

Near-IR Imaging of Atheromas in Living Arterial Tissue

Lisa A. Cassis and Robert A. Lodder*

College of Pharmacy, University of Kentucky Medical Center, Lexington, Kentucky 40536-0082

A near-IR imaging system and parallel vector supercomputer are used with a fiber-optic probe to produce chemical maps of the intimal surface of living arteries. Spectrometric information collected at hundreds of near-IR wavelengths is assembled into color pictures of the lipoprotein and apolipoprotein composition of atheromas using a vectorized 3-D cellular automaton-based algorithm that operates in parallel. The nonparametric mathematics developed to identify and quantify the constituents of each voxel in the artery wall avoid the matrix factorizations that generate excess error in other pattern recognition methods and permit analysis in a wavelength space of over 1000 dimensions using fewer than 100 calibration samples. A surface feature resolution of 5.5 μm and depth resolution of 6.5 μm are achieved with the system. Data from the fiber optics confirm the injury hypothesis of lesion formation and the differing roles of HDL and LDL in cholesterol transport. In clinical studies, approximately $1/2$ of human arterial lesions appear fibrous and contain little or no lipid. As such, these lesions would not be expected to regress in response to cholesterol-lowering agents such as lovastatin. Identification of lesion types in vivo will enhance the efficacy of treatment programs.

INTRODUCTION

Heart disease is the primary cause of death in the United States.¹ Epidemiological studies performed over a period of years have indicated that reduction of blood cholesterol levels significantly reduces the risk of atherosclerosis, ischemia, myocardial infarction, and death. For some time these data have been cited in experimental attempts to prevent arterial disease. One recent study estimates that 60 million people, or 36% of all adults between the ages of 20 and 74, have cholesterol levels high enough to warrant medical advice and intervention.²

Cholesterol is carried in lipoprotein particles of different size, density, and lipid and apolipoprotein composition.³ Cholesterol and triglycerides in different lipoprotein particles take different metabolic pathways and have different effects on the arterial disease process: a high or low total cholesterol level does not necessarily indicate degree of risk for cardiovascular disease without knowledge of the type of lipoproteins contributing to total cholesterol. Lipoprotein α , for example,

has recently been associated with thrombotic risks⁴ and progressive atherosclerosis.⁵ Some lipoproteins are thought to deposit in arterial walls, while others are not.

The variety of lipoproteins and apolipoproteins poses an analytical problem, particularly in spectrometry, where similarity in structure leads to spectral interferences. About $2/3$ of the total cholesterol in plasma is carried in low-density lipoprotein (LDL) and intermediate-density lipoprotein (IDL) particles. The principal LDL apolipoprotein is apoB (95% of apolipoprotein content), but traces of at least seven other apoproteins can be found. The second-largest carrier of plasma cholesterol is high-density lipoprotein (HDL). The major HDL apolipoprotein is apoA-I (64% of apolipoprotein content), but significant amounts of A-II, C-I, C-III, ARP, and D can also be found in HDL particles. HDLs have been identified as particles removing cholesterol from tissues (and exhibiting a protective effect against arterial disease) by a process known as *reverse cholesterol transport*.³

The presence of whole blood in arteries interferes with fluorescence spectrometry more than it interferes with near-IR spectrometry (see Figure 1, which contains spectra of rat aorta measured with whole blood and with saline in the vessel). The near-IR imaging probe and IBM 3090-600J not only identify lesions and determine their size but also determine the chemical constituents of these lesions. Chemical analysis of lesions in vivo contributes to the understanding of lesion formation and growth. For example, apolipoprotein A (associated with HDL) and apolipoprotein B (associated with LDL) are known to occur in different sections of atherosclerotic lesions⁶ and are hypothesized to contribute differently to lesion formation. One experimental treatment program that focused on the hypothesized difference in action of HDLs and LDLs reported success in the reduction of recurring lesions in patients who had already undergone coronary bypass operations.⁷ A catheter with chemical analysis capabilities could monitor these changes in vivo.

Fiber-optic catheters have been used to locate atherosclerotic lesions, but current techniques merely distinguish lesions from healthy arterial tissue. A new method, based on near-IR fiber optics, has been developed to spatially map lesions and their chemical constituents. Chemical analysis of lesions in vivo permits the kinetic study of atherogenesis and contributes to the understanding of lesion formation and growth. As certain constituents are identified as playing key roles in the progression of disease, treatment programs will be designed that focus on these constituents.

Constituent identifications in atherosclerotic lesions have been performed using HPLC,⁸ NMR chemical-shift imaging,⁹

* Author to whom correspondence should be addressed.

(1) Bell, L. P.; Hectorne, K.; Reynolds, H.; Balm, T. K.; Hunninghake, D. B. *J. Am. Med. Assoc.* 1989, 261, 3419-3423.

(2) Sempos, C.; Fulwood, R.; Haines, C.; Carroll, M.; Anda, R.; Williamson, D. F.; Remington, P.; Cleeman, J. J. *Am. Med. Assoc.* 1989, 262, 45-52.

(3) Hoeg, J. M.; Gregg, R. E.; Brewer, B. J. *Am. Med. Assoc.* 1986, 255, 512-521.

(4) Miles, L. A.; Fless, G. M.; Levin, E. G.; Scanu, A. M.; Plow, E. F. *Nature* 1989, 339, 301-303.

(5) Hijjar, K. A.; Gavish, D.; Breslow, J. L.; Nachman, R. L. *Nature* 1989, 339, 303-305.

(6) Bocan, T. M. A.; Brown, S. A.; Guyton, J. R. *Arteriosclerosis* 1988, 8, 499-508.

(7) Roberts, L. *Science* 1987, 237, 28-29.

(8) Liburdy, R. P.; Burgess, N. L.; Bartley, J. C.; Mindich, B. P. *Atherosclerosis* 1988, 71, 103-112.

(9) Maynor, C. H.; Charles, H. C.; Herfkens, R. J.; Suddarth, S. A.; Johnson, G. A. *Invest. Radiol.* 1989, 24, 52-60.

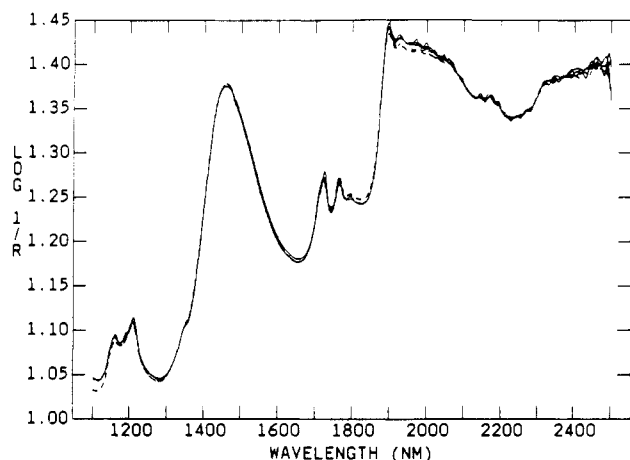


Figure 1. Near-IR spectra of the intima of the thoracic aorta, obtained with a fiber-optic probe. The solid lines depict three spectra obtained with whole blood in the vessel, and the dashed lines represent the spectra obtained with the whole blood removed and replaced with saline.

and immunocytochemistry.^{10,11} Atheroma recognition has been performed by duplex ultrasound and pneumoplethysmography,¹² atheroma classification by microscopic and histologic examinations,^{13,14} and spatial profiling by chemical analysis,¹⁴ enzymatic assays,¹⁴ and immunofluorescence techniques.¹⁵

New in vivo technologies used for removing atherosclerotic plaques require a reliable method of in vivo plaque recognition. The use of fiber-optic probes with fluorescence spectroscopy for imaging the intimal surfaces of arteries^{16,17} and guidance during laser ablation of atheromas^{18,19} has been reported. Reference 17 reports a limited usefulness of fluorescence emission for guidance of laser angioplasty because of the complexity of the emission from the broad range of atherosclerotic plaques encountered in diseased coronary arteries. Fatty, fibrous, and calcific plaque content as well as maximal intimal thickness were measured and correlated in ref 17 with fluorescence intensity ratios from emission spectra. The level of correlation was found to be dependent on the complexity of the lesion. It was suggested that plaque characteristics can be assessed by fluorescence emission but the successful implementation of spectroscopic guidance is dependent on the level of prediction error, which varies with tissue type.

A more appropriate choice of wavelengths (near-infrared wavelengths instead of visible) yields more chemical information about the lesions. Furthermore, the large range of plaque compositions is a problem that can be circumvented by using multiple calibrations (one calibration set for each

tissue type). The parallel vector algorithm currently used in our laboratory and developed specifically for spectroscopic problems of this type identifies lesions that cannot be determined adequately using a particular calibration curve²⁰ and recalibrates automatically for these lesions.

Virtually every organic compound has a near-IR spectrum that can be measured. Near-IR spectra consist of overtones and combinations of fundamental mid-IR bands, giving near-IR spectra a powerful ability to identify organic compounds while still permitting penetration of light into tissues. Quantitative determination of the components of arterial lesions can be accomplished down to the noise limit using modern mathematics and parallel pattern-recognition techniques. Furthermore, qualitative analysis, in which the constituents of arterial lesions and the vessel matrix are identified, is also possible (with benefits in analytical speed and accuracy accruing to those using near-IR spectrometry and mathematics to analyze arterial lesions). New results obtained in three specific areas demonstrate the imaging power obtained by combining the massive observation of near-IR wavelengths with parallel vector processing:

1. A near-infrared fiber-optic probe has been constructed and used for chemical analysis of arterial lesions in living tissue. The ability of the probe and supercomputer to quantify and map the positions of HDL and LDL and apolipoproteins A-I, A-II, and B at clinical concentrations has been demonstrated.

2. The transport of cholesterol by LDL to arterial walls has been observed by near-IR imaging, which has indicated steady accumulation of LDL by arterial endothelium and smooth muscle cells. LDL accumulation by endothelial-denuded vessels was greater than LDL accumulation by intact vessels, providing evidence supporting the injury hypothesis of lesion formation. HDL was not taken up by the arterial walls.

3. Oxidation of LDL has been observed in near-IR spectra, which have shown a loss of characteristic peaks for apolipoprotein B, a major constituent of LDL. Antioxidant drugs are hypothesized to inhibit the growth of atherosclerotic lesions by slowing the accumulation of LDL by vessel walls.

The results obtained demonstrate that the near-IR probe and algorithm differentiate arterial lesions from surrounding tissue on the basis of the constituents of the lesions and that the algorithm can form an image of the distribution of lipoproteins and apolipoproteins in arterial walls. The chemical imaging power of the near-IR/supercomputer technique permits the testing of important new hypotheses of lesion formation, growth, and regression.

THEORY

Near-IR spectra are used in this work both qualitatively and quantitatively. Spectra recorded at d wavelengths are represented as single points in a d -dimensional hyperspace. For simplicity, these spectra are represented as elements of a matrix T , with one row for each of n samples and one column for each of d wavelengths. Each point is translated from the origin in hyperspace in each dimension by an amount that corresponds to the magnitude of the signal observed at each wavelength. In this scheme, similar samples produce similar spectra that project as "probability orbitals" or "clusters" into similar regions of hyperspace.²¹

The Quantile Bootstrap Error-adjusted Single-sample Theory (BEST) is a new experimental clustering technique for exploring distributions of lesion spectra in hyperspace. In

(10) Hansson, G. K.; Jonasson, L.; Lojstved, B.; Stemme, S.; Kocher, O.; Gabbiani, G. *Atherosclerosis* 1988, 72, 135-141.

(11) Lamaziere, J. D.; Desmouliere, A.; Pascal, M.; Larrue, J. *Atherosclerosis* 1988, 74, 115-126.

(12) Castaldo, J. E.; Nicholas, G. G.; Gee, W.; Reed, J. F. *Arch. Neurol.* 1989, 46, 518-522.

(13) Bouissou, H.; De Graeve, J.; Pieraggi, M. Th.; Julian, M.; Thiers, J. C. *Pharmacol. Res. Commun.* 1981, 13, 241-249.

(14) Bassiouny, H. S.; Davis, H.; Massawa, N.; Gewertz, B. L.; Glagov, S.; Zarins, C. K. *J. Vasc. Surg.* 1989, 9, 202-212.

(15) Bedossa, P.; Poynard, T.; Abella, A.; Paraf, F.; Lemaigre, G.; Martin, E. *Arch. Pathol. Lab. Med.* 1989, 113, 777-780.

(16) Satori, M.; Weillbaecher, D.; Valserrama, G. L.; Kubodera, S.; Chin, R. C.; Berry, M. J.; Tittel, F. K.; Sauerbrey, R.; Henry, P. D. *Circ. Res.* 1988, 63, 1053-1059.

(17) Hoyt, C. C.; Richards-Kortum, R. R.; Costello, B.; Sacks, B. A.; Kittrell, C.; Ratliff, N. B.; Kramer, J. R.; Feld, M. S. *Lasers Surg. Med.* 1988, 8, 1-9.

(18) Leon, M. B.; Lu, D. Y.; Prevosti, L. G.; Macy, W. M.; Smith, P. D.; Granovsky, M.; Bonner, R. F.; Balaban, R. S. *J. Am. Coll. Radiol.* 1988, 12, 94-102.

(19) Gaffney, E. J.; Clarke, R. H.; Lucas, A. R.; Isner, J. M. *Lasers Surg. Med.* 1989, 9, 215-228.

(20) Lodder, R. A.; Hieftje, G. M. *Appl. Spectrosc.* 1988, 42, 1351-1365.

(21) Lodder, R. A.; Hieftje, G. M. *Appl. Spectrosc.* 1988, 42, 1512-1520.

effect, the BEST calculates the integral of a probability orbital in hyperspace by starting at the centroid of the orbital and working outward in all directions at a uniform rate. The distance between the center of a lesion orbital and a sample spectrum is proportional to the concentration(s) of the lesion constituent(s) responsible for the vector connecting the central and sample spectral points. The direction of the vector identifies the constituent(s) of the lesion. The BEST makes no assumptions about the spectral coordinate system in use or about the distribution of lesion spectral points in hyperspace, and thus the BEST does not suffer from the accuracy and precision problems associated with conventional quantitative and qualitative techniques, which rely on standard matrix algebra and parametric statistics. The BEST therefore offers superior performance as an assimilation method (a method that progressively increases its analytical performance by incorporating previously unknown samples into its calibration). The calibration samples are analyzed by another reference method (such as those listed earlier) in the same manner that Beer's law is used to develop a conventional spectrophotometric calibration.

Spectral analysis with the BEST starts with the collection of a calibration set, or "training set", of samples. The training set samples are chosen to represent as completely as possible all of the expected variations in the sample population. A bootstrap distribution of the training set is created from the training-set spectra by an operation $\kappa(\mathbf{T})$. This bootstrap distribution forms the basis for calculating directional probabilities and is calculated only once for each training set. $\kappa(\mathbf{T})$ provides the bootstrap distribution \mathbf{B} for the training set as well as the mean \mathbf{C} of the bootstrap distribution. The operation $\psi(\mathbf{T}, \mathbf{B}, \mathbf{X}, \mathbf{C})$ calculates σ_c (the BEST distance, in units of multidimensional standard deviations, or SDs) using the training set \mathbf{T} , the bootstrap distribution \mathbf{B} and center \mathbf{C} , and the test sample's spectrum \mathbf{X} . The Euclidean distance from \mathbf{C} to \mathbf{X} is scaled by σ_c to give the distance to \mathbf{X} in BEST SDs.

To calculate $\kappa(\mathbf{T})$, random selections are made from \mathbf{T} by filling \mathbf{P} with the training-set sample numbers to be used in the bootstrap sample sets $\mathbf{B}_{(s)}$,

$$\mathbf{P} = p_{ij} = r \quad (1)$$

and the values in \mathbf{P} are scaled to the training-set size n :

$$\mathbf{P} = n\mathbf{P} + 1 \quad (2)$$

A bootstrap sample $\mathbf{B}_{(s)}$ creates each row i of \mathbf{B} by setting

$$\mathbf{B}_{(s)} = t_{kj} \quad (3)$$

where k are the elements of the i th rows of \mathbf{P} . The q th row of \mathbf{B} becomes the mean of the bootstrap sample of lesion spectra

$$b_{qj} = \sum_{i=1}^n b_{(s)i,j} / n \quad (4)$$

and the mean of the bootstrap distribution is

$$c_j = \sum_{i=1}^m b_{ij} / m \quad (5)$$

$\kappa(\mathbf{T})$ is completed only once for each calibration and is used repeatedly in the routine analysis of actual lesions.

The BEST finds the hyperline connecting \mathbf{C} and \mathbf{X} and determines the probability of \mathbf{X} belonging to the same population as \mathbf{T} using the number of points (rows of \mathbf{B}) within a certain distance r_h of the hyperline (in effect, taking the points of \mathbf{B} that appear within a hypercylinder of radius r_h). The BEST estimates the point density of the hypercylinder by constructing planes connecting \mathbf{X} , \mathbf{C} , and the rows of \mathbf{B} .

The use of a series of planes allows a complex structure like a hypercylinder in d -dimensional hyperspace to be represented in a simple manner on parallel vector processors such as those in the IBM 3090-600J, regardless of the number of spatial dimensions (the number of planes is independent of the spatial dimension). Monte Carlo integration of the bootstrap distribution to give \mathbf{B} can be divided by rows independently in all equations among as many processors as are available. The use of planes in hyperspace to represent the hypercylinder makes most of the resulting calculations summations and summations of squares, an ideal situation for vector processing.

The three points that determine a spectral plane in hyperspace also specify a triangle whose sides are readily determined:

$$S_{(02)} = \left(\sum_{j=1}^d (x_j - c_j)^2 \right)^{1/2} \quad (6)$$

$$s_{(0R)i} = \left(\sum_{j=1}^d (b_{ij} - c_j)^2 \right)^{1/2} \quad (7)$$

$$s_{(2R)i} = \left(\sum_{j=1}^d (b_{ij} - x_j)^2 \right)^{1/2} \quad (8)$$

The formation of the triangle ensemble facilitates the determination of the points in \mathbf{B} that fall inside a hypercylinder of radius r_h .

$$s_{(UB)i} = (S_{(02)} + s_{(0R)i} + s_{(2R)i}) / 2 \quad (9)$$

$$a_i = (s_{(UB)i}(s_{(UB)i} - S_{(02)})(s_{(UB)i} - s_{(0R)i})(s_{(UB)i} - s_{(2R)i}))^{1/2} \quad (10)$$

$$s_{(R)i} = 2(a_i) / S_{(02)} \quad (11)$$

$$s_{(p)i} = (s_{(0R)i}^2 - s_{(R)i}^2)^{1/2} \quad (12)$$

Points in the bootstrap distribution on the same side of the plane as \mathbf{X} are assigned positive distances in $\mathbf{S}_{(p)}$ by multiplying the elements of $s_{(p)i}$ for which $\{S_{(02)}^2 + s_{(0R)i}^2 < s_{(2R)i}^2\}$ by -1 . The values of $\mathbf{S}_{(p)}$ representing points in \mathbf{B} that are outside the hypercylinder are discarded for the remainder of the calculations and n_h replicate spectra are retained:

$$n_h = m(\% \text{ of } \mathbf{B} \text{ spectra desired in hypercylinder}) / 100 \quad (13)$$

$$r_h = O(\mathbf{S}_{(R)})_{n_h} \quad (14)$$

$$\mathbf{S}_{(q)} = O(\{s_{(p)i} | s_{(R)i} < r_h\}) \quad (15)$$

To create a symmetric 1 SD contour on \mathbf{T} , $l = 0.16n_h$, $u = 0.84n_h$, and the unskewed confidence interval along the hyperline connecting \mathbf{X} and \mathbf{C} is $\{s_{(q)l} < C < s_{(q)u}\}$. The symmetric SD is not corrected for skew in the sample population, and this uncorrected σ is either calculated by

$$\frac{(|s_{(q)l}| + |s_{(q)u}|)n^{1/2}}{2} \quad (16)$$

or by calculating the standard deviation of $\mathbf{S}_{(q)}$ and multiplying it by $n^{1/2}$. The distance to the test spectrum in uncorrected BEST SDs is

$$\left(\sum_{j=1}^d (c_j - x_j)^2 \right)^{1/2} / \sigma \quad (17)$$

Of course, most biological spectral populations are skewed, especially those composed of relatively complex samples like spectra of lesions. The spectral skew of these populations arises from the nonrandom distribution of chemical compounds present in the samples and from the fact that Beer's law represents a linear transformation of sample population

information from a chemical to a spectral coordinate system. In these cases l and u must be corrected to compensate for the skew before finding the confidence limits $s_{(q)l}$ and $s_{(q)u}$. For $m \leq 1000$ the contour should be 1, so $\alpha = \Phi(-1)$. Setting $z_\alpha = \Phi^{-1}(\alpha)$ and locating the center of **T** by $c_{(T)j} = M_j(t_{ij})$ initializes the adjustment of confidence limits for skew.

$c_{(T)}$ lies in space in the direction opposite to the direction of the skew (with respect to **C**) because of the leverage effect of skewed points on the mean. The detection of leverage effects serves as the basis of the confidence-limit adjustment.

$$S_{(C0R)} = \left(\sum_{j=1}^d (c_{(T)j} - c_j)^2 \right)^{1/2} \quad (18)$$

$$S_{(C2R)} = \left(\sum_{j=1}^d (c_{(T)j} - x_j)^2 \right)^{1/2} \quad (19)$$

$$S_{(CUB)} = (S_{(02)} + S_{(C0R)} + S_{(C2R)})/2 \quad (20)$$

$$A_c = (S_{(CUB)}(S_{(CUB)} - S_{(02)})(S_{(CUB)} - S_{(C0R)})(S_{(CUB)} - S_{(C2R)}))^{1/2} \quad (21)$$

$$S_{(CR)} = 2(A_c)/S_{(02)} \quad (22)$$

$$S_{(CP)} = (S_{(C0R)}^2 - S_{(CR)}^2)^{1/2} \quad (23)$$

If $\{S_{(02)}^2 + S_{(C0R)}^2 > S_{(C2R)}^2\}$ then S is multiplied by -1 .

The $S_{(CP)}$ term is modified to $S_{(CP)}\delta + M(S_{(q)})$ to make $S_{(CP)}$ perform well in an environment where almost any axis scale or skew is possible. The addition of $M(S_{(q)})$ assures that the correction $S_{(CP)}$ and its array-analog z_o (defined below) have the same sign (direction) when $S_{(q)}$ is slightly skewed, and δ provides a skew-sensitivity adjustment. δ has a value between 0 and 1 that is set empirically for each combination of **T** and **B** to keep the absolute magnitude of the adjustment inside the actual values of $S_{(q)}$.

The z_o skew adjustment factor is calculated from $S_{(CP)}$ by

$$f_i = s_{(q)i} - S_{(CP)} \quad (24)$$

$$I_{(N)} = \{1, 2, 3, \dots, n_h\} \quad (25)$$

$$z_e = R(F(I_{(N)})) \quad (26)$$

$$z_o = \Phi^{-1}(z_e/n_h) \quad (27)$$

If $|2z_o| > |z_\alpha|$, then δ is decreased and the calculation resumed at eq 18. Otherwise, new l and u values are calculated for $S_{(q)}$:

$$l = \Phi(2z_o + z_\alpha)n_h \quad (28)$$

$$u = \Phi(2z_o - z_\alpha)n_h \quad (29)$$

The confidence interval along the line connecting **X** and **C** continues to be $\{s_{(q)l} < C < s_{(q)u}\}$. The upper confidence limit is always the one closest to the test spectrum **X**. Accordingly, σ_c is simply $s_{(q)u}$, and the distance in adjusted BEST SDs from the lesion training set **T** to the test spectrum **X** is

$$\left(\sum_{j=1}^d (c_j - x_j)^2 \right)^{1/2} / ((\sigma_c/|z_\alpha|)n^{1/2}) \quad (30)$$

The use of this equation in both *quantitative* and *qualitative analysis* of lesions provides a number of advantages over all other methods of analysis:

1. No analytical assumptions are required to make the problem solvable. Other chemometric methods assume that no discriminating variable (wavelength) is a linear combination of other discriminating variables, that the covariance matrices for all spectral groups are approximately equal, and that each group is drawn from a population that is normally distributed on the discriminating variables. None of these

assumptions is usually true, and violations of these assumptions increase the likelihood of producing incorrect quantitative and qualitative analytical results.

2. This nonparametric assimilation method can be used with *full spectra* of samples, which often include a thousand or more variables that describe each sample. The large virtual memory on the newest supercomputers has made possible the manipulation of images involving tens of thousands of spectra at thousands of wavelengths. No wavelength selection procedures or data compression techniques, such as principal-axis transformation or Fourier transformation, are required by this assimilation method in order to analyze complete spectra. Collinearity in the discriminating variables (wavelengths) does not degrade the results. (Collinearity disturbs conventional matrix techniques such as the Mahalanobis method, which relies on matrix inversion or factorization to produce a distance in SDs.)

3. The numerator of the final equation in the assimilation method provides a vector that identifies the sample components. The metric is calculated by a *highly vectorized code* that can be *distributed across as many processors as are available*. The numerator is used in computerized searches of spectral libraries for qualitative analysis of lesion samples.

4. Evaluation of the final equation of the assimilation method produces a distance value that is proportional to the concentrations of the lesion constituents. Thus, quantitative analytical capabilities are also provided by the same assimilation method, which can still be distributed across as many vector processors as are available.

5. The BEST metric is not only more accurate and precise than the Mahalanobis metric (the metric commonly used in near-IR spectrometry), the BEST metric is often calculated more rapidly as well. The matrix inversion required by the Mahalanobis metric is usually accomplished by algorithms whose complexity (in terms of number of operations required) increases as the number of wavelengths cubed. In contrast, *the complexity of the BEST metric increases linearly with the number of wavelengths*. The "overhead" in setting up the BEST is slightly greater than for the Mahalanobis; however, by the time the spectra contain 60 or more wavelengths, the BEST metric is calculated more rapidly.²² Full near-IR spectra with continuous coverage from 1100 to 2500 nm use at least 141 wavelengths on a typical commercial spectrometer. The Mahalanobis metric shows its instability through widely varying bias and RSD values. The BEST metric has a uniformly lower bias and RSD than the Mahalanobis metric. The superior bias and RSD of the BEST becomes even more notable in full spectral analysis involving hundreds of wavelengths.

EXPERIMENTAL SECTION

Instrumentation. The near-IR spectrometer and fiber-optic probe were similar to those employed previously in serial imaging experiments (experiments in which each image was formed by collecting spectra from one location at a time, assembling the image by computer after all spectra were collected).^{23,24} The images were calculated on an IBM 3090-600J parallel vector supercomputer and displayed on an IBM RS 6000 workstation. Tissue samples were maintained in Krebs physiological salt solution [containing (in mM) 130 NaCl, 47 KCl, 1.18 KH₂PO₄, 1.17 MgSO₄, 14.9 NaHCO₃, 5.5 dextrose, 0.026 Na₂EDTA, and 0.11 L-ascorbic acid] in a 37 °C Dubnoff metabolic shaker

(22) Drennen, J. K.; Lodder, R. A. *Spectroscopy* 1991, 6 (8), 34-39.

(23) Cassis, L. A.; Lodder, R. A. Near-IR Imaging of Atheromas in Living Arterial Tissue. In *Scientific Excellence in Supercomputing*; Billingsley, K. R., Brown, H. U., Derohanes, E., Eds.; Baldwin Press: Athens, GA, 1992.

(24) Lodder, R. A.; Cassis, L.; Ciurczak, E. W. *Spectroscopy* 1990, 5 (7), 12-16.

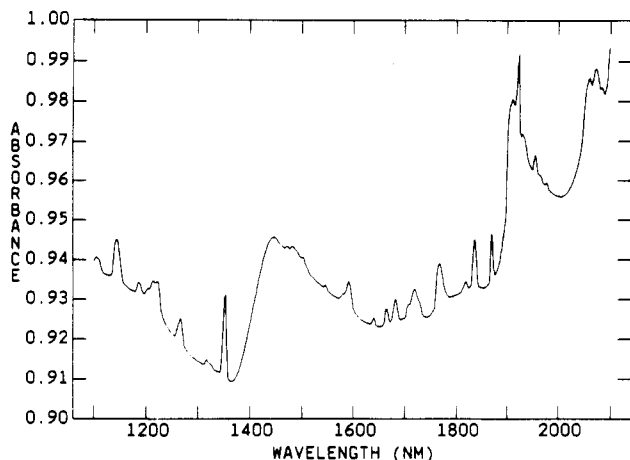


Figure 2. Near-IR spectrum of pure human apoA-I.

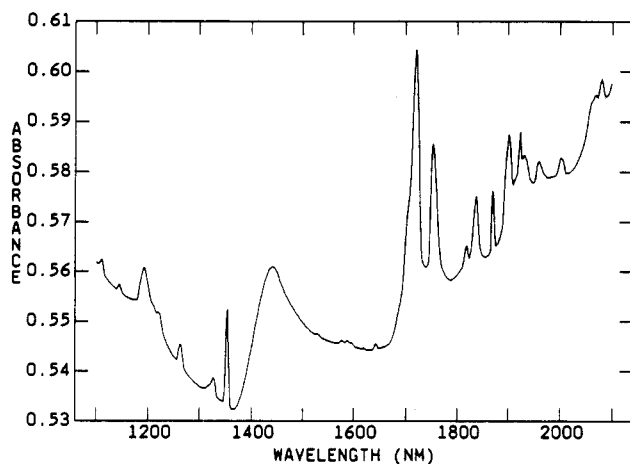


Figure 3. Near-IR spectrum of pure human apoB.

oxygenated with 95% O₂/5% CO₂ during the incubation periods. Tissue samples were removed from the bath during collection of spectra. A portion of the Krebs solution was removed from the bath each time the tissues were scanned, and the tissue was immersed in this solution during scanning. A micropositioning stage was used to position the tissue beneath the fiber-optic probe during scanning.

Spectra of pure compounds were obtained from lyophilized standards or from dialyzed solutions dried at 55 °C under reduced pressure.

Materials. Samples of HDL (purity confirmed by cellulose acetate electrophoresis with oil red O staining), LDL (purity confirmed by cellulose acetate electrophoresis with oil red O staining), apoB (97% purity by immunoelectrophoresis), apoA-I (97% purity by SDS-PAGE and immunoelectrophoresis, MW 28 300), and apoA-II (97% purity by SDS-PAGE and immunoelectrophoresis, MW 17 380) were obtained from Sigma Chemical Co. (St. Louis, MO).

RESULTS AND DISCUSSION

The roles of HDL, LDL, and apolipoproteins A-I, A-II, and B in atherosclerosis are more easily followed with near-IR fiber optics than by virtually any other existing method. HDL, LDL, and the apolipoproteins can be determined simultaneously with near-IR spectrometry in less than 1 min, making spatial maps of their distribution in arterial walls readily accessible by moving a catheter through the artery.

Successful spectroscopic identification and determination of HDL and LDL has been demonstrated using near-IR fiber optics and supercomputer algorithms. Constituents of HDL and LDL, including apolipoproteins A-I (see Figure 2), A-II, and B (see Figure 3), have also been identified using the

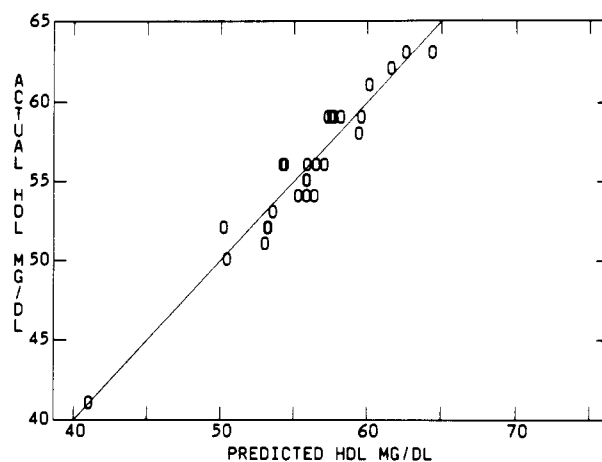


Figure 4. Calibration line for human serum HDL calculated using clinical serum samples, with validation samples shown superimposed on the line.

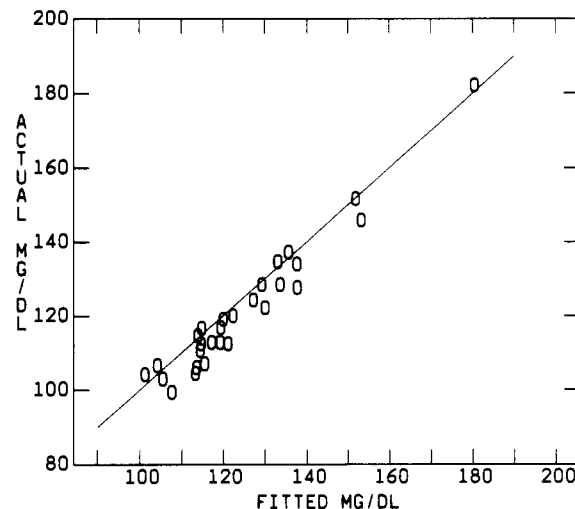


Figure 5. Calibration line for human serum LDL calculated using clinical serum samples, with validation samples shown superimposed on the line.

nondestructive near-IR fiber-optic spectrometric technique. Calibration results for human clinical HDL and LDL samples are shown in Figures 4 and 5. The standard error of prediction (SEP, calculated from cross-validation samples) for HDL in Figure 4 is only 2 mg/dL, while the SEP for LDL in Figure 5 is only 5 mg/dL.

Figure 2 shows the spectrum of solid human apolipoprotein A-I. The apolipoprotein concentration in the optical volume of the probe was ~0.5 mg/mL. The distinctive regions of this scan appear between 1900 and 2000 nm and between 2100 and 2200 nm. These regions enable the near-IR spectroscopic discrimination of apolipoprotein A-I and apolipoprotein B, which has major peaks in the 1700–1800-nm region (see Figure 3).

The spectra of the apolipoproteins and HDL and LDL can be obtained, combined into a set, and projected into hyper-space simultaneously. Figure 6 shows 10 different spectra each of apolipoprotein B, A-I, and A-II as well as LDL and HDL (giving a total of 50 scans in Figure 6). The groups of 10 scans appear as points in Figure 6 because the variability within the groups is much smaller than the differences between the groups. (The apoB group is the largest because its spectra showed the greatest variability.) Figure 6 shows that it is relatively easy to differentiate between samples of apolipoproteins and lipoproteins.

Recently a low-resolution false-color near-IR "map" of a normal human coronary artery was published along with a

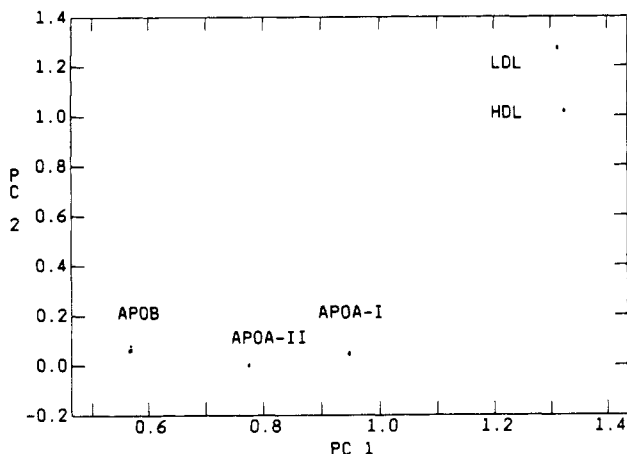


Figure 6. Fifty spectra of human serum lipoproteins and apolipoproteins, projected on a plane in hyperspace. Ten spectra are in each cluster of points. The vast differences between the spectra make the clusters appear as single points on the principal component plot.

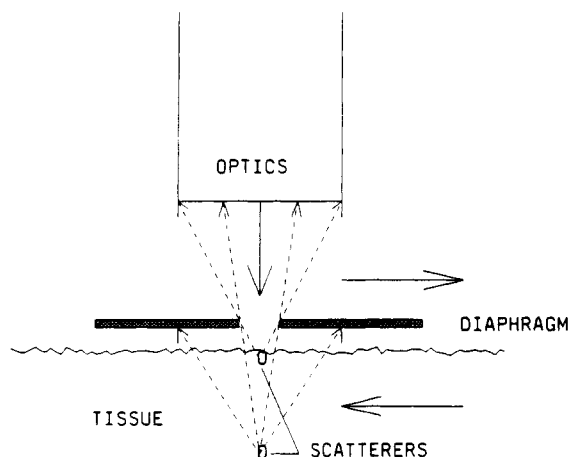


Figure 7. Detector with variable diaphragm. Controlling the solid angle and depth of field of the optics permits depth profiling of the chemical constituents of the artery wall.

map of a portion of a lipid-filled arterial lesion from a nearby vessel.²³ The ability to spatially map the distribution of apolipoproteins with a fiber-optic probe was demonstrated by comparing the maps with photomicrographs of the same area of tissue. The scans were collected over a 29×32 pixel region, and the image was formed after mapping the spectra into hyperspace and measuring interorbital distances with the BEST matrix.²³ Image processing algorithms²³ were used to produce a pixel resolution of $\sim 75 \mu\text{m}^2$. The image resolution could be improved by obtaining more spectra (more pixels).

The best resolution obtainable with the fiber-optic method in the plane (x, y) of the artery wall is $5.5 \mu\text{m}$ (with 1100-nm radiation). The resolution figure was obtained by placing a small water droplet on a glass slide at the edge of the field of view of the "camera" and moving the droplet and slide over $50 \mu\text{m}$ in $10\text{-}\mu\text{m}$ increments into the center of the field with a micropositioning stage. Regression of the spectra so obtained vs distance gave an SEE = $5.5 \mu\text{m}$ on the position of the droplet.

Controlling the solid angle of scattered light reaching the detector with a variable diaphragm provides $6.5\text{-}\mu\text{m}$ depth (z axis) resolution (see Figure 7). To use this probe, multiple passes are made over a surface, increasing the solid angle of reflected light with each pass. On the first pass, a small aperture lets only light from surface scatterers back into the detector. On later passes with larger apertures, light from the surface and subsurface layers reaches the detector.

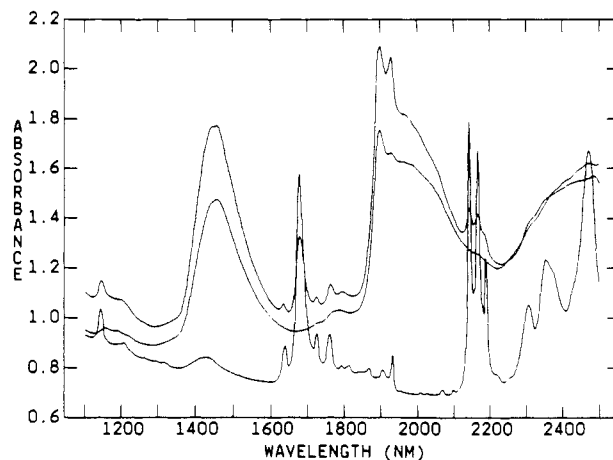


Figure 8. Spectra demonstrating the penetration ability of near-IR light. The lowest trace was generated by scanning a small polystyrene block. The center trace was generated by scanning through an aorta from outside the vessel. The upper trace was generated by placing the polystyrene block outside the aorta opposite the fiber-optic probe. Polystyrene peaks are clearly visible through the aorta.

Because the surface composition has already been calculated, the subsurface layers can be estimated by difference. A depth resolution test was performed with this probe by moving a polystyrene wafer $0\text{--}50 \mu\text{m}$ from the fiber-optic aperture. Using 1100-nm light, the regression of distance vs spectra gave a SD = $6.5 \mu\text{m}$ in the position of the polystyrene.

The intensity and solid angle of near-IR light can be limited to produce images of arterial endothelium. Equally important, this intensity can also be increased to permit light to penetrate completely through a blood vessel. Figure 8 shows (from bottom to top) spectra of a polystyrene wafer, the thoracic aorta of a rat, and polystyrene as it appears through the intervening aorta. In this experiment, a polystyrene wafer was placed behind an intact aorta, and an aluminum mirror was placed behind the polystyrene to reflect the plastic spectrum back toward the fiber-optic probe. The mirror was necessary because the polystyrene was not powdered and hence would transmit (but not appreciably scatter) the light. The spectrum of polystyrene was recorded after the light had passed through both walls of the intact aorta, through the polystyrene to the mirror, back through the polystyrene, and back through both walls of the aorta to the probe (a distance of over $4000 \mu\text{m}$). Even the largest fibrous plaques of interest in humans are seldom more than $2000 \mu\text{m}$ thick, or more than $1\text{--}2 \text{ cm}^2$ in area. The penetration of near-IR light through tissue opens up the possibility of imaging lesions from outside the artery, using a permanent implant located near a new lesion with the help of a fiber-optic catheter.

HDL and LDL are quantified spectrometrically by their distinctive apolipoprotein "tags", apoA-I and apoB. A fiber-optic reflectance probe was used to quantify lyophilized apoA-I and apoB in a human serum albumin matrix. The SEE = 1.8 mg/dL ($r^2 = 0.98$) and SEP = 1.9 mg/dL for apoA-I, and the SEE = 2.0 mg/dL and SEP = 2.1 mg/dL for apoB. Because the location and shape of near-IR peaks are always somewhat dependent on the analyte matrix, multidimensional BEST "Beer's law" calibrations for HDL and LDL in rat thoracic aorta were obtained by homogenizing the aortas from several rats with known amounts of HDL and LDL using a Polytron. (No accurate reference assay for HDL and LDL immobilized in arterial walls is available.) The SEE = 4.7 mg/dL for HDL and 7.1 mg/dL for LDL.

BEST reflectance-imaging results for actual aortas appear in Figures 9 and 10. The diagram in Figure 11 shows how the displacements of spectral points in hyperspace from the center of the spectra of normal artery are coded into colors and

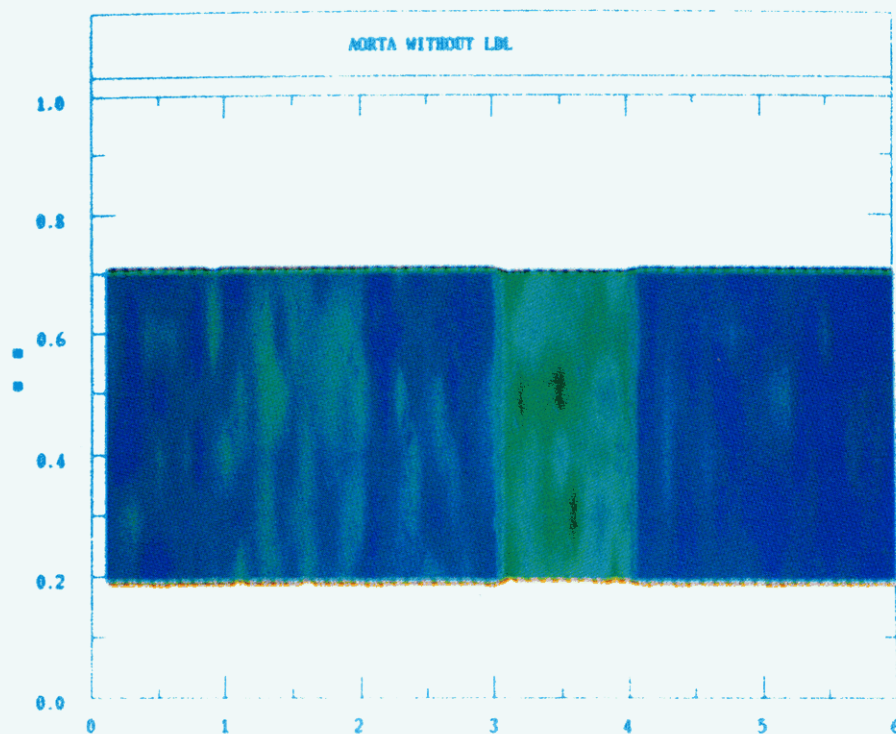


Figure 9. Near-IR spectral data collected with a fiber-optic probe to produce BEST image of an aorta showing no LDL accumulation. Axis dimensions are given in millimeters.

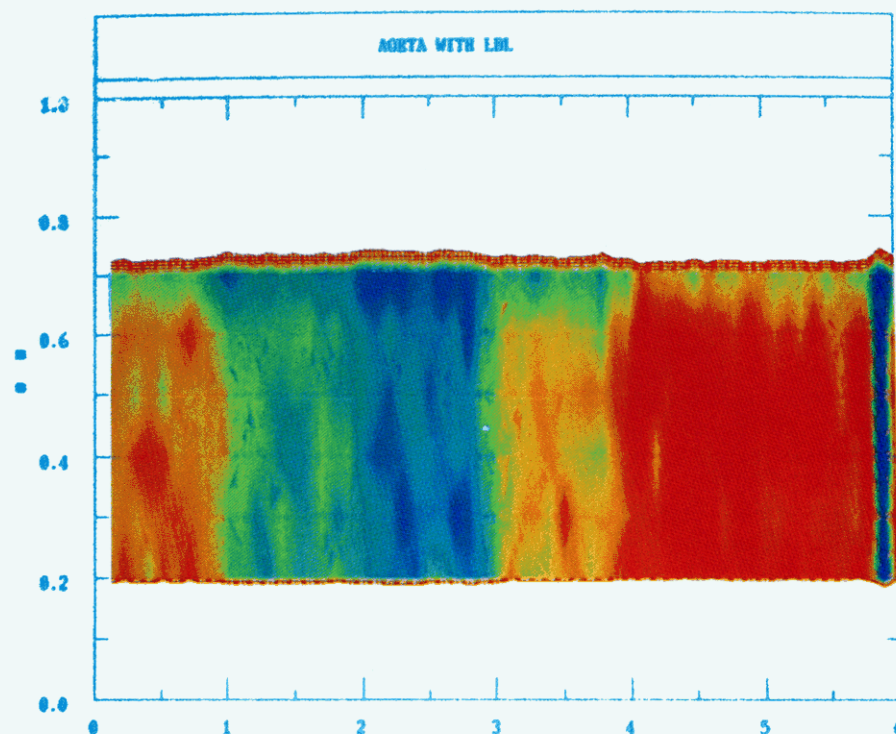


Figure 10. Near-IR spectral data collected with a fiber-optic probe to produce BEST image of an aorta showing substantial LDL accumulation. Axis dimensions are given in millimeters.

probability-density contours. The Os represent a sample of spectra taken from normal arteries. Displacements of the Os from the center of the cluster are due to minor differences in the composition of normal arteries, variations in temperature and the thickness of the arterial wall, and instrumental noise. Changes in the thickness of the arterial wall result in a baseline shift in the near-IR spectra that is coded green. The region in hyperspace immediately around the center of

normal spectra is coded dark blue, while the area slightly farther away is coded light blue.

Image reconstruction begins with the BEST comparing the spectrum obtained at each pixel in the image with the spectra obtained from normal artery walls. A displacement vector (with a direction and distance) from the center of normal artery spectra is calculated with a length given in multidimensional standard deviations (SDs), and the density

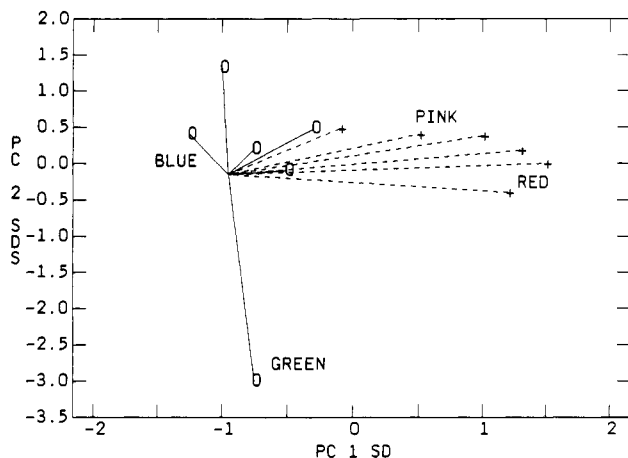


Figure 11. Principal components plot showing how colors are used to represent directions in hyperspace. Changes in the direction of spectral-point displacement denote changes in sample composition.

contours that appear in the arterial images are drawn in 0.1 SD increments. The pluses in Figure 11 represent spectra from a thoracic aorta incubated for 2 h at 37 °C in Krebs physiological salt solution and LDL (0.5 mg/mL). As the amount of LDL taken up by the artery increases, the displacement of the pluses from the centers of the normal spectra also increases. The LDL displacement is nearly orthogonal to the displacement caused by the thickness of the vessel wall. In the images of the arterial wall (Figure 10), small spectral displacements in the direction of LDL are coded pink, while larger displacements are coded red. The BEST metric acts as a "rubber yardstick" whose length can be different in every direction of hyperspace and whose stretch in SDs can be *asymmetric* (different in opposite directions). The BEST metric (or yardstick) is large in the green direction and small in the red direction, and the distance to the green spectral point is less than 3 SDs, while the distance to the farthest red spectral points is almost 20 SDs (even though the Euclidean distances are almost identical).

Figure 9 is the image of a segment of thoracic aorta obtained from a rat after incubation for 2 h in Krebs solution at 37 °C. The fiber-optic probe was moved along the arteries with a micropositioning stage to collect the spectra used to calculate each image. The arterial endothelium was denuded mechanically using a blunt probe before incubation. The green area in the image indicates a thicker portion of the vessel wall. The image is mostly blue as expected, and the contours are drawn at approximately 0.1 SD increments.

Figure 10 shows another segment of endothelium-denuded thoracic aorta (from the same rat as in Figure 9) that was incubated for 2 h in Krebs solution containing LDL (0.5 mg/mL). Accumulation of LDL is shown in shades of pink and red. The accumulation appears greater at the ends of the artery, probably because of the cutting and handling of the artery more at the ends (leading to more injury and consequently more LDL accumulation).

Experiments using the fiber-optic probe and BEST as an imaging spectrometer have verified the ability of the system to identify and quantify apolipoproteins, HDL, LDL, and cholesterol. At this point, the system was used to test a hypothesis with living tissue. It was hypothesized that an injured vessel wall would take up LDL, but not HDL. In this study, the fiber-optic probe was maintained at a single position on four segments of thoracic aorta. Two thoracic aortas were obtained from two rats and were cut in half to provide four segments. From each rat, one aortic segment was incubated in oxygenated Krebs solution at 37 °C and served as a control, while the other segment was incubated in Krebs solution

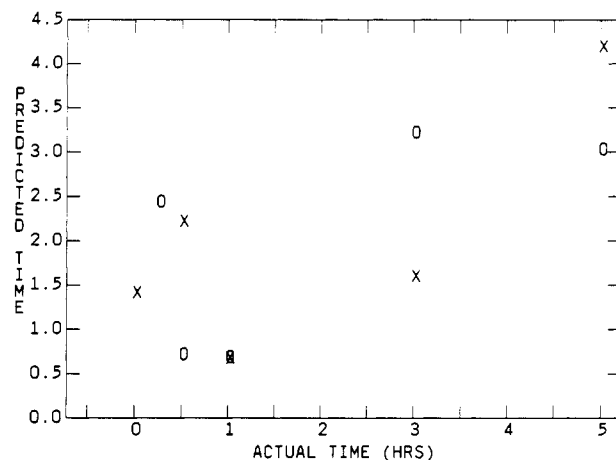


Figure 12. Results of the BEST prediction for control arteries. The supercomputer algorithm is unable to predict the incubation time of control arteries because the spectra of the control arteries do not change with time.

containing either HDL ($n = 3$ rats) or LDL ($n = 3$ rats) (0.5 mg/mL). Spectra were obtained from all of the segments immediately after addition of the lipoproteins, as well as at 15 min, 30 min, 1 h, 3 h, and 5 h afterward. Because the near-IR spectra of the lipoproteins can change somewhat with the environment (the environment of LDL changes, for example, when it moves from the Krebs to a receptor and through the cell membrane), the BEST was used to predict the amount of time each vessel had been incubating from the vessel spectra. If the vessel spectra did not change over 5 h, the BEST would detect no correlation between the vessel spectra and time. If the spectra did change with time (indicating a change in the chemical composition of the vessel wall), however, the BEST would be able to predict the amount of time the vessel had been incubating. The exact spectrum of the change that occurred could then be extracted and compared to known reference spectra.

Figure 12 shows the results of the BEST prediction for the control arteries. Before each artery segment was scanned, the segment was removed from the incubation media and washed with fresh Krebs solution to remove lipoproteins adhering nonspecifically. Additionally, during the scan the artery was maintained in fresh Krebs solution. The predicted incubation time for the control aortic segment for the HDL experiment is depicted with an O, and the control segment for the LDL is depicted with an X (Figure 12). There is little correlation between the spectra of the control arteries and time for the first 3 h of the experiment. After 3 h a very small change begins to occur in the spectra, but this change probably corresponds to the beginning of cell death in the tissue (the arteries begin to lose their resilience between 3 and 5 h).

Figure 13 shows the correlation between aorta spectra and incubation time in LDL. The correlation is high and nearly linear for the first 3 h and then begins to decrease between 3 and 5 h ($r^2 = 0.98$, SEE = 0.3 h). Figure 14 shows the correlation between the aorta spectra and the incubation time in HDL ($r^2 = 0.003$, SEE = 2.2 h). Although the endothelium of each of the four segments was partially denuded, only the LDL-incubated aortic segment showed a significant spectral change over time. The weighting spectrum (from principal-axis transformation) of the change in the LDL-incubated aortic segment appears in Figure 15. The two peaks in the 1700–1800-nm region correspond in energy to the major peaks in the apoB spectrum.

Oxidatively modified LDL (a modification suspected of increasing the risk of atherogenesis) and native LDL were easily identified using near-IR and the supercomputer. Native human LDL is shown in Figure 16, which depicts clearly the

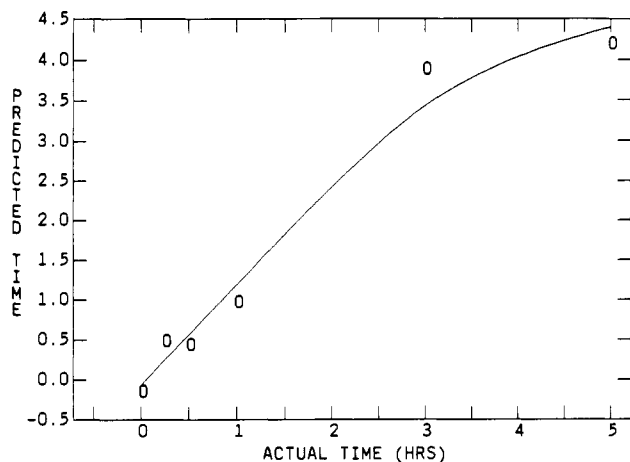


Figure 13. Correlation between aortic spectra and incubation time in LDL. LDL accumulation by arterial walls causes predictable changes in the spectra of the vessel walls over time.

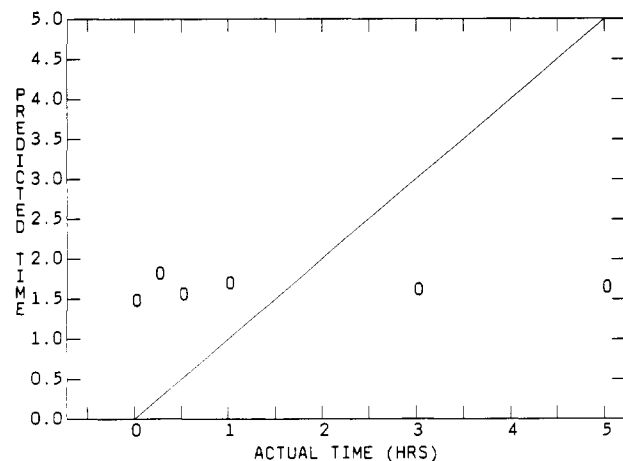


Figure 14. Correlation between aortic spectra and incubation time in HDL. HDL is not taken up by the arterial walls.

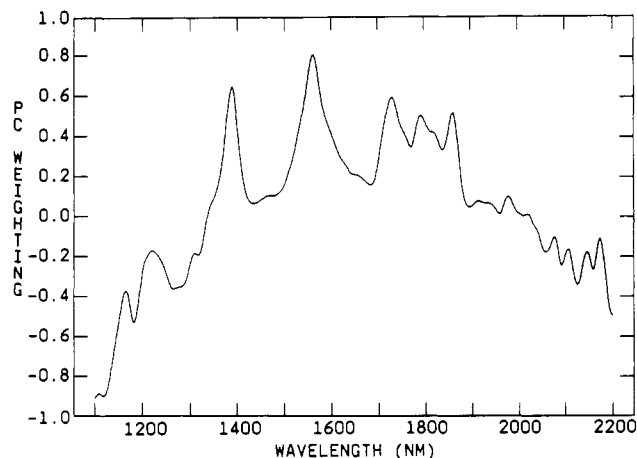


Figure 15. Weighting spectrum of the change in LDL-incubated aortic segments. The principal component that correlates to changes in LDL in arterial walls is calculated using a transformation matrix whose elements are weights applied to each wavelength in the original near-IR spectra. The two peaks in weighting in the 1700–1800-nm region correspond in energy to the major peaks in the apoB spectrum.

two major peaks of apoB between 1700 and 1800 nm. The spectrum of native LDL was obtained from an aliquot withdrawn from the bottle immediately after the bottle was opened because native LDL is known to oxidize upon exposure to the atmosphere. Figure 17 depicts the near-IR spectrum of oxidized LDL (24-h incubation with $10 \mu\text{M}$ CuSO_4 , 37°C) followed by dialysis against PBS at 4°C for 24 h. Because

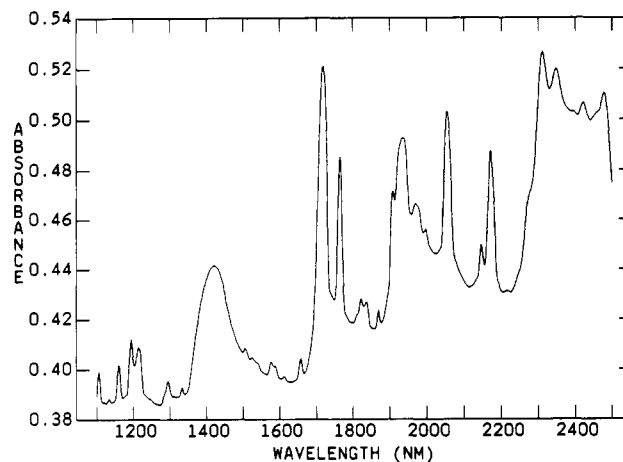


Figure 16. Near-IR spectrum of native human LDL.

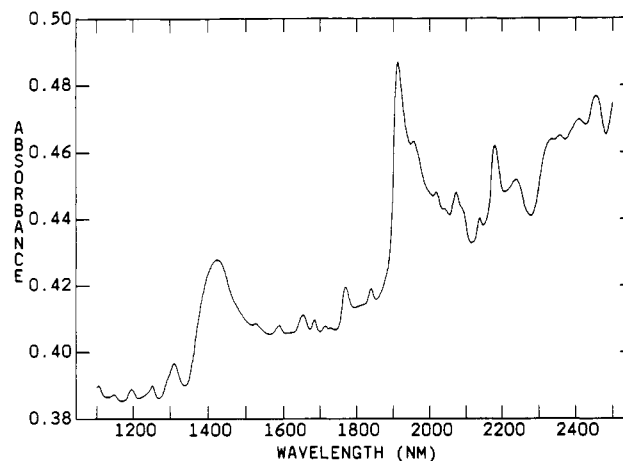


Figure 17. Near-IR spectrum of oxidatively modified human LDL.

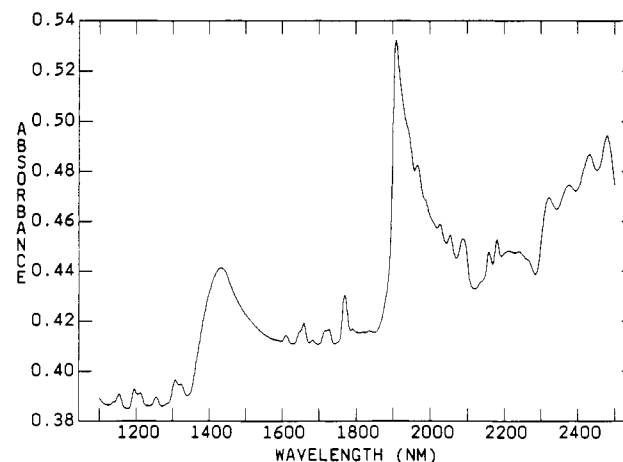


Figure 18. Near-IR spectrum of control LDL that oxidized in air.

of the known tendency of LDL to oxidize, a control solution of LDL was incubated in the same manner in the absence of CuSO_4 . The spectrum of the control LDL appears in Figure 18.

Oxidation of LDL is known to fragment apoB, and the peaks between 1700 and 1800 nm are diminished in both the control and oxidized LDL spectra. The oxidized LDL spectrum, however, shows a greater reduction in these characteristic peaks. The oxidized LDL spectrum also shows a new peak near 2240 nm that may indicate the presence of apoB fragments. A reference assay for lipid peroxidation based on thiobarbituric acid reactive substances indicated that the lipoprotein samples in Figures 17 and 18 were approximately equally oxidized.

CONCLUSION

A near-infrared fiber-optic probe has been used with a new supercomputer algorithm for chemical analysis of arterial lesions in living tissue. The ability of the probe and supercomputer to quantify and map the positions of HDL and LDL and apolipoproteins A-I, A-II, and B at clinical concentrations has been demonstrated.

The transport of cholesterol by LDL to arterial walls has been observed by near-IR imaging, which has indicated steady accumulation of LDL by arterial endothelium and smooth muscle cells. LDL accumulation by denuded vessels was greater than LDL accumulation by intact vessels, providing evidence supporting the injury hypothesis of lesion formation. HDL did not accumulate in the arterial walls. Oxidation of LDL has been observed in near-IR spectra, which have shown a loss of characteristic peaks for apolipoprotein B, a major constituent of LDL.

The combination of an imaging spectrometer and near-IR fiber-optic probe with the BEST on the supercomputer provides a powerful tool for monitoring changes in arterial wall composition, particularly in the endothelium where reflectance signals are strongest. Control of illumination through apertures gives good 3-D imaging capability, not only in the artery wall but even through the wall to surrounding tissues if desired. The low energy of near-IR photons (~ 1 eV) makes the technique nondestructive and suitable for repeated or extended use in patients or experimental animals.

ACKNOWLEDGMENT

This research was supported in part by National Science Foundation Grant STI-9108764 and by the National Institutes of Health through Grant HL45143. Portions of this paper were included in an entry to the 1990 IBM Supercomputing Competition that received first prize in the Life and Health Sciences Division.

GLOSSARY

Special Defined Operations

$M(x)$	median of x (x is a set, vector, 1- or 2-D array)
$R(f(x))$	roots of $f(x)$ by trapezoidal interpolation
r	random number on $0 < x < 1$, Monte Carlo integration of continuous uniform distribution
$\kappa(T)$	creates bootstrap distribution B for training set T and finds the center C of the distribution
$\psi(T, B, X, C)$	finds BEST distance from center C of training set T to new spectrum X using probability determined with bootstrap distribution B
$[x]$	greatest integer function of scalar, set, vector, or 1- or 2-D array
$\Phi(x)$	$1/(2\pi)^{1/2} \int_{-\infty}^x e^{-t^2/2} dt$, area from $-\infty$ to x
$\Phi^{-1}(x)$	inverse of above; i.e., given area, find x
$Q(x)$	ordered elements of x (x is a set, vector, 1- or 2-D array)
$=$	equals, or "is replaced by" when the same variable appears on both sides of $=$
$ $	"such that" qualifier on a variable, e.g., $\{x 0 < x < 1\}$ specifies the range of possible values for x
$ x $	the absolute value of x (x is a scalar)

Scalars

n	training-set size, i.e., number of samples
d	number of wavelengths
m	number of training-set replications comprising bootstrap distribution (user determined)

σ	BEST standard deviation (SD), average of upper and lower confidence limits producing a symmetric distance
σ_c	error-adjusted BEST SD, asymmetric value produced using only one confidence limit
r_h	hypercylinder radius (user-determined)
δ	skew sensitivity (user-determined)
n_h	number of spectral points falling inside a hypercylinder
l	lower confidence-limit index (index is a position in an ordered array that expresses the value of an integral from the end of the array to the index)
u	upper confidence-limit index
α	contour level specified by $\Phi(-\sigma)$, used to determine whether test spectrum is inside or outside a cluster
z_α	$\Phi^{-1}(\alpha)$
$S_{(02)}$	Euclidean distance from bootstrap-distribution center C to new spectral point X
$S_{(C0R)}$	Euclidean distance from training-set center C _(T) to bootstrap-distribution center C
$S_{(C2R)}$	Euclidean distance from training-set center C _(T) to the new spectral point X
$S_{(CUB)}$	$1/2$ the total length of the sides of a triangle specifying a particular plane in hyperspace
A_c	area of a triangle whose vertices specify a particular plane in hyperspace
$S_{(CR)}$	Euclidean distance from training-set center C _(T) to hyperline connecting C to X
$S_{(CP)}$	Euclidean distance $S_{(C0R)}$ projected on the hyperline connecting C to X
z_e	index for error adjustment in S_q
z_o	$\Phi - 1(z_e/n_h)$

Matrices, Vectors, and Arrays

B = $(b_{ij})_{m,d}$	m by d bootstrap distribution
C = $(c_j)_d$	center of the bootstrap distribution B
P = $(p_{ij})_{m,n}$	training-set sample numbers selected for bootstrap-sample sets used to calculate the bootstrap distribution
B _(S) = $(b_{(s)ij})_{n,d}$	bootstrap sample set used to calculate single rows of B
T = $(t_{ij})_{n,d}$	training-set sample spectra
X = $(x_j)_d$	test-sample spectrum
S _(0R) = $(s_{(0R)i})_m$	Euclidean distances from each element of B to C
S _(2R) = $(s_{(2R)i})_m$	Euclidean distances from each element of B to X
S _(UB) = $(s_{(UB)i})_m$	$1/2$ total length of triangle sides formed by planes in hyperspace connecting X , C , and the rows of B
A = $(a_i)_m$	areas of triangles formed by planes in hyperspace connecting X , C , and the rows of B
S _(R) = $(s_{(R)i})_m$	radial Euclidean distances from the rows of the bootstrap distribution B to the hyperline connecting X to C
S _(p) = $(s_{(p)i})_m$	Euclidean distances from C to the rows of B projected on the hyperline connecting C to X
S _(q) = $(s_{(q)i})_{n_h}$	ordered n_h elements of $\{s_{(p)i} r_h < s_{(R)i}\}$
C _(T) = $(c_{(T)j})_d$	center of training set T by $M(t_{ij})$
F = $(f_i)_{n_h}$	n_h elements of S _(q) , corrected using C _(T)
T _(N) = $(t_{(N)i})_{n_h}$	n_h independent variables of the set $\{1, 2, 3, \dots\}$ paired with F to locate the root of F

RECEIVED for review July 28, 1992. Accepted December 18, 1992.