Available online at www.sciencedirect.com
ScienceDirect
Mendeleev Commun., 2014, **24**, 75–77
Mendeleev Communications

Effects of preorganization and hydrogen bonding on intramolecular chemical ligation of (*N*)- and (*O*)-acyl isopeptides

Siva S. Panda,[a] Alexander A. Oliferenko,[a] Hadi M. Marwani[b] and Alan R. Katritzky*[a,b]

[a] *Center for Heterocyclic Compounds, Department of Chemistry, University of Florida, Gainesville, FL 32611-7200, USA. Fax: +1 352 392 9199; e-mail: katritzky@chem.ufl.edu*
[b] *Chemistry Department, King Abdulaziz University, Jeddah, 21589 Saudi Arabia*

DOI: 10.1016/j.mencom.2014.03.002

The ease of intramolecular *N*- to *N*- and *O*- to *N*-acyl transfers in (*N*)-acyl isopeptides and (*O*)-acyl isopeptides was found to be governed by preorganization rationalized in terms of geometrical and energetic characteristics of the linear isopeptide backbone, which depends on the chain length. Intramolecular hydrogen bonding was found another important factor of reactivity able to stabilize the chemical ligation transition state.

Chemical ligations have facilitated the synthesis of large peptides by linking the C-terminus of one unprotected peptide with the N-terminus of another.[1–3] Native chemical ligation (NCL), first reported by Wieland *et al*.[4] and later developed by Kent,[1,2(a)] is a chemoselective and regioselective linking of a peptide-thioester and a terminal Cys-peptide resulting in a native amide bond at the ligation site after rapid NCL *S*- to *N*-acyl transfer *via* a cyclic transition state.[1–3]

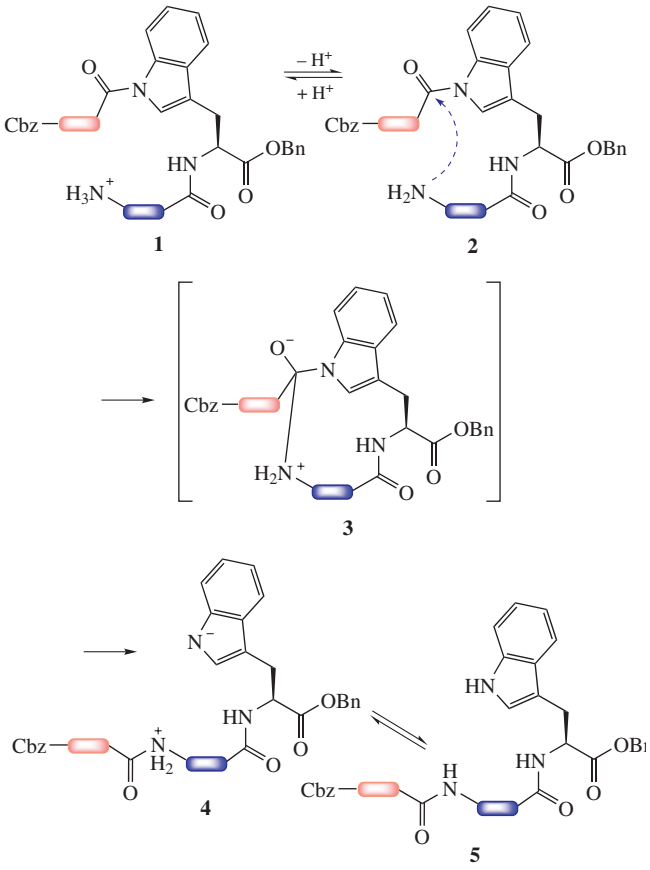
Native chemical ligation, while of great importance, is limited by (i) the requirement of a N terminal cysteine residue at the ligation site to afford a peptide containing an internal cysteine, and (ii) the low abundance of cysteine in human proteins (1.7% of the residues).[5–8] In attempts to overcome the limitation of low abundance cysteine, synthetic approaches have been developed to extend NCL to unprotected peptides at non-Cys sites. Dawson *et al*. enabled the development of ligation at Ala, combining Cys-based NCL followed by selective desulfurization to Ala on the ligated peptide construct.[9]

Recently, in order to overcome the limitation of post ligation modification we reported ligations of *N*-acyl tryptophan units[10] and *O*-acyl tyrosine[11] units *via* various cyclic transition states.

Model aminolysis reactions of oxo- and thioesters were studied computationally by Yang and Drueckhammer,[12] who highlighted the importance of including a general base (one molecule of water) into such calculations, since this significantly lowered the activation energy. Our group also studied computationally many intramolecular chemical ligation reactions including those of *S*-, *N*-, and *O*-acyl peptides.[13–15]

Herein, we computationally determine the factors governing intramolecular acyl transfers from tryptophan and tyrosine isopeptides to form the native peptide.

In our previous communications, we reported the intramolecular acyl migration of (*N*)- and (*O*)-isopeptides was carried out in the presence of base or buffer *via* spontaneous rearrangements through intramolecular N→N and O→N acyl migration to form a native amide bond at the ligation site. The N→N and O→N acyl transfer proceeds in a stepwise manner as shown in Schemes 1 and 2.[10,11] For this, reaction to be intramolecular, a cyclic transition state must be involved, within which the molecular structure is an important variable. If the molecule is able to attain a conformation at a small N–C geometrical distance with low steric hindrance and suitable hydrogen bonding, this preorganization assists the chemical ligation event. The tryptophan and tyrosine

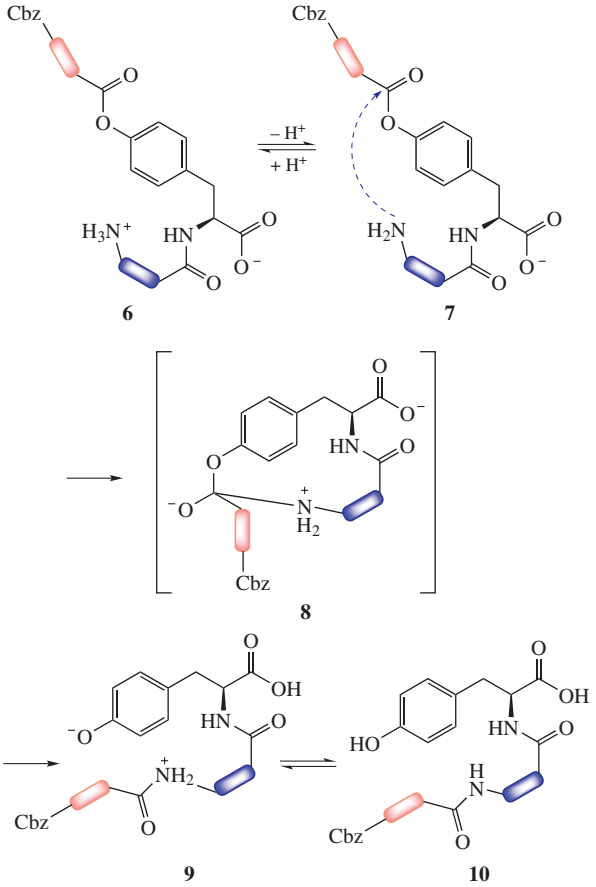


Scheme 1 Ligation pathway for intramolecular N→N acyl migration.

isopeptides **1a–d** and **6a–h**, respectively, were used for these calculations (see Figures 1 and 2).

Preorganization, or attaining an optimal conformation for binding or chemical reaction is a significant and sometimes crucial factor. Properly preorganized molecules form viable transition states at a smaller energy cost. In our case, preorganisation can be defined in terms of proximity of the nucleophile (amine nitrogen) to the electrophile (amidic or ester carbonyl), which is expressed as the geometrical distance *b*(N–C).

Full conformational searches for structures **1a–d** for tryptophan isopeptides and structure **6a–h** for tyrosine isopeptides were per-



Scheme 2 Ligation pathway for intramolecular O→N acyl migration.

formed using the MMX force field (as implemented in PCModel v. 9. software).[16] This resulted in several hundred conformations for each *S*-acyl peptide structure, which were subsequently ranked in descending order of the *b*(N–C) scoring function. The best preorganised conformers are shown in Figures 1 and 2, and the corresponding values of *b*(N–C) are given in Tables 1 and 2.

Recently Wang *et al.*[17] reported that the energy barrier of the thiol-thioester exchange step depends on side-chain steric hindrance of the C-terminal amino acid, whereas that of the acyl-transfer step

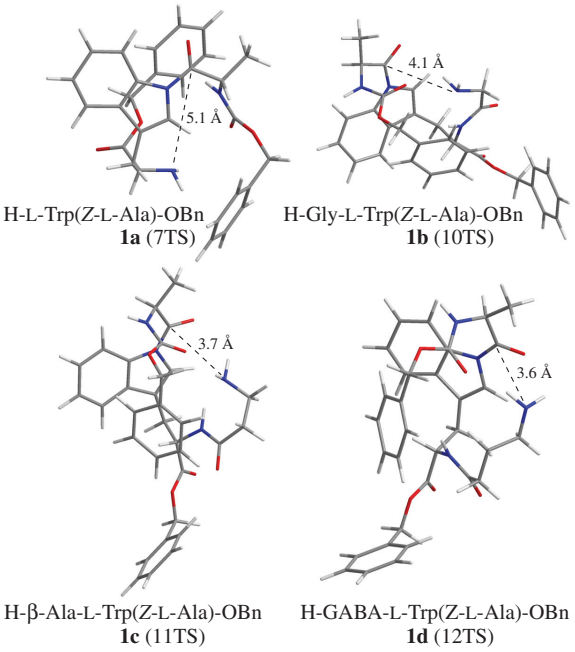


Figure 1 The best preorganized structure of tryptophan isopeptides for 7-, 10-, 11- and 12-membered cyclic transition states.

Table 1 Determination of governing parameters for tryptophan isopeptides **1a–d**.

Cyclic TS size	Ligated product (%)[a]	*b*(N–C)/Å	Hydrogen bond features		ΔE_{preorg} = $= E_{min} - E_{preorg}$/ kcal mol^{-1}
			X–H/Å	∠X–H···D/°	
7	2.0	5.061	N(3)–H(39)···O(37) 3.036	102.4	–1.475
10	44.4	4.087	N(25)–H(60)···O(29) 2.499	99.5	–2.435
11	71.4	3.657	N(24)–H(64)···O(30) 2.093	138.6	–4.616
12	99.1	3.634	N(28)–H(67)···O(31) 2.034	165.1	–4.357

[a] Relative abundance calculated from HPLC-MS with respect to bis-acylated product and starting material.

Table 2 Determination of governing parameters for tyrosine isopeptides **6a–h**.

Cyclic TS size	Ligated product (%)[a]	*b*(N–C)/Å	Hydrogen bond features		ΔE_{preorg} = $= E_{min} - E_{preorg}$/ kcal mol^{-1}
			X–H/Å	∠X–H···D/°	
12	95.3	4.851	N(22)–H(50)···N(1) 1.944	163.0	–0.021
13	85.2	4.765	N(23)–H(53)···N(1) 2.010	162.8	–0.535
14	97.0	3.779	N(22)–H(52)···N(34) 1.948	167.2	0.00
15	94.6	3.398	N(1)–H(37)···O(28) 1.885	174.8	–3.301
16	87.0	3.246	N(24)–H(56)···O(29) 1.913	171.3	–1.757
17	96.8	3.656	N(3)–H(41)···O(30) 2.088	152.5	–2.723
18	99.9	3.289	N(21)–H(55)···N(39) 1.908	163.7	–2.360
19	98.1	2.990	N(21)–H(56)···N(39) 1.970	149.1	–1.500

[a] Relative abundance calculated from HPLC-MS with respect to bis-acylated product and starting material.

depends on side-chain steric hindrance of the N-terminal amino acid. Since steric energy also influences the rate of ligation, we determine a part of the steric energy as a measure of preorganization, in addition to the geometrical scoring function *b*(N–C). This preorganization energy (ΔE_{preorg}) is calculated as the difference between the global minimum energy (E_{min}), determined through a full conformational search and the steric energy of the conformation best preorganized for ligation in terms of *b*(N–C), E_{preorg}.

$$\Delta E_{preorg} = E_{min} - E_{preorg}$$

As shown in Tables 1 and 2 for tryptophan and tyrosine isopeptides, respectively, the preorganization energy varies from zero to 4 kcal mol^{-1} depending on the cyclic TS. Generally, the smaller the molecule the smaller number of conformations exists and therefore the probability of finding a preorganized conformation close to the global minimum is higher, which is reflected by the approximately inverse relationship between *b*(N–C) and ΔE_{preorg}. Since the tyrosine derivatives are conformationally more flexible, they demonstrate smaller energy costs for preorganization comparing to the tryptophan isopeptides.

Hydrogen bonding also plays an important role in the stabilization of the transition state structure, as was computationally confirmed by us in a prior publication.[13] We showed that hydrogen bonding coupled with a proper conformation could rationalize the reactivity order in a series of isopeptides. The transition state of a chemical ligation reaction can be stabilized or destabilized by hydrogen bonding,[13] and this is what was found in the current study. This is part of the general phenomenon comprehensively

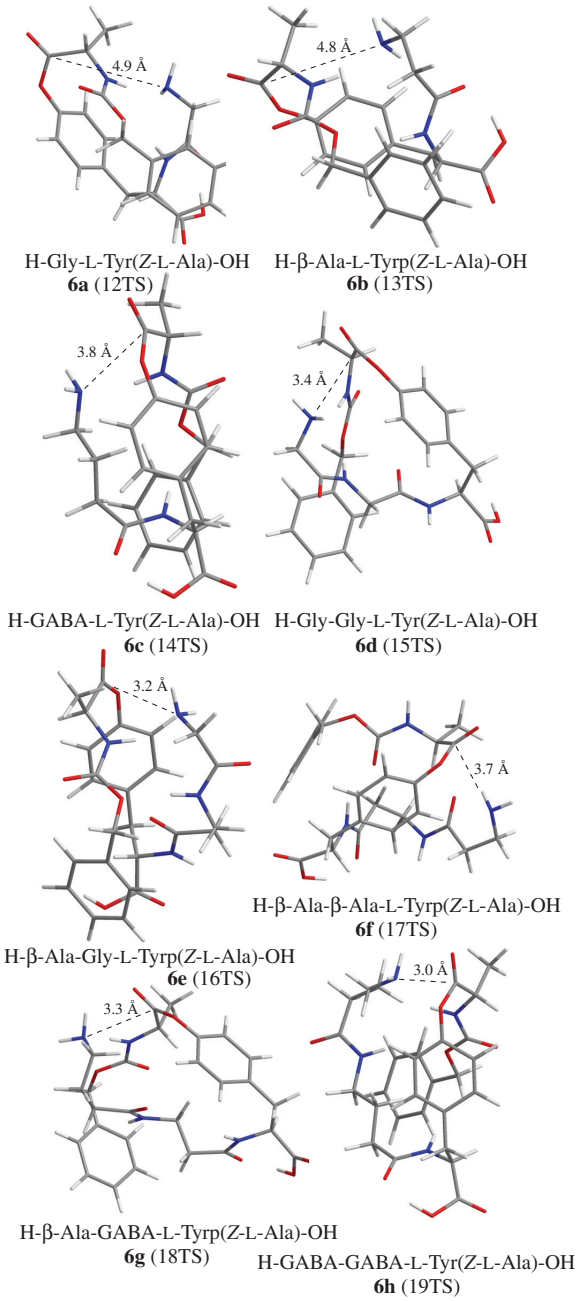


Figure 2 The best preorganized structure of tyrosine isopeptides for 12- to 19-membered cyclic transition states.

described by Toniolo.[18] A hydrogen bond can be best characterized by the H···D distance and X–H···D angle, where X is the heteroatom on the hydrogen bond acceptor side, H is the hydrogen atom, and D is the donor heteroatom. It is well-known that hydrogen bonds are directional and thus depend on the mutual orientation of the X–H dipole and the lone pair (or pairs) located on D. This can be accounted for in terms of the angle formed by these three atoms, X, H, and D. It is shown that rings of larger size are stabilized by true hydrogen bonds with short H···D distances (around 2 Å) and X–H···D angles close to 180°. The best preorganized conformers listed in Tables 1 and 2.

Intramolecular acyl transfers of tryptophan isopeptides through 10-, 11- and 12-membered transition states were favored over a 7-membered transition state and acyl migration occurred more readily in basic non-aqueous media relative to aqueous buffered conditions. As the N- to N-shift reaction occurs in the presence of piperidine, we believe general basic catalysis may be involved, because the proximity function *b*(N–C) and steric energy are high in preorganized structures of tryptophan isopeptides.

The scoring function values *b*(N–C) of tryptophan isopeptide for 7-, 10-, 11- and 12-membered cyclic transition states are 5.061, 4.087, 3.657 and 3.634 Å, respectively. The hydrogen bonding becomes stronger while decreasing the steric energy as the cyclic TS size increases. This explains the additional stabilization of the structure as well as the higher yield of the internal ligation product. The rate of ligation increases with (i) decrease in geometrical distance *b*(N–C), (ii) stronger in hydrogen bonding and (iii) lower energy of steric hindrance.

The intramolecular *O*- to *N*-acyl transfer *via* 12- to 14-membered transition states occurs in DMF–piperidine solution and 15- to 19-membered transition states in aqueous solution. The results are rationalised using a conformational search. The geometrical distance *b*(N–C) and steric energy of tyrosine isopeptides for 12- to 14-membered cyclic transition states are much higher than 15- to 19-membered cyclic transition states.

Under aqueous conditions (pH = 7.3) the 12-, 13-, and 14-membered ring transition states gave no ligation because preorganization in these three cases was poor both in terms of steric energy and preferred conformation. However in a DMF–piperidine mixture the attacking nucleophile is expected to be totally deprotonated and ligation takes place efficiently, with yield as high as 85–97%.

In summary, we report a theoretical rationalization of previous experimental results related to efficient and facile chemical ligations of (*N*)- and (*O*)-acyl di-, tri-, and tetrapeptides. It is shown that the cycle size, preorganization, and conformation strongly affect the reactivity. Supramolecular assistance to chemical ligation by formation of an intramolecular hydrogen bond is also supported by the computations.

We thank the University of Florida and the Kenan Foundation for financial support. This paper was also funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, under Grant no. (24-3-1432/HiCi). The authors, therefore, acknowledge with thanks DSR technical and financial support. We also thank Dr. C. D. Hall for helpful suggestions.

References

1 (*a*) B. Merrifield, *Science*, 1986, **232**, 341; (*b*) E. T. Kaiser, *Acc. Chem. Res.*, 1989, **22**, 47; (*c*) S. B. H. Kent, *Chem. Soc. Rev.*, 2009, **38**, 338.
2 (*a*) P. E. Dawson, T. W. Muir, I. Clark-Lewis and S. B. H. Kent, *Science*, 1994, **266**, 776; (*b*) A. El-Faham and F. Albericio, *Chem. Rev.*, 2011, **111**, 6557.
3 D. S. Y. Yeo, R. Srinivasan, G. Y. J. Chen and S. Q. Yao, *Chem. Eur. J.*, 2004, **10**, 4664.
4 T. Wieland, E. Bokelmann, L. Bauer, H. U. Lang, H. Lau and W. Schafer, *Justus Liebigs Ann. Chem.*, 1953, **583**, 129.
5 B. L. Nilsson, L. L. Kiessling and R. T. Raines, *Org. Lett.*, 2000, **2**, 1939.
6 B. L. Nilsson, L. L. Kiessling and R. T. Raines, *Org. Lett.*, 2001, **3**, 9.
7 (*a*) P. McCaldon and P. Argos, *Proteins: Struct. Funct. Genet.*, 1988, **4**, 99; (*b*) C. P. R. Hackenberger and D. Schwarzer, *Angew. Chem. Int. Ed.*, 2008, **47**, 10030.
8 R. Okamoto and Y. Kajihara, *Angew. Chem. Int. Ed.*, 2008, **47**, 5402.
9 L. Z. Yan and P. E. Dawson, *J. Am. Chem. Soc.*, 2001, **123**, 526.
10 W. Yang and D. G. Drueckhammer, *Org. Lett.*, 2000, **2**, 4133.
11 V. Popov, S. S. Panda and A. R. Katritzky, *Org. Biomol. Chem.*, 2013, **11**, 1594.
12 V. Popov, S. S. Panda and A. R. Katritzky, *J. Org. Chem.*, 2013, **78**, 7455.
13 A. A. Oliferenko and A. R. Katritzky, *Org. Biomol. Chem.*, 2011, **9**, 4756.
14 S. S. Panda, C. El-Nachef, K. Bajaj, A. O. Al-Youbi, A. Oliferenko and A. R. Katritzky, *Chem. Biol. Drug Des.*, 2012, **80**, 821.
15 J.-C. M. Monbaliu, G. Dive, C. V. Stevens and A. R. Katritzky, *J. Chem. Theory Comput.*, 2013, **9**, 927.
16 *PCModel v. 9.2*, Serena Software, www.serenasoft.com.
17 C. Wang, Q.-X. Guo and Y. Fu, *Chem. Asian J.*, 2011, **6**, 1241.
18 C. Toniolo, *CRC Crit. Rev. Biochem.*, 1980, **9**, 1.

Received: 9th January 2014; Com. 14/4283