1

Preformulation Studies as an Essential Guide to Formulation Development and Manufacture of Protein Pharmaceuticals

David B. Volkin, Gautam Sanyal, Carl J. Burke, and C. Russell Middaugh

1. INTRODUCTION

Preformulation generally refers to a process in which a bulk drug material is characterized to a sufficient extent that it can be converted to a pharmaceutically acceptable drug substance. In the case of conventional, small-molecule (nonmacromolecular) pharmaceuticals, this procedure can usually be accomplished to a high degree of scientific rigor. For example, the atomic level structures of these compounds are usually known from a combination of X-ray crystallography, nuclear magnetic resonance (NMR), and mass spectrometry. Furthermore, the sensitivity of modern high-performance chromatography systems coupled to mass spectrometric detectors often permits altered forms of these molecules to be simply detected and identified. Unfortunately, the increased complexity of proteins, arising from a combination of their large size and higher order forms (secondary, tertiary, quaternary) of structure does not yet permit

David B. Volkin and Carl J. Burke • Department of Vaccine Pharmaceutical Research, Merck Research Laboratories, West Point, Pennsylvania 19486. *Gautam Sanyal* • AstraZeneca Research and Development Boston, Waltham, Massachusetts 02451. *C. Russell Middaugh* • Department of Pharmaceutical Chemistry, University of Kansas, Lawrence, Kansas 66047.

Development and Manufacture of Protein Pharmaceuticals, edited by Nail and Akers. Kluwer Academic/Plenum Publishers, New York, 2002.

the more straightforward type of analysis utilized for small molecules to be simply applied. Nevertheless, it is still possible to treat proteins primarily as chemical (rather than biological) entities through an understanding of their structure and stability and use of a more complex array of lower resolution methodologies.

Under ideal circumstances, one would like to have available an atomic-level-resolution structure of an early protein drug candidate determined by X-ray crystallography and/or NMR. This is often not the case, however, due to the labor-intensive nature of both approaches as well as the inherent difficulties in obtaining suitable crystals for the former and size limitations by the latter. As has previously been pointed out, however, once a structure has been obtained by one of these two methods, simple comparisons of either diffraction patterns or two-dimensional NMR proton spectra can be used to establish structural identity to standard proteins (Middaugh, 1994). More importantly, modern methods of mass spectrometry, especially the electrospray and laser desorption time-of-flight varieties, now permit the molecular weight of even large proteins to be determined to a high degree of accuracy (i.e., 0.01%). Very little information about protein conformation is available, however, from such measurements. For this reason, preformulation structural analysis of proteins usually consists of use of a wide variety of methods that examine various aspects of a protein's structure. For example, the secondary structure can be estimated by both far-ultraviolet (UV) circular dichroism (CD) and Fourier-transform infrared (FTIR) spectroscopies, and tertiary structure can be probed by near-UV CD and intrinsic fluorescence studies. The oligomeric state of a protein as well as potential aggregation phenomena can be explored by a combination of molecular sieve chromatography, analytical ultracentrifugation, static and dynamic light scattering, and chemical crosslinking. Less direct, but very sensitive, techniques used primarily in a comparative mode that sense various aspects of structure include capillary electrophoresis, ion-exchange and reversed-phase high-pressure liquid chromatography (HPLC), and isoelectric focusing. When a combination of the above techniques is employed, a fairly comprehensive picture of a protein can be constructed. Most importantly, these methods can be tweaked to detect alterations in protein structure. The use of these methods is considered below in more detail.

Probably the most critical portion of protein preformulation activities is what can best be described as finding the "weak points" in the molecule. Through the use of a combination of methods, a protein can be "stressed" and its susceptibility to various forms of physical and chemical degradation evaluated. Most often, temperature, pH, and/or redox conditions are used to induce structural alterations, because these variables have often been found to significantly elevate the rate of structural changes commonly seen in proteins. Although such studies cannot absolutely predict corresponding changes in real-time stability studies, they usually provide important clues in this regard.

The most common chemical changes encountered in proteins include deamidation of asparagine (and, to a lesser extent, glutamine) residues, cleavage of peptide bonds, cystine destruction, thiol–disulfide interchange, oxidation of cysteine and methionine, photodegradation of Trp, Tyr, and Cys, and glycation and carbamylation of amino groups (Volkin *et al.*, 1995b). Such changes are usually straightforward to detect and characterize using chromatographic methods combined with mass spectrometry, especially if they are present to significant extents. Conformational alterations, especially if they are of a subtle nature, can be much more difficult to detect, because one cannot be sure that a particular method such as CD or fluorescence is sensitive to the region of the structure in which the alteration occurs. This is perhaps the biggest uncertainty in the preformulation of proteins, but can be minimized by the use of multiple techniques and different environmental stresses, which may reveal small changes undetectable in the native state. A complete thermodynamic analysis of the folding and unfolding of a protein can be particularly useful in this regard if conditions of reversibility can be found. Similarly, the effect of stress on the aggregation state of the protein as well as on its bulk solubility is often useful.

This information should be immediately helpful in devising methods to increase the stability of a protein if this is necessary. Clues to potential problems with novel delivery methods and the necessity of lyophilization or even possibly the need to reengineer the protein should be immediately addressable after these types of analyses. In particular, we note three areas of special concern. First, if a protein contains a critical residue which is unusually labile, the creation of appropriate mutant forms needs to be immediately considered, although potential problems of immunogenicity must temper any such considerations. Second, if a protein contains specific binding sites for pharmaceutically acceptable ligands, such sites need to be quickly identified and extensively characterized, because this provides an immediate opportunity to dramatically stabilize a protein through inclusion of a simple excipient. Several examples of such a situation have been described (see below). Finally, if covalent modification of the native structure is being considered for stability or bioavailability reasons, similar preformulation studies of this new entity need to be quickly initiated, because previous work on the native form may be at least partially inapplicable. We now consider various aspects of these approaches in more detail.

2. PROTEIN STRUCTURAL CHARACTERIZATION AND CONFORMATIONAL STABILITY

In designing a stable protein formulation, the integrity and stability of the folded tertiary structure of the protein must be thoroughly examined,

because proteins owe their specific biological functions to proper folding of the primary structure into a higher order structure with the minimum accessible free energy. Any successful formulation must ensure that this "native" structure is maintained for as long as possible under practical storage conditions. Ideally, to achieve this formulation goal, we should have at our disposal (a) sensitive biophysical assays for determining perturbations of native structure, (b) biological assays for assessing the significance of structural perturbations on function, and (c) methods for monitoring thermodynamics and kinetics of stabilization or destabilization of native structure. Most of the therapeutic proteins that are commercially available or in development are made using the recombinant DNA technology, and the accompanying purification process may involve unfolding and refolding of the protein. Biophysical characterization of such protein products must be performed to confirm that the correctly folded and biologically active structure is present. This may be considered either the final step in process development or the first step in preformulation research of protein therapeutics.

Several spectroscopic techniques can be used to examine the tertiary structure of proteins in solution and to assess the stability and effects of different formulation conditions on protein structure. These methods also provide the formulation scientist with tools for screening for suitable stabilizing agents if the inherent thermodynamic stability of a protein solution is inadequate for long-term storage. In the following discussion, we will focus primarily on optical techniques, i.e., UV-visible absorption, fluorescence (emission), circular dichroism (CD), and Fourier-transform infrared (FTIR) absorption spectroscopies. We will discuss the applications of these techniques to protein structure analysis and folding/unfolding studies, and discuss their relevance to pharmaceutical formulation development of proteins. In addition, the use of differential scanning calorimetry to monitor the thermal stability of protein formulations will be considered. A combination of these techniques will usually provide sufficiently detailed information on the inherent structural stability of a folded protein molecule to direct the formulation scientist to the methods which can be used to enhance a protein's stability sufficient for therapeutic use.

2.1. UV-Visible Absorption Spectroscopy

The spectral properties of the near-UV (250–320 nm) bands that originate from the three amino acids with aromatic side chains (i.e., phenylalanine, tryptophan, and tyrosine) depend upon their microenviron-

ments, which are altered when a protein's tertiary structure is perturbed. In general, near-UV spectra of proteins manifest a small blue shift upon unfolding as the polarity of the surrounding environment is increased. As a result, the absorbance at a single wavelength in the region between 280 and 295 nm (the region of maximal slope in a spectrum containing contributions from tryptophan and tyrosine residues) usually changes and this can be simply monitored as a function of the perturbing condition. The magnitude of tryptophan or tyrosine absorbance change between folded and unfolded states is usually much smaller than the change observed by fluorescence, but the high precision and accuracy of UV spectroscopy makes folding/unfolding measurements feasible. The effects of pH, temperature, or chaotropic reagents on protein structure at equilibrium can thus be studied, as was nicely demonstrated for the model enzyme ribonuclease A by Fink and Painter (1987).

In employing changes in absorption intensity as a signal for monitoring unfolding or refolding, one needs to be aware of potential light scattering contributions, which may introduce a significant inaccuracy if the protein molecules self-associate to a size comparable to the wavelength of the incident light. Aggregation is a common occurrence when proteins unfold or form a partially unfolded "molten globule" structure, and is frequently encountered in accelerated thermal stability studies. Light scattering is typically indicated by gradually decreasing optical density values in the 300- to 350-nm range, where most proteins should have no intrinsic absorbance. The Rayleigh theory of light scattering suggests that the scattering intensity is proportional to the inverse fourth power of the wavelength of light. The presence of considerable light scattering can even produce a spurious shift in the near-UV absorbance maximum, which, if not corrected for, may be misinterpreted as a change in the microenvironments of aromatic residues and consequently conformational alterations. A simple log(optical density) versus log(wavelength) relationship is often used to obtain corrected absorbance values at a particular wavelength (Mach *et al.*, 1995).

In addition to this straightforward use of UV spectroscopy, second-derivative transformations of zero-order spectra have been shown to be sensitive to changes in protein tertiary structure that alter the polarities of the microenvironments of aromatic amino acids (Ragone *et al.*, 1984, 1987). Absorbance of tryptophan residues in the electronic ground state appears to be less sensitive to solvent polarity compared to the absorbance of tyrosine residues, whereas the reverse is true for fluorescence emission, which is more affected by the environment of the electronically excited state. Thus, the two techniques are nicely complementary approaches for monitoring changes in protein structure. Deconvolution of tyrosine and tryptophan contributions to second-derivative UV absorption spectra has been achieved (Mach and

Middaugh, 1994; Mach *et al.*, 1995). Furthermore, a major second-derivative phenylalanine negative peak at 258 nm is readily resolved from tyrosine and tryptophan contributions (Mach *et al.*, 1991). Changes in the average environment of these three residues in a protein can, therefore, be simultaneously measured in both equilibrium and kinetic experiments of protein unfolding and refolding. A major advantage of using second-derivative over conventional UV zero-order spectroscopy is that baseline inconsistencies originating from light scattering are minimized.

Although peptide bonds in α-helix, β-sheet, and disordered structures give rise to distinguishable absorption spectra in the far-UV, UV absorption is not usually the method of choice for monitoring changes in the secondary structures of proteins, for the following reasons: (a) The absorption bands occur between 185 and 205 nm, a wavelength range where interferences from buffer salts and amino acid side chains make accurate measurements difficult, (b) the extinction coefficients of β-sheet and disordered structures are very similar in this wavelength range, and (3) far-UV CD and FTIR spectroscopy provide higher resolution and more accurate methods for estimation of secondary structure content and associated changes.

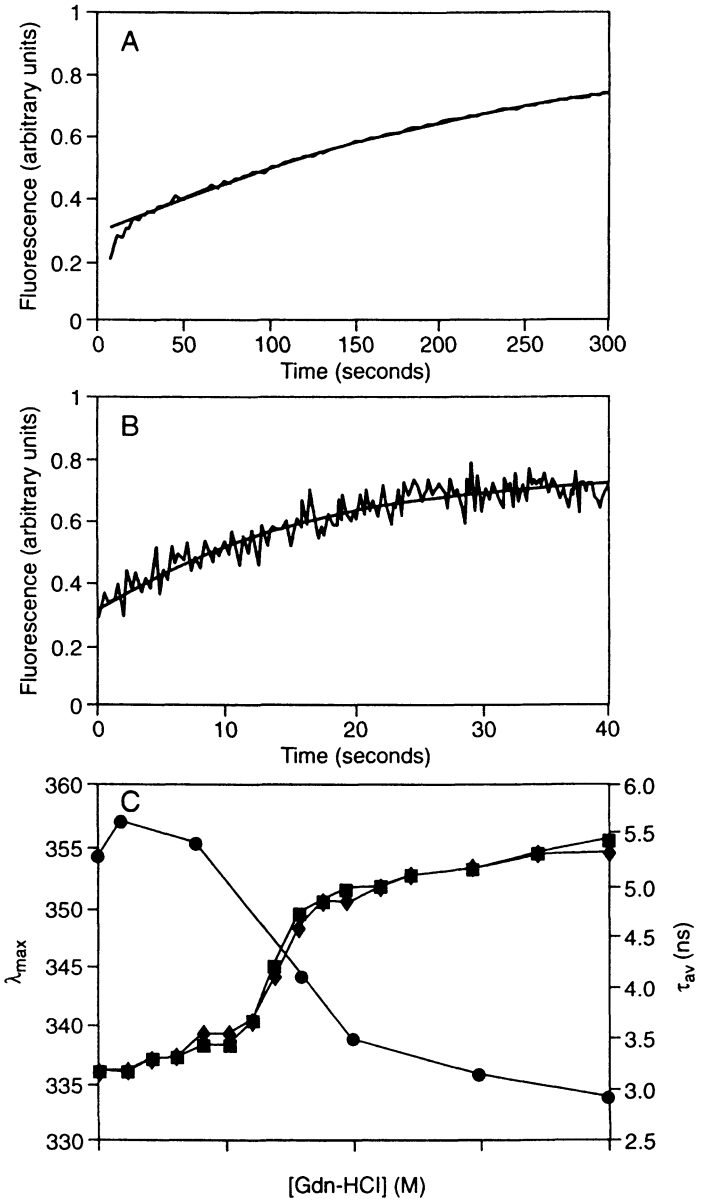
2.2. Fluorescence Spectroscopy

The intrinsic fluorescence of proteins arises from the same three amino acids that manifest near-UV absorbance. Emission from tryptophan and to a lesser extent tyrosine residues usually dominates protein fluorescence spectra, whereas the phenylalanine contribution is usually weak and therefore is not widely used in protein structural and folding studies. Consequently, like near-UV absorbance, intrinsic fluorescence also yields information on the environment of these amino acids in the tertiary structure of a protein and changes in these environments when the structure is perturbed. Unlike absorbance spectroscopy, however, in which absorption of light by the ground electronic state of molecules is observed, fluorescence entails emission of light from the excited singlet electronic state that follows excitation of molecules from their ground state upon light absorption.

The excited state of the indole side chain of tryptophan residues is particularly sensitive to solvent polarity and, when contained in a protein, to the polarity of its microenvironment. The wavelength maxima λ_{max} of tryptophan side chains in proteins usually vary from approximately 320 to 350 nm, the lower and the upper ends of the range reflecting residues buried in the hydrophobic protein interior and those completely exposed to solvent, respectively. This offers a much wider window than the absorbance method for

monitoring protein unfolding and refolding in both equilibrium and kinetic measurements as well as the ability to detect more subtle conformational changes. Furthermore, although not all of the light absorbed by a fluorophore is emitted as fluorescence, the sensitivity of steady-state fluorescence intensity measurements usually far exceeds that of absorbance measurements, thereby requiring only a fraction of the amount of the protein typically needed for near-UV absorbance experiments. For model tryptophan and tyrosine compounds, fluorescence quantum yield values are close to 0.2, suggesting that only about 20% of the photons absorbed by these fluorophores are emitted as fluorescence, the rest being lost to deexcitation of molecules by nonradiative processes. However, the measured fluorescence intensity is directly proportional to the intensity of the exciting light and to the quantum efficiency of the detector. This principle has been widely used to produce highly sensitive fluorometers that employ xenon-arc lamps or even laser sources and techniques to minimize optical loss. In many proteins, tyrosine fluorescence is quenched, in part, due to energy transfer to tryptophan, so that the effective quantum yield of tyrosine is lower. Furthermore, tyrosine fluorescence is relatively insensitive to solvent polarity. Tryptophan fluorescence spectra, in general, are much more sensitive to environment (polarity) than tyrosine spectra. Although unfolding of tertiary structure can also be studied by near-UV CD spectroscopy (250–320 nm), which also monitors changes in microenvironments of aromatic residues, the sensitivity of tryptophan and tryrosine fluorescence is much greater, allowing measurements to be made with relatively small amounts of proteins (typically a few micrograms per milliliter).

An example of using tryptophan fluorescence changes to monitor the kinetics of refolding of a potentially chemotherapeutic multidomain protein (TP40) upon dilution from a Gdn·HCl solution is shown in Fig. 1 (Gress *et al.*, 1994). TP40 is a fusion protein containing transforming growth factor-α (TGF-α) fused to a 40-kDa fragment of *Pseudomonas* exotoxin (PE). The refolding of the unfolded protein to its native state occurs via biphasic kinetics as measured by fluorescence spectroscopy (Fig. 1A). The faster phase of the apparently biphasic kinetics was quantitatively measured by stopped-flow spectrofluorometry (Fig. 1B). The tryptophan fluorescence of this protein was also utilized to monitor equilibrium unfolding as a function of Gdn·HCl concentration. Both steady-state fluorescence and fluorescence lifetime measurements were employed as complementary techniques to monitor unfolding, and very similar dependences of spectral shift and lifetime on Gdn·HCl concentration were observed (Fig. 1C). The decreases in tryptophan fluorescence intensity and lifetime as a function of increasing Gdn·HCl concentration are consistent with the tryptophan residues being exposed to a more polar environment as the protein unfolds.


A
Fluorescence (arbitrary units)
Time (seconds)
B
Fluorescence (arbitrary units)
Time (seconds)
C
λ_{max}
τ_{av} (ns)
[Gdn-HCl] (M)

The use of steady-state fluorescence to study protein folding has been greatly aided by the use of extrinsic fluorescent probes with high quantum yields. Two kinds of probes are commonly used: (a) hydrophobic probes such as 2-*p*-toludinylnaphthalene-6-sulfonate (TNS) or 8-anilino-1-naphthalene-sulfonate (ANS), which bind to hydrophobic regions of proteins, and (b) probes that can be covalently attached to specific amino acid residues. Noncovalently bound hydrophobic probes have been extensively employed for detecting either partially folded or molten globule structures in protein folding pathways. These intermediate- and often high-aggregative structures are often brought about by acidic pH, high salt concentrations, and sub-denaturing concentrations of chaotropic agents. Because in such structures the hydrophobic regions become accessible to solvent, molten globules usually bind these probes to a much greater extent than native and unfolded proteins. As an example, the effect of urea on the intrinsic tyrosine fluorescence of GroEL has been compared with urea-induced changes in the fluorescence intensity of extrinsically bound bis-ANS (Gorovits *et al.*, 1995). This study revealed that changes in the intrinsic tyrosine fluorescence reflect a two-state unfolding transition as a function of urea concentration, whereas bis-ANS fluorescence shows the presence of residual tertiary structure in one or more regions of the protein even in the presence of 3–4 *M* urea (suggesting the presence of nucleation sites for folding). Common examples of extrinsic probes used for covalent labeling of proteins are fluorescein isothiocyanate (FITC) and dansyl chloride (which react primarily with amino groups) and iodoacetamide-based probes (which preferentially label thiol groups). The approach of using extrinsic probes, especially the hydrophobic-surface-seeking probes discussed above, may be useful for detecting subtle conformational changes in protein conformations in the process of formulation development of pharmaceutical proteins, provided that the excipients or experimental conditions used do not interfere with probe fluorescence.

Steady-state fluorescence quantum yield or intensity of both intrinsic tryptophan or tyrosine residues as well as extrinsic probes is attenuated by increased temperatures. However, a transition in the intensity versus temperature plot (e.g., due to a structural change) can usually be distinguished

Figure 1. Use of intrinsic tryptophan fluorescence of proteins to monitor unfolding and refolding. (A) Kinetics of refolding of TP40, upon dilution from a largely unfolded state in 2.5 *M* Gdn·HCl to 1.0 *M* Gdn·HCl, as measured by changes in tryptophan intensity (excitation, 280 nm, emission, 335 nm). (B) The kinetics of the initial fast phase (0–40 sec) of refolding of TP40 as quantitatively measured by stopped-flow spectrofluorometry. (C) Equilibrium unfolding of TP40 and PE40δcys (a 40–kDa fragment of PE containing Ala–Cys mutations) as monitored by changes in the fluorescence emission maximum (filled diamonds, TP40; filled squares, PE40δcys), and that of TP40 monitored by tryptophan fluorescence lifetime (filled triangles). Reprinted with permission from Gress *et al.* (1994). Copyright 1994 American Chemical Society.

from the monotonic decrease in fluorescence intensity that is produced by increasing temperatures. Furthermore, the shift in the wavelength maximum of a fluorescence spectrum is usually a good indicator of the exposure of the fluorophore to altered solvent polarity and can be monitored as a function of temperature to obtain thermal stability information. For example, acidic fibroblast growth factor (aFGF) contains a single tryptophan residue, the fluorescence of which is highly quenched in the native state by neighboring histidine residues. Thus, upon excitation at 280 nm, tyrosine fluorescence is observed (Copeland *et al.*, 1991). Thermal unfolding of this protein, however, results in the generation of a tryptophan-like spectrum with concomitant quenching of tyrosine fluorescence due to relaxation of the original Trp quenching in the native state. This phenomenon has been used as a sensitive indicator of unfolding of aFGF and to study stabilization of aFGF by various polyanions (Tsai *et al.*, 1993a; Volkin *et al.*, 1993).

Changes in exposure of tryptophan or tyrosine side chains due to conformational changes can also be studied by taking advantage of the collisional quenching of their fluorescence by reagents such as iodide or acrylamide (Eftink, 1994). Surface-exposed fluorophores are quenched by iodide, which cannot usually penetrate to fluorophores buried in protein interiors. In contrast, accessibility of tryptophan or tyrosine to acrylamide, a polar, neutral molecule, varies widely depending upon the extent of burial of the residue. For example, the accessibility of the single tryptophan residue of aFGF to acrylamide was unaltered upon heparin binding to the protein, suggesting that the stabilizing interaction of this polyanion with the positive charge cluster on the surface of aFGF did not alter the local conformation around the partially buried tryptophan residue (Copeland *et al.*, 1991). This observation is consistent with the lack of an effect of heparin on either the tryptophan fluorescence lifetime or emission spectrum of aFGF.

Another powerful application of fluorescence spectroscopy lies in the determination of distances or distance distribution (for a dynamic protein conformation) between specific sites on a protein. This technique usually involves covalently labeling two sites with two different probes that can act as a donor–acceptor energy transfer pair. A decrease in emission quantum yield or lifetime of the donor (or concomitant changes in the acceptor fluorescence) depends upon the distance between the labeled sites within the protein. Conformational changes in the protein would alter this distance and be reflected in altered efficiency of energy transfer between the donor and the acceptor. In any such study, one must be certain that the labeled protein retains native structure and biological function. For example, Scheraga and co-workers site-specifically labeled RNAse A with a donor–acceptor pair (Buckler *et al.*, 1995). Time-resolved nonradiative energy transfer measurements were used to elucidate the conformation of partially

folded intermediates at sub-denaturing concentrations of Gdn·HCl. The doubly labeled RNAse was found to retain 87% of the enzymatic activity of the native RNAse. However, the midpoint of thermal unfolding of the labeled RNAse was 6°C lower than that of the wild-type protein, as detected by near-UV absorbance measurements, suggesting some destabilization of the tertiary structure upon chemical modification.

2.3. Circular Dichroism

Circular dichroism (CD) spectroscopy in the far-UV region (180–250 nm) is a powerful and convenient way to estimate the overall secondary structure content of proteins and related changes therein. Many excellent reviews have been written on this subject (e.g., Yang *et al.*, 1976; Johnson, 1990; Manning, 1989). The integrity of the native secondary structure in formulated proteins must be established, and CD spectroscopy is ideally suited for this purpose as long as the formulation excipients result in a clear solution and are not optically active in the far-UV region. Furthermore, the contribution of excipients to the optical density of the solution must be small in the wavelength range of measurement so that the dynode voltage applied to the detector stays within a reasonable range, thereby maintaining an acceptable signal-to-noise ratio.

The CD spectra results from a combination of optical absorption (electric dipole moment) and asymmetry (magnetic dipole moment). Both of these dipole moments must be nonzero for a measurable CD signal, a condition which is satisfied by peptide bonds. Furthermore, elements of protein secondary structure such as α helices, β sheets, turns, and disordered structures each produces a different magnetic dipole moment. This effect, coupled with small differences in the electronic absorption spectra of these structures, results in distinctive CD spectra from each. Different proteins contain varying contributions from these structures and therefore manifest differences in their spectra, which are made up of overlapping combinations of these individual contributions. For example, α helices manifest two distinct minima at wavelengths around 208 and 222 nm and a maximum at 192 nm. On the other hand, disordered structures produce a relatively featureless CD spectrum with a minimum around 200 nm, and β sheets yield a broad minimum at 217 nm and a maximum at 193 nm (Brahms and Brahms, 1980). Furthermore, CD intensities in the 180- to 230-nm range are different for different kinds of secondary structure. α Helices have the strongest magnetic dipole moment and therefore manifest the highest intensity bands.

When the ordered secondary structure of the peptide backbone is perturbed, such as by temperature-induced or denaturant-induced unfolding,

changes in the far-UV CD are observed. A transition from an α-helical or a β-sheet structure to a disordered structure can be easily monitored by observing spectral changes or by simply following intensities at predetermined wavelengths as a function of increasing perturbation (e.g., temperature or denaturant concentration). Simple two- and multistate models exist which allow the data to be used to determine quantitative thermodynamic parameters for the transitions detected. These measurements are extremely useful for monitoring stabilities of protein pharmaceuticals. Although such measurements made at two or three wavelengths may not permit analysis of the exact nature of a structural transition, especially for proteins that contain significant contributions from all of the different types of secondary structure, they can certainly be used to reflect order–disorder transitions and provide initial estimates of overall conformational stability versus temperature or other perturbations. Accelerated thermal stability measurements using this technique have been applied in our laboratory to screen for stabilizers of aFGF, which, in the absence of a polyanionic stabilizer, unfolds and aggregates at temperatures slightly above ambient temperature (Volkin *et al.*, 1993). aFGF is a primarily β sheet-containing protein that binds heparin (Copeland *et al.*, 1991). Accelerated stability measurements employing CD, among other techniques, have enabled us to discover a number of polyanions which can be used as stabilizers in the pharmaceutically acceptable formulations of this protein (Tsai *et al.*, 1993a; Volkin *et al.*, 1993).

Determination of overall secondary structure content from CD spectra ideally requires that data with high signal-to-noise ratio be obtained in the short wavelength range of 180–240 nm, although for proteins containing relatively high concentrations (40% or higher) of α helix, a relatively good estimate of the helix content is often obtained by simply dividing the ratio of observed mean residue ellipticity at 222 nm by 30,000 deg cm^2/dmol (Chen and Yang, 1971). Reliable estimates of the four main types of structure require analyzing the entire CD spectrum (preferably containing data from as low as 180 nm) in terms of a basis set of proteins with independently determined (by X-ray crystallography) secondary structure content, and using one or more of the following algorithms: (a) the single value decomposition method (Manavalan and Johnson, 1987); (b) the self-consistent field method (Sreerama and Woody, 1993), and (c) a convex constraint analysis (Perczel *et al.*, 1991, 1992). These analyses do not, however, necessarily ensure that accurate values for secondary structure content will be obtained. A frequently encountered problem is that spectral contributions from aromatic side chains in the far-UV region are not accounted for in the structure analysis methods (Bolotina and Lugauskas, 1985; Manning and Woody, 1989). Furthermore, there are frequently practical difficulties in obtaining CD spectra with reasonable signal-to-noise

ratio below 200 nm in the presence of salt and other excipients that must often be present in a protein formulation. This problem becomes particularly severe below 190 nm, because of high absorption from most commonly used inorganic and organic components. It should be emphasized that far-UV CD spectroscopy in the 200- to 240-nm range still provides valuable information regarding conformational changes, and this is an important tool for analyzing protein stability and developing protein formulations, even though quantitation in terms of actual secondary structure content may not be reliable.

CD spectra of proteins in the near-UV absorbance range (250–320 nm) provide information on the local environment of phenylalanine, tyrosine, and tryptophan residues in the tertiary structure of a protein. The CD intensities (ellipticity) of the bands in the ranges 250–260, 270–280, and 290–300 nm are usually indicative of the extent of induced asymmetry in the environments of phenylalanine, tyrosine, and tryptophan residues, respectively. When a protein is unfolded, these aromatic bands disappear or are greatly reduced in intensity because their immediate asymmetric environments are disrupted. Molten globule or partially folded structures of proteins usually manifest very weak or no aromatic CD bands. The lack of a near-UV CD band while the far-UV CD spectrum remains unchanged relative to the fully folded protein is an almost certain indication of the formation of a molten globule or partially folded structure. In this regard, near-UV CD and tryptophan fluorescence spectroscopy usually provide complementary information. However, CD offers an advantage, in that changes in the phenylalanine microenvironment can be readily observed. Furthermore, in most proteins containing tryptophan and tyrosine residues, tyrosine fluorescence is not clearly resolved. Near-UV CD offers a convenient means for probing the local environment of tyrosine residues for such proteins. The disadvantage of the near-UV CD technique is its low sensitivity with regard to the protein concentration requirement compared to fluorescence (one to three orders of magnitude higher concentration than the few micrograms per milliliter needed for fluorescence measurements). A good example of the complementary use of tryptophan fluorescence and near-UV CD for analyzing functionally significant subtle structural changes around tyrosine and tryptophan chromophores is described in detail for the chimeric protein TP40 (Gress *et al.*, 1994).

2.4. Fourier-Transform Infrared Absorption Spectroscopy

The amide I absorption region (1600–1700 cm^{-1}) of the IR spectra of proteins has been extensively used during the last decade to estimate the secondary structure content. The recent explosion in the use of this technique

was brought about primarily by the following events: (a) the commercial availability of FTIR instrumentation, (b) the use of band narrowing techniques to enhance the resolution of the relatively broad amide I band (Kaupinnen *et al.*, 1981), and (c) the assignment of bands, following band fitting procedures, to ordered secondary structures based on the IR spectra of proteins for which high-resolution crystal structures are available (Byler and Susi, 1986). In general, β sheets and turns can be estimated with greater accuracy from amide I IR spectroscopy than from far-UV CD spectroscopy. In contrast, bands produced by the peptide backbone in disordered and α-helical conformations overlap extensively between 1645 and 1660 cm^{-1}, often prohibiting quantitative resolution of their individual contributions to the overall structure content. Therefore, far-UV CD spectroscopy frequently provides a more accurate estimate of the α-helix content of a protein than IR studies. In most IR solution studies, D_2O has been used as solvent in order to minimize the broad water absorption bands centered near 1640 cm^{-1}. The sensitivity of this technique is rather low in conventional transmission geometries requiring high protein concentrations, typically in the range of 1–10 mg/ml when D_2O is used as solvent and 10–100 mg/ml when H_2O is employed. FTIR spectroscopy has been recently extended to the study of protein folding and stability (reviewed in Middaugh *et al.*, 1995).

The potential use of this spectroscopic technique in the pharmaceutical development of protein formulations, however, goes far beyond its simple use as another "low-resolution" solution spectroscopic technique for monitoring secondary structure content and protein folding. This is because of the applicability of IR to different physical states of a sample, including dried and lyophilized materials. The effect of lyophilization on protein structure is a critical issue in the development of lyophilized formulations. Some proteins, for example, basic fibroblast growth factor and gamma interferon, have been found to irreversibly unfold upon lyophilization. Sugars have been observed, however, to be protective against conformational damage in some cases (Prestrelski *et al.*, 1993). These studies have employed use of KBr pellets and transmission geometry. It should be noted that KBr, at the very high concentrations used, could act as a chaotropic agent and care must be taken before assigning IR changes unambiguously to drying-induced structural perturbations of the proteins (Chan *et al.*, 1996). Two alternative approaches can be employed in this regard: (a) diffuse reflectance and (b) FTIR microscopy in either the reflectance or the transmission mode. An extensive study of the real-time kinetics of conformational changes of lysozyme during dehydration has been reported that employed attenuated total reflectance spectroscopy (Carpenter *et al.*, 1994). In this method, lyophilization was performed from a D_2O solution and the amide I′ and amide II′ bands were monitored as a function of decreasing temperature and dehydration.

In testing the effects of formulation conditions on a protein's secondary structure, it is necessary to be able to rapidly and quantitatively compare spectra. A recently introduced technique for quantifying the area of overlap of second-derivative spectra is very well-suited for this purpose (Kendrick *et al.*, 1996). This method is based on comparing area-normalized second-derivative spectra and assigning a correlation coefficient (originally introduced by Prestrelski *et al.*, 1993) in addition to a visual inspection (Fig. 2A). The Gdn·HCl-induced equilibrium unfolding profile of iso-1-cytochrome monitored by this technique has been found to be strikingly similar to that determined by CD (220 nm), as shown in Fig. 2B (Bowler *et al.*, 1993). In another pharmaceutically relevant application of FTIR spectroscopy, temperature-induced aggregation of chymotrypsinogen in D_2O solution has been systematically studied by Ismail *et al.* (1991), by monitoring the appearance of a relatively strong band at 1618 cm^{-1} and a weaker one at 1685 cm^{-1}. These bands, which have consistently been seen in aggregating protein systems, have been assigned to intermolecular β sheets and their presence permits a rapid identification of this phenomenon. As another example, the temperature-induced unfolding and refolding of ribonuclease A has been thoroughly studied in both H_2O and D_2O solutions (Fabian *et al.*, 1993).

It should be noted that the general validity of the resolution enhancement (Fourier self-deconvolution) and curve-fitting techniques most commonly employed for assignment of bands and their relative contributions to particular secondary structure types has been questioned (Surewicz *et al.*, 1993). In a series of studies, Venyaminov and Kalnin (1990a, b) identified the IR signals originating from the different amino acid side chains in the amide I and II absorption regions. In some cases, spectral contributions from certain side-chain residues need to be considered in secondary structure analyses. Interference from side-chain bands with peptide bond absorption, however, remains difficult to resolve quantitatively. In view of these difficulties in determining secondary structure content in terms of absolute percentages, the direct comparison of second-derivative spectra as described above is probably the most rapid and accurate way to test the effects of different formulation conditions on a protein's secondary structure (Kendrick *et al.*, 1996).

2.5. Differential Scanning Calorimetry

Differential scanning calorimetry (DSC) is the most direct technique available for measuring thermal unfolding and stability of proteins. Although

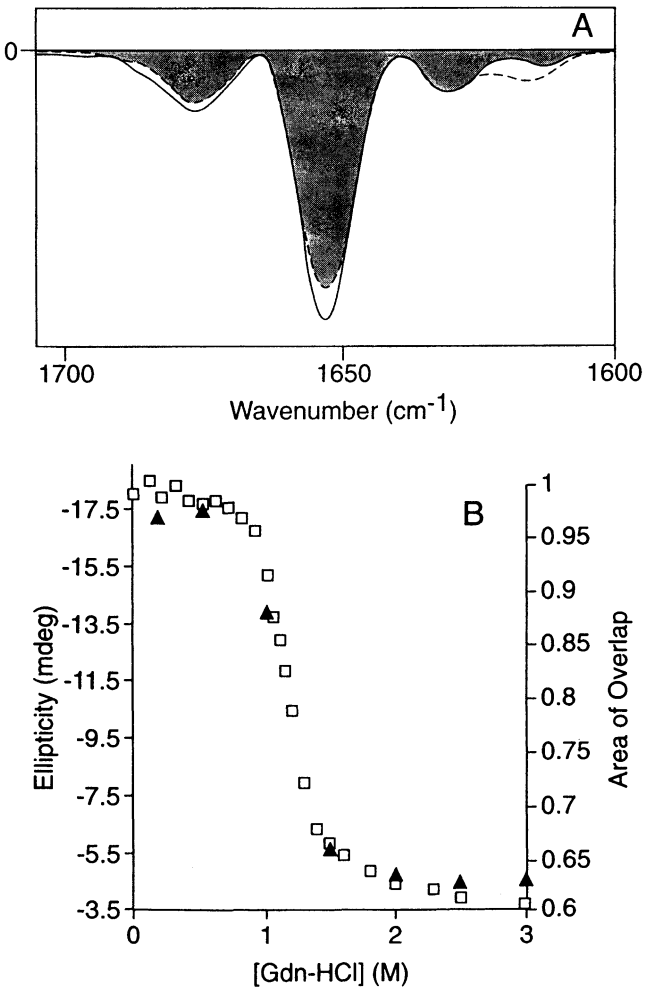


Figure 2. (A) Second-derivative area normalized FTIR spectra (amide I region) of γ-interferon in the "native" aqueous state (solid line) and rehydrated aqueous state following lyophilization (dashed line), the area of overlap being shown as the gray fill. (B) Guanidine hydrochloride-induced equilibrium unfolding transition of iso-1-cytochrome c as monitored by circular dichroism at 220 nm (open squares) and FTIR spectral area of overlap (filled triangles). Reprinted with permission from Kendrick *et al.* (1996). Copyright 1996 American Chemical Society.

thermodynamic calculations of the free energy of unfolding cannot be made when unfolding is irreversible, for example, due to aggregation of the unfolded state, a direct measure of the temperature range over which

unfolding occurs is provided by heat capacity changes as a function of changing temperature. Potential use of various excipients as protein stabilizers can be evaluated by simply monitoring their effects on the heat flow profile of a protein sample scanned from low to high temperatures. For example, the effects of different sugars, amino acids, and salts on the thermal transition temperature T_m of aFGF unfolding was studied by DSC (Tsai *et al.*, 1993a). The moderately stabilizing effect of these reagents on the thermal unfolding of this protein, as detected by DSC, is in agreement with CD results. As an example, the effect of trehalose on the DSC profile of aFGF is shown in Fig. 3. The effects of nonionic detergents on the thermal stability of recombinant human growth hormone (hGH) was recently studied by DSC (Katakam *et al.*, 1995). Surfactants did not protect against temperature-induced unfolding of hGH, as judged by the protein denaturation endotherms, although they protected the protein against interfacial aggregation. Furthermore, high concentrations of these agents actually destabilized hGH, as reflected in the decrease in the onset temperature for the peak in the endotherm.

The effect of moisture on the thermal stability of lyophilized proteins can also be examined by DSC. For example, the introduction of moisture to freeze-dried bovine somatotropin has been proposed to alter the mechanism of temperature-induced structural changes (Bell *et al.*, 1994). In this case, a kinetic DSC approach was used in which the scan rate dependence of the thermal transition was monitored. The use of equilibrium thermodynamics for interpreting calorimetric results should in general be avoided when the thermal change is irreversible, as is often the case for proteins (Lepock *et al.*, 1989; Sanchez-Ruiz *et al.*, 1988).

3. PROTEIN SIZE, QUATERNARY STRUCTURE, AGGREGATION STATE, AND SOLUBILITY

The large-scale preparation of protein pharmaceuticals requires fermentation and isolation procedures in which the protein of interest must be properly folded either during *in vivo* expression or as a separate *in vitro* refolding step. Unfortunately, aggregation reactions often compete with the proper folding of a protein, especially upon *in vitro* refolding from the unfolded or partially folded state. Protein aggregation may be mediated by covalent or noncovalent interactions, such as intermolecular disulfide bond formation and apolar associations. Aggregation is a concern not only during the expression and isolation of a protein, but also in subsequent formulation processing steps. For example, the use of lyophilization to stabilize labile

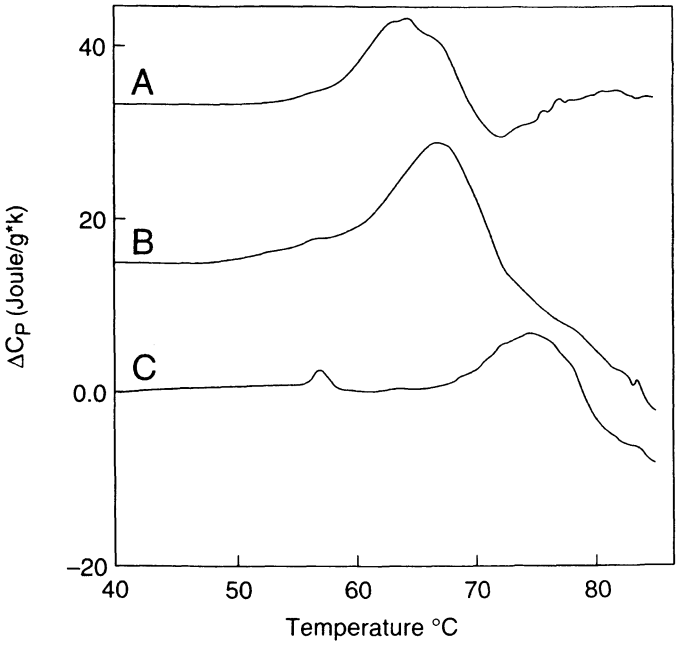


Figure 3. The effect of trehalose on the DSC endotherm of aFGF in the presence of $\frac{1}{3}\times$ heparin (by weight). All samples contained 1 mg/ml aFGF in a phosphate-buffered saline solution at pH 7.2. (A) No trehalose, (B) 0.5 *M* trehalose, (C) 1.5 *M* trehalose. Reprinted with permission from Tsai *et al.* (1993a). Copyright 1993 Plenum Press.

proteins during storage may also result in aggregation under certain conditions, because freezing and dehydration expose proteins to variations in pH and salt concentration. In addition, the aggregation of proteins storage in the dried state or during reconstitution has been observed (Constantino *et al.*, 1995). Aggregation of protein pharmaceuticals may lessen biological activity or increase the probability of an adverse antigenic response that will not occur with the native folded protein in its active monomeric or proper associative state (Cleland *et al.*, 1993). Thus, maintaining the biologically active form of a protein requires preserving its overall size, both by preventing aggregation and, in the case of oligomeric proteins, maintaining quaternary structure (Schein, 1994). A rational approach to the formulation of a stable pharmaceutical protein therefore involves a prior detailed understanding of a protein's overall size and quaternary structure as well as its association/aggregation behavior, in addition to its conformational stability and chemical integrity. Many methods are available for measuring the overall size of a protein and the

most commonly used techniques will be discussed below. Complete biophysical characterization of a native, multimeric protein or an inactive, aggregated protein will depend on obtaining data acquired from a combination of methods, because various techniques possess different detection and discrimination capabilities as well as complementarity in this regard.

3.1. Electrophoresis

Polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate (SDS–PAGE) is commonly used to assess the molecular weights of proteins. The excess of negatively charged sulfonate groups results in rodlike SDS/protein complexes of constant charge-to-mass ratio, which migrate proportionally to their molecular weight. Although this method is simple and rapid (especially with precast gels), the molecular weight estimate that is obtained is only accurate to approximately ±15% (Steele and Neilsen, 1978). If aggregation is mediated by disulfide bonds, SDS–PAGE can easily detect multimers if nonreducing conditions are employed. Noncovalent associations are difficult to detect, however, because exposure to SDS will dissociate and denature most proteins. Noncovalent aggregates can potentially be detected by covalently crosslinking the physically aggregated material with a suitable chemical reagent (e.g., glutaraldehyde) and then performing SDS–PAGE. Molecular weight estimates can be obtained with suitable standards, but discrimination of highly aggregated complexes is difficult. Large aggregates may not enter the gel matrix and may remain either undetected or merely seen as material at the gel top, whereas smaller compounds may migrate too rapidly and exit the gel undetected. Furthermore, heterogeneously aggregated proteins usually appear as a very broad, diffuse band, which also may complicate quantitation and comparisons. Because resolution and staining techniques (especially with silver) are not optimal, quantitation may also be troublesome.

Electrophoresis of chemically crosslinked oligomeric proteins has been successfully used, however, to produce a series of bands corresponding to species up to the largest oligomer present. This approach has been used to define subunit stoichiometries (Craig, 1988). For example, Herman *et al.* (1979) demonstrated varying ratios of monomer, dimer, and tetramer forms of lactate dehydrogenase by crosslinking the enzyme under different environmental conditions known to lead to inactivation or reactivation of the enzyme.

The use of capillary electrophoresis (CE) for the characterization of associative states of proteins offers many potential advantages compared to SDS–PAGE, such as attaining higher resolution in heterogeneous mixtures as well as more rapid analysis (minutes versus hours), yet in practice it also has some of the same limitations. Sample preparation is often similar, in which SDS and reducing agents are used in a manner analogous to SDS–PAGE. CE chromatography profiles can be monitored by a variety of methods, including absorbance, fluorescence, light scattering, and mass spectrometry (Kuhr, 1990). Fluorescence detection offers high sensitivity especially with the use of lasers and fluorescent derivatizing agents. Simultaneously, some limited information about protein structure can also be acquired using these types of detection (as described in previous sections) in addition to other techniques, such as Raman spectroscopy (Deyl and Struzinsky, 1991). Some currently available CE instrumentation allows collection of fractions, thus permitting identification and more extensive characterization of aggregated material. One disadvantage of CE is that the fine capillaries have the potential to become obstructed, thus preventing uniform mobility, especially if extensive protein aggregates are present. In summary, electrophoretic methods are readily performed, but yield qualitative or semiquantitative information. Therefore, further quantification and definition of protein aggregates is better suited to other analytical methods, as discussed below.

3.2. Chromatography

High-pressure size-exclusion chromatography (HP-SEC) and reversed–phase high–pressure liquid chromatography (RP-HPLC) are convenient methods for monitoring the size and quaternary structure of proteins by utilizing commonly available laboratory instrumentation. Chromatographic experiments are often well-suited to analysis of proteins in more complex formulations that contain polymers or other proteinaceous compounds that may interfere with spectroscopic methods of analysis. Under the appropriate chromatographic conditions, multimeric forms can be resolved, readily quantified, sized with the use of suitable standards, and isolated. These methods will generally yield different chromatographic profiles of the same protein preparation, because separation is based on different physical properties (i.e., size/shape and polarity).

Size-exclusion chromatography has the immediate advantage that elution position does suggest the presence of associated material. For example, as shown in Fig. 4, the HP-SEC elution of a partially aggregated

preparation of basic fibroblast growth factor (bFGF) suggests a single aggregated species, whereas the RP-HPLC elution profile of the same preparation suggests that more than one nonnative form of the protein is present (Shahrokh *et al.*, 1994). With either chromatography method, protein-containing fractions can be easily collected and then further analyzed to better define the nature of the aggregated species. In addition, amino acid analysis, N-terminal sequencing, or peptide mapping may be performed at this point to ensure identity of the individual species.

Along with CE, the most common methods of detection for chromatography are absorbance and fluorescence, with refractive index or light scattering less frequently employed. Although absorbance detection generally uses a single wavelength in either the near-UV (280 nm for aromatic side chains) or the far-UV region (215 nm for peptide bonds), sufficient sample concentrations and longer spectroscopic path lengths allow multiple-wavelength and even complete spectral data to be collected with currently available chromatography detectors. This capability enables single experiments to yield information about not only the size, but also the nature and conformational state of the molecules in the individual fractions. For complex mixtures even within individual fractions, multiple-wavelength absorbance detection may permit discrimination of proteins from other components, such as nucleic acids, or even resolution of individual proteins within single fractions if their spectroscopic properties are sufficiently different. Fluorescence monitoring employing multiple wavelengths can also yield information about the extent of protein unfolding by consideration of the position of a protein's intrinsic fluorescence emission maximum relative to that of the protein in its native state.

In addition to spectroscopic monitoring, the availability of light scattering detectors enables estimation of size parameters. Combining gel permeation chromatography with intensity-based, low-angle laser light scattering detection provides a powerful technique for quantifying and defining the size of aggregates. For example, Endo *et al.* (1992) effectively used this method to examine the effect of temperature and pH on the aggregation state of erythropoietin. In the absence of a specialized light scattering detection, a fluorescence detector can be employed by monitoring the light scattered at 90° at the incident wavelength (Dollinger *et al.*, 1992). Standards of varying molecular weight can be used for size estimation, but specific proteins behave differently than common standard proteins due to shape differences and as a result yield inaccurate size values (Middaugh *et al.*, 1992). In general, it is difficult to identify suitable standards for larger protein aggregates, polymers, or viral particles. In the case of spherical protein particles, such as some types of viral capsids, spherical latex beads of known size can be utilized. It is

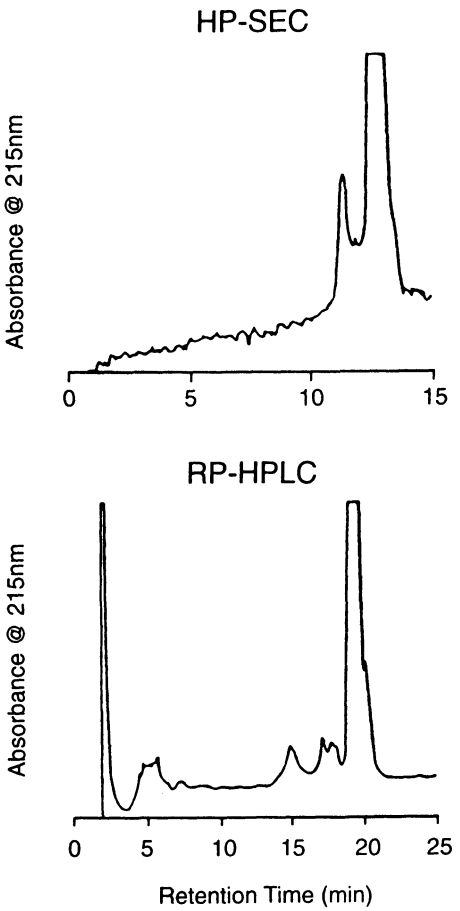


Figure 4. Elution profiles of a partially aggregated solution of bFGF using different chromatographic methods. Elution patterns corresponding to aggregate forms of the protein differ with the type of chromatography (HP-SEC or RP-HPLC) employed. Reprinted with permission from Shahrokh *et al.* (1994). Copyright 1994 American Pharmaceutical Association.

also possible to directly determine the hydrodynamic radius of a protein during chromatography by measuring fluctuations in the intensity of scattered light (dynamic light scattering) as well as radius of gyration and absolute molecular weights if the angular dependence of the scattered light is analyzed (see below). Specialized instrumentation has recently become available for both of these approaches in a chromatographic format.

3.3. UV Spectrophotometry

Although UV absorbance spectra are primarily used for protein mass quantitation, aggregation can be detected rather easily by changes in the absorbance baseline of the UV spectrum in nonabsorbing regions (e.g., >300 nm). Although it is only a qualitative technique (although in some cases the quantity "turbidity" can be quantitatively applied), the ready availability of UV spectrophotometers in many laboratories often enables one to rather easily obtain some information on the aggregation state of a protein preparation. Because light scattering is inversely proportional to the incident wavelength, protein aggregation results in an exponentially increasing baseline as the wavelength decreases (see Section 2.1). Consequently, baseline deviations due to light scattering are most evident in the UV region. One drawback is that the aggregates being monitored must be relatively large (>20 nm in diameter) before baseline changes are observed (Mach *et al.*, 1995). Kinetic data are easily collected on most common spectrophotometers, making this one of the most convenient methods for obtaining indirect information concerning the extent of protein aggregation over time (Eberlein *et al.*, 1994; Tsai *et al.*, 1993a; Volkin *et al.*, 1993). In such experiments, the appearance of optical density at an appropriate wavelength (e.g., 300–400 nm) is simply monitored as a function of time in the presence of aggregation-inducing stress, such as temperature or pH. Comparative experiments can then be performed in the presence and absence of pharmaceutical excipients to evaluate their stabilizing effects (Volkin *et al.*, 1993).

3.4. Analytical Ultracentrifugation

Although ultracentrifugation experiments are more time-consuming than the methods described above, they usually provide the most unambiguous determination of protein size and molecular weight under native solution conditions. Both sedimentation velocity and equilibrium sedimentation experiments can provide accurate measures of a protein's size and association state if the appropriate specialized instrumentation is available. Sedimentation velocity experiments can be performed in only a few hours, although absolute molecular weights cannot be obtained by this method. This technique also requires measurements (or estimates) of partial specific volume as well as diffusional or fractional coefficients. The sedimentation coefficient that is determined from such experiments also depends on the size and shape of the molecule or complex. In general,

molecular weights calculated from velocity of sedimentation involve approximations that yield data comparable in accuracy to those obtained from PAGE estimates (Van Holde, 1975). Nevertheless, velocity experiments do provide at least semiquantitative information on the aggregation state of proteins in solution. Equilibrium experiments offer a more accurate way to determine the molecular weight of a protein and its complexes, although a measure of the protein's partial specific volume is again needed.

Detection systems are available that use either optical absorbance or interferometry, such that a wide range of concentrations and sample conditions can be examined. Detection of sedimentation profiles by interferometry has the advantage of being able to monitor preparations that have little optical absorbance, but multicomponent formulations severely complicate this type of data and may make quantitative analysis impossible. In comparison, optical absorbance has the advantage of monitoring only the protein itself in a complex mixture, if none of the other components have significant absorbance in the region of interest. For example, the association of acidic fibroblast growth factor (aFGF) in the presence of the stabilizing ligand heparin was successfully monitored in an equilibrium sedimentation experiment using a wavelength of 280 nm (Mach *et al.*, 1993). Because protein association is mediated in this example by a polysaccharide, the association state of aFGF can be easily followed, because heparin does not absorb light at this wavelength. High concentrations of salts or buffers in a protein solution may prohibit using UV detection to monitor the sedimentation profile of a protein at lower, more sensitive wavelengths. Formulation excipients may also absorb strongly in other regions of interest, which may be either a hindrance or help, depending on the information desired.

In complexes which are confined to a small number of molecules, such as subunit or multimeric proteins, the extent of association and the corresponding association constants can be well defined. As mentioned above, this requires knowledge of the partial specific volume of the protein of interest, which can be estimated by calculation or directly obtained from solution density measurements. Software for such analysis is available from the instrument manufacturer or from many academic sites (contact http://bioc02.uthscsa.edu/biochem/xla.html on the World Wide Web to obtain listings for available analysis software). When the stoichiometry of the associated state is small (ca. 3), analysis is relatively straightforward and can reveal the relative amounts of monomeric and associated populations (M. L. Johnson *et al.*, 1981). For example, Green and *et al.* (1991) employed equilibrium sedimentation analysis to extensively characterize the quaternary structure of the enzyme arginase, demonstrating a reversible monomer–

trimer equilibrium with a defined dissociation constant. Interestingly, the addition of specific metal ions dramatically stabilizes the association of the monomers to the oligomeric state. Protein quaternary structure can also be modulated by less specific interactions, such as ionic strength. Sedimentation velocity experiments have also been used to show that the relative amounts of the four major oligomeric forms of casein kinase 2 can be altered by varying the sodium chloride concentration (Valero *et al.*, 1995). Analytical ultracentrifugation has been used to define the association constants, molecular weights, and shapes of immunoglobulin complexes (Liu *et al.*, 1995). When examining more heterogeneous populations of aggregates, data analysis of equilibrium sedimentation data becomes complicated, but can be simplified, assuming uniform aggregation constants for all multimers (Cole *et al.*, 1993). Finally, using equilibrium sedimentation experiments, the extent of covalent modification of proteins can also be monitored. For example, the degree of glycosylation in a preparation of human stem cell factor was determined by direct comparison to the unglycoslyated form of the protein (Arakawa *et al.*, 1991).

3.5. Light Scattering

Light scattering was described in previous sections in terms of its use as a chromatographic detection technique, but it can also offer a rapid and accurate method for characterizing the size and shape of proteins when appropriate instrumentation is employed. The most straightforward experiment (in theory and practice) employs static light scattering. The intensity of scattered light by a particle is a complex function of its size and shape, but certain limiting assumptions allow size information to be experimentally obtained. Maximum information can be deduced by measuring the scattered intensity of light as a function of angle, but it is most common to perform such studies at a fixed angle, usually either 90° or at some low angle where scattering is more intense (e.g., 135°). Although small particles show uniform scattering intensity at all angles, larger particles, such as extensively aggregated proteins, tend to produce lower scattering intensity as the detection angle is increased. Analysis of such data permits both molecular weights and radii of gyration to be obtained as well as second virial coefficients to be calculated (a useful quantity beyond the scope of our discussion here). As particles become very large, light scattering behavior becomes a very complex function of particle shape, and data interpretation becomes strongly model dependent.

In addition to static light scattering, particle size can also be determined by dynamic light scattering (also sometimes referred to as photon correlation or quasielastic spectroscopy). This technique relies on the fluctuations in time of the intensity of scattered light. The degree of correlation of these fluctuations depends on the diffusion coefficient of the scattering particle, a quantity which is simply related to hydrodynamic radius through the Stokes–Einstein equation. This mathematical treatment (a "cumulant analysis") of these data also permits the degree of homogeneity (polydispersity index) to be accessed as well as size distributions to be constructed. In ideal circumstances (particles of very different diameters), the sizes of two or three components in a mixture can be deconvoluted. A major experimental limitation in both static and dynamic light scattering experiments is interference by dust. Because scattering is dominated by larger particles, the presence of even trace quantities of such material must be avoided if interpretable data are to be obtained. The simple filtration of protein samples prior to analysis is often effective in minimizing interferences from dust particles, however, this type of sample preparation may also effect the distribution and quantity of the protein particles themselves.

Although light scattering techniques offer the ability to obtain accurate molecular weights of relatively small macromolecules as well as large aggregates, perhaps its major use in preformulation work is as an empirical tool for investigating the effect of excipients upon the aggregation state of proteins. For example, characterization of the aggregation of insulin under varying conditions of concentration, pH, and ionic strength played a key role in formulation development and even permitted an estimate of critical association constants (Kadima *et al.*, 1993). Cleland and Wang (1990) used light scattering to characterize the aggregation of carbonic anhydrase during refolding as well as to optimize refolding conditions. Similar studies have been carried out with acidic fibroblast growth factor and were a key to the development of a stable formulation of this very unstable protein (Mach *et al.*, 1993; Middaugh *et al.*, 1992). Simple titration of pharmaceutical excipients into a protein preparation and simultaneous monitoring of light scattering allows formulation components to be conveniently evaluated in a relatively short period of time for their ability to inhibit protein aggregation. In addition, this approach can be used to monitor the interaction and association of protein–ligands (Mach *et al.*, 1993). As indicated above, because light scattering is exponentially proportional to the particle diameter, very small amounts of a larger species (e.g., aggregates) can easily be detected. Conversely, small amounts of large material will dominate the scattering intensity and mask the contributions of smaller sized populations. In such cases, other methods, such as SEC-HPLC or

analytical ultracentrifugation, should be used in conjunction with light scattering measurements to ensure an accurate description of particle size distribution.

3.6. Mass Spectrometry

Compared to the methods discussed previously, mass spectrometry provides the most accurate molecular mass value. Electrospray (ES) and laser desorption time-of-flight (LD-TOF) mass spectrometry (MS) have emerged as powerful techniques well-suited for the analysis of pharmaceutical proteins (Hillenkamp and Karas, 1990; Smith *et al.*, 1991). ESMS offers higher resolution and accuracy than LD-TOF MS and is therefore the method of choice for detecting chemical modifications that result in relatively small mass changes (see Section 4.2). Covalent aggregates of large proteins, such as serum albumin, can be characterized as well (Smith *et al.*, 1991). However, a homogeneous sample is usually required to take maximum advantage of the technique. This will usually involve purification or separation before introduction into the instrument. Furthermore, instrumentation is more expensive to maintain and more difficult to operate compared to LD-TOF apparatus.

LD-TOF MS is typically more sensitive in the sense that it typically requires less sample (1–10 pmol is optimal). Moreover, sample homogeneity is not required, because only a few peaks are generated by an individual protein, in contrast to the complex charge spectrum seen in ESMS with similar samples. Furthermore, the simplicity of commercially available instrumentation and its operation has allowed this method to become widely accessible to process and formulation scientists. This method requires, however, the presence of a laser absorbing matrix to induce volatilization, which, along with the desorption process itself, could dissociate aggregated or oligomeric proteins, thereby precluding the ability to obtain the desired information on the aggregation state of the protein in solution (Hillenkamp and Karas, 1990). Nevertheless, in some cases, multimeric states can be observed with careful sample preparation and analysis. For example, as shown in Fig. 5, Suter *et al.* (1992) were able to observe tetrameric as well as smaller associated forms of interleukin-2 (IL-2). The use of crosslinking reagents should also enable noncovalently associated, multimeric states of proteins to be observed, as previously described with SDS–PAGE (Farmer and Caprioli, 1991). In general, protein aggregates having a mass of less than 200 kDa can reliably be identified with this technique, although relative peak signal intensities may not be proportional to the populations of each

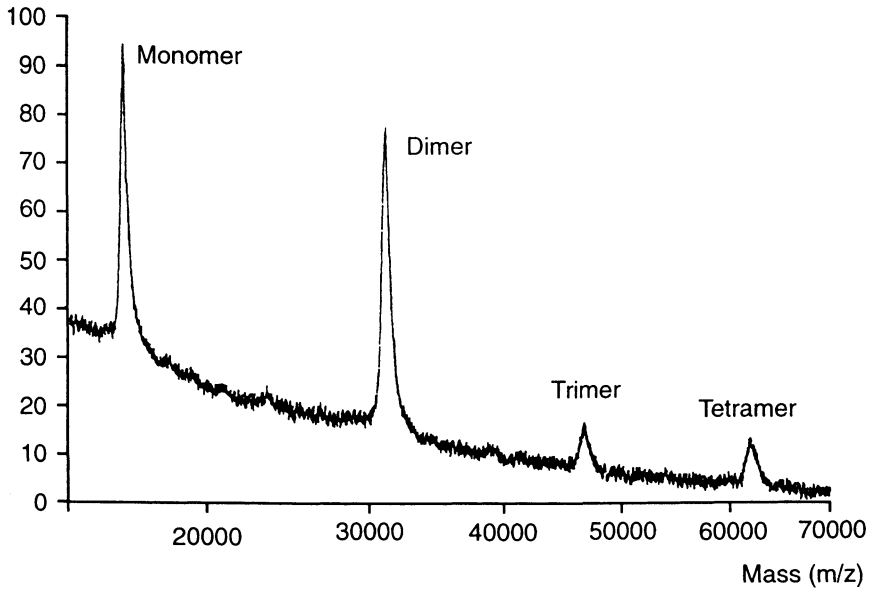


Figure 5. LD-TOF MS of aggregated forms of recombinant IL-2. The spectrum was acquired using 20 pmol of IL-2 in a sinapinic acid matrix. The oligomerized protein peaks can be eliminated by reduction and blocking of cysteine residues (not shown). Reprinted with permission from Suter *et al.* (1992). Copyright 1992 Academic Press.

particular species, because of potential differences in desorption or detection efficiency. Thus, quantitation is best evaluated by the methods previously described, such as chromatography or analytical ultracentrifugation.

3.7. Other Methods

The presence of aggregates can often be detected by a number of other optical techniques commonly used to study protein conformation and structure. The consequences of light scattering on UV absorption spectra have already been discussed, and a similar effect may be seen in circular dichroism spectra. Depending on the scattering characteristics of the high molecular weight material, spectral peaks may decrease in intensity due to absorption flattening (a particle screening effect). In the event of preferential scattering of one type of circularly polarized light, peaks may also shift position (Wallace and Mao, 1984). Fourier-transform infrared spectra of aggregated proteins usually manifest additional peaks

that can be attributed to intermolecular β structure (Dong *et al.*, 1995; van Stokkum *et al.*, 1995). In addition, fluorescence spectra may also show decreased emission due to scattering and an inner-filter effect in which densely packed chromophores prevent uniform exposure to incident light (Lakowicz, 1983). Furthermore, time-resolved fluorescence experiments can be employed to characterize protein aggregation (Brochon *et al.*, 1993).

Nonspectral methods can also be used to determine protein size and molecular weight. For instance, osmotic pressure measurements can provide an accurate measurement of molecular weight if the concentration of the protein in solution is accurately known. Another technique gaining increased acceptance is field flow fractionation (Giddings, 1989). Litzén *et al.* (1993) showed that field flow fractionation of a monoclonal antibody preparation provided better resolution of various aggregated protein species than did gel permeation chromatography.

3.8. Solubility

A critical physical property of a protein which is often neglected in a quantitative sense, but is of manifest importance in a practical pharmaceutical one is aqueous solubility. Low solubility often presents a significant obstacle during formulation development, and must be directly addressed at an early stage of development. Measurement of solubility is often complicated by the tendency of proteins to form poorly defined suspensions and gels. These phenomena do not permit an equilibrium solubility to be directly determined by simple addition of dry (e.g., lyophilized) protein to solvent until saturation is obtained (via measurement of protein concentration or, ideally, thermodynamic activity, in the saturated liquid phase). An alternative method employs sequential precipitation of protein by addition of high molecular weight (e.g., 8 kDa) polyethylene glycol (PEG). Plots of the logarithms of protein solubility versus PEG concentration are often linear and an estimate of the activity of the protein in the absence of PEG can be obtained by extrapolation to zero PEG content. For reasons described elsewhere (Middaugh and Volkin, 1992), however, this approach is best used in a comparative manner. Further information about the protein solubility problem can be found in Middaugh and Volkin (1992).

An indirect method that has been increasingly employed to monitor the effect of solution conditions on protein solubility uses turbidity (light scattering) measurements. In this approach, it is most common to follow

the rate of turbidity appearance as a function of variables such as pH, temperature, and excipient. Because only a simple absorption spectrophotometer (or fluorometer, if scattering at right angles is monitored) is necessary for such studies, this method has been extensively employed in this early screening phase of formulation development (Volkin *et al.*, 1993). Again, however, its primary use is as a comparative measure, because no absolute values of solubilities are derivable from such methods.

4. DEGRADATIVE COVALENT REACTIONS AND THEIR EFFECT ON PHYSICOCHEMICAL AND BIOLOGICAL PROPERTIES OF PROTEINS

Our discussion concerning preformulation studies of protein pharmaceuticals has focused on the effect of environmental conditions on these macromolecules' conformational integrity and stability. Exposure to elevated temperature, acidic/basic conditions, extreme redox potentials, or high ionic strength can all lead to conformational changes in protein structure (local alterations in secondary or tertiary structure), reversible unfolding (cooperative loss of most higher ordered structure), or irreversible alterations (permanent changes in the structural or chemical integrity of the protein molecule). The interplay between changes in protein structure able to facilitate chemical alterations of amino acid side chains (or peptide backbone) and the effect of degradative covalent reactions on the loss of native protein structure should be a primary focus of preformulation studies. For example, partial unfolding of a protein molecule leads to both increased solvent accessibility of previously unexposed amino acid residues as well as increased flexibility of the polypeptide chain. In many cases these types of effects facilitate the reactivity of both peptide bonds and amino acid side chains. In fact, the partial loss of conformational integrity is frequently observed to be the first step in the irreversible inactivation of proteins and enzymes at high temperatures (Volkin and Klibanov, 1989; Volkin and Middaugh, 1992). Conversely, covalent reactions altering either the amino acid side chains or the polypeptide backbone may lead to the loss of conformational stability. For example, the reduction of disulfide bonds or hydrolysis of polypeptide chains can induce protein unfolding leading to aggregation and inactivation. Obviously, a coupled interaction between these two processes can also occur and further complicate the goal of a complete temporal, physicochemical description of a protein pharmaceutical.

4.1. Covalent Alterations in Protein Structure Commonly Observed during Formulation and Storage

During formulation and *in vitro* storage, protein pharmaceuticals in liquid solution are typically exposed to pH values ranging from 4 to 8 and temperatures of 2–37°C for time periods up to 1–2 years. Under these conditions, the most commonly observed changes in amino acid side chains are deamidation and isoaspartate formation of Asn and Asp residues and the oxidation of Met and Cys residues (Powell *et al.*, 1996). The mechanistic details of these reactions have been studied extensively in both model proteins and peptides as well as protein therapeutics over the past decade. Many excellent reviews can be consulted for a detailed discussion of the reaction pathways of protein deamidation and isoaspartate formation (Aswad and Guzzetta, 1995; Bischoff and Kolbe, 1994; Clarke *et al.*, 1992; Wright, 1991) as well as protein oxidation reactions (Li *et al.*, 1995; Nguyen, 1994).

Briefly, the spontaneous deamidation of Asn residues under typical formulation conditions proceeds via an intramolecular mechanism as outlined in Fig. 6. The asparaginyl side-chain carbonyl group undergoes nucleophilic attack from the peptide bond nitrogen adjacent to the asparaginyl side chain, resulting in ring closure and release of ammonia. The resulting five-membered ring (succinimide or cyclic imide) is susceptible to hydrolysis, leading to a mixture of aspartyl and isoaspartyl residues. The isomerization of aspartyl residues to isoaspartyl can also occur via the same cyclic imide intermediate mechanism. The primary structure (presence of specific neighboring groups) as well as the conformational flexibility of the local higher ordered structure surroun-

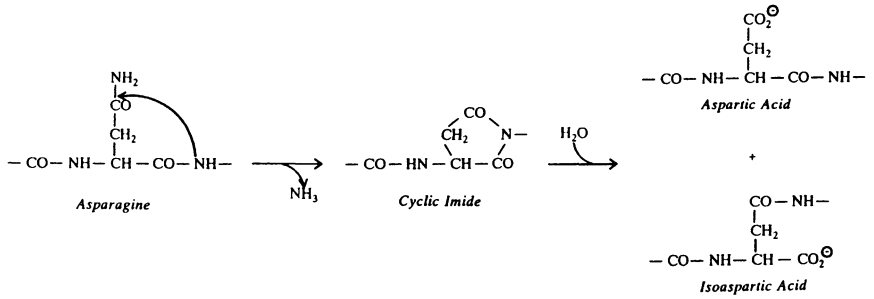


Figure 6. Deamidation and isoaspartate formation at asparginyl residues in proteins and peptides. Reprinted with permission from Volkin *et al.* (1995b). Copyright 1995 Humana Press.

ding the asparaginyl and aspartyl residues affect the rate at which deamidation and isoaspartate formation can occur. In general, Asn–Gly and Asn–Ser are the most labile sequences as a consequence of the minimum steric hindrance to cyclic imide formation provided by absent and small Gly and Ser side chains, respectively. Increased conformational flexibility also enhances deamidation by providing conformational mobility, which permits the peptide bond to more readily participate in the cyclic imide formation.

The oxidation of protein amino acid side chains under typical formulation conditions occurs primarily at the sulfur-containing amino acid residues cysteine and methionine. Protein-bound cysteine residues can oxidize to form inter- or intramolecular disulfide bonds or monomolecular byproducts, such as sulfenic and sulfinic acid (only harsher oxidation conditions can lead to irreversible formation of sulfonic acid):

$$\underset{\text{Cysteine}}{2\ \text{Protein-CH}_2\text{SH}} \rightarrow \underset{\text{Cystine}}{\text{Protein-CH}_2\text{S-SCH}_2\text{-Protein}}$$

$$\underset{\text{Cysteine}}{\text{Protein-CH}_2\text{SH}} \rightarrow \underset{\text{Sulfenic acid}}{\text{Protein-CH}_2\text{SOH}} \rightarrow \underset{\text{Sulfinic acid}}{\text{Protein-CH}_2\text{SO}_2\text{H}} \rightarrow \underset{\text{Sulfonic acid}}{\text{Protein-CH}_2\text{SO}_3\text{H}}$$

Similarly, the oxidation of protein–bound methionine residues results in the conversion of the thioether to its sulfoxide counterpart (whereas only more severe oxidation conditions lead to the formation of the sulfone):

$$\underset{\text{Methionine}}{\text{Protein-CH}_2\text{CH}_2\text{SCH}_3} \rightarrow \underset{\text{Methionine sulfoxide}}{\text{Protein-CH}_2\text{CH}_2\text{S(O)CH}_3} \rightarrow \underset{\text{Methionine sulfone}}{\text{Protein-CH}_2\text{CH}_2\text{S(OO)CH}_3}$$

Protein oxidation can occur through two general mechanisms, which involve (1) site-specific (metal-catalyzed) and (2) non-site-specific oxidation. Metal-ion-catalyzed oxidation occurs near protein metal-ion-binding sites, whereas non-site-specific oxidation is initiated by light or exogenous oxidants generated from solubilized metal ions or impurities from buffer salts or formulation excipients (Li *et al.*, 1995). The rate and extent of the metal-catalyzed oxidation of proteins are greatly influenced by the primary structure of a protein (metal-ion-binding sites). The location of non-site-specific oxidation of Met and Cys residues within proteins often correlates with surface and solvent exposure of the reactive amino acid residues.

Upon exposure to elevated temperatures at neutral pH or to acidic/basic conditions, other amino acid residues can also begin to degrade (Manning *et al.*, 1989; Volkin and Klibanov, 1989; Volkin and Middaugh, 1992; Powell, 1994; Volkin *et al.*, 1995b). At acidic pH, the hydrolysis of the peptide bond can occur nonenzymatically at Asp residues with particular lability noted at Asp–Pro sequences. In addition, deamidation of Asn

and Gln can also occur through an acid-catalyzed hydrolytic mechanism (Wright, 1991). At neutral pH, racemization at Asp-X and pyroglutamic acid formation at N-terminal Gly residues can occur, especially at elevated temperatures (Powell, 1994). In addition, nonenzymatic cyclization of the first two amino acid residues at the N-terminus and subsequent cleavage can result in the formation of a diketopiperazine and a truncated protein. Recombinant human growth hormone undergoes this reaction when incubated in the dry state at 50°C for 4 days at pH 7.4 (Battersby *et al.*, 1994). Prolonged exposure to basic conditions renders disulfide bonds chemically unstable, leading to a variety of degradation products, including beta-elimination reactions and disulfide interchange. In general, high pH accelerates both the oxidation of cysteine residues and the deamidation of Asn residues as discussed above. Exogenous oxidants can lead to the degradation of aromatic amino acid residues such as histidine, tryptophan, and tyrosine. The free amino groups of lysine residues as well as the N-terminus can react with reducing sugars to form condensation products via a series of chemical reactions known as Malliard reactions. Again, we describe these degradative covalent reactions only in passing; numerous reviews have recently appeared describing them in more depth (Aswad and Guzzetta, 1995; Bischoff and Kolbe, 1994; Clarke *et al.*, 1992; Cleland *et al.*, 1993; Li *et al.*, 1995; Nguyen, 1994; Powell, 1994; Volkin *et al.*, 1995b).

4.2. Methods for Detecting Covalently Altered Proteins and Effect of Covalent Modification on Protein Conformational Integrity and Biological Activity

Analysis of the chemical degradation of proteins can be performed at the level of the intact protein or at the individual peptide level after enzymatic cleavage of the protein. In intact proteins, one can determine to what extent and at what rate these reactions have occurred. Furthermore, one can isolate the chemically altered protein to ascertain if the chemical alteration perturbs its physical or biological properties. These types of studies should be the focus of preformulation characterization of protein pharmaceuticals. Analytical tools such as isoelectric focusing, ion exchange chromatography, and matrix-assisted laser desorption time-of-flight (MALD-TOF) mass spectrometry are commonly utilized to examine intact proteins for chemical changes. At the peptide level, identification of the specific site of the covalently altered amino acid can be more rigorously quantified (Cottrell, 1994). Proteins are enzymatically digested and the peptides analyzed directly by mass spectrometry. Alternatively, the peptides can be separated by reversed-phase

chromatography or capillary electrophoresis. The chromatographs of the individual peptides (peptide map) from native and chemically altered proteins are then compared to identify which peptide contains the altered amino acid and the peptide(s) identified by amino acid sequencing. Currently, identification of the peptide is most often made "on-line" during the RP-HPLC employing electrospray mass spectrometry detector (LC/MS) (Chait and Kent, 1992; Siuzdak, 1994). Alternatively, the intact protein can be either chemically or enzymatically modified at the site of specific types of chemical changes in the protein, permitting either direct the detection of covalent changes in the intact protein or the tagging of the corresponding peptide for identification during peptide mapping studies. For example, as discussed above, the deamidation of Asn results in the formation of both aspartic acid and isoaspartic acid residues. The carboxylic acid residues of either residue can be chemically esterified with methanolic hydrochloric acid, whereas the isoaspartate residues can be specifically labeled with S-adenosyl-L-(methyl-^{3}H)-methionine by the enzyme protein isoaspartyl methyltransferase. These approaches have been applied successfully to identify the sites of deamidation in human growth hormone and hirudin (Aswad and Guzzetta, 1995; Bischoff and Kolbe, 1994).

Preformulation studies typically emphasize the detection of chemical alterations of protein pharmaceuticals at the intact protein level, thereby allowing the effects of covalent changes on the activity and conformational integrity of proteins to be examined. In the case of deamidation of asparagine residues, a neutral asparagine residue is converted to charged aspartic acid with the concomitant release of ammonia (see above). Analytical techniques that separate protein molecules on the basis of charge are therefore commonly used to monitor deamidation. Isoelectric focusing electrophoresis (Gianazza, 1995), ion-exchange chromatography, and capillary electrophoresis are frequently used to monitor the accumulation of chemically modified charged species.

For example, a combination of these techniques has been used to characterize the deamidation-induced microheterogeneity of monoclonal immunoglobulins (Tsai *et al.*, 1993b). As can be seen in Fig. 7, the identification of varying degrees of chemically altered antibody was ascertained by a combination of anion exchange chromatography and agarose gel isoelectric focusing. The cause of microheterogeneity cannot be readily determined, because any number of posttranslational events or chemical modifications can lead to the observed charge differences. In this case, deamidation of the immunoglobulin was identified as the cause of the microheterogeneity by a variety of direct and indirect experimental observations. Differential glycosylation was ruled out by a combination of carbohydrate analysis and MALD-TOF mass determinations of the isolated

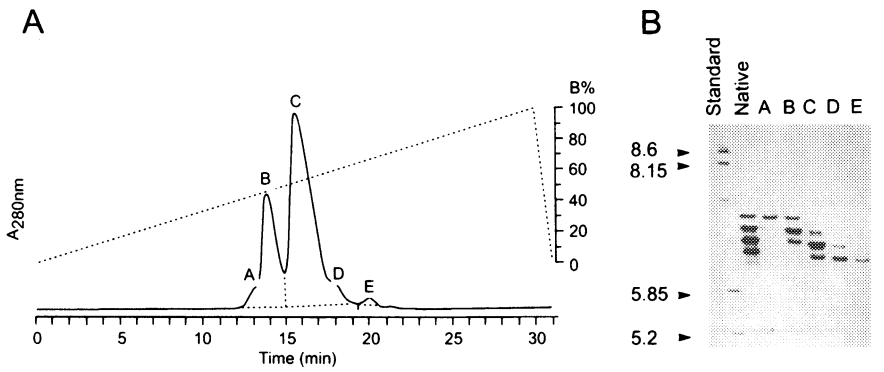


Figure 7. Deamidation-induced isoforms of h1B4 monoclonal antibody as measured by (A) chromatographic separation by anion exchange chromatography and (B) isoelectric focusing. Reprinted with permission from Tsai *et al.* (1993b). Copyright 1993 Plenum Press.

species. In addition, during incubation under conditions favoring deamidation, the isoelectric focusing bands were observed to accumulate more acidic species. This process also correlated with the kinetics of ammonia release as determined by an enzymatic assay (Tsai *et al.*, 1993b). Each charged species separated by anion exchange chromatography was collected and then analyzed by a variety of biochemical and biophysical techniques. This particular deamidation reaction was found to have no effect on either the overall conformation of the antibody as determined by circular dichroism and fluorescence spectroscopy or the binding activity of antibody as measured by a competitive binding assay.

Similar results have been observed with the deamidation of aFGF, in which the initial deamidation of Asn residues does not affect the overall conformation, conformational stability (as analyzed by thermal unfolding experiments), or mitogenic (biological) activity of the protein. One interesting challenge with this protein was that aFGF tightly binds highly negatively charged polyanions such as the sulfated polysaccharide heparin and these polyanions are required as stabilizers during formulation of the protein to prevent unfolding and aggregation during storage (Volkin and Middaugh, 1996). Due to the high charge density of the complex, conventional isoelectric focusing (IEF) cannot distinguish multiply deamidated, species, because a deamidation event does not dramatically alter the complex's charge density. An IEF-urea method was developed to unfold the protein and dissociate the aFGF-heparin complex, thus allowing quantitation of the deamidation events in aFGF during storage in the presence of polyanions (Volkin *et al.*, 1995a).

In contrast to the previous examples, deamidation of Asn residues has been shown to cause some degree of biological inactivation or loss of activity in a variety of other proteins and enzymes (Teshima *et al.*, 1995). Recombinant soluble CD4 was observed to have reduced binding capacity for the human immunodeficiency virus-1 (HIV-1) envelope glycoprotein gp120 due to a specific deamidation of Asn-52 residue in the V1 domain of the molecule. This chemical change was caused by storage of the protein in the liquid state for 6 months at 25°C at pH 7.2 (Teshima *et al.*, 1991). Interestingly, screening the deamidated protein with a panel of conformational antibodies indicated that the overall conformation of the protein was unchanged. This result suggests that Asn-52 of soluble CD4 might be directly involved in the binding of the protein to gp120. Deamidation has also been shown to contribute to the irreversible thermal inactivation of ribonuclease A, in which the enzymatic activity was shown to decrease with increasingly deamidated species (isolated from preparative IEF gels) (Zale and Klibanov, 1986). More recently, ribonuclease A specifically deamidated at Asn-67 was resolved into aspartic acid- and isoaspartic acid-containing protein fractions by hydrophobic interaction chromatography (Di Donato *et al.*, 1993). It was shown that the deamidation at Asn-67 to either Asp-67 or isoAsp-67 lowered the catalytic activity of the enzyme. The effect of deamidation on the renaturation of the unfolded and reduced enzyme was then examined. The Asp-67 deamidated species refolded at a rate equivalent to the native enzyme, whereas the Isoasp-67 molecule refolded 50% more slowly. These results directly demonstrate that deamidation of an asparagine residue can affect both the biological activity and the physical properties of a protein.

The detection of protein oxidation reactions such as the conversion of methionine to its sulfoxide or the formation of disulfide bonds from cysteine residues can also be identified at the intact protein level. Unlike the deamidation reaction described above, these oxidation reactions do not necessarily alter the charge state of the protein, but instead often lead to measurable molecular weight differences between native and chemically altered proteins. For example, intermolecular disulfide bond formation can be easily followed by accumulation of oligomer bands on unreduced SDS–PAGE. SDS–PAGE under reducing conditions is typically run at the same time to confirm the disulfide-linked nature of the higher molecular weight species. MALD-TOF mass spectrometry can also be utilized to detect the 2-Da mass change resulting from disulfide bond formation, although this analysis is typically used more for peptides than larger proteins because of its increased sensitivity for the former (Cotter, 1992). The oxidation of methionine residues results in a mass change of 16 Da, which can be more easily detected by MALD-TOF mass spectrometry (Geisow, 1992). For example, the H_2O_2-induced oxidation of the two methionine residues of relaxin was

successfully quantitatively characterized by mass spectrometric analysis of the intact protein (Nguyen *et al.*, 1993).

Protein oxidation products are typically isolated by reversed-phase or hydrophobic interaction chromatography, so that the effect of oxidation on the conformation and activity of purified, altered protein can be examined. For example, the autooxidation of three free cysteine residues of aFGF leads to the formation of intra- and intermolecular disulfide species as identified by SDS–PAGE migration patterns. Using a site-directed Cys to Ser mutant (containing two of the three cysteine residues found in native aFGF), intramolecular disulfide species were generated upon incubation with copper salts and then purified by reversed-phase HPLC (Linemeyer *et al.*, 1990). The oxidized form of aFGF was inactive in a cell-based mitogenic assay, although mitogenic activity was fully recovered upon reduction of the oxidized species with dithiothreitol. As another example, human growth hormone contains three methionine residues whose rate of reactivity toward oxidation by H_2O_2 correlates with their solvent-accessible surface area (Nguyen, 1994). Partially oxidized protein retains most of its biological activity, whereas complete oxidation of the three methionine residues leads to the loss of most of the biological activity of the protein. Interestingly, conformational studies by circular dichroism show that partially oxidized human growth hormone retains its native secondary structure, whereas fully oxidized protein exhibits an altered structure (reviewed by Nguyen, 1994). No simple or general correlation, however, exists between methionine surface exposure and reactivity toward oxidation, probably due to the fact that metal-catalyzed oxidation reactions can occur in hydrophobic protein interiors if metal-ion-binding sites are present (Li *et al.*, 1995).

It is clear from these studies that no straightforward correlation exists between the chemical lability of amino acid side chains and the conformational integrity and biological activity of proteins. In the case of Asn deamidation and Met oxidation, numerous examples of specific covalent changes have been documented in proteins, some leading to complete inactivation and others having little or no effect (Powell *et al.*, 1996). The major emphasis during the preformulation stage of protein drug development should generally be to characterize the interplay of the conformational integrity of a protein and the degradation of the labile amino acid residues described above. The effect of these covalent reactions on the biological activity must also be considered at these very early stages of preformulation studies. Although space does permit us to consider it here, the role of specific bioassays in the analysis of protein-based pharmaceuticals is critical in this regard. Even if these reactions are found to have no effect on the physicochemical or biological properties of a protein, the quantitative

characterization of the location and kinetics of deamidation and oxidation reactions occurring during storage must still be performed in order to ensure process and product consistency. Furthermore, the effect of these degradative reactions on the immunogenicity and pharmacokinetics of a protein drug as measured by animal and potentially human studies is also of utmost importance during the development of protein-based pharmaceuticals (Cleland *et al.*, 1993). These studies, however, are typically beyond the scope of the initial preformulation work.

5. CONCLUSIONS

Preformulation studies of protein pharmaceuticals entail a series of detailed biophysical and biochemical characterization experiments with the bulk drug substance in the presence and absence of various formulation excipients. In this chapter, we reviewed various experimental techniques commonly applied to quantitatively monitor the primary, secondary, tertiary, and higher order structure of protein molecules. The interplay between a protein's chemical integrity and higher order structural stability is a key parameter to be followed during these studies. By determining the effect of various environmental factors such as pH, temperature, and ionic strength on the conformational and covalent integrity of a protein molecule, an extremely useful set of protein stability "maps" can be generated, thereby providing profiles of the conformational and covalent integrity (and ideally biological activity) of a protein molecule as a function of environmental conditions. These data can be then be applied to the subsequent formulation development steps required both to determine the composition of the final drug product (dosage form, excipient composition, and container/closure) and ensure adequate storage stability (typically 2 years at 2–8°C).

A detailed determination of the biophysical and biochemical properties of a protein pharmaceutical enables the proper setup and performance of the formulation processing steps in the manufacturing environment by guiding and shaping the engineer's design criteria. Thus, carefully obtained preformulation data provide the type of information required to prevent stability problems from occurring during the large-scale production of final container-formulated protein pharmaceuticals. By investing the time and effort to obtain a detailed understanding of the causes and mechanisms of protein instability at the molecular level, the formulation scientist is, at a minimum, buying an insurance policy. This protection is quite valuable if or when something unexpected or atypical arises during formulation development scaleup work or during the actual manufacturing of a protein

pharmaceutical, because it provides the tools and information necessary to solve problems in an expeditious manner.

REFERENCES

Arakawa, T., Yphantis, D. A., Lory, J. W., Narhi, L. O., Lu, H. S., Prestrelski, S. J., Clogston, C. L., Zsebo, K. M., Mendiaz, E. A., Wypych, J., and Langley, K. E., 1991, Glycosylated and unglycosylated recombinant derived human stem cell factors are dimeric and have extensive secondary structure, *J. Biol. Chem.* **266:** 18942–18948.

Aswad, D. W., and Guzzetta, A. W., 1995, Methods for analysis of deamidation and isoaspartate formation in peptides and proteins, in: *Deamidation and Isoaspartate Formation in Peptides and Proteins* (D. W. Aswad, ed.), CRC Press, Boca Raton, PL, pp. 7–29.

Battersby, J. E., Hancock, W. S., Canova-David, E., Oeswein, J., and O'Connor, B., 1994, Diketopiperazine formation and N-terminal degradation of recombinant human growth hormone, *Int. J. Peptide Protein Res.* **44:**215–222.

Bell, L. N., Hageman, M. J., and Bauer, J. M., 1995, Impact of moisture on thermally induced denaturation and decomposition of lyophilized bovine somatotropin, *Biopolymers* **35:**201–209.

Bischoff, R., and Kolbe, H. V. J., 1994, Deamidation of asparagine and glutamine residues in proteins and peptides: Structural determinants and analytical methodology, *J. Chromatogr. B* **662:**261–278.

Bolotina, I. A., and Lugauskas, V. Yu, 1985, Detemination of the secondary structure of proteins from the circular dichroism spectra. IV. Consideration of the contribution of aromatic amino acid residues to the circular dichroism spectra of proteins in the peptide region, *Molek. Biol.* (USSR) **19:**1409–1421.

Bowler, B. E., May, K., Zaragoza, T., York, P., Dong, A., and Caughey, W. S., 1993, Destabilizing effects of replacing a surface lysine of cytochrome *c* with aromatic amino acids: implications for the denatured state, *Biochemistry* **32:**183–190.

Brahms, S., and Brahms, J., 1980, Determination of protein secondary structure in solution by vacuum ultraviolet circular dichroism, *J. Mol. Biol.* **138:**149–178.

Brochon, J. C., Tauc, P., Mérola, F., and Schoot, B. M., 1993, Analysis of arecombinant protein preparation on physical homogeneity and state of aggregation, *Anal. Chem.* **65:**1026–1034.

Buckler, D. R., Haas, E., and Scheraga, H. A., 1995, Analysis of the structure of ribonuclease A in native and partially denatured states by time-resolved nonradiative dynamic excitation energy transfer between site-specific extrinsic probes, *Biochemistry* **34:**15965–15978.

Byler, D. M., and Susi, H., 1986, Examination of the secondary structure of proteins by deconvoluted FTIR spectra, *Biopolymers* **25:**469–487.

Carpenter, J. F., Prestrelski, S. J., Anchordoguy, T. J., and Arakawa, T., 1994, Interactions of stabilizers with proteins during freezing and drying, in: *Formulation*

and Delivery of Proteins and Peptides (J. L. Cleland and R. Langer, eds.), American Chemical Society, Washington, DC, pp. 134–147.

Chait, B. T., and Kent, S. B. H., 1992, Weighing naked proteins: Practical, high accuracy mass measurement of peptides and proteins, *Science* **257:**1885–1894.

Chan, H. K., Ongpipattanakul, B., and Au-Yeung, J., 1996, Aggregation of rhDNase occurred during the compression of KBr pellets used for FTIR spectroscopy, *Pharm. Res.* **13:**238–242.

Chen, Y. H., Yang, J. T., 1971 A new approach to the calculation of secondary structures of globular proteins by optical rotatory dispersion and circular dichroism, *Biochem. Biophys. Res. Comm.* **44:**1285–1291.

Clarke, S., Stephenson, R. C., and Lowenson, J. D., 1992, Lability of asparagine and aspartic acid residues in proteins and peptides: Spontaneous deamidation and isomerization reactions, in: *Stability of Protein Pharmaceuticals; Chemical and Physical Pathways of Protein Degradation* (T. J. Ahern and M. Manning, eds.), Plenum Press, New York, pp. 1–29.

Cleland, J. L., and Wang, D. I. C., 1990, Refolding and aggregation of bovine carbonic anhydrase B: Quasi-elastic light scattering analysis, *Biochemistry* **29:** 11072–11078.

Cleland, J. L., Powell, M. F., and Shire, S. J., 1993, The development of stable protein formulations: A close look at protein aggregation, deamidation and oxidation, *Crit. Rev. Ther. Drug Carrier Sys.* **10:**307–377.

Cole, J. L., Gehman, J. D., Shafer, J. A., and Kuo, L. C., 1993, Solution oligomerization of the rev protein of HIV-1: Implications for function, *Biochemistry* **32:**11769–11775.

Constantino, H. R., Langer, R., and Klibanov, A. M., 1995, Aggregation of a lyophilized pharmaceutical protein, recombinant albumin: Effect of moisture and stabilization by excipients, *Bio/Technology* **13:**493–496.

Copeland, R. A., Ji, H., Halfpenny, A. J., Williams, R. W., Thompson, K. C., Herber, W. K., Thomas, K. A., Bruner, M. W., Ryan, J. A., Marquis-Omer, D., Sanyal, G., Sitrin, R. D., Yamazaki, S., and Middaugh, C. R., 1991, The structure of human acidic fibroblast growth factor and its interaction with heparin, *Arch. Biochem. Biophys.* **289:**53–61.

Cotter, R. J., 1992, Time-of-flight mass spectrometry for the structural analysis of biological molecules, *Anal. Chem.* **64:**1027–1039.

Cottrell, J. S., 1994, Protein identification by peptide mass fingerprinting, *Peptide Res.* **7:**115–124.

Craig, W. S., 1988, Determination of quaternary structure of an active enzyme using chemical cross-linking with glutaraldehyde, *Meth. Enzymol.* **156:**333–345.

Deyl, Z., and Struzinsky, R., 1991, Capillary zone electrophoresis: Its applicability and potential in biochemical analysis, *J. Chromatogr.* **569:**63–122.

Di Donato, A., Ciardiello, M. A., Nigris, M. D., Piccoli, R., Mazzarella, L., and D'Alessio, G., 1993, Selective deamidation of ribonuclease A: Isolation and characterization of the resulting isoaspartyl and aspartyl derivatives, *J. Biol. Chem.* **268:**4745–4751.

Dollinger, G., Cunico, B., Kunitani, M., Johnson, D., and Jones, R., 1992, Practical on-line determination of biopolymer molecular weights by high-performance liquid chromatography with classical light-scattering detection, *J. Chromatogr.* **592:**215–228.

Dong, A., Prestrelski, S. J., Allison, S. D., and Carpenter, J. F., 1995, Infrared spectroscopic studies of lyophilization- and temperature-induced protein aggregation, *J. Pharm. Sci.* **84:**415–424.

Eberlein, G. A., Stratton, P. R., and Wang, Y. J., 1994, Stability of rhbFGF as determined by UV spectroscopic measurements of turbidity, *PDA J. Pharm. Sci. Technol.* **48:**224–230.

Endo, Y., Nagai, H., Watanabe, Y., Ochi, K., and Takagi, T., 1992, Heat-induced aggregation of recombinant erythropoietin in the intact and deglycosylated states as monitored by gel permeation chromatography combined with a low-angle laser light scattering technique, *J. Biochem.* (Tokyo) **112:**700–706.

Eftink, M., 1994, The use of fluorescence methods to monitor unfolding transitions in proteins, *Biophys. J.* **66:**482–501.

Fabian, H., Schultz, C., Naumann, D., Landt, O., Hahn, U., and Saenger, W., 1993, Secondary structure and temperature-induced unfolding and refolding of ribonuclease T_1 in aqueous solution, *J. Mol. Biol.* **232:**967–981.

Farmer, T. B., and Caprioli, R. M., 1991, Assessing the multimeric states of proteins: Studies using laser desorption mass spectrometry, *Biol. Mass Spectrom.* **20:** 796–800.

Fink, A. L., and Painter, B., 1987, Characterization of the unfolding of ribonuclease A in aqueous methanol solvents, *Biochemistry* **26:**1665–1671.

Geisow, M. J., 1992, Mass measurement at high molecular weight—New tools for biotechnologists, *TIBTECH* **10:**432–441.

Gianazza, E., 1995, Isoelectric focusing as a tool for the investigation of post-translational processing and chemical modification of proteins, *J. Chromatogr. A* **705:**67–87.

Giddings, J. C., 1989, Field flow fractionation of macromolecules. *J. Chromtogr.* **470:**327–335.

Gorovits, B. M., Seale, J. W., and Horowitz, P. M., 1995, Residual structure in urea-denatured chaperonin GeoEL, *Biochemistry* **34:**13928–13933.

Green, S. M., Ginsburg, A., Lewis, M. S., and Hensley, P., 1991, Roles of metal ions in the maintenance of the tertiary and quaternary structure of arginase from *Saccharomyces cerevisiae, J. Biol. Chem.* **266:**21474–21481.

Gress, J. O., Marquis-Omer, D., Middaugh, C. R., and Sanyal, G., 1994, Evidence for an equilibrium intermediate in the folding–unfolding pathway of a transforming growth factor-α–*Pseudomonas* exotoxin hybrid protein, *Biochemistry* **33:**2620–2627.

Hermann, R., Rudolph, R., and Jaenicke, R., 1979, Kinetics of *in vitro* reconstitution of oligomeric enzymes by cross-linking, *Nature* **277:**243–245.

Hillenkamp, F., and Karas, M., 1990, Mass spectrometry of peptides and proteins by matrix-assisted ultraviolet laser desorption/ionization. *Meth. Enzymol.* **193:** 280–295.

Ismail, A. A., Mantsch, H., and Wong, P. T. T., 1992, Aggregation of chymotrypsinogen: Portrait by infrared spectroscopy. *Biochim. Biophys. Acta* **1121:** 183–188.

Johnson, M. L., Correia, J. J., Yphantis, D. A., and Halvorson, H. R., 1981, Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques, *Biophys. J.* **36:**575–588.

Johnson, W. C., 1990, Protein secondary structure and circular dichroism: A practical guide. *Proteins Struc. Funct. Genet.* **7:**205–214.

Kadima, W., Øgendal, L., Bauer, R., Kaarsholm, N., Brodersen, K., Hansen, J. F., and Porting, P., 1993, The influence of ionic strength and pH on the aggregation properties of zinc-free insulin studied by static and dynamic light scattering, *Biopolymers* **33:**1643–1657.

Katakam, M., Bell, L. N., and Banga, A., 1995, Effect of surfactants on the physical stability of recombinant human growth hormone, *J. Pharm. Sci.* **84:**713–716.

Kauppinen, J. K., Moffatt, D. J., Mantsch, H. H., and Cameron, D. G., 1981, Fourier transforms in the computation of self-deconvoluted and first-order derivative spectra of overlapped band contours, *Anal. Chem.* **53:**1454–1457.

Kendrick, B. S., Dong, A., Allison, S. D., Manning, M. C., and Carpenter, J. F., 1996, Quantitation of the area of overlap between second-derivative amide I infrared spectra to determine the structural similarity of a protein in different states, *J. Pharm. Sci.* **85:**155–158.

Kuhr, W. G., 1990, Capillary electrophoresis, *Anal. Chem.* **62:**403R–414R.

Lakowicz, J. R., 1983, *Principles of Fluorescence Spectroscopy*, Plenum Press, New York, pp. 40–47.

Lepock, J. R., Frey, H. E., Bayne, H., and Markus, J., 1989, Relationship of hypothermia-induced hemolysis of human erythrocytes to the thermal denaturation of membrane proteins, *Biochim. Biophys. Acta* **980:**191–201.

Li, S., Schoneich, C., and Borchardt, R. T., 1995, Chemical instability of protein pharmaceuticals: Mechanisms of oxidation and strategies for stabilization, *Biotechnol. Bioeng.* **48:**90–500.

Linemeyer, D. L., Menke, J. G., Kelly, L. J., DiSalvo, J., Soderman, D., Schaeffer, T., Ortega, S., Gimenez-Gallego, G., and Thomas, K. A., 1990, Disulfide bonds are neither required, present, nor compatible with full activity of human recombinant acidic fibroblast growth factor, *Growth Factors* **3:**287–298.

Litzén, A., Walter, J. K., Krischollek, H., and Wahlund, K. G., 1993, Separation and quantitation of monoclonal antibody aggregates by asymmetrical flow field flow fractionation and comparison to gel permeation chromatography. *Anal. Biochem.* **212:**469–480.

Liu, J., Lester, P., Builder, S., and Shire, S. J., 1995, Characterization of complex formation of humanized anti-IgE monoclonal antibody and monoclonal human IgE, *Biochemistry* **34:**10474–10482.

Mach, H., and Middaugh, C. R., 1994, Simultaneous monitoring of the environment of tryptophan, tyrosine, and phenylalanine residues in proteins by near-ultraviolet second-derivative spectroscopy, *Anal. Biochem.* **222:**323–331.

Mach, H., Thomson, J. A., Middaugh, C. R., and Lewis, R. V., 1991, Examination of phenylalanine microenvironments in proteins by second-derivative absorption spectroscopy, *Arch. Biochem. Biophys.* **287**:33–40.

Mach, H., Volkin, D. B., Burke, C. J., Middaugh, C. R., Linhardt, R. J., Fromm, J. R., Loganathan, D., and Mattson, L., 1993, Nature of the interaction of heparin with acidic fibroblast growth factor, *Biochemistry* **32**:5480–5489.

Mach, H., Volkin, D. B., Burke, C. J., and Middaugh, C. R., 1995, Ultraviolet absorption spectroscopy, in: *Protein Stability and Folding: Theory and Practice* (B. A. Shirley, ed.), Humana Press, Totowa, NJ, pp. 91–114.

Manavalan, P., and Johnson, W. C., Jr., 1987, Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra, *Anal. Biochem.* **167**:76–85.

Manning, M. C., 1989, Underlying assumptions in the estimation of secondary structure content in proteins by circular dichroism spectroscopy—A critical review, *J. Pharm. Biomed. Anal.* **7**:1103–1119.

Manning, M. C., and Woody, R. W., 1989, Theoretical study of the contribution of aromatic side chains to the circular dichroism of basic bovine pancreatic trypsin inhibitor, *Biochemistry* **28**:8609–8613.

Manning, M. C., Patel, K., and Borchardt, R. T., 1989, Stability of protein pharmaceuticals, *Pharm. Res.* **6**:903–918.

Middaugh, C. R., 1994, The analysis of protein pharmaceuticals: Near future advances, *Cytotechnology* **15**:187–194.

Middaugh, C. R., and Volkin, D. B., 1992, Protein solubility, in: *Stability of Protein Pharmaceuticals: Chemical and Physical Pathways of Protein Degradation* (T. J. Ahern and M. Manning, eds.), Plenum Press, New York, pp. 109–134.

Middaugh, C. R., Mach, H., Burke, C. J., Volkin, D. B., Dabora, J. M., Tsai, P. K., Bruner, M. W., Ryan, J. A., and Marfia, K., 1992, The molecular basis of the anti-growth factor activity of suramin, *Biochemistry* **31**:9016–9024.

Middaugh, C. R., Mach, H., Ryan, J., Sanyal, G., and Volkin, D., 1995, Infrared spectroscopy, in: *Protein Stability and Folding: Theory and Practice* (B. A. Shirley, ed.), Humana Press, Totowa, NJ, pp. 137–156.

Nguyen, T. H., 1994, Oxidation degradation of protein pharmaceuticals, in: *Formulation and Delivery of Proteins and Peptides* (J. L. Cleland and R. Langer, eds.), American Chemical Society, Washington, DC, pp. 59–71.

Nguyen, T. H., Burnier, J., and Meng, W., 1993, The kinetics of relaxin oxidation by hydrogen peroxide, *Pharm. Res.* **10**:1563–1571.

Perczel, A., Hollosi, M., Tusnady, G., and Fasman, G. D., 1991, Convex constraint analysis; a natural deconvolution of circular dichroism curves of proteins, *Protein Eng.* **4**:669–679.

Perczel, A., Park, K., and Fasman, G. D., 1992, Analysis of the circular dichroism spectrum of proteins using the convex constraint algorithm: A pratical guide, *Anal. Biochem.* **203**:89–93.

Powell, M. F., 1994, Peptide stability in aqueous parenteral formulations: Prediction of chemical stability based on primary sequence, in: *Formulation and Delivery of*

Proteins and Peptides (J. L. Cleland and R. Langer, eds.), American Chemical Society, Washington, DC, pp. 100–119.

Powell, M. F. *et al.*, 1996, A compendium and hydropathy/flexibility analysis of common reactive sites in proteins: Reactivity at Asn, Asp, Gln and Met motifs in neutral pH solution, in: *Formulation, Characterization and Stability of Protein Drugs:Case Histories* (Y. J. Wang and R. Pearlman, eds.), Plenum Press, New York, pp. 1–140.

Prestrelski, S. J., Arakawa, T., Wu, C-S. C., O'Neil, K. D., Westcott, K. R., and Narhi, L. O., 1992, Solution structure and dynamics of epidermal growth factor and transforming growth factor-α, *J. Biol. Chem.* **267:**319–322.

Prestrelski, S. J., Tedeschi, N., Arakawa, T., and Carpenter, J. F., 1993, Dehydration-induced conformational transitions in proteins and their inhibition by stabilizers, *Biophys. J.* **65:**661–671.

Ragone, R., Colonna, G., Balestrieri, C., Servillo, L., and Irace, G., 1984, Determination of tyrosine exposure in proteins by second-derivative spectroscopy, *Biochemistry* **23:**1871–1875.

Ragone, R., Colonna, G., Bismuto, E., and Irace, G., 1987, Unfolding pathway of myoglobin: Effect of denaturants on solvent accessibility to tyrosyl residues detected by second-derivative spectroscopy, *Biochemistry* **26:** 2130–2134.

Sanchez-Ruiz, J. M., Lopez-Lacomba, J. L. Cortijo, M., and Mateo, P. L., 1988, Differential scanning calorimetry of the irreversible thermal denaturation of thermolysin, *Biochemistry* **27:**1648–1652.

Sanyal, G., Marquis-Omer, D., and Middaugh, C. R., 1996, Biophysical characterization and formulation of TP40: A chimeric protein that requires a pH-dependent conformational change for its biological activity, in: *Formulation, Characterization and Stability of Protein Drugs: Case Histories* (Y. J. Wang and R. Pearlman, eds.), Plenum Press, New York, pp. 365–392.

Schein, C. H., 1994, Controlling oligomerization of pharmaceutical proteins. *Pharm. Acta Helv.* **69:**119–126.

Shahrokh, Z., Stratton, P. R., Eberlein, G. A., and Wang, Y. J., 1994, Approaches to analysis and demonstrating mass balance in pharmaceutical protein (basic fibroblast growth factor) formulations, *J. Pharm. Sci.* **83:**1645–1650.

Siuzdak, G., 1994, The emergence of mass spectrometry in biochemical research, *Proc. Natl. Acad. Sci. USA* **91:**11290–11297.

Smith, R. D., Loo, J. A., Barinaga, C. J., Edmonds, C. G., and Udseth, H. R., 1989, Capillary zone electrophoresis and isotachophoresis–mass spectrometry of polypeptides and proteins based upon an electrospray ionization interface, *J. Chromatogr.* **480:**211–232.

Smith, R. D., Loo, J. A., Loo, R. R. O., Busman, M., and Udseth, H. R., 1991, Principles and practice of electrospray ionization–mass spectrometry for large polypeptides and proteins, *Mass Spectrom. Rev.* **10:**359–451.

Sreerama, N., and Woody, R. W., 1993, Self-consistent algorithm for secondary structure analysis from CD data, *Anal. Biochem.* **209:**32–44.

Steele, J. C. H., and Neilsen, T. B., 1978, Evaluation of cross-linked polypeptides in SDS gel electrophoresis, *Anal. Biochem.* **84:**218–224.

Surewicz, W. K., Mantssch, H. H., and Chapman, D., 1993, Determination of protein secondary structure by Fourier transform infrared spectroscopy: A critical assessment, *Biochemistry* **32:**389–394.

Suter, M. J.-F., Moore, W. T., Farmer, T. B., Cottrell, J. S., and Caprioli, R. M., 1992, Analysis of protein–protein binding with laser desorption mass spectrometry: Can multimeric forms be studied? in: *Techniques in Protein Chemistry III* (J. Villafranca, ed.), Academic Press, San Diego, CA, pp. 447–455.

Teshima, G., Porter, J., Yim, K., Ling, V., and Guzzetta, A., 1991, Deamidation of soluble CD4 at asparagine-52 results in reduced binding capacity for the HIV-1 envelope glycoprotein gp120, *Biochemistry* **30:**3916–3922.

Teshima, G., Hancock, W. S., and Canova-David, E., 1995, Effect of deamidation and isoaspartate formation on the activity of proteins, in: *Deamidation and Isoaspartate Formation in Peptides and Proteins* (D. W. Aswad, ed.), CRC Press, Boca Raton, FL, pp. 168–191.

Tsai, P. K., Volkin, D. B., Dabora, J. M., Thompson, K. C., Bruner, M. W., Gress, J. O., Matuszewska, B., Keogan, M., Bondi, J. V., and Middaugh, C. R., 1993a, Formulation design of acidic fibroblast growth factor, *Pharm. Res.* **10:**649–659.

Tsai, P. K., Bruner, M. W., Irwin, J. I., Ip, C. C. Y., Oliver, C. N., Nelson, R. W., Volkin, D. B., and Middaugh, C. R., 1993b, Origin of the isoelectric heterogeneity of a monoclonal immunoglobulin, *Pharm. Res.* **10:**1580–1586.

Valero, E., DeBonis, S., Filhol,O., Wade, R. H., Langowski, J., Chambaz, E. M., and Cochet, C., 1995, Quaternary structure of casein kinase 2—Characterization of multiple oligomeric states and relation with its catalytic activity. *J. Biol. Chem.* **270:**8345–8352.

Van Holde, K. E., 1975, Sedimentation analysis of proteins, in: *The Proteins*, 3rd ed. (H. Neurath, R. L. Hill, and C.-L. Boeder, eds), Academic Press, New York, pp. 225–291.

Van Stokkum, I. H. M., Linsdell, H., Hadden, J. M., Haris, P. I., Chapman, D., and Bloemendal, M., 1995, Temperature-induced changes in protein structure studied by Fourier transform infrared spectroscopy and global analysis, *Biochemistry* **34:**10508–10518.

Venyaminov, S. Y., and Kalnin, N. N., 1990a, Quantitative IR spectrophotometry of peptide compounds in water (H_2O) solutions. I. Spectral parameters of amino and residue absorption bands, *Biopolymers* **30:**1243–1257.

Venyaminov, S. Y., and Kalnin, N. N., 1990b, Quantitative IR spectrophotometry of peptide compounds in water (H_2O) solutions. II. Amide absorption bands of polypeptides and fibrous proteins in α-, β- and random coil conformations, *Biopolymers* **30:**1259–1271.

Volkin, D. B., and Klibanov, A. M., 1989, Minimizing protein inactivation, in *Protein Function: A Practical Approach* (T. E. Creighton, ed.), Oxford University Press, Oxford, pp. 1–24.

Volkin, D. B., and Middaugh, C. R., 1992, The effect of temperature on protein structure, in: *Stability of Protein Pharmaceuticals: Chemical and Physical Pathways of Protein Degradation* (T. J. Ahern and M. Manning, eds.), Plenum Press, New York, pp. 215–247.

Volkin, D. B., and Middaugh, C. R., 1996, The characterization, stabilization and formulation of acidic fibroblast growth factor, in: *Formulation, Characterization and Stability of Protein Drugs: Case Histories* (Y. J. Wang and R. Pearlman, eds.), Plenum Press, New York, pp. 181–217.

Volkin, D. B., Tsai, P. K., Dabora, J. M., Gress, J. O., Burke, C. J., Linhardt, R. J., and Middaugh, C. R., 1993, The physical stabilization of acidic fibroblast growth factor by polyanions, *Arch. Biochem. Biophys.* **300:**30–41.

Volkin, D. B., Verticelli, A. M., Bruner, M. W., Tsai, P. K., Marfia, K. E., Sardana, M., and Middaugh, C. R., 1995a, Deamidation of polyanion-stabilized acidic fibroblast growth factor, *J. Pharm. Sci.* **84:**7–11.

Volkin, D. B., Mach, H., and Middaugh, C. R., 1995b, Degradative covalent reactions important to protein stability, in: *Protein Stability and Folding: Theory and Practice* (B. A. Shirley, ed.), Humana Press, Totowa, NJ, pp. 35–63.

Wallace, B. A., and Mao, D., 1984, Circular dichroism analyses of membrane proteins: An examination of differential light scattering and absorption flattening effects in large membrane vesicles and membrane sheets, *Anal. Biochem.* **142:**317–328.

Wright, H. T., 1991, Nonenzymatic deamidation of asparaginyl and glutaminyl residues in proteins, *CRC Crit. Rev. Biochem. Mol. Biol.* **26:**1–52.

Yang, J. T., Chen, G. C., and Jirgensons, B., 1976 Optical rotatory dispersion and circular dichroism of protein, in: *Handbook of Biochemistry and Molecular Biology*, 3rd ed., Vol. III (G. D. Fasman, ed.), CRC Press, Boca Raton, FL, pp. 3–140.

Zale, S.E., and Klibanov, A.M., 1986, Why does ribonuclease irreversibly inactivate at high temperatures? *Biochemistry* **25:**5432–5444.