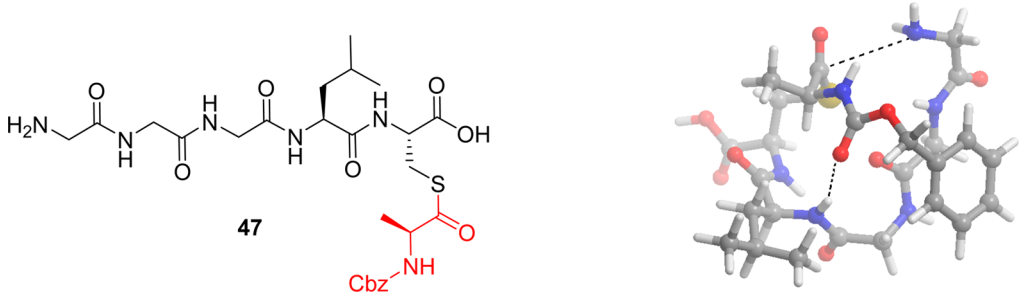


Figure 3. Preorganized conformation of *S*-acyl hexapeptide **47**.[20] Proximity between the nucleophilic and electrophilic sites as well as a stabilizing hydrogen bond are shown in dash lines.

Chemical ligation via a 7-membered cyclic transition state was investigated by microwave irradiation of **38** at pH 7.4 (Scheme 12). No ligation was observed under aqueous conditions, but HPLC-MS of the ligation mixture under basic conditions (DMF, piperidine) revealed a small amount (2%) of the ligated product **39**, together with a major peak corresponding to starting material. Recently Mhidia et al. reported an *N* to *N*-acyl shift for asparagine native peptide bond formation.[41]

When tryptophan isotripeptides **41** were irradiated under aqueous conditions, the ligated products were obtained in 29% (n = 2) and 6% (n = 3) yields via 11- and 12-membered transition states, respectively. However, under the same microwave conditions but in DMF−piperidine, ligated products were obtained via 10-, 11-, and 12-membered cyclic transition states in 44%, 71%, and 99% yields (Scheme 13).

9. SYNTHESIS OF CYCLIC PEPTIDES VIA CHEMICAL LIGATION

The topology of macrocyclic peptides decreases their susceptibility to attack by exo- and endopeptidases; thus, they play an important role in nature and are attractive targets for drug discovery. Vancomycin, cyclosporine, tyrocidine A, oxytocin, and vasopressin are well-known cyclic peptide drugs.[42]

Depending on its functional groups, a peptide can be cyclized in four different ways: head-to-tail (C-terminus to N-terminus), head-to-side chain, side chain-to-tail, or side-chain-to-side-chain. In particular, native chemical ligation offers a powerful

dx.doi.org/10.1021/ar400242q | *Acc. Chem. Res.* 2014, 47, 1076−1087

tool to couple peptide fragments under mild conditions.[6] Tam and Pallin reported cyclization of C-terminal peptide thioesters containing an N-terminal cysteine residue in a head-to-tail fashion (Figure 2a).[43] A traceless Staudinger ligation of thioester-based macrocyclization strategy was developed by Hackenberger and Kleineweischede (Figure 2b).[44]

We adapted benzotriazole methodology for the synthesis of cysteine containing macrocycles via chemical ligation.[45,46] The methodology for the regioselective syntheses of *S*- and *N*-acylcysteines was also developed recently by using *N*-acylbenzotriazoles under mild conditions.[47] Coupling of *N*-acylbis(benzotriazoles) with cysteine gave bis(*S*-acylcysteines), which on treatment with another equivalent of *N*-acylbis-(benzotriazole), afforded cyclic peptide mimetics in 64−72% yields (Scheme 14).

10. COMPUTATIONAL RATIONALIZATION OF CHEMICAL LIGATION REACTIVITY

The chemical ligations discussed above show variable yields and reactivity trends that are difficult to rationalize. For example, in a series of structurally similar isopeptides differing only in the size of a cyclic transition state, intra- versus intermolecular acylation can vary from almost a quantitative yield of native peptide to almost zero. One example of such reactivity emerged in our study of *S*- to *N*-acyl shifts in cysteine-containing di-, tri-, and tetraisopeptides.[18,19] Of those isopeptides, only the di- and tetra-analogs formed ligation products in high yields, whereas the triisopeptide formed predominantly an intermolecular acylation product. An attempt was made by Oliferenko and Katritzky[48] to rationalize this curious behavior in terms of theory and computational chemistry based on the following assumptions. First, since chemical ligation occurs through a cyclic transition state, preorganization is extremely important. The more easily the starting material achieves an appropriate cyclic conformation, the higher the probability of intramolecular reaction. Second, specific interactions can influence preorganization and hence reactivity. Thus, hydrogen bonding and NH−π interactions were identified as important factors in the stabilization of a preorganized conformer.[22] The geometric proximity of the nucleophile (amino group) to the electrophilic center (thioester carbonyl group) is an obvious criterion of preorganization. A virtual screening protocol was designed to prioritize thousands of structures generated by a full conformational search using molecular mechanics. More precise reaction energies were calculated using *ab initio* quantum chemical methods. This theoretical approach was applied to several of our chemical ligations including the di-, tri-, and tetra-*S*-acyl peptides of which the tripeptide was unreactive. Conformational analysis revealed poor preorganization of the tripeptide compared with the di- and tetrapeptides, and the geometrical distance between the nucleophilic and electrophilic sites was large (3.591 Å compared with 2.998 and 3.085 Å in the di- and tetrapeptides, respectively). Thus the curious reactivity pattern was explained by a combination of conformational analysis and quantum chemical calculations.[48] In a subsequent study, an isosteric tripeptide was synthesized in which cysteine was replaced by serine, and this tripeptide surprisingly showed an efficient *O* to *N* shift.[37] This reactivity was also explained using the same computational methodology. Long-range *S*- to *N*-migrations involving 13-, 15-, 16-, 17-, 18-, and 19-membered cyclic transition states were rationalized,[20,21] and the cyclic conformer of an S-acyl hexapeptide,[20] preorganized to react through a 17-membered transition state, is shown in Figure 3.

11. CONCLUSIONS

Microwave assisted isopeptide ligation offers the following advantages: (i) short reaction times (1−3 h) at a moderate temperature (50 °C), (ii) chemical ligation from terminal and nonterminal cysteine, serine, tyrosine, and tryptophan residues, and (iii) avoidance of ligation auxiliaries. In contrast to the classical native chemical ligation approach, the methodology allows isolation of the *S*-acyl peptide intermediates, which are potentially useful for several synthetic and biological applications.

■ AUTHOR INFORMATION

Notes

The authors declare no competing financial interest.

Biographies

Siva S. Panda was born in Odisha, India, and received his Bachelor degree (B.Pharm) from the Roland Institute of Pharmaceutical Sciences, Berhampur, India, in 2002 and Master's degree (M.Pharm) from Manipal College of Pharmaceutical Sciences, Manipal, India, in 2005. He then joined Dabur Research Foundation, Sahibabad, India, as a Research Scientist for a period of one year and worked on the synthesis of mitoxantrone, an anticancer drug. Siva obtained his Ph.D. degree in synthetic organic chemistry in 2010 under the supervision of Prof. Subhash C. Jain at University of Delhi, Delhi, India. The main focus of his doctoral research was the synthesis and biological evaluation of heterocyclic pharmacophoes, and phytochemical investigation/bioassay of medicinally important plants. Currently, he is working as a senior group leader with Prof. Alan R. Katritzky at the Center of Heterocyclic Compounds, University of Florida, Gainesville, Florida, USA, where his research involves the synthesis of novel peptides, peptide bioconjugates, peptidomimetics in solution phase, and ligation studies of various cyclic transition states in $S \rightarrow N$, $N \rightarrow N$, and $O \rightarrow N$ acyl migration.

C. Dennis Hall after retiring from his academic position at King's College, London, in 1999, joined Alan Katritzky's research group at the University of Florida where he acts as a group leader, Administrator for the on-line journal Arkivoc, and co-organizer of the Florida Heterocyclic/Synthesis conferences (Flohet). Since joining the Katritzky team, he has coauthored some 50 papers in the fields of heterocyclic chemistry, QSAR, insect control, and synthetic ion channels.

Alexander Oliferenko obtained his Ph.D. from the Moscow State University in 2000 under the supervision of Prof. Nikolai Zefirov. Since then he worked as a postdoctoral researcher at the University of Florida from 2001 to 2002, at the Pacific Northwest National Laboratory in 2004−2006, and at the Queen's University of Belfast in 2008−2011. In 2003−2004 and 2006−2010, he held permanent positions of research scientist and senior research investigator at the Moscow State University. From March 2011 to July 2013, he was working again in the Katritzky lab, now on theoretical studies of chemical ligation.

Alan R. Katritzky was educated at Oxford and Cambridge (lecturer and Founder Fellow of Churchill College). Founder Dean of the School of Chemical Sciences at East Anglia from 1962, he transferred in 1980 as inaugural Kenan Professor to the University of Florida. His research in heterocyclic chemistry covered *inter alia* N-oxides, benzotriazole methodology, electrophilic and nucleophilic substitution, computational QSPR relationships, and peptide chemistry. He holds 14 honorary doctorates from 10 Eurasian countries and associate or foreign membership of five national academies. He is Cavalieri

dx.doi.org/10.1021/ar400242q | *Acc. Chem. Res.* 2014, 47, 1076−1087

Ufficiale (Italy) and Honorary Fellow of St. Catherine's College, Oxford, and of the Polish and Italian Chemical Societies. Over 1,000 graduate students and postdoctoral associates have trained in his group. He created the nonprofit Arkat USA Inc., is organizer of the Flohet Conferences, and is publisher of the open access journal Arkivoc completely free to authors or readers. Contributions to the secondary literature include editing *Comprehensive Heterocyclic Chemistry* (40 volumes in 3 editions), *Advances in Heterocyclic Chemistry* (106 volumes), *Handbook of Heterocyclic Chemistry* (3rd edition, 2010), and *Heterocycles in Life and Society* (2nd edition, 2011).

REFERENCES

(1) (a) Merrifield, R. B. Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J. Am. Chem. Soc.* **1963**, *85*, 2149−2154. (b) Kent, S. Origin of the chemical ligation concept for the total synthesis of enzymes (proteins). *Biopolymers* **2010**, *94*, iv−ix.

(2) Muralidharan, V.; Muir, T. W. Protein ligation: An enabling technology for the biophysical analysis of proteins. *Nat. Methods* **2006**, *3*, 429−438.

(3) (a) Pattabiraman, V. R.; Bode, J. W. Rethinking amide bond synthesis. *Nature* **2011**, *480*, 471−479. (b) Hackenberger, C. P. R.; Schwarzer, D. Chemoselective ligation and modification strategies for peptides and proteins. *Angew. Chem., Int. Ed.* **2008**, *47*, 10030−10074.

(4) (a) Monbaliu, J. C.; Katritzky, A. R. Recent trends in Cys- and Ser/Thr-based synthetic strategies for the elaboration of peptide constructs. *Chem. Commun.* **2012**, *48*, 11601−11622. (b) Raibaut, L.; Ollivier, N.; Melny, O. Sequential native peptide ligation strategies for total chemical protein synthesis. *Chem. Soc. Rev.* **2012**, *41*, 7001−7015.

(5) Wieland, T.; Bokelmann, E.; Bauer, L.; Lang, H. U.; Lau, H.; Schafer, W. Polypeptide syntheses. VIII. Formation of sulfur containing peptides by the intramolecular migration of aminoacyl groups. *Justus Liebigs Ann. Chem.* **1953**, *583*, 129−149.

(6) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. Synthesis of proteins by native chemical ligation. *Science* **1994**, *266*, 776−779.

(7) Kent, S. B. H. Total chemical synthesis of proteins. *Chem. Soc. Rev.* **2009**, *38*, 338−351.

(8) Zheng, J.-S.; Tang, S.; Huang, Y.-C.; Liu, L. Development of new thioester equivalents for protein chemical synthesis. *Acc. Chem. Res.* **2013**, *46*, 2475−2484.

(9) Fang, G.-M.; Li, Y.-M.; Shen, F.; Huang, Y.-C.; Li, J.-B.; Lin, Y.; Cui, H.-K.; Liu, L. Protein chemical synthesis by ligation of peptide hydrazidesa. *Angew. Chem., Int. Ed.* **2011**, *50*, 7645−7649.

(10) Kemp, D. S.; Galakatos, N. G. Peptide synthesis by prior thiol capture. 1. A convenient synthesis of 4-hydroxy-6-mercaptodibenzofuran and novel solid-phase synthesis of peptide-derived 4-(acyloxy)-6-mercaptodibenzofurans. *J. Org. Chem.* **1986**, *51*, 1821−1829.

(11) Geurink, P. P.; Oualid, F. El; Jonker, A.; Hameed, D. S.; Ovaa, H. A general chemical ligation approach towards isopeptide-linked ubiquitin and ubiquitin-like assay reagents. *ChemBioChem* **2012**, *13*, 293−297.

(12) (a) Sohma, Y.; Sasaki, M.; Hayashi, Y.; Kimura, T.; Kiso, Y. Design and synthesis of a novel water-soluble Aß 1−42 isopeptide: An efficient strategy for the preparation of Alzheimer's disease-related peptide, Aβ1−42, via O-N intramolecular acyl migration reaction. *Tetrahedron Lett.* **2004**, *45*, 5965−5968. (b) Yoshiya, T.; Taniguchi, A.; Sohma, Y.; Fukao, F.; Nakamura, S.; Abe, N.; Ito, N.; Skwarczynski, M.; Kimura, T.; Hayashi, Y.; Kiso, Y. "*O*-Acyl isopeptide method" for peptide synthesis: Synthesis of forty kinds of "*O*-acyl isodipeptide unit" Boc-Ser/Thr(Fmoc-Xaa)-OH. *Org. Biomol. Chem.* **2007**, *5*, 1720−1730.

(13) Lewandowski, B.; De Bo, G.; Ward, J. W.; Papmeyer, M.; Kuschel, S.; Aldegunde, M. J.; Gramlich, P. M. E.; Heckmann, D.; Goldup, S. M.; D'Souza, D. M.; Fernandes, A. E.; Leigh, D. A. Sequence-specific peptide synthesis by an artificial small-molecule machine. *Science* **2013**, *339*, 189−193.

(14) (a) Liaqat, S.; Panda, S. S.; Rauf, A.; Al-Youbi, A. O.; Katritzky, A. R. Solution-phase synthesis of chiral *N*-, *O*-, and *S*-acyl isopeptides. *Synthesis* **2014**, *46*, 67−72. (b) El Khatib, M.; Jauregui, L.; Tala, S. R.; Khelashvili, L.; Katritzky, A. R. Solution-phase synthesis of chiral O-acyl isodipeptides. *MedChemComm* **2011**, *2*, 1087−1092.

(15) Panda, S. S.; Hall, C. D.; Scriven, E.; Katritzky, A. R. Aminoacyl benzotriazolides: Versatile reagents for the preparation of peptides and their mimetics and conjugates. *Aldrichimica Acta* **2013**, *46*, 43−55.

(16) Reulen, S. W. A.; Brusselaars, W. W. T.; Langereis, S.; Mulder, W. J. M.; Breurken, M.; Merkx, M. Protein−liposome conjugates using cysteine-lipids and native chemical ligation. *Bioconjugate Chem.* **2007**, *18*, 590−596.

(17) Yoshiya, T.; Ito, N.; Kimura, T.; Kiso, Y. Isopeptide method: Development of *S*-acyl isopeptide method for the synthesis of difficult sequence-containing peptides. *J. Pept. Sci.* **2008**, *14*, 1203−1208.

(18) Katritzky, A. R.; Abo-Dya, N. E.; Tala, S. R.; Abdel-Samii, Z. K. The chemical ligation of selectively *S*-acylated cysteine peptides to form native peptides via 5-, 11- and 14-membered cyclic transition states. *Org. Biomol. Chem.* **2010**, *8*, 2316−2319.

(19) Katritzky, A. R.; Tala, S. R.; Abo-Dya, N. E.; Ibrahim, T. S.; El-Feky, S. A.; Gyanda, K.; Pandya, K. M. Chemical ligation of *S*-acylated cysteine peptides to form native peptides via 5-, 11-, and 14-membered cyclic transition states. *J. Org. Chem.* **2011**, *76*, 85−96.

(20) Panda, S. S.; El-Nachef, C.; Bajaj, K.; Al-Youbi, A. O.; Oliferenko, A.; Katritzky, A. R. Study of chemical ligation via 17-, 18-, and 19-membered cyclic transition states. *Chem. Biol. Drug Des.* **2012**, *80*, 821−827.

(21) Ha, K.; Chahar, M.; Monbaliu, J.-C. M.; Todadze, E.; Hansen, F. K.; Oliferenko, A. A.; Ocampo, C. E.; Leino, D.; Lillicotch, A.; Stevens, C. V.; Katritzky, A. R. Long-range intramolecular *S*→*N* acyl migration: A study of the formation of native peptide analogues via 13-, 15-, and 16-membered cyclic transition States. *J. Org. Chem.* **2012**, *77*, 2637−2648.

(22) Bol'shakov, O.; Kovacs, J.; Chahar, M.; Ha, K.; Khelashvili, L.; Katritzky, A. R. *S*- to *N*-Acyl transfer in *S*-acylcysteine isopeptides via 9-, 10-, 12-, and 13-membered cyclic transition states. *J. Pept. Sci.* **2012**, *18*, 704−709.

(23) Hondal, R. J.; Nilsson, B. L.; Raines, R. T. Selenocysteine in native chemical ligation and expressed protein ligation. *J. Am. Chem. Soc.* **2001**, *123*, 5140−5141.

(24) Tam, A.; Soellner, M. B.; Raines, R. T. Water-soluble phosphinothiols for traceless staudinger ligation and integration with expressed protein ligation. *J. Am. Chem. Soc.* **2007**, *129*, 11421−11430.

(25) Haase, C.; Seitz, O. Internal cysteine accelerates thioester-based peptide ligation. *Eur. J. Org. Chem.* **2009**, 2096−2101.

(26) Hansen, F. K.; Ha, K.; Todadze, E.; Lillicotch, A.; Frey, A.; Katritzky, A. R. Microwave-assisted chemical ligation of *S*-acyl peptides containing non-terminal cysteine residues. *Org. Biomol. Chem.* **2011**, *9*, 7162−7167.

(27) Panda, S. S.; Katritzky, A. R. Ligation from homocysteine isopeptide via 6-membered cyclic transition state. *J. Indian Chem. Soc.* **2013**, *90*, 1519−1524.

(28) (a) Macmillan, D.; Anderson, D. W. Rapid synthesis of acyl transfer auxiliaries for cysteine-free native glycopeptide ligation. *Org. Lett.* **2004**, *6*, 4659−4662. (b) Offer, J.; Boddy, C. N. C.; Dawson, P. E. Extending synthetic access to proteins with a removable acyl transfer auxiliary. *J. Am. Chem. Soc.* **2002**, *124*, 4642−4646.

(29) Yan, L. Z.; Dawson, P. E. Synthesis of peptides and proteins without cysteine residues by native chemical ligation combined with desulfurization. *J. Am. Chem. Soc.* **2001**, *123*, 526−533.

(30) Bang, D.; Makhatadze, G. I.; Tereshko, V.; Kossiakoff, A. A.; Kent, S. B. Total chemical synthesis and X-ray crystal structure of a protein diastereomer: [D-gln 35]Ubiquitin. *Angew. Chem., Int. Ed.* **2005**, *44*, 3852−3856.

(31) Wan, Q.; Danishefsky, S. J. Free-radical-based, specific desulfurization of cysteine: a powerful advance in the synthesis of polypeptides and glycopolypeptides. *Angew. Chem., Int. Ed.* **2007**, *46*, 9248−9252.

(32) Hoffmann, F. W.; Ess, R. J.; Simmons, T. C.; Hanzel, R. S. The desulfuration of mercaptans with trialkyl phosphites. *J. Am. Chem. Soc.* **1956**, *78*, 6414.

(33) Hojo, H.; Ozawa, C.; Katayama, H.; Ueki, A.; Nakahara, Y.; Nakahara, Y. The mercaptomethyl group facilitates an efficient one-pot ligation at xaa-ser/thr for (glyco)peptide synthesis. *Angew. Chem., Int. Ed.* **2010**, *49*, 5318−5321.
(34) Li, J.; Cui, H.-K.; Liu, L. Peptide ligation assisted by an auxiliary attached to amidyl nitrogen. *Tetrahedron Lett.* **2010**, *51*, 1793−1796.
(35) Sohma, Y.; Yoshiya, T.; Taniguchi, A.; Kimura, T.; Hayashi, Y.; Kiso, Y. Development of *O*-acyl isopeptide method. *Biopolymers* **2007**, *88*, 253−262.
(36) Sohma, Y.; Sasaki, M.; Hayashi, Y.; Kimura, T.; Kiso, Y. Novel and efficient synthesis of difficult sequence-containing peptides through *O-N* intramolecular acyl migration reaction of O-acyl isopeptides. *Chem. Commun.* **2004**, 124−125.
(37) El Khatib, M.; Elagawany, M.; Jabeen, F.; Todadze, E.; Bol'shakov, O.; Khelashvili, L.; El-Feky, S. A.; Asiri, A.; Katritzky, A. R. Traceless chemical ligations from *O*-acyl serine sites. *Org. Biomol. Chem.* **2012**, *10*, 4836−4838.
(38) Popov, V.; Panda, S. S.; Katritzky, A. R. Ligations from tyrosine isopeptides via 12- to 19-membered cyclic transition states. *J. Org. Chem.* **2013**, *78*, 7455−7461.
(39) Panda, S. S.; Oliferenko, A. A.; Marwani, H. M.; Katritzky, A. R. Effects of preorganization and hydrogen bonding on intramolecular chemical ligation of (*N*)- and (*O*)-acyl isopeptides. *Mendeleev Commun.* **2014**, in press.
(40) Popov, V.; Panda, S. S.; Katritzky, A. R. Ligations of *N*-acyl tryptophan units to give native peptides via 7-, 10-, 11- and 12-membered cyclic transition states. *Org. Biomol. Chem.* **2013**, *11*, 1594−1597.
(41) Mhidia, R.; Boll, E.; Fecourt, F.; Ermolenko, M.; Ollivier, N.; Sasaki, K.; Crich, D.; Delpech, B.; Melnyk, O. Exploration of an imide capture/N,N-acyl shift sequence for asparagine native peptide bond formation. *Bioorg. Med. Chem.* **2013**, *21*, 3479−3485.
(42) White, C. J.; Yudin, A. K. Contemporary strategies for peptide macrocyclization. *Nat. Chem.* **2011**, *3*, 509−524 and references therein..
(43) Zhang, L.; Tam, J. P. Synthesis and application of unprotected cyclic peptides as building blocks for peptide dendrimers. *J. Am. Chem. Soc.* **1997**, *119*, 2363−2370.
(44) Kleineweischede, R.; Hackenberger, C. P. R. Chemoselective peptide cyclization by traceless ligation. *Angew. Chem., Int. Ed.* **2008**, *47*, 5984−5988.
(45) Berhanu, W. M.; Ibrahim, M. A.; Pillai, G. G.; Oliferenko, A. A.; Khelashvili, L.; Jabeen, F.; Mirza, B.; Ansari, F. L.; ul-Haq, I.; El-Feky, S. A.; Katritzky, A. R. Similarity analysis, synthesis, and bioassay of antibacterial cyclic peptidomimetics. *Beilstein J. Org. Chem.* **2012**, *8*, 1146−1160.
(46) Ibrahim, M. A.; Panda, S. S.; Nhon, L.; Hamed, A.; El-Feky, S. A.; Katritzky, A. R. Macrocyclic peptoids by selective S-acylation of cysteine esters. *Synthesis* **2013**, *45*, 767−772.
(47) Katritzky, A. R.; Tala, S. R.; Abo-Dya, N. E.; Gyanda, K.; El-Gendy, B. E.-D. M.; Abdel-Samii, Z. K.; Steel, P. J. Selective synthesis and structural elucidation of *S*-acyl- and *N*-acylcysteines. *J. Org. Chem.* **2009**, *74*, 7165−7167.
(48) Oliferenko, A. A.; Katritzky, A. R. Alternating chemical ligation reactivity of *S*-acyl peptides explained with theory and computations. *Org. Biomol. Chem.* **2011**, *9*, 4756−4759.