Resolving Nitrogen-15 and Proton Chemical Shifts for the Mobile Segments of Elastin

with Two-Dimensional NMR Spectroscopy

Kosuke Ohgo, Walter P. Niemczura, Brian C. Seacat, Steven G. Wise, Anthony S. Weiss, and Kristin K. Kumashiro*

Contribution from the 1Department of Chemistry, University of Hawaii, 2545 McCarthy Mall, Honolulu, Hawaii, 96822 (USA), and the 2School of Molecular Bioscience, University of Sydney, Sydney (Australia)

*Address correspondence to: Kristin K. Kumashiro, University of Hawaii, Department of Chemistry, 2545 McCarthy Mall, Honolulu, Hawaii 96822 (USA); Phone: 1-808-956-5733; Fax: 1-808-956-5908; Email: kumashir@hawaii.edu

Running title: 15N and 1H Shifts of Elastin's Mobile Regions

Summary

In this study, one- and two-dimensional NMR experiments are applied to uniformly-15N-enriched synthetic elastin, a recombinant human tropoelastin that has been crosslinked to form an elastic hydrogel. Hydrated elastin is characterized by large segments that undergo “liquidlike” motions that limit the efficiency of cross-polarization. The refocused INEPT experiment is used to target these extensive, mobile regions of this protein. Numerous peaks are detected in the backbone amide region of the protein, and their chemical shifts indicate the completely unstructured, “random coil” model for elastin is unlikely. Instead, more evidence is gathered that supports a characteristic ensemble of conformations in this rubberlike protein.

Introduction

Elasticity in vertebrate tissue is traced to elastin, an extensive biopolymer that is characterized by the predominance of small hydrophobic residues, such as glycine, alanine, valine, and proline. The elastin gene encodes tropoelastin, the soluble monomer of the mature protein. Tropoelastin is crosslinked in post-translational modification to form the amorphous (non-crystalline) and insoluble protein known more broadly as elastin. The classic picture of tropoelastin and elastin describes the monomer and the polymer as having two types of domains,
the hydrophobic and the crosslinking. The hydrophobic domains are well-known for the abundance of the repeating polytetra-, penta- and hexapeptides, of which the best-known is (VPGVG)_n. Many of the crosslinking domains are alanine-rich; lysines in these polyalanine regions are enzymatically oxidized in post-translational modification to form the heteroaromatic crosslinking moieties, desmosine and isodesmosine (1,2). The high alanine content indicates a propensity for \( \alpha \)-helix formation, which is detected in experiment (3). In contrast, the structural pictures of the extensive hydrophobic domains and, consequently, the native, intact protein are less homogeneous and also significantly more controversial (2).

Solid-state nuclear magnetic resonance (NMR) spectroscopy has been used to show that elastin peptide mimetics are neither unstructured “random coils” nor highly ordered motifs (4-6). The mimetics have sequences containing only the repeating polypentapeptides of the hydrophobic domains of elastin. These polypeptide mimetics are characterized by several conformations, lending support to models like the one of Tamburro that invokes a “conformational equilibrium”, wherein the equilibrium position shifts as the elastic fibers are stretched and then again when it contracts (7). Earlier studies also highlighted an additional consideration, i.e., the possibility that elastin’s structural distribution is unique to this protein; i.e., two-dimensional NMR data for an elastin mimetic were consistent with neither a single conformation nor a distribution based on the solved structures in databases (4). In this earlier paper, we noted that approaches based on solved structures are most appropriate for a peptides for which it is most likely that \( \alpha \)-helices or \( \beta \)-sheets are present or, alternatively, that it is unstructured or “random coil”. I.e., these commonly used, empirical approaches appeared unsuitable for peptide mimetics (and the native proteins upon which they are based) that are characterized by less common conformations such as polyproline-II or \( \beta \)-turns.

Recent two-dimensional (2D) NMR experiments provided additional insights into the multiple conformations that are present in native elastin (8,9), but these studies were conducted using unenriched elastin, as they focused on the natural-abundance \(^{13}\)C populations of the protein. This experiment, however, was more informative for the lyophilized elastin than for the hydrated form. The various segments in hydrated elastin undergo motion on different timescales, and large populations of elastin are invisible to CP (9,10). I.e., large segments of the hydrated protein undergo rapid, “liquidlike” motion. These fast, internal dynamics limit the use of the more well-established CP-based methods, like FSLG-HETCOR, for describing the conformational makeup in this protein’s mobile domains.

INEPT (Insensitive Nuclei Enhanced by Polarization Transfer) (11) is a classic solution NMR experiment that enables the observation of low-\( \gamma \) nuclei, such as \(^{15}\)N, by application of a modified spin-echo sequence. More recent accounts show that this method may be successfully applied to solid samples (12-15). A natural evolution of the methodology paired refocused INEPT (rINEPT) (16,17) with HETCOR (HETeronuclear CORrelation); rINEPT-HETCOR combines the rINEPT enhancement with the correlation information of HETCOR to yield information about targeted sites in a diverse array of samples, from silks (12) to microcrystalline proteins (13) and powder samples of, e.g., small molecules (14). rINEPT-HETCOR can be employed in both static (15) and MAS (12-14) experiments. A similar approach to the current study is the characterization of the flexible regions in a human prion peptide (18). Interestingly, the fast motions in elastin remove the need for \(^1\)H homonuclear decoupling (13,14) or ultrafast sample spinning rates (12), which were used in other studies to counter the effects of the short \(^1\)H T\(_2\)' relaxation time constants.

In this study, the rINEPT and rINEPT-HETCOR experiments are applied to a hydrated sample of uniformly \(^{15}\)N-enriched synthetic elastin, thus permitting the first detailed characterization of its extensive mobile regions. The sample was prepared by the chemical, or non-enzymatic, crosslinking of recombinant human tropoelastin (19-21). The polymeric nature of this protein was an essential element in the analysis of the \(^{15}\)N and \(^1\)H chemical shifts.
**Experimental Section**

**NMR spectroscopy.** Data were acquired on a Varian Inova NMR spectrometer, equipped with a wide-bore superconducting magnet with a $^1$H resonance frequency of 399.976 MHz. The probe used for these experiments was a 4 mm triple-resonance (HXY) T3 MAS probe (Chemagnetics/Varian NMR, Ft. Collins, CO).

For direct polarization (DP), or single-pulse excitation of the $^{15}$N, a 9.5 µs $^{15}$N 90° pulse was used with a 20 s recycle delay. Baselines of the DP spectra were corrected using the algorithm described by Golotvin and Williams (22).

For CP, a 2.9 µs $^1$H 90° pulse was followed by a 1 ms contact time with a 5 s recycle delay. The CP field strength was set to $\gamma B_{1H}/2\pi = 34.3$ kHz, matched to the (+1) Hartman-Hahn sideband condition. TPPM decoupling (23) was applied during acquisition, with an applied field strength of $\gamma B_{1H}/2\pi \approx 60$ kHz.

The $^1$H-$^{15}$N refocused INEPT spectra were obtained using 90° pulse lengths for $^1$H and $^{15}$N of 2.9 µs and 8.8 µs, respectively. Each period $\tau$ was set to 1.6 ms. A 16-step phase cycle was applied (13). A 1.5 s recycle delay was used for all experiments. For the $^1$H-$^{15}$N 2D rINEPT-HETCOR experiment (Supplemental Figure S1), the INEPT parameters were the same as described for the 1D experiments. In addition, the spectral width in the indirect dimension was 2400 Hz, with a maximum $t_1$ evolution time of 17.1 ms over 42 increments. 256 scans were acquired per $t_1$ point, such that each 2D experiment required 9.5 h of measuring time. “Semi-constant-time” $^1$H chemical shift evolution in the indirect dimension was used (18).

The overlapping signals of the Gly region in the 2D spectra were deconvolved, using two-dimensional elliptical Gaussian functions. Parameters of the function were the amplitude at the center of the Gaussian, as well as the $^1$H and $^{15}$N chemical shifts (also at center). In addition, the Gaussians were further defined by their full widths at half maximum in both dimensions ($\text{FWHM}_1$ and $\text{FWHM}_2$) and the rotation angle of the curve on the 2D plane. These parameters of amplitude, chemical shift, $\text{FWHM}_1$, $\text{FWHM}_2$, and the rotation angle were optimized for each of four elliptical Gaussian functions, applying the Levenberg-Marquardt nonlinear fitting algorithm (24) using MATLAB (MathWorks Inc., Natick, MA). These four Gaussians were used to fit the lineshape in the Gly region. Peak volumes were calculated from the product of the amplitudes, as $(\text{FWHM}_1)(\text{FWHM}_2)$.

Each rotor was sealed to preserve water content with a custom-machined spacer (Revolution NMR, Ft. Collins, CO) that was fitted with micro-o-rings (Apple Rubber Products Inc., Lancaster, NY). Fluorinated polymers were used for the spacer (Kel-F) and o-rings (fluorosilicone) to reduce background in (related) $^{13}$C CPMAS experiments (25). (To reduce losses due to the unpacking and packing of these samples, the fluorinated polymers are used here, as well.)

Data were acquired at 5 °C, 13 °C, 21 °C, 29 °C and 37 °C. The sample temperatures were calibrated using lead nitrate, Pb(NO$_3$)$_2$ (26). $^{15}$N chemical shifts were externally referenced to $^{15}$N-glycine [$\delta(^{15}$N) = 32.0 ppm at 37 °C]. For each 2D experiment, $^1$H chemical shifts were externally referenced to sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) in D$_2$O [$\delta(^1$H) = 0.0 ppm at 32 °C]. An 8.0 kHz MAS rate was used for all experiments.

**Sample preparation.** Synthetic elastin was prepared as described previously (21). Briefly, tropoelastin corresponds to amino acid residues 27-724 (gi 182,020), which is the 60 kDa mature form of the secreted protein following removal of the signal peptide (19,20). The recombinant human tropoelastin was coacervated and allowed to react with the crosslinking reagent bis(3-sulfo-N-hydroxsuccinimide ester, BS3, yielding an elastic hydrogel, i.e., “synthetic elastin”. To prepare uniformly $^{15}$N-enriched protein, the growth media was supplemented with enriched ammonium chloride ($^{15}$N, 98%+) (Cambridge Isotope Laboratories, Andover, MA). The enriched, synthetic elastin (17.5 mg) was swelled in modified phosphate buffered saline (PBS) (pH = 7.4; 83.3 mM NaCl, 40.1 mM KCl, 8.3 mM Na$_2$HPO$_4$ and 1.6 mM KH$_2$PO$_4$ overnight and then packed into the NMR rotor.
Results and Discussion

Mobile and rigid domains identified with 1D DP, rINEPT, and CP MAS NMR experiments.

The entire protein, its mobile segments, and its rigid regions were observed with DP, rINEPT, and CP, respectively, as shown in Figure 1. All experiments were done with MAS. The DP spectrum included all $^{15}$N signals. The rINEPT experiment identifies $^{15}$N nuclei that are directly bonded to one or more protons, via the $J$-coupling. Thus, it is a good selection tool for $^{15}$N-$^1$H spin pairs with weak dipolar couplings, such as the backbone amides of the extensive mobile segments of elastin. The $^{15}$N-$^1$H CPMAS experiment was used to identify regions that have dynamics more typical of a solid, where the dipolar couplings are largely intact. The sum of the CPMAS and refocused INEPT spectra appeared to be very similar to the DPMAS spectrum, with the exception of the most downfield region that contained the backbone amides of the prolines. Thus, to a rough approximation, the rINEPT and CP accounted for nearly all of the $^{15}$N nuclei in this sample.

Elastin’s extensive nature and the uniform labeling scheme did not easily facilitate site-specific assignments, as is commonly done in solid-state NMR studies of smaller proteins and peptides. However, this protein’s unique composition, its polymeric nature, and the characteristic chemical shift ranges for the various amino acid types allowed for assignments of the abundant glycines, valines, alanines, and prolines. Glycines comprise 29.8% of this protein and are primarily found in the hydrophobic domains and linker regions. Valines are also found in the hydrophobic domains and make up 13% of the protein. Alanines and prolines account for another 22.5% and 12.3%, respectively, and are found in both hydrophobic and crosslinking domain types. These four amino acids account for 78% of tropoelastin and synthetic elastin. Thus, the analysis of these spectra focused on these four amino acids, plus leucine, which is the fifth most abundant in elastin (5.7%). The chemical shifts are summarized in Table 1. Note that there are very small amounts of nitrogen-containing sidechains, such as (uncrosslinked) lysines and arginines. The crosslinks themselves also contain nitrogen, but they account for a very small part of the protein. Only 2-3 desmosines and isodesmosine per thousand residues are found in native, mature elastin, where they are accompanied by a higher population of linear allysinealdehyde and lysinorleucine crosslinks and their derivatives (1,27). This profile is correlated with a similar, very low density of BS3-derived crosslinks in synthetic elastin (28). The $^{15}$N chemical shift of the BS3 crosslink in an elastin mimic is 123.9 ppm (29), which adds a relatively minor contribution to the crowded backbone amide region, as described in more detail below.

Further downfield in the $^{15}$N DPMAS spectrum were peaks resolved at 133.7, 135.7, and 138.2 ppm at 37°C. There was also a small upfield shoulder on this envelope. These peaks have $^{15}$N chemical shifts that correspond to prolines (30). They did not appear in the rINEPT data, as expected for these nonprotonated sites. The prolines were also not observed in the CP spectra. Even though they are nonprotonated, the CP mixing times used should have permitted sufficient polarization transfer, if the Pro-rich regions were located in “rigid” domains. For instance, prolines in lyophilized elastin are clearly visible with a 0.5 to 2.5 ms contact time in a CP experiment (data not shown). Therefore, the prolines of hydrated elastin were assumed to undergo rapid motion, either in the form of ring interconversion, large-amplitude fluctuations of the backbone, or both.

Glycines were detected over the chemical shift range of ~105-112 ppm. This residue type is upfield of all other backbone amides (31), so the assignment of this region to glycine is considered relatively unambiguous. The glycine populations displayed varying degrees of heterogeneity in their dynamic profile, with some sites undergoing rapid, liquidlike motion or reorientation, whereas others were located in regions that are rigid enough to permit CP transfer. At 37°C two features were resolved at 107.6 and 110.7 ppm in the DPMAS spectrum. The high mobility of the Gly-rich regions was manifested as narrow, high-intensity peaks in the rINEPT spectra. The glycines’ CP lineshape was broader, with highest intensity found near 107.5-108 ppm.
Generally speaking, residues other than Gly and Pro have backbone amide resonances in the region of ca. 115-130 ppm (31), as was observed for the synthetic elastin. For the case of tropoelastin, much of the intensity in this region was attributed to the backbone amide nitrogens of Ala and Val. At 37°C features at 117.6, 119.4, and 121.9 ppm were most clearly resolved in the DPMAS. There was also a shoulder at ~127 ppm. As with the glycines, the positions of the resolved peaks in the DP did not vary significantly from those in the rINEPT data, although there were some minor differences in relative intensities. In the CP data, there was a broad and featureless peak in this region, with a center-of-mass at ~120 ppm. Again, this lineshape is more typical of a protein with a complex composition in the solid state. Thus, the 120 ppm resonance in the CP data is largely attributed to the crosslinking regions, which are believed to be more rigid than the hydrophobic domains.

Upfield signals from the sidechain nitrogen atoms, $^{15}\text{N}\varepsilon$-Lys (32.0 ppm), $^{15}\text{N}\eta$-Arg (71.3 ppm) and $^{15}\text{N}\varepsilon$-Arg (84.3 ppm), were also observed in the DP spectrum. These peaks were not observed in the rINEPT, presumably due to their low amounts in native elastin and the sensitivity differences between this experiment and DPMAS. The lysine sidechain resonance was observed in the CPMAS spectrum. Lysines are found almost exclusively in the crosslinking domains, so this observation is consistent with the assignment of the 120 ppm peak in the CP data described in the preceding paragraph.

An interesting experimental point arose in the consideration of the necessity of spinning this sample. Other solid-state rINEPT NMR studies utilized either ultra-fast MAS rates ($v_r = 60$ kHz) (12) or $^1\text{H}$ homonuclear dipolar decoupling techniques such as eDUMBO (13) to overcome short $^1\text{H} T_2'$ values. In the case of hydrated elastin, a moderate MAS rate ($v_r = 8$ kHz) without $^1\text{H}$-$^1\text{H}$ dipolar decoupling was sufficient to produce well-resolved spectra, presumably due to fast, large-amplitude motions in the protein. This motional narrowing might be considered as “self-decoupling”, although it was not completely effective at removing all of the broadening mechanisms, as slower spin speeds (or static experiments) led to intensity reductions (see Supplemental Figure S2). On this basis, the 8 kHz MAS spin rate was optimal for this system.

Cooling the protein to 5°C yielded subtle changes in resolution and chemical shifts. Previously, it was shown that the $^{13}\text{C}$ CPMAS spectrum of fully hydrated (unenriched) native elastin undergoes a gradual change upon cooling (10). At colder temperatures, from below 0°C to -40°C, there is a prominent backbone carbonyl peak with a center-of-mass at ~172-173 ppm and 8-9 resolved features in the aliphatic region. In addition, the spectrum of the lyophilized protein is similar, if not identical, to that of the frozen one. In contrast, warming the frozen (hydrated) sample resulted in an overall reduction in the CP-detected peak intensities across the entire $^{13}\text{C}$ spectrum. However, the resolved sites do not undergo the phase transition at the same temperature; i.e., the change from a mixture of liquidlike and solid behavior to one where the entire sample responds to cross-polarization sequences as is typical for a “true solid” does not happen at the same temperature for all resolved sites. Most notably, the $\alpha$-Gly peak was resolved at 4°C, which is a slightly higher temperature than observed for the rest of the protein. Thus, the current study included data acquired near this temperature, as shown in Figure 1B.

The $^{15}\text{N}$ CPMAS NMR spectra at 5°C and 37°C appear nearly identical at first glance, with only the expected increase in overall signal intensity in the cooled sample. A careful comparison of the DP and rINEPT spectra, however, shows that sample-cooling leads to the resolution of contributions at ~107.8 and 108.5 ppm. Similarly, in the series of peaks from 115-130 ppm, the tallest peak was still the most upfield (118.5 pm). However, features at 120.5 ppm and on the downfield shoulder (125-129 ppm) were also more apparent at the lower temperature. Generally, more peaks are resolved at the cooler temperature, and many of these $^{15}\text{N}$ chemical shifts differ from those observed at 37°C by as much as 1.1 ppm. Possibly, the loss of resolution at the warmer temperatures may signal that the protein is situated in the regime of intermediate exchange. The effect of temperature on the protein’s organization was further explored with
the increased resolution of heteronuclear correlation methods, as demonstrated in the next section.

**Variable-temperature rINEPT-HETCOR NMR resolved peaks in the crowded backbone amide region.** Frequency-switched Lee-Goldburg decoupling with heteronuclear correlation (FSLG-HETCOR) was recently used to identify multiple conformations in lyophilized and hydrated samples of unlabeled native elastin (9). The resolution in two dimensions confirmed that the broad lineshapes of the lyophilized elastin are inhomogeneous in nature, i.e., the observed lineshape in one dimension is a superposition of numerous resonances. FSLG-HETCOR yielded less structural information for the hydrated sample, however, as the lineshapes collapsed into centerband-only spectra. As with the 1D spectra illustrated above, however, selection of the numerous, mobile sites with the refocused INEPT experiment proved to be ideal for the observation of the mobile segments that are invisible to these other methods.

Tropoelastin’s 698-amino-acid sequence and overall complexity of the protein impose numerous challenges in the straightforward assignment of each of the seven resolved peaks in the rINEPT-HETCOR spectra. Thus, tentative assignments were made using the $^{15}$N chemical shifts of the elastin mimetics (VPGVG)$_n$ and (AVGVP)$_n$ (32) and the human elastin domain 26 peptide (33). In addition, the repeating polypeptides found throughout the tropoelastin sequence provide a rational basis for simplification of this complex problem. That is, the assignment of a polymer assembled from a 698-aa monomer that itself is polymerlike. As noted above, glycines are resolved from the rest of the backbone amides. In addition, there are higher propensities for any given glycine to residue next to one of the other 3 abundant amino acids, or to another glycine. For instance, glycine is preceded by proline (as PG) 59 times in the tropoelastin sequence, accounting for 28.5% of the glycines in the protein. The PG motif is found almost exclusively in the hydrophobic segments, and it is thus expected that the structure and dynamics of glycines-preceding-prolines in one region should be similar, if not identical, to PG in other segments of the protein. Similarly, the VG, GG, and AG sequences are found 35, 35, and 31 times, respectively, accounting for another ~50% of the glycines. General trends in chemical shifts for a given amino acid, coupled with the protein’s repeat-rich nature are also used to assign the downfield peaks (5)-(7), as well.

Figure 2 shows the $^{15}$N-$^1$H rINEPT-HETCOR NMR spectra of the mobile backbone region of hydrated synthetic elastin at 37°C, 29°C, 21°C, 13°C, and 5°C. Seven major features were identified through the observed temperature range, although their relative intensities and their chemical shifts show some temperature-dependence; i.e., some of the peaks were more clearly visible in the contour plots at lower temperatures. The peak positions in the $^{15}$N and $^1$H dimensions for the 37°C, 21°C, and 5°C are listed in Table 1.

Generally, there were two clusters of peaks in the contour plots, mirroring the discussion of the 1D rINEPT spectra. Peaks (1)-(4) and minor, poorly resolved contributions in this $^{15}$N chemical shift range corresponded to the backbone amide of the abundant glycines in elastin. The overlap of the peaks (1)-(3) gave rise to the broader and more intense peak centered at 107.6 ppm in the 1D MAS spectra at 37°C. At 5°C, the peak (1) was more prominent than at higher temperatures, again, consistent with the 1D spectra. Peak (4) directly corresponds to the 110.7 ppm resonance observed at 37°C in the 1D experiments.

The glycine that neighbors valine, denoted as V$_G$, has a $^{15}$N chemical shift of 110-111 ppm in the elastin mimetics (32). Valines generally induce a downfield shift, compared to the neighboring effects of Ala, Pro, and Gly (34). Thus, the downfield peak (4) at 110.7 ppm was assigned as the V$_G$ or V$_G$X population in synthetic elastin. In contrast, the assignments for the other common 2-residue sequences containing Gly were less straightforward. There is more overlap for the AG, PG, GG, and LG sequences. For instance, the peak at 107.1 could be PG, as in the (VPGVG) subunit (32), or AG, as in domain 26 of tropoelastin (33).

As with the 1D spectra, peaks (5)-(7) were due to the $^{15}$N backbone amides of Ala, Val, and, to lesser extent, Leu. Peak (7) is the most downfield, and it likely corresponds to all of the
mobile alanines. Alanines in the solution NMR spectra of the hydrophobic domain 26 were observed in the range of 122.7-126.0 ppm, for reference (33). Valines that are preceded by glycines in domain 26 and the poly(VPGVG) mimetic are all found in the range of 116-121 ppm, so the resonances at 119.4 and 117.6 ppm in the 37°C spectra are tentatively assigned to the GV backbone amides. The neighboring effect of prolines is well-known, as they induce an additional ca. 2 ppm downfield shift. So, the 119.4 and 117.6 ppm peaks might be further delineated as those of GVP and GVX (X≠P), respectively.

Temperature coefficients of the amide protons have been used to determine the propensity of hydrogen bond formation in proteins (35,36). For ¹H coefficients greater than -4.6 ppb/K, the hydrogen bond predictive value exceeds 85%. It increases to over 93% for amides in the range of -4 to -1 ppb/K (36). All peaks exhibited an upfield shift with temperature, but it was more pronounced for some peaks (See Supplemental Table S1). There was negligible change for peak (2). Thus, these glycines participate in hydrogen-bonding to an appreciable extent. The other resolved Gly cross-peaks, however, were reflective of populations that are not likely involved in H-bonding. Possibly, the sidechains of neighboring residues preclude the affected glycines from forming the hydrogen bonds. There were higher temperature coefficients for the other backbone amides, indicating that the populations corresponding to (5) and (6), especially, are not involved in hydrogen-bonding.

**Chemical shifts do not support a completely unstructured “random coil” model for elastin at 37°C.** The chemical shift of a given site is reflective of its environment, such as it is determined by steric and inductive effects. Of considerable consequence in proteins are the primary and secondary structures, i.e., sequence and conformation, respectively. Conversely, the difference between the observed chemical shift (for a given nucleus) and that of the random coil value, also known as the secondary chemical shift or chemical shift index (CSI) (37), has been used as an indicator of secondary structure. However, there is some ambiguity as to how best to treat the shift data for residues in defined secondary structures that are less common than the α-helices and β-sheets that are prevalent in most of the entries in structural databases. Most relevant here is the propensity of glycine to adopt secondary structures other than α-helix. Indeed, database entries for a 3₁₀-helical Gly (38) and β-turn (39) are often categorized in the “random coil” designation in CSI studies. Thus, the usual, empirical approach of using secondary chemical shifts must be applied judiciously to Gly-rich proteins like elastin, in which NMR (4-6,9), CD (7), and IR (40) data provide mounting evidence for a mixture of secondary structures, including the ones that are not usually represented in standard “α-helix vs. β-sheet vs. random coil” designations.

Two semi-empirical approaches were employed for the analysis of these rINEPT-HETCOR spectra. One method gave a base value for the random coil shift for each of the amino acid types X, based on solution ¹H and ¹⁵N NMR experiments of a series of the model peptides Ac-GGXGG-NH₂ in the denaturing condition of 8 M urea (41). In addition, the neighboring effects of each residue on any of the others were determined. In this manner, the chemical shifts for the central residue of a three-amino-acid sequence in a random coil could be predicted in a systematic fashion. The other method is a result of an extensive survey of database information for solved structures, i.e., the Biological Magnetic Resonance Data Bank (BioMagResBank or BMRB) and the Protein Databank (PDB). In this method, base values for the chemical shifts for each of the atoms in a given residue in random coil, β-sheet, and α-helical conformations were reported alongside smaller “corrections” for neighboring residue effects (42). In this approach, the shifts for the central residue of a 3-aa sequence in each of these three environments may be predicted.

Chemical shifts were calculated using the two approaches described above, and then they were compared to the experimental spectrum at 37°C, as shown in Figure 3. The experimental spectrum is illustrated in grey contours. The color contours are used to illustrate the intensities of the predicted chemical shifts, if the protein were a
random coil (3A-3D), a β-strand (3E-3F), or an α-helix (3G-3H). The 15N and 1H chemical shifts were calculated for Ala, Gly and Val in a given “structure” (α-helix, β-strand, or random coil) and 3-residue sequence. Each set of shifts were then plotted as a “peak”, using a Gaussian function (FWHM of 0.4 ppm for 15N and 0.08 ppm for 1H) weighted by the relative number of occurrences of that particular 3-aa sequence in tropoelastin. The various 3-aa sequences used in this calculation are summarized in Supplementary Table S2. For better viewing, only the peaks at or above the 20% contour line were used. Thus, the visible color contour peaks likely have a high number of occurrences in the protein and/or represent two or more overlapping 3-aa sequences. Particularly regarding the latter, often the residue that follows a given amino acid has less of an effect than the one that precedes it. Thus, most of the color contours are marked by a sequence like, e.g., PGX, meaning that all glycines that are preceded by proline in a random coil have nearly identical shifts, with very little effect from the residue X. Supplemental Figure S3 illustrates this representation in more detail.

The semi-empirical method utilizing the Ac-GGXGG-NH2 in denaturing conditions indicates that little random coil content is present in elastin’s mobile domains at 37 °C, as illustrated in Figures 3A and 3B. None of the shifts that are predicted for the prevalent sequences in tropoelastin overlap exactly with the 7 major peaks. Closest are the GVX (including GVP) to peak (6) (Figure 3A), which is consistent with the tentative assignment made with the 1D spectra. Another reasonably close match may be the GGX shift (Figure 3B); the calculated GGX random coil resonances are closest to peak (1). However, the AGX, VGX, and (overlapping) PGX and LGX predictions lie further away from the closest experimental peaks, with minimum 1H and 15N chemical shift differences of 0.2 ppm and 1 ppm. The differences to the Ala peaks are even more striking, with differences of 2 ppm or more in the 15N dimension.

The method based on the database information generated random coil chemical shift values that had more overlap with the experimental values than those obtained with the method using the polypeptide in urea. The random coil predictions from the database method gave chemical shifts for both 1H and 15N that located more proximally to the observed values. The calculated peak for GVX, for instance, is coincident with peak (6). The random coil shifts of GVP and very small amounts of the closely related peptides (“GVP+”), however, are further upfield in the 1H dimension and slightly downfield in the 15N scale. There is good overlap for AAX and AGX, too. However, the predicted 15N chemical shifts for the other glycine peaks are downfield by 1 ppm or more from the observed peaks. Similarly, the GAX and KAX values from experiment and calculation show poor agreement. Furthermore, the peak (5) in the observed data is not close to any of the predicted random coil peaks, just as in the comparison with the method based on the denatured peptides.

The two predictive methods have common features, regarding the prevalence of the random coil in elastin. Generally, 15N chemical shifts are observed downfield of the random coil shifts that are predicted for the bulk of the predominant glycines, alanines, and valines of elastin’s mobile domains. If chemical shifts are the primary basis for structural predictions, then these results indicate that random coils are possible or most likely for GVX. There are some conflicting results with these two methods, however.

To help resolve this conflict, data for this enriched synthetic elastin sample were compared with the spectra of elastin’s domain 26 (D26) (33) and the flexible N-terminal region of the human prion peptide (huPrP) (18). The prion segment and isolated tropoelastin domain provided additional experimental values on known random coil, or unstructured, segments, both with comparable, predominantly hydrophobic compositions. The comparison with the D26 peptide is shown in Figure 4. The prion yields a similar conclusion (not shown). All Gly 1HN shifts in the D26 peptide (“GVP+”), however, are further upfield in the 1H dimension and slightly downfield in the 15N scale. There is good overlap for AAX and AGX, too. However, the predicted 15N chemical shifts for the other glycine peaks are downfield by 1 ppm or more from the observed peaks. Similarly, the GAX and KAX values from experiment and calculation show poor agreement. Furthermore, the peak (5) in the observed data is not close to any of the predicted random coil peaks, just as in the comparison with the method based on the denatured peptides.
peaks in D26 are downfield of the observed chemical shifts. Interestingly, the values from the domain 26 data have good agreement with the method using the Ac-GGXGG-NH₂ in denaturing conditions. Certainly, the context of the Gly-rich environment appears to be significant, as such a composition is not typically well-represented in the respective databases. Generally speaking, there is no compelling evidence to support a random coil model for elastin at physiological temperatures, although there may be some segments that are unstructured.

In the spectrum comparing β structure, some of the peaks with Gly content (GGX, AGX, PGX) are close to the cluster of glycine peaks (1)-(3). However, the overlap of calculated and experimental peaks is poor in the downfield regions corresponding to the observed peaks (4)-(7). Generally, the ¹⁵N chemical shifts of the observed peaks are upfield of the predicted peaks by 2 or more ppm. Striking contrasts between predicted and experimental shifts, for instance, are seen with VGX, which is found 35 times in tropoelastin. Most of the alanines are also far from the peaks (5)-(7).

In the case of the α-helix, the alanine-centered sequences show good agreement with peaks (6) and (7), which is encouraging, as the (Ala-rich) crosslinking domains have α-helical structure. It is interesting that the α-helical Ala’s would be observed, as it has been a common conclusion that the crosslinking regions are the rigid segments. Such a finding supports a more complicated distribution of dynamics over the entire protein. That is, these data indicate that the α-helical regions are still undergoing rapid motion, retaining secondary structure while still moving rapidly enough to average the strong dipolar couplings that are typical of proteins in the solid state. As for the upfield region (Figure 3H), there is some overlap in the region near the predicted peak (2), but this result is questionable, given the predominance of glycines and the lack of such structures to assume true α-helices.

Overall, there are no predicted peaks that overlap elastin’s peaks (4) and (5) in all of the comparisons shown in Figures 3 and 4. These two peaks appear prominently in the 1D and 2D spectra, accounting for large populations in the protein.

Here we note that the above comparative analysis was also applied to data acquired at 5°C (Supplemental Figure S4). The chemical shift differences are small, but they are enough to show that the protein assumes more random coil content, with a greater number of peaks overlapping with the predicted (random coil) values.

As in other studies, there are some portions of the protein that may be reasonably described as α-helix, β-sheet, or random coil. Yet, none of these three categorizations adequately describe the great bulk of the protein. Certainly, it is possible that the other proposed structures for the protein, such as PPII or the β-turns, are a better fit, but so far, there is no predictive and systematic strategy or algorithm that allows such a comparison. Certainly, the context of the Gly-rich environment appears to be significant, as such a composition is not typically well-represented in the respective databases. Generally speaking, however, there is no compelling evidence to support a random coil model for elastin at physiological temperatures, although there may be some segments that are unstructured.

**Concluding remarks.** The structural picture of elastin has largely been unresolved, with some maintaining that it is completely unstructured (43), lacking the discrete motifs that characterize many other proteins. Certainly, there are many studies that do not support this model (4-6,44). Furthermore, none of the data on native elastin or related peptide mimetics in the solid state support the β-spiral (45) and the oiled coil (46) models, both of which call for more structured environments in this protein. As a middle ground of sorts, Tamburro proposed a conformational equilibrium, in which multiple structures were present in elastin that was “at-rest” or unstretched (7,47). In this model, when the elastin is stretched, there is a resultant shift of equilibrium position, as the various conformations are repopulated. The changes in the populations are effected by the stretching and then restoration of fiber shape, analogous to thermal effects in other systems. This study supports this model, as do various NMR experiments on native elastin and its mimetics that...
have preceded it (4-6,8,9). While a small proportion of the protein may be unstructured, there is mounting evidence, particularly in the form of multiple, resolved shifts in the $^{15}$N, $^{13}$C, and $^1$H NMR spectra, that indicate that Tamburro’s conformational equilibrium is most consistent with the experimental data of native elastin and its various synthetic analogues. Finally, these experiments demonstrate that the use of J-based methods in one and two dimensions are feasible for this system, and they also indicate that the analogous experiments targeting the $^{13}$C nuclei will be essential for determining the exact nature of the multiple conformations and populations of this complex system.

References
Acknowledgement

The authors gratefully acknowledge Dr. Joel Mackay for providing the chemical shifts from the solution NMR studies of tropoelastin’s domain 26. This work was partially supported by grants to KKK from the National Science Foundation (MCB-0344975 and MCB-1022526). ASW acknowledges support from the Australian Research Council.
Figure Captions

Figure 1. 1D $^{15}$N MAS spectra of $[U-^{15}\text{N}]$ synthetic elastin with DP (2048 scans), refocused INEPT (4096 scans), and CP (4096 scans) at (A) 37 °C and (B) 5 °C. The sum of the rINEPT and CP spectra closely approximate the DP spectrum, except for the Pro region. The spinning speed was 8 kHz. All spectra were processed with 20 Hz line broadening.

Figure 2. 2D $^1\text{H-}^{15}\text{N}$ rINEPT-HETCOR spectra of $[U-^{15}\text{N}]$ synthetic elastin taken at (A) 37 °C, (B) 29 °C, (C) 21 °C, (D) 13 °C and (E) 5 °C. The spinning speed was 8 kHz. Major features are labeled (1)-(7) and listed in Table 1. All spectra were processed with 20 Hz line broadening in the direct dimension, and 1 Hz line broadening in the indirect dimension.

Figure 3. Comparison of calculated and experimental (37 °C) chemical shifts, using the semi-empirical methods of Refs. 41 (denatured peptide measurements) and 42 (compiled database information). Comparisons are made with the random coil (3A-3D), β-strand (3E-3F), and the α-helix (3G-3H). The method using the denatured peptide (ref. 41) is used for 3A-3B, whereas the database method (ref. 42) is used for the rest. The experimental spectra are illustrated with the grey contours. Color contours denote the calculated $^{15}$N and $^1\text{H}$ chemical shifts for the Gly (G), Ala (A), and Val (V), with assigned amino acid in bold. An “X” indicates 2 or more residues in the 3rd position of the 3-aa sequence. The predicted "peaks" were plotted as Gaussians that are centered at the (predicted) shifts and weighted by the number of occurrences in the sequence of tropoelastin.

Figure 4. (A) Sequence of domain 26 in tropoelastin (ref. 33). (B) Amide region from rINEPT-HETCOR data at 37 °C. Black plus symbols (+) indicate the chemical shifts of the domain 26 peptide.
Table 1. $^{15}$N and $^1$H chemical shifts (ppm) of the backbone amide region of $[U^{15}N]$ synthetic elastin at 37 °C, 21 °C and 5 °C. These values are extracted from the DP and rINEPT-HETCOR spectra of Figures 1 and 2.

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Measurement</th>
<th>Gly NH</th>
<th>NH (except for Gly NH)</th>
<th>Pro N</th>
</tr>
</thead>
<tbody>
<tr>
<td>37 °C</td>
<td>1D, $^{15}$N</td>
<td>107.6$^{(a)}$</td>
<td>110.7</td>
<td>117.6</td>
</tr>
<tr>
<td></td>
<td>2D, $^{15}$N</td>
<td>106.3$^{(b)}$</td>
<td>107.3</td>
<td>107.9</td>
</tr>
<tr>
<td></td>
<td>2D, $^1$H</td>
<td>8.20$^{(b)}$</td>
<td>8.13</td>
<td>8.31</td>
</tr>
<tr>
<td>21 °C</td>
<td>2D, $^{15}$N</td>
<td>106.7$^{(b)}$</td>
<td>107.3</td>
<td>108.2</td>
</tr>
<tr>
<td></td>
<td>2D, $^1$H</td>
<td>8.26$^{(b)}$</td>
<td>8.15</td>
<td>8.33</td>
</tr>
<tr>
<td>5 °C</td>
<td>1D, $^{15}$N</td>
<td>107.8$^{(a)}$</td>
<td>108.5$^{(a)}$</td>
<td>111.8</td>
</tr>
<tr>
<td></td>
<td>2D, $^{15}$N</td>
<td>107.3</td>
<td>107.5</td>
<td>108.5</td>
</tr>
<tr>
<td></td>
<td>2D, $^1$H</td>
<td>8.35</td>
<td>8.19</td>
<td>8.48</td>
</tr>
</tbody>
</table>

(a) Peaks (1)-(3) were not well resolved on 1D spectrum.
(b) The top of the peak was not apparent on 2D spectrum.
Figure 1

A) 37 °C

Other residues (Ala, Val etc.)
Pro N
Gly
DP
CP
CP+ riNEPT
riNEPT

B) 5 °C

DP
CP
CP+ riNEPT
riNEPT

150 100 50 0
16N chemical shift / ppm

Arg Ne (NH)
Arg N\text{\textgreek{eta}} (NH_2)
Lys Ne (NH+)
Figure 2

A) 37 °C

B) 29 °C

C) 21 °C

D) 13 °C

E) 5 °C
Figure 3

GGXGG in urea, random coil

BMRB and PDB, random coil

BMRB and PDB, β-strand

BMRB and PDB, α-helix

16
Resolving nitrogen-15 and proton chemical shifts for the mobile segments of elastin with two-dimensional NMR spectroscopy
Kosuke Ohgo, Walter P. Niemczura, Brian C. Seacat, Steven G. Wise, Anthony S. Weiss and Kristin K. Kumashiro

*J. Biol. Chem. published online April 1, 2012*

Access the most updated version of this article at doi: [10.1074/jbc.M111.285163](http://dx.doi.org/10.1074/jbc.M111.285163)

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2012/04/01/M111.285163.DC1