Organic &
Biomolecular
Chemistry

Cite this: *Org. Biomol. Chem.*, 2011, **9**, 4756

www.rsc.org/obc

COMMUNICATION

Alternating chemical ligation reactivity of S-acyl peptides explained with theory and computations

Alexander A. Oliferenko and Alan R. Katritzky*

Received 5th April 2011, Accepted 26th April 2011
DOI: 10.1039/c1ob05536d

Previously discovered alternating reactivity of S-acyl di-, tri-, and tetrapeptide in internal chemical ligation reactions is rationalised using conformational search, virtual screening methods and quantum chemical calculations. Conformational preorganisation is shown to be the major controller of reactivity, with hydrogen bonding providing additional stabilisation for the tetrapeptide structure.

Peptide synthesis is one of the most important processes in biochemistry and life sciences. A multitude of chemical ligation techniques have been reported in the pioneering works of Kent and co-authors.[1] Some of them need no cysteine residue in the structure, nor any external chemical auxiliary. We recently described[2,3] the internal chemical ligation of S-(α-amino acyl) peptides (formed by a selective acylation of cysteine-containing peptides with N-acylbenzotriazoles) as an efficient, auxiliary-free way to native peptides. S-Acyl di-, tri-, tetra and pentapeptides were shown to undergo such S–N transfer *via* five-, eight-, eleven- and fourteen-membered cyclic transition states, respectively. However only those produced *via* five- eleven- and fourteen-membered transition states provided good yields of the desired native peptides, while the S-acyl tripeptide, needing an eight-membered transition to form the expected native tripeptide gave only a very poor yield, instead preferring to react by an intermolecular acylation.

Internal chemical ligations of this type can be considered mechanistically as intramolecular nucleophilic reactions between the thioester group and the unprotected N-terminus. Reactions of this type were studied in great detail by Connors and Bender,[4] who based on their own kinetic results and previous accounts, suggested a stepwise aminolysis mechanism assisted by general basic catalysis. The first step of this mechanism is the formation of a zwitter-ionic tetrahedral transition state, including an extraction of a proton from the attacking amino group by a general base, followed by decomposition of such transition state into amide and thiolate anion. More recently, model aminolysis reactions of oxo- and thioesters were studied computationally by Yang and Drueckhammer,[5] who highlighted the importance of including a general base (one molecule of water) into such calculations, since this significantly lowered the activation energy. In our ligation experiments general basic catalysis seems justified too, as the reaction mixture (phosphate buffer, pH 7.8) contains not only water, but much more basic anions such as HPO_4^{2-} and $H_2PO_4^-$. For this reaction to be intramolecular, a cyclic transition state must be formed, and here molecular structure is an important variable. If the molecular structure is able to attain a suitable conformation at a low energy cost, this can significantly assist the reaction (by reducing the activation entropy). The reaction can also be assisted by enthalpy gain due to assistance, such as hydrogen bonding, but can be significantly or even completely disfavoured if molecular structures are far from the optimum. The aim of the present work is thus to reveal structural features controlling the title chemical ligation and to explain the intriguing variation in the reactivity of S-acyl peptides.

Virtual screening

The S-acyl peptides under study each consist of two peptide parts: (i) a main chain C-terminal cysteine di-, tri-, tetra-, or pentapeptide featuring an unprotected N-terminus and a free COOH group, and (ii) an N-protected alanine residue linked through its CO_2H group to the cysteine sulfur atom, as shown below [eqn (1)–(4)]:

Cys(Pg-Ala)-OH (1)

Gly-Cys(Pg-Ala)-OH (2)

Gly-Leu-Cys(Pg-Ala)-OH (3)

Gly-Leu-Gly-Cys(Pg-Ala)-OH (4)

Here Pg means protecting group. In the previously reported chemical syntheses,[2,3] Fmoc was used for dipeptide **1** and benzyloxycarbonyl (Z) was used elsewhere. In our computations, all protecting groups, Fmoc and Z, are replaced by the carboxymethyl group without loss of generality, to render the calculations less time-consuming.

Preorganisation, or attaining an optimal conformation for binding or chemical reaction is a significant and sometimes crucial factor. Properly preorganised molecules can form viable transition states faster and at a smaller energy cost. In our case, preorganisation can be correctly defined in terms of proximity of the nucleophile (amine nitrogen) to the electrophile (thioester carbon). We consider here the geometrical distance *b*(N–C) as

Department of Chemistry, University of Florida, Gainesville, FL 32611-7200, USA. E-mail: katritzky@chem.ufl.edu; Fax: +1(352)392-9199; Tel: +1(352)392-0554

Table 1 Virtual screening and preorganisation results

structure	cycle size	b(N–C), Å	$E_{preorg.}$, kcal mol^{-1}
1	5	2.998	–34.68
2	8	3.591	–34.07
3	11	3.085	–28.61

a simple but relevant scoring function. Conformational searches for structures **1–3** were performed using the MMX force field (as implemented in PCModel v.9. software)[6] This resulted in a few hundreds of 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 Fig. 1, and the corresponding values of b(N–C) are given in Table 1. It is seen that structure **1**, featuring the ability to form a five-membered transition state, has the best (smallest) b(N–C) score. The next place is occupied by structure **3**, which forms the largest size cyclic transition state, whereas the intermediate structure **2**, forming an eight-membered transition state, is inferior to both **1** and **3**. This scoring is quite consistent with the observed reactivity patterns,[2,3] which suggests it can be used as a fast and inexpensive tool in computer-aided peptide design.

Apparently, structure **1** can most easily preorganise into a favorable conformation. But preorganisation always comes at a cost. To evaluate the relative thermodynamic stabilities of the selected conformers at a more accurate quantum chemical level, they were subjected to geometry optimisation using the HF/6-31+G* method (as implemented in US GAMESS/Firefly software).[7] Energies of the original preorganised conformers (E_1) and the geometry relaxed ones (E_2) are listed in Table 1. The quantum chemical preorganisation energy E_{preorg} is defined as follows: $E_{preorg} = E_2 - E_1$. The energy values in Table 1 are given in kilocalories per mole for quantum chemical conformational energies.

The E_{preorg} values are quite consistent with chemical knowledge: the shortest structure **1** has a relatively small number of conformational degrees of freedom and, accordingly, obtains the highest energy penalty for preorganisation. Structure **2** despite the longer peptide chain is just slightly better scored, with the difference between **1** and **2** only 0.6 kcal mol^{-1}. Structure **3** is noticeably less strained owing to the abundance of available conformational degrees of freedom.

Conformational analysis

Analysis of the preorganised conformers displayed in Fig. 1 can help understand how the molecular backbone takes a conformation which is favourable or unfavourable for ligation. For the purpose of conformational analysis, torsion angles φ, ψ, and ω, are defined in a similar way to that commonly used in biochemistry. All ω angles are very close to 180° and are not reported here, whereas the φ and ψ angles are listed in Table 2. As the classical propagation of peptide structure is violated in peptide thioesters, the torsion angles describing the cysteine residue are defined slightly differently: φ_{Cys}, as ∠ C_{i-1}–N–C^{α}–C^{β}, ψ_{Cys} as ∠ N–C^{α}–C^{β}–S, and φ'_{Cys} as ∠ C^{α}–C^{β}–S–C_{i+1}.

Conformational analysis of structure **1** is straightforward. The cysteine amino group is unobstructed and can easily be brought into a close proximity with the thioester moiety to afford a five-membered cyclic transition state. In the longest structure **3**, the peptide chain starts making an α-helix with Gly and Leu, but at Cys it makes an inverse γ-turn. The role of this turn and the C_7 structure formed is two-fold: (i) it brings the N-terminus in close proximity to the thioester function, a proximity that cannot be realised if the peptide chain continues the α-helix pattern; (ii) it makes a 3–1 hydrogen bond between the Cys NH and Ala

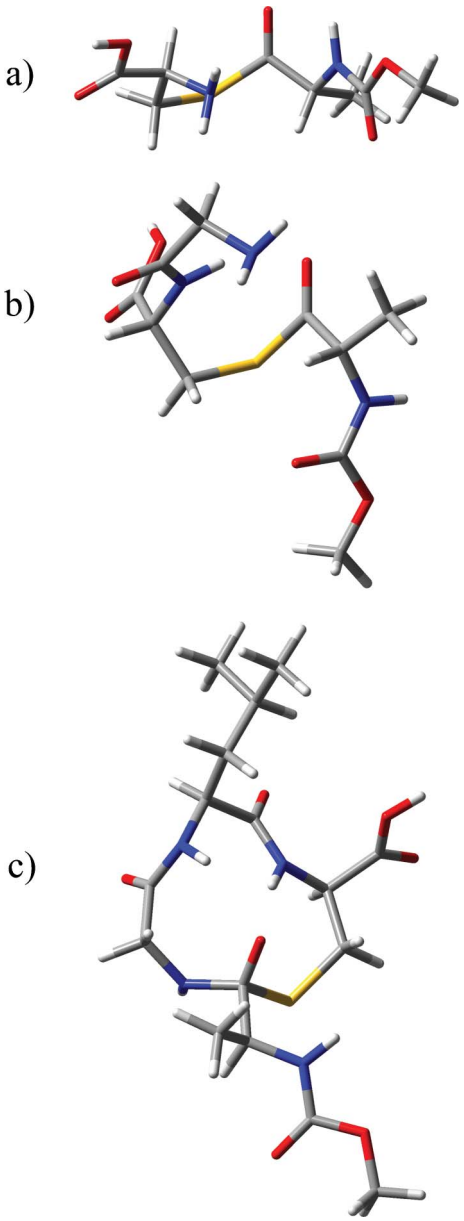

Fig. 1 Preorganised conformers of S-acyl peptides **1** (a), **2** (b), and **3** (c).

Table 2 Torsion angles of structures **1**, **2**, and **3**

structure	ψ_{Gly}	φ_{Leu}	ψ_{Leu}	φ_{Cys}	ψ_{Cys}	φ'_{Cys}
1					–67	9
2	–72			–115	–63	68
3	–124	–66	–40	–161	57	–101

Table 3 Reactivity modelling results

Structure	cycle size	b(N–C), Å	$E_{react.}$, kcal mol^{-1}
$\mathbf{1}^A$	5	2.998	−101.64
$\mathbf{2}^A$	8	3.591	−26.87
$\mathbf{3}^A$	11	3.085	−81.84

CO groups. This hydrogen bond, with the NH–O distance and the N–H–O angle equal to 1.93 Å and 150°, respectively, can provide additional stabilisation for the preorganised conformer **3**. The structure of intermediate length, structure **2**, also starts as an α-helix, but it is too short to bring the amino group close enough to the electrophilic centre. If it were an alkyl chain, a cyclic conformer could be easily formed, but a peptide chain is far less flexible. Several sets of the φ and ψ angles were tried in attempts to construct a viable eight-membered cyclic transition state, but all the structures possessed high strain energy and were seriously altered in the course of geometry optimisation with the MMX force field, and none had available favourable hydrogen bonding. These conformational features prevent the N-terminus coming close to the COS group and thus cause structure **2** to score poorly in terms of the b(N–C) function.

Quantum chemical study of reactivity

As mentioned above, the S–N transfer mechanism may be thought of as a nucleophilic substitution under general basic catalysis conditions. This mechanism implies that the free amino group is somehow deprotonated. Wang and Drueckhammer[5] assumed that a water molecule can assist the deprotonation. To simplify quantum chemical calculations of our much larger molecules compared to those used in,[5] the preorganised structures **1**, **2**, and **3** were deprotonated just by removing the proximal hydrogen atoms from the zwitterionic N-termini. To probe the reactivity, the preorganised S-acyl peptide structures modified in this way were geometry optimised at the HF/6-31+G* level of theory.

The reactivity calculation results are given in Table 3. Energies of the modified structures all have the superscript A (anionic structure). The quantum chemical reaction energy E_{react} is defined as $E_{react} = E_2{}^A - E_1{}^A$, where $E_1{}^A$ and $E_2{}^A$ are the energies of the starting and final, geometry-optimised structures, respectively. The energy values in Table 2 are given in kilocalories per mole for quantum chemical reaction energies only.

Optimised geometries of the modified structures $\mathbf{1}^A$, $\mathbf{2}^A$, and $\mathbf{3}^A$ are shown in Fig. 2. The dramatic changes that happen with structures **1** and **3** in the course of geometry optimisation are evident. Structure **1** undergoes the classical S–N transfer reaction and forms a native dipeptide, as the pendant thiolate group of Cys is clearly seen in Fig. 2a. Nucleophilic attack also occurs in $\mathbf{3}^A$, as the eleven-membered ring is formed. Although the C–S contact is visualised as a bond in Fig. 2c, it is barely a covalent bond, because the C–S separation is 2.093 Å, far beyond the range of normal C–S bonds. The ring closure also indicates the formation of the N–C amide bond with a bond length of 1.475 Å. This means that structure $\mathbf{3}^A$ undergoes an S–N transfer to afford a native tetrapeptide.

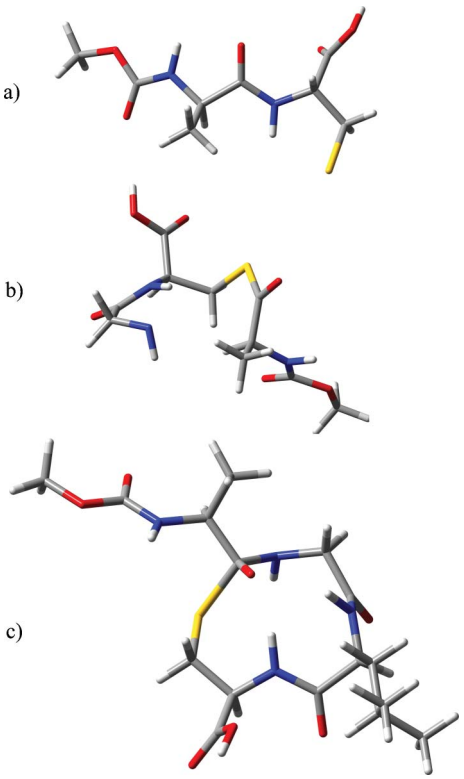

Fig. 2 S–N transfer results (a) for S-acyl peptides $\mathbf{1}^A$, (b) for $\mathbf{2}^A$, and (c) for $\mathbf{3}^A$; structure $\mathbf{2}^A$ does not afford S–N transfer.

By contrast, structure $\mathbf{2}^A$ depicted in Fig. 2b does not show any signs of a reaction leading to a native tripeptide. Instead of the expected nucleophilic attack, the structure experiences a geometry relaxation, which drives the nucleophilic –NH$^-$ moiety away from the target thioester group. The distance between those two is now 3.981 Å, *cf.* 3.591 Å before the calculation.

The quantum chemical energy data in Table 3 match very well the virtual screening results. Structure $\mathbf{1}^A$, featuring the best scoring results, is characterised with the highest reaction energy of more than −100 kcal mol^{-1}. Structure $\mathbf{3}^A$ scores next and although its E_{react} is consistently smaller (−81.84 kcal mol^{-1}), it is high enough to consider the structural change occurring with $\mathbf{3}^A$ as a chemical reaction rather than a conformational transition. The E_{react} value for structure $\mathbf{2}^A$ is −26.87 kcal mol^{-1}, which is 2–3 times smaller than for $\mathbf{1}^A$ and $\mathbf{3}^A$.

This energy range may be indicative of geometry relaxation but not a chemical reaction. Based on these computational results one can rationalise why the intramolecular chemical ligation of the S-acyl tripeptide **2** is disfavored compared to intermolecular *trans*-acylation.

The hypothesis of hydrogen bond assistance to reactivity also turns out to be valid. The amide proton of Cys and the carbonyl oxygen of Ala engage in hydrogen bonding in the preorganised structure **3**, and this hydrogen bond becomes even stronger (1.93 Å and 150° *vs.* 1.85 Å and 164°) in the product $\mathbf{3}^A$. This hydrogen bond plus a proper conformation can nicely rationalize why S-acyl peptide **3** is able to form a transition state suitable for a successful chemical ligation. Although it is known that equilibrium and rate constants can be influenced by hydrogen bonding,[8] hydrogen bonds facilitating intramolecular coupling are less well known.

One notable example of using this concept is template synthesis of macrocycles,[9,10] but until now no examples have been repeated for peptide syntheses.

Conclusions

In summary, we report a theoretical rationalisation of our previous experimental results related to efficient and facile chemical ligations of S-acyl di-, tri-, and tetrapeptides. It is shown how the length and conformation of peptide thioesters strongly affect their reactivity. Supramolecular assistance to chemical ligation by the formation of an intramolecular hydrogen bond is conjectured and supported by computations.

Notes and references

1 1 S. B. H. Kent, *Chem. Soc. Rev.*, 2009, **38**, 338–351.
2 A. R. Katritzky, N. E. Abo-Dya, S. R. Tala and Z. K. Abdel-Samii, *Org. Biomol. Chem.*, 2010, **8**, 2316–2319.
3 A. R. Katritzky, S. R. Tala, N. E. Abo-Dya, T. S. Ibrahim, S. A. El-Feky, K. Gyanda and K. M. Pandya, *J. Org. Chem.*, 2011, **76**, 85–96.
4 K. A. Connors and M. L. Bender, *J. Org. Chem.*, 1961, **26**, 2498–2504.
5 W. Yang and D. G. Drueckhammer, *Org. Lett.*, 2000, **2**, 4133–4136.
6 PCModel v. 9.2: Serena Software, www.serenasoft.com.
7 US GAMESS/Firefly: A. Granovsky, www.classic.chem.msu.ru.
8 F. Hibbert and J. Emsley in *Adv. Phys. Org. Chem.*, Ed. D. Bethell: 1991, **26**, 255.
9 P. L. Anelli, F. Montanari and S. Quici, *J. Org. Chem.*, 1985, **50**, 19.
10 R. Vilar, Hydrogen bonded template assemblies, *Struct. Bond.*, 2004, **111**, 85–137.