Advances in Amino Acid Analysis

John C. Anders

With the advent of proteomics, innovative protein-based biopharmaceuticals are entering the drug development pipeline. Compositional analysis, protein quantitation, and absorptivity determination by amino acid analysis ensures accurate characterization of recombinant proteins, amino acids, and peptides in development, bulk, and final products.

John C. Anders is a research associate in the biotechnology laboratory at AAI International, 2320 Scientific Park Drive, Wilmington, NC 28405, 910.254.7943, fax 910.254.7945, john.anders@aaiintl.com, www.aaipharma.com/aai.

Following elucidation of the genetic code by Watson and Crick in the 1950s, scientists searched for ways to understand the genome (1–3). That search branched into two broad paths: One following DNA and leading to modern molecular biology, and the other focusing on characterizing and understanding the function of proteins, the products of the genome. Genomic and protein sciences diverged into separate disciplines but were always linked by the translation of genes into proteins. The two paths converged again into a new science — proteomics — after the human genome-sequencing project was completed. The attempt to understand the complete relationship of genes to their protein products in different tissues and in both development and disease has renewed interest in protein chemistry.

Early protein chemists developed methodologies to enhance their understanding of protein structures and function. In the late 1950s and early 1960s, scientists developed techniques for determining the primary sequences and amino acid compositions of proteins. In the 1960s, Pehr Edman developed the first workable chemical process for sequencing a protein. At about the same time, Stein and Moore devised the automatic amino acid analyzer for describing proteins, earning them the 1972 Nobel Prize in chemistry (4). The underlying protein chemistry of those techniques remains basically unchanged from the early days. However, remarkable progress has made those methods more sensitive, automated, and available for use in any biochemical laboratory. This article describes amino acid analysis (AAA) and its uses in protein science, particularly in characterizing protein-based biopharmaceuticals.

Amino Acid Analysis

AAA is, in its simplest form, hydrolysis of a protein or peptide into its individual components as free amino acids (Figure 1). The method is therefore independent of shape, charge, or protein function because the protein is completely degraded. Free amino acids are subsequently separated for analysis using chromatography. Protein degradation to free the amino acids is commonly accomplished by acid hydrolysis under vacuum and heat, typically at temperatures of 110–165° C for one to 24 hours. Under such hot acid conditions, the peptide bond between amino acids is hydrolyzed (adding a molecule of water across the peptide bond), resulting in free amino acids. The temperature and time of hydrolysis depend on the nature of the desired analysis and the primary sequence of the protein sample.

Amino acid variability. The 20 naturally occurring amino acids (as listed in the box below) vary extensively in their stability and lability after exposure to harsh chemical conditions. Six amino acids — Asp, Glu, Ala, Leu, Lys, and Gly — are well recovered and commonly used for protein quantitation according to their residue yields. Actually, eight residues are well recovered because Asn and Gln are converted to Asp and Glu by deamidation under harsh acid conditions.

The Amino Acid Code

Ala	Alanine	Gly	Glycine	Pro	Proline
Arg	Arginine	His	Histidine	Ser	Serine
Asn	Asparagine	Ile	Isoleucine	Thr	Threonine
Asp	Aspartic Acid	Leu	Leucine	Trp	Tryptophan
Cys	Cysteine	Lys	Lysine	Tyr	Tyrosine
Gln	Glutamine	Met	Methionine	Val	Valine
Glu	Glutamic Acid	Phe	Phenylalanine		

Some residues, such as Cys and Trp, are almost completely degraded under acid conditions. Others, such as Ser and Tyr, are significantly (>30%) degraded during acid hydrolysis. Hydrophobic residues, such as Val, Leu, and Phe, have peptide bonds that are difficult to hydrolyze when the molecules are adjacent to other hydrophobic residues. In optimizing hydrolysis, the conditions appropriate to the primary structure of the protein sample have to be determined.

Chromatographic separation. Hydrolysis is followed by chromatographic separation of free amino acids. Chromatography requires the derivatization of the amino acid residues so that they can be detected under ultraviolet (UV) or visible (Vis) light detectors. Without derivatization, conventional UV/Vis light detectors detect amino acid residues to a limited extent.

Here amino acid analysis separates into two camps: precolumn and postcolumn derivatization methods. Precolumn methods require the derivatization of amino acids before chromatographic separation and use reagents such as phenyisothiocyanate (PITC), *o*-phthalaldehyde (OPA), or Fmoc-chloride. Precolumn methods generally use reversed-phase chromatography to separate derivatized amino acid residues on C18 columns for compositional or quantitative analysis. Many hardware platforms exist for precolumn analysis, such as the ABI430 and ABI130 from Applied Biosystems (www.appliedbiosystems.com) and the PicoTag from Waters (www.waters. com) as well as conventional gradient, high-performance liquid chromatography (HPLC) platforms.

Both precolumn and postcolumn methods are widely used for AAA of proteins. Precolumn methods are somewhat less precise than postcolumn methods but are more sensitive (picomolar quantities rather than nanomolar quantities are required). As a consequence, precolumn AAA methods are commonly used in basic and applied research laboratories in which the amount of protein sample is often limited. Postcolumn methods are often used in QA/QC laboratories in which biopharmaceutical protein samples are in greater abundance.

Early postcolumn systems employed two ion-exchange columns: the first for separation of the acidic and neutral residues and the second for separation of the later-eluting basic residues. Since the mid-1980s, a single ion-exchange column system, the Beckman 6300/7300 (www.beckman.com), has dominated postcolumn biopharmaceutical AAA. That method mixes the reagent ninhydrin with separated free amino acids in the column eluent at approximately 135 °C. Ninhydrin reacts with primary amines in free amino acids and with other amine-containing molecules (such as urea). The reaction forms a blue ninhydrin–amino acid derivative detectable at 570 nm for most residues and at 440 nm for the ninhydrin–proline derivative. Separation is achieved with cation-exchange chromatography on a sodium ion stationary phase using a citric acid–based mobile phase. Increasing pH and temperature gradients resolve all 20 naturally occurring amino acids and norleucine (a commonly used internal standard calibrant) (Figure 2). At the end of the run, the column is reequilibrated with sodium hydroxide.

In the early 1960s, ninhydrin chromatography was quite a feat and is still considered an elegant separation process for early liquid chromatographic sciences. Later, the separation was enhanced by the use of a lithium-based, mobile-phase on a lithium ion-exchange column for analysis of physiological samples containing a large number of nonnaturally occurring amino acids and other amine-containing molecules. Young and Yamamoto described more than
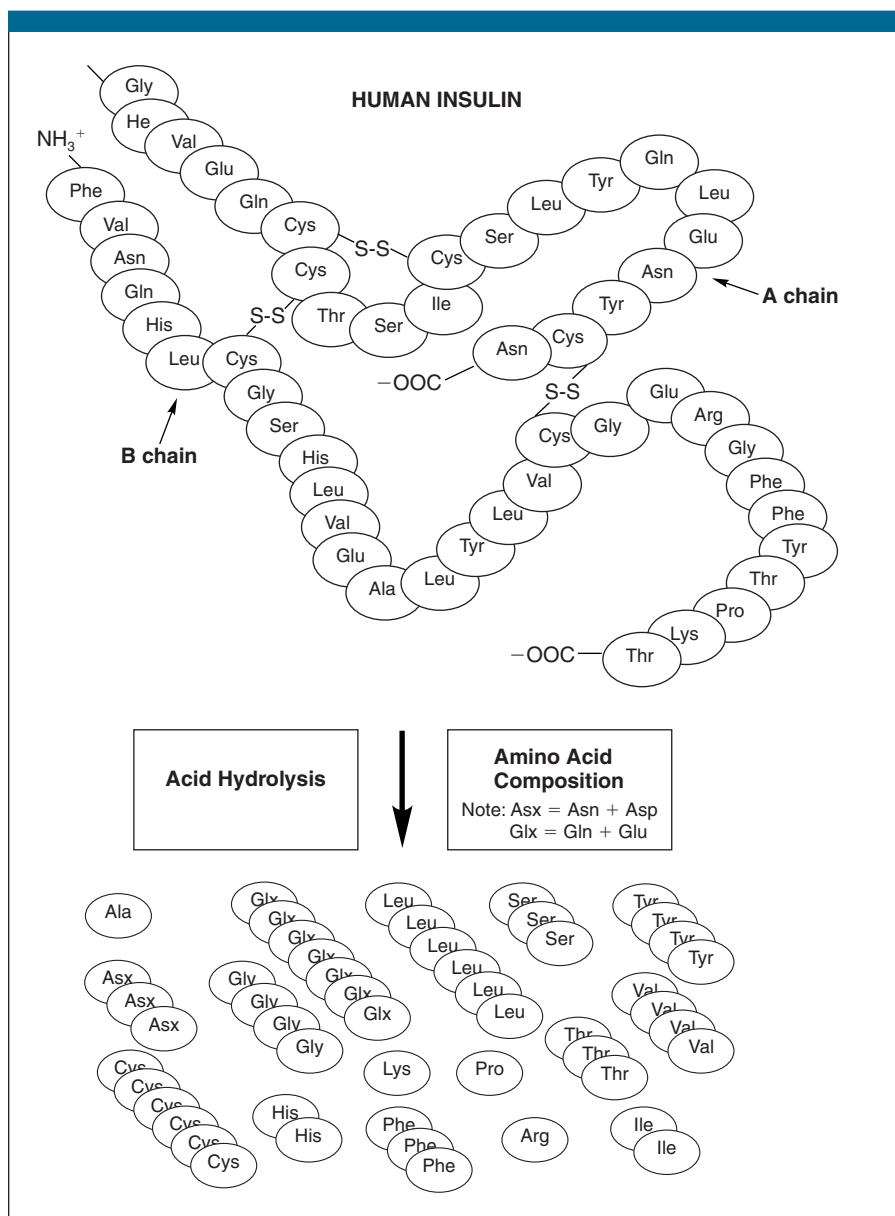

Figure 1. Amino acid analysis is hydrolysis of a protein or peptide into its individual components as free amino acids.

(a)

$$\text{Average total yield (nmol)} = \frac{\text{total nmols of all amino acids assayed}}{\text{theoretical number of residues in the protein molecule}}$$

(b)

$$\text{Number of residues of Ala in a protein} = \frac{\text{assayed yield of Ala (nmol)}}{\text{average total yield (nmol)}}$$

Equation 1. Determining the amount of a particular amino acid residue in comparison to the theoretical abundance of that same residue in the parent macromolecule; (a) calculation for the average total yield; (b) calculation for the number of residues analyzed in the protein sample for comparison

$$\text{nmols of protein} = \frac{\text{nmols of a well recovered residue (such as Ala)}}{\text{theoretical number of Ala residues in the protein}}$$

Equation 2. Determining parent protein concentration using alanine as an example

$$\varepsilon(280)\left(M^{-1}\,cm^{-1}\right) = \left(\#\,Trp\right)\left(5{,}500\right) + \left(\#\,Tyr\right)\left(1{,}490\right) + \left(\#\,Cys\right)\left(125\right)$$

converting molars to specific absortivity uses the relationship $a_m = a_s \times MW$

Equation 3. Using a statistical tool developed by Pace et al. (7) to calculate molar absorptivity constants (a_m or ϵ) based on the content of tryptophan, tyrosine, and those cystines coupled with disulfide bonds

52 ninhydrin-positive molecules separated on such columns (5). In 1995, Lowell Ericsson produced an extensive list of the postcolumn elutions for hundreds of nonnaturally occurring amino acid residues and other amines containing molecules on both sodium and lithium columns (6).

AAA Applications
AAA has many uses in biopharmaceutical production and compliance testing. These include protein and peptide quantitations, amino acid composition, protein residue and concentration calculations, and determination of protein-specific absorptivity constants.

Protein and peptide quantitation. AAA of recombinant proteins or other protein- or peptide-based biopharmaceuticals can determine the composition and quantity of naturally occurring amino acids. The quantity of peptides or proteins present in biopharmaceutical bulk or in final products can be accurately measured by analyzing well-recovered residues (Asp, Glu, Ala, Leu, Gly, and Lys). Other relatively well-recovered residues, such as Arg, Ile, Val, and Phe, are used if they are not affected adversely by excipients or located in hydrophobic regions on a protein. Labile

residues, such as Thr, Tyr, Trp, and Ser, are not recommended for quantitation. However, such residues can be analyzed using several specialized AAA methods (described below), which improve recovery over typical 6N HCl and phenol hydrolysis.

Amino acid composition. Compositional analysis of each amino acid is commonly expressed in the mole percentage (mol %) of the amino acid relative to all other residues in the protein or peptide. For example, if a protein sample has 5 nmols of Ala, and the total amount of all amino acids in that sample including Ala is 200 nmols, the mol % of Ala is 2.5. Compositional analysis provides a unique distribution profile of all amino acids in a particular protein. As such, it can be used as an identity tool for characterizing protein biopharmaceuticals.

Protein residue calculations. AAA can also be used to determine the number of residues of a particular amino acid and compare that to the theoretical abundance of that same residue in the parent protein macromolecule. This is an extension of calculating the mol %. Calculate the average yield of amino acids (Equation 1a). Then, calculate the number of residues analyzed in a protein sample (for example, Ala) for comparison with the theoretical number of residues (for example, Ala) in the parent protein (Equation 1b).

Protein concentration. A variant analysis can be used to determine the abundance of a parent protein (in nmols) based on a quantitative analysis of the well-recovered residues. Accurate determination of protein concentration requires use of well-recovered residues to prevent underestimation of the protein abundance. Theoretically, the yield of just one well-recovered amino acid present in a protein or peptide sample can be used for this purpose. However, using the average of several well-recovered residues is preferable to ensure accurate and precise quantitation of the parent protein or peptide. To calculate the amount of protein (in nmols), begin by calculating nmols of protein for a well-recovered residue, such as Ala (Equation 2); then calculate nmols of protein for other well-recovered residues and average the results. That average is the nmol abundance of the protein or peptide.

Specific absorptivity constants (a_s). Quantitative AAA can be used to determine the specific absorptivity constants (a_s) in mg/mL for protein concentration calculations. For small molecules, determining the extinction or absorptivity constant is a simple matter of correlating a known amount of material in solution with absorbance in a UV/vis spectrophotometer at the maximum wavelength. Proteins and peptides offer a different challenge. Proteins are seldom kept in a dried state unless lyophilized with a number of excipients (such as salts, simple or complex sugars, and other components) to ensure that they can be resolubilized in buffer solution when needed. Those excipients add to the total weight of the protein sample, and the excipient mass relative to the protein cannot be accurately known. Purified proteins are often available only in small quantities. Determining the amount of protein by weight, therefore, can be technically difficult.

An alternative procedure is to use a statistical tool, calculating molar absorptivity constants (a_m or ϵ) based on the content of Trp, Tyr, and a number of Cys coupled by disulfide bonds. The best summary of that process is in a 1995 article that describes a simple-to-use equation for calculating a molar absorption coefficient (Equation 3) (7).

Equation 3 is based on the 1989 work of Gill and von Hippel and the 1967 work of Edelhoch (8,9). The method is convenient and commonly used in early pharmaceutical development. However, it may require

independent verification by quantitative AAA, a method free of such statistical assumptions.

Quantitation of protein solution is accomplished using well-recovered residues and quantitative AAA. The protein is serially diluted in an appropriate buffer solution, and all dilutions are measured in a 1-cm quartz cuvette for absorbance at maximum wavelength (λ) (typically 280 nm for proteins and peptides). The dilution buffer blanks the spectrophotometer, and an investigator scans the solution at 190–400 nm to ensure measurements at maximum wavelength (approximately 280 nm) and to see whether significant light scatter exists at 340–400 nm.

Protein solutions typically reach maximum wavelength at 214 nm and 280 nm, corresponding to absorbance of the peptide bond and the aromatic residues, respectively. Measuring the maximum absorbance at 214 nm is problematic because of absorbance variability at such low wavelengths and should be avoided if possible. If significant light scatter is detected at the longer wavelengths, that amount should be subtracted from the absorbance at the maximum λ used. After UV/Vis analysis, aliquots are removed from the serially diluted samples and submitted for quantitation by AAA.

The amount of protein in each serially diluted sample is based on the yield of each well-recovered residue. That amount is converted to a concentration (mg/mL) by multiplying the number of nmols assayed by the molecular weight of the protein. A plot is constructed of UV absorbance (absorbance units or AU on the y axis) versus concentration of protein in mg/mL (x axis). The slope of the resulting line is the specific absorptivity constant (a_s) for that protein. The a_s is expressed as $AU \bullet \text{mL mg}^{-1} \text{cm}^{-1}$. Therefore, when you measure the same protein (in a 1-cm cell) in the same buffer, you simply determine the concentration as shown in Equation 4.

Determining a_s using Equation 4 will accurately measure the protein in solution. This is a convenient method of calculating

Conducting AAA Tests in Biopharmaceutical Laboratories

To meet FDA guidelines, the AAA tests conducted in biopharmaceutical laboratories should include the following steps.

Internal standards for AAAs. For all quantitative AAA, if not for all AAA procedures, an internal standard such as norleucine or norvaline is recommended. That minimizes the variability of the assay and provides an additional parameter to check in the event of method problems. Recovery of the norleucine internal standard is often the first parameter checked during method investigation because it can indicate whether pipetting errors or other procedural errors were present in the assay under investigation.

Positive controls for hydrolysis. Using a positive control for hydrolysis is recommended. Hydrolysis of a known amount of bovine serum albumin (BSA) or any other highly purified and well-characterized protein allows an investigator to calculate the recovery of the protein control and compare it to predetermined suitability

requirements. If recovery of the BSA control meets the recovery criteria, then the investigator has increased confidence that the AAA method is under control at the time of analysis. BSA is readily available from the National Institutes of Standards and Technology (NIST, www.nist.gov), providing QA/QC laboratories with a prequalified standard.

Multiple samples. Using multiple sample preparations and BSA controls provides additional acceptance control for analysis variability. If sample response variation falls within predetermined limits, the investigator can have additional confidence that the method is under control.

Choice of hydrolysis standards for use as chromatographic calibrants for all amino acids must be validated and of high quality. Using prequalified AAA hydrolysis standards in 0.1N HCl from NIST is convenient and is a prequalified calibration standard that can be used daily for quantitative AAA.

protein concentration in biopharmaceuticals during formulation development and bulk and final drug product manufacture.

QA/QC laboratory. Using AAA in development and in biopharmaceutical quality compliance testing requires some additional steps. Amino acid analyzers in biopharmaceutical QA/QC laboratories require equipment installation and operational qualification (IQ/OQ) testing. Chromatographic data acquisition systems that support such equipment must also be fully validated, including performance qualification (PQ) as well as IQ and OQ. If amino acid tests (for composition, protein or peptide concentration, or a_s determination) are used in biopharmaceutical formal release applications, those tests also require validation. An FDA guidance describes the procedure for validating analytical methods such as AAA tests (10), and those

procedures are outlined in the "Conducting AAA Tests" sidebar. Following the described validation procedures will make routine AAA more time consuming but will provide the necessary controls required for analytical testing in a modern QA/QC biopharmaceutical laboratory.

AAA Limitations

As mentioned above, several amino acid residues are heavily degraded during acid hydrolysis, and specialized methods have been developed to minimize that problem. Recovery of Ser and Thr is poor under typical 6N-HCl hydrolysis conditions. Investigators who use AAA routinely purchase highly purified HCl and evacuate hydrolysis vials under vacuum to remove as much oxygen as possible, minimizing unwanted oxidative degradants. Liquefied phenol at 1–4% v/v with 6N HCl is

$$\text{Protein concentration } (mg \, / \, mL) = \frac{\text{absorbance of protein or peptide in sample solution } (1 \, cm \, path \, length)}{a_s \left(AU \bullet mL \, mg^{-1} \, cm^{-1} \right)}$$

Equation 4. The specific absorptivity constant ($AU \bullet$ mL mg^{-1} cm^{-1}) is used to calculate the amount of protein in a solution.
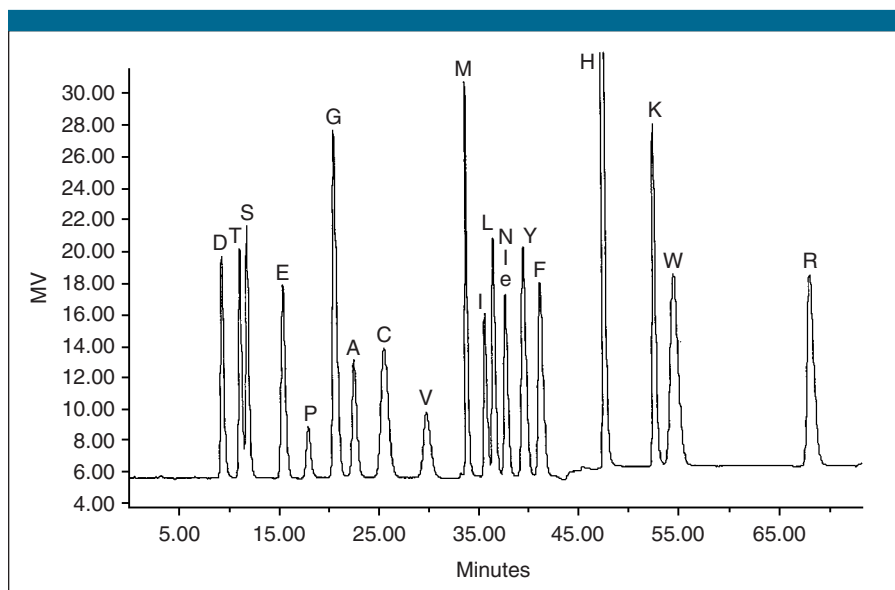
Figure 2. Increasing pH and temperature gradients result in resolution of all 20 naturally occurring amino acids and norleucine (a commonly used internal standard calibrant).

commonly used to preserve hydroxylated residues such as Ser, Thr, and Tyr.

Hydrophobic residues. Proteins with very hydrophobic residues (such as Leu, Val, and Phe) pose a problem if high recovery is required. Typically, extended hydrolysis with 6N HCl and phenol is required to hydrolyze those stubborn bonds. Long hydrolysis times can result in loss of other labile residues. Therefore, an extended hydrolysis may be required in addition to a more gentle hydrolysis to obtain a complete compositional analysis.

Cys presents a special problem. Proteins or peptides with abundant Cys (greater than approximately 3 mol %) produce several Cys degradants. Those degradants, a mixture of disulfides and other molecules, can coelute with Pro, His, and Met in cation-exchange chromatography using Beckman postcolumn systems. Degradants that elute with Pro can result in falsely high results for that residue. Met and His are similarly affected, resulting in poor precision and recovery.

Oxidative hydrolysis of Cys residues by performic acid hydrolysis or addition of 2% v/v dimethyl sulfoxide (DMSO) to the 6N HCl hydrolysis mixture is commonly used to overcome those interferences. Oxidative hydrolysis stoichiometrically converts Cys to cysteic acid, a stable form of that residue. That conversion also prevents formation of Cys degradants, allowing Pro to be accurately quantitated. Reduction and alkylation of Cys disulfides using

pyridylethylation produces nearly equivalent results.

Special conditions for Trp. Trp is undoubtedly the most labile of the residues and requires special hydrolysis conditions. Typically, Trp is hydrolyzed in either mercaptoethanesulfonic acid (MESA) or methanesulfonic acid (MSA). Following heating, the samples are titrated with base to pH 7 for MESA and pH 2 for MSA. Another method for analyzing Trp is HCl hydrolysis with dodecanethiol in thioglycolic acid. A good reference for those tests is an ABRF92-AAA study (11).

Future Innovations

With the advent of proteomics, innovative protein-based biopharmaceuticals are entering the drug development pipeline. So compositional analysis, protein quantitation, and absorptivity determination by AAA will be used for some time.

A current challenge is the obsolescence of the Beckman postcolumn instruments, production of which was halted five years ago. QA/QC investigators are searching for alternative, equivalent hardware to replace those units. The most direct replacement is the Hitachi (www.hitachi.com) 8800 amino acid analyzer. Hitachi offers simple replacement of current validated amino acid methods in use worldwide. The only drawback to the 8800 is that it is unable to separate Leu and norleucine (Nle), requiring

the use of norvaline (Nva) as an internal standard, which could affect current validated methods.

My laboratory is developing and validating a modified dual pump Hewlett Packard (www.hp.com) P1100 HPLC as an equivalent replacement for the Beckman. Dionex (www.dionex.com) now offers an amino acid analyzer using its ion-exchange HPLC. That system eliminates the need for derivatization by detecting free amino acids with an integrated pulse amperometric detector (iPAD).

References

(1) J.D. Watson and F.H.C. Crick, "Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid," *Nature* 171, pp. 737–738 (1953).

(2) F.H.C. Crick, "Note 86: Nucleic Acids," *Sci. Am.,* pp. 188–191 (1957).

(3) F.H.C. Crick, "The Structure of the Hereditary Material," *Sci. Am.,* p. 54 (1954).

(4) S. Moore and W.H. Stein, "Aminoacid Determination: Methods and Techniques," *J. Biol. Chem.* 192, 663–670 (1951). Also "Chromatographic Determination of Amino Acids by the Use of Automated Recording Equipment," *Methods of Enzymology*, Vol. 6, S.P. Colowick and N.O. Kaplan, Eds. (Academic Press, Inc., New York, 1963), pp. 819–831.

(5) J.L. Young and M. Yamamoto, "A Mixed Lithium–Sodium Gradient Buffer System for Single-Column Amino Acid Chromatography," *J. Chromatogr.* 78, 349–355 (1973).

(6) L.H. Ericsson, "Tables of Amino Acids," presented at the Association of Biomolecular Resource Facilities Workshop: Protein Society (ABRF, Boston, July 1995). This review is available from Beckman Instruments.

(7) C.N. Pace et al., "How to Measure and Predict the Molar Absorption Coefficient of a Protein," *Protein Sci.* 4, 2411–2423 (1995).

(8) S. Gill and P. von Hippe, "Calculation of Protein Extinction Coefficients from Amino Acid Sequence Data," *Anal. Biochem.* 182, pp. 319–326 (1989).

(9) H. Edelhoch, "Spectroscopic Determination of Tryptophan and Tyrosine in Proteins," *Biochemistry* 6, 1948–1954 (1967).

(10) ICH, *Guidance for Industry, Q2B Validation of Analytical Procedures: Methodology*" (FDA, Rockville, MD), November 1996. Available at www.fda.gov/cder/guidance/1320fnl.pdf.

(11) D.J. Strydom et al., "Cysteine and Tryptophan Amino Acid Analysis of ABRF92-AAA," *Techniques in Protein Chemistry IV,* R. Angeletti, Ed. (Academic Press, New York, 1993), pp. 279–288. *BP*

The lackluster performance of publicly traded contract research organizations (CROs) has disappointed shareholders during the past four years. From 1997 to 2001, the share prices of most major CROs lagged well behind those of major market indices such as the S&P 500. By the end of 2000, share prices for most CROs had sunk 50% from what they were at the end of 1996, whereas the S&P 500 had risen 75% above its 1996 close. In the meantime, the market value of the CROs' Big Pharma clients in many cases had risen 200% to 300%.

Difficult Times

The difficult times that CROs faced during 1999 and 2000 resulted from major study cancellations and slowdowns in research programs following a wave of Big Pharma mergers. Such tough times prompted the major CROs to do some important soul-searching about the reasons for their poor performance. Three factors seemed to account for their performance problems: bloated costs, a bad business model, and an unfavorable client mix.

Bloated costs. Rapid growth in the mid-1990s fueled an expansion binge at most major CROs, which spent heavily on global infrastructure, acquisitions, and business development activities. When activity slowed in 1999, the big CROs found themselves with uncompetitive cost structures and, in some cases, substantial debt. They lost contracts to smaller, privately held CROs that enjoyed lower costs and were perceived as providing better service.

A bad business model. High costs and acquisition activity often resulted from pursuing a "one-stop shop" business

In January 2001, FDA began piloting a new "systems approach" inspection program. Much like the Quality System Inspection Technique (QSIT) program for medical device inspections conducted under Center for Devices and Radiological Health (CDRH), the aim is for faster, better-focused, and more efficient GMP audits. Under the pilot, investigators perform "abbreviated" inspections that audit companies' quality systems and one or more other systems. In the face of shrinking field resources that make it impossible to inspect drug makers every two years as required by law, FDA is considering adoption of the CDER pilot program for all districts.

FDA's rationale is simple: to promote faster, more efficient inspections — and, based on preliminary metrics, more effective inspections as well. Agency officials evaluating the program found that the six district offices were able to inspect 17% more drug manufacturing facilities during the first six months of 2001 than in the same period in 2000. In addition, inspections under the pilot took less time, averaging 12 fewer hours per inspection. Twenty-three percent (35) of the pilot's 155 inspections fell into FDA's most serious category — Official Action Indicated (OAI); 62 were classified as "Voluntary Action Indicated" (VAI); and 58 were coded as "No Action Indicated" (NAI). Normally, OAI inspections are only about 5% of all inspections (1).

The systems approach, however, is much more than a force-multiplier for FDA. It serves as a tool for companies to use to their advantage. The systems paradigm is a way of thinking about the strategic environment: how to develop processes that achieve strategic goals. This paradigm is a tool that forward thinking organizations can use to design their capability to analyze tactical and strategic environments; develop and enact strategies in response to environmental demands; and sustain an adaptive and productive organizational culture.

This article highlights some of the strategic leadership opportunities that organizations can use to enhance their knowledge gathering to ensure compliance in the new environment.

Systems: Essential Properties

A system is a set of two or more elements that satisfies three conditions: The behavior of each element affects the behavior of the whole; the behavior of the elements and their effects on the whole are interdependent; and subgroups are formed, each of which affects the behavior of the whole and none of which independently affects it.