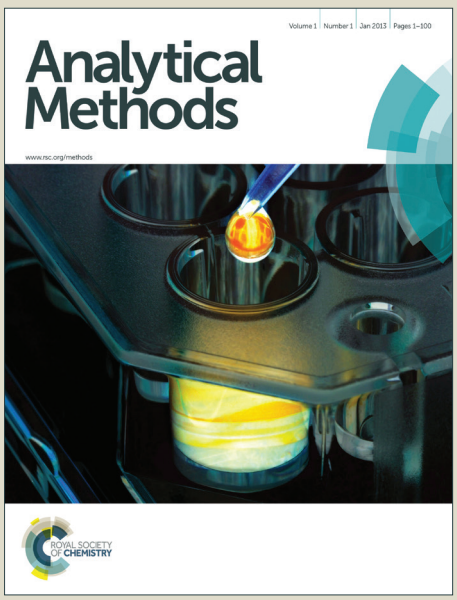
Analytical Methods

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/methods

Investigating the synthetic mechanism of 3, 5-diamino-1, 2, 4-triazole by using fibre optic ATR-IR spectroscopy combined with kernel independent component analysis

Hua Li,[*,1] Jun Gao,[1] Hongsheng Tang, [1] Jiguang Li,[2] Tianlong Zhang[1] and Xiaofeng Yang[1]

[1]Institute of Analytical Science, College of Chemistry & Materials Science, Northwest University, Xi'an, 710069, China

[2]School of Chemistry & Chemical Engineering, Ningxia University, Yinchuan 750021, China

Author to whom correspondence should be addressed:

Dr. Hua Li

Institute of Analytical Science, College of Chemistry & Materials Science, Northwest University

Tel: 86-29-88302635

Fax: 86-29-88303527

E-mail: nwufxkx2012@126.com

Abstract

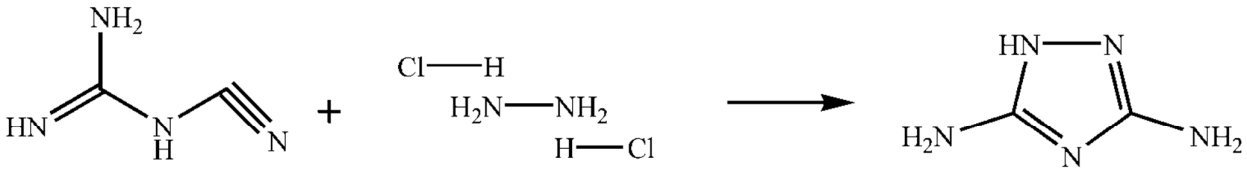
In-line attenuated total reflection infrared (ATR-IR) combined with kernel independent component analysis (KICA) was applied to investigate the synthetic mechanism of 3, 5-diamino-1, 2, 4-triazole (DAT). In the proposed method, the reaction process was firstly monitored by an in-line IR probe to acquire IR spectra at a precise period of time, and then KICA was used to directly estimate the spectrum of each component in the mixture without pre-separation and other prior information. The estimated IR spectra were used to monitor the reaction and deduce the reaction mechanism. The intermediates involved in the reaction process were further validated by the density functional theory (DFT) at B3LYP level. The computational results obtained by the KICA approach are in good agreement with those by the quantum chemical calculation method, thereby demonstrating the reliability of the KICA method. The results show that in-line ATR-IR spectroscopy combined with KICA can be applied to study the synthetic mechanism of 3, 5-diamino-1, 2, 4-triazole effectively.

Keywords: Kernel independent component analysis; In-line FT-IR; Density functional theory; Synthetic mechanism

1. Introduction

3, 5-Diamino-1, 2, 4-triazole (DAT) is an important fine chemical intermediate that has been widely applied in biological and pharmaceutical fields[1], and it has been recommended as kinase inhibitors, cathode catalyst, ligands to transition metals[2-6]. DAT is also an important organic synthetic intermediate for several high energy density materials (HEDMs), such as 5-amino-3-nitro-1, 2, 4-triazole, 4, 6-bis (5-amino-3-nitro-1, 2, 4-triazol-1-yl)-5-nitropyrimidine and 3, 6-bis (4-amino-3-nitro-1, 2, 4-triazol-1-yl)-1, 2, 4, 5-tetrazine[7, 8.] In order to overcome the difficulty of obtaining raw materials, reduce the synthetic cost and improve the yield of DAT, it is urgent to investigate the synthetic mechanism of DAT which has a brightly prospect in industrialized utilization.

The synthetic route of DAT is shown in Scheme 1.



cyanoguanidine hydrazine dihydrochloride DAT

Scheme 1 Synthetic route of DAT.

Interest in real-time signals, either improving yields or as controlling reaction steps, drive plenty of chemists make extensive efforts to the in-line reaction monitoring. They employ several equipments to study in-line reaction monitoring[9-11]. Among these, the in-line attenuated total reflection (ATR) IR probe, which can obtain qualitative and quantitative information simultaneously without additional reagents, has been proved suitable for both kinetic and mechanistic studies. ATR is a sampling technique used in conjunction with infrared spectroscopy which can directly examine samples in the solid or liquid state without further preparation. The sample is in contact with the ATR crystal. A beam of infrared light undergoes multiple internal reflections in the crystal. Then the beam of light is collected by a detector. Generally, an in-line ATR-IR probe has two main advantages: one is in-line monitoring which means the detection takes place inside the reaction system directly, that can minimize the risk of contamination and loss during probing time; the other is that it is suitable for investigating unstable intermediates[12, 13]. Moreover, this approach can greatly reduce the consumption of chemicals, least waste generation and fast and accurate following conversion of the reaction in-line to the conventional ones. In addition, chemometrics can process the large-scale data matrix, which makes the analytical procedure convenient and speedy.

Independent component analysis (ICA) is a statistical method that has the potential ability for blind source separation (BSS) and feature extraction by estimating independent components (ICs) from the

measured overlapping signals of mixtures without the prior information about the mixtures[14, 15]. As a novel signal processing method, the concept of ICA was first proposed by Herault and Jutten in 1986. The main idea of ICA is to find a mathematical transformation of the data into a linear combination of statistically independent components. Since it was developed in the 1990s, ICA has attracted great interests of researchers and has been applied in chemical studies, e.g., monitoring dynamic and batch processes[16], processing ultraviolet-visible (UV-Vis) and near-infrared (NIR) spectral data of mixtures[17, 18], extraction of pure mass spectra from overlapping gas chromatographic-mass spectrometric (GC-MS) data[19], identifying constituents in commercial gasoline combining with IR[20], Raman[21], Nuclear magnetic resonance (NMR) and mass spectrometry (MS)[22, 23].

The conventional ICA algorithms adopts a single nonlinear contrast function to estimate source signals: the FastICA algorithm computes sequentially components using a deflation scheme[24]; the JADE algorithm uses joint diagonalization of a set of fourth-order cumulant matrices[25]; the extended Infomax algorithm is a variation of the Infomax algorithm[26, 27], which can analyze either super-Gaussian or sub-Gaussian signal. The mean-field ICA (MF-ICA) algorithm finds the mean of the sources and their covariance matrix, and uses them to describe the sources, mixing matrix and the noise covariance matrix[28]. All these ICA algorithms are reported to be successfully applied in certain application areas. Kernel ICA (KICA) is one of ICA method that based on kernel-based measures of dependence. It estimates the source signal by using contrast functions based on canonical correlation. Wang[29, 30] has reported that KICA is more robust than conventional non-kernel ICA algorithms when use them to isolate source spectral features from multiclass spectral data of mixtures.

In the present study, KICA was used to estimate original IR spectra profiles from IR dataset of DAT that produced by the reaction between cyanoguanidine and hydrazine dihydrochloride. The estimated

individual spectrum of each component in the mixture can be applied to deduce the synthetic mechanism of DAT. The intermediates involved in the reaction process were further validated by the density functional theory (DFT) at B3LYP level. The above experimental results show that in-line ATR-IR spectroscopy combined with KICA could be applied to study the synthetic mechanism of DAT successfully. To the best of our knowledge, fewer examples of such applications in studying synthesis mechanism of HEDMs have been reported up to date.

2. Theory and algorithm

2.1 Kernel independent component analysis

The FT-IR spectra data $\mathbf{X}$ is a bilinear matrix. It can be expressed simply as Eq. (1):

$$\mathbf{X} = \mathbf{C}\mathbf{S}^{\mathbf{T}} + \mathbf{E} \quad (1)$$

Where $\mathbf{X}(m \times n)$ is an observed data matrix (absorbance matrix); m denotes the number of data channels associated with the experimental FT-IR wavenumber range and interval taken; n is the number of scan; $\mathbf{C}(m \times d)$ is a mixing matrix; $\mathbf{S}^{\mathbf{T}}(d \times n)$ is a denoting source matrices (pure components spectra matrix); $\mathbf{E}(m \times n)$ is the error matrix (often omitted); d is the number of components included in the bilinear decomposition of Eq. (1).

ICA aims to find a maximum likelihood matrix of $\mathbf{C}$, when $\mathbf{W}$ is the pseudoinverse of the mixing matrix $\mathbf{C}$, i.e., $\mathbf{W} = \mathbf{C}^{-1}$, so that $\hat{\mathbf{S}} = \mathbf{W}\mathbf{X} = \mathbf{W}\mathbf{C}\mathbf{S} = \mathbf{S}^{\mathbf{T}}$. Corresponding to a kernel K, there is a map, $\emptyset$, from input space X to feature space H such that:

$$K(x_i, x_j) = < \emptyset(x_i) \cdot \emptyset(x_j) > (i, j = 1, 2, \cdots, n) \quad (2)$$

It means that the kernel can be used to evaluate an inner product in the feature space, which often referred to as the "kernel trick"[31]. In KICA space, it is known that canonical correlation analysis (CCA) is a classical linear model for finding dependencies[32]. It is truly that the least dependence of the variables is equivalent to their maximum independence. Therefore, KICA can be performed using kernel-CCA (KCCA) procedure, and the inefficiency of the conventional ICA algorithms led by using

single contrast function can be avoided by using the alternative kernel functions with optimal parameters setting.

For estimation of source spectra profiles from the spectral data, the below steps are adopted:

(1) Input the raw spectral data. The raw data were stored in an absorbance matrix x (n spectra with m components).

(2) Set the initial kernel function of KICA. The Gaussian kernel, $K(x_i, x_j) = exp[-(x_i - x_j)^2/2\sigma^2]$, was adopted with its width initialized in this work.

(3) Raw spectral data were whiten, $\hat{x} = Px$.

(4) Let, $s_i = Wx_i$ for each i, we obtained a set of estimated source vector $\{s_1, s_2, \cdots, s_n\}$, then calculated the center Gram matrix $K_1, K_2, \cdots, K_m$, which depends on these vectors.

(5) Calculate the smallest eigenvalue in (3) and according to the formula (4), minimizing the contrast function $C(W)$ in the direction of W, thus we obtained W.

$$\begin{pmatrix} K_1^2 & K_1K_2 & \cdots & K_1K_m \\ K_2K_1 & K_2^2 & \cdots & K_2K_m \\ \vdots & \vdots & \vdots & \vdots \\ K_mK_1 & K_mK_1 & \cdots & K_m^2 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix} = \lambda \begin{pmatrix} K_1^2 & 0 & \cdots & 0 \\ 0 & K_2^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & K_m^2 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix} \quad (3)$$

$$C(W) = I_\lambda(K_1, \cdots, K_m) = -1/2\, log\, \lambda_k\, (K_1, \cdots, K_m) \quad (4)$$

(6) Output the W, estimated source signal, $\hat{S} = W\hat{x}$.

2.2 Subspace comparison method

The subspace comparison method (SCM) has been proposed by Liang[33]. For a mixed data matrix, the principle component vectors are determined by different methods. The main subspace of a same system is consistent. The selected principle component vectors are derived from the same analysis system, and every other alternate vector can be mutually representing linearly. However, if it contains excess noise factor, the subspace calculations by different algorithms will not represent linearly. According to this feature, we can design a appropriate space function to get chemical ranks of this system easily.

It needs two different kinds of algorithms to calculate base vectors. Herein, we adopt simple to-use interactive self-modeling mixture analysis (SIMPLISMA) and singular value decomposition (SVD) to get two groups of principle component base vectors. Then one group is represented as $F = [f_1, f_2, \cdots, f_A]$, the other is represented as $G = [g_1, g_2, \cdots, g_A]$ via the gram-schmidt process. The i, j vector in F, G is represented as f_i, g_j, respectively. Subspace dimension is A. And we can use $\{g_j\}$ as base to present f_i: $f_i = \sum_{j=1}^{A} g_j \left(g_j^T f_i\right) + e_i$. The vector e_i is present that missing part of f_i orthogonalize $\{g_j\}$. After square modulus operations for f_i are performed in the above equation, we calculated using the following equations.

$$1 = (f_i^T f_i) = \sum_{j,k}^{A} \left(f_i^T g_j\right) \left(g_j^T g_k\right) \left(g_k^T f_i\right) + (e_i^T e_i) = \sum_{j=1}^{A} \left(f_i^T g_j\right)^2 + (e_i^T e_i) \quad (5)$$

$$T(A) = \sum_{i,j}^{A} \left(f_i^T g_j\right)^2 = Tr(F^T G G^T F), (i = 1,\ 2, \cdots, A;\ j = 1, 2,\ \cdots,\ A) \quad (6)$$

$Tr(\cdot)$ represents the trace of matrix, scilicet the sum of the diagonal elements of the matrix $(F^T G G^T F)$. Therefore, we can define the subspace difference degree as: $D(A) = A - T(A)$. $D(A)$ denotes the function of subspace difference.

3. Experimental section

3.1 Reagents and instrument

Cyanoguanidine (99.5% purity, Acros Organics) and hydrazine dihydrochloride (99% purity, Shanghai Meryer Chemical Technology Co., Ltd.) were analytical-reagent grade and were used as supplied. IR spectra were recorded on a Vertex 70 spectrometer (Bruker Optics, Ettlingen, Germany), equipped with an ATR probe (IN350-T, Germany) made of diamond combined with bundles of mid-IR optical fibers and a liquid nitrogen cooled mercury-cadmium-telluride (MCT) detector (Bruker Optics, Ettlingen, Germany) in a range of 4000-600 cm^{-1}. A 1.5 m silver halide fiber with an ATR probe was inserted in the reaction media from the top of three-necked flask. All spectra were recorded with a 4 cm^{-1} spectral resolution. The scanning frequency of background and the sample were 32 and 16 scans per seconds, respectively.

3.2 Synthesis of DAT

A 50 mL three-necked flask to which an in-line ATR-IR probe was inserted was charged with 30 mL of deionized water. The background was then scanned at 20.0 ℃. After that, cyanoguanidine (2.50 g, 29.73mmol) was dissolved in the above water, and then start to monitor, hydrazine dihydrochloride (3.75 g, 35.75 mmol) was slowly added to the above solution. The above mixture was kept at 48-50 ℃ for 90 min, and then the pH of the solution was adjusted to 10.5 with NaOH (3 mol/L). Finally, evaporation of the solvent under reduced pressure gave the crude product, which was further purified by recrystalization from methanol to afford the desired product (2.59 g, 26.14 mmol, 88% yield).

3.3 Dataset

The spectra of DAT synthesis were observed on VERTEX 70 connected to a in-line ATR-IR fiber probe. When the reaction process finished, a three-dimensional diagram was shown in Fig. 1 and a data matrix **X** (1763×1299) which consisted of 1763 wavelength points and 1299 scan number was output by the OPUS 6.5 software package (Bruker Optics, Ettlingen, Germany).

3.4 Software

All computations were performed with subroutines developed under Matlab (ver.7.0) on a PC. The subroutine of ICA was downloaded from the website[34]. The PCA method was obtained from the PLS-Toolbox-5.5 (Eigenvector Research, Inc., USA).

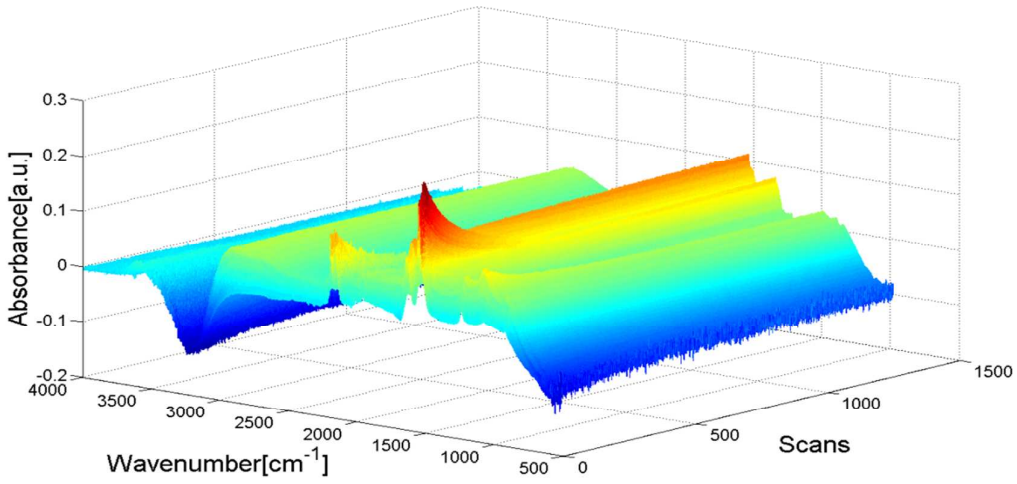


Fig. 1 In-line IR spectra (4000-600 cm^{-1}) were recorded during the synthesis of DAT

starting from cyanoguanidine and hydrazine dihydrochloride.

4. Results and discussion

4.1 Determine the component number

4.1.1 Estimation of the ICs using PCA method

Initially, the principle component analysis (PCA) was employed to estimate the principle componets (PCs). The statistically mutual independent of ICs is a much stricter property than the irrelevance of PCs determination in PCA algorithm. However, in this study, PCs and ICs were consistent because the mutual independence property of each substance and the respective spectrum.

The calculation is realized by arraying the covariance matrix eigenvalues in descending order and calculating the eigenvalues logarithm, and the result is shown in Table 1. The eigenvalue associated to signals should be significantly larger than the eigenvalue associated to noise. In practical situations, the decrease of the eigenvalue is generally smooth, leading to ambiguous rank estimation.

Table 1 Results of principal component analysis

PC	Eigenvalue of covariance matrix	Variance of the PC(%)	Cumulative variances (%)
1	3.45	98.42	98.42
2	4.33×10^{-2}	1.23	99.66
3	5.87×10^{-3}	0.17	99.82
4	1.73×10^{-3}	0.05	99.87
5	3.71×10^{-4}	0.01	99.88
6	1.37×10^{-5}	0.00	99.89

From Table 1, it's difficult to tell the exact number of the ICs in the system. For the first component's variance is captured to 98.42%, beyond 95%. In fact, the composition of this system cannot be only one. And it is clearly that the differences between the forth and the fifth component's variance was not significant. Therefore, SCM method was adopted to determine the chemical ranks in multi-variable statistical analysis.

4.1.2 Estimation of the ICs using SCM method

SCM is chosen as a proper solution for chemical rank determination. Two algorithms (SIMPLISMA and SVD) are employed to calculate key vectors from the obtained spectra matrix as their respective

base vectors. Each group of vectors is linearly correlated. Once a wrong rank number is chosen, the degree of subspace difference would significantly increase due to the broken of linear connection. Therefore, the right component number appears with a minimum subspace difference degree. The calculating results were shown in Fig. 2.

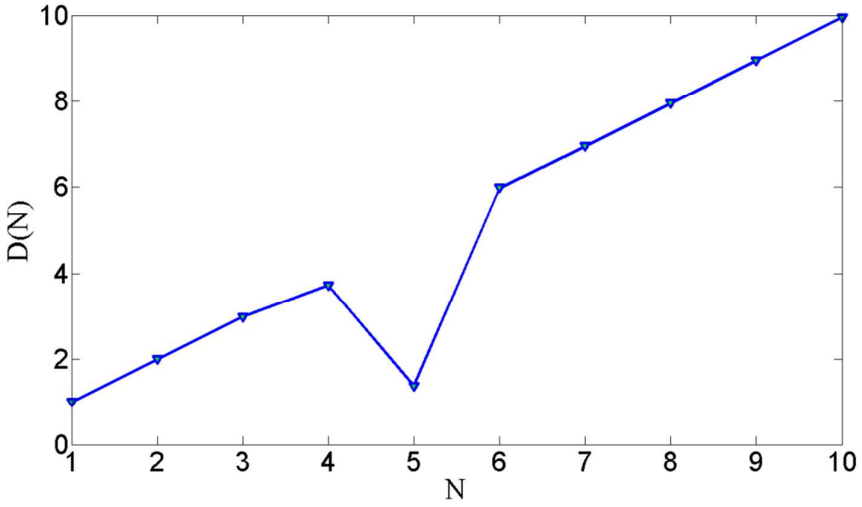


Fig. 2 Number of independent components was determined by SCM method.

From Fig. 2, the number of ICs is 5 by this method in the system. The experiments in Liang's research and our results all reveal that the proposed method is reliable.

4.2 Estimation of IR Spectra Profiles by KICA Algorithm

The background of raw experimental IR spectra data matrix was deducted firstly. Then, KICA algorithm was adopted to process the de-noised data matrix. To begin to calculate, the new estimates of source IR spectra matrix $\mathbf{S}^{\mathbf{T}}$ would be obtained after each iterative calculation. The iterative calculations were repeated until convergence to obtain the real meaningful IR spectrum of each component (Fig. 3). The real spectra of cyanoguannidine, hydrazine dihydrochloride and DAT were measured by the ATR Probe with the solvent of water. The spectra over 2800 cm^{-1} were cut out because the background correction generated negative peaks around 3000 cm^{-1}.

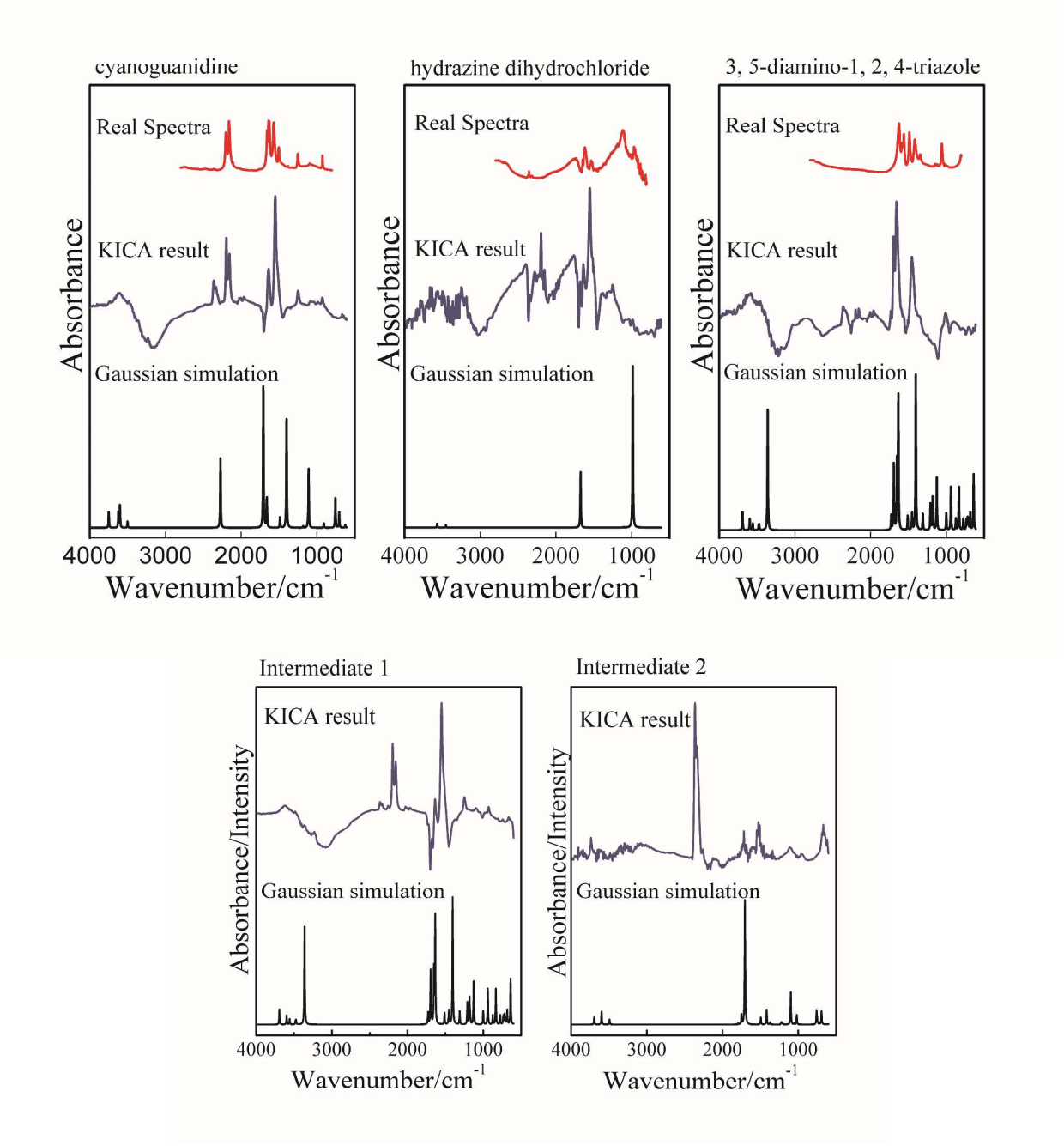


Fig. 3 IR spectra of reactants (cyanoguanidine and hydrazine dihydrochloride), intermediates (1 and 2), and the product (DAT) obtained from KICA algorithm, ATR probe and Gaussian simulation respectively.

It can be seen from Fig. 3 that cyanoguanidine (~2156, 2198 cm^{-1}), hydrazine dihydrochloride (~1454 cm^{-1}), and DAT (1600~1450 cm^{-1}) have unique band relatively free of spectral interference from other species. And the shape of some characteristic bands was reversal, especially intermediate **1** (~1652, 1704 cm^{-1}). In ICA calculation, there are the threefold indeterminacy of permutation, sign and shape[35]. The troubles caused by these indeterminacies include the rearrangement of all the output

components after each calculation, dealing with difficulty. Intermediate **1** and **2** are the hypothetic compounds according to the possible synthetic mechanism between cyanoguanidine and hydrazine dihydrochloride. Therefore, whether the hypothetic compounds exist or not, it needs further to validate.

4.3 Validation of Intermediates by Density Functional Theory

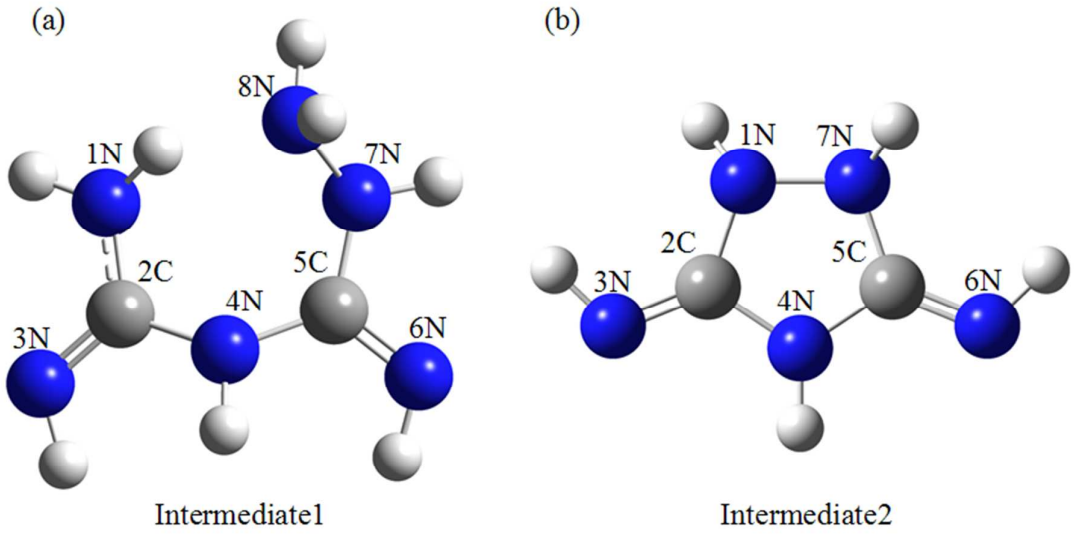


Fig. 4 Geometric (a) and IR spectra (b) of intermediate **1** and **2** fully optimized at the B3LYP/6-31G++ level.

Density functional theory (DFT) is now the electronic structure method of choice for fast but also accurate calculations of the properties of even large molecules; it is reliable and widely used for the computation of organic compounds and intermediates. Therefore it is suitable for the elucidation of reaction paths and mechanisms[36-38]. The commonly used B3LYP functional was applied within this study, as it was reported that the obtained geometries and properties are in good accordance with the experimental results[39]. In our study, in order to verify the correctness of each intermediate, the geometric configuration of each intermediate was fully optimized and the vibrational frequencies computed at the B3LYP/6-31G++ level. Polarized continuum model (PCM) is applied in consideration of solvent effects. All the calculations were carried out using the Gaussian 03 program package[40].

The optimized geometric configurations of intermediates are shown in Fig. 4. The vibrational analysis of optimized geometric configurations shows no imaginary frequency (which means all

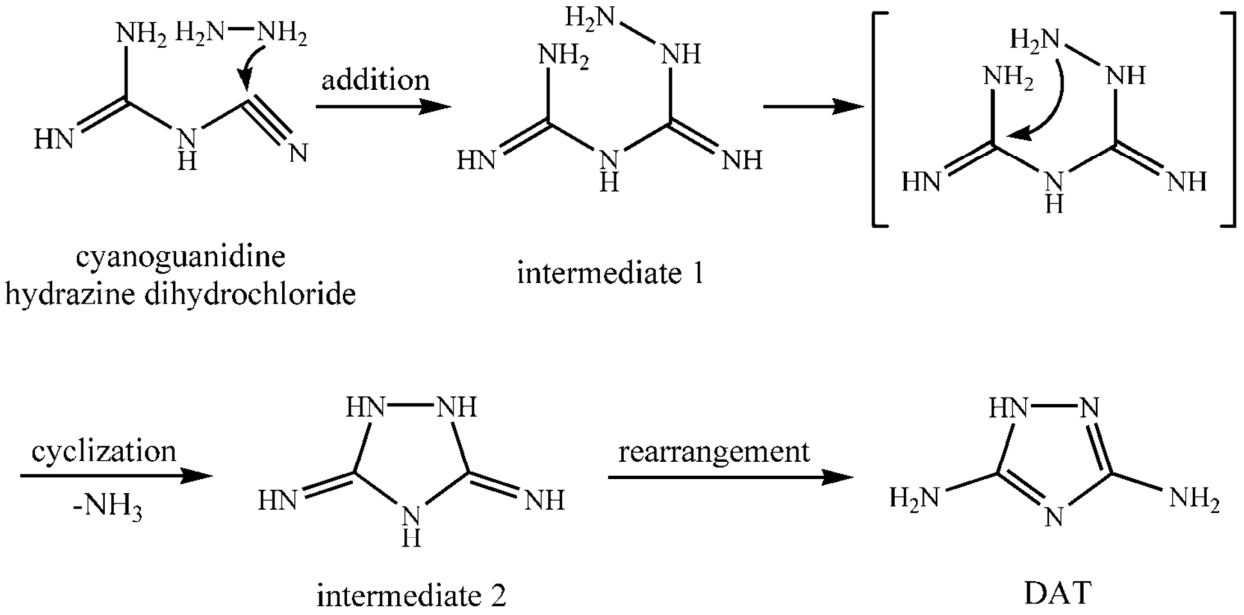
frequencies have a positive sign), which ensures that this is a stable point on the potential energy surfaces (which is a local minimum). The result demonstrated that the optimized geometries were relatively stable.

The second layer for each component in Fig. 3 depicts the Gaussian simulation results between the vibration frequency and the intensity. Intermediate **1** mainly has the following strong absorption peaks. The peaks at 1199 and 1120 cm^{-1} belong to the stretching vibration of C-N groups. The peaks at 1409 and 1479 cm^{-1} are bending vibration of N-H group. The peaks at 1732 and 1672 cm^{-1} are both the stretching vibration of C=N groups. Intermediate **2** exhibits several strong absorption peaks. The peaks at 1053 and 1088 cm^{-1} attribute to the stretching vibration of C-N groups. The peak at 1768 cm^{-1} corresponds to the stretching vibration of C=N group. The peak at 1448 cm^{-1} is the bending vibration of N-H group.

The fitting out frequencies by Gaussian calculation is different from infrared spectra, but the position of each peak is broadly similar. In this study, IR spectrum of each substance is calculated by KICA from the mixed reaction system, and therefore the deviation occurs when compared with the spectrum of the pure substance. In addition, quantum mechanical computations will slightly bias by inadequate simulation. However, each of calculated characteristic peaks are within the scope of the respective ranges. The previous studies indicated that the results of vibrational frequencies and IR spectra analyzed by DFT/B3LYP are reliable[39, 42]. The calculated peak positions by B3LYP/6-31G++ are agreed with those of the estimated spectra obtained by KICA (only IR characteristic peaks were compared due to the complicated experimental conditions). The estimated results prove the reliability of the KICA method used in our study.

4.4 Synthetic mechanism of DAT

The possible synthetic mechanism (Scheme 2) of DAT was deduced by combining the changes of functional groups in IR spectra with the further qualitative theoretical analyses of intermediates. According to the IR spectra of reactants, intermediates and product (shown in Fig. 3), the synthetic mechanism of DAT can be described as follows.



Scheme 2 Possible synthetic mechanism of DAT.

With the addition of hydrazine dihydrochloride, the carbon atom of the C≡N group from cyanoguanidine was attacked by one of the nitrogen atom under acidic conditions. By comparison the spectra of cyanoguanidine with intermediate **1**, the peak intensities of C≡N (2156, 2198 cm^{-1}) decreased due to the transfer of electrons in the molecule of cyanoguanidine, which formed new C-N and C=N groups. The absorption peak intensities of the groups of C-N (1350-1100 cm^{-1}) and C=N (1704, 1652 cm^{-1}) were increased, showing cyanoguanidine being transformed to intermediate 1.

As the reaction continued, another nitrogen atom attacked the carbon atom of the C=N group. It shows that the absorption peak intensities of the C≡N group was further decreased in comparison with the IR spectra of the intermediate **1**, which demonstrating that intermediate **1** has transformed to intermediate **2** and the skeleton vibration of heterocyclic compounds (1600~1450 cm^{-1}) has preliminary formed. Finally, the IR spectrum of product shows that the absorption peaks of 2156 and 2198 cm^{-1} disappeared, which proves that the bond of C≡N was not in existence, and the reaction was essentially completed. Thus, intermediate **2** rearranges spontaneously to give DAT as the fianal product.

In conclusion, DAT was synthesized by nucleophilic addition, deamination, cyclization and rearrangement upon reaction of hydrazine dihydrochloride with cyanoguanidine.

5. Conclusions

The in-line ATR-IR fibric probe was used to monitor the synthetic process of DAT. The KICA algorithm has been successfully employed to estimate the spectral data, and the IR spectra of reactants, intermediates and product were obtained, respectively. Furthermore, the intermediates involved in the reaction process were validated by the DFT B3LYP method. It indicates that the results obtained by KICA method are consistent with the calculation results of quantum chemical calculation method. Therefore, the chemometric resolution methods utilized in the present study are reliable. The results of all experiments show that in-line ATR-IR spectroscopy combined with KICA method can be applied to investigate the synthetic mechanism of DAT successfully. In addition, no trap agent is needed to identify the intermediates of the reaction process. KICA method possessed distinctly characteristic peaks and exhibited the variation of peaks' intensity accompanied with the structure changed. KICA method offered a good separation outcome as well as fast processing speed (within 1 minute). And KICA, as a well-known and powerful method, would provide significant guidance to investigate the synthetic reaction mechanism.

Acknowledgments

This research was financially supported by National Natural Science Foundation of China (21175106) and the specialized research Fund for the Doctoral Program of Higher Education (No. 20126101110019).

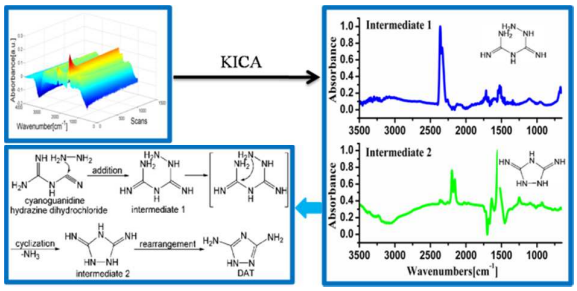
References

1. S. S. Kumar and H. P. Kavitha, *Mini-Rev. Org. Chem.*, 2013, **10**, 40-65.
2. X. M. Deng, J. H. Wang, J. M. Zhang, T. Sim, T. Sasaki, N. D. Kim, T. Sasaki, W. Luther, R. E. George, P. A. Jänne and N. S. Gray, *Med. Chem. Lett.*, 2011, **2**, 379-384.
3. F. R. Brushett, M. S. Thorum, N. S. Lioutas, M. S. Naughton, C. Tornow, H. R. M. Jhong, A. A. Gewirth and P. J. A. Kenis, *J. Am. Chem. Soc.*, 2010, **132**, 12185-12187.

4. J. Hernández-Gil, N. Ovèjak, S. Ferrer, F. Lloret and A. Castiñeiras, *Inorg. Chem.*, 2013, **52**, 2289-2291.
5. M. S. Thorum, J. Yadav, A. A. Gewirth, *Angew. Chem. Int. Ed.*, 48 (2009) 165-167.
6. A. Zarrouk, B. Hammouti, S. S. Al-Deyab, R. Salghi, H. Zarrok, C. Jama and F. Bentiss, *Int. J. Electrochem. Sci.*, 2012, **7**, 5997-6011.
7. V. Thottempudi, H. Gao and J. M. Shreeve, *J. Am. Chem. Soc.*, 2011, **133**, 6464-6471.
8. K. Y. Lee, C. B. Storm, M. A. Hiskey and M. D. Coburn, *J. Energ. Mater.*, 1991, **9**, 415-428.
9. P. Novak, A. Kišić, T. Jednačak, S. Milijanić and G. J. Verbanec, *Pharmaceut. Biomed.*, 2011, **54**, 660-666.
10. B. J. McCullough, T. Bristow, G. O'Connor and C. Hopley, *Rapid Commun. Mass Spectrom.*, 2011, **25**, 1445-1451.
11. G. M. Hamminga, G. Mul and J. A. Moulijn, *Appl. Spectrosc.*, 2007, **61**,470-478.
12. D. Lumpi, W. Christoph, M. Schöpf, E. Horkel, G. Ramer, ~~B. Lendlb~~ B. Lendl and J. Fröhlich, *Chem. Commun.*, 2012, **48**, 2451-2453.
13. E. Karabudak, B. L. Mojet, S. Schlautmann, G. Mul and H. J. G. E. Gardeniers, *Anal. Chem.*, 2012, **84**, 3132-3137.
14. J.G. Li, J. Gao, H. Li, X.F. Yang and Y. Liu, *Anal. Methods*, 2014, **6**, 4305-4311.
15. A. Hyvärinen and E. Oja, *Neur. Netw.*, 2000, **13**, 411-430.
16. C. K. Yoo, J. M. Lee, P. A. Vanrolleghem and I. B. Lee, *Chemom. Intell. Lab. Syst.*, 2004, **71**, 151-163.
17. G. Q. Wang, Z. Y. Hou, Y. X. Tang, J. B. Zhao, Y. A. Sun and D. X. Fu, *Anal. Chim. Acta*, 2010, **679**, 43-48.
18. X. G. Shao, W. Wang, Z. Y. Hou and W. S. Cai, *Talanta*, 2006, **69**, 676-680.
19. X. G. Shao, G. Q. Wang, S. F. Wang and Q. D. Su, *Anal. Chem.*, 2004, **76**, 5143-5148.
20. N. Pasadakis and A. A. Kardamakis, *Anal. Chim. Acta*, 2006, **578**, 250-255.
21. X. G. Shao, Z. C Liu and W. S. Cai, *Analyst*, 2009, **134**, 2095-2099.
22. F. Szabo de Edelenyi, A. W. Simonetti, G. Postma, R. Hou and L. M. C. Buydens, *Anal. Chim. Acta*, 2005, **544**, 36-46.
23. M. J. Zhang, P. J. Tong, W. M. Wang, J. P. Geng and Y. P. Du, *Chemome. Intell. Lab. Syst.*, 2011, **105**, 207-214.
24. A. Hyvärinen and E. Oja, *Neur. Comput.*, 1997, **9**, 1483-1492.
25. A. Hyvarinen, *IEEE Sig. Proc. Lett.*, 1999, **6**,145-147.
26. A. J. Bell and T. J. Sejnowski, *Neur. Comput.*, 1995, **7**, 1129-1159.
27. J. F. Cardoso, *IEEE Sig. Proc. Lett.*, 1997, **4**, 112-114.
28. P. A. Højen-Sørensen, O. Winther, and L. K. Hansen, *Neur. Comput.*, 2002, **14**, 889-918.

29. G. Q. Wang, Y. A. Sun, Q. Z. Ding, C. H. Dong, D. X. Fu, and X. L. Sun, *Spectrochim. Acta A*, 2008, **70**, 571-576.
30. G. Q. Wang, Q. Z. Ding, Y. A. Sun, L. H. He and X. L. Sun, *Anal. Chim. Acta*, 2007, **594**, 101-106.
31. B. Schölkopf, A. Smola, and K. R. Müller, *Neur. Comput.*, 1998, **10**, 1299-1319.
32. C.O. Sakar and O. Kursun, *Expert Syst. Appl.*, 2012, **39**, 3333-3344.
33. H. L. Shen, X. N. Li, and Y. Z. Liang, *Chinese Sci. Bull.*, 2000, **45**, 1368-1373.
34. http://www.cis.hut.fi/projects/ica/fastica/.
35. A. Ciaramella, and R. Tagliaferri, *Neur. Netw.*, 2003, **1**, 708-713.
36. M. W. Wong, *Chem. Phys. Lett.*, 1996, **256**, 391-399.
37. H.Y. Gu, X.Y. Huang, L.Y. Yao, E. Teye and Y. Wen, *Anal. Methods*, 2014, **6**, 3360-3364.
38. J. Zhang, H. Du, F. Wang, X. D. Gong, and Y. S. Huang, *J. Phys. Chem. A*, 2011, **115**, 6617-6621.
39. R. H. Nobes, W. R. Rodwell, L. Radom, *J. Comput. Chem.*, 1982, **3**,561-564.
40. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Jr. Montgomery, T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochter-ski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Sal-vador, J. J. Dannenberg, G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian 03, Revision D.01, Gaussian Inc., Wallingford CT, 2004.
41. N. Tarannum, M. Singh, *Crystallogr. Rep.*, 2014, **59**, 987-992.
42. N. Sundaraganesan, S. Kalaichelvan, C. Meganathan, B. Dominic Joshuaa, and J. Cornardb, *Spectrochim. Acta. A*, 2008, **71**, 898-906.
43.

Graphical Abstract



Investigating the synthetic mechanism of 3, 5-diamino-1, 2, 4-triazole by using fibre optic ATR-IR spectroscopy combined with kernel independent component analysis