# Fluorine-19 Nuclear Magnetic Resonance Spectroscopic Study of Fluorophenylalanine- and Fluorotryptophan-Labeled Avian Egg White Lysozymes†

Chenyang Lian,‡,⊥ Hongbiao Le,‡,⊥ Bernard Montez,‡ Jessica Patterson,‡,∥ Shannon Harrell,‡ David Laws,‡,∥ Ichiro Matsumura,§ John Pearson,‡,▽ and Eric Oldfield*,‡

*Department of Chemistry, University of Illinois at Urbana-Champaign, 505 South Mathews Avenue, Urbana, Illinois 61801, and Department of Molecular and Cell Biology, University of California at Berkeley, Berkeley, California 94720*

*Received November 24, 1993; Revised Manuscript Received March 3, 1994*®

ABSTRACT: We report the 470-MHz (11.7 T) $^{19}F$ solution nuclear magnetic resonance (NMR) spectra of 2-, 3-, and 4-fluorophenylalanine incorporated into the egg white lysozymes (EC 3.2.1.17) of chicken, pheasant, and duck, as well as spectra of 4-fluorotryptophan incorporated into chicken, California valley quail, and Bob White quail and 5- and 6-fluorotryptophan-labeled chicken lysozyme. The $^{19}F$ solution NMR spectrum of [4-F]Phe hen egg white lysozyme (HEWL) consists of three sharp resonances, which span a total chemical shift range of 4.8 ppm (at $p^2H = 6.1$). For [3-F]Phe HEWL, the chemical shift range is much smaller, 1.0 ppm (at $p^2H = 5.9$), due presumably to the occurrence of fast phenyl ring flips about the $C^\beta$–$C^\gamma$ bond axis. For [2-F]Phe HEWL, six resonances are observed, spanning a chemical shift range of 7.4 ppm (at $p^2H = 5.8$), due to slow $C^\beta$–$C^\gamma$ ring flips, i.e., both ring-flip isomers appear to be "frozen in" because of steric hindrance. Rotation of the [2-F]Phe residues remains slow up to 55 °C ($p^2H = 4.7$). With the [F]Trp-labeled proteins, we find a maximal 14.6-ppm shielding range for [4-F]Trp HEWL but only a 2.8- and 2.4-ppm range for [5- and 6-F]Trp HEWL, respectively, due presumably to increased solvent exposure in the latter cases. Guanidinium chloride denaturation causes loss of essentially all chemical shift nonequivalence, as does thermal denaturation. Spectra recorded as a function of pH show relatively small chemical shift changes (<1.4 ppm) over the pH range of ~1.2–7.8. In addition, spectra of highly acetylated [4-F]Phe and [4-F]Trp HEWLs, in which most lysine side chains are converted to (neutral) acetamides (as determined by electrospray ionization mass spectrometry) also show only minor chemical shift changes, although Phe-3 (which is 3.71 Å from the N-terminal lysine) becomes shielded by ~1.5 ppm on acetylation. About 1–1.5-ppm shielding changes were also seen among the [4-F]Trp lysozymes of hen, California valley quail, and Bob White quail and appear to be due to minor side-chain differences (e.g., Val↔Ile, Ser↔Thr) rather than to surface charge field modifications (Gln→His). These results suggest that surface charge fields make only a small contribution to $^{19}F$ shielding. Preliminary assignments of [4-F]Trp HEWL expressed in *Saccharomyces cerevisiae* have been made by using W62Y and W63Y mutants, and $^2H$ solvent-induced shifts were consistent with these assignments. Iodine and *N*-bromosuccinimide oxidation and TEMPO acetamide and $Gd^{3+}$ binding cause line-broadening, which yields tentative assignments for some of the other peaks. Finally, we investigated the effects of inhibitor binding to [4-F]Trp HEWL. We find fast, intermediate, and slow chemical exchange behavior, respectively, on binding *N*-acetyl-D-glucosamine, *N*,*N*′-diacetylchitobiose, and *N*,*N*′,*N*″-triacetylchiototriose ($(NAG)_3$) inhibitors. There are modest (~2 ppm) shielding changes for two resonances, tentatively assigned to Trp-63 and Trp-108, with the 16.8-ppm $^{19}F$ chemical shift range for [4-F]Trp HEWL/$(NAG)_3$ being the largest observed so far in proteins. Overall, our results indicate that $^{19}F$-labeled amino acids can be readily incorporated (within a few days) into avian lysozymes, that spectra can begin to be assigned by means of interspecies comparisons and site-directed mutagenesis, that ortho fluorine substitution presents a large steric hindrance to phenyl ring rotation, and that surface charge fields play only a small role in $^{19}F$ shielding, while (neutral) inhibitor binding or small amino acid side-chain changes appear to cause larger shielding effects than do surface charge field modifications.

$^{19}F$ NMR studies of proteins began in the late 1960s and early 1970s with use of $^{19}F$ probes such as $CF_3CO$ or $CH_2F$ [see, e.g., Gerig (1989) and Phillips et al. (1991)]. Subsequently, a number of groups used bacterial and animal systems to produce proteins containing $^{19}F$-labeled aromatic amino acids. For example, Kimber et al. labeled dihydrofolate reductase with [6-$^{19}F$]Trp (Kimber et al., 1977, 1978); Hull and Sykes investigated [*m*-F]Tyr-labeled alkaline phosphatase (Hull & Sykes, 1975, 1976); Browne and Otvos studied [4-$^{19}F$]-Trp-labeled alkaline phosphatase (Browne & Otvos, 1976); Chaiken et al. investigated semisynthetic ribonuclease-S′ labeled with [4-F]Phe (Chaiken et al., 1973) and [4-F]His (Taylor et al., 1981); Ho et al. investigated [5-F]Trp J-protein from *Salmonella typhimurium* (Robertson et al., 1977), as well as [4-, 5-, and 6-F]Trp-labeled lactate dehydrogenase (Peersen et al., 1990; Ho et al., 1989); Gerig et al. investigated [4-F]Phe-labeled hemoglobins and carbonic anhydrase (Gamc-

† This work was supported in part by the United States Public Health Service (National Institutes of Health Grant HL-19481) and in part by the Director, Office of Energy Research, Office of Basic Energy Science, Divisions of Material Sciences and Energy Biosciences of the U.S. Department of Energy, under contract DE-AC03-76SF0098 (Jack F. Kirsch, Principal Investigator).

‡ University of Illinois at Urbana-Champaign.
§ University of California at Berkeley.
⊥ University of Illinois Graduate Fellow.
∥ University of Illinois Undergraduate Fellow.
▽ United States Public Health Service Postdoctoral Fellow (NIH Grant GM-14545).

® Abstract published in *Advance ACS Abstracts,* April 15, 1994.

0006-2960/94/0433-5238$04.50/0 

sik et al., 1986, 1987; Gerig et al., 1983; Gamcsik & Gerig, 1986); and, most recently, Luck and Falke have produced an elegant series of papers on [3-F]Phe- and [5-F]Trp-labeled D-galactose and D-glucose chemosensory receptors and the transmembrane aspartate transducer (Luck & Falke, 1991a,b,c; Falke et al., 1992).

In many of these studies, no specific assignments were possible, while in the others, there was no analysis of the observed shifts in terms of molecular structure. Thus, a small protein, $^{19}$F-labeled at many sites, stable over a wide pH range, whose detailed X-ray, neutron, and $^1$H NMR solution structures have been extensively studied, available either from many different species (for assignment purposes and investigation of mutation–shift effects) or from genetically engineered derivatives, expressed in a suitable host, containing a large number of aromatic amino acids (HEWL[1] has six Trp, three Phe, three Tyr, and one His), and without a heme group (to eliminate ambiguities from heme ring current paramatrization), would be desirable in order to further pursue theories of $^{19}$F chemical shifts in proteins (de Dios et al., 1993; Pearson et al., 1993). The ability to bind a substrate or an inhibitor is an additional desirable feature, as this would generate even more shielding data, as is the ability to reversibly thermally denature the protein. In our opinion, avian lysozymes appear to be good candidates, since as we show below they can be labeled either biosynthetically or via expression in a suitable fungal (yeast, *Aspergillus*) system (Archer et al., 1990). In what follows, we present results on [$^{19}$F]Phe- and [$^{19}$F]Trp-labeled lysozymes which serve as an entry point to investigating in more detail $^{19}$F shielding, electrostatics, and, potentially, the topic of protein–protein interaction (Sheriff et al., 1987).

## EXPERIMENTAL METHODS

*NMR Aspects.* Fluorine NMR spectra were obtained at 470.3 MHz on a "home-built" Fourier transform NMR spectrometer, which consists of an 11.7-T, 2.0-in. bore superconducting solenoid magnet (Oxford Instruments, Osney Mead, Oxford, U.K.), a Nicolet (Madison, WI) Model 1280 computer, RF Powers Labs (Kirkland, WA) Model 505 and Henry Radio (Los Angeles, CA) Model 701D02L radio frequency amplifiers, and a Cryomagnet Systems (Indianapolis, IN) 5-mm $^{19}$F NMR probe. Spectral widths were typically ±7 kHz, using 2 × 8192 data points, and recycle delay times were typically 1.5 s. Chemical shifts are referenced with respect to an external standard of trifluoroacetic acid at 0 ppm, using the convention that high-frequency, low-field paramagnetic or deshielded values are taken as positive (IUPAC $\delta$ scale).

*Mass Spectrometry.* Electrospray mass spectra of fluorinated lysozymes were obtained on Finnegan-MAT (San Jose, CA) Model TSQ 700, Sciex (Thornhill, Ontario, Canada) Model API III, or VG Biotech (Altringham, U.K.) Quattro Model mass spectrometers.

*Chemical Aspects.* *o*-, *m*-, and *p*-fluorophenylalanines, [4-, 5-, and 6-F]tryptophans, *N*-acetylglucosamine (NAG), *N*,*N*′-diacetylchitobiose ($(NAG)_2$), *N*,*N*′,*N*″-triacetylchitotriose ($(NAG)_3$), *N*-bromosuccinimide, gadolinium chloride, guanidinium chloride (reagent grade), deuterium oxide, and CM-cellulose were all from Sigma Chemical Co. (St. Louis, MO). 4-Acetamido-2,2,6,6-tetramethylpiperidine-1-oxyl (TEMPO acetamide) was from Frinton Laboratories (Vineland, NJ). Acetic anhydride was from Fisher Scientific Co. (Fair Lawn, NJ).

*Poultry Science.* Fluorinated amino acid incorporation was achieved by feeding birds a normal "laying mash" diet, enriched with 0.1% of the requisite amino acid, which was dispersed into the feed using a Wesson (Fullerton, CA) corn oil carrier phase. Chickens were obtained from the University of Illinois Poultry Farm (Champaign, IL), turkeys from Toubl Game Bird Farms (Beloit, WI), pheasants from Loney Sharp (Doyle, CA), ducks from Whistling Wings, Inc. (Hanover, IL), California valley quail from J. Reddy (Bishop, CA), and Bob White quail from Muddy Valley Game Farm (Ingraham, IL). Good quality $^{19}$F NMR spectra were typically obtained from the egg whites of birds which had been on labeled feed for 7–10 days.

*Recombinant HEWL Expression.* The W62Y and W63Y mutants of HEWL were made and expressed using the constructs and strains previously described (Malcolm et al., 1989). The transformed *Saccharomyces cerevisae* strains were grown in histidine-supplemented SD media [see Sherman et al. (1986)] for 48 h and then transferred to 1% Difco Bacto-yeast extract, 8% dextrose, and 1× amino acid supplements (except tryptophan) plus 20 mg/L [4-F]Trp. Ethanol was added to the cultures to 1% daily. After 4 days, the secreted protein was purified from the supernatant as described by Malcolm et al. (1989).

*Lysozyme Preparation.* Lysozyme was isolated from fresh eggs according to Arnheim, Prager, and Wilson (1969) using a CM-cellulose adsorption and elution technique. Sample purity was verified by sodium dodecyl sulfate–polyacrylamide gel electrophoresis using a Bio-Rad (Richmond, CA) Mino-Protean II apparatus and 14% acrylamide gels, stained with Coomassie Brilliant Blue (Laemmli, 1970).

*Chemical Modification.* Acetylation of lysozyme was according to Yamasuki, Hayashi, and Funatsu (1968a,b) using acetic anhydride. The reaction of lysozyme with iodine was carried out as described by Imoto and Rupley (1973) and by Norton and Allerhand (1976).

## RESULTS AND DISCUSSION

We show in Figure 1A the 470-MHz $^{19}$F NMR spectrum of a 5.6 mM solution of [4-F]Phe HEWL, p$^2$H = 4.7, at room (bore) temperature (23 °C). As expected on the basis of previous studies of [4-$^{19}$F,$^3$H-ul)]Phe incorporation into chicken minced oviduct (Vaughan & Steinberg, 1960), there is adequate [4-F]Phe incorporation for NMR purposes. We find, as expected, three resonances of equal intensity, corresponding (but not necessarily on a 1:1 basis) to Phe-3, Phe-34, and Phe-38 in the HEWL sequence (Canfield, 1963). Fluorine incorporation can be followed, in principle, by NMR spin-counting techniques (Luck & Falke, 1991a) or by use of mass spectrometry; by using electrospray ionization mass spectrometry, we find that on average, each HEWL molecule contains about 1.2 fluorine atoms.

Now, in order to begin to assign each of the three resolved resonances in Figure 1A to a specific fluorophenylalanine, we incorporated [4-F]Phe into a number of other avian species, each of which has one or more Phe→Tyr mutations in its primary sequence (Araki et al., 1991; Hermann & Jolles, 1970). Pheasant has Phe-3→Tyr, while duck has Phe-3,-34→Tyr. $^{19}$F NMR spectra of each of the additional species investigated—pheasant and duck—are shown in Figure 1B and C. As may be seen from these results and those presented in Table 1, the most highly deshielded residue in chicken may

[1] Abbreviations: HEWL, hen egg white lysozyme; CQL, California valley quail lysozyme; BQL, Bob White quail lysozyme; NAG, *N*-acetylglucosamine; $(NAG)_2$, *N*,*N*′-diacetylchitobiose; $(NAG)_3$, *N*,*N*′,*N*″-triacetylchitotriose; NBS, *N*-bromosuccinimide.

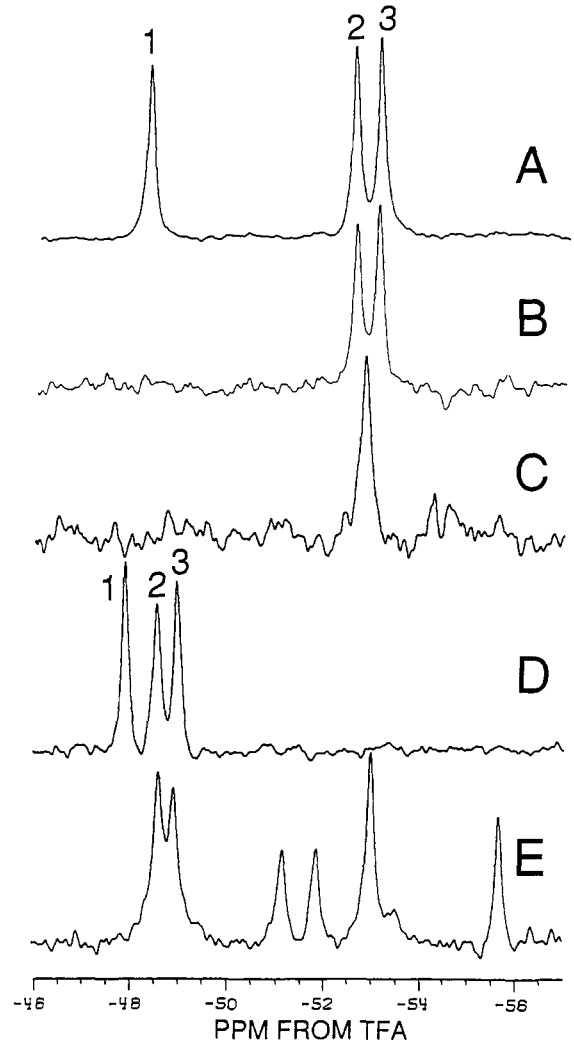


FIGURE 1: 470-MHz $^{19}F$ NMR spectra of fluorophenylalanine-labeled avian lysozymes in $^2H_2O$ at 23 °C. (A) [4-F]Phe HEWL, p$^2$H = 4.47, 5.6 mM, 3152 scans, 40-Hz line-broadening due to exponential multiplication. (B) [4-F]Phe pheasant egg white lysozyme, 4.7 mM, 30 000 scans, 40-Hz line-broadening. (C) [4-F]-Phe duck egg white lysozyme, p$^2$H = 5.7, 10 mM, 400 scans, 40-Hz line-broadening. (D) [3-F]Phe HEWL, p$^2$H = 1.3, 10 mM, 436 scans, 20-Hz line-broadening. (E) [2-F]Phe HEWL, p$^2$H = 1.57, 10 mM, 428 scans, 10-Hz line-broadening.

Table 1: Fluorine NMR Chemical Shifts for [4-$^{19}$F]Phe Avian Lysozymes

| | chemical shift (ppm from TFA) | | |
|---|---|---|---|
| labeled site | hen | pheasant | duck |
| Phe-3 | −48.3 | | |
| Phe-34 | −52.6 or −53.1 | −52.6 or −53.0 | |
| Phe-38 | −53.1 or −52.6 | −53.0 or −52.6 | −52.9 |

be assigned to the nonconserved Phe-3. The overall [4-F]Phe HEWL chemical shift range of 4.8 ppm is relatively small for [F]Phe in a protein, where 8–12-ppm values are more typical (Gamcsik et al., 1986), although this may simply be a reflection of the very small number of sites sampled.

We also incorporated meta and ortho fluorophenylalanines into HEWL, and the 470-MHz $^{19}F$ NMR spectra obtained are shown in Figure 1D and E. With the *meta*-labeled species, Figure 1D, the chemical shift range is exceedingly small, only 1.0 ppm, due presumably to fast axial rotation about $C^\beta$–$C^\gamma$. This will effectively average the fluorine shielding and is in sharp contrast to the *o*-fluorophenylalanine result shown in Figure 1E. Here, instead of three resonances corresponding to the three phenylalanines, we observe *six* major peaks. Based on previous $^1H$ NMR work on phenylalanine rotation in proteins (Wuthrich, 1986), it is logical to infer that ortho fluorine substitution greatly increases the barrier to rotation about $C^\beta$–$C^\gamma$, resulting in two frozen ring-flip conformers for each of the three [2-F]Phe residues. The overall shielding range is 7.4 ppm, considerably larger than that of the para position, and we find (data not shown) even on heating of the sample to 55 °C (at p$^2$H = 4.7) that the shielding pattern remains essentially the same. Thus, $^{19}F$ incorporation in the ortho position causes a significant intraresidue structural change, not observable in the meta position. However, such an effect is reasonable based on the large steric repulsion between $F^\delta$ and $H^\beta$s and could even result in a modification of ring orientation.

[2-F]Phe in solution gives only a single resonance, so the problem can best be thought of as a two-term equation for the barrier height:

$$\Delta E^{TOTAL} = \Delta E^{INTRA} + \Delta E^{INTER}$$

The total barrier to rotation is made up of an intraresidue and an interresidue interaction. For [2-F]Phe in the protein, both terms are large, so we see two peaks. In solution, $\Delta E^{INTER}$ is small, thus one peak. For [3-F]Phe in the protein, $\Delta E^{INTRA}$ is small, thus again only one peak/site is seen. Molecular modeling shows close $H^\beta$–$F^\delta$ interactions in the [2-F]Phe protein, consistent with these ideas.

Because of their relatively small shielding ranges, the *o*-, *m*-, and *p*-fluorophenylalanyl HEWL spectra are expected to be very difficult to compute theoretically using either the *ab initio* or weak electrical interaction models described elsewhere (Pearson et al., 1993; Hinton et al., 1992; de Dios et al., 1993), and, in addition, the [2-F]Phe result implies—certainly for this modification—that it will be essential to independently determine side-chain structure using, e.g., $^1H$,$^{13}C$ NMR methods. However, such spectra are nonetheless very useful for testing ideas about the *types* of interaction which may be important in fluorine shielding (see below).

We have also begun to investigate the $^{19}F$ NMR spectra of the six tryptophan sites in HEWL, which occur at positions 28, 62, 63, 108, 111, and 123 in the primary sequence. Tryptophan is likely to be a better target amino acid with which to investigate electrostatic field effects in proteins, because the indole ring size should impede the sorts of ring-flip processes which can occur readily with the phenyl groups in Phe (and Tyr). Figure 2 shows the 470-MHz $^{19}F$ NMR spectra of HEWL labeled with [4-, 5-, or 6-F]Trp, and there are remarkably large differences between the spectra of the three different [F]Trp-labeled proteins. Such effects have not been observed previously in any other protein. While six Trp resonances can be accounted for in each spectrum, the *range* of chemical shifts in the case of [4-F]Trp is 14.6 ppm, while in the [5- and 6-F]Trp systems, the total shielding ranges are only 2.8 and 2.4 ppm, respectively. It is also clear that the individual component line widths for the [4-F]Trp residues are about 3 times as large as the line widths observed in the case of the [5- or 6-F]Trp species. Such large differential line widths do not appear to arise from spin–spin coupling effects, as deduced from solution NMR studies of the individual amino acids (data not shown), but rather appear to "scale" with the overall shielding range. It is thus necessary to investigate which factors are likely to influence $^{19}F$ shielding in lysozyme and to ask whether these factors could contribute to the large line width differences.

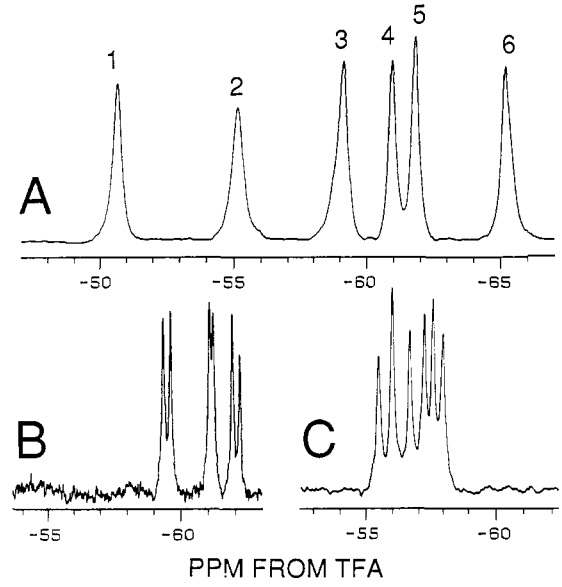


FIGURE 2: 470-MHz $^{19}$F NMR spectra of [4-, 5-, and 6-F]tryptophan-labeled hen egg white lysozymes in $D_2O$ at 23 °C. (A) [4-F]Trp HEWL, 10 mM, p$^2$H = 5.7, 5500 scans, 40-Hz line-broadening. (B) [5-F]Trp HEWL, 10 mM, p$^2$H = 3.54, 400 scans, 10-Hz line-broadening. (C) [6-F]Trp HEWL, 3.5 mM, p$^2$H = 1.21, 4950 scans, 40-Hz line-broadening. Note that the abscissae are all on the same scale.

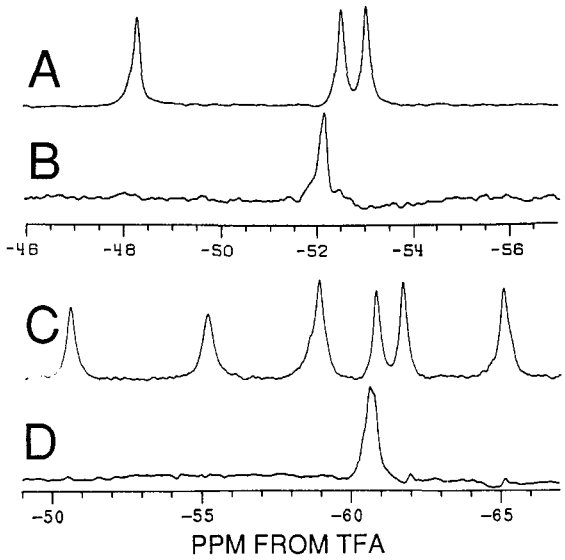


FIGURE 3: 470-MHz $^{19}$F NMR spectra of native and guanidinium chloride denatured 4-fluorophenylalanine- and 4-fluorotryptophan-labeled hen egg white lysozymes in $^2H_2O$ at 23 °C. (A) [4-F]Phe HEWL, 5.6 mM, p$^2$H = 4.47, 3152 scans, 40-Hz line broadening. (B) 6.2 M guanidinium chloride-denatured [4-F]Phe HEWL, p$^2$H = 4.0, 2800 scans, 40-Hz line-broadening. (C) [4-F]Trp HEWL, 5.8 mM, p$^2$H = 4.26, 500 scans, 40-Hz line-broadening. (D) 6.2 M guanidinium chloride-denatured [4-F]Trp HEWL, p$^2$H = 4.0, 10 000 scans, 50-Hz line-broadening.

First, we find that virtually all of the chemical shift nonequivalence observed in the $^{19}$F NMR spectra of *o*-, *m*-, and *p*-fluorophenylalanine and [4-, 5-, and 6-F]Trp-labeled HEWL is due to folding of the protein into its native structure, as determined by chemical (guanidinium chloride) denaturation experiments, as shown in Figure 3, as well as by thermal denaturation (data not shown). As can be seen from Figure 3 and Table 2, the denatured protein chemical shift is generally slightly to lower field than that of the most shielded residue in the protein, and in previous work we have suggested that electrostatic field effects are likely to play a major role in

Table 2: Fluorine NMR Chemical Shifts of Native and Guanidinium Chloride (8 M) Denatured Hen Egg White Lysozymes Containing [*p*-F]Phe, [*m*-F]Phe, [4-F]Trp, [5-F]Trp, and [6-F]Trp Residues at 23 °C[a]

| | chemical shift (ppm from TFA) | | | chemical shift (ppm from TFA) | |
|---|---|---|---|---|---|
| site | native | denatured | site | native | denatured |
| [*p*-F]Phe-1 | −48.3 | −52.2 | [5-F]Trp-1 | −59.4 | −60.9, −61.0 |
| [*p*-F]Phe-2 | −52.6 | | [5-F]Trp-2 | −59.6 | |
| [*p*-F]Phe-3 | −53.1 | | [5-F]Trp-3 | −61.1 | |
| [*m*-F]Phe-1 | −48.1 | −49.6 | [5-F]Trp-4 | −61.2 | |
| [*m*-F]Phe-2 | −48.8 | | [5-F]Trp-5 | −61.9 | |
| [*m*-F]Phe-3 | −49.1 | | [5-F]Trp-6 | −62.2 | |
| [4-F]Trp-1 | −50.6 | −60.6, −60.8 | [6-F]Trp-1 | −55.5 | −57.8, −58.0 |
| [4-F]Trp-2 | −55.2 | | [6-F]Trp-2 | −56.0 | |
| [4-F]Trp-3 | −59.0 | | [6-F]Trp-3 | −56.7 | |
| [4-F]Trp-4 | −60.9 | | [6-F]Trp-4 | −57.2 | |
| [4-F]Trp-5 | −61.8 | | [6-F]Trp-5 | −57.5 | |
| [4-F]Trp-6 | −65.2 | | [6-F]Trp-6 | −57.9 | |

[a] Peak designations refer to line numbers (Figures 1 and 2). Specific and tentative peak assignments are discussed in the text.

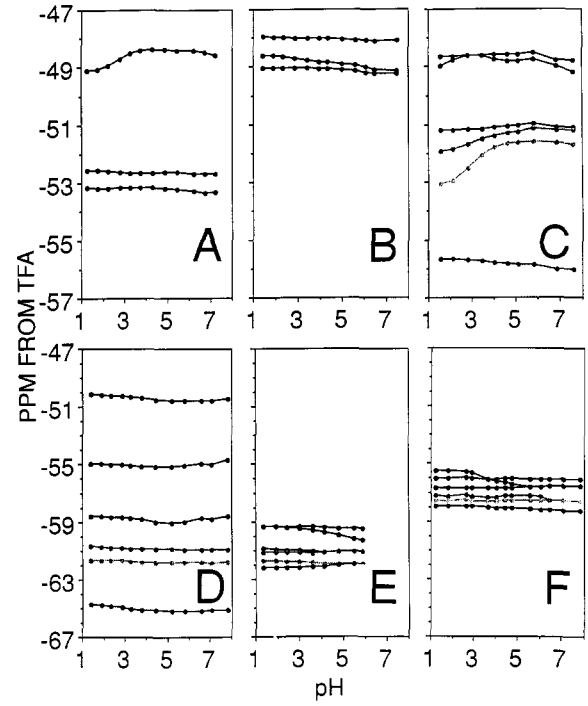


FIGURE 4: $^{19}$F NMR chemical shifts of F-Phe- and F-Trp-labeled hen egg white lysozymes as a function of p$^2$H. Typical concentrations were ≈10 mM, and 200 scans were accumulated per data point, using a 1.5-s recycle time, 90° pulse excitation, and 40-Hz line-broadening. The lines are drawn simply to join the points. (A) [4-F]Phe HEWL. (B) [3-F]Phe HEWL. (C) [2-F]Phe HEWL. (D) [4-F]Trp HEWL. (E) [5-F]Trp HEWL. (F) [6-F]Trp HEWL.

shielding of $^{19}$F resonances in proteins (Augspurger et al., 1992, 1993; Pearson et al., 1993; de Dios et al., 1993). The following question then arises: Where do these electrostatic field effects come from?

One possible candidate is the presence of numerous charged residues on the surface of the protein. We have investigated this possible contribution to shielding in three different ways. First, we have varied surface charge by investigating $^{19}$F spectra over a wide range of pH values (from pH = 1.2 to pH = 7.8). Second, we have removed most lysine surface charges by means of acetylation. Third, we have investigated a number of different avian species, including charge variants near sites of fluorine incorporation. Figure 4 shows chemical shift versus pH titration curves for each [F]Phe and [F]Trp HEWL

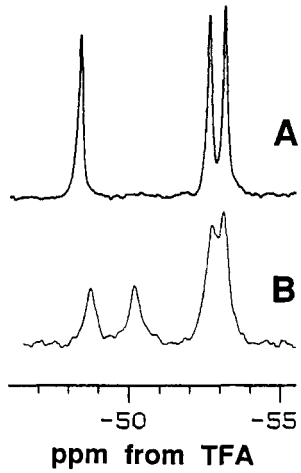


FIGURE 5: 470-MHz $^{19}F$ NMR spectra of native and acetylated [4-F]Phe-labeled hen egg white lysozyme in $^2H_2O$ at 23 °C. (A) Native sample, 8 mM, $p^2H$ = 5.0, 7455 scans, 1.5-s recycle time, 40-Hz line-broadening. (B) Acetylated sample, spectral conditions basically same as in spectrum A.

investigated. Clearly, while there *are* a number of chemical shift changes as a function of pH, they are typically $\lesssim$1.4 ppm—a small effect given a total shielding range of 14.6 ppm for [4-F]Trp. Thus, pH or surface charge field effects are unlikely to be the major cause of the large fluorine chemical shift nonequivalencies, since seven carboxylates are expected to titrate in this pH range (Imoto et al., 1972). This view is reinforced by the results of experiments aimed at removing surface charge from lysine residues by means of acetylation, $-^+NH_3 \rightarrow -NHCOCH_3$. Figure 5 shows the 470-MHz $^{19}F$ NMR spectra of [4-F]Phe HEWL as both the native protein, Figure 5A, and after acetylation (with acetic anhydride; Yamasaki et al. 1968a,b), Figure 5B. Acetylation was verified by means of electrospray ionization mass spectrometry, and typical results for native and an acetylated [4-$^{19}F$]Phe lysozyme are shown in Figure 6A and B. As can be seen in Figure 6B, acetylation results in a series of M + 42*n* peaks, where *n* = 3–7. This is the result expected for acetylation of the six surface lysines plus the terminal amino group, basically as reported elsewhere by others using classical analytical techniques. Electrospray ionization mass spectrometry permits, however, a much more accurate and rapid product distribution analysis, and we deduce from Figure 6 that ~80% of all lysines are acetylated in this particular sample.

In the $^{19}F$ NMR spectra of Figure 5, it is clear that there are no significant shielding changes for Phe-34 and Phe-38, although the shielding of some Phe-3 residues has increased by ~1.2 ppm. Upon inspection of the X-ray crystallographic structure(s) of HEWL (Blake et al., 1967; Bernstein et al., 1977; Abola et al., 1987), we find that $F^\zeta$ of Phe-3 is only ~3.7 Å from the N-terminal nitrogen. This is a very close contact—much less than the lysine $^+NH_3$–F distances of 6.0–30.0 Å observed for all other Phe (and Trp) pairs in HEWL, accounting in at least a qualitative fashion for the observation of a shift at this site (which may even be structural in nature) on acetylation. Similar small shielding changes are observed in [4-F]Trp HEWL upon acetylation (data not shown), consistent with the fluorophenylalanine results shown in Figure 5.

The possible effect of surface charge upon shielding was also investigated by comparing $^{19}F$ NMR spectra of several almost identical [4-F]Trp avian lysozymes. The amino acid sequences of the lysozymes from chicken, California valley quail (CQL), and Bob White quail (BQL) are very similar, differing at only five positions between the three proteins, as shown in Table 3. Three of the amino acid variations are $CH_2$ group additions/subtractions; the fourth substitutes lysine for arginine, while the fifth substitutes histidine for glutamine (see Table 3), and we expect His-121 to be protonated in the pH range 1–5 (since in model-building experiments it appears exposed). The 470-MHz $^{19}F$ NMR spectra of each of the three species are shown in Figure 7. Four of the sites have essentially identical chemical shifts, while the chemical shifts

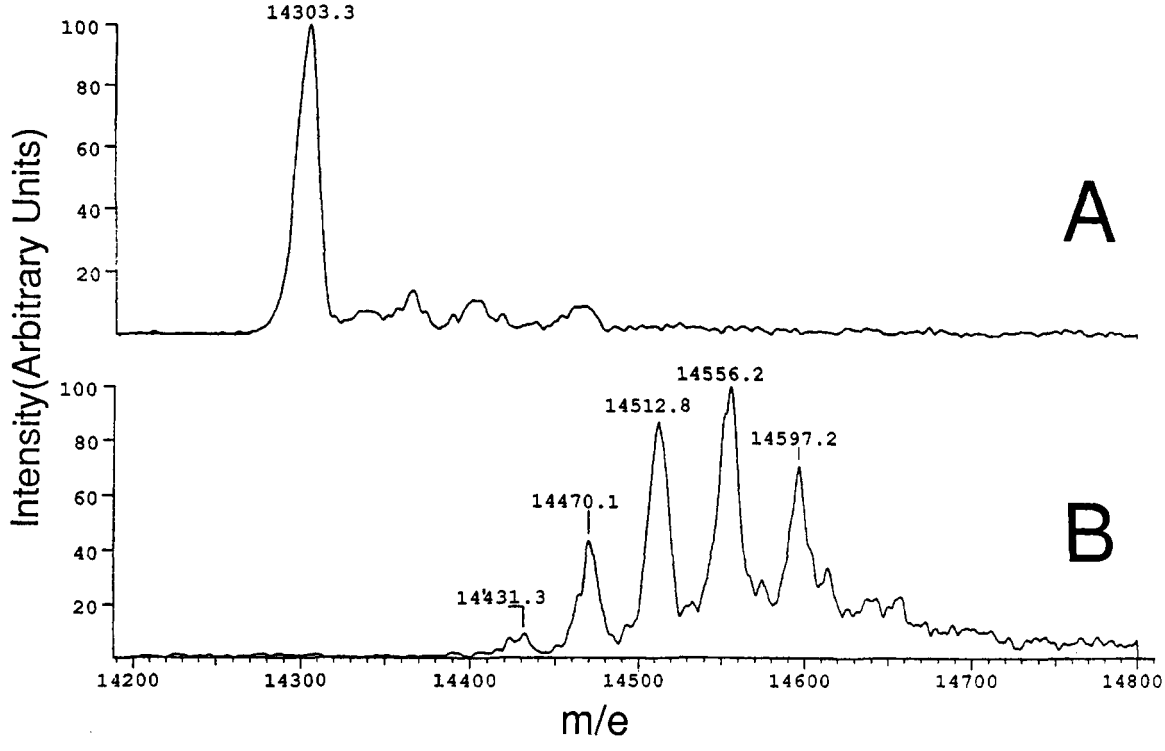


FIGURE 6: Electrospray ionization mass spectra of native and acetylated hen egg white lysozyme. (A) Native HEWL. (B) Acetylated [4-F]Phe HEWL.

Table 3: Amino Acid Sequence Differences between Hen Egg White Lysozyme (HEWL), California Valley Quail Egg White Lysozyme (CQL),[a] and Bob White Quail Egg White Lysozyme, (BQL)[a]

| | residue | | | | |
|---|---|---|---|---|---|
| protein | 40 | 55 | 68 | 91 | 121 |
| HEWL | Thr | Ile | Arg | Ser | Gln |
| CQL | Ser | Val | Arg | Thr | His |
| BQL | Ser | Val | Lys | Thr | Gln |

[a] Sequence data obtained from Prager et al. (1972) and Ibrahim et al. (1979)

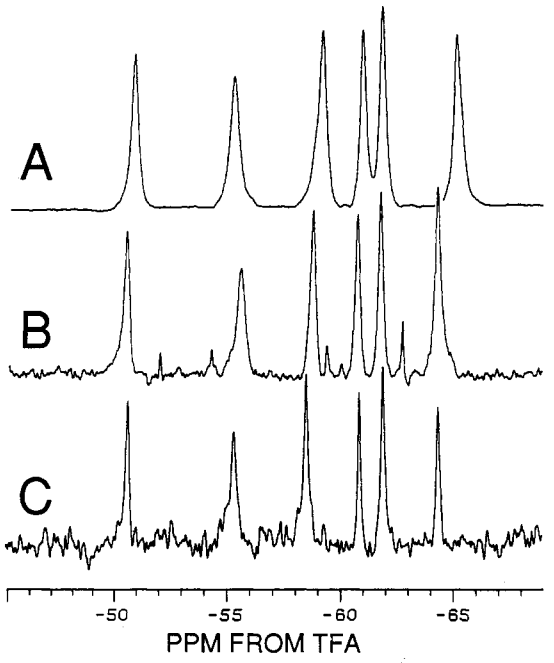


FIGURE 7: 470-MHz $^{19}$F NMR spectra of three [4-F]Trp-labeled avian lysozymes in $^2H_2O$ at 23 °C. (A) HEWL, 8.0 mM, p$^2$H = 4.75, 5500 scans, 40-Hz line-broadening. (B) California valley quail egg white lysozyme, 8.4 mM, p$^2$H = 5.09, 34 432 scans, 40-Hz line-broadening. (C) Bob White quail egg white lysozyme, 8.0 mM, p$^2$H = 4.77, 5900 scans, 40-Hz line-broadening.

of two of the sites vary among the three proteins. These changes could, in principle, be due to either the methylene additions/subtractions which occur at residues 40, 55, and 91 or the change in charge field due to the substitution of histidine for glutamine at residue 121. To determine which type of effect dominates, we measured the $^{19}$F spectrum of [4-F]Trp BQL at pH = 4.8 and compared this spectrum to those of [4-F]Trp HEWL at pH = 4.8 and [4-F]Trp CQL at pH = 5.0, as shown in Figure 7. The observation that the BQL spectrum resembles more closely the CQL spectrum (root mean square shielding difference = 0.16 ppm) than the HEWL spectrum (root mean square shielding difference = 0.39 ppm) suggests that alkylation changes between the various proteins have a much stronger influence upon chemical shifts, due to differences in side-chain packing, than do changes in surface charge fields. Of course, if the His-121 were not protonated in the pH range investigated, then alkylation changes would still be the cause of the hen/quail shielding differences.

Thus, the pH titration, acetylation, and interspecies comparison results strongly suggest that surface charge fields play only a small role in determining $^{19}$F shielding in proteins. In contrast, even minor changes in amino acid side-chain structure cause 1–1.5-ppm changes in shielding, and as we have suggested elsewhere (Pearson et al., 1993), our results are consistent with the idea that the electrostatic interactions

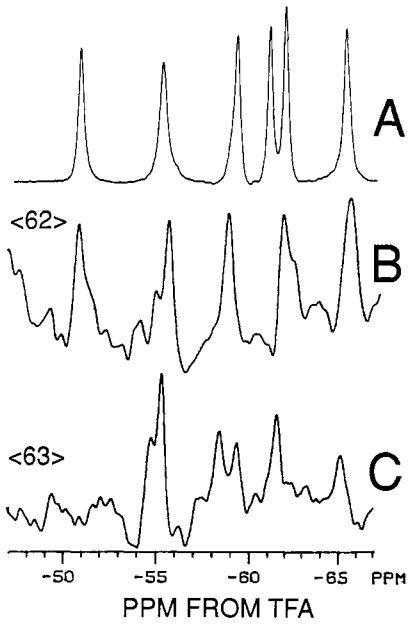


FIGURE 8: 470-MHz $^{19}$F NMR spectra of normal and mutant [4-F]-Trp-labeled hen egg white lysozymes. (A) Normal HEWL. (B) W62Y mutant, 229 000 scans. (C) W63Y mutant, 373 900 scans.

of major importance are due to (1) fixed dipoles within the protein interior and, in some cases, (2) interactions with solvent (water) molecules. If this is correct, we expect major shielding changes for interactions which would modify 1 or 2 above. In lysozyme, this might be accomplished by, for example, inhibitor or immunoglobin binding. Of course, for our results to be really useful, we need to specifically assign and eventually be able to compute the various fluorine chemical shifts observed above, and we report our initial steps in this direction in what follows.

As shown previously by Luck and Falke (1991a), such assignments are best made by using proteins prepared via site-directed mutagenesis. Recombinant strains of yeast expressing the W62Y and W63Y mutants of HEWL were thus grown in [4-F]Trp-supplemented media. The yeast were apparently able to incorporate [4-F]Trp in place of the natural analog. Expression of recombinant fluorinated protein was low, but as shown in Figure 8, the peaks attributable to Trp-62 and Trp-63 can be at least tentatively assigned. Given these tentative assignments, we then investigated the $^2$H solvent isotope shift, which is generally thought to cause a ~0.1–0.25-ppm upfield shift for solvent-exposed residues, in $^2H_2O$. For peaks 1–6 shown in Figure 2A, we find ($\delta_{H_2O}$–$\delta_{D_2O}$) solvent shifts of –0.02 (peak 1), –0.02 (peak 2), +0.04 (peak 3), +0.11 (peak 4), +0.03 (peak 5), and –0.02 (peak 6). Based on the tentative assignments deduced from Figure 8, the $^2$H SIS of Trp-62 is thus the largest (0.11 ppm), consistent with its generally larger solvent accessibility in both X-ray and solution NMR structures.

On binding saccharide inhibitors to [4-F]Trp HEWL, we observe that the shieldings of three peaks change, Figure 9. These peaks should be attributable, based on the X-ray and previous NMR studies, to Trp-62, -63, and -108 (Dobson & Williams, 1975; Sykes & Parravano, 1969; Thomas, 1967), and indeed the results of Figure 8 appear to be consistent with this expectation.

We find that binding of the mono- and disaccharide inhibitors *N*-acetylglucosamine (NAG) and *N*,*N*′-diacetylchitobiose ($(NAG)_2$), Figure 9B and C, results in generally similar shielding patterns, with fast exchange behavior for

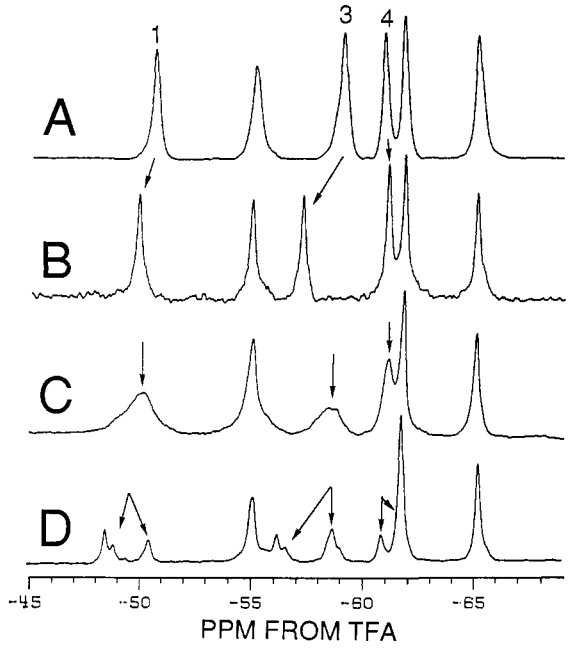


FIGURE 9: 470-MHz $^{19}F$ NMR spectra of [4-F]Trp-labeled hen egg white lysozyme without and with various *N*-acetylglucosamine (NAG) inhibitors in $^2H_2O$ at 23 °C. (A) 5.8 mM HEWL, $p^2H = 4.26$, 500 scans, 40-Hz line-broadening. (B) 4 mM HEWL + 0.6 M NAG, $p^2H = 5.14$, 9000 scans, 40-Hz line-broadening. (C) 5.8 mM HEWL + 3 mM $(NAG)_2$, $p^2H = 5.43$, 28 000 scans, 40-Hz line-broadening. (D) 5.85 mM HEWL + 4.0 mM $(NAG)_3$, $p^2H = 4.0$, 11 500 scans, 40-Hz line-broadening. The arrows trace the chemical shift connectivities on inhibitor binding.

NAG and intermediate exchange for $(NAG)_2$. For $(NAG)_3$, Figure 9D, slow exchange behavior is seen. The exchange rates clearly anticorrelate with the hydrogen-bonding properties of the respective saccharides, and as can be seen from Figure 9, it is a simple matter to follow the shielding changes for peaks 1, 3, and 4 from HEWL to HEWL/NAG to HEWL/$(NAG)_2$ to HEWL/$(NAG)_3$. In addition, since the Trp-62 and Trp-63 assignments can be tentatively deduced from Figure 8, it seems logical to suggest that the remaining peak which undergoes a large shift on $(NAG)_n$ binding arises from Trp-108, although additional experiments with a specific Trp-108 mutant will be required in order to test this hypothesis.

We also found (data not shown) that the HEWL/$(NAG)_2$ spectrum can be readily converted from the intermediate exchange regime (at 35 °C) to slow exchange (5 °C) or fast exchange (70 °C), consistent with the results of previous workers, again confirming the correlations shown in Figure 9. Also of interest in the slow exchange spectrum (Figure 9D) is the observation that the two peaks that shift most on $(NAG)_3$ binding have well-resolved "doublet" splittings, and since $(NAG)_3$ can exist in both $\alpha$- and $\beta$-anomeric forms, it seems likely that these small additional shielding nonequivalencies have their origins in sugar anomerism.

Finally, in a further attempt to obtain peak assignments, we carried out a series of "standard" chemical modification and additive experiments which have been used previously to make peak assignments in HEWL [see, e.g., Allerhand et al. (1977)]. Specifically, we investigated $Gd^{3+}$ binding (Campbell et al., 1973), TEMPO acetamide binding (Wien et al., 1972), *N*-bromosuccinimide oxidation (Hayashi et al., 1965), and iodine oxidation (Hayashi et al., 1968a,b). $Gd^{3+}$ broadened peaks 3 and 4 (numbering given in Figure 3A), and TEMPO acetamide broadened peak 4, consistent with the assignments suggested above. NBS and $I_2$ oxidation results were more complex, due, we believe, to the decreased reactivity of the [F]Trp nucleus. NBS caused additional peaks in the region of peak 1 (Trp-63), consistent with preferential oxidation of unlabeled Trp-62 sites (adjacent Trp-63). Clean NBS oxidation to the Trp-62 oxindole was thus not accomplished. $I_2$ oxidation of Trp-108 was also retarded by fluorine substitution, although at pH 8.5, iodination of Tyr-20, -23 caused broadening of peak 2, consistent with a tentative assignment to $F^\epsilon$ of Trp-111, which is 6.0 Å from $H/I^\epsilon$ of Tyr-23.

## CONCLUSIONS

The results we have reported above are of interest for several reasons. First, we have obtained the first $^{19}F$ NMR spectra of a variety of $^{19}F$-labeled amino acids incorporated into lysozymes from several different avian species. Second, our results are consistent with rapidly rotating *m*- and *p*-fluorophenylalanine groups in HEWL, while in the *o*-fluorophenylalanine-labeled system, fluorine substitution appears to lead to a locking-in of ring-flip side-chain conformers. Here, fluorine clearly cannot be regarded as a nonperturbing probe. Third, our results with [4-, 5-, and 6-F]-Trp-labeled HEWL show remarkable differences. The [4-F]-Trp HEWL spectrum has a 14.6-ppm shielding range and ~150-Hz line widths, while the [5- and 6-F]Trp species have only 2.8- and 2.4-ppm ranges of chemical shift nonequivalence, respectively, and ~50-Hz line widths. One possible explanation for this is that the 5- and 6-F sites are in regions of higher dielectric constant and thus have weaker electric field-induced shifts (de Dios et al., 1993; Pearson et al., 1993). Fourth, our results show that surface charge fields have little effect on $^{19}F$ chemical shifts, as deduced from the results of pH titration, acetylation, and interspecies comparisons. Fifth, our results show that only a few minor side-chain changes cause 1–1.5-ppm shielding differences for four out of six [4-F]Trp residues—due possibly to more global readjustments of the (buried) peptide charge field on substitution. Sixth, as anticipated, we find substantial shielding changes for three residues on inhibitor binding and slow, intermediate, and fast exchange behaviors are observed with $(NAG)_3$, $(NAG)_2$ and NAG, respectively. Seventh, we have obtained $^{19}F$ NMR spectra of W62Y and W63Y recombinant proteins, expressed in *S. cerevisiae*, which permit tentative assignment of Trp-62, Trp-63, and, by inference, Trp-108.

## ACKNOWLEDGMENT

We thank Professor J. Kirsch for providing laboratory facilities (I.M.), Professor T. Imoto for helpful comments and assistance, Dr. Richard Milberg for obtaining the electrospray mass spectra, Dr. B.-R. Fang for carrying out some of the inhibitor binding experiments, Ellen Chien for solvent accessibility calculations, and William L. Heckt (Agricultural Experiment Station) for help with lysozyme production.

## REFERENCES

Abola, E., Bernstein, F. C., Bryant, S. H., Koetzle, T. F., & Weng, J. (1987) Protein Data Bank, in *Crystallographic Databases Information Content, Software Systems, Scientific Applications*, (Allen, F. H., Bergerhoff, G., & Sievers, R., Eds.) pp 107–132, Data Commission of the International Union of Crystallography, Bonn/Cambridge/Chester.

Allerhand, A., Norton, R., & Childers, R. (1977) *J. Biol. Chem. 252*, 1786–1794.

Araki, T., Kuramoto, M., & Torikata, T. (1991) *Agric. Biol. Chem. 55*, 1707–1713.

Archer, D. V., Jeenes, D. J., MacKenzie, D. A., Brightwell, G., Lambert, N., Lowe, C., Radford, S. E., & Dobson, C. M. (1990) *Bio/Technology 8*, 741–745.
Arnheim, N., Prager, E. M., & Wilson, A. D. (1969) *J. Biol. Chem. 244*, 2085–2094.
Augspurger, J., Pearson, J., Oldfield, E., Dykstra, C. E., Park, K. D., & Schwartz, D. (1992) *J. Magn. Reson. 100*, 342–357.
Augspurger, J. D., Dykstra, C. E., Oldfield, E., & Pearson, J. P. (1993) in *Nuclear Magnetic Shieldings and Molecular Structure*, (Tossell, J., Ed.) pp 75–94, Kluwer Academic Publishers, Dordrecht, The Netherlands.
Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) The Protein Data Bank: A Computer-Based Archival File for Macromolecular Structures, *J. Mol. Biol. 112*, 535–542.
Blake, C. C. F., Johnson, L. N., Mair, G. A., North, A. C. T., Phillips, D. C., & Sarma, V. R. (1967) *Proc. R. Soc. London, Ser. B 167*, 378–388.
Browne, D. T., & Otvos, J. D. (1976) *Biochem. Biophys. Res. Commun. 68*, 907–913.
Campbell, I. D., Dobson, C. M., Williams, R. J. P. & Xavier, A. V. (1973) *J. Magn. Reson. 11*, 172–181.
Canfield, R. (1963) *J. Biol. Chem. 238*, 2698–2707.
Chaiken, I. M., Freedman, M. H., Lyerla, J. R., Jr., & Cohen, J. S. (1973) *J. Biol. Chem. 248*, 884–891.
de Dios, A. C., Pearson, J. G., & Oldfield, E. (1993) *Science 260*, 1491–1496.
Dobson, C. M., & Williams, R. J. P. (1975) *FEBS Lett. 56*, 362–365.
Falke, J. J., Luck, L. A., & Scherrer, J. (1992) *Biophys. J. 62*, 82–86.
Gamcsik, M. P., & Gerig, J. T. (1986) *FEBS Lett. 196*, 71–74.
Gamcsik, M. P., Gerig, J. T., & Swenson, R. B. (1986) *Biochim. Biophys. Acta 874*, 372–377.
Gamcsik, M. P., Gerig, J. T., & Gregory, D. H. (1987) *Biochim. Biophys. Acta 912*, 303–316.
Gerig, J. T. (1989) *Methods Enzymol. 177*, 3–23.
Gerig, J. T., Klinkenborg, J. C., & Nieman, R. A. (1983) *Biochemistry 22*, 2076–2087.
Hayashi, K., Imoto, T., Funatsu, G., & Gunatsu, M. (1965) *J. Biochem. (Tokyo) 58*, 227–235.
Hayashi, K., Shimoda, T., Yamada, K., Kumai, H., & Funatsu, M. (1968a) *J. Biochem. (Tokyo) 64*, 239–245.
Hayashi, K., Shimoda, T., Imoto, T., & Funatsu, M. (1968b) *J. Biochem. (Tokyo) 64*, 365–370.
Hermann, J., & Jolles, J. (1970) *Biochim. Biophys. Acta 200*, 178–179.
Hinton, J. F., Guthrie, P., Pulay, P., & Wolinski, K. (1992) *J. Am. Chem. Soc. 114*, 1604–1605.
Ho, C., Pratt, E. A., & Rule, G. S. (1989) *Biochim. Biophys. Acta 988*, 173–184.
Hull, W. E., & Sykes, B. D. (1975) *J. Mol. Biol. 98*, 121–153.
Hull, W. E., & Sykes, B. D. (1976) *Biochemistry 15*, 1535–1546.
Ibrahim, I. M., Prager, E. M., White, T. J., & Wilson, A. C. (1979) *Biochemistry 18*, 2736–2744.
Imoto, T., & Rupley, J. A. (1973) *J. Mol. Biol. 80*, 657–667.
Imoto, T., Johnson, L. N., North, A. C. T., Phillips, D. C., & Rupley, J. A. (1972) in *The Enzymes*, (Boyer, P. D., Ed.) Vol. 7, pp 666–868, Academic Press, New York.
Kimber, B. J., Griffiths, D. V., Birdsall, B., King, R. W., Scudder, P., Feeney, J., Roberts, G. C. K., & Burgen, A. S. V. (1977) *Biochemistry 16*, 3492–3500.
Kimber, B. J., Feeney, J., Roberts, G. C. K., Birdsall, B., Griffiths, D. V., & Burgen, A. S. V. (1978) *Nature 271*, 184–185.
Laemmli, U. K. (1970) *Nature (London) 227*, 680–685.
Luck, L. A., & Falke, J. J. (1991a) *Biochemistry 30*, 4248–4256.
Luck, L. A., & Falke, J. J. (1991b) *Biochemistry 30*, 4257–4261.
Luck, L. A., & Falke, J. J. (1991c) *Biochemistry 30*, 6484–6490.
Malcolm, B. A., Rosenberg, S., Corey, M. J., Allen, J. S., de Baetselier, A., & Kirsch, J. F. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 133–137.
Norton, R. S., Allerhand, A. (1976) *J. Biol. Chem. 251*, 6522–6528.
Pearson, J. G., Oldfield, E., Lee, F. S., & Warshel, A. (1993) *J. Am. Chem. Soc. 115*, 6851–6862.
Peersen, O. B., Pratt, E. A., Truong, H.-T. N., Ho, C., & Rule, G. S. (1990) *Biochemistry 29*, 3256–3262.
Phillips, L., Separovic, F., Cornell, B. A., Barden, J. A., & dos Remedios, C. G. (1991) *Eur. Biophys. J. 19*, 147–155.
Prager, E. M., Arnheim, N., Mross, G. A., & Wilson, A. C. (1972) *J. Biol. Chem. 247*, 2905–2916.
Robertson, D. E., Kroon, P. A., & Ho, C. (1977) *Biochemistry 16*, 1443–1451.
Sheriff, S., Silverton, E. W., Padlan, E. A., Cohen, G. H., Smith-Gill, S. J., Finzel, B. C., & Davies, D. R. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8075–8079.
Sherman, F., Fink, G. R., & Hicks, J. B. (1986) in *Methods in Yeast Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Sykes, B. D., & Parravano, C. (1969) *J. Biol. Chem. 244*, 3900–3904.
Taylor, H. C., Richardson, D. C., Richardson, J. S., Wlodawer, A., Komoriya, A., & Chaiken, I. M. (1981) *J. Mol. Biol. 149*, 313–317.
Thomas, E. W. (1967) *Biochem. Biophys. Res. Commun. 29*, 628–634.
Vaughan, M., & Steinberg, D. (1960) *Biochim. Biophys. Acta 40*, 230–236.
Wien, R. W., Morrisett, J. D., & McConnell, H. M. (1972) *Biochemistry 11*, 3707–3716.
Wüthrich, K. (1986) in *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.
Yamasaki, N., Hayashi, K., & Funatsu, M. (1968a) *Agric. Biol. Chem. 32*, 55–63.
Yamasaki, N., Hayashi, K., & Funatsu, M. (1968b) *Agric. Biol. Chem. 32*, 64–68.