Journal of Biomolecular NMR, 4 (1994) 543–551
ESCOM

J-Bio NMR 189

Estimates of ϕ and ψ torsion angles in proteins from one-, two- and three-bond nuclear spin–spin couplings: Application to staphylococcal nuclease

Arthur S. Edison[a], Frank Weinhold[a,b], William M. Westler[c] and John L. Markley[a,c,*]

[a]*Graduate Biophysics Program,* [b]*Theoretical Chemistry Institute and Department of Chemistry and* [c]*Department of Biochemistry and National Magnetic Resonance Facility at Madison, University of Wisconsin at Madison, Madison, WI 53706, U.S.A.*

Received 12 July 1993
Accepted 14 January 1994

Keywords: Dihedral angles; Calculated coupling constants; Protein structure

SUMMARY

Calculated coupling constants ($^3J_{H^N H^\alpha}$, $^1J_{C^\alpha H^\alpha}$, $^2J_{C'H^\alpha}$, $^1J_{C^\alpha N}$ and $^2J_{C^\alpha N}$) from our accompanying paper [Edison, A.S. et al. (1994) *J. Biomol. NMR*, **4**, 519–542] have been used to generate error surfaces that can provide estimates of the ϕ and ψ angles in proteins. We have used experimental coupling data [$^3J_{H^N H^\alpha}$: Kay, L.E. et al. (1989) *J. Am. Chem. Soc.*, **111**, 5488–5490; $^1J_{C^\alpha H^\alpha}$: Vuister, G.W. et al. (1993) *J. Biomol. NMR*, **3**, 67–80; $^2J_{C'H^\alpha}$: Vuister, G.W. and Bax, A. (1992) *J. Biomol. NMR*, **2**, 401–405; $^1J_{C^\alpha N}$ and $^2J_{C^\alpha N}$: Delaglio, F. et al. (1991) *J. Biomol. NMR*, **1**, 439–446] to create error surfaces for selected residues of the protein staphylococcal nuclease. The residues were chosen to include all those with five experimental couplings, as well as some with four experimental couplings, to demonstrate the relative importance of $^3J_{H^N H^\alpha}$ and $^1J_{C^\alpha H^\alpha}$. For most of the cases, we obtained good agreement between the X-ray structure [Loll, P.J. and Lattman, E.E. (1989) *Protein Struct. Funct. Genet.*, **5**, 183–201] and the NMR data.

INTRODUCTION

The calculation of protein structures from NMR nuclear Overhauser enhancement (NOE) distance constraints is well established (Wagner et al., 1992 and references therein). Angular constraints obtained from coupling constants occasionally supplement the NOE distances, but the angular constraints have been limited primarily to very small (< 4 Hz) or very large (> 9 Hz) values of $^3J_{H^N H^\alpha}$. More recently, stereospecific assignments of prochiral groups and estimates of

*To whom correspondence should be addressed.

Abbreviations: CUPID, ContinUous ProbabIlity Distribution analysis of rotamers; $^nJ_{AB}$, single-bond (n = 1), geminal (n = 2), or vicinal (n = 3) coupling constant between nuclei A and B; NOE, nuclear Overhauser enhancement; r^2, correlation coefficient.

torsion angles have been made on the basis of vicinal couplings across χ^1 (Wagner et al., 1992 and references therein). With the use of isotopic labeling (Markley and Kainosho, 1993) and new multidimensional experimental techniques (Bax and Grzesiek, 1993), many new coupling constants have become experimentally accessible in large (20–30 kDa) proteins (Neuhaus et al., 1984; Montelione et al., 1989; Wider et al., 1989; Chary et al., 1991; Delaglio et al., 1991; Edison et al., 1991; Bax et al., 1992; Blake et al., 1992; Griesinger and Eggenberger, 1992; references contained in Wagner et al., 1992; Vuister et al., 1993).

Our preceding paper (Edison et al., 1994) demonstrated ab initio calculations and experimental correlations of five coupling constants, $^3J_{H^NH^\alpha}$, $^1J_{C^\alpha H^\alpha}$, $^2J_{C'H^\alpha}$, $^1J_{C^\alpha N}$ and $^2J_{C^\alpha N}$, as functions of ϕ and ψ angles in a peptide (see Fig. 1 in Edison et al., 1994), with good correlations for $^3J_{H^NH^\alpha}$, $^1J_{C^\alpha N}$ and $^2J_{C^\alpha N}$ and fair correlations for $^1J_{C^\alpha H^\alpha}$ and $^2J_{C'H^\alpha}$. In this paper we demonstrate an application of these calculations to the estimation of ϕ and ψ angles in the protein staphylococcal nuclease. The methods we use are similar to those developed by other groups for the use of $^3J_{H^NH^\alpha}$ and $^1J_{C^\alpha H^\alpha}$ couplings as restraints in molecular dynamics calculations (Kim and Prestegard, 1990; Mierke et al., 1992; Torda et al., 1993).

This paper does not consider the effects of conformational averaging. However, time dependence could be introduced in a straightforward fashion through restrained molecular dynamics methods (Torda et al., 1993). Recently, our laboratory has developed a method called CUPID which explicitly accounts for conformational averaging across a single dihedral angle by finding Fourier coefficients of the probability distribution from experimental data (Džakula et al., 1992a,b). However, we have not implemented a CUPID-like approach for the 2D problems presented here, because the large numbers of experimental constraints required to obtain 2D Fourier coefficients are lacking.

The plan of this paper is as follows. In the Methods section we describe the procedure used to estimate the ϕ and ψ angles from coupling constant data, measured in the 149-residue protein staphylococcal nuclease ($^3J_{H^NH^\alpha}$: Kay et al., 1989; $^1J_{C^\alpha H^\alpha}$: Vuister et al., 1993; $^2J_{C'H^\alpha}$: Vuister and Bax, 1992; $^1J_{C^\alpha N}$ and $^1J_{C^\alpha N}$: Delaglio et al., 1991). Subsequently, we give the results of our calculations, presented in the form of contour plots, for the 24 residues with the largest number of experimentally measured couplings. In these contour plots, regions in ϕ–ψ space which show the best agreement with the experimental couplings are shown in white. As a comparison, we include the ϕ and ψ angles from the X-ray structure (Loll and Lattman, 1989). We must emphasize that the calculated couplings used to create the contour plots were linearly fit to the same experimental data (along with data from several other proteins) in the previous paper (Edison et al., 1994). To verify the accuracy of our theoretical functions, an independent data set must be examined with the existing parameters. Currently, however, no other protein has such a complete data set. It will be seen that most of the residues presented in this study are constrained to regions that are in good agreement with the X-ray dihedral angles. This demonstrates that the complete (or nearly complete) set of the five couplings considered in this study, $^3J_{H^NH^\alpha}$, $^1J_{C^\alpha H^\alpha}$, $^2J_{C'H^\alpha}$, $^1J_{C^\alpha N}$ and $^2J_{C^\alpha N}$, can provide an important constraint on ϕ and ψ angles in proteins.

Apart from the 24 residues shown in this paper, we have calculated ϕ–ψ plots for 99 additional residues in staphylococcal nuclease, for which there is at least one experimental coupling. The results of the complete set of 123 residues with experimentally measured couplings indicate that fewer than four or five couplings can often constrain ϕ–ψ space and have good agreement with X-ray dihedral angles. However, as the number of couplings decreases, the area of ϕ–ψ space that

is consistent with all the data increases. We usually find good agreement with the X-ray angles in β-sheet and α-helical regions, but often see differences between X-ray and NMR structures in regions connecting secondary structures. With this technique, we are unable to determine whether these differences are due to different static structures, different dynamic states, or limitations of our method. The complete set of contour plots for 123 residues is available as supplementary material.

METHODS

In the case of a single conformation about ϕ and ψ, the most probable set of dihedral angles can be found from the minimum of

$$f(\phi,\psi) = \sum_{i}^{nexp} c_i \{ J_i^{fit}(\phi,\psi) - J_i^{exp} \}^2 \quad (1)$$

where c_i is a weighting coefficient for coupling type i and nexp is the number of measured couplings for a particular amino acid residue. In the work shown below, we have chosen the c_i values to be proportional to the relative values of r^2, obtained from the experimental correlations described in the previous paper (Edison et al., 1994). In principle, the global minimum of $f(\phi,\psi)$ would provide the single pair of dihedral angles that best fit the data. In practice, however, we have found it most useful to make a contour plot of $f(\phi,\psi)$, so that other minima and the general shape of the function can be ascertained.

Others have used expressions similar to Eq. 1 with $^3J_{H^N H^\alpha}$ (Kim and Prestegard, 1990; Torda et al., 1993) and $^1J_{C^\alpha H^\alpha}$ (Mierke et al., 1992) by converting $f(\phi,\psi)$ to energy units for restrained molecular dynamics calculations. Our results are directly convertible to the form of a 'rotational pseudopotential' function. In this work we simply present the comparison of the X-ray-derived angles with the contour plots of $f(\phi,\psi)$, in order to most easily evaluate the utility of this approach.

RESULTS AND DISCUSSION

We have used the empirically fitted values of $^3J_{H^N H^\alpha}$, $^1J_{C^\alpha H^\alpha}$, $^2J_{C'H^\alpha}$, $^1J_{C^\alpha N}$ and $^2J_{C^\alpha N}$ from Tables 2 and 3 of the previous paper (Edison et al., 1994) to construct error surfaces according to Eq. 1. The weighting coefficients c_i were chosen by the ratio $c_i = r_i^2/r_{max}^2$, where r_i^2 is the regression coefficient from the experimental correlation for coupling constant i and r_{max}^2 is that from the coupling with the best fit to experimental data ($^3J_{H^N H^\alpha}$ in this case). In practice, $f(\phi,\psi)$ is not very sensitive to other choices of c_i, for example an arbitrary setting of all values to 1.0 (data not shown). The weighting coefficients used in this study were: 1.0, 0.63, 0.41, 0.98 and 0.83 for $^3J_{H^N H^\alpha}$, $^1J_{C^\alpha H^\alpha}$, $^2J_{C'H^\alpha}$, $^1J_{C^\alpha N}$ and $^2J_{C^\alpha N}$, respectively.

The X-ray structure of staphylococcal nuclease (1.65 Å resolution) used for this study was that from Loll and Lattman (1989). The NMR data were from Bax's group at NIH ($^3J_{H^N H^\alpha}$: Kay et al., 1989; $^1J_{C^\alpha H^\alpha}$: Vuister et al., 1993; $^2J_{C'H^\alpha}$: Vuister and Bax, 1992; $^1J_{C^\alpha N}$ and $^2J_{C^\alpha N}$: Delaglio et al., 1991).

As an illustrative example, Fig. 1 shows the five separate components and the sum of all the couplings of $f(\phi,\psi)$ for Glu^{10} of staphylococcal nuclease. It is clear from this figure that each of the five couplings contributes to the definition of ϕ–ψ space, but that one angle often is dominant.

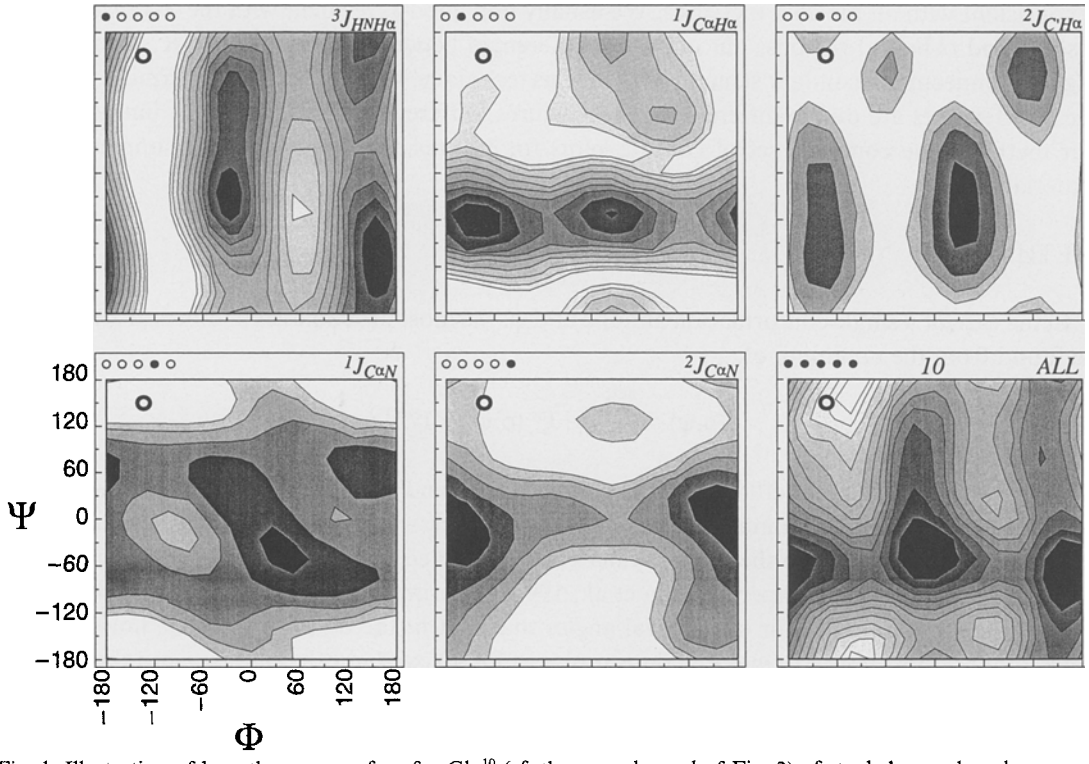


Fig. 1. Illustration of how the error surface for Glu[10] (cf. the second panel of Fig. 2) of staphylococcal nuclease was constructed from the five couplings used in this study. The circles on top of the plots indicate which experimental coupling was used (filled) or not used (empty). They correspond (left to right) to $^3J_{H^NH\alpha}$, $^1J_{C\alpha H\alpha}$, $^2J_{C'H\alpha}$, $^1J_{C\alpha N}$ and $^2J_{C\alpha N}$. Contour levels are shaded from white (best agreement with experiment) to black (worst agreement with experiment). The white contours, which indicate regions with the lowest values of Eq. 1, indicate the most likely ϕ and ψ angles. The first five panels show error surfaces for the individual experimental couplings. The final panel shows the combined error surfaces for all five couplings and the X-ray dihedral angles (indicated by the 'bull's-eye').

In Figs. 1 and 2, the circles at the top of each panel successively represent (from left to right) $^3J_{H^NH\alpha}$, $^1J_{C\alpha H\alpha}$, $^2J_{C'H\alpha}$, $^1J_{C\alpha N}$ and $^2J_{C\alpha N}$. They are filled if the experimental coupling was available and empty if not. The residue number is at the top of each panel. The contour values (in increasing shades of gray) were drawn at: 1, 2, 4, 7, 11, 16, 22, 29, 37, 46, 56 and 67 Hz^2. The lowest (whitest) contours represent the smallest value of Eq. 1 and thus the most probable region for a static conformation. We have drawn 'bull's-eyes' with about 10° radii which represent about one-half of the minimum expected experimental uncertainty in the dihedral angles determined from X-ray analysis (Edison et al., 1994).

Figure 2 includes all residues with the complete set of the five experimental couplings considered in this paper (residues 10, 17, 18, 24, 26, 28, 34, 69, 70, 71, 73, 74, 75, 76, 77, 90, 94 and 101). Moreover, we have included six additional residues, each with four experimental couplings, that demonstrate some of the effects of missing data (residues 25, 82, 83, 111, 119 and 121).

In analyzing the error surfaces of Fig. 2, it is important to realize that these surfaces were generated by using a static model; more sophisticated approaches to conformational averaging

have been proposed (Džakula et al., 1992a,b; Torda et al., 1993). In residues with disagreement between the X-ray and NMR data, it is difficult to distinguish real physical differences (which would indicate structural differences in the crystalline and solution states) from artifacts of the X-ray structure, the measured couplings, or the calculated couplings used in Eq. 1.

First, we consider residues from Fig. 2 with all five experimental couplings. With the exception of residues 17, 18, 69 and 70, the estimates of dihedral angles from the NMR coupling data overlap with the X-ray angles. The lowest contour region usually contains, or is within about 15° of, the X-ray angles. Additionally, our analysis is sensitive to small changes within a region of secondary structure, as evidenced by the stretch from residues 71 to 77. However, two general problems with this analysis are clear by looking at almost every residue plot. The minima often are somewhat broad (e.g., residues 10 and 24), and other regions of the ϕ–ψ map contain local minima. The broad minima generally have the absolute minimum in the region of the 'correct' X-ray angles, but the shapes of the wells indicate that these five couplings alone would generally constrain the angles to within no more than about $\pm 30°$ in either direction. However, the best ultimate use of this data will be in conjunction with NOE and other NMR data. The second problem of local minima is potentially more disturbing. Many regions of local minima occur outside of energetically 'allowed' ϕ–ψ space. For example, many of the plots have minima in the region around $\phi = 60°$, $\psi = -150°$, but this is a high-energy region and could be dismissed by physical arguments. However, a second common region around $\phi = 60°$, $\psi = 60°$ is often a local minimum and could cause some uncertainty, but when all five couplings were used, the other local minima usually had values higher than the region around the X-ray angles. Also, when the 'correct' answer is in the region around $\phi = 60°$, $\psi = 60°$, there seems to be less ambiguity in energetically allowed regions (e.g., residue 28).

Four residues that have all five experimental couplings do not agree with X-ray data: 17, 18, 69 and 70. Residues 17 and 18 are at the transition from a β-sheet to a reverse turn (Loll and Lattman, 1989; Wang et al., 1990). We often find a disagreement between our analysis and the X-ray structure at such interfaces (data available as supplementary material). This disagreement might be due to conformational averaging, different static structures, or inadequate theoretical treatment of these regions. Also, Loll and Lattman report two unusual β-bulges at residues 15, 16 and 24 and residues 18, 19 and 22 (Loll and Lattman, 1989). Residues 69 and 70 do not agree well with the X-ray dihedral angles. In the crystal structure, Loll and Lattman report two intermolecular hydrogen bonds from Lys^{70} and Lys^{71} into the active site of an adjacent molecule (Loll and Lattman, 1989). We show good agreement with residue 71, but is is likely that this region will generally differ in the solution and crystal structures.

We have shown six examples of our calculations with a single coupling constant missing. From the complete set of calculations, we have generally found that $^3J_{H^N H^\alpha}$ is the most important coupling to define ϕ and $^1J_{C^\alpha H^\alpha}$ is the most important to define ψ. Although the other couplings ($^2J_{C'H^\alpha}$, $^1J_{C^\alpha N}$ and $^2J_{C^\alpha N}$) all contribute to both ϕ and ψ, they tend to be less important. The typical response to a missing $^3J_{H^N H^\alpha}$ can be seen by the comparison of residues 24, 25 and 26. These all have about the same X-ray dihedral angles and general features, but residue 25 (which lacks $^3J_{H^N H^\alpha}$) is clearly much broader along ϕ. The difference between missing $^2J_{C^\alpha N}$ and $^3J_{H^N H^\alpha}$ can be seen by comparing residues 82 and 83. In residue 82, the lack of $^2J_{C^\alpha N}$ is hardly noticeable, but without $^3J_{H^N H^\alpha}$, residue 83 has almost no definition along ϕ. Residues 119 and 121 illustrate the relative importance of $^1J_{C^\alpha H^\alpha}$ and $^2J_{C^\alpha N}$. Clearly, without $^1J_{C^\alpha H^\alpha}$, 119 has almost no definition

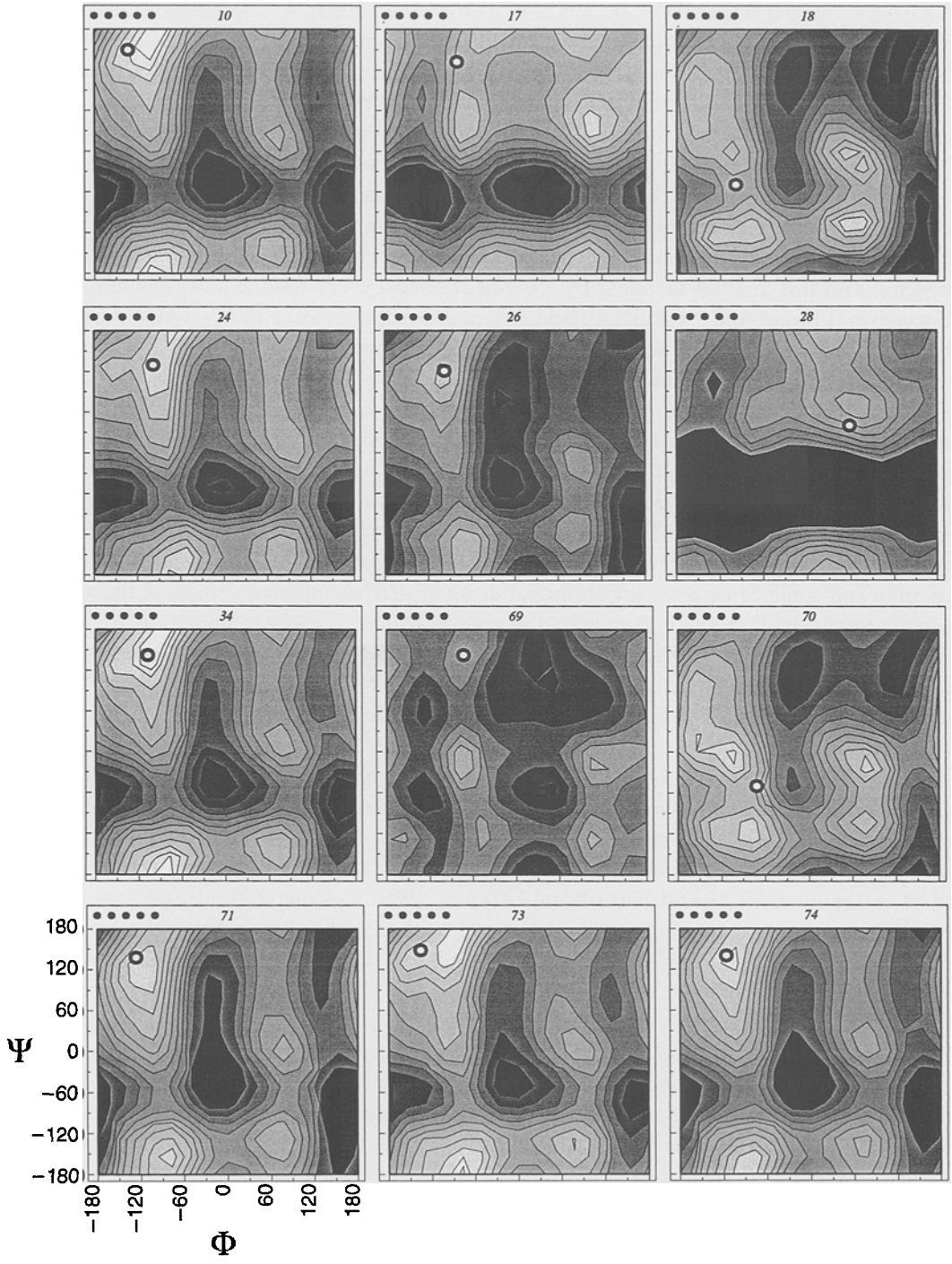
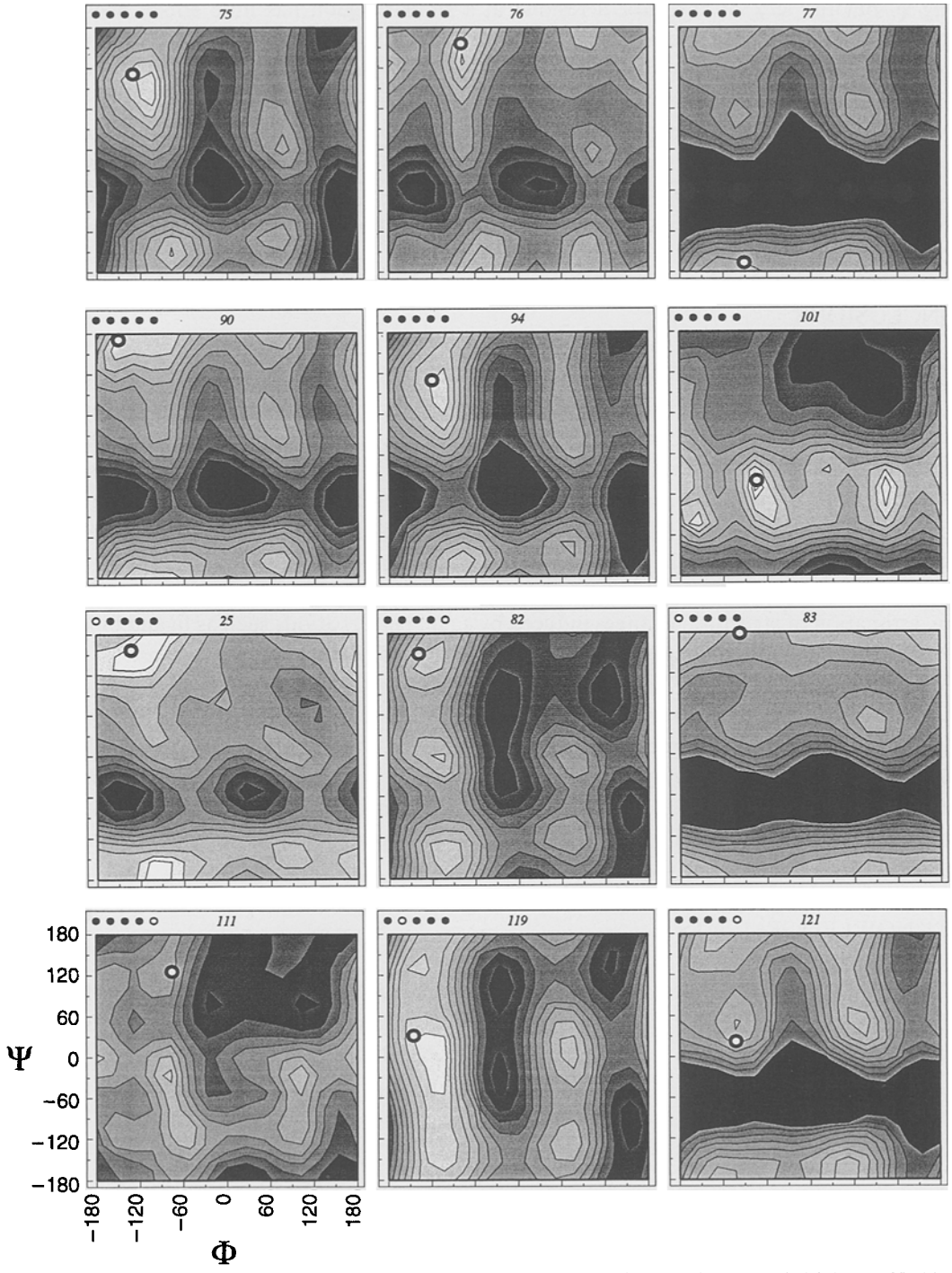


Fig. 2. Error surfaces derived for selected residues of staphylococcal nuclease, for which four or five experimental couplings were available. The circles at the top correspond to $^{3}J_{H^{N}H^{\alpha}}$, $^{1}J_{C^{\alpha}H^{\alpha}}$, $^{2}J_{C'H^{\alpha}}$, $^{1}J_{C^{\alpha}N}$ and $^{2}J_{C^{\alpha}N}$, respectively, and are filled when the experimental coupling constant was available and empty when there was no data (see the Results



and Discussion section for experimental NMR references). The contours were drawn at 1, 2, 4, 7, 11, 16, 22, 29, 37, 46, 56 and 67 Hz^2, with white corresponding to the lowest error (most likely region for the dihedral angles). The 'bull's-eyes' drawn on the plots represent the (ϕ,ψ) angles for that residue from the X-ray crystal structure (Loll and Lattman, 1989).

along ψ, but the lack of $^2J_{C\alpha_N}$ for a residue in a similar region has little effect. Most of our calculations are relatively insensitive to $^2J_{C'H\alpha}$, $^1J_{C\alpha_N}$ and $^2J_{C\alpha_N}$. A possible exception to this is residue 111. Like residues 82 and 121, 111 is missing $^2J_{C\alpha_N}$. However, unlike these other cases, we see a rather large discrepancy for 111 which cannot be explained by crystal contacts. Like residues 17 and 18, 111 is in a region between elements of secondary structure (Loll and Lattman, 1989) but we currently do not know the basis of differences often found in these regions. Note that our calculations for $^1J_{C\alpha H\alpha}$ are considerably more qualitative than those for $^3J_{HNH\alpha}$ (Edison et al., 1994). The discrepancy in residue 111 is along ψ, so it is possible that this is an artifact of our calculations.

CONCLUSIONS

Empirical potential energy functions have many shortcomings in molecular dynamics calculations. Any experimental data that can be interpreted theoretically should be used to supplement (or correct) empirical energy potentials. Some groups have made progress in implementing equations similar to Eq. 1 as restraints into molecular dynamics calculations (Kim and Prestegard, 1990; Mierke et al., 1992; Torda et al., 1993). Our approach described in this and the previous paper should provide a useful addition to the restrained molecular dynamics effort*. Although we have presented this work in the form of contour plots to predict static conformations, we think that the practical implementation of these methods will be in the prediction of secondary structure, estimation of structural changes induced by small perturbations such as ligand binding, and as restraints in simulated annealing or restrained molecular dynamics calculations.

We must stress, however, that errors in experimental X-ray dihedral angles and measured couplings (described in the previous paper) used to calibrate *any* theoretical function can lead to large errors in the implementation of the function. Thus, even for the best behaved theoretical couplings (such as the Karplus equation, or our results for $^3J_{HNH\alpha}$ or $^1J_{C\alpha_N}$), we would suggest that potentials generated from Eq. 1 should have no penalties in at least a 30° range and relatively soft penalties outside this range. For the more qualitative couplings ($^1J_{C\alpha H\alpha}$ and $^2J_{C'H\alpha}$), substantially larger regions without penalty would be appropriate.

ACKNOWLEDGEMENTS

We thank Dr. Željko Džakula for helpful discussions. Drs. Geerten Vuister and Ad Bax kindly supplied us with results prior to publication. A.S.E. is a trainee of an NIH Molecular Biophysics Training Grant (GM 08293). J.L.M. is supported by NIH Grant RR 02301.

REFERENCES

Bax, A., Max, D. and Zax, D. (1992) *J. Am. Chem. Soc.*, **114**, 6923–6925.
Bax, A. and Grzesiek, S. (1993) *Acc. Chem. Res.*, **26**, 131–138.

---

*A *Mathematica* (Wolfram, 1991) package containing the calculated coupling surfaces from the previous paper (Edison et al., 1994) and a routine to calculate error surfaces according to Eq. 1 is available upon request from: Operations Assistant, NMRFAM, Biochemistry Department, University of Wisconsin-Madison, 420 Henry Mall, Madison, WI 53706, U.S.A.

Blake, P.R., Lee, B., Summers, M.F., Adams, M.W.W., Park, J.-B., Zhou, Z.H. and Bax, A. (1992) *J. Biomol. NMR*, **2**, 527–533.

Chary, K.V., Otting, G. and Wüthrich, K. (1991) *J Magn. Reson.*, **93**, 218–224.

Delaglio, F., Torchia, D.A. and Bax, A. (1991) *J. Biomol. NMR*, **1**, 439–446.

Džakula, Ž., Westler, W.M., Edison, A.S. and Markley, J.L. (1992a) *J. Am. Chem. Soc.*, **114**, 6195–6199.

Džakula, Ž., Edison, A.S., Westler, W.M. and Markley, J.L. (1992b) *J. Am. Chem. Soc.*, **114**, 6200–6207.

Edison, A.S., Westler, W.M. and Markley, J.L. (1991) *J. Magn. Reson.*, **92**, 434–438.

Edison, A.S., Markley, J.L. and Weinhold, F. (1994) *J. Biomol. NMR*, **4**, 519–542.

Griesinger, C. and Eggenberger, U. (1992) *J. Magn. Reson.*, **97**, 426–434.

Kay, L.E., Brooks, B., Sparks, S.W., Torchia, D.A. and Bax, A. (1989) *J. Am. Chem. Soc.*, **111**, 5488–5490.

Kim, Y. and Prestegard, J.H. (1990) *Protein Struct. Funct. Genet.*, **8**, 377–382.

Loll, P.J. and Lattman, E.E. (1989) *Protein Struct. Funct. Genet.*, **5**, 183–201.

Markley, J.L. and Kainosho, M. (1993) In *NMR of Biological Molecules*, (Ed., Roberts, G.C.K.) Practical Approach Series, Oxford Press, New York, NY, pp. 101–152.

Mierke, D.F., Grdadolnik, S.G. and Kessler, H. (1992) *J. Am. Chem. Soc.*, **114**, 8283–8284.

Montelione, G.T., Winkler, M.E., Rauenbuhler, P. and Wagner, G. (1989) *J. Magn. Reson.*, **82**, 198–204.

Neuhaus, D., Wagner, G., Vasák, M., Kägi, J.H.R. and Wüthrich, K. (1984) *Eur. J. Biochem.*, **143**, 659–667.

Torda, A.E., Brunne, R.M., Huber, T., Kessler, H. and Van Gunsteren, W.F. (1993) *J. Biomol. NMR*, **3**, 55–66.

Vuister, G.W. and Bax, A. (1992) *J. Biomol. NMR*, **2**, 401–405.

Vuister, G.W., Delaglio, F. and Bax, A. (1993) *J. Biomol. NMR*, **3**, 67–80.

Wagner, G., Hyberts, S.G. and Havel, T.F. (1992) *Annu Rev. Biophys. Biomol. Struct.*, **21**, 167–198.

Wider, G., Neri, D., Otting, G. and Wüthrich, K. (1989) *J. Magn. Reson.*, **85**, 426–431.

Wolfram, S. (1991) *Mathematica: A System for Doing Mathematics by Computer*, 2nd ed., Addison-Wesley, Redwood City, CA.