Journal of Biomolecular NMR, **13:** 181–185, 1999.
KLUWER/ESCOM

Analysis of deuterium relaxation-derived methyl axis order parameters and correlation with local structure

Anthony Mittermaier[a,b], Lewis E. Kay[b,c] & Julie D. Forman-Kay[a,b]
[a]*Hospital for Sick Children, Structural Biology and Biochemistry Programme, 555 University Avenue, Toronto, ON, Canada M5G 1X8*
[b]*Department of Biochemistry, University of Toronto, Toronto, ON, Canada M5S 1A8*
[c]*Protein Engineering Network Centers of Excellence and Departments of Medical Genetics and Chemistry, University of Toronto, Toronto, ON, Canada M5S 1A8*

Received 11 September 1998; Accepted 20 October 1998

Key words: order parameters, packing density, protein dynamics, surface accessibility

Abstract

Methyl axis (S^2_{axis}) and backbone NH (S^2_{NH}) order parameters derived from eight proteins have been analyzed. Similar distribution profiles for Ala S^2_{axis} and S^2_{NH} order parameters were observed. A good correlation between the two S^2_{axis} values of Val and Leu methyl groups is noted, although differences between order parameters can arise. The relation of S^2_{axis} or S^2_{NH} to solvent accessibility and packing density has also been investigated. Correlations are weak, likely reflecting the importance of collective, non-local motions in proteins. The lack of correlation between these simple structural parameters and dynamics emphasizes the importance of motional studies to fully characterize proteins.

Proteins are not rigid. While a great deal can be learned from the static structures obtained using NMR and X-ray crystallography, these models do not provide information on transiently populated conformations which may be biologically relevant (Alber et al., 1983; Varley and Pain, 1991; Zavodsky et al., 1998) nor do they describe the residual entropy of the folded state which is important for protein stability (Karplus et al., 1987). Consequently, a great deal of effort has been devoted to the characterization of protein dynamics using a wide variety of techniques (Brooks et al., 1988; Palmer, 1997).

Over the past several years NMR-based methods have emerged for studying protein dynamics over timescales ranging from ps (Palmer, 1997) to ms (Farrow et al., 1994; Tolman et al., 1997). ^{15}N spin relaxation studies (Kay et al., 1989), for example, have become a routine part of NMR structural analyses and yield information about the motion of backbone as well as sidechain NH bond vectors on the ps–ns timeframe. Typically ^{15}N relaxation data is interpreted in terms of an order parameter, S, describing the amplitude of ps–ns bond vector motions along with an effective correlation time, τ_e, related to the timescale of the motions (Lipari and Szabo, 1982). More recently, methods have been developed for the study of sidechain dynamics in proteins (Muhandiram et al., 1995; LeMaster and Kushlan, 1996; Yang et al., 1998). In one such approach proteins are labeled uniformly with ^{13}C and fractionally deuterated, allowing measurement of T_1 and $T_{1\rho}$ relaxation times of deuterons in CH_2D (Muhandiram et al., 1995) and CHD (Yang et al., 1998) groups. This method is particularly attractive because the relaxation of the deuteron is dominated by the well-understood quadrupolar interaction. Dynamics of methyl groups in a number of proteins have now been measured using the ^{2}H-based approach. In the present communication we use a database of eight proteins for which structural information, ^{15}N-^{1}HN S^2 values and methyl sidechain dynamics are available (see legend to Figure 1) in order to establish whether there are correlations between

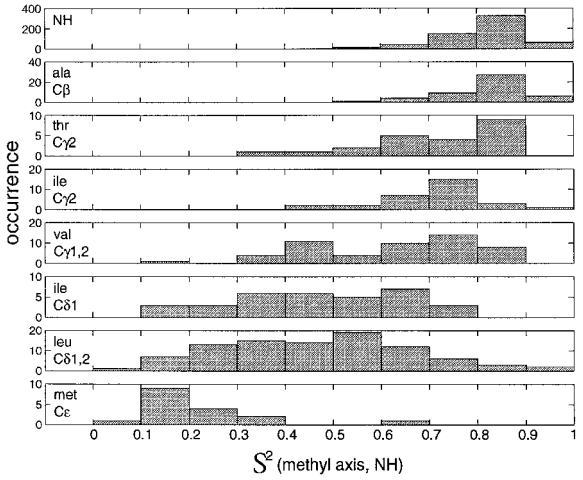

Figure 1. Distribution of order parameters for the backbone amides (S^2_{NH}) and Ala, Thr, Ile, Val, Leu and Met methyl axes (S^2_{axis}) of eight proteins. Values of S^2_{axis} were determined from CH_2D moieties using methods described previously (Muhandiram et al., 1995). Sample sizes are NH: 583, Ala: 46, Thr: 22, Ile $C^{\gamma 2}$: 30, Val $C^{\gamma 1,\gamma 2}$: 52, Ile $C^{\delta 1}$: 33, Leu $C^{\delta 1,\delta 2}$: 92, Met C^{ϵ}:17. The proteins studied were apo-calmodulin [3CLN (Babu et al., 1988)], the N-terminal cellulose binding domain from endoglucanase C [1ULO (Johnson et al., 1996)], the C-terminal SH2 domain from phospholipase Cγ1 [2PLE (Pascal et al., 1994)], the N-terminal SH3 domain from drk, Ca^{2+}-bound staphylococcal nuclease [1SNC (Loll and Lattman, 1989)], the N-terminal SH2 domain from Syp tyrosine phosphatase [1AYD (Lee et al., 1994)], the N-terminal domain of troponin C [5TNC (Herzberg and James, 1988)) and ubiquitin [1UBQ (Vijay-Kumar et al., 1987)]. Data from flexible N- and C-termini and from the domain linker region of calmodulin were not included in the analysis. In the cases of the C-terminal SH2 domain from phospholipase Cγ1 (Kay et al., 1996), the N-terminal SH3 domain from drk (Yang and Kay, 1996), the N-terminal SH2 domain from Syp tyrosine phosphatase (Kay et al., 1998) and the N-terminal domain of troponin C (Gagne et al., 1998), the ^2H S^2_{axis} values have been published (references are listed adjacent to the protein).

backbone and sidechain order parameters as well as between sidechain mobility and the degree of solvent exposure and sidechain packing.

Figure 1 illustrates the amide and methyl axis order parameter distributions for the backbone and for each methyl-containing residue. As expected, there is a general decrease in the value of the order parameter as the separation from the backbone increases (Kay et al., 1996; LeMaster and Kushlan, 1996; Wynn et al., 1996). With the exception of methionine, there is also a general increase in the width of each of the distributions with increasing separation from the mainchain. It is noteworthy that NH order parameters (S^2_{NH}) and alanine methyl axis (C_α-C_β) order parameters (S^2_{axis}) have similar distributions. Indeed, the two samples can be described with a single distribution function [established by the Kolmogorov–Smirnov test (Press et al., 1988)]. This is in direct contrast to results from a ^{13}C relaxation study of staphylococcal nuclease which found that Ala S^2_{axis} values were significantly smaller than measured $S^2_{C\alpha H\alpha}$ values (Nicholson et al., 1996). Interestingly, the correlation coefficient, r (Zar, 1984), between an Ala S^2_{axis} value and the backbone NH order parameter of the same residue is only 0.34, with *p*, the probability of the two samples derived from uncorrelated populations, equal to 0.022. This low correlation is likely due to the intervening dihedral angle, ϕ. One would expect the correlation between alanine S^2_{axis} and $S^2_{C\alpha H\alpha}$ to be much higher, although sufficient data is not available at present to establish whether this is in fact the case. Statistically significant correlations ($p < 0.05$) were not observed between S^2_{axis} and S^2_{NH} for any of the other residues, suggesting that the degree of bond vector mobility is (weakly) correlated over a separation of one dihedral angle but not, in general, much further. Finally, although the S^2_{axis} distributions for Ile $C^{\delta 1}$ and Leu $C^{\delta 1,\delta 2}$are similar, the profiles for Ile $C^{\gamma 2}$ and Val $C^{\gamma 1,\gamma 2}$ are statistically distinct suggesting that the presence of the $C^{\delta 1}$ carbon in Ile has the effect of reducing mobility at the $C^{\gamma 2}$ position.

Since the two methyl groups of Val and Leu residues belong to a single isopropyl moiety, they must have essentially the same mobility. As noted previously, however, the order parameters of the methyl groups can differ if for example the effective averaging axis for the isopropyl unit makes different angles with the two methyl threefold axes (LeMaster and Kushlan, 1996; Yang et al., 1998). In Figure 2, the S^2_{axis}values for the two methyls of Val (a), Leu (b) and Ile (c) are compared. In the case of Val and Leu the agreement between intra-residue S^2_{axis}values is close, with correlation coefficients of 0.89 and 0.86, respectively. About half the methyl pairs for these residues are equal to within experimental error. The $C^{\gamma 2}$ and $C^{\delta 1}$ methyl axis order parameters of Ile correlate less well ($r = 0.44$, $p = 0.016$). This level of correlation is similar to that observed between $S^2_{axis}$and S^2_{NH} for Ala.

Buck et al. (1995) have found that Asn and Gln NH_2 order parameters in lysozyme correlate well with their degree of burial. These residues can be found in both highly solvent exposed as well as buried positions and when buried they often participate in hydrogen bonds or polar interactions. A study which compared the motion of nitroxide spin labels introduced at com-

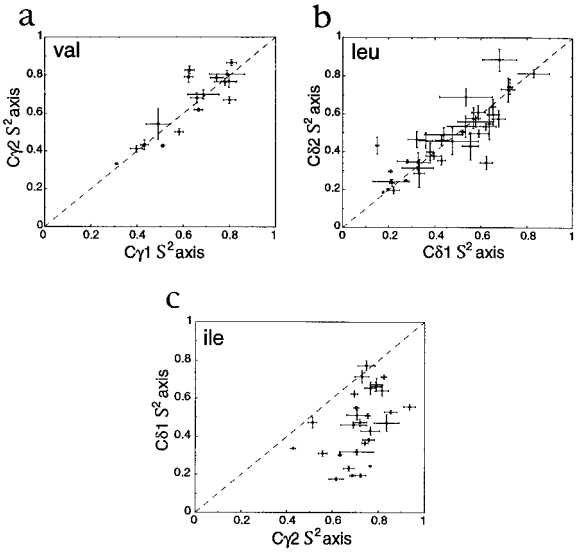

Figure 2. Comparison of the two methyl axis order parameters of Val (a), Leu (b) and Ile (c). Since S^2_{axis} values could only be obtained for one of the methyls in some cases, the number of points relative to Figure 1 is reduced. Sample sizes are: Val 18, Leu: 38, Ile: 29.

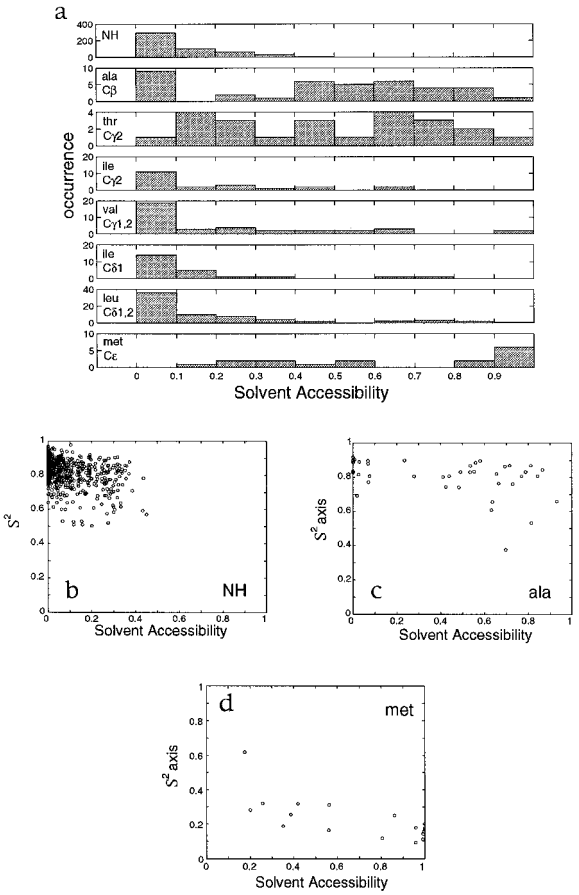

Figure 3. (a) Distribution of fractional solvent accessibilities calculated on a per methyl basis. For the backbone, fractional solvent accessibility of the (N, HN, C^{α}, H^{α}, C, O) unit was calculated. Values of (r,*p*) are (-0.32,$1.2e^{-15}$), (-0.37,0.011) and (-0.71,$8.7e^{-4}$) for NH (b), Ala (c) and Met (d), respectively.

pletely buried or exposed sites also noted a correlation with solvent accessibility (Mchaourab et al., 1996). Fractional solvent accessibilities for each of the proteins considered here were calculated on a per-atom basis with the program MOLMOL (Koradi et al., 1996) using a probe radius of 1.4 Å. The methyl solvent accessibility is defined as the exposed surface area of the methyl group in the context of the entire protein normalized by the exposed surface area of the methyl group in the residue alone. As shown in Figure 3a, the accessibility of Ile, Leu and Val methyl groups is almost always low while Thr, Ala and Met display a greater range of accessibilities. Backbone amide order parameters were found to have a very small, $r = -0.32$, but statistically significant, $p = 1.2e^{-15}$, anti-correlation with the solvent accessibility of the local backbone atoms (Figure 3b). The correlation of S^2_{axis} with methyl solvent accessibility was also found to be weak for Ala (Figure 3c), with $r = -0.37$ and $p = 0.011$, and stronger for methionine (Figure 3d) with $r = -0.71$ and $p = 8.7e^{-4}$. Statistically significant correlations ($p < 0.05$) were not found for Thr, Ile, Val or Leu. If similar sample sizes are selected at random from the S^2_{NH} vs. solvent accessibility profile and correlation statistics are performed, *p* is greater than the cutoff value of 0.05 in approximately one third of the cases. This suggests that even if weak correlations on the order of those observed for the backbone were to exist for these residues, they could pass undetected due to the smaller sample sizes.

Normalized packing values were calculated for the methyl groups as described by DeDecker et al. (1996) and Pattabiraman et al. (1995) and interpreted as the local packing density. A value of zero indicates complete isolation while a value of one implies contact with other atoms over the entire van der Waals surface of the methyl group. Local packing density and solvent accessibility are highly anti-correlated, with average values of $r = -0.88$, $p = 4.7e^{-98}$ for all of the methyls in the eight proteins in our data base. Figures 3a and 4a indicate, however, that important differences in the packing density and solvent accessibility profiles do exist. In the case of Val, Ile and Leu, which are all predominantly buried, a wide range of packing densities are observed. Of note, no statis-

tically significant correlations ($p < 0.05$) between S^2_{axis} and local packing density were found for these residues. As for solvent accessibilities, a very weak but significant correlation was detected between backbone NH order parameters and local packing density (r = 0.35, $p = 6.3e^{-18}$; Figure 4b). Ala (r = 0.34, $p = 0.021$; Figure 4c) and to a larger extent Met (r = 0.68, $p = 0.0017$; Figure 4d) S^2_{axis} values also showed weak correlations with packing density. The greater dependence of S^2_{axis} with packing density in the case of Met relative to other residues may be due to its frequent occurrence in partially solvent-exposed positions. In this environment it would have both conformational (between rotamers) and vibrational (within each rotamer) degrees of freedom and might therefore be more sensitive to increases in local packing than a predominantly buried residue which is already restricted to a single conformer.

The relationship between τ_e and local structure was also investigated. It is noteworthy that the interpretation of methyl τ_e values is not straightforward since they depend on both the rate of methyl rotation and the rate of reorientation of the methyl axis. Moreover, τ_e can be related to microscopic rate constants describing the motion only within the framework of a particular motional model (Lipari and Szabo, 1982). With these caveats in mind, statistically significant (negative) correlations between τ_e and solvent accessibility were found for Ala, Thr and Met, while positive correlations were observed between τ_e and packing density for Ala and Thr. The strongest correlations were noted for Thr, with (r,p) = (-0.59, 0.024) and (0.66,$5.0e^{-4}$) for solvent accessibility and packing density, respectively.

The most notable result from this analysis is not the presence of correlations but the fact that they are so weak. While solvent-exposed residues are on average less conformationally restrained than deeply buried ones (Buck et al., 1995; LeMaster and Kushlan, 1996), our results show that between these two extremes sidechain mobility is governed only to a small extent by the degree of solvent accessibility or the extent of local packing. This is perhaps not surprising in lieu of results from molecular dynamics simulations which show that collective motions account for most of the variation in mobility between atoms (Swaminathan et al., 1982). Differences in the amount of movement cannot, therefore, be considered an entirely local phenomenon, with the motion of each small portion of a protein likely affected by both proximal and distal interactions in a context dependent manner.

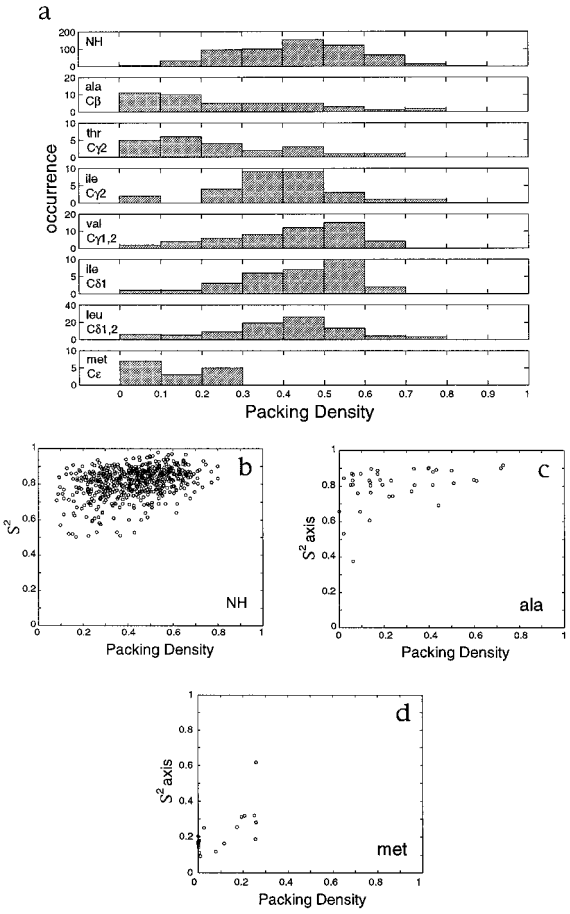

Figure 4. (a) Distribution of local packing density calculated on a per methyl basis. For the backbone, the packing density of the (N, Cα, C, O) unit was calculated using the program *OS* (DeDecker et al., 1996; Pattabiraman et al., 1995). Values of (r,p) are (0.35, $6.3e^{-18}$), (0.34,0.021) and (0.68,0.0017) for NH (b), Ala (c) and Met (d), respectively.

This emphasizes the importance of dynamics studies and underscores the need to develop methodology to probe the codependence of conformational fluctuations. Although such information is not available from a single measure of order along a chain, it is possible to extract correlation parameters from a simultaneous analysis of order parameters measured at each position along a sidechain (Daragan and Mayo, 1998; LeMaster and Kushlan, 1996). As the amount of dynamics data increases and the methodology develops further there will undoubtedly be a corresponding improvement both in the microscopic description of protein thermodynamics and in our intuitive understanding of protein behavior.

In summary, we have analyzed the deuterium relaxation-derived methyl axis order parameters for eight proteins. Of note, Ala S^2_{axis} and S^2_{NH} values were found to have very similar distributions, although the correlation between the two order parameters for a given residue is greatly attenuated by the presence of an intervening dihedral angle. The two methyl axis order parameters of Val and Leu were found to be highly correlated although, due to dependence on the orientation of the effective averaging axis, the values are not always equal. The relationship between methyl axis order parameters, solvent accessibility and packing density has also been investigated. Correlations were found to be consistently very weak, likely the result of the highly cooperative nature of protein dynamics. The fact that the structural properties of a protein cannot easily be used to predict the dynamic properties underscores the fact that experimental dynamics data provide unique information which is crucial to the comprehensive understanding of proteins.

Acknowledgements

This research was supported by the Medical Research Council of Canada (L.E.K. and J.F.K.). L.E.K. is an International Howard Hughes Research Scholar. T.M. is supported by a post-graduate fellowship from the Natural Sciences and Engineering Research Council of Canada.

References

Alber, T., Gilbert, W.A., Ponzi, D.R. and Petsko, G.A. (1983) *Ciba Found. Symp.*, **93**, 4–24.

Babu, Y.S., Bugg, C.E. and Cook, W.J. (1988) *J. Mol. Biol.*, **204**, 191–204.

Brooks, C.L., Karplus, M. and Pettitt, B.M. (1988) *Proteins, A Theoretical Perspective of Dynamics, Structure and Thermodynamics*, Wiley, New York, NY.

Buck, M., Boyd, J., Redfield, C., MacKenzie, D.A., Jeenes, D.J., Archer, D.B. and Dobson, C.M. (1995) *Biochemistry*, **34**, 4041–4055.

Daragan, V.A. and Mayo, K.H. (1998) *J. Magn. Reson.*, **130**, 329–334.

DeDecker, B.S., O'Brien, R., Fleming, P.J., Geiger, J.H., Jackson, S.P. and Sigler, P.B. (1996) *J. Mol. Biol.*, **264**, 1072–1084.

Farrow, N.A., Zhang, O., Forman-Kay, J.D. and Kay, L.E. (1994) *J. Biomol. NMR*, **4**, 727–734.

Gagne, S.M., Tsuda, S., Spyracopoulos, L., Kay, L.E. and Sykes, B.D. (1998) *J. Mol. Biol.*, **278**, 667–686.

Herzberg, O. and James, M.N.G. (1988) *J. Mol. Biol.*, **203**, 761–779.

Johnson, P.E., Joshi, M.D., Tomme, P., Kilburn, D.G. and McIntosh, L.P. (1996) *Biochemistry*, **35**, 14381–14394.

Karplus, M., Ichiye, T. and Pettitt, B.M. (1987) *Biophys. J.*, **52**, 1083–1085.

Kay, L.E., Muhandiram, D.R., Farrow, N.A., Aubin, Y. and Forman-Kay, J.D. (1996) *Biochemistry*, **35**, 361–368.

Kay, L.E., Muhandiram, D.R., Wolf, G., Shoelson, S.E. and Forman-Kay, J.D. (1998) *Nat. Struct. Biol.*, **5**, 156–163.

Kay, L.E., Torchia, D.A. and Bax, A. (1989) *Biochemistry*, **28**, 8972–8979.

Koradi, R., Billeter, M. and Wüthrich, K. (1996) *J. Mol. Graph.*, **14**, 51–55.

Lee, C.H., Kominos, D., Jacques, S., Margolis, B., Schlessinger, J., Shoelson, S.E. and Kuriyan, J. (1994) *Structure*, **2**, 423–438.

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

Lipari, G. and Szabo, A. (1982) *J. Am. Chem. Soc.*, **104**, 4546–4559.

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

Mchaourab, H.S., Lietzow, M.A., Hideg, K. and Hubbard, W.L. (1996) *Biochemistry*, **35**, 7692–7704.

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

Nicholson, L.K., Kay, L.E. and Torchia, D.A. (1996) In *NMR Spectroscopy and its Applications to Biomedical Research* (Sarkar, S.K., Ed.), Elsevier, New York, NY, pp. 241–279.

Palmer, A.G. (1997) *Curr. Opin. Struct. Biol.*, **7**, 732–737.

Pascal, S.M., Singer, A.U., Gish, G., Yamazaki, T., Shoelson, S.E., Pawson, T., Kay, L.E. and Forman-Kay, J.D. (1994) *Cell*, **77**, 461–472.

Pattabiraman, N., Ward, K.B. and Fleming, P.J. (1995) *J. Mol. Recogn.*, **8**, 334–344.

Press, W.H., Flannery, B.P., Teukolsly, S.A. and Vetterling, W.T. (1988) (Eds.) *Numerical Recipes in C*, Cambridge University Press, Cambridge, pp. 490–494.

Swaminathan, S., Ichiye, T., van Gunsteren, W. and Karplus, M. (1982) *Biochemistry*, **21**, 5230–5241.

Tolman, J.R., Flanagan, J.M., Kennedy, M.A. and Prestegard, J.H. (1997) *Nat. Struct. Biol.* **4**, 292–297.

Varley, P.G. and Pain, R.H. (1991) *J. Mol. Biol.* **220**, 531–538.

Vijay-Kumar, S., Bugg, C.E. and Cook, W.J. (1987) *J. Mol. Biol.*, **194**, 531–544.

Wynn, R., Harkins, P.C., Richards, F.M. and Fox, R.O. (1996) *Protein Sci.* **5**, 1026–1031.

Yang, D. and Kay, L.E. (1996) *J. Mol. Biol.*, **263**, 369–382.

Yang, D., Mittermaier, A., Mok, Y.K. and Kay, L.E. (1998) *J. Mol. Biol.*, **276**, 939–954.

Zar, J.H. (1984) Simple Linear Correlation Biostatistical Analysis, Prentice-Hall, Englewood Cliffs, N.J., pp. 306–327.

Zavodsky, P., Kardos, J., Svingor, A. and Petsko, G.A. (1998) *Proc. Natl. Acad. Sci. USA*, **95**, 7406–7411.