Serpersu, E. H., McCracken, J., Peisach, J., & Mildvan, A. (1988) *Biochemistry 27*, 8034-8044.
Shimizu, T., Mims, W. B., Peisach, J., & Davis, J. L. (1979) *J. Chem. Phys. 70*, 2249-2254.
Stezowski, J. C., & Hoard, J. L. (1984) *Isr. J. Chem. 24*, 323-334.
Strosberg, D. A., Buffard, D., Lauwereys, M., & Foriers, A. (1986) in *The Lectins: Properties, Functions, and Applications in Biology and Medicine* (Liener, I. E., Sharon, N., & Goldstein, I. J., Eds.) pp 249-264, Academic Press, Orlando.
Tichy, M., Ticha, M., & Kocourek, J. (1971) *Biochim. Biophys. Acta 229*, 63-67.
Tipton, P., & Peisach, J. *Biochemistry* (in press).
Trowbidge, I. S. (1974) *J. Biol. Chem. 249*, 6004-6012.
VanCamp, H. L., Sands, R. H., & Fee, J. A. (1981) *J. Chem. Phys. 75*, 2098-2107.
Yariv, J., Kalb, A. J., & Levitzki, A. (1968) *Biochim. Biophys. Acta 165*, 303-305.
Yudanov, Y. F., Salikhov, K. M., & Zhidomirov, G. M., Tsvetkov, Y. D. (1969) *Zh. Strukt. Khim. 10*, 732-734.

Uniform ^{13}C Isotope Labeling of Proteins with Sodium Acetate for NMR Studies: Application to Human Carbonic Anhydrase II†

Ronald A. Venters,‡ Tiffany L. Calderone,‡ Leonard D. Spicer,*,‡,§ and Carol A. Fierke*,‡
Department of Biochemistry and Department of Radiology, Duke University Medical Center, Durham, North Carolina 27710
Received December 4, 1990; Revised Manuscript Received January 8, 1991

ABSTRACT: Uniform double labeling of proteins for NMR studies can be prohibitively expensive, even with an efficient expression and purification scheme, due largely to the high cost of [$^{13}C_6$, 99%]glucose. We demonstrate here that uniformly (>95%) ^{13}C and ^{15}N double-labeled proteins can be prepared for NMR structure/function studies by growing cells in defined media containing sodium [1,2-$^{13}C_2$, 99%]acetate as the sole carbon source and [^{15}N, 99%]ammonium chloride as the sole nitrogen source. In addition, we demonstrate that this labeling scheme can be extended to include uniform carbon isotope labeling to any desired level (below 50%) by utilizing media containing equal amounts of sodium [1-^{13}C, 99%]acetate and sodium [2-^{13}C, 99%]acetate in conjunction with unlabeled sodium acetate. This technique is less labor intensive and more straightforward than labeling using isotope-enriched algal hydrolysates. These labeling schemes have been used to successfully prepare NMR quantities of isotopically enriched human carbonic anhydrase II. The activity and the ^1H NMR spectra of the protein labeled by this technique are the same as those obtained from the protein produced from media containing labeled glucose; however, the cost of the sodium [1,2-$^{13}C_2$, 99%]acetate growth media is considerably less than the cost of the [$^{13}C_6$, 99%]glucose growth media. We report here the first published ^{13}C and ^{15}N NMR spectra of human carbonic anhydrase II as an important step leading to the assignment of this 29-kDa zinc metalloenzyme.

The size of proteins whose backbone assignment and subsequent structure determination can be effectively elucidated by high-field NMR spectroscopy has increased rapidly as advances in pulse sequences, probe design, and instrumentation have been made. One major contributing factor to this advance has been the ability to utilize ^{13}C and ^{15}N isotopically labeled proteins in both backbone and side-chain assignment strategies (Marion et al., 1989; Wang et al., 1990; Westler et al., 1988; Bax et al., 1990; Fesik et al., 1990; Kay et al., 1990; Ikura et al., 1990).

The traditional first step in a protein NMR solution structure study is to assign the backbone protons as completely as possible. For modestly sized proteins, this is usually accomplished by using homonuclear ^1H 2-D methodologies including through-bond or COSY-type experiments in conjunction with through-space or NOESY experiments. The data sets generated are subsequently analyzed and assigned on the basis of published strategies (Wüthrich, 1986; Englander & Wand, 1987). These strategies break down, however, as the size of the protein exceeds 10 kDa due to spectral overlap, a decrease in sensitivity resulting from line broadening, decreases in T_2 relaxation times, and a marked increase in assignment ambiguities. Most of these problems can be overcome by recently developed heteronuclear experiments, including heteronuclear multiple quantum coherence experiments (HMQC, HMQC-COSY, NOESY-HMQC, and TOCSY-HMQC) (Marion et al., 1989; Wang et al., 1990; Westler et al., 1988), ^{13}C-^{13}C magnetization-transfer experiments (HC-CH) (Bax et al., 1990; Fesik et al., 1990; Kay et al., 1990), and the 3-D experiments [HNCO, HNCA, HCACO, HCA(CO)N] introduced by Ikura et al. (1990). All of these experiments derive sequential connectivities through one-bond J-couplings and are, therefore, independent of the local geometry of the molecule. In addition, assignments using these experimental methods are less ambiguous because they do not depend on NOE information and because they offer several independent pathways in determining sequential connectivities. Moreover, since only large couplings (H-C, ~125-160 Hz; C-C, ~33-45 Hz; C-N, ~7-15 Hz; and H-N, ~90 Hz) are

†Supported by grants to C.A.F. from the National Institute of Health (R29GM40602) and from the American Cancer Society (JFRA246) and to L.D.S. from the National Institute of Health (R01GM41829) and Glaxo Pharmaceuticals, Inc.
*Corresponding authors.
‡Department of Biochemistry.
§Department of Radiology.

0006-2960/91/0430-4491$02.50/0

used, these methods can be utilized on larger molecules with shortened relaxation times.

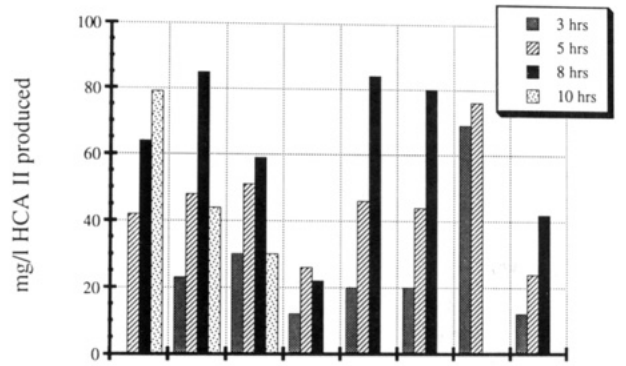
In order for these experiments to achieve common use, investigators must be able to isotopically label a protein, both uniformly and type specifically, in an efficient and cost-effective manner and to purify it in milligram quantities. Efficient type-specific labeling methods have recently been discussed (Hoffman & Spicer, 1991), and we focus here on a new approach to uniform ^{13}C labeling and ^{13}C and ^{15}N double labeling.

If the protein under consideration for NMR study can be cloned and expressed in *Escherichia coli*, uniform labeling with ^{15}N is relatively straightforward and inexpensive by using defined media containing [^{15}N, 99%]ammonium chloride as the nitrogen source. Uniform (>95%) labeling with ^{13}C is also relatively straightforward but can be prohibitively expensive if defined media containing [$^{13}C_6$, 99%]glucose are used. Unfortunately, in addition to its high cost, the availability of [$^{13}C_6$, 99%]glucose depends on the demand cycle. We describe here a technique using defined media containing sodium [1,2-$^{13}C_2$, 99%]acetate as the sole carbon source to uniformly (>95%) ^{13}C label proteins that is just as simple as labeling with glucose. Labeled sodium acetate is three to four times lower in price than labeled glucose and is readily available. In addition, by using equal amounts of sodium [1-^{13}C, 99%]acetate and sodium [2-^{13}C, 99%]acetate in combination with unlabeled sodium acetate in growth media, it is possible to uniformly label carbon to any desired level below 50%.

This method was developed in our laboratories for NMR structure/function studies of human carbonic anhydrase II (HCAII), a 29-kDa monomeric zinc metalloenzyme. Due to the relatively large size of the protein, these studies depend on selective and uniform labeling with ^{15}N and ^{13}C. Optimizing the growth conditions for carbon labeling by utilizing either sodium acetate or glucose provided yields of 70–80 mg of HCAII per liter of media. Since optimum labeling with glucose requires 2 g/L, whereas labeling with sodium acetate requires 3 g/L, the actual cost benefit for labeling with sodium [1,2-$^{13}C_2$, 99%]acetate is approximately 2-fold. The 1-D and 2-D NMR spectra obtained from the protein purified from sodium acetate growths are identical to those obtained from the protein purified from glucose growths. In both cases, the quality and the resolution in the ^1H, ^{13}C, and ^{15}N spectra appear to be adequate for further 2-D and 3-D experimentation and assignment purposes. We also show the more general applicability of the approach by carbon labeling the methionine repressor protein MetJ.

Materials and Methods

High-level expression of either HCAII (Fierke et al., 1991) or the methionine repressor protein MetJ (Tom Kirby, personal communication) in *E. coli* has been achieved by the construction of vectors containing the protein gene subcloned behind a phage T7 RNA polymerase promoter vector (Rosenberg et al., 1987). Protein production is then achieved by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG), which induces a chromosomal copy of T7 RNA polymerase (behind a *lac* UV promoter) in the cell line BL21(DE3), which in turn starts transcription of the protein gene (Studier & Moffatt, 1986). With this promoter system, 10–15% of the total cellular protein is HCAII or MetJ after induction. Purification procedures for HCAII routinely used are a slight modification of those published (Khalifah et al., 1977) and rely primarily on a sulfonamide affinity column with azide elution to selectively isolate HCAII from the remainder of the cellular proteins. HCAII activity was measured by assaying



Temp., °C	30	34	37	37	34	34	37	34
[Ac], g/l	3	3	3	2	--	--	--	--
[13-C-Ac], g/l	--	--	--	--	3	4	--	--
[Glu], g/l	--	--	--	--	--	--	2	2

FIGURE 1: Graph of HCAII expression as a function of both time after induction with IPTG (3, 5, 8, and 10 h) and growth conditions, including temperature (columns 1, 2, and 3), acetate concentration (columns 3, 4, 5, and 6), and glucose concentration (columns 7 and 8). The cells were lysed by the addition of lysozyme and Triton X-100 (Cull & McHenry, 1990). The concentration of HCAII was determined from the Diamox-inhibitable *p*-nitrophenyl acetate esterase activity in the supernatant, with $k_{cat}/K_M = 2500\ M^{-1}\ s^{-1}$ at pH 8.0. These growth conditions were repeated several times, and a 10–15% variability in HCAII expression was observed.

enzyme-catalyzed hydrolysis of *p*-nitrophenyl acetate at 348 nm (Verpoorte et al., 1967), while the cellular concentration of the MetJ protein was estimated from Coomassie blue stained SDS–polyacrylamide gels.

Direct comparison between sodium acetate and glucose as the sole carbon sources in cell growth and protein production was carried out at a variety of conditions in order to optimize both systems (Figure 1). The growth media consisted of M9 salts (Maniatis et al., 1986) to which 1 μM $FeCl_3$, 25 μM $ZnSO_4$, 5 mg/L thiamine, 10 mL/L vitamin mixture (containing 10 mg/100 mL each of biotin, choline chloride, folic acid, niacinamide, D-pantothenate, and pyridoxal and 1 mg/100 mL riboflavin), and 50 μg/mL ampicillin are added. For the MetJ protein, the media contained 20 μM or 1 mM methionine instead of zinc. Cell growth was initiated by inoculating M9 salts/glucose media with a single colony from a freshly streaked plate, growing at 37 °C until mid-log phase (A_{600} = 0.3–0.5), and then diluting the mixture 1:50 into prewarmed media. This procedure is crucial for obtaining efficient cell growth and protein production. The doubling time for BL21(DE3) in minimal/acetate media is 3–4 h. Conditions varied included temperature, pH, acetate concentration, glucose concentration, rifampycin addition, and induction time.

Sodium [1,2-$^{13}C_2$, 99%]acetate was obtained from three sources (Cambridge Isotope Laboratories, Isotec Inc., and MSD Isotopes). Stock solutions were prepared and immediately adjusted to pH 9. Labeling efficiency of material from these sources was compared. Sodium [1-^{13}C, 99%]acetate, sodium [2-^{13}C, 99%]acetate, and [^{15}N, 99%]ammonium chloride were obtained from Cambridge Isotope Laboratories. All NMR experiments were performed on a GN500 General Electric 500-MHz spectrometer equipped with an inverse-detection probe.

Results and Discussion

Figure 1 shows a comparison of the time course of HCAII production in minimal/glucose and minimal/acetate media at a variety of conditions. Surprisingly, under optimal conditions, the HCAII production in minimal/acetate (3 g/L) media is equal to or higher than that in minimal/glucose

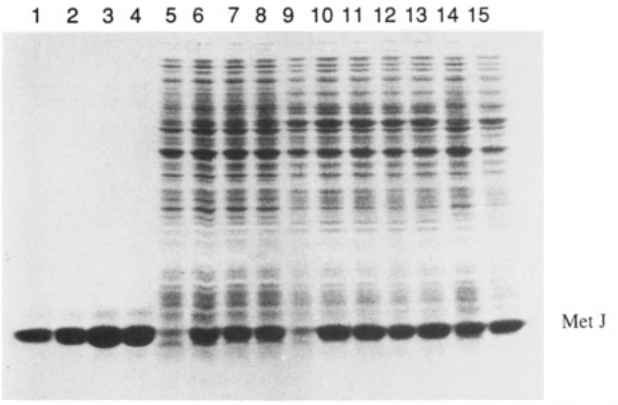


FIGURE 2: Coomassie blue stained SDS–polyacrylamide gel of *E. coli* extracts containing the MetJ repressor protein. A total of 250 μL of cell culture was pelleted, resuspended in 50 μL of gel loading buffer (0.2% sodium dodecyl sulfate, 0.2% 2-mercaptoethanol, 5% glycerol, 0.05 M Tris, pH 6.8), lysed by boiling, and run on a 12% SDS–polyacrylamide gel (Laemmli, 1970), stained with Coomassie blue. The samples were loaded as follows: lanes 1–4 are 10 μg, 15 μg, 20 μg, and 25 μg of the purified MetJ protein. In the remaining lanes the cells were grown under the following conditions: lanes 5–8, 2 g/L glucose at 34 °C incubated for 0, 3, 5, and 7 h after induction, respectively; lanes 9–12, 3 g/L acetate at 34 °C incubated for 0, 3, 5, and 7 h after induction, respectively; and lanes 13–15, 4 g/L acetate at 34 °C, 2 g/L glucose at 37 °C, and 3 g/L acetate at 37 °C, respectively, incubated for 5 h after induction.

media. The sodium acetate concentration in the media was varied in order to determine the minimum amount required for reasonable HCAII yields; protein production was minimal at 2 g/L acetate but increased more than 2-fold at a level of 3 g/L (Figure 1) with no further improvement observed at higher concentrations. In contrast, maximum expression with glucose is attained at 2 g/L, with overall expression similar to that obtained at optimal sodium acetate levels. These data indicate that, even after factoring in the need for increased sodium [1,2-$^{13}C_2$, 99%]acetate concentrations to obtain similar expression yields, uniform labeling of carbon (>95%) with acetate is 2 times less expensive than labeling with glucose.

Temperature variations caused perhaps the largest perturbations in cell growth and subsequent protein production. As can be seen in Figure 1, maximum protein production in acetate increases by about one-third as the induction temperature is lowered from 37 to 34 °C. Further reduction in the temperature to 30 °C provides the same overall protein yield as that obtained at 34 °C; however, induction times are longer. It should be noted here that the optimal induction time for HCAII production increases slightly when sodium acetate is present as the only carbon source compared to when glucose is present.

Other parameters, including the pH of the media, the cell density at induction, the addition of rifampycin (50 μg/mL 2 h after induction), and the concentrations of Zn^{2+}, IPTG, and vitamins, were also varied as part of the overall optimization process. During induction, the pH of the acetate-based media increased; maintenance of the pH at 7.0 by addition of monobasic sodium phosphate caused a decrease in HCAII yield (up to 5-fold). The cells were induced in late log phase (A_{600} = 0.7) for most cases, although no difference in overall production was observed if IPTG was added at an A_{600} between 0.6 and 0.9. Varying the concentrations of Zn^{2+} (0.5–2 mM), IPTG (0.1–0.5 mM), and vitamins (0–1×) or the addition of rifampycin had a minimal effect on overall HCAII yields, although cell growth without vitamins was slowed slightly.

Production of the MetJ protein in *E. coli* grown on minimal/acetate media was then pursued by using the optimized

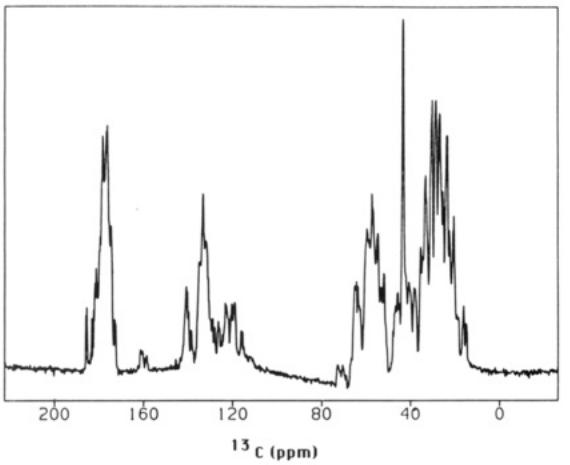


FIGURE 3: 500-MHz ^{13}C NMR spectrum of uniformly (>95%) ^{13}C- and ^{15}N-enriched human carbonic anhydrase II. The sample contains 0.76 mM protein in a >99% D_2O solvent at pH 6.5 in 100 mM phosphate buffer. A total of 10 000 scans of 16K data points were collected at 38 °C.

conditions for the HCAII protein (34 °C and 3 g/L acetate) with the exception that 1 mM methionine was substituted for 25 μM zinc. As shown in Figure 2, we obtained similar (or slightly increased) production of MetJ in minimal/acetate media compared to minimal/glucose media, which is estimated from standards to be about 50–75 mg of MeJ/L of media. In this case, optimal production occurred at 3–5 h after induction. These data suggest that production of proteins in minimal/acetate media with the T7 RNA polymerase vector system in BL21(DE3) cells can be used as a general technique.

With the procedures outlined above, two 1 mM uniformly (>95%) ^{13}C and ^{15}N double-labeled human carbonic anhydrase II samples were prepared for high-resolution NMR studies from 1 L of defined media containing 1 g/L [^{15}N, 99%]ammonium chloride and 3 g/L sodium [1,2-$^{13}C_2$, 99%]acetate. The samples contained 100 mM phosphate buffer at pH 6.5 in a >99% D_2O solvent. Figure 3 presents the 500-MHz proton-decoupled 1-D ^{13}C spectrum of this protein, which clearly shows uniform labeling of all carbons throughout the protein. This spectrum contains resonances from carbonyl, aromatic, α, and aliphatic carbons at intensities consistent with complete uniform labeling. The 500-MHz proton-decoupled nitrogen spectrum (not shown) contains resonances from amide nitrogens in addition to side-chain nitrogen resonances from histidines, arginines, and lysines. The 1-D proton spectrum of this protein is identical with the spectrum of unlabeled HCAII if ^{13}C is decoupled during acquisition (with the exception of splittings due to ^{1}H–^{15}N coupling), indicating that the growth of the protein on sodium [1,2-$^{13}C_2$, 99%]acetate does not significantly alter the protein conformation versus the conformation of HCAII grown on [$^{13}C_6$, 99%]glucose.

Figure 4 presents the backbone amide region of the proton detected ^{1}H–^{15}N HMQC spectrum of ^{13}C- and ^{15}N-labeled HCAII and clearly indicates uniform ^{15}N labeling and no apparent differences in the labeling distribution when sodium acetate is used in the growth medium. Since, at present, we do not have three-channel capability, the ^{13}C interactions have not been decoupled. Due to the ^{13}C interactions present, the cross-peak resonances in this coupled spectrum are broad (average of 62.5 Hz) compared with those of the HMQC spectrum obtained on a ^{15}N, ^{12}C sample.

We have also prepared a 4 mM sample of HCAII containing 26% ^{13}C and 95% ^{15}N uniform labeling by growing cells in defined media containing 0.78 g/L sodium [1-^{13}C,

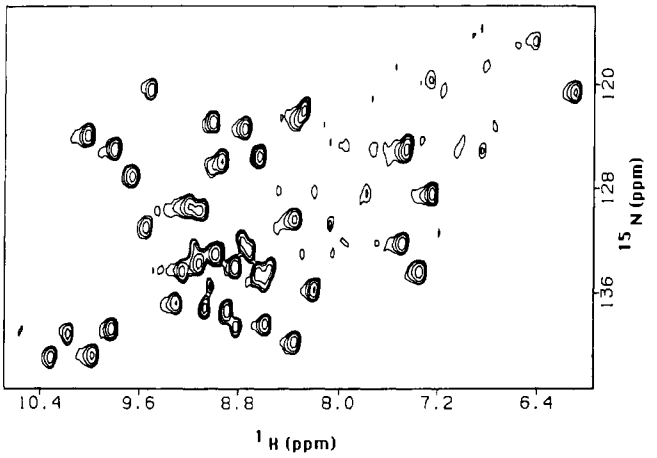


FIGURE 4: The amide region of a 500-MHz proton detected ^{1}H–^{15}N HMQC spectrum of uniformly (>95%) ^{13}C- and ^{15}N-enriched human carbonic anhydrase II. The sample contains 0.76 mM protein in a >99% D_2O solvent at pH 6.5 in 100 mM phosphate buffer. The spectrum was recorded with 512 blocks and 128 scans of 2K data points per block at 38 °C. ^{15}N frequencies were decoupled during acquisitions but ^{13}C frequencies were not.

99%]acetate, 0.78 g/L sodium [2-^{13}C, 99%]acetate, 1.44 g/L unlabeled sodium acetate, and 1 g/L [^{15}N, 99%]ammonium chloride. The ^{1}H-decoupled ^{13}C spectrum obtained from this sample is identical with the spectrum obtained from the >95% labeled sample shown in Figure 3. This technique for generating proteins with lower levels of ^{13}C isotope enrichment is more straightforward and less labor intensive than labeling with enriched algal hydrolysates, as is currently employed. Although the use of sodium [1-^{13}C, 99%]acetate and sodium [2-^{13}C, 99%]acetate in equal molar amounts leads to uniform labeling, the use of either one of these labeled acetates independently of the other in growth media leads to nonuniform carbon labeling. We prepared growth media containing 0.22 g/L sodium [1-^{13}C, 99%]acetate and 2.78 g/L unlabeled sodium acetate as the only carbon sources. The ^{13}C spectrum of the protein purified from this growth media contains resonances from all carbon types; however, the carbonyls are labeled preferentially. In contrast, the protein isolated from growth media containing 0.22 g/L sodium [2-^{13}C, 99%]acetate and 2.78 g/L unlabeled sodium acetate exhibits preferential ^{13}C labeling at the aromatic and aliphatic carbons.

These data indicate an alternative, relatively low cost, and time-efficient procedure for labeling proteins expressed in *E. coli* to any desired level with ^{13}C and for double labeling with ^{13}C and ^{15}N. The ability to label proteins cost effectively for NMR studies allows researchers to take advantage of the elegant heteronuclear experiments recently developed. The utilization of these advances often simplifies small protein NMR structure determination and substantially extends the upper size limit of proteins whose full assignment and subsequent structure determination can be elucidated by NMR. This, in turn, greatly enhances the value of NMR spectroscopic techniques for macromolecular structure/function studies in biochemical research.

When sodium [1,2-$^{13}C_2$, 99%]acetate is used in place of [$^{13}C_6$, 99%]glucose as the sole carbon source during growth, uniform ^{13}C labeling of proteins can be realized at lower isotope cost. The use of sodium [1-^{13}C, 99%]acetate and sodium [2-^{13}C, 99%]acetate in combination with unlabeled sodium acetate in the growth media allows for uniform carbon labeling to any desired level. The protein produced is identical with that obtained from glucose without a reduction in overall yield and without a substantial increase in the induction time required. It should be noted that the conditions given here have been optimized for a specific protein utilizing a specific expression system. We have demonstrated that the general procedure is transferable to other systems by expressing the methionine repressor protein MetJ according to the optimum procedure developed here for HCAII. Optimal conditions generally must be determined independently, however, for MetJ and for other systems of interest.

ACKNOWLEDGMENTS

The Duke University NMR Center was established with grants from the NIH, NSF, and the North Carolina Biotechnology Center, which are gratefully acknowledged. We thank Dr. Tom Kirby for his generous gift of the *metJ*-expression plasmid. We greatly appreciate the assistance of Mr. David Hyre and Dr. Ronald C. Greene.

REFERENCES

Bax, A., Clore, M., Driscoll, P. C., Gronenborn, A. M., Ikura, M., & Kay, L. E. (1990) *J. Magn. Reson. 87*, 620–627.

Cull, M., & McHenry, C. S. (1990) *Methods Enzymol. 182*, 147–153.

Englander, S. W., & Wand, A. J. (1987) *Biochemistry 26*, 5953–5958.

Fesik, S. W., Eaton, H. L., Olejniczak, E. T., Zuiderweg, E. R. P., McIntosh, L. P., & Dahlquist, F. W. (1990) *J. Am. Chem. Soc. 112*, 886–888.

Fierke, C. A., Krebs, J. F., & Venters, R. A. (1991) *International Workshop on Carbonic Anhydrase: From Biochemistry and Genetics to Physiology and Clinical Medicine*, Verlag-Chemie, Heidelberg.

Hoffman, D. W., & Spicer, L. D. (1991) *Techniques in Protein Chemistry II* (Villafranca, J. J., Ed.) pp 409–416, Academic Press, New York.

Ikura, M., Kay, L. E., & Bax, A. (1990) *Biochemistry 29*, 4659–4667.

Kay, L. E., Ikura, M., & Bax, A. (1990) *J. Am. Chem. Soc. 112*, 888–889.

Khalifah, R. G., Strader, D. J., Bryant, S. H., & Gibson, S. M. (1977) *Biochemistry 16*, 2241–2247.

Laemmli, U. K. (1970) *Nature 227*, 680–685.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1986) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., & Clore, G. M. (1989) *Biochemistry 28*, 6150–6156.

Rosenberg, A. H., Lade, B. N., Chui, D. S., Lin, S. W., Dunn, J. J., & Studier, F. W. (1987) *Gene 56*, 125–135.

Studier, F. W., & Moffatt, B. A. (1986) *J. Mol. Biol. 189*, 113–130.

Verpoorte, J. A., Mehta, S., & Edsall, J. J. (1967) *J. Biol. Chem. 242*, 4221–4229.

Wang, J., Hinck, A. P., Loh, S. N., & Markley, J. L. (1990) *Biochemistry 29*, 102–113.

Westler, W. M., Stockman, B. J., & Markley, J. L. (1988) *J. Am. Chem. Soc. 110*, 6256–6258.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*. Wiley, New York.