Journal of Biomolecular NMR, **21:** 167–171, 2001.
KLUWER/ESCOM

Optimized labeling of $^{13}CHD_2$ methyl isotopomers in perdeuterated proteins: Potential advantages for ^{13}C relaxation studies of methyl dynamics of larger proteins

Rieko Ishima[a], John M. Louis[b] & Dennis A. Torchia[a,*]
[a]*Molecular Structural Biology Unit, National Institute of Dental and Craniofacial Research and* [b]*Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, U.S.A.*

Received 29 June 2001; Accepted 7 August 2001

Key words: conformational change, internal motion, isotope labeling, model-free, NMR

Abstract

$^{13}CHD_2$ methyl isotopomers are particularly useful to study methyl dynamics in proteins because, as compared with other methyl isotopomers, the ^{13}C relaxation mechanism for this isotopomer is straightforward. However, in the case of proteins, where $(\omega\tau)^2 \gg 1$, the refocused INEPT pulse sequence does not completely suppress unwanted $^{13}CH_3$ signals. The presence of weak $^{13}CH_3$ peaks is usually not a serious problem for smaller proteins because there are relatively few methyl signals and they are sharp; however, signal overlap becomes more common as the size of the protein increases. We overcome this problem by preparing a protein using a 98% D_2O cell culture medium containing 3-^{13}C pyruvic acid, 50–60% deuterated at the 3-position, and 4-^{13}C 2-ketobutyric acid, 98% and 62% deuterated at the 3- and 4-positions, respectively. This approach significantly reduces the population of the CH_3 isotopomer while optimizing the production of $^{13}CHD_2$, the isotopomer desired for ^{13}C relaxation measurements. In larger proteins where the deuterium T_2 may be too short to measure accurately, we also suggest the alternative measurement of the proton T_2 of the $^{13}CH_2D$ methyl isotopomer, because these protons are well-isolated from other protons in these highly deuterated samples.

Communication

There is great interest in elucidating the dynamics of methyl groups in the large hydrophobic aliphatic sidechains of Ile, Leu and Val residues, because these sidechains often contribute to hydrophobic cores of proteins and are involved in hydrophobic interactions with target molecules. Uniform ^{13}C labeling in conjunction with ^{1}H-^{13}C heteronuclear pulse methodology has greatly facilitated detection of methyl signals in multidimensional NMR spectra. However, interpretation of ^{13}C methyl relaxation data of a uniformly ^{13}C labeled and protonated protein is complicated by ^{13}C-^{13}C dipolar and J couplings, and by multi-exponential decay caused by ^{13}C-^{1}H dipolar cross correlation (Kay et al., 1992). To overcome these problems, labeling approaches that produce proteins containing high populations of ^{12}C-$^{13}CHD_2$ methyl sites have been developed. This was first done in the case of thioredoxin, by growing a modified *E. coli* strain, DL323, in a 50% D_2O medium containing appropriately ^{13}C labeled glycerol and sodium carbonate (LeMaster, 1999; LeMaster and Kushlan, 1996). More recently, we expressed the HIV-1 protease in *E. coli* using a 98% D_2O culture media containing 3-^{13}C pyruvate (Ishima et al., 1999). In this sample, methyl groups were partially protonated because the pyruvate was protonated, while all other hydrogen sites in the protein were highly deuterated (Rosen et al., 1996). In addition, the Ala, Leu, Val and Ile ($C^{\gamma 2}$ only) methyl carbons were labeled with ^{13}C (Lee et al., 1997). The HIV-1 protease labeled by these methods contained high populations

*To whom correspondence should be addressed. E-mail: dtorchia@dir.nidcr.nih.gov

of both ^{12}C-$^{13}CHD_2$ and ^{12}C-$^{13}CH_2D$ methyl groups and was therefore well-suited to measure both ^{13}C and ^{2}H relaxation rates (Ishima et al., 2001a). The ^{13}C relaxation rates provided information about methyl motions on both ps-ns and ms-μs time scales, while the deuterium relaxation rates, measured with high resolution and without a constant time period (Ishima et al., 2001a), provided information about ps-ns time scale motions (Muhandiram et al., 1995).

In addition to the methyl isotopomers discussed in the previous paragraph, CD_3 and CH_3 methyl groups are also present in the sample (Rosen et al., 1996). A pulse sequence that exclusively selects CHD_2methyl signals has been discussed, and is particularly useful when constant time carbon evolution is required (Muhandiram et al., 1995). Because C^{methyl}-C J-couplings are absent in our sample, carbon constant time evolution is not required and CHD_2 signals were selected in ^{13}C relaxation experiments using refocused INEPT magnetization transfer (Ishima et al., 1999). During INEPT transfer, CHD_2, CH_2D, and CH_3 methyl ^{13}C signals evolve under J coupling to one, two and three protons respectively. Theory shows that if relaxation effects are negligible, refocused INEPT suppresses all but CHD_2 methyl signals when the INEPT pulse spacing equals $1/4J_{CH}$ (Sorensen and Ernst, 1983). In the case of a typical protein where $(\omega_H\tau_c)^2 \gg 1$, the ^{13}C spins in the CH_3 3/2 and 1/2 spin manifolds relax at significantly different rates, and as a consequence CH_3 signals are incompletely suppressed by the refocused INEPT scheme (Ishima et al., 1999). Since the methyl ^{13}C relaxation experiments did not require the use of a constant time period, the signals of both CHD_2 and CH_3 peaks were sharp and had little overlap. However, in the case of measurements of ^{13}C methyl relaxation rates at fields below 500 MHz or in either large proteins or partially unfolded proteins, where signals are not well dispersed, overlap of CHD_2 and CH_3 methyl signals in different residues will likely be a more serious problem.

We overcame this problem by expressing HIV-1 protease in *E. coli* grown on a 98% D_2O medium containing partially deuterated sodium pyruvate-3-^{13}C to reduce the population of CH_3 methyl isotopomers. Although methyl protons of protonated pyruvate slowly exchange with the D_2O in the culture medium, we have found it difficult to rely upon direct exchange with solvent to control the level of deuteration of the methyl isotopomers in the amino acid sidechains because of the variable time period between initiation of the cell-growth and protein-induction phases of the culture. In addition to adding the partially deuterated pyruvate, we also added partially deuterated sodium 4-^{13}C 2-ketobutyrate to the culture media, in order to obtain ^{13}C labeled Ile$C^{\delta1}$, as described previously (Gardner and Kay, 1997). Unlike pyruvate, the methyl protons of the 2-ketobutyrate do not exchange with the D_2O in the culture medium, and partial deuteration of the 2-ketobutyrate is required to obtain significant populations of Ile $^{13}C^{\delta1}HD_2$ and $^{13}C^{\delta1}H_2D$ isotopomers.

The specific protocol to obtain the labeled protein is as follows. HIV-1 protease containing five mutations (Q7K, L33I, L63I, C67A, C95A) (Louis et al., 1999; Mahalingam et al., 1999) was expressed in *E. coli* BL21 (DE3) in 98% D_2O minimal medium at 37 °C. A 0.5 l culture containing 1 gm of 3-^{13}C,3,3,3-D_3 pyruvate (50-60% 3-D_3; 99% 3-^{13}C; sodium salt; Isotec Inc, Miamisburg, OH, U.S.A.) as the sole carbon source was inoculated (A_{600} 0.075) with an overnight *E. coli* culture that was grown in the same medium. When the culture reached an optical density of $\sim$ 0.6 (roughly 18 h), protein expression was induced by the addition of 2 mM IPTG and was allowed to proceed for 5 h. Approximately 1 h before induction, 1 gm of the 3-^{13}C, 3,3,3-D_3 pyruvate and 100 mg of 4-^{13}C,3,3,4,4,4-D_5 2-ketobutyrate (98% 3-D_2; 62% 4-D_3; 99% 4-^{13}C; sodium salt; Isotec Inc, Miamisburg, OH, U.S.A.) were added to the culture. The protease was purified from inclusion bodies as described previously (Weber et al., 1997). A wet cell pellet of 1.8 gms derived from 0.5 l culture yielded $\sim$ 5 mg of pure protease.

Replacement of protonated pyruvate with partially deuterated pyruvate resulted in a substantial decrease of the CH_3/CHD_2 signal intensity ratio. In the ^{13}C-^{1}H HSQC spectrum of the protease prepared using the protonated pyruvate (Ishima et al., 2001a), CH_3 signal intensities of Leu, Val, Ile $C^{\gamma2}$ were 70%, 70%, 56% of the respective CHD_2 intensities. In contrast, in the ^{13}C-^{1}H HSQC spectrum of the protease prepared using the partially deuterated pyruvate (Figure 1A), CH_3 signal intensities of Leu, Val, Ile $C^{\gamma2}$ were only 12%, 14%, 20% of CHD_2 intensities, respectively. We also measured signal volumes because they are approximately proportional to the number of carbon-attached protons in the simple HSQC spectrum. Assuming that each amino acid side chain was randomly deuterated and using the observed range of CH_3/CHD_2 signal volume ratio of 0.12–0.25, we calculated that the deuterated pyruvate (as supplied, 55% D_3) was effectively further enriched by exchange with

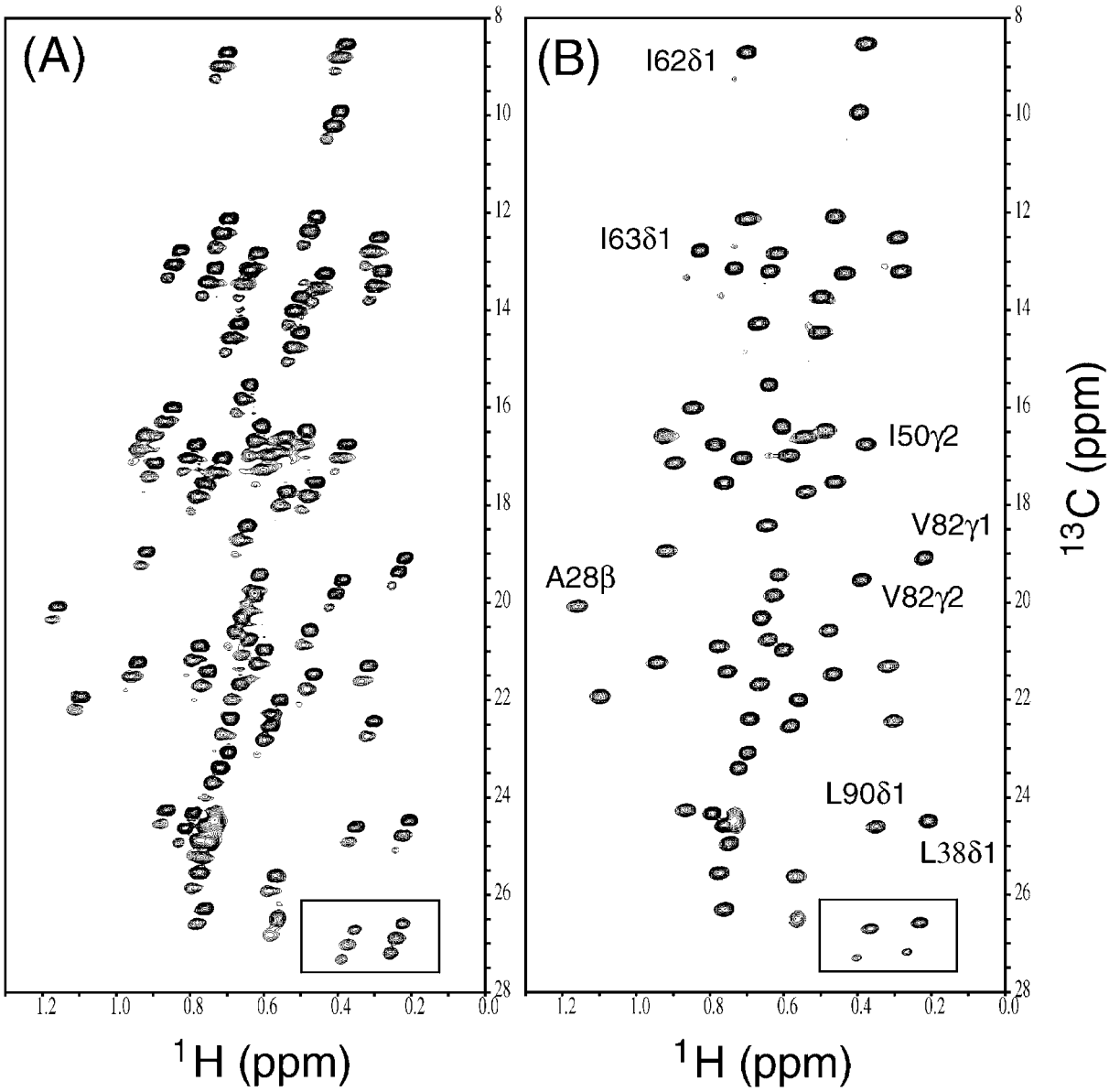

Figure 1. Comparison of methyl regions of ^{13}C-H HSQC spectra of the HIV-1 protease recorded using (A) simple INEPT and (B) refocused INEPT transfer. $^{13}CH_3$, $^{13}CH_2D$, $^{13}CHD_2$ methyl signals of Ala, Ile, Leu and Val residues are observed in (A) whereas only $^{13}CHD_2$ methyl signals of these residues have significant intensity in (B). The protein sample was prepared from *E. coli* grown on a D_2O medium containing 3-^{13}C,3,3,3-D_3 pyruvate and 4-^{13}C,3,3,4,4,4-D_5 2-ketobutyrate as described in the text. For purposes of comparison, the signals of L90 and L38 observed in spectra of a protein sample prepared (Ishima et al., 2001a) by growing *E. coli* on a D_2O medium containing fully protonated, 99% 3-^{13}C pyruvate are enclosed within boxes at the bottom of each panel. All spectra were recorded on a Bruker DMX 500 MHz spectrometer at 20 °C, with an INEPT delay of 1.9ms, using samples of 0.5 mM protease bound to a potent inhibitor, DMP323, prepared as described previously (Ishima et al., 2001a).

solvent to level of ca. 65–75% D_3. When protonated 2-ketobutyrate was included in the 98% D_2O culture medium in order to label Ile $C^{\delta 1}$, only the $C^{\delta 1}H_3$ isotopomer of Ile was observed in the HSQC spectrum (data not shown) of the protease. This observation shows that H-D exchange with solvent is at most a few percent at the methyl position of the 2-ketobutyrate. On the basis of this observation, the relative values of the Ile $C^{\delta 1}H_3$:$C^{\delta 1}H_2D$:$C^{\delta 1}HD_2$ signal volumes, in the HSQC spectrum of the protease expressed using the 62% 4-D_3 2-ketobutyrate (Figure 2A), are predicted to be 0.38:1.23:1. This prediction is in reasonable agreement with the observed average values of 0.32:1.03:1, when one considers that calculation does not account for the fact that transverse relaxation during the INEPT transfer periods attenuates the $C^{\delta 1}H_3$ and $C^{\delta 1}H_2D$ signals more than the $C^{\delta 1}HD_2$ signals.

In the refocused INEPT spectrum, Figure 1B, the ratios of CH_3 to CHD_2 intensities of Leu, Val, Ile $C^{\gamma 2}$, Ile $C^{\delta 1}$ were 3.2%, 2.7%, 3.8%, 7.6% respectively, ca. four times smaller than in the simple INEPT spectrum, Figure 1A. The ratio of Ile $C^{\delta 1}H_3$ to $C^{\delta 1}HD_2$ signal intensity is about twice that of the other methyl groups. A two-fold reduction of Ile $C^{\delta 1}H_3$ to $C^{\delta 1}HD_2$ intensity ratio could be achieved by using ca. 70% deuteration of the methyl position of 2-ketobutyrate. However, the linewidths of the methyl signals are small, so even at their present intensity levels the chemical shifts of the CH_3 signals would have to be within 5–10 Hz of the chemical shifts of the CHD_2 signals in order to significantly affect the measurement of the large CHD_2 intensity. In addition to the methyls of the large hydrophobic amino acids, the methyl group of Ala was also labeled, Figure 1. Except for a somewhat higher level of ^2H enrichment, the pattern of isotope incorporation in the Ala methyl was similar to that observed for Leu, Val and Ile $C^{\gamma 2}$.

A crucial step of our labeling approach is the addition of a half of the pyruvate one hour before induction. When we added all of partially deuterated pyruvate at the beginning of the cell growth, the intensities of the Leu, Val, and Ile$C^{\gamma 2}$ CHD_2 isotopomers were 4–20 times smaller than those of the Ile $C^{\delta 1}$ (data not shown). This observation is an apparent consequence of exchange of the pyruvate methyl protons with the D_2O medium during the 18 h cell growth period (Rosen et al., 1996). Because of exchange with solvent, one could reduce the cost of labeling the protein, by adding protonated pyruvate at the initiation of cell growth (provided that the growth rate is slow, as is typical in the 98% D_2O medium) and the partially deuterated pyruvate 1 h before induction of protein expression.

Selective methyl protonation in perdeuterated proteins has been used to determine the structures of larger proteins by detecting methyl and amide NOEs while suppressing proton spin-flips (Mueller et al., 2000). In structural applications it is essential to label CH_3 isotopomers exclusively in order to maximize sensitivity and minimize signal overlap in NOESY spectra. To produce only CH_3 isotopomers of Leu, Val and Ile, Kay's group used a growth medium containing specifically ^{13}C labeled isovalerate and 2-ketobutyrate (Gardner and Kay, 1997; Goto et al., 1999). In contrast, dynamics studies require high levels of $^{13}CHD_2$ and $^{13}CH_2D$ isotopomers. It would be possible to make $^{13}CHD_2$ and $^{13}CH_2D$ isotopomers by modifying the protocol that produces only $^{13}CH_3$ isotopomers using partially-deuterated ketoisovalerate and 2-ketobutyrate in the growth medium.

The inclusion of pyruvate and 2-ketobutyrate in the growth medium yields Ile side chains labeled at both methyl positions, permitting dynamics of Ile $C^{\gamma 2}$ and $C^{\delta 1}$ sites to be studied in one sample. However, the small $^3J_{C\delta C\gamma}$ coupling (< 3.5 Hz) can affect T_2 measurements carried out using weak effective fields in CPMG experiments. Numerical simulations show an effective field CPMG RF of 500Hz ($\tau_{CPMG} = 0.5$ ms) suppresses the effect of the $^3J_{C\delta C\gamma}$ coupling on the measurement of both the Ile $C^{\gamma 2}$ and $C^{\delta 1}$ T_2, provided that both methyl signals are within 800 Hz of the RF carrier frequency. This condition is satisfied for most Ile residues at 500 MHz (typical range of Ile $C^{\gamma 2}$ and $C^{\delta 1}$ signals in proteins is 9–20 ppm) if the carrier is set to 15 ppm. However, if one uses an effective RF field of 500 Hz ($\tau_{CPMG} = 1.0$ ms) only the T_2's of the Ile $C^{\gamma 2}$ and $C^{\delta 1}$ signals within 400Hz of the carrier (12–18 ppm at 500 MHz) can be reliably measured. The use of a medium containing ketoisovalerate plus 2-ketobutyrate would avoid this problem, but has the disadvantage that dynamics of the Ile$^{\gamma 2}$ and Ala methyls cannot be studied because they are not ^{13}C labeled.

Because our approach maximizes the yield of the CHD_2 isotopomer, the population of CH_2D isotopomer is about 50% of the maximum attainable. This implies that the sample is not optimally labeled for ^2H relaxation measurements. However, because ^2H methyl relaxation experiments have excellent sensitivity (Muhandiram et al., 1995), the sub-optimal CH_2D isotopomer population should not prevent acquisition of high quality ^2H relaxation data, particularly as high

sensitivity cold probes are becoming widely available. Therefore in the case of small to intermediate size proteins labeled using partially deuterated pyruvate, it should be possible to collect accurate ^{13}C and ^{2}H relaxation data on a single labeled sample.

In the case of larger proteins, M.W. > 25 kDa, the small methyl deuteron T_2, < 5 ms, may become a significant impediment to accurate ^{2}H T_2 measurements. Carbon T_2 measurements present no such problems in larger proteins, because the $^{13}CHD_2$ T_2 is ca. 25 times larger than that of the methyl deuteron. Nonetheless, if possible, it is beneficial to augment ^{13}C methyl relaxation data with that of another spin to study sub-ns time scale motion. This is the case because the methyl ^{13}C T_2 is sensitive to chemical exchange (Ishima et al., 1999; LeMaster and Kushlan, 1996), which if not correctly identified, can cause errors in model-free parameters derived from of relaxation data. In contrast, the methyl ^{1}H T_2 is much less sensitive to chemical exchange than the ^{13}C T_2 (Ishima et al., 2001b). Therefore combining $^{13}CHD_2$ ^{13}C T_1 and T_2 measurements with CH_2D deuteron T_1 and proton T_2 measurements, may improve the reliability of model-free parameters of methyl groups in larger proteins. Straightforward interpretation of ^{1}H T_2 measurements will be feasible for a protein having a high level of overall deuteration, because the equivalent protons in the CH_2D isotopomer are well isolated from other protein protons and relax via local dipolar interactions.

Acknowledgements

We thank DuPont Pharmaceuticals for DMP323 and Lewis E. Kay for alerting us to the potential problem posed by ^{13}C ^{3}J coupling in CPMG experiments. This work was supported by the Intramural AIDS Targeted Anti-Viral Program of the Office of the Director of the National Institutes of Health.

References

Gardner, K.H. and Kay, L.E. (1997) *J. Am. Chem. Soc.*, **119**, 7599–7600.

Goto, N.K., Gardner, K.H., Mueller, G.A., Willis, R.C. and Kay, L.E. (1999) *J. Biomol. NMR*, **13**, 369–374.

Ishima, R., Louis, J.M. and Torchia, D.A. (1999) *J. Am. Chem. Soc.*, **121**, 11589–11590.

Ishima, R., Louis, J.M. and Torchia, D.A. (2001a) *J. Mol. Biol.*, **305**, 515–521.

Ishima, R., Petkova, A.P., Louis, J.M. and Torchia, D.A. (2001b) *J. Am. Chem. Soc.*, **123**, 6164–6171.

Kay, L.E., Bull, T.E., Nicholson, L.K., Griesinger, C., Schwalbe, H., Bax, A. and Torchia, D.A. (1992) *J. Magn. Reson.*, **100**, 538–558.

Lee, A.L., Urbauer, J.L. and Wand, A.J. (1997) *J. Biomol. NMR*, **9**, 437–440.

LeMaster, D.M. (1999) *J. Am. Chem. Soc.*, **121**, 1726–1742.

LeMaster, D.M. and Kushlan, D.M. (1996) *J. Am. Chem. Soc.*, **118**, 9255–9264.

Louis, J.M., Clore, G.M. and Gronenborn, A.M. (1999) *Nat. Struct. Biol.*, **6**, 868–875.

Mahalingam, B., Louis, J.M., Reed, C.C., Adomat, J.M., Krouse, J., Wang, Y.F., Harrison, R.W. and Weber, I.T. (1999) *Eur. J. Biochem.*, **263**, 238–245.

Mueller, G.A., Choy, W.Y., Yang, D., Forman-Kay, J.D., Venters, R.A. and Kay, L.E. (2000) *J. Mol. Biol.*, **300**, 197–212.

Muhandiram, D.R., Yamazaki, T., Sykes, B.D. and Kay, L.E. (1995) *J. Am. Chem. Soc.*, **117**, 11536–11544.

Rosen, M.K., Gardner, K.H., Willis, R.C., Parris, W.E., Pawson, T. and Kay, L.E. (1996) *J. Mol. Biol.*, **263**, 627–636.

Sorensen, O.W. and Ernst, R.R. (1983) *J. Magn. Reson.*, **51**, 477–489.

Weber, I.T., Wu, J., Adomat, J., Harrison, R.W., Kimmel, A.R., Wondrak, E.M. and Louis, J.M. (1997) *Eur. J. Biochem.*, **249**, 523–530.