

Analytical Methods

Accepted Manuscript



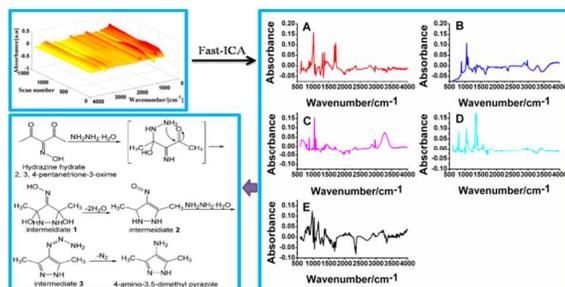
This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

A table of contents entry



The synthesis mechanism of 4-amino-3,5-dimethyl pyrazole was investigated by using in-line FT-IR spectroscopy combined with Fast-ICA algorithm.

1 **Study on the synthesis mechanism of 4-amino-3, 5-dimethyl**
2 **pyrazole: by fibre optic in-line FT-IR spectroscopy combined**
3 **with independent component analysis**

4
5 Jiguang Li, Jun Gao, Hua Li*, Xiaofeng Yang, Yu Liu

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

7 *Xi'an, 710069, China*

8
9 Author to whom correspondence should be addressed:

10 Dr. Hua Li

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

12 Tel: 86-29-88302635

13 Fax: 86-29-88303527

14 *E-mail:* nwufxkx2012@126.com

Abstract

The application of Fourier transform infrared (FT-IR) spectroscopy for in-line monitoring the synthesis process of 4-amino-3, 5-dimethyl pyrazole is described. The obtained data matrix of IR was analyzed by a well known chemometric method, independent component analysis (ICA), which determined the concentration profiles and the spectra of reactant, intermediates and product. The geometric configurations of the intermediates were fully optimized and the vibrational frequencies computed at the density functional theory (DFT) B3LYP/6-31G level of theory. The computational results by the ICA method are in good agreement with those by the quantum chemical calculation method, which demonstrated the reliability of the proposed ICA method. The possible synthesis mechanism was deduced, and the results indicated that ICA combined with in-line FT-IR spectroscopy can be used to study the synthesis mechanism of 4-amino-3, 5-dimethyl pyrazole successfully.

Keywords: Independent component analysis; In-line FT-IR; Density functional theory; Synthesis mechanism

1. Introduction

1H, 4H-3, 6-dinitropyrazolo [4, 3-c] pyrazole (DNPP) is a novel pyrazolo [4, 3-c] pyrazole multi-nitro-heterocyclic energetic compound. Its unique molecular structure endows the compound with excellent thermal stability, good oxygen balance and high standard enthalpy molar of formation. The energy of DNPP is predicted about 85% octogen (HMX).¹ Therefore, it has a broad development prospect and is an important research topic in the field of energetic materials. 4-Amino-3, 5-dimethyl pyrazole is the key intermediate of DNPP and thus the aim of studying the synthesis mechanism is to further improve the yield of DNPP and reduce its cost. The synthetic route is shown in Scheme 1.

“Here Scheme 1”

Herein we provided a safe and fast in-line monitoring technique based on process analytical technology (PAT) for monitoring the synthesis process of 4-amino-3, 5-dimethyl pyrazole. This method employed Fourier transform infrared (FT-IR) spectroscopy coupled with attenuated total reflectance (ATR) to track the depletion of reactants, the change of intermediates and the formation of resultant.² The major advantage of this in-situ technique versus at-line ATR-IR-spectroscopy is that monitoring takes place inside the reaction system, thus eliminating the risks of sample alteration during probing. This method was proved to be a versatile and effective tool to observe the highly reactive intermediates under strongly corrosive conditions, high pressure and high/low temperature,^{3,4} which were close to real experimental conditions, so as to provide the real-time and dynamic information of the reaction components. Furthermore, it's used to collect the data on the multi-dimensional complicated reactions in a short time. This technique provides such a method for further investigating the progress of complicated reactions that can either be determined by tracking changes in absorbance values at selected wavenumbers or applying modern chemometric methods,

1 which process the entire spectra information. Among these multivariate curve resolutions, alternating
2 least squares (MCR-ALS) should be particularly mentioned.⁵ This chemometric resolution method
3 decomposes the recorded data set into smaller matrices containing information on the spectra and the
4 concentration profiles of each component involved in the reaction. Generally the application of this
5 technique needs for combining with other chemometric methods, such as principal component
6 analysis (PCA),⁶ evolving factor analysis (EFA)⁷ or ALS constrained,⁸ which has become possible to
7 analyze multidimensional process data and has been widely applied in IR data analyses.⁹⁻¹¹

8 However, independent component analysis (ICA) is a statistical technique that aims to separate the
9 unobserved independent source variables from the observed variables that are the combinations (or
10 mixtures) of these source variables.^{12,13} ICA relies on the assumption of the source signals are
11 statistically independent and non-Gaussian distributed with the emphasis on the separated components
12 being mutually independent. ICA is not the same as statistical methods based on the second-order
13 moment (e.g., PCA and factor analysis (FA), which are applied for analysis of one matrix), or partial
14 least squares regression (PLSR) and projection pursuit (PP), which relate two matrices. ICA is not the
15 same as the MCR method. It is a statistical approach based on the fourth order moment of the signals
16 with the latent variables produced being chemically interpretable. In recent years, ICA has attracted
17 great interests of chemists and has been applied in chemical studies, e.g., monitoring dynamic and
18 batch processes,^{14,15} processing ultraviolet-visible (UV-Vis) and near-infrared (NIR) spectral data of
19 mixtures,^{16,17} extraction of pure mass spectra from overlapping gas chromatographic-mass
20 spectrometric (GC-MS) data,¹⁸ identifying constituents in commercial gasoline combining with
21 FT-IR,¹⁹ Raman,²⁰ nucleic magnetic resonance (NMR),²¹ electron paramagnetic resonance (EPR)²²
22 and mass spectrometry (MS).²³ To the best of our knowledge, few reports focused on this method in

1 studying synthesis mechanism.

2 In this article, we try to investigate the performance of the combination of in-line IR monitoring
3 reaction process with ICA resolution method to study the synthesis mechanism of 4-amino-3,
4 5-dimethyl pyrazole. The concentration profiles and IR spectra of the components were determined by
5 analyzing the spectral data with ICA method, and the correctness of intermediates was verified by
6 density functional theory (DFT) B3LYP/6-31G level of theory. Then the possible synthesis
7 mechanism was deduced. We compared the analyzed results with Liu²⁴, in which the synthesis
8 mechanism of 4-amino-3, 5-dimethyl pyrazole by MCR-AIS method was studied. We found that the
9 ICA resolution method combined with in-line FT-IR spectroscopy can also be applied to study the
10 synthesis mechanism of 4-amino-3, 5-dimethyl pyrazole effectively. In addition, ICA needed shorter
11 calculation time and shown higher accuracy in comparison with MCR-ALS method. And the above
12 mentioned works were expected to provide significant guidance to the investigation of the reaction
13 mechanism in future.

14 2. Theory

15 A data represented by **D** obtained from FT-IR spectrometer is a bilinear matrix. According to the
16 Lambert–Beer law, it can be expressed simply as follows:

$$17 \mathbf{D} = \mathbf{CS}^T + \mathbf{E} \quad (1)$$

18 Where **D** ($m \times n$) is an observed data matrix (absorbance matrix), m denotes the number of data
19 channels associated with the experimental FT-IR wavenumber range and interval taken, n is the
20 number of scan (one point about 6 seconds). **C** ($m \times d$) is a mixing matrix (concentration matrix), and
21 **S^T** ($n \times d$) is a denoting source matrices (pure components spectra matrix). In addition, **E** ($m \times n$) is the
22 error matrix, often omitted. m and n are respectively the numbers of rows and columns of the data

1 matrix \mathbf{D} , and d is the number of components included in the bilinear decomposition of Equation (1).

2 The basic problem of ICA is to estimate both the mixing matrix \mathbf{C} and the corresponding
3 realizations of source matrix \mathbf{S} , given only observations the matrix \mathbf{D} , on the assumptions that the
4 components in \mathbf{S}^T are mutually statistically independent and the mixing vectors in \mathbf{C} are linearly
5 independent. Therefore, it is equivalent to find a maximum likelihood of \mathbf{C} , when \mathbf{W} is the
6 pseudoinverse of the mixing matrix \mathbf{C} , $\mathbf{W} = \mathbf{C}^{-1}$, the estimated source signal will be equal to the
7 original source signal \mathbf{S}^T :

$$8 \quad \mathbf{S}_{\text{new}}^T = \mathbf{W}\mathbf{D} = \mathbf{W}\mathbf{C}\mathbf{S} = \mathbf{S}^T \quad (2)$$

9 Using the ICA algorithm, we can obtain the rows of \mathbf{S} whose norm is 1. Compared to the PCA, the
10 \mathbf{S} and \mathbf{W} matrix in Eq. (2) may be considered as a loading matrix and a score matrix, i.e. \mathbf{W} can be
11 regarded as the score matrix \mathbf{T} , while \mathbf{S} can be treated as loading matrix \mathbf{P} .²⁵

12 Considering the accuracy and computational complexity, Fast-ICA is one of the most popular
13 algorithms for linearly ICA, which has a fast convergence and it can deal with sub-Gaussian and
14 super-Gaussian signal.²⁶⁻²⁷ And ICA is based on a fixed-point iteration scheme for finding a maximum
15 of the non-Gaussianity of components sequentially using a deflation scheme.²⁸ Therefore, this
16 algorithm was adopted to analyze the observed data matrix. The basic form of the one-unit version of
17 the Fast-ICA algorithm is as follows:

18 (1) Choose an initial (e.g., random) weight vector, w_q .

19 (2) Let $w_q = E\{zg(w_q^T X)\} - E\{g'(w_q^T z)\}w_q$, Here, g is the first derivative and g' is the second
20 derivative of G , $G_1(u) = \frac{1}{a_1} \log \cosh(a_1 u)$.

21 (3) Let $w_q = w_q / \|w_q\|$

22 (4) If not converged, go back to (2).

Note that convergence means that old and new values of w_q are in the same direction, i.e. their dot-product is (almost) equal to 1. It is unnecessary for the vector to converge to a single point.

3. Experiment

3.1. Reagents and solutions

All solutions were prepared with double-distilled water. 2, 3, 4-Pentanetrione-3-oxime was prepared according to the literature.¹ Hydrazine hydrate and acetyl acetone (all are analytical-reagent grade and purchased from Chengdu Kelong Chemical Reagent Factory, Sichuan, China).

3.2. Instrument

IR spectra were recorded on a FT-IR spectrometer (Bruker Optics Vertex70, Germany) equipped with an ATR probe (IN350-T, Germany) made of diamond combined with bundles of mid-IR optical fibers and a mercury cadmium telluride (MCT) detector in a range of 4000-600 cm^{-1} . A 1.5 m long silver halide fiber with a two bounce ZnSe ATR element fully immersed in the solution during the reaction. All spectra were recorded with an accumulation of 16 scans and 4 cm^{-1} spectral resolution.

3.3. Synthesis of 4-amino-3, 5-dimethyl pyrazole

2, 3, 4-Pentanetrione-3-oxime (3.6 g) and ethanol (15 ml) were added to a three-necked flask (100 ml), to which the in-line ATR probe was inserted. The background was then scanned. After that hydrazine hydrate (3 ml) was added to the above solution slowly that the temperature would not exceed 5 $^{\circ}\text{C}$ for 50 min. Then the solution was heated to 80 $^{\circ}\text{C}$ and kept for 100 min. Finally, the solvent was evaporated and the desired product was collected by filtration and dried in air.

3.4. Data-sets

The in-line IR monitoring spectra of synthesizing 4-amino-3, 5-dimethyl pyrazole were collected by Bruker Vertex 70 spectrometer. After the reaction process finished, a three-dimensional diagram was

1 shown in Fig. 1. Three-dimensional diagram was converted into a two-dimensional observed matrix
2
3
4
5
6 **D** (1763×1441) which consisted of 1763 wavelength points and 1441 scan number by the Bruker
7
8 software package OPUS.
9

10 **3.5. Software**

11 All programs were performed in MATLAB (Mathworks Inc., Natick, MA, USA) for windows. The
12
13 PCA method was obtained from the PLS-Toolbox-5.5 (Eigenvector Research, Inc., USA), and
14
15
16
17
18
19 Fast-ICA was downloaded from the website.²⁹ Other in-house software written in MATLAB was used
20
21 for other data manipulations.
22

23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
“Here Fig. 1”

“Here Table 1”

11 **4. Results and discussion**

12 **4.1. Data processing**

13 4.1.1 .Determination of the principal component number

14 The veracity of the data was easily affected by the direct analysis as the original data we obtained
15 usually included the baseline noise. Therefore, under the premise that the accuracy of the analysis was
16 not significantly influenced, we filtered the above two-dimensional matrix **D**. PCA can reduce the
17 calculation errors, because it has the advantage of determining the component number in the reaction
18 before using the ICA resolution method. The eigenvalues and accumulated contribution efficiency
19 were pretreated by PCA procedure as shown in Table 1. The percentage of the variance captured by
20 PCA is useful in assessing the importance of principal components. We applied the obtained results to
21 the above synthetic route (Scheme 1) and compared the number of factors 4, 5, 6, 7. Results showed
22 that the accumulated contribution efficiency reached 99.90% when the number of factor was selected

1 as 5, which meant that the 5 principal components (PCs) could basically represent the major
2 information of the reaction process.

3 4.1.2. Analysis by Fast-ICA and MCR-ALS

4 The filtered two-dimensional matrix **D** was input to the Fast-ICA program and MCR-ALS,
5 respectively. To begin to calculate, the new estimates of spectral matrix **S** and concentration matrix
6 **C** would be obtained after each iterative calculation. The iterative calculations were repeated until
7 convergence to obtain the real meaningful concentration curve (Fig. 2) and IR spectrum of each
8 component (Fig. 3).

9 **“Here Fig. 2”**

10 **“Here Fig. 3”**

11 Compared Fig 3a with Fig. 3b, it can be seen that the shape of some characteristic bands was no
12 difference from reactant (*A*) and product (*E*). However, the shape of some characteristic bands was
13 reversal from Fig. 3b, especially for the intermediate **1**(*B*), **2**(*C*) and **3**(*D*) ($\sim 1000\text{ cm}^{-1}$), and the
14 characteristic bands of intermediate **1** and **2** have no significantly difference. These troubles could
15 have an influence on deducing the accuracy of synthesis mechanism. Therefore, we determine to
16 choose the estimated results by the Fast-ICA algorithm to deduce the synthesis mechanism of
17 4-amino-3, 5-dimethyl pyrazole. In addition, another advantage of the Fast-ICA method is that it
18 needs only about 10 seconds to perform the estimation of the matrix **D**, while MCR-ALS method
19 needs 1.5 hours because the calculation of EFA consumes most of the time (the specific calculation
20 time based on the size of a matrix).

21 **4.2. Theoretical calculation of various substances**

1 DFT is now the electronic structure method of choice for accurate calculations of the properties of
2 large molecules, it is reliable and widely used algorithm in the reaction process.³⁰⁻³¹ It was reported
3 that the molecular structure and performance calculated by DFT B3LYP were closer to the
4 experimental observation.³² In this study, in order to verify the correctness of each intermediate, the
5 geometric configuration of each intermediate were fully optimized and the vibrational frequencies
6 computed at the B3LYP/6-31G level of theory. All the calculations was carried out using the program
7 package Gaussian 03.³³

8 4.2.1. Optimized geometric configurations

9 The optimized geometric configurations of intermediates are shown in Fig. 4. The calculation
10 results indicated that the vibration frequencies of the optimized geometric configurations were all
11 positive, which meant that they were the lowest energy point on the potential energy surfaces. The
12 result demonstrated that the optimized geometries were relatively stable.

13 **“Here Fig. 4”**

14 4.2.2. Vibrational frequencies and IR spectra

15 Fig. 5 depicts the relationship between the vibration frequency (wavenumber/cm⁻¹) and **the**
16 intensity of each intermediate. The force fields calculated at the B3LYP/6-31G level of theory were
17 scaled down via a single scale factor, 0.9613, which is considered to be the best at the level.³⁴
18 Intermediate **1(a)** mainly has the following strong absorption peaks. The peaks at 1116 and 1191 cm⁻¹
19 belong to the stretching vibration of C-N groups. The peaks at 1381 and 1440 cm⁻¹ are assigned to the
20 symmetrical and asymmetric bending vibration of C-H groups, respectively. The peaks at 1518 and
21 1533 cm⁻¹ are both in-plan bending vibration of N-H groups. Intermediate **2(b)** exhibits several strong
22 absorption peaks. The peak at 1235 cm⁻¹ attributes to the stretching vibration of C-N group, and 1303

1
2
3 1 cm^{-1} the stretching vibration of N=O group. The peaks at 1444 and 1456 cm^{-1} are assigned to the
4
5 2 symmetrical and asymmetric bending vibration of C-H groups, respectively. The peaks at 1539 and
6
7 3 1573 cm^{-1} are both in-plan bending vibration of N-H groups. Intermediate **3(c)** has several strong
8
9 4 absorption peaks as well. The peak at 1240 cm^{-1} is due to the stretching vibration of C-N group. The
10
11 5 peaks at 1484 and 1647 cm^{-1} correspond to the stretching vibration of N=N and C=N group,
12
13 6 respectively. The peak at 1519 cm^{-1} is resulted from in-plan bending vibration of N-H group. And the
14
15 7 peak at 1597 cm^{-1} is in-plane bending vibration of secondary amine N-H.

21 8 **“Here Fig. 5”**

22
23
24 9 The calculated example indicated that the results of vibrational frequencies and IR spectra analyzed
25
26 10 by B3LYP/6-31+G are reliable.^{35,36} The peak positions of the calculated spectra by B3LYP/6-31+G
27
28 11 agrees well with those of the analyzed spectra obtained by Fast-ICA (only IR characteristic peaks
29
30 12 were compared due to the complicated experimental environment). These analyzed results prove the
31
32 13 reliability of the ICA resolution method we used in this paper.

36 14 **4.3. Synthesis mechanism of 4-amino-3, 5-dimethyl pyrazole**

37
38
39 15 As experimental data was obtained during a dynamic reaction process, IR spectrum of each
40
41 16 substance analyzed by Fast-ICA was different from the standard sample. In addition, the experimental
42
43 17 materials and standard samples were measured under different conditions. Therefore, it could not be
44
45 18 fully compared with the standard spectrum (KBr pellets), and only characteristic peaks were used in
46
47 19 comparison with the peaks of each pure substance. The possible synthesis mechanism (Scheme 2) of
48
49 20 4-amino-3, 5-dimethyl pyrazole was deduced by combining the changes of functional groups in IR
50
51 21 spectra with the further qualitative theoretical analyses of intermediates.

52
53
54
55
56 22 From the above concentration profiles and IR spectra of reactant, intermediates and product are
57
58
59
60

1 shown in Fig. 2a and Fig. 3a, respectively. The synthesis mechanism of 4-amino-3, 5-dimethyl
2 pyrazole could be derived as follows.

3 Fig. 3a (A) shows the strong stretching vibration absorption peaks of C=O (1722 cm^{-1}) and C=N
4 groups (1685 cm^{-1}), this indicates that it is the reactant IR spectrum of 2, 3, 4-pentanetrione-3-oxime.
5 When diamid hydrate was slowly added into 2, 3, 4-pentanetrione-3-oxime, a lone pair electrons was
6 one of amino nitrogen that attacked a carbon atom on the carbonyl of 2, 3, 4-pentanetrione-3-oxime,
7 which weakened the stretching vibration absorption intensity of C=O (1722 cm^{-1}) and C=N (1685
8 cm^{-1}) groups. As the reaction continued, another lone pair electrons of diamid hydrate attacked the
9 remained carbon atom of 2, 3, 4-pentanetrione-3-oxime. Fig. 3a (B) shows that the peaks at C=O
10 (1722 cm^{-1}) completely disappeared, while the C-N ($1350\sim 1000\text{ cm}^{-1}$) groups appeared and the bond
11 of C=N was shift to 1647 cm^{-1} . By comparison Fig. 3a (A) and Fig. 3a (B), it demonstrated that
12 intermediate **1** was formed. Subsequently, intermediate **1** removed two molecules of water under
13 alkaline condition (hydrazine hydrate), and rearranged into the stable intermediate **2** which is aromatic.
14 And the absorption peak of N=O (about 1329 cm^{-1}) appeared, the corresponding IR spectrum can be
15 seen in Fig. 3a (C). When the temperature of the reaction system risen, the bond of N=O was
16 transformed to the N=N ($1330\sim 1500\text{ cm}^{-1}$), and the absorption peaks of N-H ($3140\sim 3340\text{ cm}^{-1}$) and
17 C=N (1685 cm^{-1}) were appeared, indicating that intermediate **2** was transformed to azo compound
18 intermediate **3** (Fig. 3a (D)) due to the strong reducibility of hydrazine hydrate. The absorption
19 intensities of N=N ($1300\sim 1500\text{ cm}^{-1}$) decreased, and the absorption peak of $-\text{NH}_2$ (3337 cm^{-1}) and
20 pyrazole ring skeleton vibration (1607 cm^{-1}) appeared, so it can be seen that intermediate **3** removed
21 one molecule of N_2 under reflux conditions. Finally, 4-amino-3, 5-dimethyl pyrazole was formed (Fig.
22 3a (E)). The resolved results are consistent with the previous studies.^{1,24}

1
2
3
4 **1 References**

- 5
6 1. Y. F. Luo, Z. X. Ge, B. Z. Wang, H. H. Zhang and Q. Liu, *Chin. J. Energ. Mater.*, 2007, **15**, 205-207.
- 7
8
9 2. Z. H. Lin, L. L. Zhou, A. Mahajan, S. Song, T. Wang, Z. H. Ge and D. Ellison, *J. Pharm. Biomed. Anal.*, 2006, **41**,
10
11 99-104.
- 12
13
14 3. D. Lumpi, W. Christoph, M. Schöpf, E. Horkel, G. Ramer, B. Lendlb and J. Fröhlich, *Chem. Commun.*, 2012, **48**,
15
16 2451-2453.
- 17
18
19 4. D. Lumpi, C. Braunschier, C. Hametner, E. Horkel, B. Zachhuber, B. Lendl and J. Fröhlich, *Tetrahedron Letters.*, 2009,
20
21 **50**, 6469-6471.
- 22
23
24 5. M. Garrido, F. X. Rius and M. S. Larrechi, *Anal. Bioanal. Chem.*, 2008, **390**, 2059-2066.
- 25
26
27 6. S. Wold, K. Esbensen and P. Geladi, *Chemom. Intell. Lab. Syst.*, 1987, **2**, 37-52.
- 28
29
30 7. H. R. Keller and D. L. Massart, *Chemom. Intell. Lab. Syst.*, 1992, **12**, 209-224.
- 31
32
33 8. Jr. J. Workman, B. Lavine, R. Chrisman and M. Koch., *Anal. Chem.* 2011, **83**, 4557-4578.
- 34
35
36 9. M. Garrido, I. Lázaro, M. S. Larrechi and F. X. Rius., *Anal. Chim. Acta.* 2004, 515, 65-73.
- 37
38
39 10. F. Zhang, Y. Chen and H. Li, *Electrophoresis.*, 2007, **28**, 3674-3683.
- 40
41
42 11. T. Azzouz and R. Tauler, *Talanta.*, 2008, **74**, 1201-1210.
- 43
44
45 12. P. Comon, *Signal Process.*, 1994, **36**, 287-314.
- 46
47
48 13. A. Hyvärinen and E. Oja, *Neural Networks.*, 2000, **13**, 411-430.
- 49
50
51 14. J. M. Lee, C. K. Yoo and I. B. Lee, *Chem. Eng. Sci.*, 2004, **59**, 2995-3006.
- 52
53
54 15. C. K. Yoo, J. M. Lee, P. A. Vanrolleghem and I. B. Lee, *Chemom. Intell. Lab. Syst.*, 2004, **71**, 151-163.
- 55
56
57 16. 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.
- 58
59
60 17. X. G. Shao, W. Wang, Z. Y. Hou and W. S. Cai, *Talanta.*, 2006, **69**, 676-680.
18. X. G. Shao, G. Q. Wang, S. F. Wang and Q. D. Su, *Anal. Chem.*, 2004, **76**, 5143-5148.

- 1
2
3
4 19. N. Pasadakis and A. A. Kardamakis, *Anal. Chim. Acta.*, 2006, **578**, 250-255.
- 5
6 20. Jr. L. Silveira, Jr. A. R. de Paula, C. A. Pasqualucci and T. P. Marcos Tadeu, *Instrum. Sci. Technol.*, 2008, **36**,
7
8 134-145.
- 9
10 21. F. Szabo de Edelenyi, A. W. Simonetti, G. Postma, R. Hou and L. M. C Buydens, *Anal. Chim. Acta.*, 2005, **544**,
11
12 36-46.
- 13
14 22. C. Q. Chang, J. Y. Ren, Peter C. W. Fung, Y. S. Hung, J. G. Shen and Francis H. Y. Chan. *J. Magn. Reson.*, 2005,
15
16
17
18
19 **175**, 242-255.
- 20
21 23. M. Zhang, P. Tong, W. Wang, J. P. Geng and Y. P. Du, *Chemom. Intell. Lab. Syst.*, 2011, **105**, 207-214.
- 22
23 24. Y. Liu, X. F. Xiong, Y. F. Luo, B. Z. Wang, K. Wang and H. Li, *Chin. J. Comput. Appl. Chem.*, 2012, **29**, 527-531.
- 24
25 25. C. K. Yoo, D. S. Lee and P. A. Vanrolleghem, *Water. Res.*, 2004, **38**, 1715-1732.
- 26
27 26. P. Tichavsky, Z. Koldovsky and E. Oja, *Signal Process.*, 2006, **54**, 1189-1203.
- 28
29 27. X. Giannakopoulos, J. Karhunen and E. Oja, *Int. J. Neural. Syst.*, 1999, **9**, 99-114.
- 30
31 28. A. Hyvarinen, *IEEE Trans. Neural Networks.*, 1999, **10**, 626-634.
- 32
33 29. <http://www.cis.hut.fi/projects/ica/fastica/>.
- 34
35 30. M. W. Wong, *Chem. Phys. Lett.*, 1996, **256**, 391-399.
- 36
37 31. J. Zhang, H. Du, F. Wang, X. D. Gong and Y. S. Huang, *J. Phys. Chem. A.*, 2011, **115**, 6617-6621.
- 38
39 32. B. Lambie, R. Ramaekers and G. Maes, *J. Phys. Chem. A.*, 2004, **108**, 10426-10433.
- 40
41 33. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Jr. Montgomery, T.
42
43 Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G.
44
45 Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida,
46
47 T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J.
48
49 Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y.
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3
4 1 Ayala, K. Morokuma, G. A. Voth, P. Sal-vador, J. J. Dannenberg, G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C.
5
6 2 Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S.
7
8
9 3 Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T.
10
11 4 Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W.
12
13
14 5 Wong, C. Gonzalez and J. A. Pople, Gaussian 03, Revision D.01, Gaussian Inc., Wallingford CT, 2004
15
16 6 34. A. P. Scott and L. Radom, *J. Phys. Chem.*, 1996, **100**, 16502-16513.
17
18
19 7 35. S. L. Zhang, X. F. Xiong, T. Yu, Y. B. Wang, B. Z. Wang, Z. X. Ge, G. H. Zhai and H. Li, *Chem. J. Chin. Univ.*,
20
21 8 2012, **33**, 1444-1449.
22
23
24 9 36. W. P. Lai, P. Lian, B. Z. Wang, Y. F. Luo, Z. X. Ge, Z. Z. Zhang and Y. Q. Xue, *Chin. J. Comput. Appl. Chem.*, 2007,
25
26 10 **24**, 1025-1028.
27
28
29 11
30
31 12
32
33
34 13
35
36 14
37
38
39 15
40
41 16
42
43
44 17
45
46 18
47
48
49 19
50
51 20
52
53
54 21
55
56 22
57
58
59
60

1 Captions

2 **Scheme 1.** Synthetic route of 4-amino-3, 5-dimethyl pyrazole.

3 **Scheme 2.** The possible synthesis mechanism of 4-amino-3, 5-dimethyl pyrazole.

4 **Fig.1.** The three-dimensional diagram of in-line FT-IR spectra (4000-600 cm^{-1}) recorded during the
5 synthesis of 4-amino-3, 5-dimethyl pyrazole by 2, 3, 4-pentanetrione-3-oxime and hydrazine hydrate.

6 **Fig.2.** Concentration profiles of reactant (a), intermediates **1(b)**, **2(c)** and **3(d)** and product (e)
7 analyzed by Fast-ICA (a) and MCR-ALS (b).

8 **Fig.3.** IR spectra of reactant (A), intermediates **1(B)**, **2(C)** and **3(D)** and product (E) analyzed by
9 Fast-ICA (a) and MCR-ALS (b).

10 **Fig.4.** Geometric configuration of intermediates **1(A)**, **2(B)** and **3(C)** fully optimized at the DFT
11 B3LYP/6-31G level of theory.

12 **Fig.5.** IR spectra of intermediates **1(a)**, **2(b)** and **3(c)** computed at the DFT B3LYP/6-31G level of
13 theory.

1 **Table 1** Results of principal component analysis

Principal component number	Eigenvalue of covariance (D)	Variance captured this PC (%)	Variance captured total (%)
1	5.33	81.89	81.89
2	1.02	15.66	97.92
3	1.04×10^{-1}	1.59	99.12
4	4.52×10^{-2}	0.69	99.81
5	5.72×10^{-3}	0.09	99.90
6	4.73×10^{-3}	0.07	99.98
7	5.63×10^{-4}	0.01	99.99

2

3

4

5

6

7

8

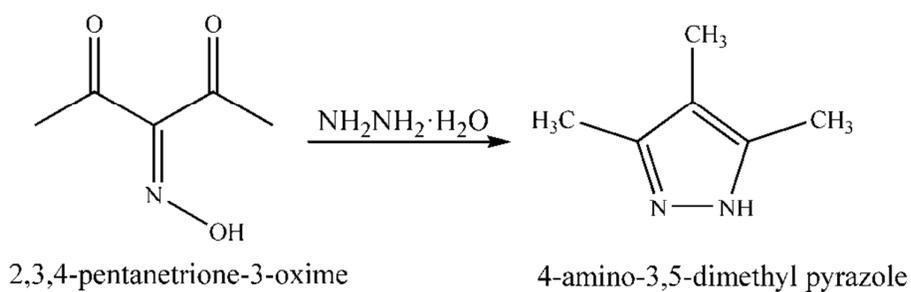
9

10

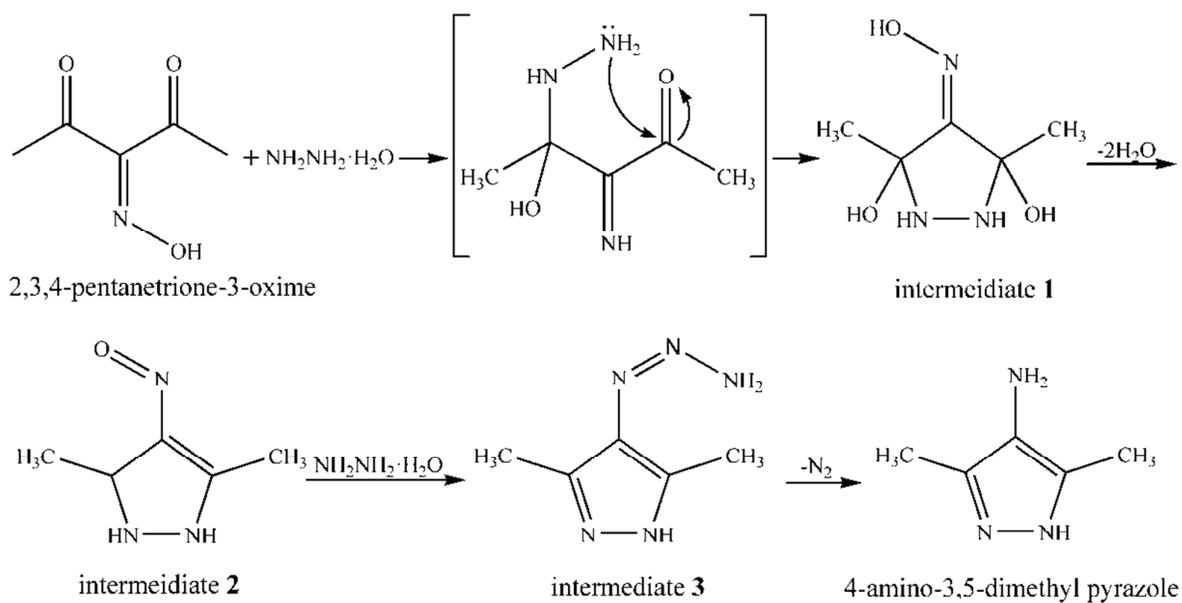
11

12

19



Scheme 1



Scheme 2

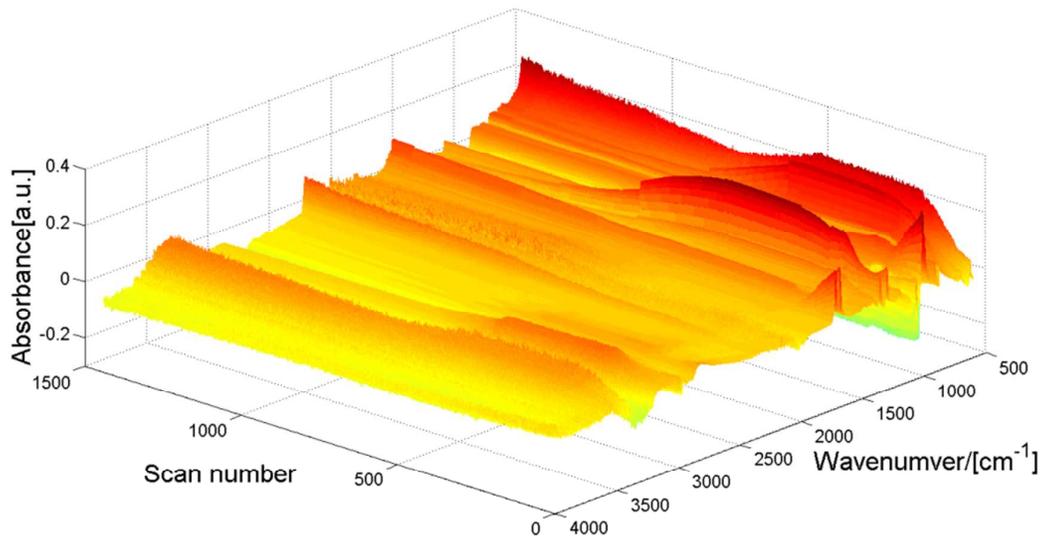


Fig. 1

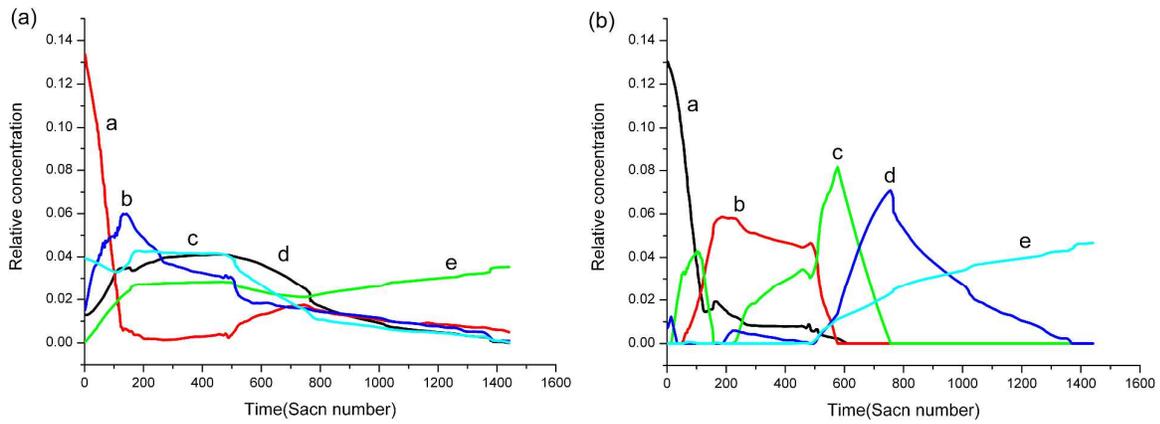


Fig. 2

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

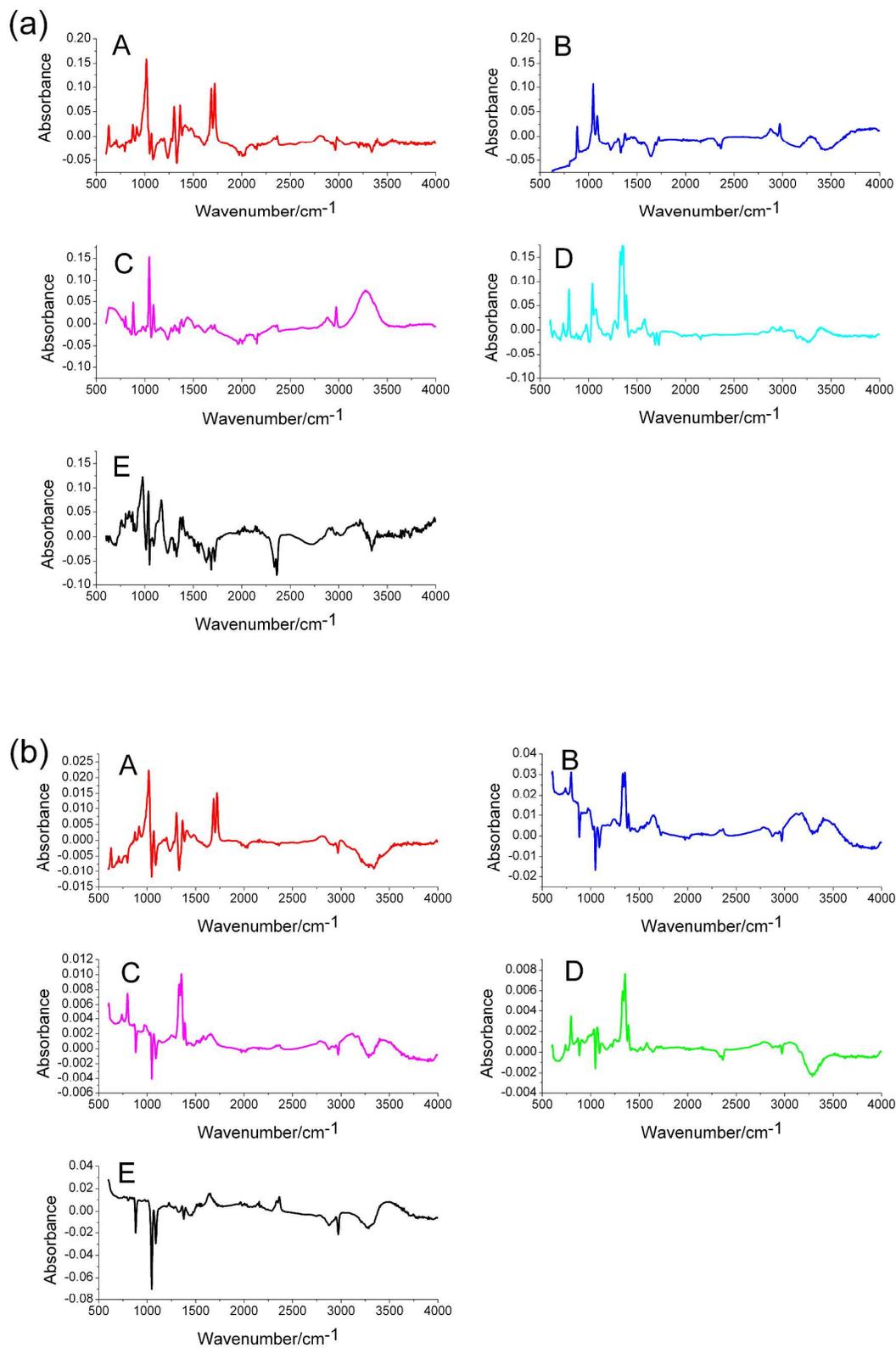


Fig. 3

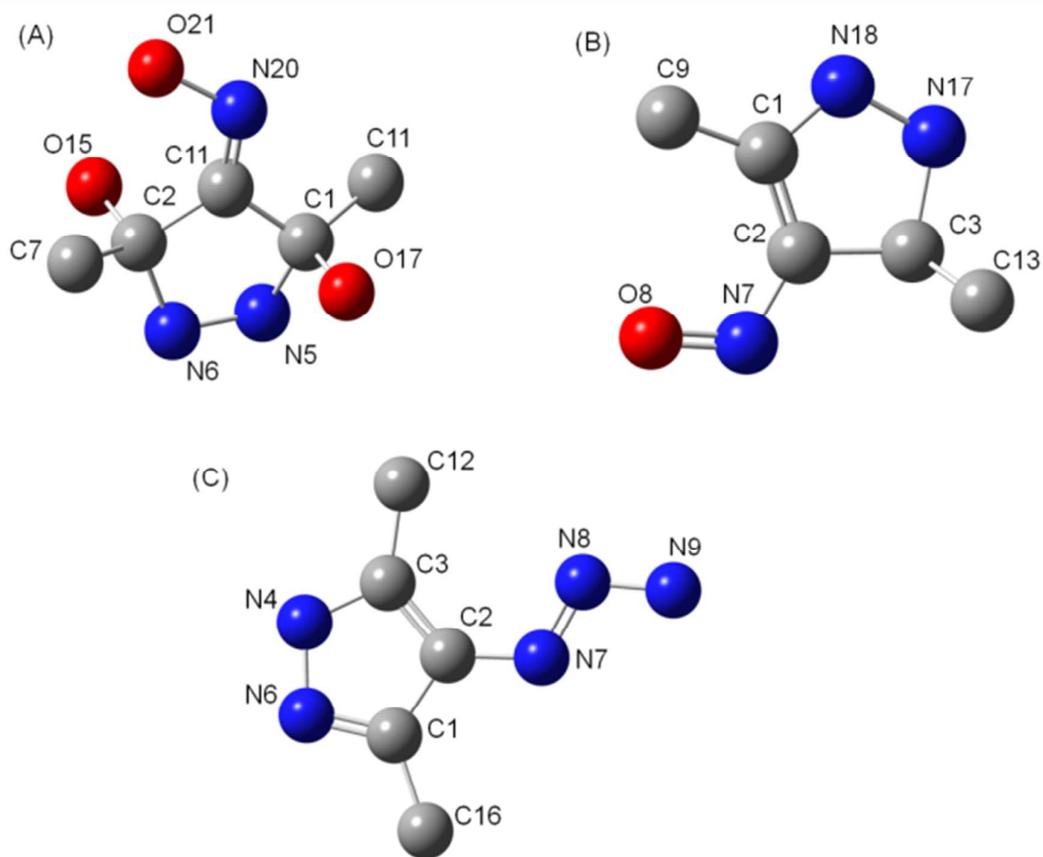


Fig. 4

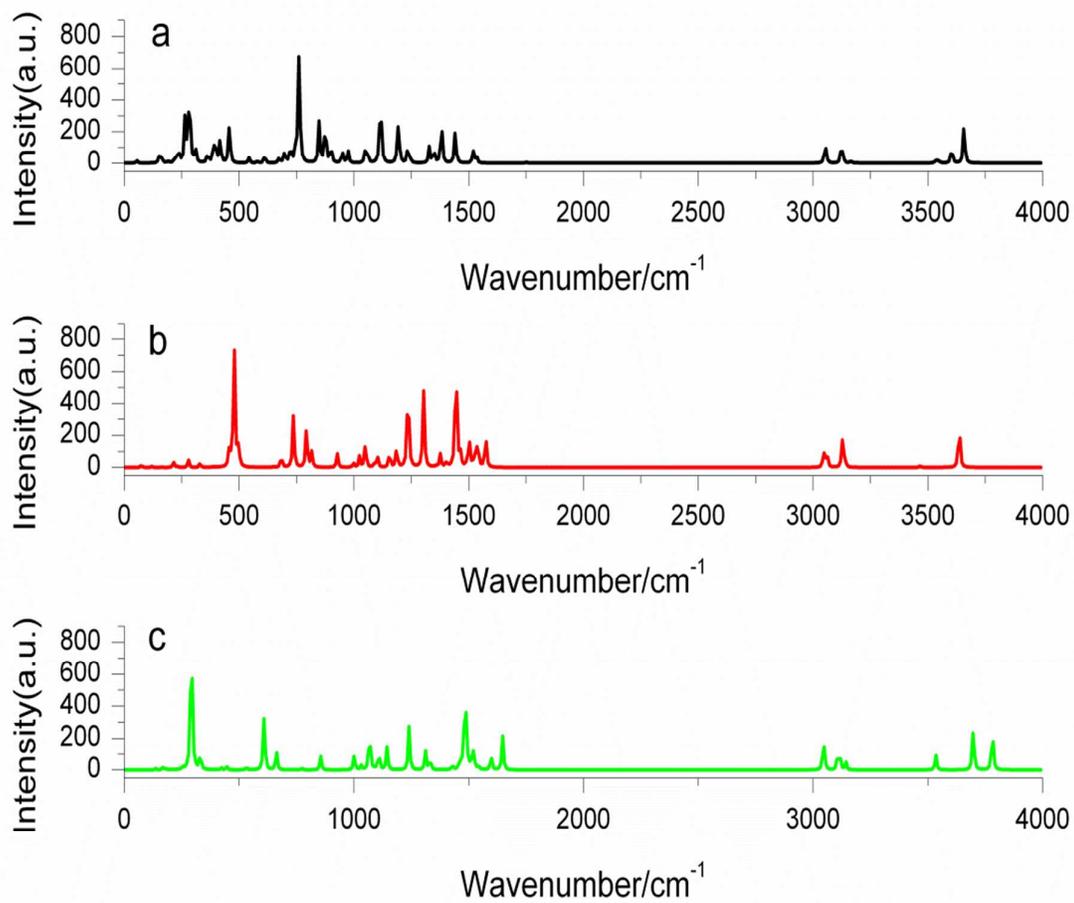


Fig. 5