Journal of Industrial Microbiology, 6 (1990) 71–75
Elsevier

SIM 00234

Evaluation of infrared spectroscopy as a bacterial identification method

Alan H. Lipkus [1], Krishnan K. Chittur [2], Stephen J. Vesper [3], Jayne B. Robinson [4] and George E. Pierce [5]

Battelle Columbus Division, 505 King Avenue, Columbus, OH 43201, U.S.A.

(Received 16 December 1988; revised 4 June 1989; accepted 6 June 1989

Key words: Bacterial identification; Infrared spectroscopy

SUMMARY

A study motivated by the recent revival of interest in the use of IR spectroscopy to identify bacteria is reported. A library of FT-IR spectra of dried bacterial films was compiled using 16 different strains. A test set was compiled from spectra of the same strains grown several months later. The test set was quantitatively compared with the library on the basis of spectral similarity in the region 980–1190 cm^{-1}. Six of the strains in the test set were not matched with the correct strain in the library despite efforts to reproduce the conditions under which cells were grown and prepared. The results suggest that reproducibility of the bacterial spectra is a potential difficulty that must be addressed by any attempts to develop FT-IR spectroscopy as a bacterial identification method.

INTRODUCTION

The procedures needed to identify many microorganisms by conventional biochemical and serological tests are often laborious [2]. As a result, a variety of alternative approaches incorporating chromatographic and spectroscopic techniques have been developed. One of the most commonly used of these approaches is based on pyrolysis of the cells followed by gas-liquid chromatography or mass spectrometry [17]. Another approach uses extraction of fatty acids followed by gas chromatography [10]. Other methods for microbial identification, purely spectroscopic in nature, are also being evaluated. One attraction of these methods is that they can operate on whole, even living, cells. Among the techniques being studied in this regard are Fourier transform infrared (FT-IR) [5,6,12,13], fluorescence [3,15], and resonance Raman [4,7] spectroscopies.

Infrared spectroscopy has several attractive features as a possible physical method for bacterial identification. It examines the total chemical composition of the sample, so it is not necessary that the bacteria incorporate any special probe molecule, e.g. a chromophore. Also, no extensive preparation of the bacteria for spectroscopic analysis is required. Investigations into the use of infrared spectroscopy for bacterial identification began in the early 1950's when it was recognized that there can be significant differences between the IR spectra of different bacterial species [16]. Despite much effort in this area [14], however, no practical identification system was developed.

One obstacle to this early work was the limited data-handling and computing capabilities then available. This restriction was effectively removed with the introduction of computer-interfaced FT-IR spectrometers. These instruments can perform sophisticated data manipulations routinely. Furthermore, the equipment is relatively inexpensive, and spectral acquisition is rapid. These attributes have led to a recent revival of interest in the use of IR spectra as bacterial 'fingerprints' [5,6,12,13].

Correspondence: ([2] Present address): K.K. Chittur, Biomedical Engineering Department, Case Western University, Cleveland, OH 44106, U.S.A.
Present addresses:
[1] Research Department, Chemical Abstracts Service, Columbus, OH 43210, U.S.A.
[3] Emery Chemicals, 4900 Este Avenue, Cincinnati, OH 45232, U.S.A.
[4] Dept of Microbiology, Ohio State University, Columbus, OH 43210, U.S.A.
[5] Celgene Corporation, 7 Powder Horn Drive, Warren NJ 07059, U.S.A.

We now report on a study designed to evaluate the prospects for a practical system of rapid bacterial identification based on FT-IR spectroscopy. This study involves the analysis of a relatively large database of FT-IR spectra obtained from various bacterial strains. For each strain, the database includes spectra of samples grown at two separate times, from 5 to 11 months apart. We are thus able to test the long-term reproducibility of the spectra, a crucial matter in regard to practical applications of a spectral library.

MATERIALS AND METHODS

Bacterial cultivation

The 16 strains used in this study are listed in Table 1; they are all type strains from the American Type Culture Collection (ATCC). With one exception, the liquid media and growth conditions used were those recommended for each strain by the ATCC [Catalogue of Strains I (ATCC, Rockville, MD, ed. 15, 1982)]; the exception was *A. simplex*, which was grown on tryptic soy agar. To ensure purity and identity of all type strains, representative clones were subjected to standard biochemical testing prior to use [2]. The results of these tests were consistent with their ATCC identification. Each strain was grown in duplicate (in flasks referred to as 1 and 2). Flasks were inoculated using a fresh starter culture (2.5% inoculum) and were incubated overnight ($\sim$16 h) in a shaking water bath at the optimum temperature for that strain. Cell suspensions were washed twice with 0.9% NaCl. The cell pellets were resuspended in 2 ml of 0.9% NaCl. Isolates were stored between testing periods on agar slants at refrigerator temperatures (contamination was detected in the stored isolates of *A. simplex* and *B. subtilis*, so these were reordered from ATCC). Each strain was grown in duplicate again (in flasks to be referred to as 3 and 4) by the same procedure several (5 to 11) months later.

Sample preparation

For each of the washed cell suspensions, prepared as described above, the cell density was determined microscopically with a Petroff-Hausser counter, and the suspension was adjusted, as necessary, to contain 10^{10}–10^{11} cells/ml. From each of these final suspensions, two samples in the form of dried films were prepared for spectroscopic analysis. (Bacterial samples have been dried in most previous work [5,12–14,16], but cells in aqueous solution have been used recently [6]). Each film was prepared by placing one drop of the cell suspension on a clean CaF_2 window, which was then dried in a vacuum desiccator. In all, 128 samples (i.e. 16 strains × 4 flasks × 2 windows) were prepared in this manner.

TABLE 1

Bacterial strains used

Bacteria	ATCC number
Gram-negative:	
Enterobacter cloacae	13047
Escherichia coli	11775
Escherichia coli	23545
Escherichia coli	25922
Escherichia coli	31616
Klebsiella pneumoniae	8044
Proteus vulgaris	13315
Pseudomonas aeruginosa	10145
Pseudomonas fluorescens	13525
Serratia marcescens	13880
Gram-positive:	
Arthrobacter crystallopoietes	15481
Arthrobacter simplex	6946
Bacillus subtilis	6051
Rhodococcus erythropolis	4277
Rhodococcus rhodochrous	12674
Staphylococcus epidermidis	14990

Spectral acquisition

The mid-IR absorbance spectrum of each sample was obtained on a Digilab model FTS-60 spectrometer equipped with a mercury-cadmium-telluride detector. The CaF_2 window was placed in a standard transmission cell for measurement. Spectra were acquired at a resolution of 4 cm^{-1} by coaddition of 512 scans. Interferograms were processed using triangular apodization and one level of zero filling. After thorough cleaning and drying, a spectrum of the window was obtained to serve as the reference for absorbance calculations.

Data processing

The strain specificity of bacterial spectra can potentially be degraded by sources of noise. Baseline variation is a common source of (low-frequency) noise in biological FT-IR spectroscopy. We assumed that over the spectral region of interest (see below) the baseline variation is linear and we applied to each of the spectra a recently developed filtering method that compensates for linear

variation in the baseline [9]. In order to compare the spectra quantitatively, a measure of the similarity between two spectra must be defined. The measure used in this study regards the (filtered) spectra as vectors and takes the cosine of the angle between the vectors as their similarity [1]. Thus, the similarity between spectra $\vec{A}_m$ and $\vec{A}_n$ is

$$S_{mn} = \vec{A}_m \cdot \vec{A}_n / |\vec{A}_m| |\vec{A}_n| \quad (1)$$

Because it is an angular measure, this similarity is invariant to changes in the magnitudes of the spectral vectors. Thus, this measure automatically compensates for proportional absorbance differences arising from unavoidable fluctuations in the number of cells in each sample of bacterial film. Since all of the bacterial spectra are very much alike in terms of their basic 'waveform', their similarity usually falls between 0.9 and 1.0, which is the maximum value of this measure.

RESULTS AND DISCUSSION

A spectral region that is highly variable among bacteria is the region just below 1200 cm^{-1}. This part of the IR spectrum has been examined before as a possible fingerprint for bacteria [5,12]. Several of the FT-IR spectra in the database compiled for this study are shown in Fig. 1 over the region used in this work, 980 to 1190 cm^{-1}. The vibrational modes of several common groups contribute to this region. The most significant of these groups in biological systems are P=O and C-O. These groups appear in many different molecular structures in virtually all of the constituents of a bacterium, though it has been proposed that polysaccharides associated with the bacterial cell wall make the dominant contribution in this region [5]. A conventional assignment of a bacterial spectrum cannot be made in detail, but this is unimportant since assignment is unnecessary in using these spectra for identification purposes.

The database of bacterial spectra was analyzed in two ways. The first analysis was designed to be a 'control', which would avoid the problem of long-term reproducibility but address the reproducibility between spectra of cells grown at the same time. Each pair of spectra obtained for every strain from cells grown in flask 1 was put into a library, while each pair of spectra obtained from cells grown in flask 2 was put into a test set. Since flasks 1 and 2 were grown at the same time, variations in cell cultivation and sample preparation procedures within each strain should be minimal between the two groups. The test set was then used to challenge the library. For each strain, the two spectra in the test set were compared with the two spectra for every strain in the library. The four similarity values between the two test spectra and the two library spectra were calculated, and the highest of these four values was retained as a measure of the match between those test and library strains.

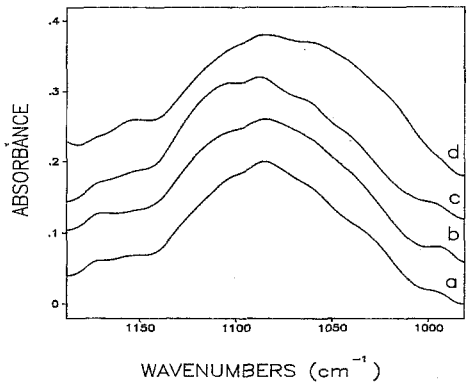

Fig. 1 FT-IR spectra of several bacteria in the region 980–1190 cm^{-1} (a) *E. cloacae*, (b) *P. fluorescens*, (c) *A. simplex*, (d) *B. subtilis*. The spectra are offset and normalized for clarity.

The best matches for each test strain are given in Table 2 (analysis A). To make the presentation easier, each value S in Table 2 is represented in scaled form: $10^4 \times S - 9000$. These results show that for each test strain the best match (highest similarity) is correct; i.e. the matching strain in the library is the same as the test strain. Thus, when there is a minimum of procedural variation within each strain, the FT-IR spectra of the particular strains studied do indeed act like unique fingerprints. This suggests that if appropriate reference strains were grown at the same time as the unknown strains, reliable identification by spectrum matching might be generally feasible.

In the second analysis of the database, the long-term reproducibility of the bacterial spectra was examined in terms of its effect on their fingerprinting ability. All of the spectra obtained from flasks 1 and 2 were put into a new library, and all the spectra obtained from flasks 3 and 4 grown several months later were put into a new test set. As before, the test set was used to challenge the library. For each strain, the four spectra in the test set were compared with the four spectra for every strain in the library. In each such comparison, all of the 16 similarity values were calculated, and the highest was retained.

The best matches for each test strain are shown in

TABLE 2

Similarity matching of bacterial FT-IR spectra

Test strain[b]	Best matches[a]							
	A: Library = Flask 1 Test set = Flask 2				**B:** Library = Flasks 1 and 2 Test set = Flasks 3 and 4			
Gram-negative								
1. *E. cloacae*	997 (1)	974 (9)	962 (6)	962 (4)	976 (1)	936 (4)	932 (9)	925 (6)
2. *E. coli* (11775)	997 (2)	995 (9)	986 (4)	973 (1)	974 (2)	968 (4)	959 (9)	958 (1)
3. *E. coli* (23545)	996 (3)	984 (4)	975 (2)	949 (8)	930 (13)	920 (4)	920 (3)	886 (5)
4. *E. coli* (25922)	996 (4)	990 (3)	979 (2)	966 (8)	996 (4)	979 (2)	978 (3)	966 (8)
5. *E. coli* (31616)	999 (5)	924 (13)	844 (3)	795 (4)	948 (5)	939 (13)	863 (6)	763 (4)
6. *K. pneumoniae*	995 (6)	920 (1)	893 (2)	890 (13)	978 (6)	962 (1)	925 (2)	923 (9)
7. *P. vulgaris*	1000 (7)	971 (9)	970 (8)	964 (10)	977 (4)	954 (2)	952 (3)	948 (1)
8. *P. aeruginosa*	997 (8)	982 (4)	979 (9)	973 (2)	991 (8)	987 (9)	986 (7)	980 (10)
9. *P. fluorescens*	998 (9)	978 (8)	975 (2)	973 (7)	993 (9)	985 (2)	974 (1)	967 (4)
10. *S. marcescens*	998 (10)	977 (7)	972 (1)	967 (9)	992 (10)	987 (1)	977 (9)	976 (7)
Gram-positive								
11. *A. crystallopoietes*	994 (11)	949 (8)	940 (4)	937 (14)	903 (13)	804 (6)	791 (3)	778 (5)
12. *A. simplex*	995 (12)	981 (14)	960 (8)	953 (7)	981 (14)	971 (8)	967 (10)	959 (11)
13. *B. subtilis*	998 (13)	924 (5)	829 (3)	813 (6)	994 (13)	911 (5)	874 (6)	831 (4)
14. *R. erythropolis*	999 (14)	979 (8)	968 (12)	946 (9)	991 (14)	982 (12)	970 (8)	961 (11)
15. *R. rhodochrous*	995 (15)	903 (12)	789 (14)	763 (7)	972 (11)	962 (14)	935 (8)	915 (12)
16. *S. epidermidis*	997 (16)	851 (10)	841 (1)	809 (6)	934 (6)	925 (13)	923 (2)	921 (1)

[a] The four highest similarity values are shown for each of the two analyses, **A** and **B**. Similarities are scaled as described in text. For each similarity value, the identification number of the corresponding library strain is given in parenthesis (same numbering as in column 1).

[b] Different *E. coli* strains are distinguished by their ATCC numbers.

Table 2 (analysis B). For six of the strains (numbers 3, 7, 11, 12, 15, and 16) the best match is incorrect. These incorrect matches are not evenly distributed with regard to Gram stain classification: there are two among the ten Gram-negative strains, but four among the six Gram-positive strains. For the strains most closely related by conventional taxonomy, i.e. the four *E. coli* strains, one (ATCC 23545) is incorrectly matched; this is the only one of the six incorrectly matched strains in which the correct match is found among the four best matches.

The preponderance of incorrect matches among the Gram-positive strains appears to be significant: if six errors were distributed among the 16 strains at random, four or more would fall among the six Gram-positive strains only 9% of the time. The specific causes of this apparent correlation of spectral reproducibility with Gram stain classification are unclear. It may originate in the major differences between the chemical composition of Gram-positive and -negative cell walls, e.g. the higher lipid content of the latter.

The problems with spectral reproducibility can be illustrated by comparing a strain that was correctly matched with one that was incorrectly matched. In Fig. 2 are shown the most similar pairs of library and test spectra for *S. marcescens* (correctly matched) and *P. vulgaris* (incorrectly matched). Examination of the *P. vulgaris* spectra, as well as other pairs of incorrect matches, does not indicate the physical reasons for these failures of reproducibility. Since the microbial cultures were stored between tests on agar slants, it is possible that in some cases natural genetic drift has taken place and resulted in slightly different biochemical characteristics. However, we believe that this type of handling procedure reflects the most typical use. It may be that freeze-drying the cultures between tests would have provided greater stability in the population.

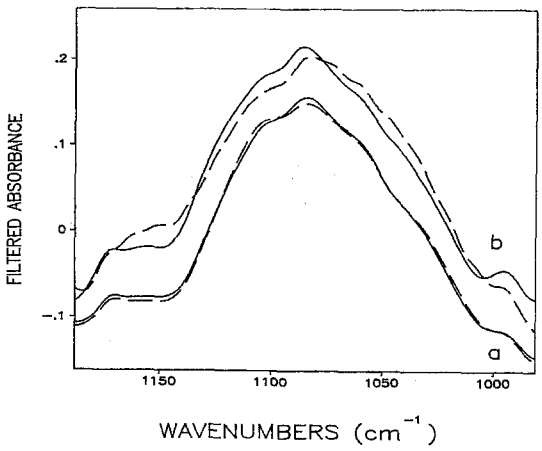

Fig. 2 Illustration of spectral reproducibility. (a) Library spectrum (solid curve) and test spectrum (dashed curve) of *S. marcescens*; (b) Library spectrum (solid) and test spectrum (dashed) of *P. vulgaris*. All spectra are baseline filtered and normalized [9]; the top pair is shown offset for clarity.

The fact that the number of incorrect matches made in this study was relatively high suggests that the development of FT-IR spectroscopy as a general method for rapid bacterial identification has some potential difficulties and possible limitations. These incorrect matches occurred despite what we thought were reasonable efforts to provide identical conditions for cell cultivation and sample preparation for both the library and test set. This work makes clear that the reproducibility of bacterial FT-IR spectra cannot be taken for granted and must be addressed by any attempts to develop FT-IR spectroscopy as a microbial identification method.

ACKNOWLEDGEMENTS

We thank S. Moore, S. Hoover, D. McCallum, C. Hines, and T. Linkfield for their technical assistance. This work was supported by the National Institutes of Health, Division of Research Resources, under Grant Number RR-01367, establishing the National Center for Biomedical Infrared Spectroscopy.

REFERENCES

1 Anderberg, M.R. 1973. Cluster Analysis for Applications. pp. 71–73. Academic Press, New York.

2 Bergey's Manual of Systematic Bacteriology. 1984. (Holt, J.G., ed.), Williams and Wilkins, Baltimore.

3 Brahma, S.K., M.P. Baek, D. Gaskill, R.K. Force, W.H. Nelson and J. Sperry. 1985. The rapid identification of bacteria using time-resolved fluorescence and fluorescence excitation spectral methods. Appl. Spectrosc. 39: 869–872.

4 Dalterio, R.A., W.H. Nelson, D. Britt and J.F. Sperry. 1987. An ultraviolet (242 nm excitation) resonance Raman study of live bacteria and bacterial components. Appl. Spectrosc. 41: 417–422.

5 Giesbrecht, P., D. Naumann, H. Labischinski and G. Barnickel. 1985. A new method for the rapid identification and differentiation of pathogenic microorganisms using Fourier transform infrared spectroscopy. In: Rapid Methods and Automation in Microbiology and Immunology. (Habermehl, K.-O., ed.), pp. 198–206, Springer-Verlag, Heidelberg.

6 Hopkinson, J.H., J.E. Newbery, D.M. Spencer and J.F. Spencer. 1988. Differentiation between some industrially significant yeasts through the use of Fourier transform-infrared spectroscopy. Biotechnol. Appl. Biochem. 10: 118–123.

7 Howard, Jr. W.F., W.H. Nelson and J.F. Sperry. 1980. A resonance Raman method for the rapid detection and identification of bacteria in water. Appl. Spectrosc. 34: 72–75.

8 Kenner, B.A., J.W. Riddle, S.W. Rockwood and R.H. Bordner. 1958. Bacterial identification by infrared spectrophotometry II. Effect of instrumental and environmental variables. J. Bacteriol. 75: 16–20.

9 Lipkus, A.H., T.J. Lenk, K.K. Chittur and R.M. Gendreau. 1988. Cluster analysis of protein Fourier transform infrared spectra. Biopolymers. 27: 1831–1838.

10 Moss, C.W. 1981. Gas-liquid chromatography as an analytical tool in microbiology. J. Chromatogr. 203: 337–347.

11 Naumann, D., 1984. Some ultrastructural information on intact, living bacterial cells and related cell-wall fragments as given by FTIR. Infrared Phys. 24: 233–238.

12 Naumann, D. 1986. The ultra rapid differentiation and identification of pathogenic bacteria using FT-IR techniques. In: Proceedings of the 1985 Conference on Fourier and Computerized Infrared Spectroscopy. (Grasselli, J.G., D.G. Cameron, eds.), SPIE, pp. 268–269.

13 Naumann, D., V. Fijala, H. Labischinski. 1987. The differentiation and identification of pathogenic bacteria using FR-IR and multivariate statistical analysis. In: Book of Abstracts, 6th International Conference on Fourier Transform Spectroscopy. Vienna. p. 178.

14 Norris, K.P. 1959. Infra-red spectroscopy and its application to microbiology. J. Hyg. 54: 326–345.

15 Pau, C.-P., G. Patonay, C.W. Moss, G.M. Carlone, T.M. Rossi and I.M. Warner. 1986. A rapid enzymatic procedure for 'fingerprinting' bacteria by using pattern recognition of two-dimensional fluorescence data. Clin. Chem. 32: 987–991.

16 Stevenson, H.J.R. and O.E.A. Bolduan. 1952. Infrared spectrophotometry as a means for identification of bacteria. Science. 116: 111–113.

17 Wieten, G., J. Haverkamp, L.G. Berwald, D.G. Groothuis and P. Draper. 1984. Analytical pyrolysis in clinical and pharmaceutical microbiology. In: GC/MS Applications in Microbiology (Odham, G., L. Larsson and P.A. Mardh, eds.), pp. 335–380, Plenum, New York.