Combination of Micro-PIXE with the Pattern Recognition Technique for the Source Identification of Individual Aerosol Particles

P. GUO,* J. WANG, X. LI, J. ZHU, T. REINERT, J. HEITMANN, D. SPEMANN, J. VOGT, R.-H. FLAGMEYER, and T. BUTZ

Shanghai Institute of Nuclear Research, CAS, P.O. Box 800-204, 201800 Shanghai, China (P.G., J.W., X.L., J.Z.); and Fakultät für Physik und Geowissenschaften, Universität Leipzig, Abteilung Nukleare Festkörperphysik, Linnestr. 5, 04103 Leipzig, Germany (J.Z., T.R., J.H., D.S., J.V., R.-H.F., T.B.)

In order to identify the sources of individual aerosol particles, we developed an identification system based on the combination of the micro-PIXE (particle-induced X-ray emission) technique with the pattern recognition (PR) technique. It uses the micro-PIXE spectrum of an individual aerosol particle as its fingerprint to identify the origin of the particles. A total of 294 micro-PIXE spectra of individual aerosol particles, collected at Shanghai City in China, were applied to test the PR system. These particles were collected from several pollution sources and a local environment monitor site. Most of the environment particles were identified, and the most probable sources were assigned by the PR system. It was found that cement factories and vehicle exhaust are the major contributors. About 14% of the particles from the environment monitor site could not be identified by their spectra. They might have come from some other sources we have not yet investigated. The validity and the efficiency of the system are evaluated and discussed. It is demonstrated that the combination of micro-PIXE with the pattern recognition technique offers a new, powerful tool for the source identification of individual aerosol particles.

Index Headings: Individual aerosol particles; Pattern recognition technique (PR); Micro-PIXE.

INTRODUCTION

One of the major problems in air pollution studies is the source identification of airborne particles. Although bulk analytical methods, such as broad-beam particle-induced X-ray emission (PIXE), using multivariate statistics techniques constitute a useful tool for identifying sources of the aerosol particles, they still have some limitations that can be avoided by single-particle analysis. The association of certain elements on a particular particle can be observed directly by single-particle measurement. Some elements tend to occur at relatively high concentration in a small particle. However, in a bulk measurement the concentration of these elements may be at a level lower than the detection limit. If these particles are measured individually, the associated elements can be determined easily. More advantages of the single particle analysis have been discussed by Van Espen,[1] Orilic,[2] and Grime.[3] Some successful analyses of individual airborne particles by nuclear microprobes have been reported.[4–7] The difficulty in accurate quantification of the element concentration in a single particle has been discussed by Jaksic et al.,[8] Bogdanovic et al.,[9] and Orilic et al.[10] To facilitate the source identification, a large-scale investigation of individual aerosol particles should be undertaken. Converting these spectra into exact element concentration is a difficult and time-consuming task. An ideal method should be rapid, automated, and relatively inexpensive.

Two samples with similar chemical composition should have similar spectral patterns. Therefore, the micro-PIXE spectra of individual aerosol particles can be used as the fingerprints of their chemical composition. In the source identification of an individual aerosol particle, one is mostly interested in whether a given particle has a spectral pattern similar to that of a known source particle, rather than its quantitative chemical composition. From this point of view, we propose the application of a pattern recognition technique for the identification of individual aerosol particles by using their micro-PIXE spectra. This approach essentially emulates visual inspection of the spectral features by a spectroscopist, but places the inspection on a more quantitative basis. It can be easily implemented with a personal computer. The pattern recognition approach has been adopted in practical applications of many analytical spectroscopy techniques.[11–13]

A set of individual aerosol particles from seven pollution sources was analyzed by a nuclear microprobe. Their micro-PIXE spectra were used to establish a fingerprint database for the pollution sources. Other individual aerosol particles were collected from a local environment monitor site and were analyzed with the same facilities. These particles were identified for their origins by comparing their micro-PIXE spectrum patterns with those in the fingerprint database by using the pattern recognition technique.

EXPERIMENTAL

Sample Collection and Preparation for Nuclear Microprobe Analysis. The pollution source samples of individual aerosol particles were collected from seven industrial plumes, such as ferrous smelters, oil and coal combustors, cement factories, construction sites, vehicle exhaust, and soil dusts. They contribute the majority of the aerosol particles floating in the atmosphere over Shanghai City. The environmental monitor samples were collected at the city center.

A cascade impact sampler (Sierra Anderson Instruments, Inc., Model HY-1) was used, and the aerosol particles were collected on thin polystyrene fiber filters. Ninety-seven percent of the collected particles are smaller
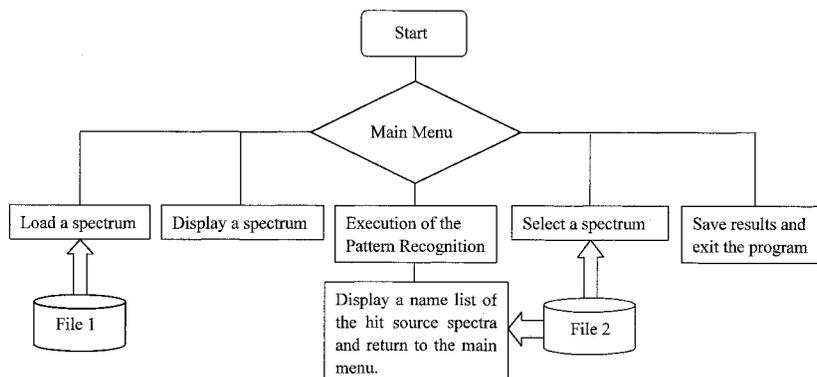
FIG. 1. The flow chart of the spectrum pattern recognition system. File 1: the spectrum files of individual aerosol particles; File 2: the spectrum files in the database of pollution sources.

than 10 μm and 70% of them are smaller than 3 μm. Several methods for the preparation of individual aerosol particles have been reported in our earlier work.[14] However, none of them could separate small particles (<3 μm) clearly, and only the groups of the aerosol particles were measured.[15] In order to measure the single particles one by one, it is essential to isolate the particles from each other on a thin foil. Furthermore, the isolated particles should be close enough to be found easily during a beam scanning. Keeping these requirements in mind, a new method for the preparation of individual aerosol particles was developed.

A solution of nylon powder in iso-butyl alcohol was prepared at a temperature of 80 °C. A droplet of the solution was dropped onto a surface of deionized water contained in a rotating beaker at room temperature. The droplet stretched, and a very thin (0.2 μm) nylon foil formed on the water surface. Just after the formation of the foil (in a few seconds), the collected single particles were dropped dipersively from the sampling filter onto the sticky foil by a small shaker. After five minutes of solidification, the foil attached on a stainless steel frame was taken out of the water. The single particles were separately embedded in the thin, taut, and flat nylon foil. After 24 h drying at a temperature of 60 °C, the individual aerosol particle was ready for single particle analysis by nuclear microprobes.

Nuclear Microprobe Irradiation and Spectra Acquisition. A reliable and efficient experiment for individual aerosol particle analysis demands good performance from the nuclear microprobe. Because most of the particles are smaller than 3 μm, both high spatial resolution and beam current are needed. The Leipzig nuclear microprobe, LIPSION, satisfied the experimental requirements. A 100 μm object diaphragm and a 100 μm aperture diaphragm were used to produce a focused beam spot with 1 μm size and 80 pA current for most single-particle measurements. The particles were so small that they could be considered as thin targets. The 2.25 MeV protons could penetrate all the particles. The beam charge
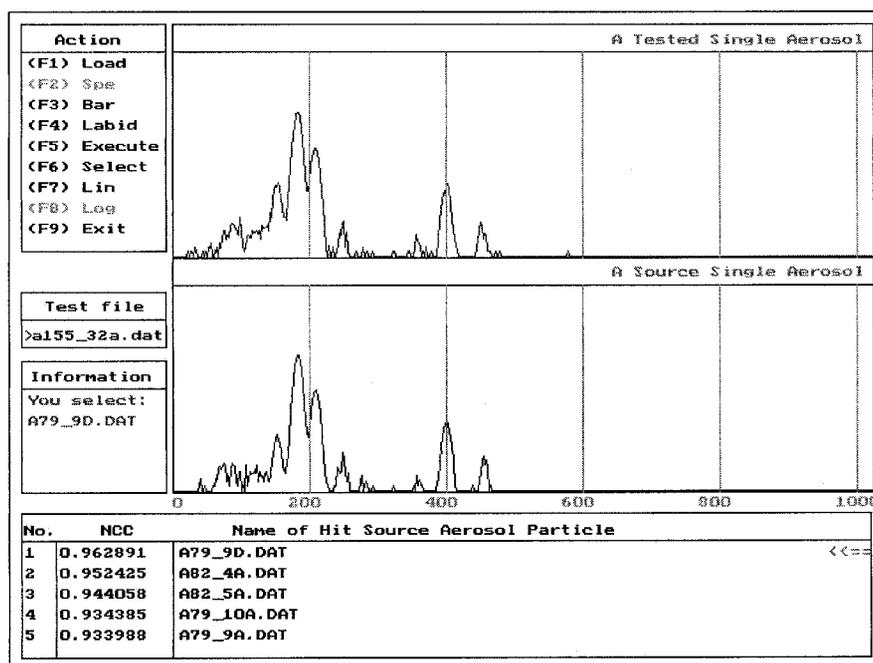

FIG. 2. The program menu on the screen. NCC in the bottom window shows the values of cos θ.
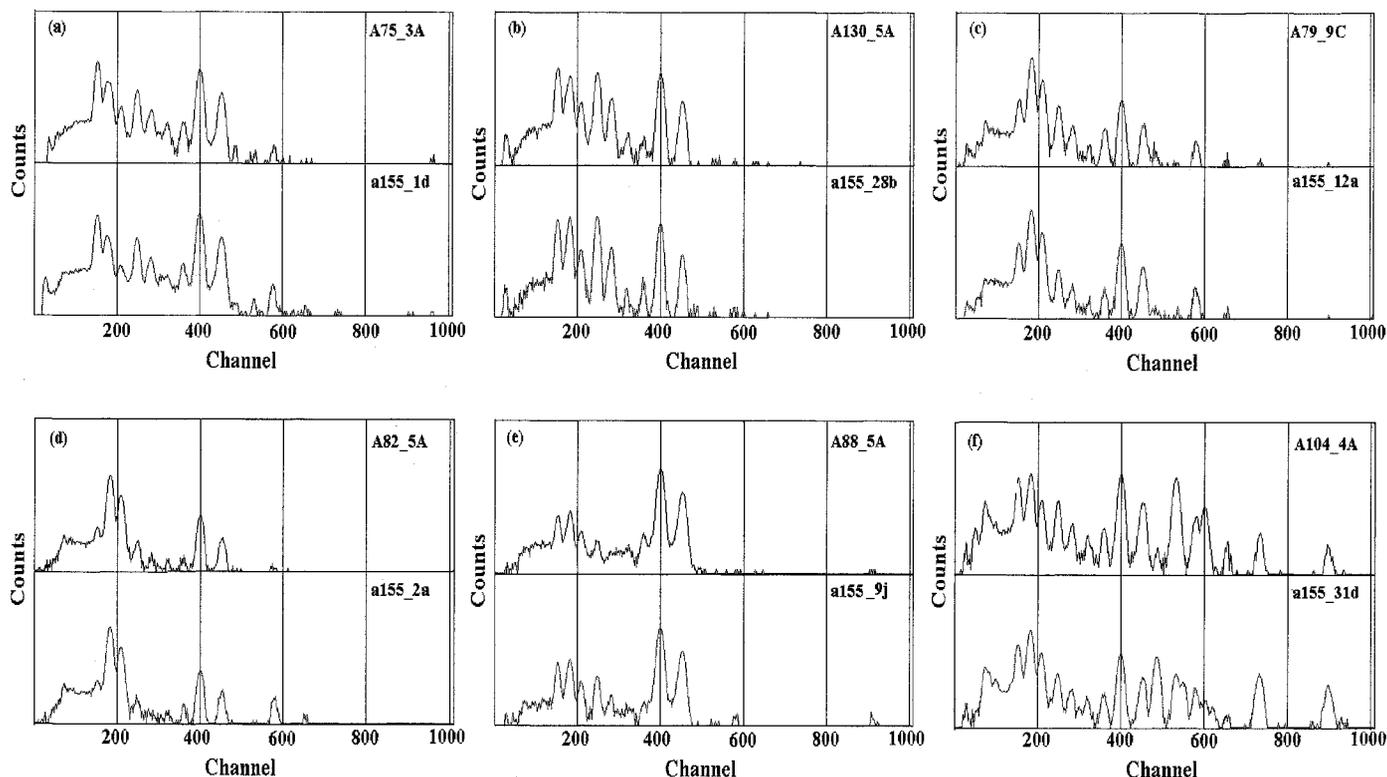
FIG. 3. Six pairs of the PIXE spectra of the particles from the pollution sources (top) and an unknown individual aerosol particle (bottom), which is recognized by this system as a particle from the top pollution source. (**a**) A75_3A is from soil dusts; (**b**) A130_5A is from coal combustors; (**c**) A79_9C is from cement factories; (**d**) A82_5A is from construction sites; (**e**) A88_5A is from ferrous smelters; and (**f**) A104_4A is from vehicle exhaust.

TABLE I. Values of cos θ derived from the micro-PIXE spectra of 12 typical individual aerosol particles. All spectra were taken under identical conditions and with similar counting statistics. Here part A is for the shape method and part B is for the peak intensity method. Two spectra having very similar patterns should have a value of cos θ near 1. A discriminate threshold of 0.95 is used in this table to indicate the strong correlation between two samples.[a]

	A75_3a	A155_1d	A130_5a	A155_28b	A79_9c	A155_12a	A82_5a	A155_2a	A88_5a	A155_9j	A104_4a	A155_31d
(A) Values of cos θ for the shape method												
a75_3a												
A155_1d	**0.98**											
A130_5a	0.94	0.94										
A155_28b	0.93	0.92	**0.98**									
A79_9c	0.92	0.90	0.93	0.93								
A155_12a	0.93	0.90	0.93	0.93	**0.98**							
A82_5a	0.86	0.82	0.87	0.87	0.93	0.94						
A155_2a	0.85	0.81	0.85	0.85	0.94	0.93	**0.98**					
A88_5a	0.93	0.93	0.88	0.85	0.82	0.85	0.80	0.78				
A155_9j	0.93	0.93	0.91	0.89	0.85	0.88	0.82	0.80	**0.97**			
A104_4a	0.86	0.87	0.84	0.83	0.84	0.85	0.79	0.80	0.80	0.81		
A155_31d	0.84	0.83	0.82	0.82	0.86	0.86	0.81	0.82	0.76	0.77	**0.95**	
(B) Values of cos θ for the method of peak intensities												
A75_3a												
A155_1d	**0.97**											
A130_5a	0.93	0.92										
A155_28b	0.94	0.93	**0.96**									
A79_9c	0.93	0.93	0.91	0.93								
A155_12a	0.93	0.91	0.90	0.91	**0.98**							
A82_5a	0.87	0.88	0.87	0.87	0.93	0.93						
A155_2a	0.83	0.85	0.82	0.83	0.91	0.93	**0.98**					
A88_5a	0.91	0.90	0.87	0.90	0.89	0.91	0.91	0.86				
A155_9j	0.93	0.93	0.92	0.93	0.93	0.94	0.93	0.91	**0.97**			
A104_4a	0.90	0.89	0.85	0.86	0.93	0.91	0.81	0.79	0.80	0.85		
A155_31d	0.90	0.87	0.83	0.83	0.93	0.91	0.80	0.78	0.78	0.82	**0.95**	

[a] Values of cos θ that are above the discriminate threshold appear in boldface.
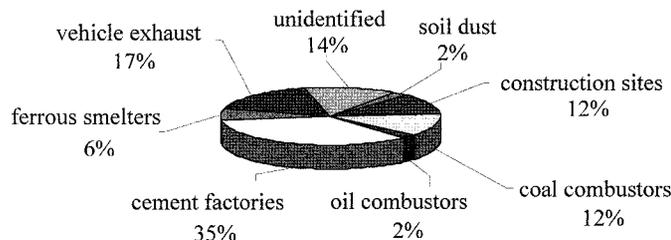
FIG. 4. The contribution of the pollution sources to the local atmosphere in Shanghai.

FIG. 5. The clustering results of some unidentified particles.

was collected by a downstream Farady cup. There was no conductive coating on the sample surface. No data correction for the effects of the particle matrix and shape was applied. A 20 nC integrated beam charge was used for each particle measurement. Much experimental time was spent on searching for isolated small particles. A proper searching procedure was important for the high efficiency of this work. At first, a large area scan (70 μm) was needed to find areas where the isolated small particles were present. The coordinates of each small particle in the scanning area were digitized and recorded for further small area scans (6–10 μm). The three-dimensional movement of the sample stage was digitally controlled. The accuracy of the sample displacement was better than 1 μm. One or more isolated small particles were covered by the beam within the scanning area. We measured 294 isolated individual aerosol particles. Each single particle measurement took about 15 min. The accelerator and the microprobe of LIPSION performed under a very stable working condition. The high stability of the beam and the high precision of the sample displacement made the individual aerosol particle measurements very efficient. The characteristic spectrum of each single particle was extracted from the data according to the particle shape in the small scanning area.

PATTERN RECOGNITION SYSTEM

Principle of the System. The principle of the pattern recognition technique has been described in detail elsewhere.[13] Here we only briefly introduce the method and its application to PIXE spectra.

A PIXE spectrum can be treated as an *n*-component column vector in a Euclidean space. For the pattern recognition, two spectrum vectors are considered similar if the Euclidean angle between the two vectors is small enough. The discriminative function of two spectrum vectors is defined as below:

$$\cos \theta = \vec{X}' \cdot \vec{Y} / [(\vec{X}' \cdot \vec{X})(\vec{Y}' \cdot \vec{Y})]^{1/2} \qquad (1)$$

Here $\vec{X}'$ and $\vec{Y}'$ are the transposes of $\vec{X}$ and $\vec{Y}$, respectively. The numerator is the dot product of the two vectors, and the denominator is the product of their magnitudes. The value of $\cos \theta$ will be close to 1 for two PIXE spectra with similar features, and it is independent of intensities.

There are two ways to transform a spectrum into a Euclidean vector; one is according to the spectrum shape, and the other is by the intensity of each peak in the spectrum. A set of 25 peaks was set up for the vector in our work. The peak intensities were acquired by using the
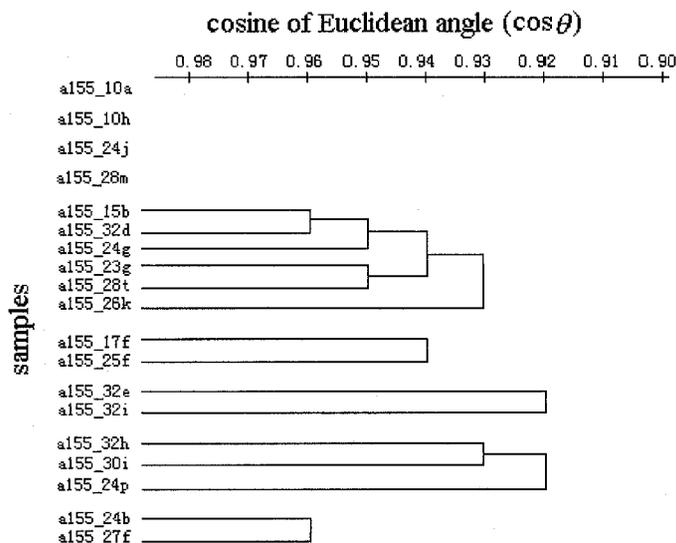
program AXIL. If a peak in the spectrum is lower than the detection limit, the corresponding intensity is zero. The method of the spectrum shape is simple and quick; however, the peak intensity method is favorable to some trace element features.

Implementation. A program to identify a PIXE spectrum of an individual aerosol particle has been developed on a personal computer. The flow chart is shown in Fig. 1. The identification of an individual aerosol particle by its observed PIXE spectrum pattern follows this process. At the start, the main menu (Fig. 2) is shown on the screen. The user first selects a file name with the function key F1 "Load" to enter a spectrum, which will be displayed on the screen instantly. Next, F5 "Execute" can be used to calculate the values of cos θ between the loaded spectrum and each of the spectra from the fingerprint database. Then a name list of the hit source spectra will be presented at the bottom of the screen. Finally, the user can use F6 "Select" to choose a satisfactory source spectrum from the list and to display the spectrum on the screen for comparison. The system contains all the spectra of the source particles as a fingerprint database. The user can add or delete any of the source spectra in the database. Both logarithmic and linear scales are used for the spectrum display and comparison. The logarithmic scale is more sensitive to trace element features of the particles. The PIXE spectrum has high background at the low-energy region because of *bremsstrahlung*. The pattern recognition algorithm may misunderstand a spectrum with high background, especially when the peak signals are very low. To tackle this problem, we used a background subtraction procedure before the spectrum comparison.

RESULTS AND DISCUSSION

We measured 294 individual aerosol particles. Among them, 74 particles are from seven pollution sources. The other particles are from a local environment monitor site. The spectral patterns can be considered as the fingerprints of the individual aerosol particles.

Figure 3 shows six pairs of spectra for comparison.

Each pair contains two spectra; one is a typical spectrum from a pollution source, and the other is an unknown spectrum from the monitor site. By visual inspection, it is obvious that the two spectra in each pair have similar features, which implies similar chemical composition. The spectra in different pairs have different features of their sources.

To illustrate the validity and efficiency of the pattern recognition technique, we applied Eq. 1 to some typical PIXE spectra of particles. The values of $\cos \theta$ were evaluated for all possible pairs of the spectra and are presented in a triangular matrix shown in Table I. Here part A is for the shape method and part B is for the peak intensity method. A discriminate threshold of 0.95 is used in this table to indicate the strongly correlated samples. For example, the value of $\cos \theta$ between a75_3a and a155_1d is 0.98, which means that the particle a155_1d came from the pollution source a75_3a. The result of the calculation coincides with the visual inspection in Fig. 3.

The results of source identification are shown in Fig. 4. The cement factory is the primary source of pollution. This is a result of municipal construction in this region of Shanghai in recent years. The second large pollution source in the region is the heavy traffic. About 14% of the aerosol particles have not been identified by this system. They do not originate from the pollution sources we have investigated. This observation suggests that they might come from other pollution sources. If we set the threshold of discrimination (value of $\cos \theta$) to 0.92, some unidentified particles are clustered into nine categories by means of the pattern recognition system. Figure 5 shows the clustering results. These nine sources were not defined in our study. This result is useful in the search for new pollution sources.

CONCLUSION

A graphical user interface program for the identification system of aerosol particles based on PIXE spectra has been developed. The combination of the micro-PIXE with the pattern recognition technique for the source identification of individual aerosol particles is efficient and easy to operate even for an inexperienced person. According to the results of identification, we could see that seven types of pollution sources were not enough for the identification of all aerosol particles in Shanghai. Some new pollution sources can be found by clustering the unidentified particles.

1. P. Van Espen, Mikrochim. Acta **114/115,** 129 (1994).
2. I. Orlic, Nucl. Instr. And Meth. B **104,** 602 (1995).
3. G. W. Grime, X-ray Spectrom. **27,** 221 (1998).
4. N. E. G. Löwenstam, J. Pallon, and H. C. Hansson, J. Aerosol Sci. **19,** 1031 (1988).
5. P. Artaxo, W. Maenhout, H. Storms, and R. Van Grieken, J. Geophys. Res. **95,** 16971 (1990).
6. P. Artaxo, M. L. C. Rabello, F. Watt, and G. Grime, Nucl. Instr. Meth. B **75,** 521 (1993).
7. G. W. Grime and F. Watt, Nucl. Instr. Meth. B **75,** 495 (1993).
8. M. Jaksic, I. Bogdanovic, E. Gereda, S. Fazinic, and V. Valkovic, Nucl. Instr. Meth. B **77,** 505 (1993).
9. I. Bogdanovic, M. Jaksic, S. Fazinic, and V. Valkovic, Nucl. Instr. Meth. B **85,** 732 (1994).
10. I. Orlic, T. Osipowicz, F. Watt, and S. M. Tang, Nucl. Instr. Meth. B **104,** 630 (1995).
11. E. K. Kemsley, S. Ruault, and R. H. Wilson, Food Chem. **54,** 321 (1995).
12. M. Twomey, G. Downey, and P. B. McNulty, J. Sci. Food Agric. **67,** 77 (1995).
13. L. I. Yin and S. M. Seltzer, Advances X-Ray Anal. **33,** 603 (1990).
14. Y. Sha, Y. Gu, G. Liu, A. Wang, P. Zhang, J. Zhu, J. Shi, S. Yang, P. Liu, and X. Li, Nucl. Instr. Meth. B **109/110,** 79 (1996).
15. Y. Sha, J. Shi, Y. Gu, J. Zhu, and X. Li, Nucl. Sci. Techn. (in Chinese), **7,** 61 (1996).